

Tesis Doctoral

Del átomo a la célula: determinantes moleculares que definen la función de galectinas humanas

Guardia, Carlos M.A.

2014-05-30

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Guardia, Carlos M.A.. (2014-05-30). Del átomo a la célula: determinantes moleculares que definen la función de galectinas humanas. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Guardia, Carlos M.A.. "Del átomo a la célula: determinantes moleculares que definen la función de galectinas humanas". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2014-05-30.

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires

Del átomo a la célula: determinantes moleculares que definen la función de galectinas humanas

RESUMEN

Las galectinas constituyen una familia de proteínas altamente conservadas a través de la evolución con capacidad de unir carbohidratos presentes en diferentes glicoproteínas de la membrana plasmática y la matriz extracelular. Estas proteínas reconocen en forma específica unidades repetidas de *N*-acetil lactosamina [LacNAc, Gal β 1-4-NAcGlc] presentes en *N*- y *O*-glicanos. El énfasis del trabajo se centró en determinar los detalles estructurales que diferencian a cada una de estas proteínas, buscando correlaciones entre la dinámica molecular, la estabilidad y la selectividad de unión a glicanos para desencadenar la función celular para la que han sido evolutivamente seleccionadas.

A través de una aproximación multidisciplinaria, se empleó una variedad de técnicas de simulación computacional y diversas estrategias experimentales para estudiar los mecanismos y determinantes moleculares de regulación de la unión a glicanos en galectinas y posterior activación de su función en diversos cultivos celulares. Las metodologías empleadas consistieron desde simulaciones de dinámica molecular clásica hasta la expresión y purificación de proteínas recombinantes y ensayos de unión a ligando, cinéticas de oxidación y apoptosis *in vitro* mediante citometría de flujo y diversas técnicas de microscopía óptica y electrónica.

En primer lugar, este estudio se ha centrado en la caracterización y comparación estructural y dinámica de cada galectina y su especificidad por diversos glicanos fisiológicamente relevantes. En segundo lugar, se ha estudiado el mecanismo de oxidación y formación de enlaces disulfuro en galectina-1, proponiendo la existencia de un interruptor molecular presente sólo en este miembro de la familia. El mismo explicaría el cambio de función al cambiar de estado redox. Por último, se estudió el impacto de la estructura cuaternaria en la unión a glicanos específicos en glicoproteínas y la formación de redes supramoleculares de glicanos en la superficie de la célula.

Palabras clave: galectina, dinámica molecular, oxidación, dicroísmo circular, fluorescencia, glicosilación.

From the atom to the cell: molecular determinants that define the function of human galectins'

ABSTRACT

Galectins are members of a family of highly conserved proteins characterized by the ability to bind saccharide motifs in various cell surface glycoproteins and the extracellular matrix. These proteins specifically recognize repeating units of *N*-acetyl lactosamine [LacNAc, Gal β 1-4-GlcNAc] present in *N*- and *O*-glycans. The aim of this work was to define the details and structural motifs that distinguish each individual member of the galectin family, seeking for correlations between the molecular dynamics, stability and selectivity for glycans displayed by these lectins and their proposed functional activity.

Using a multidisciplinary approach, in the present study we used a wide variety of computational simulation tools and different experimental approaches to explore the mechanisms and molecular determinants that regulate glycan binding of galectins in cell cultures. The employed methodologies ranged from classical molecular dynamics simulations to the expression and purification of recombinant proteins and ligand-binding assays, oxidation kinetics and apoptosis assays using flow cytometry and various optical and electronic microscopy techniques.

First, we have focused on the structural characterization and comparison of molecular dynamics of each galectin in the vicinity of their ligand binding sites and their specificity for a set of physiologically relevant glycans. Secondly, we studied the mechanism of oxidation and disulfide bond formation in galectin-1 protein, suggesting the presence of a molecular switch that explains the function change during redox state turnover. Finally, we studied the impact of the quaternary structure in the binding properties of galectins to specific glycans and analyzed the ability of galectin-1 to form supramolecular membrane structures called lattices when they interact with specific glycans on the surface of target cells.

Keywords: galectin, molecular dynamics, oxidation, circular dichroism, fluorescence, glycosylation.

A mis padres y mi hermana.

A mis queridos maestros:

Marta González, Marina Nicelli, Jorge Martínez y Mary Oieni.

*“Siempre nos hemos definido por la habilidad de superar lo imposible. Y contamos esos momentos...momentos en los que nos atrevemos a apuntar más alto, a romper barreras, a estirarnos hacia las estrellas...**hacer conocido lo desconocido**. Contamos estos momentos como los logros que más nos enorgullecen. ¿Hemos perdido todo esto? O quizás sólo hemos olvidado que todavía somos pioneros y que apenas hemos empezado, y que nuestros más grandes logros no pueden estar detrás de nosotros. Porque nuestro destino está por encima.”*

Trailer oficial de la película “*Interstellar*” (2014)

INDICE GENERAL

Resumen	-
Abstract	-
I Introducción general	1
I.1 Biología estructural: donde las disciplinas convergen	1
I.1.1 ¿Qué es la biología estructural?	1
I.1.2 Biología estructural: historia, actualidad y futuro	3
I.2 Glicosilación y lectinas	5
I.2.1 Vías de glicosilación	7
I.2.2 Descubrimiento de las lectinas animales	9
I.2.3 Clasificación actual de las lectinas animales	11
I.2.4 Galectinas: introducción	13
I.3 Objetivos y lineamientos	18
I.4 Referencias	19
II Métodos	24
II.1 Métodos computacionales	24
II.1.1 Métodos basados en la Mecánica Clásica	25
II.1.1.1 Campos de fuerza clásicos	25
II.1.2 Modelos para el solvente	28
II.1.2.1 Solvente implícito: PCM, PB y GB	29
II.1.2.2 Solvente explícito: TIP3P y TIP4P	30
II.1.3 Explorando la superficie de energía potencial	31
II.1.3.1 Minimización de la energía	32
II.1.3.2 Simulaciones de Dinámica Molecular	33
RMSD, RMSF y distribución radial	37
Dinámica esencial	38
Entropía conformacional	40
Energía libre de interacción	42
II.1.4 Modelado por homología	45
II.2 Métodos experimentales	47
II.2.1 Reactivos y anticuerpos	47
II.2.2 Producción de galectina-1 recombinante	48
II.2.3 Líneas celulares	54
II.2.4 Instrumentación	54
II.2.4.1 Espectrofotometría	54
II.2.4.2 Fluorescencia: espectroscopía y microscopía	55
II.2.4.3 Dicroísmo circular	57
II.2.4.4 Microscopía de transmisión electrónica	61
II.2.4.5 Citometría de flujo	63
II.3 Referencias	65

III La relación estructura/función en la familia de las galectinas humanas	69
III.1 Introducción	69
III.2 Métodos computacionales	72
III.3 Resultados	78
III.3.1 Comparación global de las galectinas	78
III.3.2 Comportamiento dinámico de los diferentes DRC	83
III.3.3 Comparación detallada del LBG de todos los DRC	86
III.3.4 Efecto de la unión a ligando en el DRC	89
Estructura y dinámica	89
Termodinámica de los LBG	94
Selectividad por ligandos	95
III.4 Discusión y conclusiones	104
III.5 Referencias	108
 IV Rol de las cisteínas en la función y dinámica de galectina-1 y su familia	 112
IV.1 Introducción	112
IV.2 Métodos experimentales y computacionales	114
IV.3 Resultados	120
IV.3.1 Algunas cisteínas están involucradas en el proceso de oxidación de Gal-1	120
IV.3.2 La oxidación y reducción de Gal-1 es un proceso reversible	122
IV.3.3 Las Cys2, 16 y 88 juegan un rol clave	124
IV.3.4 Oxidación de Gal-1 con H ₂ O ₂	125
IV.3.5 Gal-1 oxidada no une lactosa	129
IV.3.6 Gal-1 oxidada no modifica la viabilidad de células T	131
IV.3.7 Gal-1, cisteínas y sus hermanas	133
IV.4 Discusión y conclusiones	135
IV.5 Referencias	139
 V Explorando la estructura cuaternaria de las galectinas: ¿existen los <i>lattices</i>?	 142
V.1 Introducción	142
V.2 Métodos experimentales y computacionales	147
V.3 Resultados	150
V.3.1 Análisis de la interfase de dimerización de Gal-1	150
V.3.2 Del modelado molecular a las células: ensayo de unión a lectinas	155
V.3.3 Gal-1 induce la agregación del VEGFR2	157
V.4 Discusión y conclusiones	163
V.5 Referencias	165
 VI Conclusiones Generales	 169
VI.1 Conclusiones Generales	169
VI.2 Perspectivas	172
 VII Publicaciones	 174
 VIII Agradecimientos	 175
 Fe de erratas	 181



INTRODUCCION GENERAL

I.1 Biología estructural: donde las disciplinas convergen

I.1.1 ¿Qué es la biología estructural?

La **biología estructural**, como lo han propuesto Hendrickson y Klug [1], es una rama de la biología molecular, la bioquímica y la biofísica que estudia la estructura de macromoléculas biológicas (tales como las proteínas y los ácidos nucleicos), el origen de esta estructura y cómo las alteraciones en la misma afectan su función. Es notable destacar el interés de esta disciplina, en particular en la biología, dado que estas macromoléculas son encargadas de la mayoría de las funciones vitales de las células y porque es sólo a través de su estructura tridimensional particular que son capaces de llevar a cabo dichas funciones.

Las biomoléculas tienen como máximo un tamaño del orden de 100 Å. Por ello, son muy pequeñas para ser vistas y estudiadas en detalle incluso a través del más avanzado de los microscopios ópticos. Así, los métodos experimentales usados por los biólogos estructurales generalmente involucran medidas promediadas de miles de moléculas idénticas al mismo tiempo, utilizando diferentes estrategias, principalmente espectroscópicas. Algunos de los métodos más importantes son:

- Cristalografía de rayos X
- Resonancia magnética nuclear (RMN)
- Resonancia paramagnética electrónica (RPE)
- Microscopía crío-electrónica (Cryo-EM, del inglés *cryo-electron microscopy*)
- Resonancia de efecto Raman
- Dispersión dinámica de luz (DLS, del inglés *dynamic light scattering*)

- Dispersión de rayos X de ángulo pequeño (SAXS, del inglés *small-angle X-ray scattering*)
- Dicroísmo circular e interferometría de polarización dual

La mayoría de los investigadores utilizan estas técnicas para estudiar el estado “nativo” de las macromoléculas, resultando igualmente poderosas también para observar otros estados como el naciente o el desnaturalizado y la **dinámica molecular** (DM) involucrada en la transición entre los distintos estados. Es aquí donde sólo algunas técnicas experimentales pueden ofrecer información de este tipo (RMN, principalmente) y la necesidad de un nuevo enfoque permitió la introducción de la DM por simulación computacional al ámbito de la biología estructural. Originalmente concebida dentro de la física teórica, y en especial para las ciencias de materiales [2-5], la DM ha logrado ampliar la visión bioquímica del funcionamiento de las biomoléculas gracias a la elevada resolución espacial (descripción atómica de sistemas de entre 10.000 y 100.000 átomos) y temporal (desde los picosegundos hasta los milisegundos [6]), gracias al crecimiento exponencial de los avances en *hardware* en las últimas décadas [7-10]. Y como no podía ser de otro modo, el campo de la DM es multidisciplinario. Las leyes y teorías aplicadas provienen de las matemáticas, física y química, así como también del uso de algoritmos recientes de las ciencias de la computación y de la teoría de la información [11]. Comprender el origen, el uso, el alcance y las perspectivas del campo de las simulaciones computacionales requiere necesariamente una interacción dinámica entre todas esas ramas de la ciencia moderna.

Finalmente, es necesario considerar otra disciplina clave en el entendimiento del problema de la biología estructural: la **bioinformática**. La arquitectura de las biomoléculas (su estructura terciaria) depende, de manera todavía muy difícil de predecir, de la estructura primaria, es decir, de la secuencia lineal particular de aminoácidos/nucleótidos [12-14]. Es por ello que la bioinformática, analizando los diferentes patrones en las secuencias de estos bloques constructores trata de predecir la forma, topología, sitios de interacción y características fisicoquímicas relevantes de las macromoléculas a partir del estudio minucioso de la estructura primaria. El aumento de la calidad de dichas predicciones, tanto en tamaño molecular como en complejidad del sistema, será sin lugar a dudas el gran desafío de la biología estructural del futuro que pondrá a prueba a viejas y nuevas hipótesis de la biología [15].

I.1.2 Biología estructural: historia, actualidad y futuro

Los primeros intentos de explorar las propiedades de los materiales biológicos más allá del límite del ojo desnudo se realizaron gracias a los microscopios de R. Hooke (1635-1703) y A. van Leeuwenhoek (1632-1723). Hacia finales del siglo XIX, la microscopia óptica había alcanzado su límite teórico de resolución dado por E. Abbé y otros [16]. Para alcanzar resolución atómica, se necesita normalmente del uso de radiación cuya longitud de onda sea similar a la dimensión de los átomos que deseamos observar ($\sim 0,2$ nm). La longitud de onda de la luz visible es varios órdenes de magnitud mayor (~ 500 nm) que la necesaria para cumplir este requisito. Afortunadamente, desarrollos varios en el área de la física arrojaron nuevas fuentes de longitudes de onda más cortas: W. Röntgen descubrió los rayos X en 1895 y J. J. Thomson, los electrones en 1897. Los descubrimientos de los neutrones por J. Chadwick en 1932 y del efecto de RMN en 1946 completaron la fertilización del campo que daría nacimiento a todas las técnicas modernas de espectrometrías utilizadas en la resolución de estructuras de biomoléculas.

Hace tan sólo 100 años se sabía muy poco acerca de la forma de las biomoléculas debido a la escasez de herramientas y técnicas para el estudio de sus estructuras. En la actualidad, la biología estructural es una ciencia madura [17]. Las coordenadas atómicas de más de 96.000 proteínas, ácidos nucleicos y sus complejos están hoy depositadas en la *Protein Data Bank* (PDB; ver **Figura I.1**). El estado avanzado del campo en el presente se debe principalmente a los logros tecnológicos que han producido herramientas sofisticadas y comerciales, aunque todavía muy costosas. El primer método usado para la determinación de la estructura de macromoléculas fue la interpretación de los patrones de difracción que se producen cuando cristales de proteínas son irradiados con rayos X. A la cristalografía de rayos X y su técnica relacionada, la cristalografía de neutrones, se les han unido la microscopia electrónica y la resonancia magnética nuclear, así como también los desarrollos en materia de tecnología computacional, los cuales han impactado de lleno en todas las técnicas instrumentales de hoy en día. La **Figura I.2** muestra la cronología de algunos de los avances más importantes de la biología estructural de los últimos cien años.

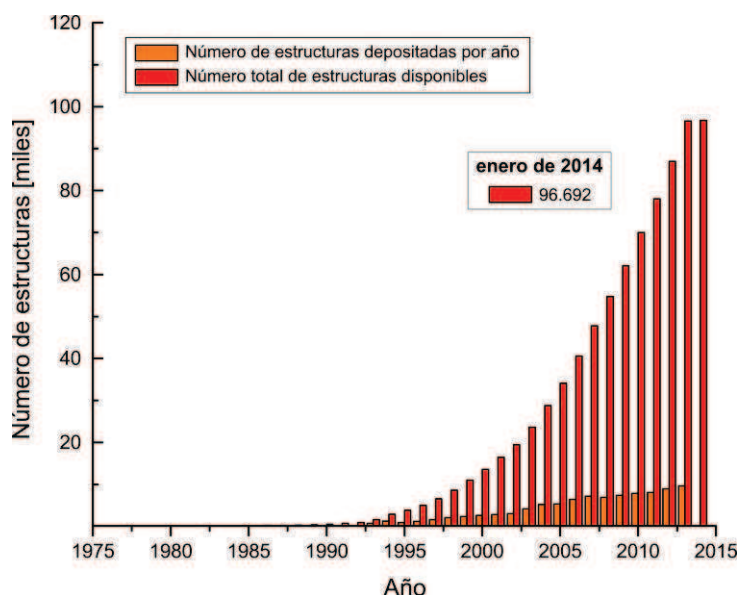


Figura I.1. Protein Data Bank (PDB). La *Protein Data Bank* es un repositorio mundial único para el procesamiento y distribución de estructuras tridimensionales de macromoléculas biológicas. La figura muestra la evolución en el número de estructuras depositadas a lo largo de la vida del proyecto (20/10/1971-06/01/2014). Datos provistos por la PDB (<http://www.rcsb.org/>).

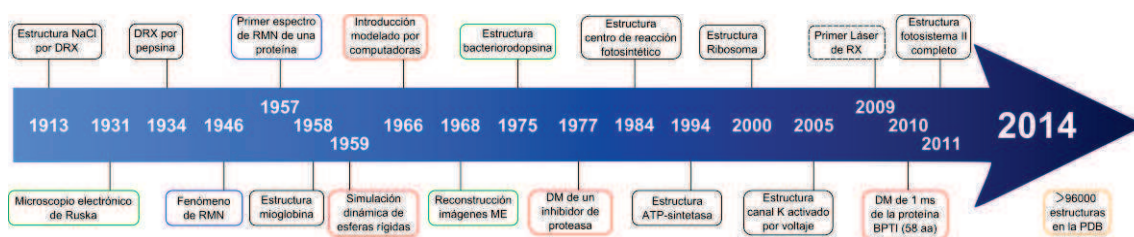


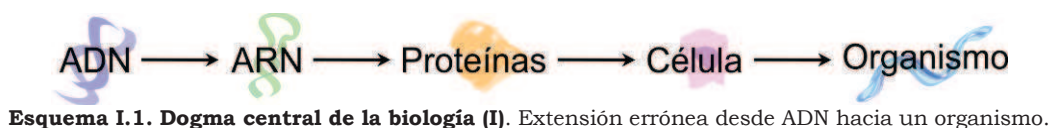
Figura I.2. La evolución de la biología estructural. La línea de tiempo muestra algunos de los desarrollos clave en cristalografía (negro), microscopía electrónica (ME, verde), resonancia magnética nuclear (RMN, azul) y métodos computacionales (rojo). DRX, difracción de rayos X; DM, dinámica molecular; PDB, *Protein Data Bank*.

Es notable el desarrollo particular de la biología estructural en Sudamérica, a pesar del esfuerzo constante que implica el avance de grandes plataformas regionales. Un gran hito local ha sido la creación del Centro de Biología Estructural del Mercosur (CeBEM) en el año 2011. Se trata de un proyecto iniciado por nueve grupos de Brasil, Argentina, Uruguay y Paraguay, destacados por su nivel de excelencia y reconocimiento internacional (www.cebem.org.ar). Su objetivo principal es el de asociar a estos grupos en una red regional, integrando competencias e infraestructura, promover el intercambio de investigadores y graduados, y formar así un aumento de masa crítica en el área que logre la creación, puesta en marcha y finalización de proyectos de investigación de alto impacto científico y biotecnológico. En particular, en nuestro país, son tres los centros de referencia en biología estructural: Fundación Instituto Leloir (Ciudad Autónoma de Buenos Aires), la Facultad de Ciencias Exactas y

Naturales de la Universidad de Buenos Aires (Ciudad Autónoma de Buenos Aires) y el Instituto de Biología Molecular y Celular de Rosario (Rosario, Santa Fe). La necesidad de ampliar la diversidad de técnicas disponibles en la región, así como también la profundización e integración de cada una de ellas, será fundamental para lograr un crecimiento local sostenido de una disciplina que ha venido para instalarse como un eslabón esencial en la cadena de la comprensión de la biología molecular y celular.

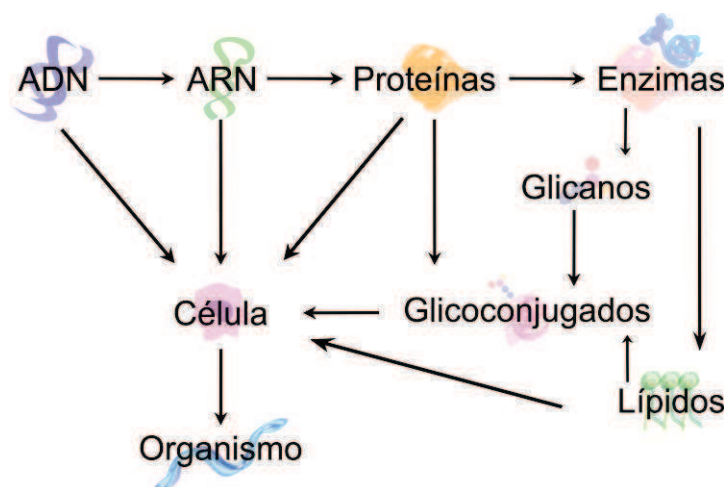
I.2 Glicosilación y lectinas

El paradigma central de la biología molecular moderna indica que la información biológica fluye desde el ADN hacia el ARN y luego hacia las proteínas. El poder de este concepto no sólo radica en la precisión con la que este flujo se da, sino también en la posibilidad que nos ofrece de manipular cualquier clase de biomoléculas basándonos en el conocimiento de otra y en los patrones de homología de secuencia y relación que predicen funciones y relaciones evolutivas [18]. Con el descubrimiento de la secuencia completa del genoma humano y de otras especies comúnmente estudiadas como organismos modelo, se han anticipado descubrimientos aún más espectaculares en la comprensión de los sistemas biológicos. Sin embargo, es frecuente encontrar una acepción errónea en la extensión del dogma central (**Esquema I.1**).



De hecho, para crear una célula se requieren otras dos clases importantes de moléculas: lípidos e hidratos de carbono. Estas moléculas sirven como intermediarios energéticos, moléculas señalizadoras o componentes estructurales de la célula. El rol estructural de los carbohidratos ha sido esencial en la evolución hacia complejos multicelulares, órganos y organismos, lo que requiere de la comunicación entre células entre sí y con la matriz que las rodea. Aún más, todas las células y muchas macromoléculas biológicas llevan unido en forma covalente un complejo y denso arreglo de azúcares (llamados oligosacáridos o glicanos). Dado que la mayoría de los glicanos se encuentran en la superficie externa de las biomoléculas de la célula y en aquellas que son secretadas, estos azúcares pueden modular o mediar una gran

variedad de eventos de comunicación célula-célula y célula-matriz que son cruciales para el desarrollo y la función de un organismo multicelular complejo, como puede ser el cuerpo humano. Incluso, estos glicanos son capaces de permitir la interacción entre organismos (por ejemplo, entre huésped y parásitos durante una infección). Además, es muy común encontrar glicanos más simples y extremadamente móviles unidos a proteínas en el núcleo y en el citoplasma, donde parecen servir como puntos clave de regulación. Por lo tanto, una visión más amplia y completa del paradigma de la biología molecular debería verse como el que se muestra en el **Esquema I.2.**



Esquema I.2. Dogma central de la biología (II). Extensión correcta de un paradigma molecularmente más completo.

En la primera parte del siglo XX, la química y la bioquímica de los carbohidratos fueron definitivamente temas de interés. Sin embargo, durante la fase inicial de la revolución de la biología molecular moderna, el estudio de los glicanos quedó relegado frente a aquellas clases de biomoléculas que recibieron toda la atención (ADN, ARN, proteínas). Algunas causas de este fenómeno se debieron a la inherente complejidad estructural, la dificultad de determinar su secuencia y el hecho de que su biosíntesis no se puede predecir directamente del molde de ADN de las biomoléculas a la cual modifica con su presencia. Fue recién a finales de la década de los '80 que el desarrollo de una variedad de nuevas técnicas para explorar las estructuras de estas cadenas de glicanos abrió un nuevo horizonte en la biología molecular, que incluso ha acuñado su propio nombre: **glicobiología** [19]. Hoy en día, la glicobiología permite un área de encuentro para las disciplinas tradicionales de la química y bioquímica de los carbohidratos con el estudio moderno de la biología celular y molecular de los glicanos. De hecho, el término “glicobiología” ha ganado su lugar en el mundo

científico: cuenta con dos revistas biomédicas prestigiosas, una sociedad científica internacional en constante crecimiento, una *Gordon Research Conference* y un *Keystone Symposium* que llevan su nombre. Investigar en glicobiología requiere no sólo de los aportes clásicos de la nomenclatura, biosíntesis, estructura, síntesis orgánica y función de glicanos complejos, sino también de disciplinas generales de la genética molecular, la biología y fisiología celular y la química de proteínas que permitan entender en profundidad el contexto biológico donde estos hidratos de carbono cumplen su función.

I.2.1 Vías de glicosilación

La superficie de las células de mamíferos y las proteínas de la matriz extracelular con la cual interaccionan se hallan cubiertas por diversos tipos de glicanos unidos a proteínas (glicoproteínas) y lípidos (glicolípidos), cuya composición varía de acuerdo al estado de diferenciación y activación de las células. El proceso de biosíntesis de glicoproteínas está mediado por glicosiltransferasas y glicosidasas que actúan en forma coordinada a medida que las proteínas se sintetizan en el retículo endoplásmico (RE) y son transportadas por el aparato de Golgi hacia la superficie celular [20, 21]. Las glicosiltransferasas son enzimas que catalizan la transferencia de un monosacárido de un nucleótido dador a un sustrato aceptor. Las glicosidasas catalizan la hidrólisis de un enlace glicosídico en un glicano precursor. Estas enzimas actúan en conjunto con chaperonas y cofactores en el RE/Golgi a los fines de agregar y/o eliminar azúcares de una glicoproteína o glicolípido y así determinar su estructura final [18].

La maquinaria de glicosilación representa un 1% del total del genoma y a la fecha se han clonado más de 100 glicosiltransferasas y glicosidasas [22]. Curiosamente, el perfil de esta maquinaria de glicosilación se halla altamente regulado y varía sensiblemente frente a numerosos estímulos externos. En consecuencia, una misma célula en dos contextos diferentes puede presentar variaciones en el patrón de glicosilación que afectan y modulan la funcionalidad de glicoproteínas de la membrana, lo que constituye un nivel de regulación adicional de los procesos celulares. Es interesante destacar la complejidad del sistema, ya que debemos no sólo considerar la variabilidad a nivel de la estructura proteica (proteómica) sino a su vez la variabilidad creada a nivel del patrón de glicosilación (glicómica). En este sentido, durante el proceso de biosíntesis de un glicano, la diversidad de monosacáridos

(fucosa, galactosa, *N*-acetilgalactosamina, glucosa, *N*-acetilglucosamina, ácido glucurónico, manosa, ácido siálico y xilosa), las distintas uniones posibles entre dos monosacáridos y la potencialidad de ramificaciones en la estructura sacarídica resultan en una inmensa variedad de estructuras de glicanos posibles [18]. Dentro de las distintas vías de glicosilación en mamíferos, podemos destacar la *N*- y la *O*-glicosilación. Los *N*-glicanos están covalentemente unidos al residuo asparragina de la secuencia consenso Asn-X-Ser/Thr. Por el contrario los *O*-glicanos se hallan unidos covalentemente a un residuo Ser/Thr de la secuencia primaria y no están determinados por la presencia de secuencias consenso [23].

Una de las funciones biológicas más relevantes de los *N*-glicanos está relacionada con la regulación y el control del plegado de las proteínas que se sintetizan en el RE. En este sentido, cabe destacar que todas las células eucarióticas sintetizan *N*-glicanos y que los primeros pasos de la vía de biosíntesis se hallan altamente conservados. No obstante, en las últimas etapas de dicha vía, los *N*-glicanos adquieren una gran diversidad estructural a medida que transitan por el aparato de Golgi. En los vertebrados, las proteínas secretadas o de la membrana plasmática pueden presentar *N*-glicanos del tipo alta manosa, híbridos o complejos, siendo estos últimos los más abundantes (**Figura I.3**) [24].

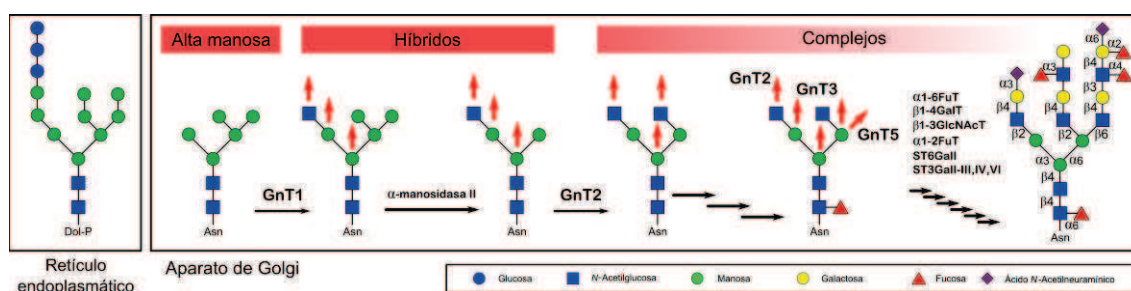


Figura I.3. Biosíntesis y diversidad estructural de *N*-glicanos. Se esquematizan *N*-glicanos de alta manosa, híbridos y complejos. A su vez se detallan las distintas enzimas que participan en la síntesis. Adviértase que en un *N*-glicano determinado no se encuentran todos los productos de las enzimas detalladas, ya que algunas compiten por un mismo sustrato. Dol-P, dolicol-fosfato; GnT1, 2, 3 y 5, GlcNAc transferasas; FuT, fucosil transferasas; GalT, galactosil transferasas; ST6Gal, α 2,6-sialil transferasas; ST3Gal, α 2,3-sialil transferasas.

Las vías de *O*-glicosilación comienzan con la unión de un monosacárido como α -GalNAc a un residuo de serina o treonina. A partir de este monosacárido pueden sintetizarse diversas estructuras como los core 1, core 2, core 3 y core 4 *O*-glicanos. A su vez estos core *O*-glicanos pueden ser extendidos para formar estructuras más complejas. Por lo tanto, el número de posibles sitios de *N*- y *O*-glicosilación en una

secuencia primaria, así como las enzimas presentes en el sistema RE/Golgi, son elementos clave que determinan el grado y tipo de glicosilación de una proteína [25].

Considerando la amplia diversidad y distribución que poseen los glicoconjugados (glicoproteínas, glicolípidos y proteoglicanos), resulta sencillo visualizar que estas estructuras median una amplia variedad de efectos biológicos. Estos efectos en general se pueden agrupar en dos tipos: (1) funciones estructurales del glicano en sí mismo o la modulación de la función de la molécula a la cual están covalentemente unidos y (2) funciones relacionadas al reconocimiento de la estructura sacarídica por proteínas con dominios de reconocimiento de carbohidratos (lectinas). Se han descrito diversos tipos de lectinas involucradas en la regulación de múltiples procesos biológicos (fertilización, embriogénesis y reconocimiento de patógenos). Respecto a su participación en procesos inmunológicos, se ha observado que diferentes familias de lectinas (colectinas, selectinas, siglecs, galectinas) cumplen papeles críticos en procesos diversos como reconocimiento de patógenos, activación y diferenciación linfocitaria y procesos de regulación homeostática. En el presente trabajo de Tesis se focalizó nuestra atención en la familia de las galectinas.

I.2.2 Descubrimiento de las lectinas animales

Las lectinas se definen como aquellas proteínas o glicoproteínas que unen carbohidratos y que no cumplen funciones de anticuerpo o de enzima [26]. En la **Tabla I.1** se muestran los principales descubrimientos en la investigación de lectinas que catapultaron el campo de la glicobiología a la era moderna de la ciencia. Aunque las lectinas fueron descubiertas hace más de 100 años en plantas [27], hoy en día sabemos que se encuentran presentes en toda la Naturaleza (incluso en el mundo de los microorganismos, donde reciben otros nombres como hemaglutininas, adhesinas, y toxinas [28-30]). Las interacciones dependientes de glicanos entre células animales fueron descubiertas al observar el fenómeno de reagregación de células de esponjas marinas disociadas, en una especie de reconocimiento específico entre ellas. Asimismo, se contaba con evidencia de la existencia de hemaglutininas (actividad de aglutinación de glóbulos rojos) en los fluidos corporales de varios crustáceos y arácnidos. Sin embargo, no fue hasta la década de los '60 que dichas interacciones fueron descritas en animales multicelulares como los vertebrados. La primera evidencia directa de lectinas de mamífero apareció fortuitamente durante el trabajo de G. Ashwell y col.,

quienes estudiaban los mecanismos que controlan el reciclaje de glicoproteínas en la circulación sanguínea [31-34]. El resultado fue el descubrimiento del “receptor de asialoglicoproteína”, un complejo proteico de la membrana de los hepatocitos que reconoce de manera específica residuos de β -Gal o GalNAc terminales presentes en glicoproteínas o células circulantes. Posteriormente, se reveló que la actividad de esta lectina se inhibe por el plasma de aves y reptiles. De este modo, se descubrió la lectina hepática de pollo, la cual presenta afinidad preferencial por residuos terminales de GlcNAc. Por lo tanto, los pollos tienen un receptor hepático GlcNAc-específico en lugar de uno específico para residuos de Gal/GalNAc, lo que permite la circulación de asialoglicoproteína, pero no así de glicoproteínas a las que se les han removido los residuos de Gal terminales. Un magnífico código de barras molecular de reconocimiento especie-dependiente.

Año	Descubrimiento
1888	Descripción de la primera fitoaglutinina (ricina) por parte de P.H. Stillmark en su Tesis doctoral en la Universidad de Dorpat (Estonia)
1891	P. Ehrlich aplica ricina y abrina (hemaglutinina tóxica de las semillas del regaliz americano <i>Abrus precatorius</i>) como modelo antigénico en estudios inmunológicos
1899	Descripción de la primera lectina animal de la glándula de albúmina del caracol <i>Helix pomatia</i> (M.L. Camus)
1902	Descripción de la primera aglutinina bacteriana (R. Kraus)
1907	Descripción de las aglutininas no tóxicas de planta (K. Landsteiner y H. Raubitschek)
1919	J.B. Sumner cristaliza concanavalina A del grano de <i>Canavalia ensiformis</i> , obteniendo por primera vez una hemaglutinina pura
1936	Primera evidencia de que los carbohidratos en la superficie de los glóbulos rojos podrían ser los ligando causantes de la hemaglutinación inducida por concanavalina A
1941	Primera descripción de una aglutinina viral (G.K. Hirst)
1947	Descripción de la especificidad de ciertas hemaglutininas por diferentes grupos sanguíneos (W.C. Boyd y K.O. Renkonen)
1952	Descubrimiento de la naturaleza sacarídica del grupo sanguíneo H por W.J.T. Morgan y W.M. Watkins, mediante el uso de aglutinina de anguila eléctrica
1954	W.C. Boyd y E. Shapleigh introducen el término “lectina” (del Latin <i>lectus</i> , elegir)
1959	N. Sharon descubre la presencia de amino-hexosas, primeros pasos en la demostración de los procesos de glicosilación de proteínas
1960	P.C. Nowell describe la actividad estimuladora de linfocitos de la fitohemaglutinina de la semilla de <i>Phaseolus vulgaris</i> (frijol)
1963	I.J. Goldstein introduce la cromatografía de afinidad para la purificación de lectinas
1972	Resolución por primera vez de la estructura 3D de una lectina: concanavalin A (G.M. Edelman, K.D. Hardman y C.F. Ainsworth)
1974	Primera obtención de un lectina de mamífero pura (receptor de asialoglicoproteína) a partir de hígado (G. Ashwell)
1980	I.J. Goldstein y col. definen a las lectinas como proteínas de unión a carbohidratos de un origen no inmunológico que aglutinan células
1988	S.H. Barondes amplía la definición de lectina: son proteínas de unión a carbohidratos diferentes a los anticuerpos y las enzimas

Tabla I.1. Principales descubrimientos en el área de las lectinas

A partir de estos descubrimientos, el ritmo de la investigación en lectinas se aceleró drásticamente, hecho que se debe principalmente a la introducción de la

cromatografía por afinidad por Agarwal y Goldstein [35]. Desde el moho mucilaginoso *Dictyostelium discoideum* hasta tejidos de mamíferos, se han descubierto una gran variedad de lectinas específicas por galactosa. La mayoría de ellas demostraron ser muy diferentes del receptor original de asialoglicoproteína, siendo solubles en agua y de menor peso molecular. Hoy en día, dichas lectinas se conocen como galectinas. Mientras tanto, Hill y col. reportaban otro tipo de sistema hepático de reconocimiento basado en residuos de fucosa [36-39]. A comienzos de los '70, Neufeld y col. reportaron un sistema de captura de enzimas lisosomales a partir de fibroblastos dependiente de carbohidratos [40, 41]. En 1977, Sly y col. demostraron el bloqueo específico de este sistema por parte del monosacárido manosa-6-fosfato (Man-6-P). En conjunto con los trabajos de los grupos de Kornfeld, Jourdain, Von Figura y otros se arribó al descubrimiento de los receptores de Man-6-P, los cuales reconocen oligosacáridos fosforilados de alta manosa que son expresados selectivamente en las enzimas lisosomales. Usando modificaciones de este tipo particular de sistemas de captura y rastreo de enzimas glicosiladas se descubrió el receptor de manosa en macrófagos [42, 43]. A comienzo de los '80, el concepto de lectinas en animales capaces de reconocer ligandos sacarídicos endógenos era un hecho firmemente establecido. Durante la preparación de esta Tesis, una búsqueda en PubMed en el sitio de Internet del National Center of Biotechnology Information (www.ncbi.nlm.nih.gov) arrojó casi 96.000 resultados para la búsqueda de “lectins”, mostrando el crecimiento exponencial del campo.

I.2.3 Clasificación actual de las lectinas animales

Luego de su descubrimiento, las lectinas animales fueron clasificadas de acuerdo con las secuencias de carbohidratos a las cuales se unían preferencialmente (por ejemplo, lectinas de unión a β -galactósidos). Pero no fue hasta el advenimiento de las técnicas modernas de la biología molecular y la cristalización de proteínas que una clasificación más consistente se desarrollaría, basándose en la homología de secuencias de aminoácidos y relación evolutiva de dichas lectinas (**Tabla I.2** y **Figura I.4**). La primera de dichas clasificaciones fue propuesta por Drickamer [44], analizando dos grupos mayoritarios de lectinas descubiertos hasta la fecha: el primer grupo requería Ca^{2+} para el reconocimiento y fue así llamado **lectinas del tipo-C**, y el otro grupo requería los tioles “libres” (reducidos) de los residuos de cisteína para ser estables y se

denominó **lectinas del tipo-S** (actualmente **galectinas**). Las lectinas que reconocían Man-6-P se secuenciaron y se encontraron homólogas pero distintas a los otros dos grupos, por lo que se agruparon como **lectinas del tipo-P** (o también M6PR, del inglés *Man-6-P receptors*). Mientras que las lectinas del tipo P y S reconocen una clase particular de azúcares (Man-6-P y β -galactósidos, respectivamente), las lectinas del tipo-C agrupan proteínas con diferente especificidad por glicanos aunque similares motivos estructurales. Un descubrimiento clave ocurrió cuando al clonar en forma independiente tres receptores de adhesina vascular homólogos, éstos mostraron un motivo estructural del tipo-C en su extremo *N*-terminal. Hoy en día estos tres dominios forman parte de la familia de **selectinas** (**Figura I.4**). Del mismo modo, las colectinas y dos lectinas que unen calcio (calnexina y calreticulina) actualmente se incorporan también al grupo de las lectinas del tipo-C. Estudios en la década de los '90 revelaron que algunos miembros de la superfamilia de inmunoglobulinas pueden reconocer glicanos, dando origen así a un nuevo grupo llamado **lectinas del tipo-I**. Un subgrupo de estas lectinas que reconocen específicamente ácidos siálicos ha sido recientemente designado como las "**siglecs**" (del inglés *sialic acid/immunoglobulin superfamily/lectins*). Otra clase de lectinas solubles evolutivamente muy antiguas llamadas **pentraxinas** se incluye como grupo no por su homología de secuencia, sino por una organización estructural pentamérica consistente y por un probable rol común en la respuesta inmune primaria del huésped frente a infecciones [45, 46].

Grupo de Lectina	Ca ²⁺	Especificidad	Características estructurales que las definen
Tipo-C (selectinas, colectinas,)	☑	Variable	Motivo tipo-C
Tipo-S (galectinas)	☒	β-galactósidos	Motivo tipo-S
Tipo-P (MPRs)	Variable	Man-6-P	Motivo tipo-P
Tipo-I (siglecs)	☒	Variable	Dominios tipo-Ig
Pentraxinas	Mayormente	Galactósidos	Motivos multiméricos
Calnexina, calreticulina	☑	Glc ₁ Man ₉ en el RE	Motivo de calnexina
Tipo de unión a heparina	☒	Heparina/heparan-SO ₄ ²⁻	Aminoácidos básicos en cluster
Tipo-F	Variable	L-fucosa	Motivo tipo-F
Tipo-M	☑	Man ₈	Motivo tipo-M
Tipo-L	☑	Variable	Motivo tipo-L
Tipo-R	☒	Variable	Motivo tipo-R
Caja-F	☒	GlcNAc ₂	Motivo caja-F
Ficolinas	☑	GlcNAc, GalNAc	Motivo de ficolina
Tipo-quitinasa (quilectinas)	☒	Quito-oligosacáridos	Motivo de triosa-P isomerasa
Tipo-X (intelectinas)	☑	Gal, galactofuranosa, pentosas	Motivo de intelectina

Tabla I.2. Clasificación de las lectinas animales

Esta agrupación está basada principalmente en homología de secuencias y probables relaciones evolutivas, e incluye a la mayoría de las lectinas animales hasta

hoy conocidas. Sin embargo muchas otras, como las proteínas de unión a glicosaminoglicanos (GAG), y muchas nuevas “lectinas”, no forman parte de ninguna de las familias *bona fide* hasta aquí descritas. Las mismas constituyen un grupo de proteínas que quedan por estudiar y brindar nuevos paradigmas de relación estructura-función en el campo de la glicobiología.

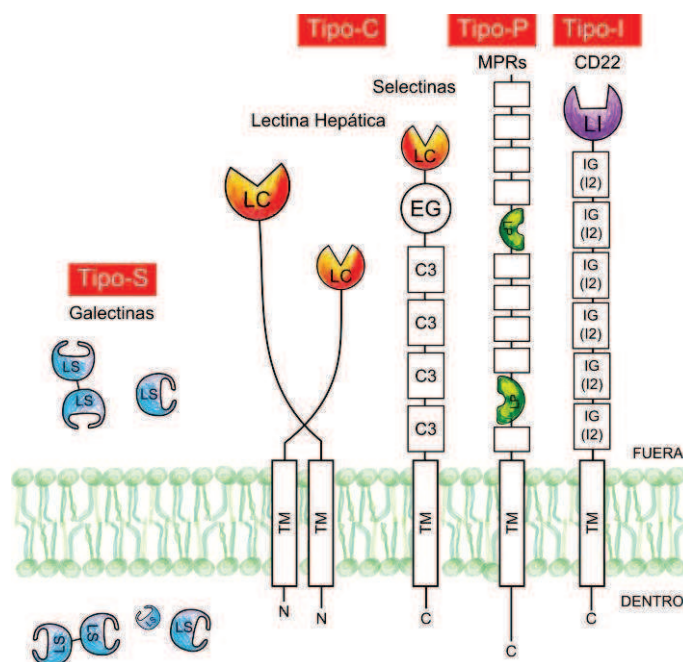


Figura I.4. Ejemplos esquemáticos de las lectinas animales, basados en la estructura y topología. LC, lectina tipo-C; LS, lectina tipo-S; LP, lectina tipo-P; LI, lectina tipo-I. Otros dominios marcados son: EG, tipo-EGF (endotelial growth factor); IG (I2), dominio tipo inmunoglobulina-2; TM, dominio transmembrana, C3, dominio regulatorio del complemento.

I.2.4 Galectinas: introducción

Luego del descubrimiento de las aglutininas en plantas y de las lectinas en *Dictyostelium discoideum* (discoidina I) a principios de los '70 por Barondes y Rosen y col., muchos investigadores salieron a la búsqueda de lectinas en tejidos animales. Como se indicó en el apartado anterior, la primera lectina animal reportada fue el receptor de asialoglicoproteína hepático, una lectina del tipo-C. La siguiente lectina descubierta en animales fue una proteína reconocida hoy como la primera del tipo-S, la primera galectina. Esta fue originalmente descrita por Teichberg y col. durante sus estudios sobre la posible presencia de lectinas en los órganos eléctricos de las anguilas eléctricas [47]. La proteína, nombrada por aquel entonces electrolectina, poseía actividad de hemaglutinina inhibida por β -galactósidos y pudo ser aislada por

cromatografía de afinidad en soportes funcionalizados con β -galactósidos. Sorprendentemente, la correcta actividad de esta lectina requería la presencia de β -mercaptoetanol en los buffers de corrida, sugiriendo la presencia de una o más cisteínas en forma libre (reducida). La electrolectina tiene un peso de aproximadamente 15 kDa y se la encuentra como un homodímero no covalente.

La primera lectina del tipo-S hallada en mamíferos, hoy en día llamada galectina-1, fue aislada en 1976 por Kornfeld y col. a partir de extractos de corazón y pulmón de ternero [48]. Al igual que con la electrolectina de Teichberg, se requirieron lactosa y condiciones reductoras para su correcta extracción y purificación a partir de dichos tejidos. En el mismo año, Barondes y col. identificaron también una lectina del tipo-S en extractos de músculo de pollo [49]. A comienzos de 1980, Wang y col. hallaron una proteína de unión a carbohidratos de 35 kDa de características similares, hoy conocida como galectina-3, en fibroblastos humanos y murinos que también tiene afinidad por β -galactósidos [50]. La naturaleza soluble de todas estas proteínas, la homología de secuencias y los requerimientos por condiciones reductoras e independencia de Ca^{2+} para retener su actividad, permitieron a Drickamer agruparlas como lectinas del tipo-S para remarcar su dependencia de -SH libres y su solubilidad. La nomenclatura para las galectinas ha sido sistematizada en 1994 [51]. La primera galectina hallada (electrolectina, lectina de unión a β -galactósidos, galaptina, L-14, etc., dependiendo del origen) se llama **galectina-1** (Gal-1). Su homólogo más cercano, **galectina-2** (Gal-2). La lectina de 35 kDa se llamó **galectina-3** (Gal-3) y los otros miembros de esta familia fueron así numerados en forma consecutiva según su orden de descubrimiento.

A la fecha, se han identificado 15 miembros en la familia de las galectinas en vertebrados superiores [52], mostrando diferentes tipos de arquitecturas de un dominio de reconocimiento de carbohidratos (DRC) canónico de aproximadamente 130 aminoácidos [53] (**Figura I.5**). Las Gal-1, -2, -5, -7, -10, -11, -13, -14 y -15 son ejemplos de galectinas **prototipo**, las cuales se pueden encontrar en forma monomérica o como homodímero no covalentemente unido. Gal-4, -6, -8, -9 y -12 son las llamadas galectinas de **repeticiones en tándem**, donde cada una de ellas se encuentra formada por dos DRC diferentes pero vinculados por un péptido conector, haciéndolas dímeros permanentes. Finalmente Gal-3 es la única galectina tipo **quimera**, ya que es la única que posee tres dominios estructurales distintivos: un dominio no relacionado (BH3) de

12 aminoácidos *N*-terminal, una secuencia del tipo colágeno, rica en prolina y glicina y un dominio *C*-terminal que contiene el DRC característico [54].

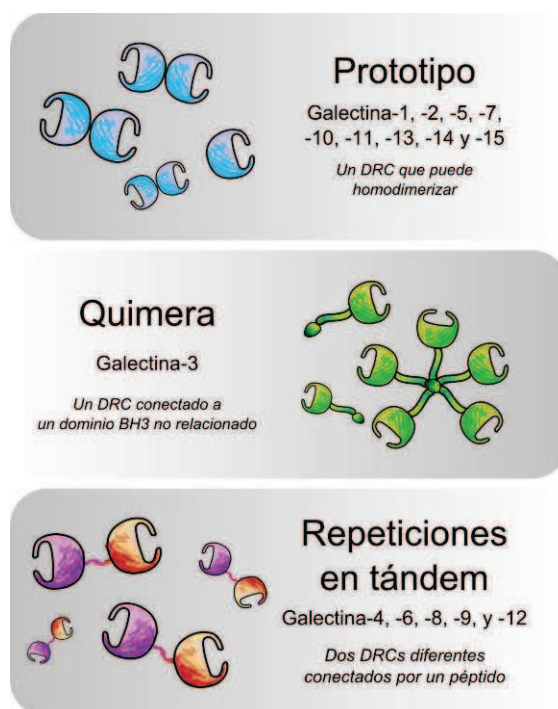


Figura I.5. Clasificación estructural de las galectinas. De acuerdo a su estructura bioquímica las galectinas se clasifican en tres grupos diferentes. Las galectinas prototipo contienen un sitio de reconocimiento de carbohidratos (DRC) por subunidad que pueden homodimerizar. Galectina-3 (la única tipo quimera) posee un DRC en el extremo *C*-terminal y un dominio no relacionado (BH3) en el extremo *N*-terminal. Este último es responsable de los efectos intracelulares de esta lectina. Las galectinas del tipo “repeticiones en tándem” son un grupo de galectinas bivalentes formadas por dos monómeros con CRDs de distinta afinidad unidos por un péptido conector.

Todos los miembros de la familia de las galectinas carecen de péptido señal o dominios de anclaje a membrana y parecen ser sintetizadas en polisomas libres en el citoplasma y acumularse allí previos a su secreción [55, 56]. Las galectinas recientemente sintetizadas en el citosol celular han demostrado ser funcionales en cuanto a la unión de β -galactósidos. La complejidad de la biosíntesis, secreción y oligomerización de las galectinas se encuentra ilustrada en la **Figura I.6**. Curiosamente, la exportación de las galectinas al exterior celular no involucra en forma directa el movimiento a través del aparato secretor de la célula. El proceso por el cual estas proteínas son secretadas ha sido explorado en varias ocasiones [57-60], pero el mecanismo básico es actualmente desconocido. Se conoce de otros factores de crecimiento pequeños y ciertos componentes de la matriz extracelular que son excretados mediante mecanismos no clásicos, pero si se tratan de aquellos utilizados por las galectinas, es todavía incierto. Por otro lado, la mayoría de las galectinas pierden rápidamente su actividad en un ambiente no-reductor, como el que existe

fuera de la célula en condiciones normales. Sin embargo, algunos estudios sobre la biosíntesis de Gal-1 hallaron que la proteína recientemente exportada es muy inestable en ausencia de ligandos, pero que cuando se encuentran disponibles ligandos de alta afinidad en el medio la proteína es estable. Así, la secreción regulada y la disponibilidad de ligandos podrían regular la actividad y estabilidad de las galectinas a lo largo de su vida intra y extracelular.

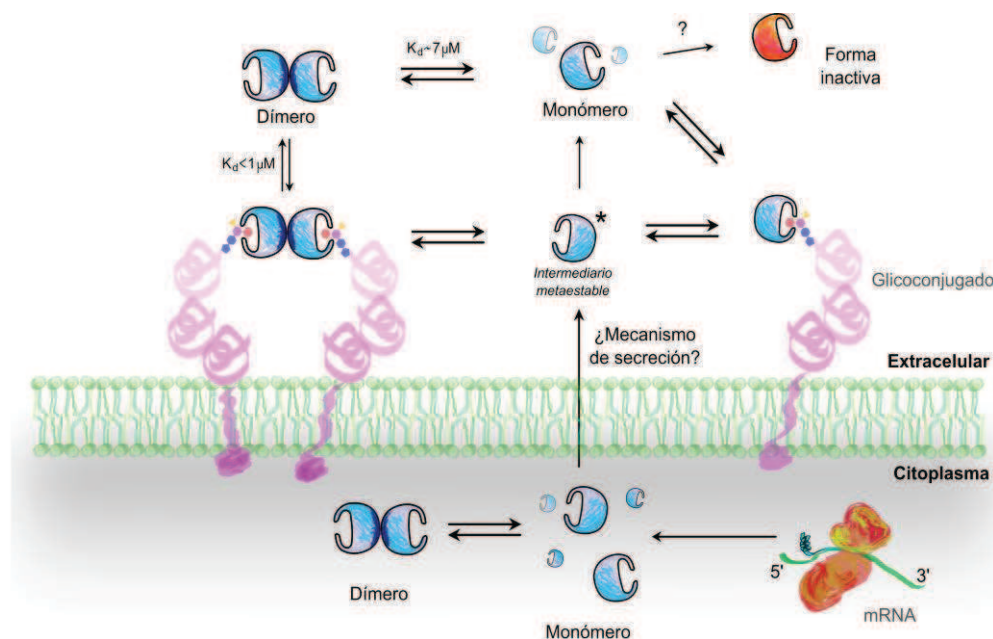


Figura I.6. Rutas biosintéticas posibles de Gal-1. La proteína es traducida en polisomas libres en el citoplasma y así como se sintetiza, tiene capacidad de unión a hidratos de carbono u otras proteínas en la célula. Luego de su secreción o exportación por un mecanismo aún desconocido, la galectina recientemente sintetizada es inestable, pero se asocia con glicanos para aumentar su estabilidad frente a la inactivación por variados y también poco explorados mecanismos. La proteína monomérica estable puede dimerizar e interactuar con ligandos en la superficie de la célula y en la matriz extracelular.

Mientras que algunas galectinas (como Gal-1 y -3) son ampliamente expresadas en diferentes tejidos y compartimientos celulares de diferentes especies animales [20, 61-64], otros miembros de la familia tienen una localización mucho más restringida. Por ejemplo, Gal-7 se encuentra preferencialmente en la piel [65-67], Gal-4 en el tracto gastrointestinal [68, 69], Gal-5 sólo se expresa en reticulocitos de rata [70, 71], Gal-12 en tejido adiposo [72-75] y Gal-10 está presente en eosinófilos humanos, pero no así en los murinos [76]. A pesar de esta distribución todas, de una u otra manera, se encuentran involucradas en procesos de destino y desarrollo celular [77], homeostasis de la respuesta inmune [20], progresión tumoral [78] y neovascularización [79, 80]. Incluso una sola galectina puede afectar distintas células dependiendo del tipo y las circunstancias. Por ejemplo, Gal-1 puede estimular o inhibir la proliferación celular [81-

83] así como también algunos procesos de adhesión celular a la matriz extracelular [84, 85]. Existe también evidencia de que las galectinas pueden simultáneamente poseer diferentes funciones extra e intracelulares. Ejemplo de ello son los casos donde Gal-1 y -3 demostraron estar implicadas en procesos de splicing del pre-mRNA [86, 87]. Un resumen de las principales funciones de cada galectina se puede encontrar en la **Figura I.7**.

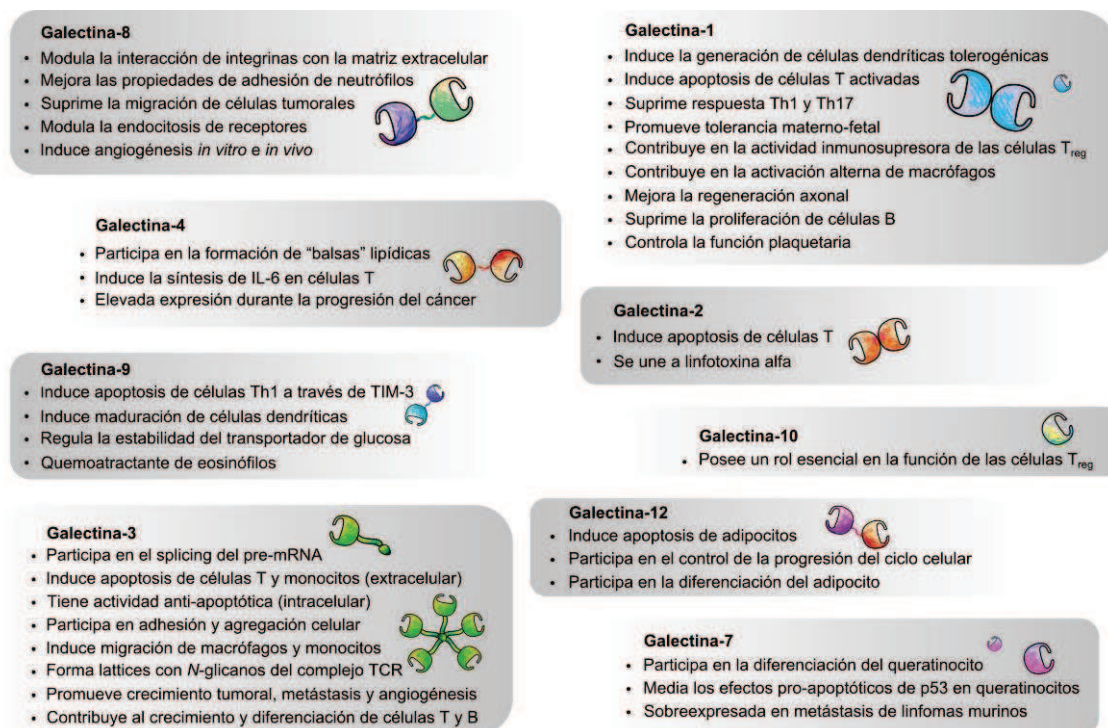


Figura I.7. Efectos de diversas galectinas en la fisiología del sistema inmune y los tumores. Diferentes funciones han sido atribuidas a los diversos miembros de esta familia de proteínas de unión a carbohidratos. En la figura se resumen algunas de las más relevantes. (Adaptado de [61]).

La presencia de tantas galectinas en gran cantidad de especies evolutivamente divergentes sugiere, por un lado, un rol en procesos y funciones celulares básicas para los organismos vivos. Por otra parte, la evidencia de que hay docenas de galectinas dentro de una misma especie implica que han evolucionado para participar en una variedad importante de funciones cada vez más específicas, tal como se ha puesto en evidencia en el párrafo anterior. Sin embargo, a pesar del crecimiento exponencial del área de estudio de estas lectinas, la comprensión de la relación entre la estructura y las funciones bioquímicas ha sido escasamente explorada. La naturaleza molecular del reconocimiento de cada glicano por cada galectina, la especificidad de las señales intracelulares observadas y las diferentes funciones que se desencadenan por estas

interacciones representan un conjunto grande de preguntas abiertas en un campo en constante crecimiento.

I.3 Objetivos y lineamientos

En este trabajo de Tesis se estudió la familia completa de las galectinas humanas utilizando un enfoque multidisciplinario. Basándonos en la hipótesis general de que la estructura y dinámica de una proteína determina gran parte de su función, nos abocamos al estudio comparativo de esta familia con el fin de aportar una visión estructural a los mecanismos y determinantes moleculares de regulación de la unión a glicanos en galectinas y posterior activación de su función en diversos cultivos celulares. Para ello, se emplearon una variedad de simulaciones de dinámica molecular clásica así como también experimentos de mesada como ser la expresión y purificación de proteínas recombinantes y ensayos de unión a ligando, cinéticas de oxidación y apoptosis *in vitro* mediante citometría de flujo y diversas técnicas de microscopía óptica y electrónica.

Luego de esta introducción general, se incluye un apartado de Metodología en el cual se presenta una descripción de los métodos y materiales utilizados. Como todo estudio moderno de la biología estructural, los métodos escogidos pueden clasificarse en dos grandes áreas: aquellos que proveen una representación de los sistemas moleculares dinámicos a partir de simulaciones, y aquellos estrictamente experimentales. En el primer caso, se ofrece una descripción de los métodos de simulación computacional; en el segundo, se tratan brevemente los fundamentos de la expresión y purificación de Gal-1 recombinante, proteína modelo utilizada en los experimentos, así como también una breve introducción a las técnicas espectroscópicas de fluorescencia y dicroísmo circular. Los resultados obtenidos durante el trabajo de Tesis se desarrollarán en tres capítulos (III a V). Dada la diversidad de técnicas y disciplinas que se exploran en cada capítulo y con el fin de mantener siempre en claro el origen de la problemática, se decidió incluir una introducción específica en cada uno de esos capítulos para guiar al lector en los antecedentes más importantes del tema a desarrollar y las metodologías y ensayos particulares empleados en dicha sección para explorar las hipótesis de estudio. Por último, en el capítulo VI se resumen las principales conclusiones de esta Tesis, junto a las perspectivas futuras de los temas investigados.

Como el título de esta Tesis bien define, es mi deseo que durante la lectura de este trabajo uno pueda recorrer y comprender con mayor profundidad el mundo de las galectinas: desde cada átomo de la proteína, hasta su solvatación e interacción con los de los glicanos, pasando por agentes reductores y oxidantes a los que se enfrenta, y finalmente visualizarla interactuando en cada célula, siempre volviendo a la estructura como si tuviéramos un resorte que nos atrae nuevamente a ella.

I.4 Referencias

1. Hendrickson, W.A. and Klug, S.A. (1991) Editorial. *Current Opinion in Structural Biology* 1, 1-2
2. Rahman, A. (1964) Correlations in the Motion of Atoms in Liquid Argon. *Physical Review* 136, A405-A411
3. Stillinger, F.H. and Rahman, A. (1974) Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics* 60, 1545-57
4. Alder, B.J. and Wainwright, T.E. (1959) Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics* 31, 459-66
5. Alder, B.J. and Wainwright, T.E. (1957) Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* 27, 1208-09
6. Piana, S., Lindorff-Larsen, K., and Shaw, D.E. (2013) Atomic-level description of ubiquitin folding. *Proceedings of the National Academy of Sciences* 110, 5915-20
7. Stone, J.E., Phillips, J.C., Freddolino, P.L., Hardy, D.J., Trabuco, L.G., and Schulten, K. (2007) Accelerating molecular modeling applications with graphics processors. *Journal of Computational Chemistry* 28, 2618-40
8. Yasuda, K. (2008) Accelerating Density Functional Calculations with Graphics Processing Unit. *Journal of Chemical Theory and Computation* 4, 1230-36
9. Amisaki, T., Toyoda, S., Miyagawa, H., and Kitamura, K. (2003) Development of hardware accelerator for molecular dynamics simulations: A computation board that calculates nonbonded interactions in cooperation with fast multipole method. *Journal of Computational Chemistry* 24, 582-92
10. Shaw, D.E., Deneroff, M.M., Dror, R.O., Kuskin, J.S., Larson, R.H., Salmon, J.K., Young, C., Batson, B., Bowers, K.J., Chao, J.C., Eastwood, M.P., Gagliardo, J., Grossman, J.P., Ho, C.R., Ierardi, D.J., Kolossváry, I., Klepeis, J.L., Layman, T., McLeavey, C., Moraes, M.A., Mueller, R., Priest, E.C., Shan, Y., Spengler, J., Theobald, M., Towles, B., and Wang, S.C. (2008) Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* 51, 91-97
11. Leach, A.R. (2001) *Molecular Modelling: Principles and Applications* Pearson Education EMA
12. Chothia, C. (1992) One thousand families for the molecular biologist. *Nature* 357, 543-44
13. Gibrat, J.-F., Madej, T., and Bryant, S.H. (1996) Surprising similarities in structure comparison. *Current Opinion in Structural Biology* 6, 377-85
14. Bowie, J.U., Luthy, R., and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164-70
15. Mount, D.W. (2001) *Bioinformatics. Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press
16. Bracegirdle, B. (1989) Microscopy and comprehension: the development of understanding of the nature of the cell. *Trends in Biochemical Sciences* 14, 464-68
17. Campbell, I.D. (2002) The march of structural biology. *Nat Rev Mol Cell Biol* 3, 377-81
18. *Essentials of Glycobiology*. (2009) Editado por Ajit Varki, Richard D Cummings, Jeffrey D Esko, Hudson H Freeze, Pamela Stanley, Carolyn R Bertozzi, Gerald W Hart, and Marilynn E Etzler. Cold Spring Harbor Laboratory Press
19. Rademacher, T.W., Parekh, R.B., and Dwek, R.A. (1988) Glycobiology. *Annual Review of Biochemistry* 57, 785-838

20. Rabinovich, G.A. and Toscano, M. (2009) Turning "sweet" on immunity: Galectin-glycan interactions in immune tolerance and inflammation. *Nature Rev Immunol* 9, 338-52
21. van Kooyk, Y. and Rabinovich, G.A. (2008) Protein-glycan interactions in the control of innate and adaptive immune responses. *Nat Immunol* 9, 593-601
22. Ohtsubo, K. and Marth, J.D. (2006) Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* 126, 855-67
23. Bertozzi, C.R., Kiessling, and Laura, L. (2001) Chemical Glycobiology. *Science* 291, 2357-64
24. Paulson, J.C., Blixt, O., and Collins, B.E. (2006) Sweet spots in functional glycomics. *Nat Chem Biol* 2, 238-48
25. Gagneux, P. and Varki, A. (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9, 747-55
26. Barondes, S.H. (1988) Bifunctional properties of lectins: lectins redefined. *Trends in Biochemical Sciences* 13, 480-82
27. Boyd, W.C. and Shapleigh, E. (1954) Specific Precipitating Activity of Plant Agglutinins (Lectins). *Science* 119, 419-19
28. Karlsson, K.-A. (1995) Microbial recognition of target-cell glycoconjugates. *Current Opinion in Structural Biology* 5, 622-35
29. Patti, J.M. and Höök, M. (1994) Microbial adhesins recognizing extracellular matrix macromolecules. *Current Opinion in Cell Biology* 6, 752-58
30. Mouricout, M. (1997) Interactions between the Enteric Pathogen and the Host. In *Mechanisms in the Pathogenesis of Enteric Diseases* (Paul, P., et al., eds), pp. 109-23, Springer US
31. Hudgin, R.L., Pricer, W.E., Ashwell, G., Stockert, R.J., and Morell, A.G. (1974) The Isolation and Properties of a Rabbit Liver Binding Protein Specific for Asialoglycoproteins. *Journal of Biological Chemistry* 249, 5536-43
32. van den Hamer, C.J.A., Morell, A.G., Scheinberg, I.H., Hickman, J., and Ashwell, G. (1970) Physical and Chemical Studies on Ceruloplasmin: IX. THE ROLE OF GALACTOSYL RESIDUES IN THE CLEARANCE OF CERULOPLASMIN FROM THE CIRCULATION. *Journal of Biological Chemistry* 245, 4397-402
33. Morell, A.G., Gregoriadis, G., Scheinberg, I.H., Hickman, J., and Ashwell, G. (1971) The Role of Sialic Acid in Determining the Survival of Glycoproteins in the Circulation. *Journal of Biological Chemistry* 246, 1461-67
34. Ashwell, G. and Harford, J. (1982) Carbohydrate-Specific Receptors of the Liver. *Annual Review of Biochemistry* 51, 531-54
35. Agrawal, B.B.L. and Goldstein, I.J. (1965) Specific Binding of Concanavalin A to Cross-Linked Dextran Gels. *Biochem J* 96, 23C-25C
36. Haltiwanger, R.S., Lehrman, M.A., Eckhardt, A.E., and Hill, R.L. (1986) The distribution and localization of the fucose-binding lectin in rat tissues and the identification of a high affinity form of the mannose/N-acetylglucosamine-binding lectin in rat liver. *Journal of Biological Chemistry* 261, 7433-39
37. Lehrman, M.A., Haltiwanger, R.S., and Hill, R.L. (1986) The binding of fucose-containing glycoproteins by hepatic lectins. The binding specificity of the rat liver fucose lectin. *Journal of Biological Chemistry* 261, 7426-32
38. Lehrman, M.A. and Hill, R.L. (1986) The binding of fucose-containing glycoproteins by hepatic lectins. Purification of a fucose-binding lectin from rat liver. *Journal of Biological Chemistry* 261, 7419-25
39. Lehrman, M.A., Pizzo, S.V., Imber, M.J., and Hill, R.L. (1986) The binding of fucose-containing glycoproteins by hepatic lectins. Re-examination of the clearance from blood and the binding to membrane receptors and pure lectins. *Journal of Biological Chemistry* 261, 7412-18
40. Sando, G.N. and Neufeld, E.F. (1977) Recognition and receptor-mediated uptake of a lysosomal enzyme, $\pm$ -L-iduronidase, by cultured human fibroblasts. *Cell* 12, 619-27
41. Rome, L.H., Weissmann, B., and Neufeld, E.F. (1979) Direct demonstration of binding of a lysosomal enzyme, alpha-L-iduronidase, to receptors on cultured fibroblasts. *Proceedings of the National Academy of Sciences* 76, 2331-34
42. Tietze, C., Schlesinger, P., and Stahl, P. (1982) Mannose-specific endocytosis receptor of alveolar macrophages: demonstration of two functionally distinct intracellular pools of receptor and their roles in receptor recycling. *The Journal of Cell Biology* 92, 417-24

43. Stahl, P. and Gordon, S. **(1982)** Expression of a mannosyl-fucosyl receptor for endocytosis on cultured primary macrophages and their hybrids. *The Journal of Cell Biology* 93, 49-56
44. Drickamer, K. **(1988)** Two distinct classes of carbohydrate-recognition domains in animal lectins. *Journal of Biological Chemistry* 263, 9557-60
45. Inforzato, A., Bottazzi, B., Garlanda, C., Valentino, S., and Mantovani, A. (2012) Pentraxins in Humoral Innate Immunity. In *Current Topics in Innate Immunity II* (Lambris, J.D. and Hajishengallis, G., eds), pp. 1-20, Springer New York
46. Du Clos, T.W. **(2013)** Pentraxins: Structure, Function, and Role in Inflammation. *ISRN Inflammation* 2013, 22
47. Teichberg, V.I., Silman, I., Beitsch, D.D., and Resheff, G. **(1975)** A beta-D-galactoside binding protein from electric organ tissue of *Electrophorus electricus*. *Proceedings of the National Academy of Sciences* 72, 1383-87
48. de Waard, A., Hickman, S., and Kornfeld, S. **(1976)** Isolation and properties of beta-galactoside binding lectins of calf heart and lung. *Journal of Biological Chemistry* 251, 7581-87
49. Nowak, T.P., Haywood, P.L., and Barondes, S.H. **(1976)** Developmentally regulated lectin in embryonic chick muscle and a myogenic cell line. *Biochemical and Biophysical Research Communications* 68, 650-57
50. Roff, C.F. and Wang, J.L. **(1983)** Endogenous lectins from cultured cells. Isolation and characterization of carbohydrate-binding proteins from 3T3 fibroblasts. *Journal of Biological Chemistry* 258, 10657-63
51. Barondes, S.H., Cooper, D.N., Gitt, M.A., and Leffler, H. **(1994)** Galectins. Structure and function of a large family of animal lectins. *Journal of Biological Chemistry* 269, 20807-10
52. Ahmed, H., Allen, H.J., Sharma, A., and Matta, K.L. **(1990)** Human splenic galactin: carbohydrate-binding specificity and characterization of the combining site. *Biochemistry* 29, 5315-19
53. Cooper, D.N.W. **(2002)** Galectinomics: finding themes in complexity. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1572, 209-31
54. Gong, H.C., Honjo, Y., Nangia-Makker, P., Hogan, V., Mazurak, N., Bresalier, R.S., and Raz, A. **(1999)** The NH₂ Terminus of Galectin-3 Governs Cellular Compartmentalization and Functions in Cancer Cells. *Cancer Research* 59, 6239-45
55. Sato, S., Burdett, I., and Hughes, R.C. **(1993)** Secretion of the Baby Hamster Kidney 30-kDa Galactose-Binding Lectin from Polarized and Nonpolarized Cells: A Pathway Independent of the Endoplasmic Reticulum-Golgi Complex. *Experimental Cell Research* 207, 8-18
56. Kuchler, K. and Thorner, J. **(1992)** Secretion of Peptides and Proteins Lacking Hydrophobic Signal Sequences: The Role of Adenosine Triphosphate-Driven Membrane Translocators. *Endocrine Reviews* 13, 499-514
57. Cho, M. and Cummings, R.D. **(1995)** Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. I. Physical and chemical characterization. *Journal of Biological Chemistry* 270, 5198-206
58. Cooper, D.N. and Barondes, S.H. **(1990)** Evidence for export of a muscle lectin from cytosol to extracellular matrix and for a novel secretory mechanism. *The Journal of Cell Biology* 110, 1681-91
59. Mehul, B. and Hughes, R.C. **(1997)** Plasma membrane targetting, vesicular budding and release of galectin 3 from the cytoplasm of mammalian cells during secretion. *Journal of Cell Science* 110, 1169-78
60. Cleves, A.E., Cooper, D.N., Barondes, S.H., and Kelly, R.B. **(1996)** A new pathway for protein export in *Saccharomyces cerevisiae*. *The Journal of Cell Biology* 133, 1017-26
61. Yang, R.Y., Rabinovich, G.A., and Liu, F.T. **(2008)** Galectins: structure, function and therapeutic potential. *Expert Rev Mol Med* 10, e17
62. Camby, I., Le Mercier, M., Lefranc, F., and Kiss, R. **(2006)** Galectin-1: a small protein with major functions. *Glycobiology* 16, 137R-57R
63. Domic, J., Dabelic, S., and Flögel, M. **(2006)** Galectin-3: An open-ended story. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1760, 616-35
64. Rabinovich, G.A., Toscano, M., Jackson, D.A., and Vasta, G. **(2007)** Functions of cell surface galectin-glycoprotein lattices. *Curr Opin Struct Biol* 17, 513-20

65. Bernerd, F., Sarasin, A., and Magnaldo, T. (1999) Galectin-7 overexpression is associated with the apoptotic process in UVB-induced sunburn keratinocytes. *Proceedings of the National Academy of Sciences* 96, 11329-34
66. Sato, M., Nishi, N., Shoji, H., Kumagai, M., Imaizumi, T., Hata, Y., Hirashima, M., Suzuki, S., and Nakamura, T. (2002) Quantification of Galectin-7 and Its Localization in Adult Mouse Tissues. *Journal of Biochemistry* 131, 255-60
67. Gendronneau, G., Sidhu, S.S., Delacour, D., Dang, T., Calonne, C., Houzelstein, D., Magnaldo, T., and Poirier, F. (2008) Galectin-7 in the Control of Epidermal Homeostasis after Injury. *Molecular Biology of the Cell* 19, 5541-49
68. Chiu, M.L., Parry, D.A., Feldman, S.R., Klapper, D.G., and O'Keefe, E.J. (1994) An adherens junction protein is a member of the family of lactose-binding lectins. *Journal of Biological Chemistry* 269, 31770-76
69. Danielsen, E.M. and van Deurs, B. (1997) Galectin-4 and Small Intestinal Brush Border Enzymes Form Clusters. *Molecular Biology of the Cell* 8, 2241-51
70. Barrès, C., Blanc, L., Bette-Bobillo, P., André, S., Mamoun, R., Gabius, H.-J., and Vidal, M. (2010) Galectin-5 is bound onto the surface of rat reticulocyte exosomes and modulates vesicle uptake by macrophages. *Blood* 115, 696-705
71. Gitt, M.A., Wiser, M.F., Leffler, H., Herrmann, J., Xia, Y.-R., Massa, S.M., Cooper, D.N.W., Lusi, A.J., and Barondes, S.H. (1995) Sequence and Mapping of Galectin-5, a β -Galactoside-binding Lectin, Found in Rat Erythrocytes. *Journal of Biological Chemistry* 270, 5032-38
72. Yang, R.-Y., Hsu, D.K., Yu, L., Ni, J., and Liu, F.-T. (2001) Cell Cycle Regulation by Galectin-12, a New Member of the Galectin Superfamily. *Journal of Biological Chemistry* 276, 20252-60
73. Yang, R.-Y., Hsu, D.K., Yu, L., Chen, H.-Y., and Liu, F.-T. (2004) Galectin-12 Is Required for Adipogenic Signaling and Adipocyte Differentiation. *Journal of Biological Chemistry* 279, 29761-66
74. Yang, R.-Y., Havel, P.J., and Liu, F.-T. (2012) Galectin-12: A protein associated with lipid droplets that regulates lipid metabolism and energy balance. *Adipocyte* 1, 96-100
75. Hotta, K., Funahashi, T., Matsukawa, Y., Takahashi, M., Nishizawa, H., Kishida, K., Matsuda, M., Kuriyama, H., Kihara, S., Nakamura, T., Tochino, Y., Bodkin, N.L., Hansen, B.C., and Matsuzawa, Y. (2001) Galectin-12, an Adipose-expressed Galectin-like Molecule Possessing Apoptosis-inducing Activity. *Journal of Biological Chemistry* 276, 34089-97
76. Swaminathan, G.J., Leonidas, D.D., Savage, M.P., Ackerman, S.J., and Acharya, K.R. (1999) Selective Recognition of Mannose by the Human Eosinophil Charcot-Leyden Crystal Protein (Galectin-10): A Crystallographic Study at 1.8 Å Resolution. *Biochemistry* 38, 13837-43
77. Hernandez, J.D. and Baum, L.G. (2002) Ah, sweet mystery of death! Galectins and control of cell fate. *Glycobiology* 12, 127R-36R
78. Liu, F.-T. and Rabinovich, G.A. (2005) Galectins as modulators of tumour progression. *Nat Rev Cancer* 5, 29-41
79. Cárdenas Delgado, V.M., Nugnes, L.G., Colombo, L.L., Troncoso, M.F., Fernández, M.M., Malchiodi, E.L., Frahm, I., Croci, D.O., Compagno, D., Rabinovich, G.A., Wolfenstein-Todel, C., and Elola, M.T. (2011) Modulation of endothelial cell migration and angiogenesis: a novel function for the "tandem-repeat" lectin galectin-8. *The FASEB Journal* 25, 242-54
80. Markowska, A.I., Jefferies, K.C., and Panjwani, N. (2011) Galectin-3 Protein Modulates Cell Surface Expression and Activation of Vascular Endothelial Growth Factor Receptor 2 in Human Endothelial Cells. *Journal of Biological Chemistry* 286, 29913-21
81. Wells, V. and Mallucci, L. (1991) Identification of an autocrine negative growth factor: Mouse α -galactoside-binding protein is a cytostatic factor and cell growth regulator. *Cell* 64, 91-97
82. Adams, L., Kenneth Scott, G., and Weinberg, C.S. (1996) Biphasic modulation of cell growth by recombinant human galectin-1. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1312, 137-44
83. Yamaoka, K., Ingendoh, A., Tsubuki, S., Nagai, Y., and Sanai, Y. (1996) Structural and Functional Characterization of a Novel Tumor-Derived Rat Galectin-1 Having Transforming Growth Factor (TGF) Activity: The Relationship between Intramolecular Disulfide Bridges and TGF Activity. *Journal of Biochemistry* 119, 878-86
84. Cooper, D.N., Massa, S.M., and Barondes, S.H. (1991) Endogenous muscle lectin inhibits myoblast adhesion to laminin. *The Journal of Cell Biology* 115, 1437-48

-
85. Vandenbrule, F.A., Buicu, C., Baldet, M., Sobel, M.E., Cooper, D.N.W., Marschal, P., and Castronovo, V. **(1995)** Galectin-1 Modulates Human Melanoma Cell Adhesion to Laminin. *Biochemical and Biophysical Research Communications* 209, 760-67
86. Dagher, S.F., Wang, J.L., and Patterson, R.J. **(1995)** Identification of galectin-3 as a factor in pre-mRNA splicing. *Proceedings of the National Academy of Sciences* 92, 1213-17
87. Vyakarnam, A., Dagher, S.F., Wang, J.L., and Patterson, R.J. **(1997)** Evidence for a role for galectin-1 in pre-mRNA splicing. *Molecular and Cellular Biology* 17, 4730-37



MÉTODOS

II.1 Métodos computacionales

Desde principios del siglo XX, y luego de los descubrimientos realizados por Planck, Young y Einstein, entre otros, la mecánica cuántica se posicionó como la teoría adecuada para la descripción de partículas de dimensiones atómicas. En este contexto, debe recurrirse a metodologías basadas en ella para estudiar sistemas moleculares.

La aplicación de métodos basados en la mecánica cuántica a la química, hasta aproximadamente mediados del siglo XX, se limitaba al estudio de las propiedades de átomos aislados o moléculas pequeñas, debido a su alto costo computacional. Con el intenso desarrollo de la tecnología informática, en las últimas décadas la mecánica cuántica ha sido aplicada en la resolución de una gran diversidad de problemas químicos. Sin embargo, aún hoy resulta imposible el estudio de un sistema de miles de átomos, como por ejemplo una proteína, con un elevado nivel de teoría dentro del contexto de la mecánica cuántica. Debido a esto, en conjunto con el desarrollo de los métodos de simulación basados en la mecánica cuántica, se han desarrollado metodologías basadas en la mecánica clásica, y métodos híbridos cuántico-clásicos.

La utilización de la mecánica cuántica resulta indispensable cuando se desea abordar un sistema en el que se requiere detalle de su estructura electrónica, en particular para cualquier problema en el que se estudie la formación o ruptura de enlaces químicos. Sin embargo, existen otro tipo de situaciones de interés, en especial en macromoléculas, en los que se desean estudiar cambios conformacionales, aspectos relacionados con las fluctuaciones y el comportamiento dinámico de la molécula. En estos casos, los métodos basados en la mecánica clásica resultan de gran utilidad.

A continuación se explicarán solamente los métodos basados en la mecánica clásica, dado que los estudios realizados no requirieron del detalle cuántico de la estructura electrónica de los sistemas. Luego, se explicarán los diversos esquemas utilizados en este trabajo para explorar la superficie de energía potencial.

II.1.1 Métodos basados en la Mecánica Clásica

Muchos sistemas de interés en química y biología requieren del estudio molecular de sistemas de gran cantidad de átomos, como por ejemplo las proteínas. Para este tipo de sistemas resulta prácticamente imposible, aún con el poder computacional existente hoy en día, su tratamiento completo con mecánica cuántica. Es por esto que en gran cantidad de problemas en los que no se requiere detallar la distribución electrónica, se utilizan métodos basados en la mecánica clásica, lo que se denomina mecánica molecular. En estos métodos se ignora el movimiento de los electrones y se calcula la energía exclusivamente en función de las posiciones de los núcleos. Sin embargo, la mecánica cuántica se utiliza para el desarrollo de los distintos parámetros requeridos para el cálculo, que se explicarán en la próxima sección.

II.1.1.1 Campos de fuerza clásicos

En los métodos de mecánica molecular, la energía potencial viene dada por lo que se denomina campo de fuerza, que no es más que una expresión de la energía potencial dependiente de las coordenadas de los núcleos y una serie de parámetros [1]. La expresión del campo de fuerza AMBER [2, 3] – el cual fue utilizado para los cálculos realizados en este trabajo – contiene dos clases de contribuciones, las contribuciones de unión y las de no unión. Las de unión se computan para átomos formando un enlace o como máximo a 2 átomos de distancia, e incluyen penalidades para las desviaciones respecto de su valor de equilibrio de distancias de enlace, ángulos y ángulos diedros. Mientras que las dos primeras se representan a través de un potencial armónico centrado en la distancia de equilibrio, los ángulos diedros o torsiones están descriptos con una función periódica. Las contribuciones de no unión, se computan entre átomos que están en diferentes moléculas o a más de 3 enlaces de distancia, y se dividen en interacciones electrostáticas basadas en la ley de Coulomb, por un lado, y un potencial de Lennard-Jones que describe en forma aproximada las interacciones dispersivas y la

repulsión originada por el principio de exclusión de Pauli a distancias muy cortas. La expresión de la energía potencial resulta:

$$E(r^N) = \sum_{\text{enlaces}} \frac{k_{b,i}}{2} (r_i - r_{i0}) + \sum_{\text{ángulos}} \frac{k_{\theta,i}}{2} (\theta_i - \theta_{i0}) + \sum_{\text{torsiones}} \frac{V_n}{2} \cos(n\varpi - \gamma) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left\{ \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \quad (\text{II.1.1.1})$$

El primer término refiere a la contribución de estiramiento de los enlaces, el cual se describe a través de un potencial armónico, donde $k_{b,i}$ corresponde a la constante de fuerza asociada a la unión i -ésima y r_{i0} a la distancia de equilibrio de la misma unión. El potencial armónico resulta una buena y sencilla aproximación al potencial real en la zona cercana al mínimo. El segundo término corresponde a las contribuciones dadas por las flexiones angulares. Estas también se representan con un potencial armónico de constante $k_{\theta,i}$ y valor de equilibrio θ_i . Normalmente, la energía requerida para modificar un ángulo es menor a la requerida para variar la distancia entre dos átomos, por lo que los valores de $k_{\theta,i}$ suelen ser menores que los de las constantes $k_{b,i}$. El término asociado a las torsiones o ángulos diedros presenta tres parámetros: V_n corresponde a la barrera energética asociada a la rotación entre dos mínimos sucesivos, n representa la multiplicidad del potencial, es decir indica el número de mínimos encontrados entre 0 y 360°, y por último γ corresponde a la fase de la función sinusoidal. El último término de la ecuación II.1.1.1 contiene las contribuciones de no unión, y depende de la distancia entre pares de átomos, r_{ij} . La primera parte corresponde al potencial de Van der Waals, y contiene los parámetros ϵ_{ij} y σ_{ij} , asociados a la profundidad del pozo de energía y a la distancia entre los átomos correspondiente al mínimo de energía, respectivamente. La parte electrostática contiene como parámetros las cargas sobre los átomos. En el campo de fuerzas de AMBER, la distribución de cargas es representada a través de cargas puntuales q_i ubicadas en las coordenadas de cada átomo. Estas cargas se obtienen de manera de reproducir el potencial electrostático de la molécula, calculado a partir de cálculos cuánticos.

En las simulaciones en fase condensada, con el fin de representar un sistema infinito se utilizan las llamadas **condiciones periódicas de borde** que son, simplemente, la replicación en todas direcciones de una caja (**Figura II.1**). El objetivo de este esquema es lograr que cualquier partícula dentro de esta caja tenga el mismo entorno que el que tendría en el seno de un sistema macroscópico. El problema es ahora cómo calcular las interacciones, ya que si el sistema es infinito, las interacciones

también lo son. Para lidiar con esta suma infinita de términos se utilizan dos estrategias diferentes. Para los términos de no unión, se utiliza un **radio de corte**, tal que para los átomos que se encuentran a una distancia mayor que este radio estas interacciones no se computan. En el caso de las interacciones de Lennard-Jones, al ser de corto alcance decaen mucho más rápido ($1/r^6$) de lo que crece el número de interacciones (r^2), y esto no genera un problema siempre que el radio de corte sea lo suficientemente grande. En el caso de las interacciones Coulómbicas, el problema es más complejo ya que este tipo de interacciones decrece como la inversa de la distancia, resultando de mayor alcance.

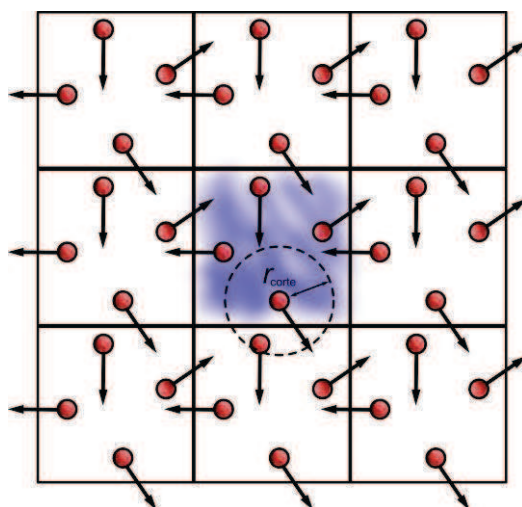


Figura II.1. Condiciones periódicas de borde y radio de corte. Representación esquemática de un sistema con condiciones periódicas de borde y radio de corte, donde la celda central se repite en todas las direcciones y una partícula dada interactúa (términos de no unión) sólo con aquellas dentro del radio de corte.

Para resolver esta situación, se utiliza una metodología denominada sumas de Ewald [4], un método originalmente desarrollado para el estudio de cristales iónicos, que en su implementación se denominaba *Particle Mesh Ewald* [5, 6]. Esta metodología se utiliza para calcular las interacciones electrostáticas totales entre los átomos presentes en una celda unidad de forma muy eficiente. La energía electrostática se calcula en dos partes: una primera parte en la cual a las interacciones electrostáticas asociadas a las cargas puntuales se les agrega una distribución de carga Gaussiana, que neutraliza cada carga. Esta parte se calcula en el espacio real y a diferencia de la suma directa de las interacciones electrostáticas, la expresión resultante converge en forma mucho más rápida. Y en una segunda etapa, las correcciones necesarias debido a la incorporación de las densidades “neutralizantes” se calculan en el espacio recíproco. Las sumas asociadas a los cálculos en el espacio recíproco también presentan una rápida convergencia. De esta manera se logra tener en cuenta las interacciones

electrostáticas de largo alcance más allá del radio de corte de las interacciones de no unión en forma altamente eficiente [7].

Finalmente, puede observarse que la mera expresión de la energía potencial no es suficiente para definir un campo de fuerza. Se requiere además una gran cantidad de parámetros asociados a cada uno de los sumandos en la energía potencial. Para el caso de simulación de proteínas, el campo de fuerza de AMBER contiene los parámetros asociados a los aminoácidos naturales (en sus distintos estados de protonación, según corresponda) presentes en la mayoría de las proteínas. En este contexto, puede hablarse de la transferibilidad del campo de fuerza, es decir, el hecho de que el mismo conjunto de parámetros puede utilizarse para moléculas asociadas. Por ejemplo, los parámetros para un residuo de arginina serán los mismos en todas las proteínas estudiadas, y no será necesaria su parametrización en cada proteína que se desee estudiar. En el campo de fuerza de AMBER los átomos dentro de cada residuo reciben un nombre de átomo y un tipo de átomo. Los tipos de átomo permiten que no sea necesario incluir parámetros para cada átomo incluido en cada uno de los aminoácidos. Distintas especies pueden poseer el mismo tipo de átomo, y los parámetros asociados a las uniones, los ángulos y los diedros se dan por tipo de átomo, en lugar de para cada átomo en particular. Si bien en el campo de fuerza se incluyen parámetros para todos los aminoácidos naturales y otras moléculas comúnmente encontradas en los sistemas proteicos, para incluir grupos prostéticos o azúcares o sus sustituyentes, por ejemplo, será necesario obtener los parámetros necesarios para realizar el cálculo.

II.1.2 Modelos para el solvente

La gran mayoría de las reacciones químicas de interés ocurren en solución acuosa; no es de extrañar entonces el esfuerzo que químicos y físicos han hecho para desarrollar modelos que permitan simular esta vital sustancia.

Dentro de los ingredientes utilizados para modelar al agua se encuentran fluidos dieléctricos, cargas puntuales fijas, cargas fluctuantes, dipolos puntuales, potenciales tipo Lennard-Jones, potenciales armónicos de enlace, entre otros. La utilización de dichos ingredientes, así como la parametrización de los mismos, dependen en buena medida de qué propiedad interesa reproducir y con qué nivel de precisión en relación al costo computacional asociado. A continuación se describirán

tres de los modelos de solventes más conocidos e implementados, ya que fueron los utilizados en esta Tesis para describir la solvatación en diferentes simulaciones.

II.1.2.1 Solvente implícito: PB y GB

La energía libre de solvatación (ΔG_{sol}) se puede descomponer en tres contribuciones:

$$\Delta G_{sol} = \Delta G_{elec} + \Delta G_{cav} + \Delta G_{vdW} \quad (II.1.2.1)$$

El primer componente es el electrostático, el segundo es el costo de cavitación (trabajo necesario para generar un cavidad para el soluto en contra de la presión del solvente y de la penalidad entrópica asociada a la reorganización de las moléculas de solvente), y el último se refiere a las interacciones de van der Waals entre solvente-soluto, el cual puede ser dividido en un termino repulsivo (ΔG_{rep}) y otro de dispersión atractivo (ΔG_{disp}). El término crítico es el primero, relacionado con la contribución electrostática a la solvatación.

Para la contribución electrostática, uno de los modelos que más se aplica en la descripción de la solvatación de biomoléculas en forma implícita son los de Poisson-Boltzmann (PB) y de Generalized Born (GB). La ecuación de PB describe en forma muy eficiente el entorno electrostático de un soluto inmerso en un solvente con iones [8]. La misma tiene la siguiente forma

$$\nabla \cdot [\varepsilon(r) \nabla \Psi(r)] = -\rho_s(r) - \sum_i c_i^\infty z_i q \lambda(r) e^{-\frac{z_i q \Psi(r)}{k_B T}} \quad (II.1.2.2)$$

donde $\varepsilon(r)$ representa el dieléctrico en función de la posición, $\Psi(r)$ representa el potencial electrostático, $\rho_s(r)$ la densidad de carga del soluto, c_i^∞ la concentración del ion i -ésimo a una distancia infinita del soluto, z_i es la valencia del ion i , q es la carga del protón, k_B la constante de Boltzmann, T la temperatura y $\lambda(r)$ es un factor que tiene en cuenta la accesibilidad de los iones en solución en función de la distancia al soluto (generalmente cercano a 1).

A pesar de tratarse de una ecuación ampliamente estudiada y reformulada para ser aplicable a diferentes problemas de la fisicoquímica de las interacciones, resulta ser computacionalmente muy costosa de implementar sin realizar aproximaciones [9-11]. Por ello, la performance de los métodos basados en la ecuación II.1.2.2 aún no han logrado la obtenida por los métodos más comúnmente empleados de GB, los cuales

son una aproximación/linealización de la ecuación de PB. El modelo de GB está basado en una representación del soluto como un grupo de esferas cuya constante dieléctrica difiere de la externa propia del solvente [12]. La fórmula que permite tener en cuenta las interacciones resultantes es

$$G_s = \frac{1}{8\pi} \left(\frac{1}{\varepsilon_0} - \frac{1}{\varepsilon} \right) \sum_{i,j}^N \frac{q_i q_j}{f_{GB}} \quad (\text{II.1.2.3})$$

donde

$$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2} e^{-D} \quad (\text{II.1.2.4})$$

y

$$D = \left(\frac{r_{ij}}{2a_{ij}} \right)^2, \quad a_{ij} = \sqrt{a_i a_j} \quad (\text{II.1.2.5})$$

siendo ε_0 la permitividad del vacío, ε la constante dieléctrica del solvente modelado, q_i la carga de la partícula i -ésima, r_{ij} la distancia entre las partículas i y j y a_n una cantidad con unidades de distancia conocida como el radio efectivo de Born [12]. El radio efectivo de Born de un átomo describe el grado de inmersión del mismo dentro del soluto; cualitativamente se podría pensar como la distancia del centro del átomo a la superficie del soluto. Por lo tanto, una descripción precisa de dichos parámetros resultan claves en el éxito de la descripción en el modelo de GB [13].

El segundo y el tercer término de la ecuación II.1.2.1, correspondientes a la pérdida de entropía para generar una cavidad en el solvente donde se aloja el soluto y la estabilización debida a interacciones de van der Waals (balance entre repulsión y dispersión), respectivamente, son los términos más sencillos de calcular. La mayoría de los modelos continuos utilizan parámetros similares a los empleados en Mecánica Clásica y hoy en día existen parámetros adecuados para la mayoría de los átomos de la tabla periódica y dependen ampliamente de su polarizabilidad y tamaño atómico [14].

II.1.2.2 Solvente explícito: TIP3P y TIP4P

Modelar de forma explícita el solvente consiste en considerar a cada una y todas las moléculas de agua que constituyen el sistema de forma explícita, considerando su estructura molecular con el grado de descripción que el modelo imponga. En las simulaciones clásicas en las que estos modelos explícitos son ampliamente empleados, la estructura molecular del agua significa dar coordenadas a los núcleos de los átomos

que la forman, distancias y ángulos de equilibrio, cargas puntuales, es decir, parámetros. Dos modelos bajo la misma filosofía de parametrización se han utilizado en la presente Tesis. Uno de ellos es el TIP3P [15]. Este modelo presenta tres sitios interactuantes, que corresponden a las posiciones de los núcleos de la molécula de agua. A cada átomo se le asigna una carga puntual ($-0,834\text{ e}$ para el oxígeno, y $+0,417\text{ e}$ para cada uno de los átomos de hidrógeno), y el átomo de oxígeno toma además los parámetros de Lennard-Jones ($\sigma = 3,15061\text{ \AA}$, $\varepsilon = 636,4\text{ J mol}^{-1}$) (**Figura II.2A**). El otro modelo empleado es el TIP4P [16]. En este modelo se representa a la molécula de agua también como un cuerpo rígido (sin grados de libertad internos) pero esta vez con cuatro centros, uno centrado en cada átomo y uno extra en el eje C_{2v} situado en la bisectriz del ángulo H-O-H (sitio M) (**Figura II.2B**). Las moléculas de agua TIP4P interactúan entre sí mediante interacciones Coulómbicas entre cargas situadas en los hidrógenos ($+0,520\text{ e}$) y el sitio M ($-1,040\text{ e}$) y mediante un potencial de tipo Lennard-Jones centrado solamente en el átomo de oxígeno sin carga ($\sigma = 3,15365\text{ \AA}$, $\varepsilon = 648,0\text{ J mol}^{-1}$). En ambos modelos las distancias interatómicas O-H se fijan en $0,9572\text{ \AA}$, al igual que el ángulo HOH, en $104,52^\circ$, correspondiéndose con las observadas experimentalmente [1, 15].

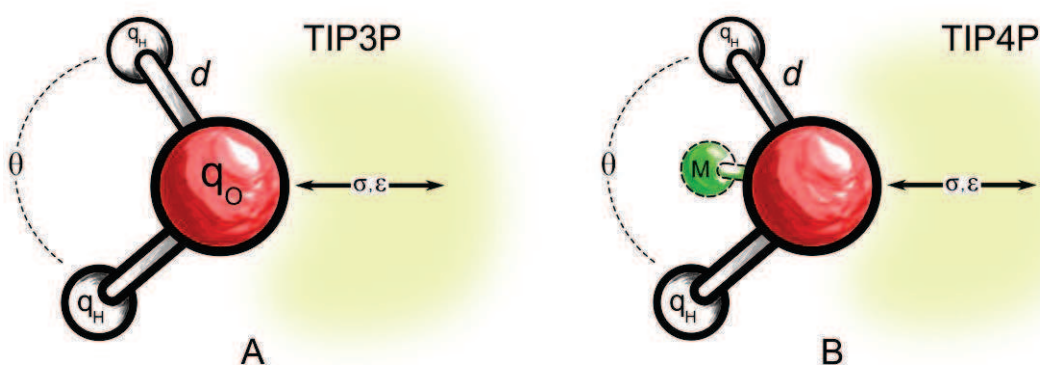


Figura II.2. Modelos explícitos de agua. Representación esquemática de distribución de sitios, cargas, ángulos y distancias para los modelos A) TIP3P y B) TIP4P.

II.1.3 Explorando la superficie de energía potencial

La superficie de energía potencial (SEP), definida en el contexto de la aproximación de Born–Oppenheimer [17, 18], es una función escalar que define el valor de la energía potencial del sistema en función de las coordenadas de los núcleos. Por ejemplo, en un sistema de N átomos depende de $3N$ variables, asociadas a las coordenadas cartesianas de cada uno de los núcleos. El conocimiento de la superficie

de energía potencial permite obtener información cinética y termodinámica de un sistema. Exceptuando sistemas extremadamente simples, la SEP resulta compleja y es imposible conocerla en su totalidad. En esta sección nos centraremos en distintos métodos desarrollados para explorar la SEP de forma eficiente, que permiten concentrar la atención en el área de la SEP que se desea explorar. Si bien existe una gran cantidad de metodologías para realizar esta tarea, se desarrollarán solamente las metodologías aplicadas en los cálculos realizados en este trabajo.

II.1.3.1 Minimización de la energía

En cualquier problema de interés químico siempre va a resultar importante encontrar los puntos mínimos de la SEP. En los mínimos de la SEP se encuentran las configuraciones o conformaciones estables, las más relevantes para el sistema de estudio. Para encontrar los mínimos de la SEP debe aplicarse un algoritmo de minimización. Debido a la complejidad con la que varía la SEP con las coordenadas cartesianas, se utilizan métodos numéricos para obtener los mínimos de la función. Dentro de los diferentes métodos, pueden distinguirse los métodos que usan las derivadas de la función a minimizar, y aquellos que no. Dentro de los primeros, existen métodos que utilizan solo la derivada primera (que es nula en la posición del mínimo), llamados de primer orden, y los que utilizan además las derivadas segundas (deben ser positivas), llamados de segundo orden. Los métodos utilizados en este trabajo corresponden a métodos de derivadas de primer orden, en particular los métodos llamados *Conjugate Gradients* y *Steepest Descents* [1]. La inclusión de la derivada primera en la determinación del mínimo de energía resulta de mucha utilidad, dado que el gradiente de la SEP indica la dirección hacia donde se encuentra el mínimo, y su valor absoluto otorga la pendiente en el punto. El gradiente de la SEP corresponde al opuesto de la fuerza, por lo que siguiendo la dirección de la fuerza sobre cada átomo puede gradualmente reducirse la energía del sistema.

El método *Steepest Descents* realiza movimientos en dirección paralela a la fuerza neta con pasos de una longitud arbitraria fijada al inicio del cálculo. En caso de que se realice un paso que produzca una reducción de la energía, la longitud del paso se incrementa por un factor. En caso de producirse un incremento de la energía se procede a reducir la longitud del paso.

El método *Conjugate Gradients*, por su parte, se diferencia del método de *Steepest Descents* en que la dirección en la que se realiza el movimiento corresponde a la

dirección de la fuerza neta más la dirección anterior multiplicada por un factor. Este factor se calcula como el cociente entre el módulo del gradiente del paso actual y el paso anterior. De esta manera, no se observa el fenómeno oscilatorio que se observa en el método de *Steepest Descents* en las zonas donde hay mínimos muy angostos.

Por último, debe tenerse en cuenta que estos métodos apuntan siempre ir a la reducción de la energía, por lo que el resultado final dependerá fuertemente de la configuración de partida. En otras palabras, simplemente a través de optimizaciones de geometría como las descritas en esta sección no se podrán atravesar barreras que conecten distintos mínimos, y el mínimo que se obtenga será el más cercano a la estructura de partida. Resulta entonces imprescindible la utilización de otras metodologías que permitan un muestreo más amplio de la SEP.

II.1.3.2 Simulaciones de Dinámica Molecular

En los casos en que se desea estudiar un sistema complejo, como en el caso de este trabajo en el que se realizan simulaciones de sistemas proteicos, la complejidad de la superficie de energía potencial involucra la existencia de gran cantidad de mínimos locales. Esto implica que no existe un solo mínimo representativo del sistema, sino que existen una serie de conformaciones, que estarán presentes en mayor o menor proporción según sus energías relativas.

Cuando se realiza una medida experimental, todas las conformaciones relevantes estarán representadas en el ensamble de moléculas presente en el sistema. Una manera de representar este ensamble es a través de la reconstrucción de la evolución temporal de la dinámica del sistema, asumiendo que el sistema ha evolucionado un tiempo suficiente para que haya podido explorar las distintas conformaciones disponibles. Justamente, la *hipótesis ergódica* [1] postula que el promedio temporal puede reemplazarse por el promedio sobre las configuraciones de un sistema dado, dando origen así a las simulaciones basadas en evolución temporal, diferenciándose de aquellas que buscan explorar el colectivo de configuraciones espaciales. En una simulación de dinámica molecular (DM), se realiza la reconstrucción de la evolución temporal de la dinámica de un sistema a través de la resolución de las ecuaciones de movimiento de Newton. En otras palabras, la trayectoria se obtiene a través de la resolución de las ecuaciones diferenciales asociadas a:

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (\text{II.1.3.1})$$

donde se describe el movimiento de una partícula de masa m_i en la coordenada x_i , con F_{x_i} la fuerza en esa dirección.

La fuerza sobre cada una de las partículas cambia cuando la partícula modifica su posición, y es función del movimiento de todas las otras partículas. Esto convierte el problema en no resoluble analíticamente por lo que debe recurrirse a métodos numéricos. El método implementado en el programa AMBER se denomina Algoritmo de Verlet [19], basado en el método de las diferencias finitas. En éste método, la integración se parte en pequeños pasos, separados por un intervalo de tiempo fijo δt . Se calcula la fuerza en cada partícula como la suma de las interacciones con el resto de las partículas, y a partir de la fuerza puede calcularse la aceleración, y con ella la nueva posición de las partículas en el siguiente paso. Durante el δt , la fuerza es considerada constante. En el algoritmo de Verlet se utilizan estratégicamente las expansiones de Taylor de las posiciones a diferentes tiempos, según

$$r(t + \delta t) = r(t) + v(t)\delta t + \frac{a(t)}{2}\delta t^2 + \dots \quad (\text{II.1.3.2})$$

$$r(t - \delta t) = r(t) - v(t)\delta t + \frac{a(t)}{2}\delta t^2 + \dots \quad (\text{II.1.3.3})$$

Sumando miembro a miembro las expresiones II.1.3.2 y II.1.3.3, y despreciando los términos de orden superior al segundo, se obtiene:

$$r(t + \delta t) \approx 2r(t) - r(t - \delta t) + a(t)\delta t^2 \quad (\text{II.1.3.4})$$

Las velocidades, necesarias para la estimación de la energía cinética, no aparecen explícitamente, pero pueden obtenerse de la resta de las mismas expresiones de antes:

$$v(t) \approx \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \quad (\text{II.1.3.5})$$

Debe prestarse especial atención a la elección del intervalo de tiempo δt , denominado *time step*. Cuanto menor sea el *time step* mayor será la cantidad de pasos que necesitaremos para realizar una simulación de una dada cantidad de tiempo, con lo que incrementaremos el costo computacional. Sin embargo, un intervalo de tiempo demasiado grande dará origen a inestabilidades en el algoritmo de integración. Por otra parte, debe elegirse un *time step* que sea menor que las escalas de tiempo asociadas a los movimientos que se desean observar, y que permita describir el movimiento de

los átomos en la molécula. En la práctica se utiliza un intervalo de tiempo que sea aproximadamente un décimo del movimiento característico más rápido del sistema. En el caso de un sistema de átomos simulado clásicamente, los movimientos más rápidos están asociados a las vibraciones de los enlaces, que se encuentran en el orden de los fs ($1 \text{ fs} = 10^{-15} \text{ s}$). Por lo tanto, un intervalo de tiempo de 0,1-0,2 fs resulta apropiado. Las vibraciones más rápidas están asociadas a los enlaces en los que uno de los átomos involucrados en la unión es un átomo de hidrógeno. Una estrategia para aumentar el *time step* y así acelerar los cálculos consiste en utilizar el algoritmo SHAKE [20], que permite fijar estas uniones en sus posiciones de equilibrio.

Normalmente, es deseable realizar las simulaciones de DM a temperatura constante. La temperatura está relacionada con la energía cinética (K), de la siguiente manera:

$$K = \sum_{i=1}^N \frac{m_i v_i^2}{2} = \frac{3k_B T}{2} N \quad (\text{II.1.3.6})$$

donde v_i y m_i corresponden a la velocidad y la masa de la partícula i y N el número de partículas del sistema. Por lo tanto, una manera simple de mantener la temperatura constante consiste en escalar las velocidades. Sin embargo, este método no refleja la naturaleza fluctuante de la temperatura alrededor de un promedio, y por ello no resulta ser el mejor de los enfoques de control térmico de las simulaciones.

Uno de los métodos más utilizados para regular la temperatura en una simulación de DM es el algoritmo de Berendsen [21]. El mismo consiste en acoplar el sistema a un baño externo cuya temperatura se fija en la temperatura deseada, que remueve o entrega calor según sea necesario. Supongamos que a un tiempo dado t el sistema posee una temperatura $T(t)$. Puede verse de la ecuación II.1.3.6 que si las velocidades se escalan por un factor λ , el cambio asociado en la temperatura resulta:

$$\Delta T = (\lambda^2 - 1)T(t) \quad (\text{II.1.3.7})$$

En este método, las velocidades se escalan en cada paso, de manera que el cambio en la temperatura es proporcional a la diferencia entre la temperatura del sistema y la temperatura del baño:

$$\frac{dT(t)}{dt} = \frac{1}{\tau} [T_{\text{baño}} - T(t)] \quad (\text{II.1.3.8})$$

donde τ es un parámetro de acoplamiento que determina cuan fuertemente el baño está acoplado al sistema. El cambio de temperatura entre dos pasos sucesivos separados por un *time step* δt , corresponde a:

$$\Delta T = \frac{\delta t}{\tau} [T_{\text{baño}} - T(t)] \quad (\text{II.1.3.9})$$

Con esta expresión, el factor de escalamiento de las velocidades λ resulta:

$$\lambda^2 = 1 + \frac{\delta t}{\tau} \left[\frac{T_{\text{baño}}}{T(t)} - 1 \right] \quad (\text{II.1.3.10})$$

Cuanto más grande sea este parámetro, más débil será el acoplamiento, por lo que la fluctuación en la temperatura será mayor. Normalmente, la utilización de un factor de acoplamiento entre 100 y 1000 veces más grande que el *time step* elegido otorga buenos resultados para los sistemas estudiados como en esta tesis.

Por otra parte, es también común realizar simulaciones de DM en condiciones de presión constante. Para mantener la presión constante, el sistema debe variar el volumen de la celda en la que se realiza la simulación. Berendsen propuso entonces un esquema similar al planteado para la regulación de la temperatura, por el cual se acopla el sistema a un baño equilibrado a una cierta presión [21]. El volumen de la celda se escala entonces en un factor γ , lo que equivale a escalar las coordenadas atómicas en $\gamma^{1/3}$. Resulta entonces:

$$\gamma = 1 + \kappa \frac{\delta t}{\tau_P} (P - P_{\text{baño}}) \quad (\text{II.1.3.11})$$

donde $\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T$ es la compresibilidad isotérmica, y τ_P corresponde al factor de acoplamiento del baño al sistema.

Luego de realizar una simulación de DM, el resultado principal que se obtiene es la trayectoria de cada uno de los átomos que conforman el sistema a lo largo del tiempo simulado, el cual típicamente no excede los 100 ns. La visualización de la misma a través de algún programa de representación gráfica aporta gran cantidad de información sobre el sistema. Con ella se pueden analizar la evolución de parámetros geométricos claves, como distancias, ángulos, ángulos diedros y otros parámetros de relevancia estructural como interacciones de puente hidrógeno e hidrofóbicas. Por otro lado, en el caso de una macromolécula, puede observarse la dinámica de la estructura terciaria, aunque los movimientos globales asociados a ella muchas veces requieren un

largo tiempo de simulación para ser observados. El programa de visualización utilizado durante este trabajo fue el programa Visual Molecular Dynamics (VMD), el cual presenta una gran cantidad de herramientas muy útiles para el análisis de una simulación de DM [22].

RMSD, RMSF y distribución radial. Un parámetro que permite estudiar la estabilidad de la estructura a lo largo de la simulación es la raíz cuadrada de la desviación cuadrática media (RMSD, del inglés *root mean squared deviation*) respecto a una estructura de referencia, la cual se calcula de la siguiente manera:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2} \quad (II.1.3.12)$$

En la expresión II.1.3.12, N corresponde al número de átomos en el sistema, y δ_i a la distancia entre las posiciones del átomo i en dos estructuras distintas, una de ellas será la estructura de referencia. Realizando este cálculo para cada una de las “fotos” de la simulación, se obtiene el RMSD en función del tiempo. La estructura de referencia puede ser la estructura cristalográfica, por ejemplo, o la estructura promedio, lo que otorga una noción de la magnitud de las fluctuaciones del sistema a lo largo de la dinámica, es decir, como una función del tiempo.

Ahora, si el promedio es sobre el tiempo, se puede tener una noción de las fluctuaciones pero por partícula:

$$RMSF = \sqrt{\frac{1}{T} \sum_{t_j=1}^T [x_i(t_j) - \tilde{x}_i]^2} \quad (II.1.3.13)$$

donde T es el tiempo total de promedio y $\tilde{x}_i$ es la posición de referencia de la partícula i . Esta es la expresión para calcular la raíz cuadrada de la fluctuación cuadrática media (RMSF), la cual resulta muy útil para comparar las magnitudes de las fluctuaciones de cada partícula del sistema en un rango de tiempo de simulación determinado, es decir, como una función de cada partícula.

Otra herramienta analítica muy útil para obtener información estructural de sistemas atómicos y moleculares (especialmente líquidos) es la llamada función de distribución radial, $g(r)$. Esta función no es otra cosa que la probabilidad de encontrar una partícula a una distancia r de otra, y se define como sigue:

$$g(r) = \frac{P(r, r + dr)}{V(r, r + dr)} \quad (II.1.3.14)$$

donde $P(r, r+dr)$ es la probabilidad de encontrar una partícula encerrada en el volumen $V(r, r+dr)$ (Figura II.3A).

Esta función contiene valiosa información acerca del sistema. En un gas ideal esta función es constante: la probabilidad de encontrar una partícula en un lugar del espacio es constante y a través de ella se pueden calcular muchas de las propiedades termodinámicas de los gases [23]. En un sólido consta de una serie infinita de crestas y valles debido a la periodicidad del mismo y puede ser experimentalmente determinada en experimentos de difracción de neutrones o rayos X [24, 25]. En un líquido esta función resulta intermedia, por lo general tiene una serie de máximos y mínimos cada vez menos pronunciados hasta que se vuelve constante (Figura II.3B).

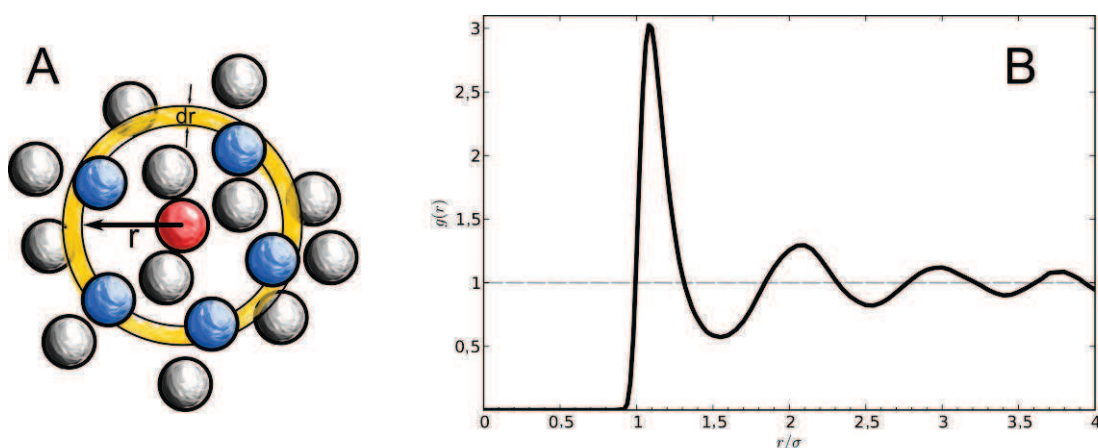


Figura II.3. Función de distribución radial. A) Representación esquemática del cálculo de la función de distribución radial; B) Forma gráfica de una función de $g(r)$ para un líquido ideal de esferas rígidas.

Dinámica esencial. En sistemas complejos como los sistemas proteicos, resulta comúnmente difícil por simple observación de la trayectoria lograr detectar movimientos globales de la estructura, como ser la variación en la distancia entre dos elementos de estructura secundaria o el movimiento de un *loop*. La determinación de la dinámica esencial (*Essential Dynamics*) [26] permite determinar los movimientos principales que ocurren en el sistema, extrayendo principalmente los movimientos de baja frecuencia, que ocurren globalmente a lo largo de la simulación e involucran normalmente gran cantidad de átomos.

Técnicamente, el cálculo implica la diagonalización de la matriz de covarianza de las posiciones, lo que corresponde a una matriz de $3N \times 3N$, siendo N la cantidad de átomos del sistema. Si definimos X como un vector de dimensión $3N$ que contiene las

coordenadas x, y, z , de los N átomos presentes en el sistema, la matriz de covarianza de las posiciones en el lugar ij (C_{ij}) está definida por la siguiente expresión:

$$C_{ij} = \langle (X_i - \langle X_i \rangle) (X_j - \langle X_j \rangle) \rangle \quad (\text{II.1.3.15})$$

donde $\langle \rangle$ denota el promedio temporal. Los autovectores de esta matriz representan los movimientos esenciales, mientras que los autovalores indican el peso de cada uno de ellos en la varianza estructural global. En la mayoría de los casos, se observa que con no más de diez modos esenciales puede obtenerse una representación de más del 60% de la varianza estructural observada en la simulación [26]. La animación de la trayectoria en la dirección dada por los autovectores de mayor peso permite observar los movimientos globales asociados a los modos obtenidos.

Los autovectores de una trayectoria pueden ser comparados con los de otra trayectoria, lo que permite obtener una medida cuantitativa de la similitud entre dos trayectorias independientes, de la siguiente manera:

$$\gamma_{AB} = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^n (v_i^A \cdot v_i^B)^2 \quad (\text{II.1.3.16})$$

En la expresión II.1.3.16, γ_{AB} representa el coeficiente de similitud entre los modos esenciales de dos trayectorias distintas A y B , mientras que v_i^A y v_i^B corresponden a los autovectores i y j de las trayectorias A y B , respectivamente. n es el número de vectores a utilizar en la comparación, elegido de forma tal de representar un porcentaje (elegido por el usuario) de la variación estructural global.

Normalmente, resulta muy útil para el análisis de los modos esenciales evaluar la proyección de cada uno de ellos sobre la trayectoria global. Si se grafica la proyección de cada uno de los modos en función del tiempo de la simulación, se puede tener una idea aproximada de la cantidad de veces que fue explorado el movimiento asociado al modo. Se observará un comportamiento aproximadamente ondulatorio, donde cada ciclo representa un movimiento completo. En especial, para los modos de muy baja frecuencia puede observarse si el movimiento fue realizado en forma completa al menos una vez, indicando que el mismo fue muestreado correctamente.

Por último, al utilizar la técnica de dinámica esencial, no debe olvidarse que la misma es muy sensible a la longitud de la trayectoria obtenida. Por otra parte, los movimientos esenciales que se observarán serán exclusivamente los que ocurrieron durante la simulación, por lo que movimientos globales que involucren largas escalas de tiempo (que excedan el tiempo de simulación) no serán observados con esta técnica.

Entropía conformacional. La variación de entropía conformacional al cambiar entre dos estados de un sistema resulta de gran interés. Por ejemplo, en sistemas proteicos puede evaluarse el cambio de entropía conformacional asociado al cambio entre dos estados conformacionales de la proteína o a la unión de un determinado ligando, lo que puede modificar las fluctuaciones del sistema significativamente. Schlitter en 1993 propuso un método [27], luego mejorado por Andricioaei y Karplus en el 2001 [28], para estimar la entropía conformacional a partir de la diagonalización de la matriz de covarianza de las posiciones. Este cálculo se basa en la utilización de la derivación de la entropía para un oscilador armónico, por lo que es denominada aproximación cuasiarmónica.

El desarrollo de Schlitter se basa en la utilización de la entropía calculada para un oscilador armónico cuántico, es decir:

$$S_{OA} = \frac{k\alpha}{e^\alpha - 1} - \ln(1 - e^{-\alpha}) \quad (\text{II.1.3.17})$$

donde $\alpha = \hbar\omega/k_B T$, y ω la frecuencia del oscilador armónico. La frecuencia ω es un parámetro conectado con la varianza de las posiciones, o sea con $\langle \Delta x^2 \rangle$. En una DM, se obtiene la varianza clásica, que llamaremos $\langle \Delta x^2 \rangle_{cl}$. En el límite clásico:

$$\langle \Delta x^2 \rangle_{cl} = \frac{k_B T}{m\omega^2} \quad (\text{II.1.3.18})$$

con lo que resulta:

$$\frac{1}{\alpha^2} = \frac{mk_B T}{\hbar^2} \langle \Delta x^2 \rangle_{cl} \quad (\text{II.1.3.19})$$

En esta última expresión queda entonces relacionado el valor de α con la varianza de las posiciones, y por lo tanto la entropía también lo está (ecuación II.1.3.17).

Inicialmente, Schlitter introduce una aproximación para simplificar el cálculo y obtiene la siguiente expresión:

$$S_{OA} < S = \frac{k}{2} \ln \left(1 + \frac{e^2}{\alpha^2} \right) \quad (\text{II.1.3.20})$$

Generalizando este problema para N partículas, debe recurrirse a la matriz de covarianza de las posiciones, ya definida en II.1.3.15. En este caso, utilizaremos la matriz de covarianza de las posiciones pesada por las masas, es decir $\mathbf{C}' = \mathbf{M}^{1/2} \mathbf{C} \mathbf{M}^{1/2}$ donde $\mathbf{M}$ es una matriz diagonal que contiene las masas de las

partículas. De esta forma, se obtiene la expresión para la entropía conformacional de un sistema de N átomos, según el desarrollo de Schlitter:

$$S_{sch} = \frac{k}{2} \ln \det \left(\mathbf{I} + \frac{k_B T e^2}{\hbar^2} \mathbf{C}' \right) \quad (\text{II.1.3.21})$$

donde $\mathbf{I}$ corresponde a la matriz unitaria.

Luego, Andronaeci y col. demostraron que no es necesario realizar la aproximación que deriva en la ecuación II.1.3.18, y obtuvieron la siguiente expresión para la entropía:

$$S_{and} = k \sum_i^{3N-6} \left[\frac{\hbar \omega_i / k_B T}{e^{\hbar \omega_i / k_B T} - 1} - \ln(1 - e^{-\hbar \omega_i / k_B T}) \right] \quad (\text{II.1.3.22})$$

En la expresión anterior, los valores de ω_i se obtienen de los autovalores (λ_i) de la matriz $\mathbf{C}'$, dado que $\omega_i = \sqrt{k_B T / \lambda_i}$.

Para obtener la entropía conformacional a partir de una simulación de DM, debe entonces en primer lugar calcularse y diagonalizarse la matriz de covarianza de las posiciones, pesada por las masas $\mathbf{C}'$. Una vez obtenidos los autovalores de la matriz $\mathbf{C}'$, se calcula la entropía utilizando la ecuación II.1.3.21 (o la II.1.3.22).

Finalmente, debemos tener en cuenta que el valor de la entropía conformacional se incrementa a medida que aumenta el tiempo de simulación, dado que aumenta la exploración del espacio de las fases del sistema. Es por ello que sería deseable obtener una estimación de la entropía a tiempo infinito (S_∞). Para lograrlo, se calcula la entropía para intervalos crecientes de tiempo, y luego se ajusta alguna función $S(t)$ que contenga una asíntota horizontal y otorgue el valor de S_∞ extrapolando para $t \rightarrow \infty$. La función elegida para el ajuste fue [29]:

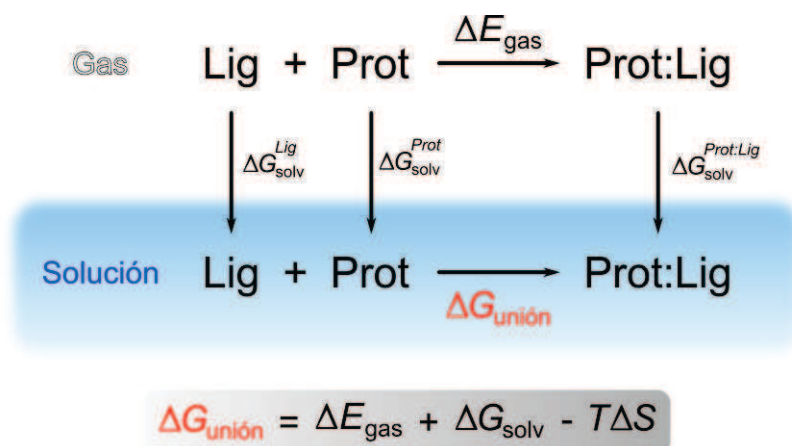
$$S(t) = S_\infty - \frac{a}{t^{2/3}} \quad (\text{II.1.3.23})$$

Para realizar entonces el cálculo de la entropía conformacional es conveniente dividir la trayectoria en intervalos de tiempo de tamaño creciente, y calcular la entropía para cada uno de los intervalos. Luego se grafica la entropía calculada para cada intervalo en función de la longitud de tiempo del intervalo, y se ajusta la función dada por la ecuación II.1.3.23. Una alternativa al último ajuste, y aprovechando la presencia de un valor asintótico de la entropía cuando la trayectoria tiende a infinito, es una linealización por inversa de la abscisa, donde $S(t)$ en función de $1/t$ genera una curva a la que se le puede ajustar una recta y rescatar de la ordenada al origen la S_∞ .

Energía libre de interacción. Los cálculos de energía libre y su variación a lo largo de un proceso otorgan información esencial de un sistema químico. La energía libre de Gibbs es la energía libre calculada en condiciones de número de partículas, temperatura y presión constantes (*NTP*), las cuales son normalmente utilizadas en las simulaciones de DM, además de ser las condiciones experimentales más frecuentes.

En muchos casos resulta necesario obtener una estimación del cambio en energía libre asociado a la unión de un ligando en una proteína, por ejemplo, cuando se desea comparar la afinidad por distintos sustratos o inhibidores de una enzima. Para ello, existe un modulo implementado en el paquete de programas de AMBER denominado MM-PBSA (*Molecular Mechanics Poisson Boltzmann Solvent Accesible Surface Area*) o MM-GBSA (*Molecular Mechanics generalizad Born Solvent Accesible Surface Area*). Estos módulos se utilizan luego de una simulación de DM de la cual se colectan un número (alto) de estructuras o *snapshots* del sistema. Este método combina el cálculo de energía potencial dado por el campo de fuerzas de AMBER con un método de solvente continuo, que puede tratarse del método de Poisson-Boltzmann (PB) o el método de Born generalizado (GB) (ver sección II.1.2). Por otro lado, la contribución a la solvatación no electrostática (o no polar) se incorpora con el cálculo de la SASA (*Solvent Accesible Surface Area*). Originalmente descrito por Lee y Richards [30], los valores de SASA se expresan comúnmente en Å². La SASA se calcula generalmente mediante el algoritmo de Shrake y Rupley [31] en donde una esfera (de solvente) con un radio de prueba determinado (1,4 Å es el valor usado para probar agua como solvente) se hace “rodar” a lo largo de toda la superficie de la molécula. Los valores de SASA fueron calculados mediante el plugin vmdICE [32], implementado en el programa de visualización VMD [33]. Así, con este valor calculado sobre nuestro sistema es posible estimar la energía libre de solvatación no electrostática que se la define linealmente dependiente de la superficie accesible al solvente. Los resultados de estos cálculos si bien en forma aislada cuantitativamente pueden resultar en valores no consistentes con los valores experimentales, permiten obtener tendencias comparativas muy útiles en el estudio de las interacciones proteína-ligando [34-36].

Para realizar el cálculo de $\Delta G_{\text{unión}}$, explícitamente se requiere calcular todos los términos del ciclo termodinámico mostrado en el Esquema II.1.



Esquema II.1. Cálculo de $\Delta G_{\text{unión}}$. Representación esquemática del ciclo termodinámico que permite el cálculo del cambio de energía libre de Gibbs de unión ($\Delta G_{\text{unión}}$) de un ligando (Lig) a una proteína (Prot). ΔE_{gas} representa el cambio de energía potencial de unión en vacío y ΔG_{solv} , el cambio de energía libre para el proceso de solvatación de cada molécula involucrada en la unión. Finalmente, $-T\Delta S$ tiene en cuenta los cambios de entropía

En primer lugar, se realiza una DM larga para el complejo proteína-ligando a estudiar (en general alrededor de los 20-30 ns, dependiendo del sistema). Luego, debe generarse un conjunto de entre 100 y 1000 fotos del sistema seleccionadas de manera aleatoria, o bien fotos consecutivas pero separadas como mínimo cada 4 ps [37]. A continuación, los módulos de AMBER se encargan de remover automáticamente el solvente explícito y finalmente a partir de la DM del complejo se construyen tres trayectorias nuevas (la del complejo Lig:Prot, la del receptor Prot y la de ligando Lig). Una vez que estos nuevos conjuntos de fotos son generados, se calcula el valor de ΔE_{gas} como el promedio de la diferencia en la energía potencial derivada de la evaluación de cada estructura del sistema, utilizando el correspondiente campo de fuerzas (sección II.1.1.1), como se muestra en las ecuaciones II.1.3.24 y II.1.3.25.

$$\Delta E_{\text{gas}} = E_{\text{Prot:Lig}} - E_{\text{Prot}} - E_{\text{Lig}} \quad (\text{II.1.3.24})$$

$$E = E^{\text{unión}} + E^{\text{no-unión}} = E^{\text{unión}} + (E_{\text{Elec}} + E_{\text{vdW}}) \quad (\text{II.1.3.25})$$

El uso de la aproximación de “trayectoria simple” [38, 39] (es decir, sólo se calcula la trayectoria del complejo y en base a ésta se construyen las trayectorias para el receptor y el ligando), en el cálculo de ΔE_{gas} , todos los términos de unión se cancelan entre sí y el único aporte a ΔE_{gas} proviene de los términos de no-unión E_{Elec} y E_{vdW} . Las ventajas de la aproximación de trayectoria simple para los cálculos de energía libre de unión radican en un menor costo de computo (al calcular menos trayectorias de DM), la convergencia de los resultados se alcanza más rápido y se disminuye la dispersión de los mismos ya que las diferencias conformacionales que se generan en regiones lejanas

al sitio de unión para las biomoléculas (en particular para el receptor y la porción proteica del complejo) no se cancelan cuando el cálculo se hace a partir de tres DM generadas de manera independiente, una para el complejo y otras para el receptor y ligando, respectivamente.

El próximo paso consiste en evaluar la solvatación de las estructuras provenientes de las fotos derivadas de la aproximación de una trayectoria y calcular así la contribución del ΔG_{solv} a la energía de unión global, que puede ser separada en dos términos según

$$\Delta G_{solv} = \Delta G_{solv}^{polar} + \Delta G_{solv}^{no\ polar} = \Delta G_{solv}^{PB(GB)} + \Delta G_{solv}^{SASA} \quad (II.1.3.26)$$

donde el primero corresponde a la componente polar y el segundo a la no polar de ΔG_{solv} . La contribución polar al término de solvatación se calcula utilizando el modelo de PB [40] o GB [41] (sección II.1.2), y el término no polar incluye un término de cavitación que tiene en cuenta el efecto hidrofóbico de la desolvatación cuando se forma el complejo a partir de sus componentes libres en solución. En particular, este efecto no polar es debido mayoritariamente al cambio en el área expuesta a solvente cuando se forma el complejo y se suele estimar considerándolo directamente proporcional al cambio de área expuesta a solvente luego de la formación del complejo, utilizando parámetros optimizados para cada tipo de solvatación [42] (ecuación II.1.3.27).

$$\begin{aligned} \Delta G_{solv}^{SASA} &= \gamma \cdot \Delta SASA + \beta = \\ &= \gamma \cdot (SASA_{Prot:Lig} - SASA_{Prot} - SASA_{Lig}) + \beta \end{aligned} \quad (II.1.3.27)$$

Un trabajo reciente de Hou y col. [34] muestra una comparación sistemática entre MM-PBSA y MM-GBSA para el cálculo de la energía libre de unión de un conjunto de complejos proteína-ligando, mostrando que si bien MM-PBSA funciona mejor para predecir el valor absoluto de la energía libre, MM-GBSA es capaz de predecir mejor la afinidad relativa entre complejos. Considerando que MM-GBSA es computacionalmente menos costoso, este método surge como una herramienta promisorio para ser utilizada en el diseño asistido de drogas y explicar selectividades diferenciales según valores de afinidad, características determinantes a la hora de seleccionar el mejor ligando entre muchos.

Por último, la entropía S se descompone en contribuciones traslacionales, rotacionales y vibracionales [43] (ecuación II.1.3.28).

$$S = S^{traslacion\ al} + S^{rotacional} + S^{vibracional} \quad (II.1.3.28)$$

$$-T\Delta S = -T(S_{Prot:Lig} - S_{Prot} - S_{Lig}) \quad (II.1.3.29)$$

Los dos primeros términos de la ecuación II.1.3.28 se computan por expresiones estándar de la termodinámica estadística, y el último se calcula típicamente utilizando algún análisis de modos normales (armónicos o cuasiarmónicos) [44, 45]. En la práctica, la mayoría de las implementaciones de software actuales determinan las tres contribuciones a S como parte de un análisis armónico de modos normales de vibración. El término correspondiente al cambio de la entropía del proceso (ecuación II.1.3.29) es necesario si se está en búsqueda del orden de magnitud correcto del valor absoluto del cambio de energía libre. Sin embargo, en algunos casos, no es necesario para poder estimar cambios de energía libre relativos (cuando por ejemplo, los complejos estudiados son muy similares entre ellos) [41, 46].

II.1.4 Modelado por homología

Dentro de las herramientas más importantes de la bioinformática, el modelado por homología es quizá la más importante de las más comúnmente utilizadas en el rubro de la biología estructural. Modelar por homología consiste en predecir la estructura de una proteína cuya estructura experimental no se conoce, utilizando como molde la estructura experimental de otra proteína. Este método se basa en que ambas secuencias son homólogas, es decir, ambas proteínas comparten un antecesor común en la evolución de las mismas. La aproximación está basada principalmente en dos observaciones:

- La estructura de una proteína es determinada únicamente por su secuencia de Aminoácidos [47].
- Durante la evolución, los cambios en la estructura proteica tienen una frecuencia menor que los cambios en la secuencia. Ergo, a pesar de que dos proteínas relacionadas/homólogas posean secuencias distintas, pueden seguir manteniendo estructuras similares [48].

El modelado por homología es la única manera de obtener información estructural cuando las técnicas experimentales fallan. Esto se debe a que muchas proteínas son demasiado grandes para un análisis de RMN y/o resultan muy difíciles de cristalizar para su análisis de rayos X, como por ejemplo, las proteínas de transmembrana.

El modelado por homología involucra tres etapas:

1. **Alineamiento de secuencia.** La secuencia de aminoácidos de la proteína cuya estructura se quiere modelar (objetivo o *target*) se alinea con otra cuya estructura experimental se conoce (templado o *template*). El alineamiento se lleva a cabo utilizando una matriz de sustitución que da un peso probabilístico al alinear dos pares de aminoácidos de cada secuencia. Por ejemplo, alinear dos aminoácidos similares como una serina con una treonina, tendrá una probabilidad mayor que si se alinea esa serina con una fenilalanina. Para este propósito las matrices de sustitución más conocidas son PAM y BLOSUM. El alineamiento secuencial es realizado por programas como Clustal-W [49], BLAST [50], PSI-BLAST [51], los cuales también son capaces de buscar en base de datos secuencias homólogas a la de interés y establecer un orden de similitud en forma más automatizada.
2. **Construcción del modelo.** El alineamiento entre el objetivo y el templado sirve de entrada para el programa que realiza la construcción del modelo. En el caso de esta Tesis, el programa utilizado fue Modeller [52, 53], el cual aplica restricciones espaciales obtenidas de dos maneras: 1) a partir del alineamiento y 2) restricciones estereoquímicas obtenidas del campo de fuerza CHARMM-22 y de datos estadísticos obtenidos de un conjunto representativo de estructuras proteicas conocidas. Luego, empleando gradiente conjugado y DM, se optimiza el modelo aplicando restricciones espaciales, que pueden ser de origen experimental [54].
3. **Evaluación del modelo.** Luego de la construcción del modelo, éste debe ser evaluado a fin de asegurar que las características estructurales del modelo sean consistentes con el sentido fisicoquímico. Una alternativa, puede ser la evaluación y corrección de anormalidades en los ángulos ϕ - ψ (gráfico de Ramachandran), en la longitud de los enlaces, etc. Otra opción es comparar ciertos parámetros estadísticos obtenidos a partir del análisis de estructuras determinadas experimentalmente con aquellos existentes en el modelo. De esta forma, es posible reconocer qué regiones del modelo están plegadas correctamente y cuales no, y determinar si es necesario algún tipo de refinamiento estructural. Entre las herramientas virtuales más conocidas se encuentran: Procheck [55], ANOLEA [56], WHATIF [57] y Verify3D [58].

II.2 Métodos experimentales

Simular es experimentar con un modelo. Pero cuando el modelo es la realidad, el experimento cobra otro sentido. Por ende, para acercarse a la simulación, el modelo y la realidad a un único punto de encuentro, es necesario también realizar experimentos de mesa que sean útiles para poner a prueba las hipótesis de trabajo en un entorno lo más cercano posible al escenario real. Para ello, siempre es valiosa la colaboración entre grupos con experiencia en diferentes disciplinas. En ese sentido, luego de aprendidas las principales técnicas computacionales se ha tratado de emplearlas también en la asistencia a la resolución de dos problemas concretos de la biología de las galectinas pero abordándolos asimismo con técnicas experimentales bioquímicas-biofísicas en una manera interdisciplinaria. Para el autor de esta Tesis, el desafío y la motivación son dobles dado que se trata de traer grupos con diferentes enfoques de estudio a sentarse a discutir sobre un mismo problema, coordinando experimentos *in silico* e *in vitro* de manera estratégica y aportando a la visión de la relación estructura-función en la familia de las galectinas humanas. A continuación, entonces, se introducen y resumen las técnicas, ensayos y reactivos empleados en los ensayos experimentales de esta Tesis.

II.2.1 Reactivos y anticuerpos

Todos los experimentos *in vitro* se realizaron a 25°C en buffer fosfato salino 100 mM (PBS) conteniendo 0,1 mM de ácido dietilentriaminopentaacético (DPTA) como quelante de trazas de metales y a pH 7,4, salvo que se indique lo contrario.

Se utilizaron diferentes reactivos químicos cuyo empleo ha de describirse en cada capítulo en particular. A continuación se citan para tener un registro de ellos: β -lactosa (Sigma), catalasa y seroalbúmina bovina (BSA) liofilizadas (Sigma), polimerasa KOD Hot Start (Novagen), ligasa T4 DNA (Promega), H_2O_2 (Mallinckrodt Chemicals), ácido 5,5'-ditiobis-(2-nitrobenzónico) (DTNB) (Sigma), dodecilsulfato de sodio (SDS), 2-mercaptoetanol (ME) (Gibco), ditioneitol (DTT), paraformaldehído (PFA) y glutaraldehído de calidad microscopía electrónica (Electron Microscopy Science), antibióticos (penicilina y estreptomycin) (Gibco), tripsina-EDTA (Gibco), Triton X-100 (Pierce), polivinilpirrolidona y sacarosa (Sigma), estreptavidina-PE (Sigma), lectinas biotiniladas (Sigma): fitohemaglutinina-L (L-PHA), lectina de *Lycopersicon esculentum*

(tomate) (LEL), lectina de *Sambucus nigra* (saúco) (SNA) y lectina II de *Maackia amurensis* (MAL-II).

Para los ensayos de apoptosis se utilizaron buffer de tinción (100 mM HEPES, 1,4 M NaCl, 25 mM CaCl₂), anexina V conjugada a FITC y yoduro de propidio.

Para los ensayos de microscopía, los anticuerpos anti-VEGFR2 monoclonal (55B11 C-terminal) y anti-VEGFR2 policlonal (AF357 N-terminal) fueron obtenidos de Cell Signaling y R&D Systems, respectivamente. Como anticuerpos secundarios se utilizaron el anti-IgG_{conejo}-Alexa488® (Invitrogen) y anti-IgG_{cabra} (Jackson ImmunoResearch) según el caso. Para la tinción nuclear se utilizó 4',6-diamino-2-fenilindol (DAPI). Las muestras se montaron con el reactivo ProLong® Gold Antifade (Life Technologies). Para la inmunocitoquímica, se empleó proteína-A conjugada con nanopartículas de oro (5 y 10 nm de diámetro) (Utrecht University, Netherlands). Se empleó acetato de uranilo para el contraste de membranas y metilcelulosa y sacarosa en todas las soluciones como criopreservantes.

II.2.2 Producción de galectina-1 recombinante

Una parte esencial del trabajo experimental de esta Tesis fue la optimización del protocolo de producción de galectina-1 (Gal-1) humana recombinante [59] y su ejecución en forma repetida y sostenida a lo largo del desarrollo de los diferentes experimentos. Para tener una noción de tiempos, cada tanda de producción llevó alrededor de siete días de trabajo con el actual equipamiento del laboratorio. En cada producción se obtuvieron, en promedio, unos 20-30 mg de Gal-1 pura. Si bien para los ensayos con células se requieren cantidades del orden de los nanogramos, ese no es el caso de los ensayos *in vitro*, en donde se requieren usualmente mil veces más de proteína recombinante para lograr señales óptimas en los ensayos espectroscópicos aquí presentados. Eso involucró muchos meses de trabajo solamente dedicados a la optimización de la producción y llevarla a cabo sucesivas veces para contar con el reactivo en tiempo y forma para los posteriores ensayos. Es por ello que es pertinente resaltar este apartado casi al nivel de una etapa más de logros y resultados de esta Tesis. A continuación se citarán los procedimientos implementados y optimizados en el laboratorio en más de quince producciones espaciadas en el tiempo de trabajo.

Generación bacterias competentes, método químico. Cultivar durante la noche (DN, 16 h aprox.) a 37°C y con agitación orbital, 10 mL de bacterias *Escherichia coli* BL21

(DE3) en medio líquido Luria Broth (LB) (10 g triptona de soja, 5 g de extracto de levadura, 10 g NaCl y agua destilada c.s.p 1L; esterilizar en autoclave y llevar a pH 7,4 bajo esterilidad). Realizar una dilución 1:100 del inóculo obtenido (volumen final 100 mL) y se espera en agitación y 37 °C hasta que el medio adquiriera una densidad óptica a 600 nm de 0,6 (2-3 horas). Transferir a un tubo de centrifuga previamente enfriado y mantener en hielo 15 min. Centrifugar 10 min a 3000 rpm a 4°C. Resuspender en 35 mL de CaCl₂ 0,1 M y mantener en hielo durante 1 h, con agitación ocasional. Centrifugar nuevamente 10 min a 3000 rpm en centrífuga refrigerada. Resuspender el pellet en 1 mL de CaCl₂ 0,1M y 15% glicerol, todo estéril. Alicuotar la suspensión resultante en tubos eppendorf de 0,5 mL (20 µL/tubo), usar para transformar o congelar a -70°C.

Transformación de las bacterias. En el tubo eppendorf conteniendo las bacterias competentes (en baño de hielo), agregar 2 µg de vector de expresión (plásmido pET22b (Novagen) conteniendo el gen codificante de Gal-1 humana gentileza de la Dr. Linda Baum, UCLA, USA). Incubar en forma secuencial: a) 2 min en hielo; b) 40 s a 42°C y c) 2 min en hielo. Agregar 80 µL de medio LB rico en glucosa (LB con glucosa 1M). Incubar durante 1 h a 37°C en baño o en incubador con agitación. Sembrar 30 µL del cultivo líquido con rastrillo de vidrio en placas con LB agar con ampicilina (Amp, 100 µg/mL). Incubar DN a 37°C en estufa para bacterias. Realizar los controles adecuados para probar la incorporación del plásmido, la viabilidad de las bacterias y el funcionamiento del antibiótico selector (placas con bacterias sin transformar en LB agar con y sin ampicilina). Inocular 100 mL de medio de cultivo líquido LB con Amp con una UFC transformada proveniente de las placas correspondientes. Incubar DN a 37°C en agitador.

Preparación de la columna de afinidad. Para la separación de la Gal-1 recombinante del resto del lisado celular proveniente del cultivo de las bacterias productoras de proteína recombinante, se utilizó una estrategia de recuperación por afinidad. El relleno de una columna de afinidad se preparó desde cero siguiendo el protocolo original [59]. En forma resumida, en una columna de vidrio de 30 x 2,5 cm (GE Healthcare) rellenar con 150 mL de Sepharose 6B (Sigma), lavar 3 veces con agua destilada, luego 3 veces con Na₂CO₃ 0,5 M pH 11,0 y descartar el líquido. Agregar 60 mL de una solución al 40% v/v de divinilsulfona en Na₂CO₃ 0,5 M pH 11,0. La sefarosa se oscurece y toda la suspensión adquiere un color marrón claro. Agitar a temperatura

ambiente durante 1,5 h. Luego de que la reacción de activación se completa, lavar el relleno con 5 volúmenes de agua destilada y luego con 2 volúmenes de una solución de Na_2CO_3 0,5 M pH 10,0. Agregar 60 mL de β -lactosa (Sigma) al 40 % m/v en Na_2CO_3 0,5 M pH 10,0. Dejar en constante agitación durante 3 días a 4 °C. Remover el sobrenadante y agregar 60 mL de Na_2CO_3 0,5 M pH 8,8, 1 mL de ME y permitir agitación durante 3 h a temperatura ambiente. Lavar el relleno cuatro veces con agua destilada y reservar a 4 °C hasta su uso, equilibrando el relleno con PBS pH 7,5 y 0,02% NaN_3 (**Figura II.4**). Cada vez que se utiliza esta columna hay que seguir los siguientes pasos para prepararla antes de cargarla con el lisado celular: 1) Lavar la columna con 500 mL de buffer de lavado; 2) Bloquear la unión inespecífica haciendo circular 300 mL de buffer de bloqueo y dejando reposar la columna con este buffer DN a 4°C en cámara fría; 3) Lavar la columna con suficiente buffer de lavado (4 volúmenes aproximadamente) hasta que el eluido presente $\text{DO}_{280 \text{ nm}}$ igual a cero.



Figura II.4. Columna de afinidad, rellena con lactosil-sefarosa. Al costado de la columna, se muestra la bomba peristáltica utilizada para hacer circular los diferentes buffers a través de la misma.

Inducción de la expresión de Gal-1. A partir del inóculo, transferir 10 mL por Erlenmeyer de 2 L con 750 mL de medio de cultivo LB Amp en su interior. Utilizar por producción tres de estos Erlenmeyers para alcanzar las cantidades y rendimientos manifestados en este protocolo. Incubar a 37 °C y con agitación hasta una $\text{DO}_{600 \text{ nm}}$ de 0,4 (inicio de la fase exponencial de crecimiento del cultivo). El vector de expresión utilizado es un plásmido que posee el gen *LGALS1* inserto bajo el operón *lac* el cual se activa en la presencia de lactosa. La estrategia más común es utilizar un análogo no-hidrolizable en lugar de lactosa: isopropil- β -D-1-tiogalactopiranosido (IPTG). Por lo

tanto, al alcanzar la DO indicada, agregar IPTG (concentración final 100 μ M) a cada frasco Erlenmeyer, incubar 4 h a 37 °C con agitación y centrifugar (5000 rpm, 10 min a 4°C).

Preparación del lisado celular. Para recuperar la proteína recombinante del interior de las bacterias, es necesario lisar las células y liberar todo el contenido de las mismas a la suspensión. En el laboratorio utilizamos dos estrategias diferentes, según la disponibilidad de reactivos: a) empleo del buffer de lisis B-PER (Pierce) según indicaciones del fabricante, suplementado con ME 4mM; b) sonicado del pellet en buffer PBS suplementado con ME y un inhibidor de proteasas (1 mM de fluoruro de metilfenilsulfonio (PMSF) recién preparado), en frío y empleando 4 pulsos de 30 s cada uno (80% potencia), esperando 30 s entre cada pulso para permitir la disipación del calor. En la experiencia propia de la puesta a punto de este protocolo particular, se recomienda el lisado químico con el buffer de lisis comercial por sobre el lisado ultrasónico, ya que arroja mejores rendimientos y resulta mucho más fácil y seguro su uso e implementación, siendo el costo del mismo su única desventaja. Una vez lisado el pellet de bacterias, centrifugar a 10.000 rpm por 15 min a 4 °C. Transferir el sobrenadante a un tubo limpio y proseguir inmediatamente a la carga de la columna de afinidad.

Elución de la columna de afinidad. Dejar fluir el buffer de lavado de la columna hasta llegar casi al límite del relleno. Cargar el sobrenadante del paso anterior y permitir que ingrese en su totalidad por gravedad al interior de la columna. Agregar suficiente buffer de lavado para evitar que el relleno se seque. Cerrar la columna para evitar movimiento del líquido y dejar en reposo DN a 4 °C. Lavar la columna con buffer de lavado hasta $DO_{280\text{ nm}}$ igual a cero. Cambiar al buffer de elución y comenzar a eluir la Gal-1 de la columna. Guardar los primeros 30 mL para chequear ausencia de Gal-1 en esta fracción, e ir chequeando las fracciones (3 mL por fracción) ya sea midiendo $Abs_{280\text{ nm}}$ o mediante la reacción de Bradford (**Figura II.5**). Una vez que haya eluido toda la Gal-1, lavar la columna haciendo circular 500 mL de buffer de stripping y luego 500 mL de buffer de lavado. Cerrar la columna y reservar a 4 °C hasta próximo uso.

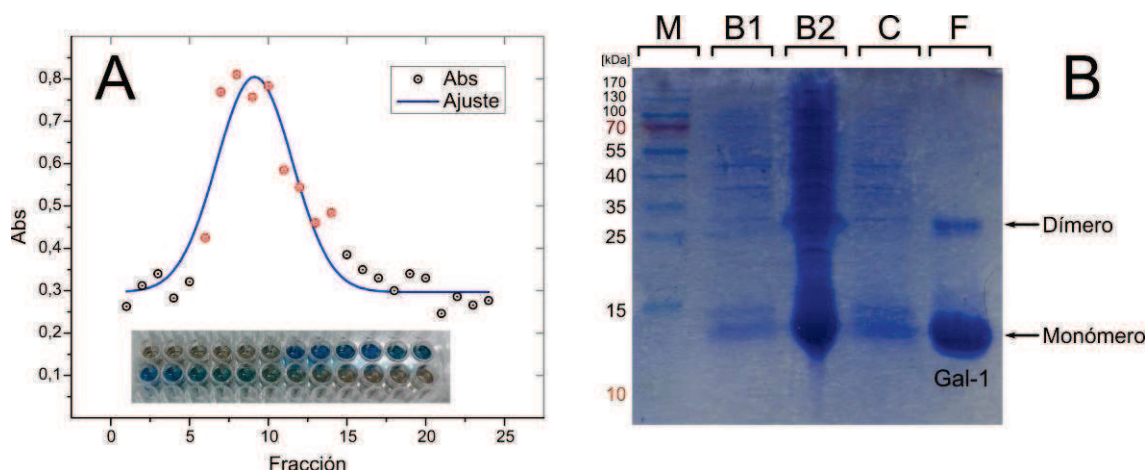


Figura II.5. Elución de Gal-1 de la columna de afinidad. A) Perfil de elución de Gal-1 utilizando el reactivo de Bradford (azul = presencia de proteína) y midiendo la absorbancia de cada fracción eluida (~3 mL por fracción). Los puntos en rojo corresponden a las fracciones utilizadas para los pasos siguientes de purificación, los que contienen altas concentraciones de Gal-1 recombinante; B) Comprobación mediante electroforesis de la efectividad de cada paso de la purificación; calles: M: marcador de peso molecular; B1: lisado de bacterias antes de la inducción con IPTG; B2: lisado de bacterias luego de la inducción; C: lavado de la columna antes de la elución; F: fracción final conteniendo la Gal-1 recombinante pura. Observar la ausencia de Gal-1 durante el lavado de la columna de afinidad (C) y la pureza lograda del eluato (F), así como también la presencia de una banda característica de dímero de Gal-1 tanto en B2 como en F.

Concentración de la Gal-1 recombinante. Luego de todo proceso cromatográfico, las sustancias eluidas sufren dilución con respecto a la concentración que traían originalmente antes de ser sembradas en la columna. Para ello y para contar con volúmenes y concentraciones adecuadas de trabajo, se requiere un paso previo de concentración de las fracciones eluidas de la columna de afinidad. En nuestro caso, empleamos una membrana de diálisis (SnakeSkin, Pierce MWCo 7000 Da) que se carga con la solución de Gal-1. Cerrar los extremos herméticamente con dos broches o clamps y cubrirla con Aquacid II (Calbiochem) en un recipiente que pueda permanecer a 4 °C. Este reactivo no es más que la sal sódica de la carboximetilcelulosa con una gran capacidad de absorción de agua a través de la bolsa de diálisis (1 g de Aquacid puede absorber 5 mL de agua en 1 h). Reducir hasta obtener un 20% del volumen original (6 h aproximadamente). Lavar el exterior de la bolsa y abrir uno de los extremos de la misma para remover con pipeta la solución concentrada de Gal-1 en presencia de lactosa. Almacenar a 4 °C.

Filtrado molecular. El próximo paso de purificación implica un tamiz molecular de Sephadex preparativa HiPrep Sephacryl S-100 HR (GE Healthcare). Para ello, la columna se lava con 3 volúmenes de agua MilliQ y luego se deja DN a 4 °C haciéndole circular buffer de elución. Cargar no más de 5 mL de la solución concentrada de Gal-1 y hacer correr buffer de elución. El volumen muerto de la columna utilizada es de

alrededor 30 mL. Unos mililitros antes de ese volumen, comenzar la recolección de las fracciones y chequear la presencia de Gal-1. Juntar las fracciones positivas y reservar a 4°C. Lavar la columna de filtración molecular con 300 mL de buffer de elución, seguido de 30 mL NaOH 0,4 M y finalmente 4 volúmenes de agua MilliQ. Mantener la columna en un volumen de etanol acuoso al 20% y a 4°C hasta su próximo uso.

Diálisis final, depleción de endotoxinas y liofilización. Luego de la elución final a través de la columna de tamiz molecular, a pesar de encontrarse muy pura, Gal-1 sigue unida a lactosa y en presencia de azida sódica. Para recuperar la forma libre de la proteína, se vuelven a utilizar las bolsas de diálisis y esta vez se realiza una extensa diálisis frente a 2L de PBS con ME 4 mM a 4°C y agitación. Cambiar 3 veces el buffer en un período de 2 días. Para depletar el lipopolisacárido (LPS) y otras endotoxinas provenientes de las bacterias utilizadas para la producción de la proteína recombinante, sólo para aquellas tandas de galectinas que se emplearon para ensayos *in vitro* con células (la presencia de trazas de LPS no afecta significativamente los ensayos biofísicos), se utilizaron unas columnas de Polimixina-B acopladas a una matriz de agarosa diseñadas para tal fin (Sigma) según las indicaciones del fabricante. Finalmente, alicuotar la Gal-1 recuperada en tubos eppendorf de 1,5 mL llenados a la mitad, cerrar, perforar en la tapa y liofilizar DN. La proteína liofilizada se guarda a -20 °C (o a -70°C en caso de almacenamiento prolongado). La pureza e identidad del producto se verificó frente a su comportamiento en un protocolo de *western blot* (**Figura II.6**) y un análisis de espectrometría de masas (LANAIS-UBA).

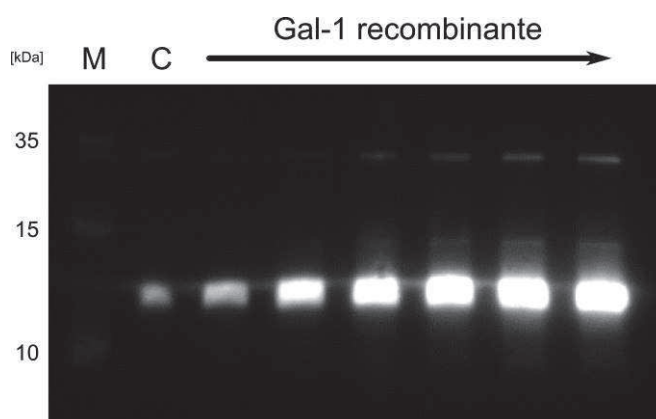


Figura II.6. Análisis de western blot para la Gal-1 recombinante purificada. Las cantidades sembradas por calle fueron 10, 20, 40, 60, 80 y 100 ng. M: marcadores de peso molecular; C: 20 ng de Gal-1 control. Dilución 1:1000 anticuerpo primario anti-Gal-1.

II.2.3 Líneas celulares

Cultivo de Jurkat. Se empleó la línea celular de células T activadas Jurkat (donación Dr. Norberto Zwirner, Laboratorio de Fisiopatología de la Inmunidad Innata, IByME, Buenos Aires) en medio RPMI (Gibco) suplementado con suero fetal bovino (SFB) al 10% y antibióticos/antimicóticos (Gibco).

Cultivo de B16. La línea celular de células melanoma murino no metastático (B16-F0 ATCC CRL-6322) se cultivó en medio DMEM (Gibco), 15% SFB y antibióticos/antimicóticos.

Cultivo de HUVEC. Se emplearon células endoteliales (HUV-ECs, del inglés, *human umbilical vein endothelial cells*) de cultivo primario (donación Dr. Silvio Gutkind, NIH) en medio F-12K (ATCC 30-2004) suplementado con 20% SFB, antibióticos, un cóctel de crecimiento para células endoteliales basado en VEGF (ATCC PCS-100-041), L-glutamina y heparina (ambos de Sigma). Las placas se recubrieron con colágeno I (BD) (solución al 0,1% en ácido acético 0,02 M, 1 h de tratamiento a temperatura ambiente, enjuagadas posteriormente con PBS estéril y reservadas hasta su uso).

En todos los casos, medio completo y DMSO estéril (ATCC 4-X) al 5% y una cámara de 2-propanol fueron utilizados para criopreservar en nitrógeno líquido las células y sus pasajes en crioviales de 2 mL (Corning) mientras no se utilizaban.

II.2.4 Instrumentación

II.2.4.1 Espectrofotometría

La espectroscopía UV-visible utiliza radiación electromagnética de las regiones visible, ultravioleta cercana (UV) e infrarroja cercana del espectroelectromagnético. La energía de la radiación de esta región del espectro es absorbida por las moléculas, lo que provoca diferentes transiciones electrónicas y así, contar con una herramienta de caracterización y cuantificación molecular. Los espectros típicos que se obtienen con esta técnica representan la absorbancia medida a cada longitud de onda. La absorbancia de un compuesto (Abs) a cada longitud de onda (λ) se relaciona con su concentración molar (C), el camino óptico (l) y la absorptividad molar (ϵ) a través de la ley de Lambert-Beer, válida para absorbancias menores a la unidad:

$$\text{Abs}_{\lambda} = \epsilon_{\lambda} l C \quad (\text{II.2.4.1})$$

II.2.4.2 Fluorescencia: espectroscopía y microscopía

La fluorescencia refiere al proceso mediante el cual una molécula absorbe y subsecuentemente emite luz, en un intervalo de tiempo, entre la absorción de la luz de excitación y la emisión de la luz fluorescente, usualmente de pocos nanosegundos. La disponibilidad de equipos capaces de detectar la emisión por fluorescencia permite estudiar materiales fluorescentes, ya sea de manera natural (materiales autofluorescentes) o tratados con sondas fluorescentes. Dos son los equipos particulares que poseen esta capacidad y han sido utilizados en esta Tesis: espectrofluorímetro y microscopio de fluorescencia.

El espectrofluorímetro consiste en una fuente de luz de excitación que puede ser una lámpara de UV o láser y un juego de lentes para dirigir el haz de luz, primero hasta el monocromador de excitación que selecciona la longitud de onda de la luz que incidirá en la muestra, y luego otro que realiza el mismo trabajo para llevar la luz emitida hacia el sistema de detección (estratégicamente ubicado para evitar la llegada de luz de la fuente de excitación). Éste último procesa la radiación recibida y envía dicha información a una computadora donde un software especial grafica el espectro de excitación o emisión correspondiente. De manera similar que para la espectrofotometría, la gráfica del espectro de emisión es la representación de la intensidad de luz emitida contra la longitud de onda de la misma. Para obtener este espectro, se fija la longitud de onda en el monocromador de excitación y se hace un barrido en un intervalo de longitudes de onda para determinar los picos característicos de emisión del material. El equipo de fluorescencia empleado para los ensayos en esta Tesis fue un Jasco FP-6500 (colaboración establecida con el Dr. Julio Caramelo, director del Laboratorio de Biología Estructural y Celular de la Fundación Instituto Leloir, Buenos Aires, Argentina).

La microscopía de fluorescencia es una herramienta de inestimable valor ya que permite alcanzar altos niveles de sensibilidad y resolución, permitiendo una apreciación diferente de la información que se puede obtener de los especímenes y que generalmente pasa desapercibida, siendo hoy una técnica indispensable en biología celular. La principal diferencia del microscopio de fluorescencia con un microscopio óptico convencional es que permite irradiar al espécimen con la luz de excitación y separar la luz fluorescente emitida, que es mucho más débil, de la luz de excitación. Así, sólo la luz emitida por el espécimen es detectada por el ojo o el detector

(usualmente una cámara digital). Como resultado, las partes fluorescentes de la muestra brillan contra un fondo (*background*) oscuro con suficiente contraste como para permitir la detección. Sin embargo, para muestras gruesas (es decir, más grandes que el tamaño de la profundidad de campo, unos 2-3 μm para fluorescencia), las imágenes que se obtienen se presentan muy borrosas debido a la recolección luz de emisión fuera del plano focal. Para el microscopio de fluorescencia convencional utilizado en esta Tesis, una corrección matemática de las imágenes es posible gracias al desarrollo y aplicación de la deconvolución de la imagen mediante procesamiento digital incorporado en el microscopio (DeltaVision Olympus IX70 Deconvolution RT, gentileza PhD. Tatiana Karpova del *NCI Core Fluorescence Imaging Facility*, Building 41, Room C615, NIH, Bethesda, MD, USA) (**Figura II.7**). En otros casos, el uso de técnicas como la microscopía confocal o de dos fotones es altamente recomendable para evitar este inconveniente típico de esta microscopía, sin la necesidad de recurrir al post-procesamiento de la imagen.

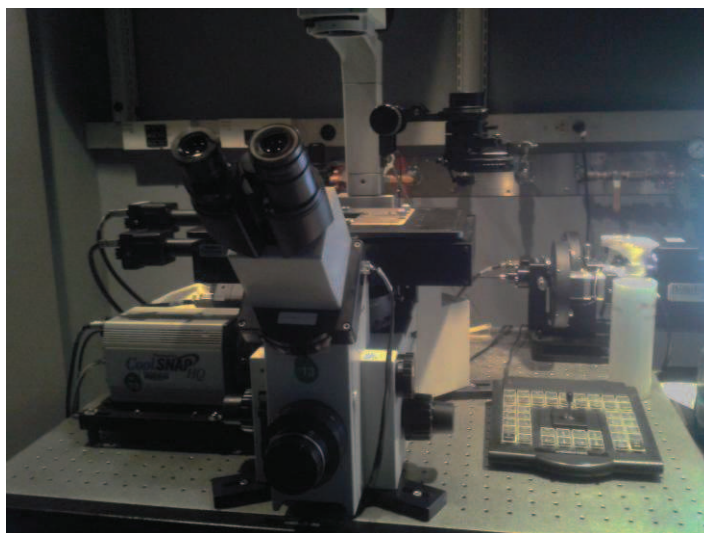


Figura II.7. Microscopio de Fluorescencia. DeltaVision Olympus IX70 Deconvolution RT.

El estudio de fluoróforos intrínsecos (emisión de aminoácidos aromáticos en las proteínas como Phe, Trp y Tyr y cofactores como el NAD, FAD y FMN) y fluoróforos extrínsecos (proteínas verdes fluorescentes y sondas y sensores fluorescentes libres o acoplados a proteínas/anticuerpos) ha permitido identificar células, componentes celulares sub-micrométricos y otras entidades con un alto grado de especificidad [60]. Más aún, esta técnica permite la detección del material fluorescente con una altísima sensibilidad: puede ser detectada una cantidad muy pequeña de moléculas fluorescentes (unas 50 moléculas por μm^3). En una misma muestra, diferentes sondas

fluorescentes permiten apreciar distintos tipos de moléculas gracias a las técnicas de marcado específico.

II.2.4.3 Dicroísmo circular

El dicroísmo circular (DC) es una técnica espectroscópica que se emplea principalmente en el análisis conformacional de sustancias ópticamente activas.

La luz polarizada circularmente tiene naturaleza quiral, y se puede considerar formada por dos componentes que son imágenes especulares una de la otra y que pueden ser diferenciadas por su interacción con moléculas quirales. Esta interacción puede detectarse por dos maneras: por una diferencia en la velocidad de los dos haces de luz a través de la muestra —que da lugar a la llamada dispersión óptica rotatoria—, o bien por la diferente absorción de cada haz por parte de la muestra —originando el dicroísmo circular—. El gran desarrollo y aplicabilidad de la técnica en el estudio de sistemas biológicos permite interpretar directamente los cambios en la estructura secundaria de un péptido o una proteína. En particular, el espectro de DC en la región del ultravioleta lejano de proteínas es muy sensible a cambios en la estructura proteica.

La expresión del dicroísmo circular. Las medidas de DC se realizan determinando la diferencia de absorbancias entre la luz polarizada circularmente hacia la derecha (AD) y hacia la izquierda (AI). Teniendo en cuenta la ley de Lambert-Beer (II.2.4.1), puede expresarse esta diferencia de absorbancias como:

$$\Delta \text{Abs} = \text{AI} - \text{AD} = (\epsilon_I l C) - (\epsilon_D l C) \quad (\text{II.2.4.2})$$

donde nuevamente C es la concentración molar del soluto quiral, ϵ es el coeficiente de extinción molar y l es la longitud del paso de luz [61]. De aquí, se define $\Delta\epsilon = \epsilon_I - \epsilon_D$, de donde resulta

$$\Delta \text{Abs} = \Delta\epsilon l C \quad (\text{II.2.4.3})$$

Al atravesar un medio quiral, los vectores eléctricos (E) de la luz polarizada describen una elipse cuyo eje mayor se mueve hacia un nuevo ángulo de rotación. Cuando los vectores eléctricos de las dos componentes circulares se encuentran en la misma dirección, la suma de sus magnitudes proporciona el semieje mayor de la elipse ($E_I + E_D$); cuando, por el contrario, se encuentran en direcciones opuestas, la diferencia de sus magnitudes proporciona el semieje menor de la elipse ($E_I - E_D$). El DC se define,

precisamente, a partir de la relación entre los semiejes mayor y menor. Esta relación es la tangente del ángulo θ , conocido como *elipticidad* (Figura II.8).

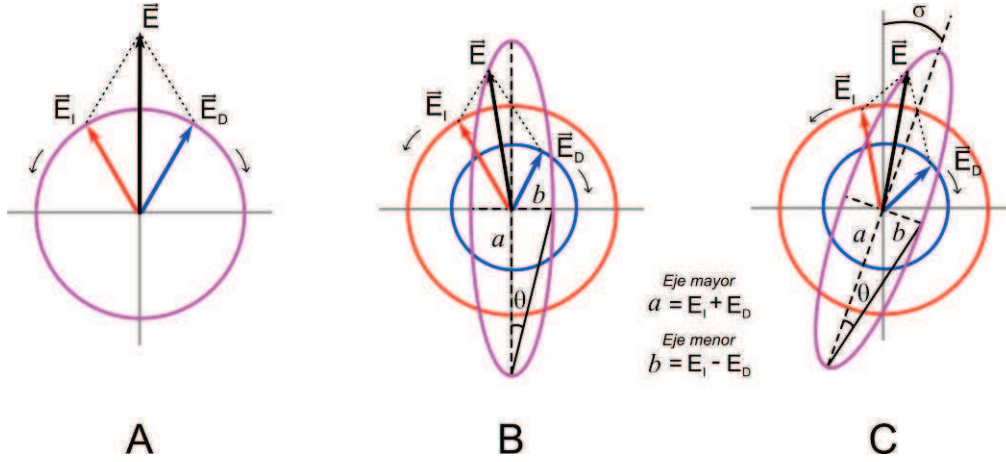


Figura II.8. Representación del fenómeno de dicroísmo circular. A) Luz circularmente polarizada hacia la derecha (E_D) y hacia la izquierda (E_I). Al pasar por una muestra ópticamente activa, la recomposición de los haces produce luz elípticamente polarizada. La diferencia entre la absorción de ambos componentes es la señal de dicroísmo circular (θ). B) Generación de luz elípticamente polarizada sin corrimiento de fase, y C) con un ángulo σ de corrimiento (aquí no sólo la absorción es diferencial en una dirección sino también la fase de determinado vector campo-eléctrico, lo que da origen a este corrimiento angular).

Debido a que este ángulo es por lo general muy pequeño, la tangente de θ es aproximadamente igual a θ , por lo que:

$$\tan \theta (\text{radianes}) \approx \theta (\text{radianes}) = \frac{I_I^{1/2} - I_D^{1/2}}{I_I^{1/2} + I_D^{1/2}} \quad (\text{II.2.4.4})$$

donde se ha tenido en cuenta que la intensidad de la radiación es proporcional al cuadrado de la amplitud del vector eléctrico E . Si se sustituye la intensidad por la definición de absorbancia ($I = I_0 e^{-\text{Abs} \ln 10}$) en II.2.4.4, la elipticidad ahora puede reescribirse como:

$$\theta (\text{radianes}) = \frac{\left(e^{-\frac{\text{Abs}_I}{2} \ln 10} - e^{-\frac{\text{Abs}_D}{2} \ln 10} \right)}{\left(e^{-\frac{\text{Abs}_I}{2} \ln 10} + e^{-\frac{\text{Abs}_D}{2} \ln 10} \right)} = \frac{e^{\frac{\Delta \text{Abs} \ln 10}{2}} - 1}{e^{\frac{\Delta \text{Abs} \ln 10}{2}} + 1} \quad (\text{II.2.4.5})$$

Como $\Delta \text{Abs} \ll 1$, la expresión exponencial se puede expandir en serie de Taylor y truncarla hasta el primer orden. Agregando un factor de conversión de radianes a grados ($180/\pi$), se llega a la expresión:

$$\theta (\text{grados}) = \Delta \text{Abs} \left(\frac{\ln 10}{4} \right) \left(\frac{180}{\pi} \right) \quad (\text{II.2.4.6})$$

Se define la *elipticidad molar* como aquella que no depende de la concentración ni del camino óptico, y que incluye arbitrariamente un factor de 100 para manejar números razonables. Utilizando además la expresión II.2.4.3, queda definida como

$$[\theta] = \frac{100 \theta \text{ (grados)}}{l C} = 100 \Delta\epsilon \left(\frac{\ln 10}{4} \right) \left(\frac{180}{\pi} \right) = 3298 \Delta\epsilon \quad (\text{II.2.4.7})$$

Históricamente, la elipticidad molar se expresa en $^{\circ} \text{ cm}^2 \text{ dmol}^{-1}$. Para calcularla a partir de las mediciones de $\Delta\epsilon$ es necesario conocer la concentración de la muestra (g L^{-1}), el camino óptico l (cm) y la masa molar de la sustancia (g mol^{-1}). Si el compuesto de estudio es una proteína, se suele utilizar la masa molar media residual (promedio de la masa molar de todos los aminoácidos que la componen) en lugar de la masa molar de la proteína, quedando las unidades ahora como $^{\circ} \text{ cm}^2 \text{ dmol}^{-1} \text{ residuo}^{-1}$.

La estructura secundaria/terciaria en espectros de dicroísmo circular. Los cromóforos de una estructura química que presentan un entorno asimétrico interaccionan con la luz polarizada de forma tal que presentan un espectro de DC. En las proteínas y péptidos, los cromóforos mayoritarios son los grupos amida de los enlaces peptídicos (UV-lejano 180-260 nm) y las cadenas laterales de los aminoácidos aromáticos y enlaces disulfuro entre cisteínas (UV-cercano 260-320 nm). Es importante notar que la mayoría de estos grupos químicos no son ópticamente activos por sí mismos, pero es el entorno quiral en el que se encuentran lo que hace que puedan ser detectables por DC.

Las estructuras peptídicas poseen regiones donde los grupos cromóforos se encuentran en disposición altamente ordenada, tales como las hélices α o las hojas β . El enlace amida absorbe la luz en la región del ultravioleta lejano, a una longitud de onda de entre 185 y 250 nm, y en su espectro pueden identificarse dos transiciones electrónicas: a 220 nm, una transición $n \rightarrow \pi^*$, y a 190 nm la correspondiente $\pi \rightarrow \pi^*$. Por lo tanto, es esperable que sea este cromóforo el que domine en el espectro de DC de longitudes de onda del UV-lejano, dando así información sobre la estructura secundaria de la proteína.

Dependiendo del ordenamiento de los enlaces peptídicos, las transiciones ópticas del enlace amida pueden dividirse en múltiples transiciones de diferente frecuencia e intensidad y, en consecuencia, las estructuras secundarias descritas (ya sea hélice α , lámina β , giro β o desordenada) tendrán espectros de DC característicos

(Figura II.9). Se describe a continuación sintéticamente el espectro característico para cada una de estas estructuras:

- *Hélice- α* . Es la estructura secundaria más común, observada con frecuencia en proteínas globulares. El espectro presenta tres bandas características que corresponden a las transiciones $n \rightarrow \pi^*$ (banda negativa a 220-222 nm), $\pi \rightarrow \pi^*$ paralela (banda negativa a 207-209 nm) y $\pi \rightarrow \pi^*$ perpendicular (banda positiva a 191-193 nm). La intensidad de estas tres bandas refleja la proporción de elipticidad en los péptidos analizados.
- *Lámina- β* . El espectro muestra una banda negativa ancha sobre los 216 nm, una banda positiva entre 190 y 200 nm, y otra banda negativa alrededor de 175 nm. Estos espectros presentan mayor variabilidad que los de una hélice α , tanto en la amplitud como en la posición de las bandas.
- *Giro- β* . Las transiciones características de un giro β son la $n \rightarrow \pi^*$, que muestra una banda negativa débil alrededor de los 220-230 nm, y la $\pi \rightarrow \pi^*$ que muestra una banda positiva a 200-215 nm y una banda negativa a 180-200 nm.
- *Desordenada*. El espectro típico de una estructura desordenada —en inglés, ampliamente usado, *random coil*— se caracteriza por una banda negativa intensa alrededor de los 190 nm y otra banda muy débil, negativa o positiva, en la región de los 220-230 nm.

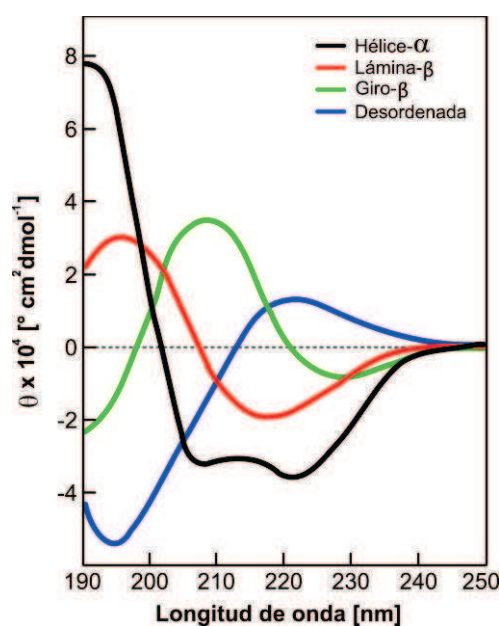


Figura II.9. Dicroísmo circular en proteínas. Apariencia general de los espectros de dicroísmo circular en la región del ultravioleta lejano para las conformaciones más comunes que presentan las proteínas.

Si bien estas bandas han sido caracterizadas para cada tipo de estructura secundaria, en la práctica es difícil interpretar con exactitud el espectro de DC para una proteína. Sin embargo, cuando en una proteína predomina un único tipo de estructura secundaria, el DC permite evaluarlo con certeza.

Cuando se investiga el espectro de DC utilizando radiación UV-cercana, es posible obtener como resultado una medida empírica de la estructura terciaria y conformación de la proteína de estudio, gracias a la absorción diferencial de las cadenas laterales de los aminoácidos Phe (250-270 nm), Tyr (270-290 nm) y Trp (280-310 nm), y también de los enlaces disulfuro entre Cys, que arrojan una señal ancha y débil a lo largo de todo el espectro. Más allá de estas longitudes de onda, algunos grupos prostéticos (como ser el grupo *hemo*) son capaces de mostrar actividad de DC también. A diferencia de los espectros en el UV-lejano, los espectros de DC en el UV-cercano no pueden ser asignados a una estructura 3D proteica en particular. Sin embargo, pueden ser útiles para confirmar un correcto plegamiento de la proteína: si la proteína posee elementos de estructura secundaria pero no presenta una estructura tridimensional definida (mal plegada o de glóbulo-fundido), la señal en esta región del espectro será cercana a cero. Por otro lado, la presencia de una buena señal en esta región espectral es un buen indicador de que la proteína se encuentra plegada en una conformación bien definida. La intensidad en el UV cercano es mucho menor que en el lejano debido a que en una proteína hay menos residuos aromáticos que enlaces peptídicos (los cromóforos en el UV-lejano) y la relación señal/ruido es muy débil para soluciones diluidas. Por ello, estos experimentos requieren mayor concentración de proteína, generando en algunos casos inconvenientes prácticos debido a la solubilidad o agregación de las mismas. Este tipo de DC es como una huella dactilar de la proteína en condiciones nativas (plegada) y ha sido útil en el estudio de interacción proteína-proteína y en el seguimiento de cambios conformacionales y de plegamiento.

El equipo de DC empleado para los ensayos en esta Tesis fue un Jacso J-815 equipado con un controlador de temperatura basado en el efecto Peltier en la celda en donde se coloca la muestra (colaboración establecida con el Dr. Julio Caramelo).

II.2.4.4 Microscopía de transmisión electrónica

La microscopía de transmisión electrónica se basa en la formación de imágenes por contraste producido al incidir sobre la muestra con un haz de electrones, que tiene

la capacidad de atravesar una lámina extremadamente delgada del material en estudio. Dichos contrastes pueden deberse simplemente a la absorción de electrones (modo de operación de campo claro), a la difracción de electrones o bien a diferencias de fases de los electrones causadas por la interacción con la muestra.

La resolución de este tipo de microscopios es notablemente superior a la de un microscopio óptico, ya que en ambos casos el limitante de la resolución está dado por la longitud de onda del haz incidente, siendo la longitud de onda de De Broglie de los electrones menor a la longitud de onda de la luz empleada en los microscopios ópticos. También depende de la potencia del haz incidente. Así, es posible observar muestras delgadas con una resolución muy cercana a los nanómetros de forma aceptable.

El haz de electrones se genera por emisión termoiónica de algún filamento (en general de tungsteno). Los electrones se aceleran por un potencial característico del equipo, viajan a través de una cámara de vacío e inciden sobre la muestra, que debe ser delgada (espesor menor a 200 nm). Los electrones se enfocan con lentes electromagnéticos.

En esta Tesis se utilizó esta técnica para la obtención de imágenes de muestras biológicas (receptores y nanopartículas de oro acopladas a anticuerpos) *in situ* en preparados provenientes de células (ver sección V.3). Estas imágenes fueron recolectadas en las muestras adecuadamente preparadas sobre grillas Quantifoil® Multi A (Electron Microscopy Sciences) de cobre u oro, según el caso, en donde la malla de metal es recubierta con un patrón de diferentes formas y tamaños de hoyos en una delgada capa de carbono (Figura II.10).

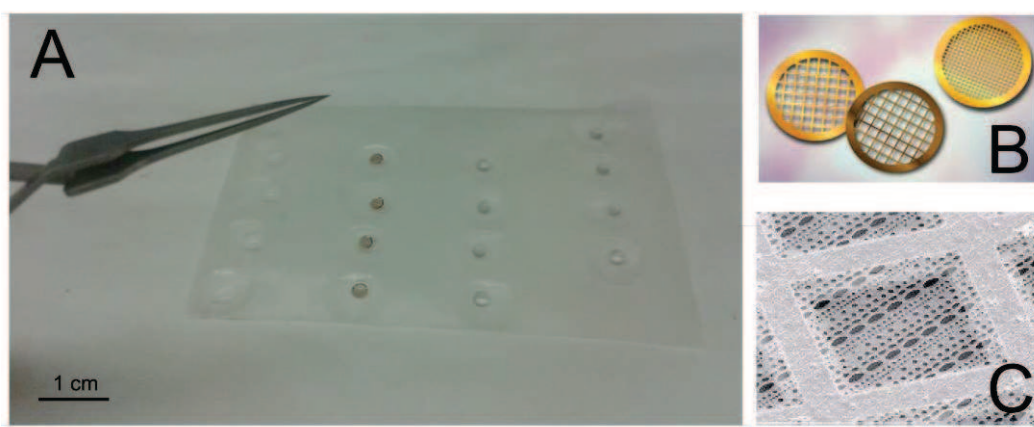


Figura II.10. Grillas de Cu ó Au Quantifoil® Multi A. A) Fotografía de las grillas utilizadas; B) Representación esquemática de la grilla metálica; C) Imagen microscópica del detalle de la malla de carbono que recubre la grilla metálica
(Tomada de <http://www.emsdiasum.com/microscopy/products/grids/quantifoil.aspx>).

Se han obtenido imágenes utilizando dos equipos distintos, ambos ubicados en Biophysics Section and Electron Microscopy Core, Laboratory of Cell Biology, en el National Institute of Health (NIH) Bldg. 50, Bethesda, Maryland, Estados Unidos, como parte de una estadía de tres meses (Beca Fulbright – Bunge y Born):

- Microscopio de Transmisión Tecnai G2 Polara (FEI, The Netherlands) operado a 200 kV y trabajando siempre a temperaturas menores a los 100 K (**Figura II.11A**), equipado con una cámara CCD 2K X 2K al final del sistema de filtración GIF 2000 (Gatan Inc. Pleasanton, CA, USA). En estas condiciones se alcanzan magnificaciones nominales de 30.000X, resultando en un tamaño de pixel de aproximadamente 5 Å en el plano de la muestra.
- Microscopio de Transmisión Tecnai T-12 (FEI, The Netherlands) operando a 120 kV (**Figura II.11B**).

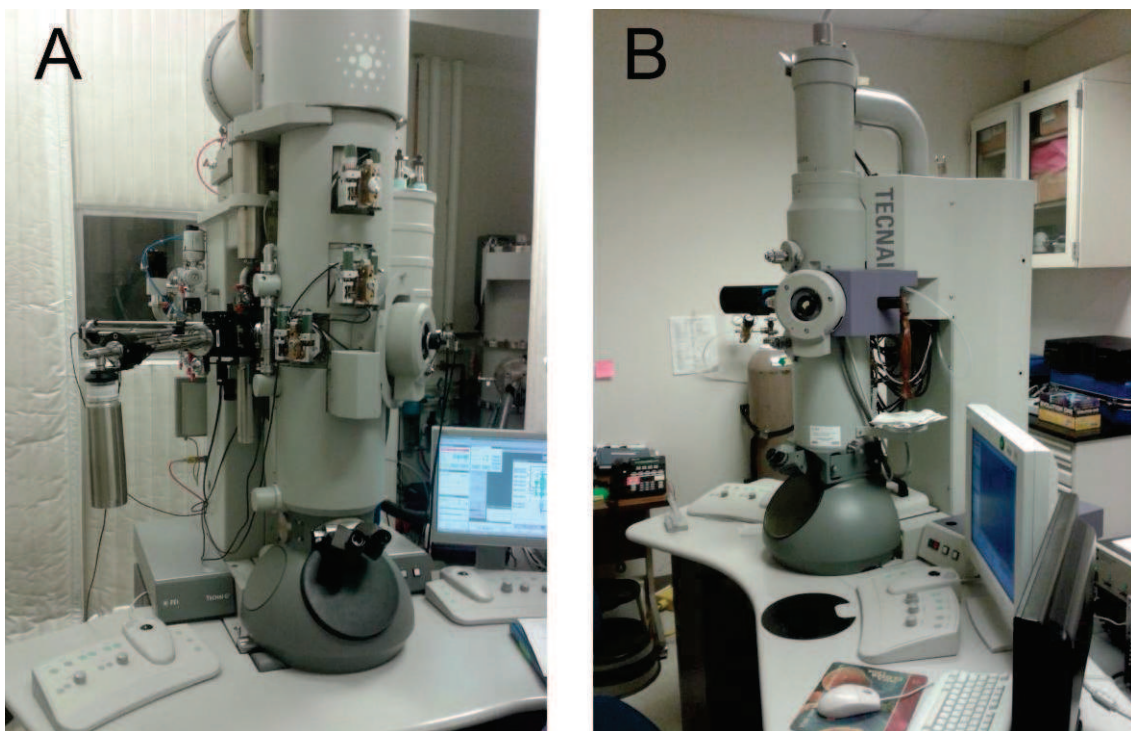


Figura II.11. Microscopios de Transmisión Electrónica. A) Tecnai G2 Polara; B) Tecnai T-12.

II.2.4.5 Citometría de flujo

La citometría de flujo es una técnica muy poderosa que permite evaluar determinadas características físicas y química de las células en suspensión mediante la interacción con la luz. Entrega información sobre el tamaño celular, la granularidad o complejidad interna, así como también de la intensidad de fluorescencia basal que

poseen las células en estudio. Esta tecnología hace posible la medición de los más diversos parámetros celulares como presencia y concentración de antígenos de superficie, citoplasmáticos y nucleares, ácidos nucleicos, actividad enzimática, flujo de calcio, potencial de membrana, pH, entre otros [62].

El principio básico de la citometría de flujo es la interacción de una fuente de luz, específicamente un rayo láser, con células en suspensión, las cuales son canalizadas en un flujo laminar de un fluido soporte y analizadas individualmente por el láser, recogiendo la información obtenida de dicha interacción por un arreglo de detectores y procesamiento computacional de las señales. De este modo, es posible la medición de un gran número de células, en forma individual, en un periodo de tiempo muy corto y detectando más de un parámetro en simultáneo mientras éstas se desplazan en el sistema de flujo constante (llegando hasta valores de 100.000 células por segundo). Cada célula pasa por un punto donde son impactadas por un láser de longitud de onda característico y cuya luz es desviada o alterada de acuerdo a la interacción característica que resulte con cada célula. La variación de longitud de onda así producida es captada y depurada por un complejo sistema de lentes, espejos y detectores, los cuales codifican la señal fotoeléctrica en una señal digital que es analizada por una computadora provista con el software adecuado. Existen equipos que cuentan con uno, dos o más rayos láseres, siendo el más común un láser de Ar. El citómetro empleado para los ensayos en esta Tesis fue un FACSCanto (BD Biosciences).

La interacción de una célula con la luz del láser puede producir el desvío de ésta última en un ángulo menor a 10° lo que se conoce como dispersión frontal o FSC (del inglés, *forward scatter*) y a su vez la desviación de la luz en más de 10° , se conoce como dispersión lateral o SSC (del inglés, *side scatter*). Es así como FSC y SSC son parámetros intrínsecos de la célula, siendo la primera una medida proporcional al tamaño de la partícula que produce la dispersión y la segunda proporcional a la complejidad de la estructura interna de la misma.

Si a la suspensión celular se agrega un fluoróforo excitable con el láser del citómetro que tenga afinidad por algún componente celular, es posible detectar la presencia de dicho componente (y su concentración) mediante el registro de la emisión del fluorocromo. Estratégicamente, y si el equipo posee más de un canal de detección de fluorescencia, es posible la identificación de más de una molécula utilizando la combinación de diferentes fluoróforos para una misma célula, y así reconocer subpoblaciones celulares dentro de la muestra analizada. En los citómetros de flujo

modernos y mayoritariamente empleados en la clínica, es posible hoy en día de reconocer hasta 12 canales de fluorescencia diferentes en forma simultánea, lógicamente usando fluorocromos de diferentes espectros de emisión. Dados los grandes avances en la tecnología de anticuerpos y proteínas conjugadas, y la constante aparición de nuevos colorantes con mejores rendimientos cuánticos, estabilidad y versatilidad óptica, la fenotipificación de células con una aplicación clínica concreta está sobrepasando límites antes impensados [62-65].

En particular, en esta Tesis, se ha utilizado la técnica de citometría de flujo para la medición de apoptosis celular. La anexina-V es una proteína que se une específicamente y con alta afinidad al fosfolípido fosfatidilserina, el cual se encuentra en la cara interna de la membrana plasmática y es expuesto en la cara externa de dicha membrana cuando comienza el proceso de apoptosis. Para la unión de la proteína con el fosfolípido es requerida la presencia de calcio, por lo que será necesario tener en cuenta este requerimiento durante la preparación de la suspensión celular. Uniendo a la anexina-V un marcador fluorescente, por ejemplo isotiocianato de fluoresceína (FITC), podemos detectar la apoptosis por citometría de flujo o por microscopía de fluorescencia. La integridad de la membrana puede ser también evaluada con el marcador catiónico llamado ioduro de propidio (IP) que se intercala en la doble cadena de ADN, permitiendo la diferenciación entre apoptosis temprana (IP negativo, membrana celular aún intacta) y apoptosis tardía/necrosis celular (IP positivo, membrana celular comprometida).

II.3 Referencias

1. Leach, A.R. (2001) *Molecular Modelling: Principles and Applications*. Pearson Education EMA
2. Ponder, J.W., Case, D.A., and Valerie, D. (2003) Force Fields for Protein Simulations. In *Advances in Protein Chemistry*, pp. 27-85, Academic Press
3. Case, D.A., Cheatham, T.E., Darden, T.A., Gohlke, H., Luo, R., Merz, K.M.J., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005) The Amber biomolecular simulation programs. *J. Computat. Chem.* 26, 1668-88
4. Ewald, P.P. (1921) Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik* 369, 253-87
5. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics* 98, 10089-92
6. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G. (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics* 103, 8577-93
7. Darden, T., Perera, L., Li, L., and Pedersen, L. (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* 7, R55-R60

8. Fogolari, F., Brigo, A., and Molinari, H. (2002) The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *Journal of Molecular Recognition* 15, 377-92
9. Shestakov, A.I., Milovich, J.L., and Noy, A. (2002) Solution of the Nonlinear Poisson-Boltzmann Equation Using Pseudo-transient Continuation and the Finite Element Method. *Journal of Colloid and Interface Science* 247, 62-79
10. Lu, B., Zhang, D., and McCammon, J.A. (2005) Computation of electrostatic forces between solvated molecules determined by the Poisson-Boltzmann equation using a boundary element method. *The Journal of Chemical Physics* 122, -
11. Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* 98, 10037-41
12. Still, W.C., Tempczyk, A., Hawley, R.C., and Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *Journal of the American Chemical Society* 112, 6127-29
13. Onufriev, A., Case, D.A., and Bashford, D. (2002) Effective Born radii in the generalized Born approximation: The importance of being perfect. *Journal of Computational Chemistry* 23, 1297-304
14. Langlet, J., Claverie, P., Caillet, J., and Pullman, A. (1988) Improvements of the continuum model. 1. Application to the calculation of the vaporization thermodynamic quantities of nonassociated liquids. *The Journal of Physical Chemistry* 92, 1617-31
15. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* 79, 926-35
16. Jorgensen, W.L. and Madura, J.D. (1985) Temperature and size dependence for Monte Carlo simulations of TIP4P water. *Molecular Physics* 56, 1381-92
17. Levine, I.N. (2001) *Quantum Chemistry*. Pearson Education S.A.
18. Szabo, A. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publ Inc
19. Verlet, L. (1967) Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* 159, 98
20. Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327-41
21. Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81, 3684-90
22. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - Visual Molecular Dynamics. *J. Molec. Graphics* 14, 33-38
23. Chandler, D. (1987) *Introduction to Modern Statistical Mechanics*. Oxford University Press
24. Filipponi, A. (1994) The radial distribution function probed by X-ray absorption spectroscopy. *Journal of Physics: Condensed Matter* 6, 8415
25. Sorenson, J.M., Hura, G., Glaeser, R.M., and Head-Gordon, T. (2000) What can x-ray scattering tell us about the radial distribution functions of water? *The Journal of Chemical Physics* 113, 9149-61
26. Amadei, A., Linssen, A.B.M., and Berendsen, H.J.C. (1993) Essential dynamics of proteins. *Proteins: Structure, Function, and Genetics* 17, 412-25
27. Schlitter, J. (1993) Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chemical Physics Letters* 215
28. Andricioaei, I. and Karplus, M. (2001) On the calculation of entropy from covariance matrices of the atomic fluctuations. *Journal of Chemical Physics* 115
29. Harris, S.A., Gavathiotis, E., Searle, M.S., Orozco, M., and Laughton, C.A. (2001) Cooperativity in drug-DNA recognition: A molecular dynamics study. *Journal of the American Chemical Society* 123, 12658-63
30. Lee, B. and Richards, F.M. (1971) The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology* 55, 379-IN4
31. Shrake, A. and Rupley, J.A. (1973) Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *Journal of Molecular Biology* 79, 351-71

32. Knapp, B., Lederer, N., Omasits, U., and Schreiner, W. (2010) vmdICE: A plug-in for rapid evaluation of molecular dynamics simulations using VMD. *Journal of Computational Chemistry* 31, 2868-73
33. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14, 33-38
34. Hou, T., Wang, J., Li, Y., and Wang, W. (2010) Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* 51, 69-82
35. Kongsted, J., Söderhjelm, P., and Ryde, U. (2009) How accurate are continuum solvation models for drug-like molecules? *Journal of Computer-Aided Molecular Design* 23, 395-409
36. Swanson, J.M.J., Henchman, R.H., and McCammon, J.A. (2004) Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophysical Journal* 86, 67-74
37. Genheden, S. and Ryde, U. (2010) How to obtain statistically converged MM/GBSA results. *Journal of Computational Chemistry* 31, 837-46
38. Gohlke, H. and Case, D.A. (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry* 25, 238-50
39. Page, C.S. and Bates, P.A. (2006) Can MM-PBSA calculations predict the specificities of protein kinase inhibitors? *Journal of Computational Chemistry* 27, 1990-2007
40. Srinivasan, J., Cheatham, T.E., Cieplak, P., Kollman, P.A., and Case, D.A. (1998) Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *Journal of the American Chemical Society* 120, 9401-09
41. Gohlke, H., Kiel, C., and Case, D.A. (2003) Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *Journal of Molecular Biology* 330, 891-913
42. Rastelli, G., Rio, A.D., Degliesposti, G., and Sgobba, M. (2010) Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of Computational Chemistry* 31, 797-810
43. Tidor, B. and Karplus, M. (1994) The Contribution of Vibrational Entropy to Molecular Association: The Dimerization of Insulin. *Journal of Molecular Biology* 238, 405-14
44. Karplus, M. and Kushick, J.N. (1981) Method for estimating the configurational entropy of macromolecules. *Macromolecules* 14, 325-32
45. Brooks, B.R., Janežič, D., and Karplus, M. (1995) Harmonic analysis of large systems. I. Methodology. *Journal of Computational Chemistry* 16, 1522-42
46. Wang, W. and Kollman, P.A. (2000) Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. *Journal of Molecular Biology* 303, 567-82
47. Epstein, C.J., Goldberger, R.F., and Anfinsen, C.B. (1963) The Genetic Control of Tertiary Protein Structure: Studies With Model Systems. *Cold Spring Harbor Symposia on Quantitative Biology* 28, 439-49
48. Chothia, C. and Lesk, A.M. (1986) The relation between the divergence of sequence and structure in proteins. *The EMBO Journal* 5, 823-26
49. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-80
50. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403-10
51. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389-402
52. Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.-y., Pieper, U., and Sali, A. (2006) Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*
53. Martí-Renom, M.A., Stuart, A.C., Fiser, A., Sánchez, R., Melo, F., and Šali, A. (2003) Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure* 29, 291-325

-
54. Fiser, A. and Šali, A. (2003) Modeller: Generation and Refinement of Homology-Based Protein Structure Models. In *Methods in enzymology* (Charles W. Carter, Jr. and Robert, M.S., eds), pp. 461-91, Academic Press
55. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26, 283-91
56. Melo, F. and Feytmans, E. (1998) Assessing protein structures with a non-local atomic interaction energy. *Journal of Molecular Biology* 277, 1141-52
57. Vriend, G. (1990) WHAT IF: A molecular modeling and drug design program. *Journal of Molecular Graphics* 8, 52-56
58. Eisenberg, D., Lüthy, R., and Bowie, J.U. (1997) [20] VERIFY3D: Assessment of protein models with three-dimensional profiles. In *Methods in enzymology* (Charles W. Carter Jr, R.M.S., ed), pp. 396-404, Academic Press
59. Pace, K.E., Hahn, H.P., and Baum, L.G. (2003) Preparation of recombinant human galectin-1 and use in T cell death assays. *Methods Enzymology* 363, 499-518
60. Lakowicz, J. (1999) Principles of Fluorescence Spectroscopy. Plenum, New York
61. Schellman, J.A. (1975) Circular dichroism and optical rotation. *Chemical Reviews* 75, 323-31
62. Krutzik, P.O., Irish, J.M., Nolan, G.P., and Perez, O.D. (2004) Analysis of protein phosphorylation and cellular signaling events by flow cytometry: techniques and clinical applications. *Clinical Immunology* 110, 206-21
63. Meyer, M., Scheper, T., and Walter, J.-G. (2013) Aptamers: versatile probes for flow cytometry. *Applied Microbiology and Biotechnology* 97, 7097-109
64. Pedreira, C.E., Costa, E.S., Lecrevisse, Q., van Dongen, J.J.M., and Orfao, A. (2013) Overview of clinical flow cytometry data analysis: recent advances and future challenges. *Trends in biotechnology* 31, 415-25
65. Laerum, O.D. and Farsund, T. (1981) Clinical application of flow cytometry: A review. *Cytometry* 2, 1-13



LA RELACION ESTRUCTURA-FUNCION EN LA FAMILIA DE LAS GALECTINAS HUMANAS

III.1 Introducción

Hasta la fecha, las galectinas más estudiadas y caracterizadas a nivel de estructura han sido galectina-1 (Gal-1) [1, 2], galectina-3 (Gal-3) [3, 4] y galectina-9 (Gal-9) [5-7]. Se han obtenido varios cristales de estas proteínas en diferentes condiciones, es decir, unidos a diferentes ligandos y de varias mutantes. En la mayoría de los casos, Gal-1 es siempre tomado como referencia para comparar la actividad de los otros miembros de esta familia.

Recientemente, la estructura cristalina del dominio *N*-terminal de Gal-9 (Gal-9N) en presencia de poli-*N*-acetil-lactosamina [6] y del extremo *C*-terminal de galectina-9 (Gal-9C) en complejo con un oligosacárido biantenarico y sialillactosa ha sido reportada [7], sumando nuevos aportes al entendimiento del proceso de unión en galectinas. También, la estructura del dominio *C*-terminal de una proteína humana tipo-galectina (hGRPC, del inglés, *human galectin-related protein C-terminus*, anteriormente conocida como HSPC159 por *hematopoietic stem cell precursor*) se obtuvo en un estado libre (sin ligandos) [8, 9]. Es interesante que no se ha encontrado actividad de lectina para esta proteína en particular, a pesar de poseer el plegado característico de galectinas. Con respecto a los otros miembros de la familia de galectinas humanas, el dominio *N*-terminal de la lectina tipo repeticiones en tándem galectina-4 (Gal-4N) de ratón ha sido cristalizado en una iniciativa de proteómica [10] mientras que la estructura del extremo *C*-terminal fue obtenida por espectroscopía de RMN [11]. Una situación similar se observa con galectina-8 (Gal-8), donde la estructura del extremo *N*-terminal (Gal-8N) se encuentra disponible en presencia [12] y ausencia [13] de lactosa, mientras que el dominio *C*-terminal (Gal-8C) se obtuvo por RMN [14]. Las estructuras de galectina-2 (Gal-2) [15], galectina-7 (Gal-7) [16] y galectina-10 (Gal-10) [17] también

han sido reportadas. Ninguna de todas estas estructuras ha sido descripta en detalle o comparadas con todas las otras estructuras hasta el momento en que se ha comenzado este trabajo de Tesis. Y eso es debido a que no se encuentran estructuras experimentales para los miembros restantes de la familia de galectinas humanas, las cuales incluyen Gal-4N, Gal-12 completa, Gal-13 (o también PP13, del inglés *placental protein 13*) y Gal-14 (o PPL13, del inglés *PP13-like protein*). De hecho, los intentos para completar los estudios estructurales de los miembros recientemente descubiertos sólo han sido posibles gracias al uso de las herramientas de simulación computacional (cuando no existe aun información experimental directa), demostrando la importancia de los estudios sistemáticos de estructura-función. Éste es el caso de Gal-13 (PP13) [18, 19] donde modelos por homología son utilizados como única aproximación posible al problema del análisis de la estructura.

Siguiendo la propuesta de Hirabayashi [20], las galectinas han sido clasificadas entre los tres tipos discutidos en la sección I.2.4, según la organización de los Dominios de Reconocimiento de Carbohidratos (DRC) en la estructura global de la proteína. Aunque es operativa, esta clasificación no tiene en cuenta la filogenia entre los distintos miembros de la familia. Este tema particular ha sido considerado en un estudio reciente [21], donde la comparación de secuencias y un análisis completo de la posición de intrones/exones revelan una clasificación natural de las galectinas en dos grupos: DRC del tipo **F3** y **F4** (**Figura III.1**). De este análisis surge el resultado interesante de que las galectinas tipo repeticiones en tándem están compuestas por un DRC de cada tipo, formando así las galectinas del tipo **bi-DRC**. El mapa filogenético proveniente de este estudio e integrando todas las galectinas humanas y la hGRPC se muestra en la **Figura III.2**.

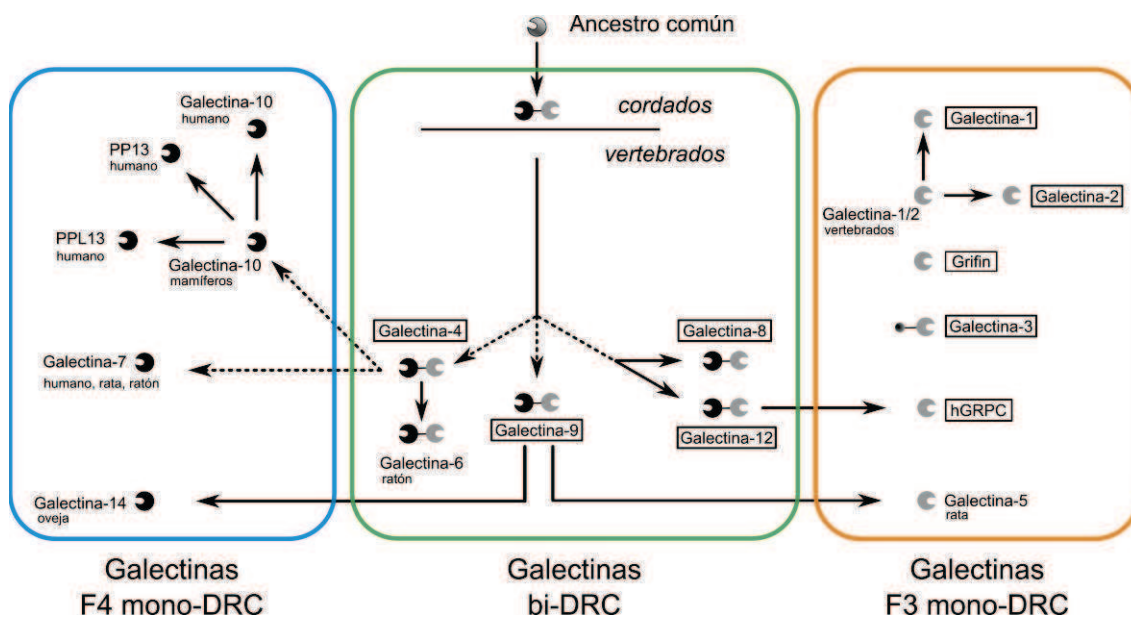


Figura III.1. Relaciones filogenéticas propuestas entre galectinas, basadas en ubicaciones de los genes, estructura de intrones-exones y árbol filogenético de secuencias. Las flechas de puntos indican desconocimiento en la secuencia de divergencia. Los dos tipos de DRC se muestran en negro (F4) o en gris (F3). Las galectinas encerradas en recuadros están identificadas en una gran variedad de especies. Aquellas que no están recuadradas sólo han sido descritas en un subtipo particular de especies (indicadas bajo el nombre de cada galectina). Adaptado de [21].

Como ha sido mencionado previamente, de un total de 16 secuencias conocidas que codifican galectinas, solo 9 han sido estructuralmente descritas en humanos. Es más, no se ha realizado aún un estudio estructural a nivel comparativo del plegamiento completo ni de los surcos de unión a ligando (LBG, del inglés *Ligand Binding Groove*) de cada galectina. Es por eso que hemos encontrado en ese vacío de información la motivación de los estudios realizados usando herramientas computacionales que han sido aplicadas durante las últimas décadas como herramienta fundamental complementaria a la información experimental [22, 23]. Partiendo de las estructuras disponibles de las galectinas humanas y realizando modelos por homología (sección II.1.6) para las restantes, se llevaron a cabo simulaciones de DM en ausencia y presencia de diferentes ligandos, con el fin de analizar las interacciones proteína-glicano, la dinámica de dicha unión y caracterizar la unión mediante estudios de energía libre de unión.

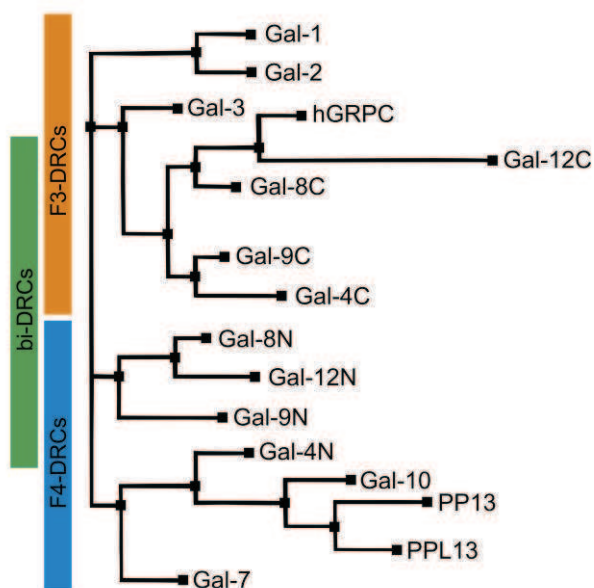


Figura III.2. Árbol filogenético de las galectinas humanas, basado en análisis de secuencias. Este es un esquema cualitativo reducido del árbol filogenético completo presentado en [21]. La figura es usada para presentar la familia de galectinas humanas, su clasificación y las relaciones filogenéticas entre ellas.

III.2 Métodos computacionales

Estructuras iniciales. Las estructuras iniciales de cada DRC de todas las galectinas humanas se obtuvieron de la PDB (<http://www.pdb.org>) (Tabla III.1) en caso de estar disponibles.

Galectina	Grupo	PDB ID	Resolución [Å] ^a
Gal-1	F3 mono-DRC	1GZW	1,65
Gal-2	F3 mono-DRC	1HLC	2,90
Gal-3	F3 mono-DRC	1A3K	2,10
Gal-4C	bi-DRC	1X50	N.A.
Gal-7	F4 mono-DRC	1BKZ	1,90
Gal-8N	bi-DRC	2YV8	1,92
Gal-8C	bi-DRC	2YRO	N.A.
Gal-9N	bi-DRC	2ZHM	1,84
Gal-9C	bi-DRC	3NV1	1,50
Gal-10	F4 mono-DRC	1LCL	1,80
hGRPC	F3 mono-DRC	3B9C	1,90

Tabla III.1. Estructuras de galectinas humanas disponibles, determinadas por experimentos de rayos X o RMN. ^aN. A. No Aplica, experimento de RMN.

Para aquellas galectinas cuyo DRC no estaba disponible, sus estructuras se construyeron por modelado por homología (sección II.1.4) usando el programa Modeller. Para obtener el templado (*template*) adecuado para cada proteína objetivo (*target*), estructuras similares se han buscado usando el servidor ModWeb para modelado de estructuras proteicas. (<http://modbase.compbio.ucsf.edu/scgi/modweb.cgi>) (Tabla III.2). Alineamientos de todas las secuencias de galectinas se realizaron utilizando el método y programa

CLUSTAL-W, usando la matriz de sustitución BLOSUM62. Comenzando con estos *pairwise target templates*, los alineamientos fueron determinados y refinados, poniendo especial atención al plegamiento particular del DRC de las galectinas. Todos los árboles de alineamiento fueron construidos con el método de la matriz de distancia, usando como distancia entre dos DRCs, su valor mínimo de RMSD (sección II.1.3.2).

Galectina objetivo	Galectina de templado [PDB ID]	Grupo
Gal-4N	Gal-4N de ratón [2DYC]	bi-DRC
Gal-12N	Gal-8N [2YV8]	bi-DRC
Gal-12C ^a	-	bi-DRC
PP13	Gal-10 [1G86]	F4 mono-DRC
PPL13	Gal-10 [1G86]	F4 mono-DRC

Tabla III.2. Pares templado-objetivo elegidos para modelar por homología las estructuras de las galectinas humanas faltantes. ^aDado que el extremo C-terminal de Gal-12 presenta menos del 12% de identidad de secuencia con respecto a las otras galectinas humanas (**Figura III.2**) y los modelos por homología obtenidos mostraron alta flexibilidad en el loop S5S6 (debido a la presencia de dos triptófanos muy cercanos entre sí), no se ha podido obtener en forma satisfactoria una estructura estable para estudiar con el resto de los DRCs.

Simulaciones de DM. A partir del cristal, estructura por RMN o de la estructura modelada por homología, cada DRC fue sujeto al siguiente protocolo: los átomos de H se agregaron automáticamente. El estado de protonación estándar a pH fisiológico se aplicó a todos los aminoácidos protonables. Las estructuras se solvataron en una caja de solvente explícito octaédrico con moléculas de agua tipo TIP3P (sección II.2.3), localizando los límites de la caja a 10 Å de la superficie de la proteína. La cantidad total de átomos en cada sistema simulado va desde 15000 a 18000 átomos aproximadamente, teniendo en cuenta el solvente. Simulaciones de DM se llevaron a cabo a 1 atm y 300 K (barostato y termostato de Berendsen) (sección II.1.3.2), usando condiciones periódicas de borde y suma de Ewald para tratar las interacciones de largo alcance con un radio de corte de 10 Å y contabilizar interacciones directas. El algoritmo SHAKE se aplicó a todos aquellos enlaces involucrando átomos de H, permitiendo así el uso de un time-step de 2 fs para la integración de las ecuaciones de Newton, parámetros recomendados en el paquete de programas de AMBER. El campo de fuerzas f99SB se utilizó para todos los residuos proteicos [24]. Para los ligandos (glicanos) se emplearon los parámetros provenientes de la base de datos GLYCAM [25, 26]. Para equilibrar el sistema se realizaron 8000 pasos de minimización, seguidos de 500 ps de una corrida a volumen constante para termalizar el sistema lentamente hasta 300 K. Finalmente, una corrida de 1 ns de duración se llevó a presión constante para alcanzar la densidad adecuada (1 g mL⁻¹). Para las estructuras experimentalmente determinadas las DMs de producción fueron de 50 ns, mientras que para los modelos

construidos por homología, estas DMS duraron 60 ns. Fotos (set de coordenadas y velocidades a un tiempo particular) se colectaron a intervalos de 1 ps y se utilizaron para el análisis posterior.

Dinámica esencial y Entropía conformacional. Estos métodos fueron descriptos en las secciones II.1.3.2.

Construcción *in silico* del complejo Gal-9N con un octasacárido. Nagae y col han reportado diferentes estructuras cristalinas del extremo *N*-terminal de Gal-9 en complejo con diferentes oligómeros de *N*-acetillactosamina (LacNAc) [6]. Ellos encontraron en particular dos cristales con diferentes estructuras: dímeros de LacNAc en el LBG (PDBIDs 2ZHK y 2ZHL) y trímeros de LacNAc (PDBIDs 2ZHM y 2ZHN). En la forma cristalina 2ZHM, encontraron cuatro complejos proteína-ligando (designados A-D) que fueron divididos en dos grupos según la forma en que el ligando interactúa con la proteína: las moléculas de Gal-9N en los complejos A y B reconocen el extremo reductor del azúcar LacNAc, mientras que en los complejos C y D interactúan con la porción central de la molécula de LacNAc. En la estructura cristalina 2ZHN, sólo se observa un complejo donde la proteína reconoce el extremo no reductor del trímero de LacNAc. De este modo, alineando con el programa VMD el complejo A del cristal 2ZHM con el complejo del cristal 2ZHN, se pudo recuperar un **modelo de unión a poli-*N*-acetillactosamina que ocupa el LBG completo de una galectina**. Las estructuras de las proteínas se superpusieron con un valor de RMSD de 0,27 Å, mientras que las unidades del extremo no reductor del complejo A de 2ZHM con la primera y segunda unidad de LacNAc del extremo no reductor del complejo de 2ZHN se superpusieron casi perfectamente (**Figura III.3**). A partir de esta superposición estructural, se almacenó la posición de cada átomo en un único archivo para generar un modelo *in silico* de Gal-9N unido al polisacárido tetra-LacNAc que ocupa el LBG completo.

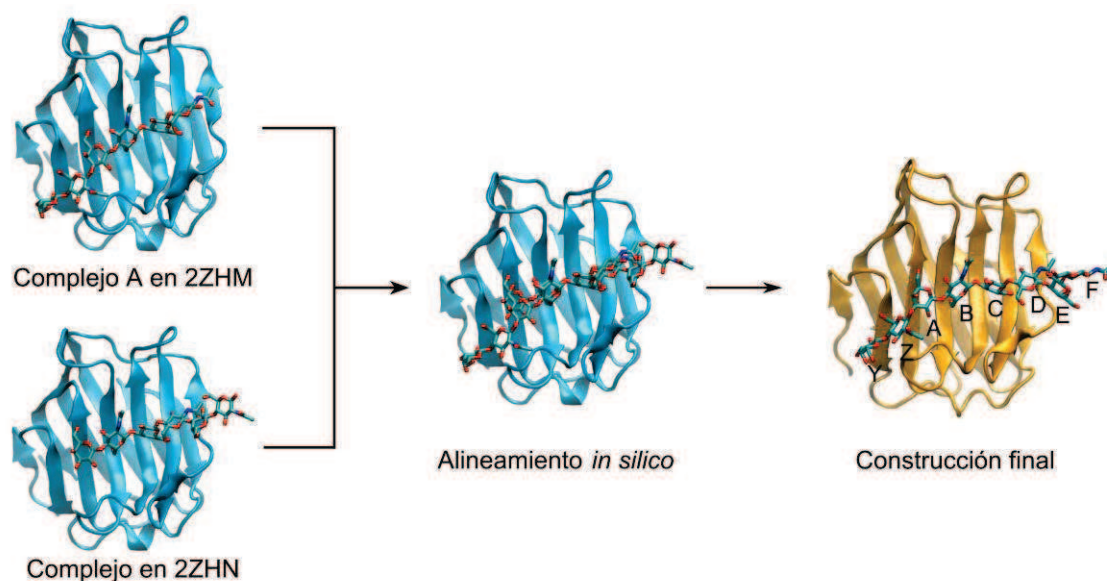


Figura III.3. Construcción *in silico* del complejo tetra-LacNAc unido a Gal-9N. Elección de los complejos adecuados de Gal-9N con sus ligandos y estructura final resultante del alineamiento estructural, mostrando la total ocupación del LBG y los MBSs Y-F. El ligando octamérico consiste en cuatro subunidades repetidas del disacárido *N*-acetilactosamina (LacNAc: β -D-galactopiranosil-(1 $\rightarrow$ 4)-*N*-acetil- D-glucosamina), unidas por un enlace β ,1 $\rightarrow$ 3.

Definiendo los sitios individuales de unión a monosacáridos (MBS, del inglés *monosaccharide binding sites*) y cálculos de energía libre de unión. Dos enfoques se emplearon en este apartado. Uno involucra la definición topológica de los MBS a partir del modelo *in silico* de Gal-9N unido a tetra-LacNAc y la comparación de ocupación de cada MBS por las otras galectinas. El otro involucra el desarrollo de una metodología de comparación basada en la construcción de otros modelos de unión (utilizando un approach de superposición de estructuras muy similar al empleado arriba), enfrentando todas las galectinas con un set de ligandos biológicamente relevantes al mejor estilo de un *in silico glyco-array*. Se describirán por separado y en detalle ambos enfoques a continuación:

- Los ocho MBS se denominan Y, Z, A, B, C, D, E y F, siendo el sitio D en el cual generalmente se ocupa el extremo reductor del galactósido (**Figura III.3**). Típicamente, las unidades de LacNAc se unen a las galectinas ocupando los sitios C y D, como se ha observado para LacNAc en la mayoría de las estructuras en complejo para diferentes galectinas [27]. Los sitios E y F están casi completamente fuera de la proteína e inmersos en el solvente, tal como puede observarse en la **Figura III.3**; estos sitios no se incluyeron en el análisis de la estructura pero sí en el de los cálculos de energía libre. Para definir qué aminoácido contribuye a cada MBS, se alinearon todos los DRCs con el complejo de Gal-9N/octasacárido y se definieron con el siguiente

criterio topológico: un residuo contribuye al *i*-ésimo MBS cuando la distancia entre cualquier átomo pesado del residuo (en el DRC particular de estudio) se encuentra a menos de 5 Å de cualquier átomo pesado del ligando del complejo de Gal-9N. Los parámetros termodinámicos para cuantificar esta interacción a partir de la simulación por DM fueron obtenidos a partir del método MM-GBSA (sección II.1.5.2). Se registraron las energías electrostáticas (ELE) y de van der Walls (vdW), así como la contribución a la energía libre por solvatación (GB_{solv}) y el resultado de la contribución total del modelo a la unión ($\Delta G_{\text{unión}}$).

- Para poder comprender en profundidad el origen molecular de la selectividad entre galectinas, hemos desarrollado un *screening* de cada LBG frente a la capacidad de unión con un set variado y determinado de ligandos (*glyco-array in silico*). Partiendo de una estructura base/molde de la molécula de LacNAc, se han ido reemplazando estratégicamente determinadas posiciones de éste disacárido por otros monosacáridos o grupos sulfato, para contar con un total de 11 ligandos de relevancia biológica y con los cuales todas las galectinas, en principio, deberían interactuar de una u otra forma (son todos derivados β -galactósidos) (**Figura III.4**).

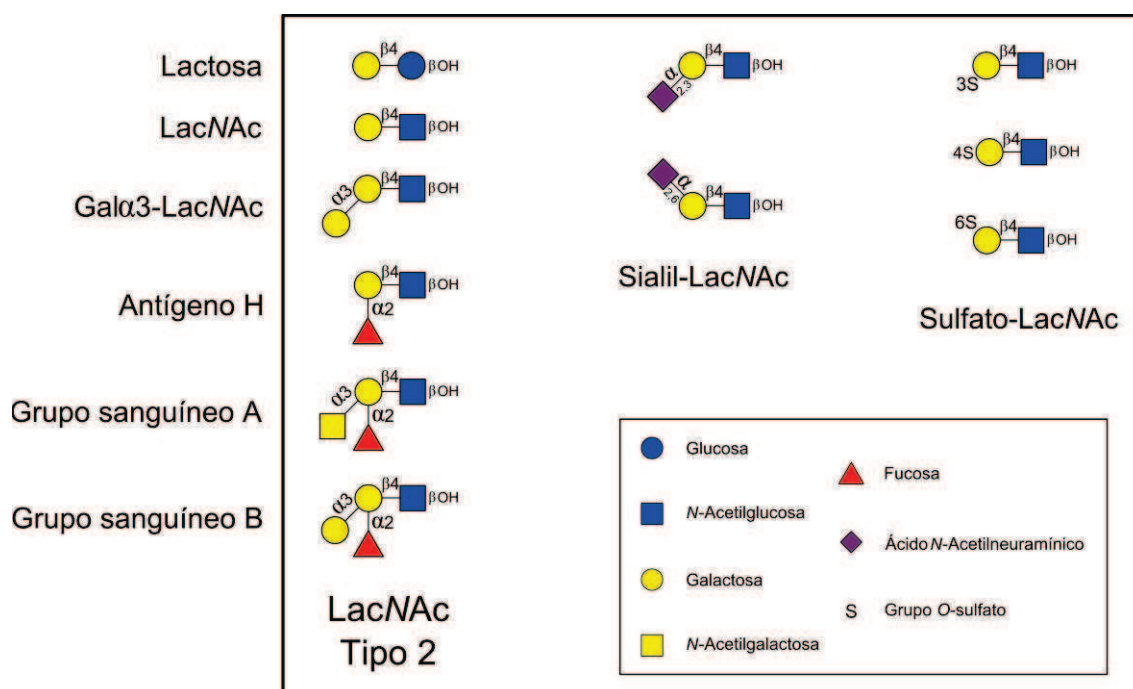


Figura III.4. Representación esquemática de los glicanos utilizados en el *glycoarray in silico*. Se utilizaron sólo los derivados tipo 2 de LacNAc por disponibilidad de estructuras tridimensionales experimentales sólo para estos glicanos. En un recuadro interior se encuentra la simbología empleada para cada monosacárido y en cada estructura está detallado el tipo de enlace que los conecta.

Hemos seleccionado estos ligandos basándonos en un criterio puramente biológico: estos son los β -(1-4)-galactósidos más comunes y presentes en el cuerpo humano [28], y por ende potenciales ligandos *in vivo* para todas las galectinas. En todos los casos, los conformeros más estables de estos glicanos fueron elegidos a partir de los criterios impuestos en la base de datos GLYCAM (ángulos y torsiones más estables en vacío). Una vez que se obtuvo la estructura de cada azúcar derivada de la molécula de LacNAc, y partiendo de algún complejo Gal: β -galactósido experimental, se realizó un alineamiento/superposición estructural de este complejo proteico con el azúcar que se deseó estudiar. Entonces, con la ayuda del programa VMD, fue posible “acomodar” el nuevo azúcar en la posición adecuada del LBG de la galectina bajo estudio, eliminar el ligando original y dejar en su posición, como resultado del mejor alineamiento estructural posible, el nuevo ligando que se desea estudiar (de algún modo similar a lo realizado en la sección anterior para construir el complejo Gal-9N:tetraLacNAc). En los casos donde no se contaba con una estructura experimental de partida de la galectina con un β -galactósido (Gal-8C y Gal-10), se alineó también la estructura proteica al molde de Gal-1:Lactosa. De esta manera, se construyeron *in silico* (y una por una) las 99 estructuras tridimensionales de partida (9 galectinas x 11 ligandos), se solvataron con moléculas de agua explícitas TIP3P y se las preparó para realizar una DM de cada una de ellas.

Dado que esta especie de docking manual puede no arrojar la estructura del complejo real y además, sabiendo que las lectinas pueden unir el ligando pero en una conformación ligeramente diferente de la que los glicanos poseen en vacío/solución, se utilizó un protocolo particular de minimización/termalización diferente del anteriormente descrito y utilizado para el análisis de la estructura secundaria de las galectinas. Aquí se emplea el protocolo de *simulated annealing* [29, 30]. Luego de unos 5000 pasos de minimización del sistema completo (*steepest descent* los 10 primeros pasos y luego *conjugate gradient*, ver sección II.1.3.1), el protocolo implica un calentamiento rápido del sistema (Gal-X:Lig Y + solvente explícito TIP3P) de 1 a 450 K en 2500 pasos, aplicando una restricción al movimiento del esqueleto (*backbone*: átomos pesados del enlace amídico (-N-CO-) y el C $_{\alpha}$ de cada cadena lateral) de la proteína y permitiendo que el sistema permanezca a esta temperatura durante 45 ps de simulación. De esta manera, la restricción evita la desnaturalización de la proteína durante el sobre-calentamiento del sistema y al mismo tiempo, dada la elevada temperatura asignada en este intervalo de tiempo de simulación, las cadenas laterales

de los aminoácidos y todo el azúcar poseen la suficiente energía cinética para visitar muchísimas conformaciones y tener más oportunidades de encontrarse con la mejor conformación para el complejo. Inmediatamente después de transcurrida esta etapa de calentamiento, el sistema se fuerza a un enfriamiento lento y controlado (3 K/ps) hasta la temperatura final de simulación (300 K). A continuación, se corre 1 ns de estabilización del sistema a NPT y finalmente se registran 3 ns de DM en las mismas condiciones para cada complejo, con el fin de obtener la estructura final del complejo y suficiente estadística para realizar los estudios de comparación de unión utilizando MM-GBSA como criterio de discriminación de la afinidad.

III.3 Resultados

Para poder lograr una conexión entre estructura y función de diferentes miembros de la familia de galectinas humanas, en primer lugar comparamos la estructura global de plegamiento de los DRCs de todas las galectinas. Luego, analizamos el comportamiento dinámico de los mismos y finalmente, se realizó un estudio comparativo detallado de los LBG y selectividad de cada DRC.

III.3.1 Comparación global de las galectinas

La estructura global del DRC en una galectina consiste en 135-140 aminoácidos en un arreglo tipo β -sándwich a nivel estructura terciaria, formado por dos láminas β sutilmente dobladas (**Figura III.5**). El sitio cóncavo de la proteína (**cara S**) está constituido por seis hebras llamadas S1-S6 y la cara convexa (**cara F**) por cinco hebras llamadas F1-F5, excepto en algunos miembros de la familia que poseen una hebra F extra al comienzo de la secuencia proteica (Gal-3 y Gal-8N, por ejemplo). A lo largo de la secuencia primaria, las hebras están organizadas según S1-F2-S3-S4-S5-S6-F3-F4-F5-S2-F1, con loops de longitud variable interconectándolas. Los loops serán identificados como las dos hebras beta que el loop conecta (por ejemplo loop "S5S6" es el loop que conecta a la hebra S5 con la S6, siempre yendo desde el extremo N- al C-terminal). La cara S forma el "surco" en el cual reposa el carbohidrato mientras se une a la lectina (LBG).

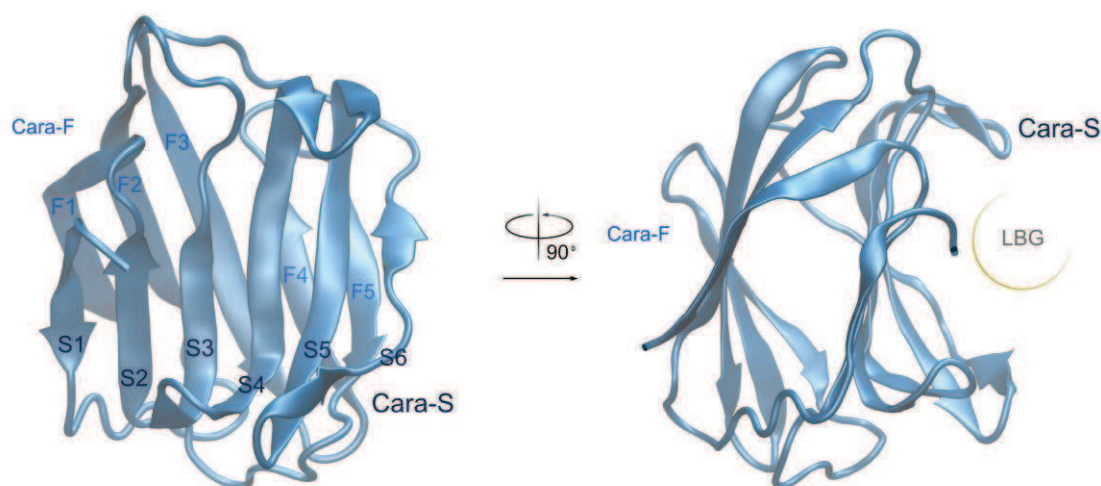


Figura III.5. Representación esquemática de la estructura tridimensional de Gal-1. La estructura de Gal-1 se muestra como un ejemplo del plegamiento conservado para las galectinas humanas. Se muestra también la nomenclatura utilizada para cada hebra- β y la ubicación del LBG.

Para proveer una primera comparación de todos los DRCs se realizó un alineamiento de secuencia múltiple (Figura III.6) con el programa CLUSTALW. En promedio, encontramos un 29% de identidad y un 47% de similitud (sustituciones conservativas) de secuencia. En este sentido, las galectinas humanas son otro ejemplo más de una familia de proteínas con relativamente baja identidad de secuencia pero una elevada similitud estructural [31, 32].

		20	40	60	
Gal-1	>....ACGLVASNLNKPGECLRVRG..EVAP.DAKSFVLNLGKDSNN....LCLHFNPRFNAHGDNANTIVONSKDG				
Gal-2	>.....ELEVKNMDMKPGSTLKITG..SIAD.GTDGFEVINLGQGTDK....LNLHFNPRFSES...TIVONSLDG				
Gal-3	>..LIVPYNLPLPGGVVPRMLITILG..TVKP.NANRIALDFQRGND....VAFHFNPRFENNRR..VIVONTKLD				
Gal-4C	>..PVVPYGRLOQGLTARRTITIKG..YVPP.TGKSEAINFKVGSSG...DIALHINPRMGNG...TVVRNSLLN				
Gal-4N	>YNPTLFPYQPIPGGLNVGMSVYIQC..VASE.HMKREFVNFVVGQDPG..SDVAFHFNPREDDGWD...KVVFNILQG				
Gal-7	>..SNVPHKSSLPEGIRPQTVLRIQG..LVPP.NASRFHVNLCCGEEQG...SDAALHFNPRLDTS...EVVFNSEKQ				
Gal-8C	>.....PFAARLNTPMGPGRTVVVKG..EVNA.NAKSFNVDLLAGSKD...TALHLNPRNLK...AFVRNSFLQ				
Gal-8N	>..PVIPIYGTIPDQLDPGLTIVICG..HVPS.DADREFQVDLQNGSSVKPRADVAHFHFNPRFKRAG...CHVONTLIN				
Gal-9C	>...MPFITTLGGGLYPSKSLILSG..TVLP.SAQRFHINLCSGNH....IAFHLNPREDEN...AVVRNTQID				
Gal-9N	>..PAVPFSGTILQGGQLDGLQITVNG..TVLSSSGTRFAVNFQTGFSGN...DIAFHNPRFEDGG...YVVRNTQON				
Gal-10	>..SLLPVPYTEAASLSTGSTVTKGRPLVCFNLPEYLQVDFHTEMKEE..SDIVFHFQVCFGR...VVMNSREY				
Gal-12N	>.....PYVTITFGGLHAGKMMVLQG..VPL.DAHRFCQVDFQCGCSLCPRPDIAHFHFNPRFHTTKP..HVIONILHG				
PPI3	>..SSLPVYKLPVSLVSGSVIITKCTPIHSFINDPQLQVDFYTDMDDED..SDIAFRFRVHFEGNH....VVMNRREF				
PPL13	>..SSLPVYPTLPVSLPVGSCVITGTPIITFVKDPQLEVNFYTGMDDED..SDIAFQFRLHFHGH...AIMNSCVF				
hGRPC	>.....PFCGHIKGMRPGKKVLMVG..IVDL.NPESFAISLTCGDSSEDPADVAIELKAVETDR...QLLRNGCIS				

	80	100	120	140	
Gal-1	>GAWGTEQREAV..FPFQPGSVAEVCTITFDQANLTKLPDGYEFKFPNRLN.LEALINMAADGDFKIKCVAFD....				
Gal-2	>SNWGCEQREDH..LCFSPGSEVKFTVTFESDKFKVKLPDGHELTSNRLG.HSHSYLSYRGGFNMSSFKLKE....				
Gal-3	>NNWGREERQSV..FPFESGKPKFKIQLVPEPDHFKVAVNDHALLQVNHRYKKLNEHSLKLGISGIDILTSASYTMI...				
Gal-4C	>GSWGSEEEKITH.NPPFGPQPFDLISRCGLDRFKVYANGQHLFDFAHRLSAFQRVDTLEITCGDVTLSYVQISGPSSG				
Gal-4N	>GKWGSEERKRS..MPFKGAAPFELVFIvlaehykvVNGNPFYEGHRLP.LQMVTHLQVDDGLQLQSINFIGGQP...				
Gal-7	>GSWGSEERGP..VPFQGPQPFELVLIASDDGFKAVVGDAQYHHGRHRLP.LARVRLVEVGGDVQLDSVRIF....				
Gal-8C	>ESWGSEERNIT.SPPFSPGMFYFEMIIYCDVREFKVAVNGVHSLEYKHFELSSIDTLEINGDIHLEVRWS....				
Gal-8N	>EKWGSEETIDY..TPPKREKSFETVIMVLKDKFQVAVNGKHTLLGHRTG.PEKIDTLGITGVKNVHSIGFSGPSSG				
Gal-9C	>NSWGSEERSLPRKMPFVRGQSPSVWILCEAHCLKVAVDQHLFEYHRLNRLPTNRLVGGDIQLT.....				
Gal-9N	>GSWGSEERKTH..MPFQKMPFDLCFLVQSSDFKVMVNGILFVQYFHRVP.FHRVDTISVNGSVCLSYISFQ....				
Gal-10	>GAWKQVESKN..MPFQDQEFELSI SVLPDKYQVMVNGQSSYTFDHRIP.PEAVKMQVWRDLSITKPNVSYLKR...				
Gal-12N	>GRWQREARWPH..LALRRGSSFLILFLFGNEEVKSVNGQXFLHFRYRLP.LSHVDTLGIQFGLVEAVGF....				
PPI3	>GIWMLEETDY..VPFEDGKQFELCTVYHNEYEIKVNGIRIYGFVHRIP.PSFVKMQVSRDLSIT.....				
PPL13	>GINRYEEKCY..LPFEDGKPFELCTVYRHKYKVMVNGQRIYNAHRFP.PASVKMLQVFRDISITRVLIS....				
hGRPC	>GERSEECASIP.YPFPFIPDQPFVEILCEYPRFRVFDGHQLFDFYHRIQTLSAIDTIKINGDIQITKLG.....				

Figura III.6. Alineamiento múltiple de secuencias para todas las galectinas estudiadas. Gal-4C, -8N, -9N y 12-N fueron truncadas por ocho residuos al comienzo de la secuencia para inducir un alineamiento correcto en la subrutina CLUSTALW implementada en VMD. Residuos estrictamente conservados (fondo azul) y residuos homólogos (85% de conservación en fondo cian y 70% en fondo verde) se encuentran resaltados. La figura se realizó utilizando GeneDoc (<http://www.nrbcs.org/gfx/genedoc/>).

De los estudios cristalográficos previos se ha sugerido que un grupo de aminoácidos altamente conservados son los responsables de la capacidad de selección y unión a glicanos. Para Gal-1 ellos son His44, Asn46, Arg48, Val59, Asn61, Trp68, Glu71 y Arg73. En el alineamiento mostrado en la **Figura III.6** resulta evidente que sólo cuatro residuos están altamente conservados **en toda la familia** (G25, N73, W80, E83 y R125, numerados según posiciones del alineamiento), mientras que el resto varía considerablemente. De manera interesante, estudios de mutagénesis en Gal-1 han demostrado que el residuo de triptófano conservado en el LBG no es esencial y que puede ser sustituido por Tyr o Phe [33, 34]. Como es evidente, un anillo aromático en dicha posición es fundamental para asegurar la unión a ligando, garantizando el correcto apilamiento entre el glicanos y las cadenas laterales de los aminoácidos. Sin embargo, la naturaleza ha seleccionado en todos los casos de galectinas que éste sea un triptófano.

Con respecto a la comparación estructural de todos los DRCs, en primer lugar se realizó un alineamiento estructural basado en RMSD para todas las galectinas. Para ello, comenzando por las estructuras experimentalmente descritas o por los modelos por homología creados, se llevaron a cabo 50 ns de DM y se recolectaron las estructuras promedio de dicho muestreo para cada DRC. El árbol basado en el RMSD entre dichas estructuras promedio se muestra en la **Figura III.7A**. En la **Figura III.7B** se observa un árbol similar pero considerando solamente las estructuras experimentales.

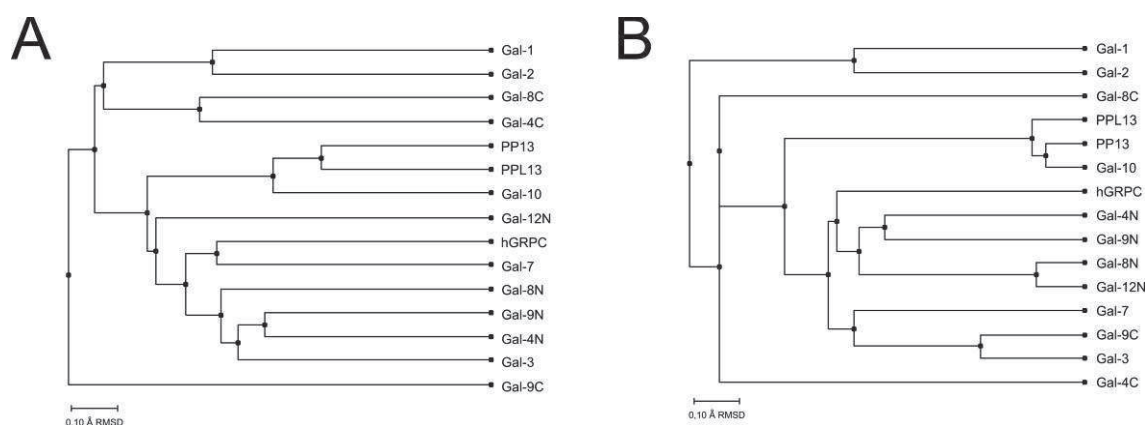


Figura III.7. Alineamiento múltiple estructural, usando A) las estructuras promedio de las simulaciones de DM de cada galectina y B) las estructuras experimentales y modelos por homología construidos. Representaciones de árbol basadas en RMSD provenientes de alineamientos *in silico* de las estructuras secundarias en cuestión.

La primera conclusión importante es que todos los DRC obtenidos experimentalmente son muy similares: el máximo valor de RMSD es de 2,16 Å, entre Gal-1 y Gal-10. Estas pequeñas diferencias son aún mantenidas luego de la relajación inducida por la DM (RMSD máximo = 2,18 Å). Durante la escala de tiempo explorada en nuestras simulaciones, todas las proteínas se mantuvieron muy cerca de la posición original de sus coordenadas a pesar de la relajación, termalización y solvatación de la estructura (RMSD máximo = 1,87 Å). De manera interesante, para aquellas proteínas modeladas por homología uno espera que la estructura final tenga mucha relación con el templado utilizado. Sin embargo, durante las simulaciones de DM, cualquier tipo de nexo estructural es perdido ya que los modelos no muestran necesariamente alta similitud estructural (basado en RMSD) con respecto a sus templados (ni a las otras estructuras en particular). De esta manera, estos resultados verifican que el proceso de relajación y muestreo de DM fue satisfactorio y los modelos por homología finales no se vieron afectados por la elección particular de un templado.

La comparación entre el árbol basado en estructura (**Figura III.7**) y el basado en el alineamiento filogenético (**Figura III.2**) muestra algunas tendencias cualitativas similares pero también diferencias notables. Por ejemplo, el grupo de las galectinas F3-DRC Gal-1, -2, -8C y -4C se encuentra en una rama diferente de la observada para el resto de las proteínas, principalmente del tipo F4-DRC, pero algunos miembros que esperamos que se agruparan de igual manera están claramente ubicados en otras ramas del árbol estructural (Gal-3, hGRPC). También es interesante que Gal-9C diverge del comportamiento esperado si sólo analizamos su filogenia y se ubica en una rama particular del árbol estructural, separado del resto de la familia. Por lo tanto, la primera conclusión del presente análisis es que la estructura global de todos los DRCs es muy similar (RMSD máx. de sólo 2,18 Å, un valor que es del mismo orden de magnitud que algunas de las resoluciones cristalográficas) pero las similitudes estructurales no parecen seguir ni la historia filogenética ni la clasificación original de la arquitectura del dominio.

Aunque los datos hasta aquí analizados demuestran que las estructuras de los DRCs de las galectinas se encuentra altamente conservadas desde un punto de vista global, algunas diferencias importantes existen entre cada DRC, las cuales están localizadas en regiones específicas. La **Figura III.8** muestra un alineamiento estructural de todas las galectinas, coloreadas según RMSD (usando Gal-1 como plegamiento de referencia). La figura muestra claramente que la mayor parte de la estructura terciaria

se superpone casi perfectamente, pero que existen diferencias significativas en las regiones de los loops que conectan las hebras de cada DRC.



Figura III.8. Comparación estructural de todas las galectinas. Las regiones coloreadas de azul corresponden a los valores más bajos de RMSD (usando la estructura de Gal-1 como referencia), mientras que las regiones rojas a los valores más altos. Los loops más relevantes están expresamente resaltados.

Cuando comparamos estos loops, los que poseen más variabilidad son los loops F5S2, F2S3, S4S5 y S3S4, ordenados de menor a mayor RMSD. La diferencia en el loop F5S2 es debida a que Gal-1 y -2 son las únicas galectinas que se desvían significativamente del resto de las conformaciones de los DRCs. Este corrimiento de estas dos galectinas en particular (42% identidad, 59% similitud de secuencia) es posiblemente debido al tipo de aminoácidos involucrados en la formación de este loop que no permite la formación de una hélice- α muy corta, como la observada en la mayoría de las otras galectinas (**Tabla III.3**). Para el loop F2S3, su abanico de estructuras puede dividirse en tres grupos diferentes: Gal-9N por un lado, luego Gal-10, PP13 y PPL13 por el otro y el resto de las galectinas entre ambas conformaciones. El grupo formado por las tres galectinas del tipo F3-DRC es consecuencia de que comparten la misma pequeña hélice- α en esta porción particular de la secuencia. Para Gal-9N, se presenta un loop de características similares, aunque la presencia de un aminoácido más fuerza la secuencia a adoptar una conformación totalmente diferente. Para el loop S4S5 encontramos grandes diferencias, entre DRCs que presentan este loop apuntando al LBG y otros que apuntan en una dirección totalmente opuesta, hacia el solvente. Finalmente, para el loop S3S4 la variabilidad es sustancial. Cinco grupos diferentes pueden distinguirse: (i) proteínas que presentan el loop S3S4 más pequeño de todos: Gal-1 y -2, plegados en la misma forma; Gal-3 y -9C, con pequeñas diferencias; (ii) Gal-10 y PP13, y una pequeña diferencia para PPL13; (iii) en Gal-4N y -

8C este loop apunta al solvente; (iv) el loop S3S4 de seis aminoácidos en Gal-4C, -7 y -9N es prácticamente hidrofílico y ofrece otro plegamiento particular; y (v) el loop S3S4 más largo se observa para Gal-8N, -12N y hGRPC, cuyas conformaciones también se diferencian de las observadas para los otros DRCs.

Galectina	Residuos	Triptófanos	Cisteínas	Giros- β^a	Puentes- β^a	Hélices $3_{10}^{a,b}$
Gal-1	134	1	6	11	3	0
Gal-2	129	1	2	10 (11)	1 (3)	0 (1)
Gal-3	137	1	1	12 (10)	1	1 (0)
Gal-4C	140	1	1	11 (9)	1 (2)	1 (0)
Gal-4N	142	1	0	13 (10)	1	1
Gal-7	135	1	1	11	1	1
Gal-8N	133	2	1	10 (8)	2	0 (2)
Gal-8C	139	1	3	10 (8)	1	1
Gal-9N	127	2	3	12 (11)	2	1
Gal-9C	136	1	2	11 (10)	1	0
Gal-10	141	2	2	10	1	2
Gal-12N	145	2	4	10	1	1
PP13	132	1	2	8	0	2
PPL13	137	1	4	8 (9)	0	2
hGRPC	134	0	4	12 (10)	1	1

Tabla III.3. Datos estructurales relevantes, obtenidos del análisis de los experimentos de rayos X o RMN y los modelos por homología. ^aEn paréntesis, se muestran los cambios en estructura secundaria luego de haber realizado la DM sobre dicha estructura. ^bPor hélice 3_{10} se entiende a aquella estructura secundaria en forma helicoidal y sentido horario en donde el grupo N-H de un aminoácido forma un enlace de hidrógeno con el grupo C=O del aminoácido tres residuos anterior ($i + 3 \rightarrow i$), a diferencia de la típica hélice- α ($i + 4 \rightarrow i$).

Estos datos estructurales muestran que el plegamiento del DRC en galectinas humanas se encuentra altamente conservado. Las principales diferencias están localizadas en loops específicos, cuyas composiciones aminoacídicas y conformaciones promedio adoptadas durante las DMs permiten su agrupación en diferentes DRCs según comportamientos similares. Sin embargo, como se mencionó previamente, no hay correlación de estos resultados con lo que se espera según la clasificación filogenética.

III.3.2 Comportamiento dinámico de los diferentes DRC

Con el fin de acercar una primera descripción a la dinámica del DRC, se computó el RMSF de cada miembro. Como se indica en la sección II.1.5.2, los valores de RMSF muestran la fluctuación en el tiempo de la posición de un átomo (el carbono α , en este caso) a lo largo de toda la simulación y se grafica contra el número de residuo. Estos valores muestran la movilidad/flexibilidad de cada segmento en la estructura de la proteína. Un primer acercamiento a los resultados arroja, como era de esperar, que las regiones más flexibles de los DRCs son los loops. Sin embargo, la

magnitud de esta fluctuación difiere significativamente entre los loops de una dada galectina, resultando en diferentes patrones para cada DRC (**Figura III.9**).

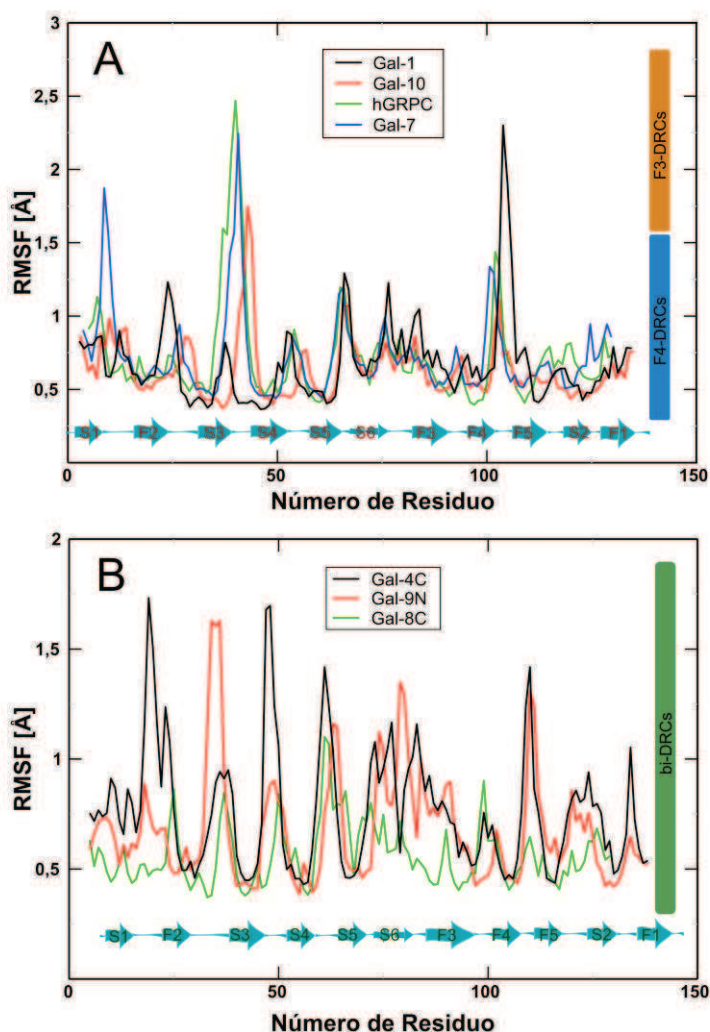


Figura III.9. Gráficos de RMSF versus residuo, para galectinas seleccionadas. Galectinas tipo A) mono-DRC y B) bi-DRC. Se agrega en cada gráfica la localización de cada hebra- β y loop para más claridad.

Se observa una fluctuación principal en todas las galectinas debida a la flexibilidad del loop S3S4 (residuos *ca.* 40-50) excepto para Gal-1, donde la principal fluctuación está corrida al loop F4F5 (*ca.* residuo 100). Fuera de estos casos, el patrón de fluctuaciones de los loops parece ser distintivo de cada DRC. Todos los DRCs de galectinas del tipo bi-DRC muestran una flexibilidad similar distribuida entre todos los loops que cae por debajo de los 2,0 Å, mientras que para galectinas del tipo F3- o F4-DRC se tiene entre una y tres movilidades principales (de hasta 2,5 Å), y además otros segmentos más rígidos.

Siguiendo con el análisis dinámico de estos dominios, se calcularon los modos esenciales (ME) según el análisis de dinámica esencial (sección II.1.5.2) para todos los casos. Las **Tablas III.4** y **III.5** muestran las contribuciones a los tres primeros MEs al movimiento total de la proteína.

Galectina	Gal-1	Gal-2	Gal-3	hGRPC	Gal-7	Gal-10	PP13	PPL13
Tipo de DRC	F3				F4			
ME1	33	24	31	29	27	18	18	19
ME2	12	16	15	13	11	12	9	16
ME3	8	7	10	8	8	10	7	11
#ME > 50%	3	4	4	5	7	7	6	4

Tabla III.4. Contribuciones (%) de cada modos normal al movimiento global en las galectinas del tipo mono-DRC.

Galectina	Gal-4C	Gal-8C	Gal-9C	Gal-4N	Gal-8N	Gal-9N	Gal-12N
Tipo de DRC	F3			F4			
ME1	25	17	24	26	25	19	37
ME2	11	11	17	10	11	14	13
ME3	9	10	13	7	8	9	9
#ME > 50%	6	8	6	4	6	6	2

Tabla III.5. Contribuciones (%) de cada modos normal al movimiento global en las galectinas del tipo bi-DRC.

Como puede observarse de las tablas, resulta muy difícil obtener alguna tendencia general. Tal vez el resultado más interesante es que para todas las galectinas F3-DRC se requieren menos MEs para describir la dinámica de la proteína, ya que más de la mitad de las fluctuaciones estructurales se explican con menos de cinco MEs, mientras que entre seis y ocho MEs son necesarios para las galectinas mono-F4 y las bi-DRCs. Los patrones de distribución son similares siendo el ME1 aquel que aporta entre 20 y 40% del movimiento total, el ME2 un 15% y el ME3 entre 5 y 10%. El primer ME muestra generalmente movimientos de uno o dos loops, mientras que modos de baja amplitud involucran muchos más loops. Sin embargo, qué loop y cómo se mueve es una característica de cada DRC, apuntando a un patrón de comportamiento dinámico completamente diferente para cada galectina. Por ejemplo, en Gal-1 el ME1 y ME2 están localizados en el loop F4F5. Las otras galectinas mono-F3 (Gal-2, -3 y hGRPC) presentan movimientos principales en el loop S3S4, acompañados por un desplazamiento del loop S5S6, imitando un movimiento de apertura/cierre del LBG. Para las galectinas del tipo mono-F4 (Gal-7, -10 y PPL13) los principales desplazamientos se observaron en el loop S3S4. En Gal-7 además acompaña el loop S1F2 y en PPL13 los loops F2S3 y F3S6. En Gal-13 todos sus loops participan de los principales MEs. Para las galectinas del tipo bi-DRC no se observa ningún patrón claro. Gal-4C posee su movimiento principal focalizado en los loops S1F2 y S3S4, mientras que el mismo ME en Gal-4N se encuentra en los loops F4F5, S4S5 y F5S2. La

flexibilidad del loop S3S4 se hace presente en Gal-9N, -12N y particularmente en Gal-8N, en donde éste loop es notablemente más largo. Los otros MEs son, sin embargo, diferentes para cada sistema. Finalmente, Gal-8C y -9C son realmente rígidas y los principales movimientos están localizados en los loops S5S6 y F5S2, respectivamente. En conclusión, cada DRC parece moverse de manera diferente y el único loop protagonista en la mayoría de estos movimientos es el S3S4.

III.3.3 Comparación detallada del LBG de todos los DRC

Para comprender el detalle particular de cada surco de unión, lo que primero hicimos fue comparar el LBG en las proteínas libres de ligando y luego analizar el efecto de la unión sobre la estructura y dinámica de aquellos complejos galectina:ligando cuya estructura era conocida experimentalmente.

Descripción global del LBG. Como se mencionó anteriormente, la cara S del DRC está determinada por las hebras S1-S6, formando el LBG, un surco o bolsillo donde los glicanos se unen. En la mayoría de los complejos galectina:ligando determinados hasta la fecha, el glicano (usualmente un disacárido) se une a uno de los extremos del LBG, formado por las hebras S4-S6, justamente donde los aminoácidos conservados están localizados. Sin embargo, algunos DRCs son conocidos por unir oligosacáridos y por lo tanto todo el LBG debe estar involucrado en el proceso de reconocimiento del ligando [6]. Para comparar en detalle cada DRC LBG, se dividió el LBG en ocho MBS diferentes, denominados Y, Z, A, B, C, D, E y F (**Figura III.3**). Se eligió una clasificación ordenada de los datos, basados en la topología estructural del LBG, para poder resaltar la relevancia de la posición de un aminoácido en particular en una hebra particular. Los resultados se muestran en la **Figura III.10**.

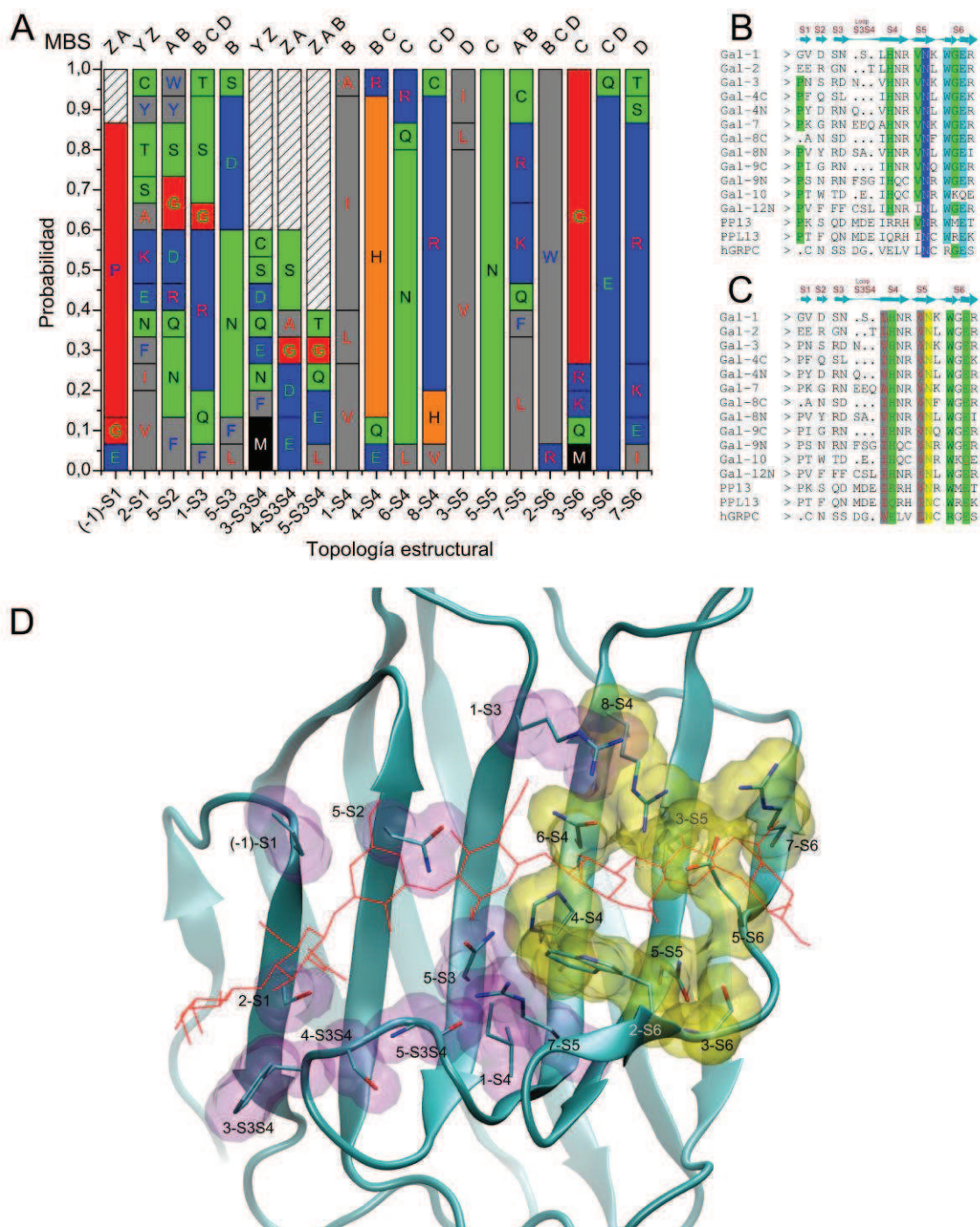


Figura III.10. Composición de aminoácidos de los LBGs basada en una organización topológica de todas las galectinas. A) Gráfico de barras que resume los resultados del análisis topológico de todas las galectinas humanas. Cada residuo está coloreado según su característica fisicoquímica, usando el código de una sola letra. La topología estructural corresponde a la porción de estructura secundaria (sitio ocupado por un aminoácido) bajo estudio, precedida por un número que indica la posición relativa de dicha porción con respecto a la estructura secundaria de la que forma parte (por ejemplo, 3-S3S4 indica la posición 3 del loop S3S4, contando desde el comienzo del loop S3S4, y yendo en la dirección natural de traducción). El índice de probabilidad fue calculado contando el número de veces que un residuo en particular aparece en una posición topológica en particular, dividido el total de DRCs estudiados. B) Resultados del alineamiento estructural de todas las galectinas con el templado del complejo tetra-LacNAc unido a Gal-9N, como se explicó en la sección III.2. La topología del LBG se resalta en la parte superior del gráfico para (continúa...)

(...) mayor claridad. Se utiliza el mismo código de colores que el de la **Figura III.6**. C) Alineamiento estructural similar al realizado en B), en donde se resaltan aquellos aminoácidos con idénticas propiedades fisicoquímicas. D) Representación 3D de la construcción *in silico* del complejo Gal-9N:tetra-LacNAc en donde se han resaltado los aminoácidos más importantes. Basados en el análisis global en A), se han coloreado con sombra amarilla los aminoácidos más conservados y con sombra violeta los más variables para galectinas seleccionadas.

Los resultados mostrados en la **Figura III.10A** claramente revelan la existencia de posiciones topológicas altamente conservadas, mientras que otras parecen permitir la presencia de cualquier tipo de residuo. A medida que nos dirigimos desde la hebra S1 hacia la S6, los aminoácidos y su posición topológica se hacen cada vez más similares entre cada galectina, marcando la importancia de esa región para la selectividad basal de β -galactósidos por parte de todos los miembros de la familia. Es más, incluso cuando sólo se encontraron dos o tres residuos diferentes por posición topológica, o en casos donde un residuo conservado es reemplazado por otro, no se observó correlación de este fenómeno con la definición filogenética o según la arquitectura de la galectina en cuestión. Es importante resaltar también que los MBS Y, Z y A son menos conservados que los B, C y D (**Figura III.10B y C**). Además, el MBS E se encuentra casi afuera del DRC y no pudo ser definido correctamente con este conjunto de estructuras analizadas. Este sitio suele ser relevante cuando se estudia la interacción de las galectinas con ligandos que se encuentran en el extremo reductor del azúcar (en nuestro caso, sólo se encuentra el -OH anomérico, son todos azúcares reductores), como otros monosacáridos en glicanos complejos, lípidos o aminoácidos en proteínas glicosiladas. Es más, el tipo de unión del disacárido (1-3 ó 1-4) en el sitio C-D determina la dirección del enlace glicosídico y puede direccionar parte del ligando hacia la dirección del sitio E y permitiendo que este sitio ahora sea relevante [27]. Las posiciones más conservadas son 5-S5 (Gln) y las hidrofóbicas 1-S4 y 3-S5 (con Ile, Val o Leu). También están muy conservadas las posiciones 2-S6 (con Trp, salvo para hGRPC), 5-S6 (con Glu, excepto para Gal-10 que lleva una Gln) y -1-S1 (ocupada por Pro, excepto en Gal-1 y -2). Una situación similar se observa para la His en la posición 4-S4, Gln en posición 6-S4 y Arg en posición 8-S4, salvo para hGRPC, PP13 y PPL13, donde otro residuo diferente ocupa dicha ubicación. hGRPC también difiere en la Arg conservada en posición 7-S6 junto con Gal-8N, -10 y PP13. La Gly en posición 3-S6 es reemplazada por residuos más grandes en Gal-12N, -10, PP13 y PPL13. Otras posiciones, como 1-S3, 5-S3 y 7-S5, muestran también cierto grado de conservación.

En particular, cada DRC posee al menos dos, uno o ningún cambio en las posiciones conservadas, excepto para hGRPC, Gal-10, PP13 y PPL13, las cuales muestran más desvíos respecto del consenso. Como este análisis está enfocado en los residuos del LBG, es de esperar que estas diferencias puedan explicar porqué Gal-10 no posee afinidad por β -galactósidos pero al parecer puede unir manosa en una forma particular [35]. Un comportamiento similar pero inesperado se encuentra para PP13 y PPL13, cuyos LBGs son también bien diferentes. Finalmente, la ausencia de residuos claves en el LBG de la proteína tipo galectina hGRPC, como el Trp en la posición 2-S6, también podría explicar porqué esta proteína no muestra afinidad alguna por muchos de los glicanos estudiados [8, 9]. Por lo tanto, de esta parte del estudio, hemos puesto de manifiesto la presencia de algunos residuos altamente conservados en posiciones topológicas determinadas, con diferencias particulares entre galectinas, sin encontrar una relación clara con la organización filogenética o estructural de las mismas. Gal-10, PP13, PPL13 y especialmente hGRPC son los DRCs con mayor divergencia de estos residuos conservados, lo que explicaría justamente el porqué de las singularidades encontradas en bibliografía de estas galectinas frente a la unión a diferentes glicanos.

III.3.4 Efecto de la unión a ligando en el DRC

Estructura y dinámica

Como se mencionó en la introducción del capítulo, muchas de las estructuras de los distintos DRCs se han obtenido en presencia de un carbohidrato. Los detalles de las interacciones que ocurren en ese proceso de unión han sido extensamente analizados en cada caso (por ejemplo [1, 4, 36]) y no es el objetivo de esta Tesis una comparación individual sino la búsqueda de similitudes y diferencias globales. Para ello, se realizaron simulaciones de DM de los complejos DRC:glicano disponibles a la fecha (PDB ID de la estructura entre paréntesis): Gal-1:lactosa (1GZW), Gal-2:lactosa (1HLC), Gal-3:LacNAc (1AK3), Gal-7:LacNAc (5GAL), Gal-8N:lactosa (2YXS) y Gal-9N:tetra-LacNAc (construcción particular de esta Tesis, a partir de las estructuras 2ZHN y 2ZHM, ver sección III.2). El complejo Gal-10:manosa (1QKQ) se ha procesado como el resto de los sistemas proteína:ligando, pero dado que a los pocos nanosegundos de simulación de DM el ligando abandona el LBG (debido tal vez a interacciones muy débiles), no se pudo analizar por falta de suficiente acumulación de datos del ligando en el sitio de unión y lograr así un resultado estadísticamente relevante.

Se analizó el efecto de la unión a ligando en la dinámica del DRC a través de la comparación de las RMSF de las estructuras con y sin ligando (**Figura III.11**). Una apreciación general de los resultados indica que la unión al ligando disminuye la movilidad de loops específicos de la estructura pero, curiosamente, no siempre aquellos cercanos al LBG. Para algunos casos particulares, aunque en menor proporción, la unión del ligando a la proteína provocaba el aumento de la movilidad de ciertos loops, efecto estudiado experimentalmente para Gal-1 [37, 38].

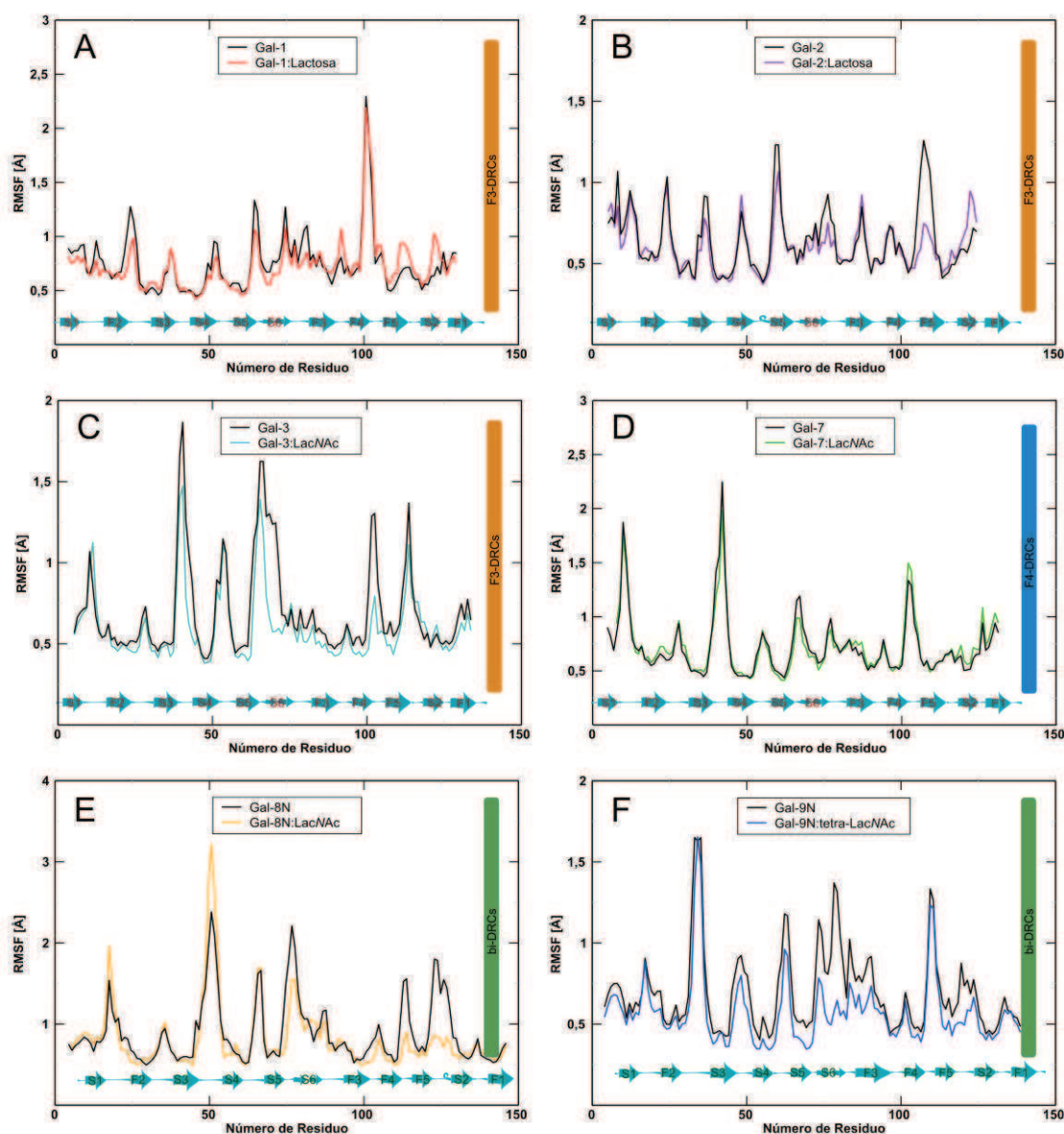


Figura III.11. Comparación de las fluctuaciones estructurales asociadas al proceso de unión. A) RMSF en función del número de residuo de Gal-1 y el complejo Gal-1:lactosa. B) RMSF para Gal-2 y Gal-2:lactosa. C) RMSF para Gal-3 y Gal-3:LacNAc. D) RMSF para Gal-7 y el complejo Gal-7:LacNAc. E) RMSF de Gal-8N y Gal-8N:LacNAc. F) RMSF para Gal-9N y la construcción *in silico* Gal-9N:poli-LacNAc.

Para ilustrar y desarrollar en profundidad estas tendencias, los resultados particulares para Gal-1 (**Figura III.11A**) demuestran que no ocurren cambios particulares en los loops F4F5, S6F3 y S3S4, siendo los últimos dos los más cercanos al LBG. Los loops F2S3 y S6F3, en contacto uno con el otro, reducen su movilidad al unir lactosa. También, el loop corto S5S6 y parte de la hebra S6 sufren una disminución en la movilidad debido a que se encuentran directamente involucrados en la interacción con el ligando. De forma interesante, regiones de contacto en la cara F del DRC, como ser el loop F3F4 y el comienzo del loop F5S2 así como también el comienzo del loop S2F1, aumentan significativamente su movilidad al unir el ligando. Estos resultados son consistentes con aquellos encontrados por un trabajo reciente donde se estudia la unión de lactosa a Gal-1 por métodos espectroscópicos [37].

Para Gal-2 la unión a lactosa también reduce la movilidad de la proteína, pero ahora está principalmente focalizado en el loop F5S2 (**Figura III.11B**). Para Gal-3, la unión de lactosa o LacNAc reduce significativamente la movilidad de regiones específicas de la proteína (**Figura III.11C**), principalmente sobre la hebra S6 y el loop F4F5. El principal contacto responsable de esta disminución de las fluctuaciones en la hebra S6 es el residuo Glu184 (indexado según 1AK3.pdb), mientras que el efecto en el loop F4F5 es sutil y no puede ser adjudicado a una interacción específica. Para Gal-8N (**Figura III.11E**), la disminución de la movilidad se observa en el pequeño loop F5F6 y en el gran loop F6S2. La reducción de la movilidad en el loop más corto es debida a un aumento en la interacción entre los residuos Asn126-Glu111 (índices del 2YV8.pdb) que no se encuentra definida en la proteína libre. Para el loop F6S2, se observa directamente un cambio de estructura del loop, el cual adopta una conformación de hélice en el estado unido a ligando. Para Gal-9N (**Figura III.11F**), la hebra S6 y el loop S6F3 son los más afectados por la unión al oligosacárido. Los efectos más importantes son evidentes en áreas cercanas a la Pro84 y la Arg125 (índices de la cadena A del cristal 2HZM.pdb). Ambos efectos pueden ser causados por la unión del ligando, el cual induce nuevas conformaciones de la Asn75 y la Pro84 permitiendo múltiples interacciones con los grupos NH de la Arg125. La cadena lateral de la Arg125 se congela en una conformación más cercana a la cara S que lo evidenciado en Gal-9N libre, donde la Arg125 está mucho más cerca de la cara F e interactuando fuertemente con Glu86 y Phe123:O, aunque con una conformación tal que el backbone de la proteína es mucho más libre para moverse. De forma interesante, la unión a ligando por parte de Gal-7 no reduce significativamente la movilidad de ninguna parte de la

proteína (**Figura III.11D**) y esto ha sido recientemente comprobado experimentalmente [39]. En resumen, a pesar de la tendencia general que muestra una reducción generalizada de la movilidad de los loops de cada DRC luego de la unión a ligando, la región particular que se ve afectada es característica de cada DRC y difícilmente sea fácil de predecir para otros complejos y otras galectinas cuyas estructuras aún no han sido experimentalmente descritas.

A continuación se calcularon los correspondientes valores de entropía conformacional para la proteína libre y unida a ligando, como se describe en la sección II.1.5.2. La entropía conformacional resume en un número la flexibilidad (o la amplitud del espacio conformacional de la proteína) donde valores grandes indican más flexibilidad y/o un espacio conformacional accesible más grande. El análisis llevado a cabo a partir de los 50-60 ns de simulación de DM arroja una buena convergencia de la entropía en la mayoría de las estructuras estudiadas (**Figura III.12**), obteniéndose valores absolutos alrededor de las 2,5 kcal K⁻¹ mol⁻¹. A partir de estos valores en la proteína libre y unida a ligando, calculamos el cambio en entropía conformacional del DRC por unión a ligando (ΔS). Estos valores se resumen en la **Tabla III.6**.

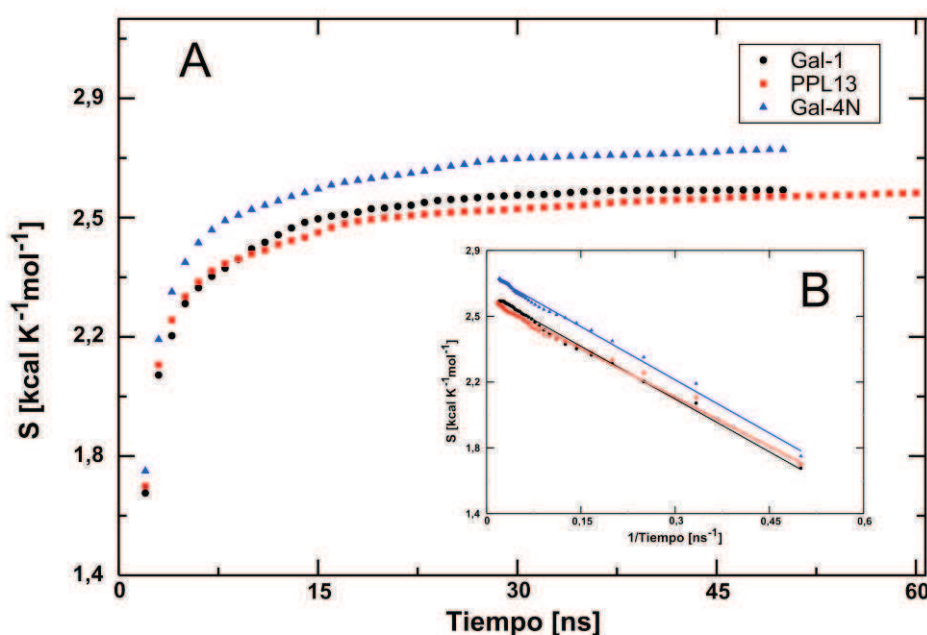


Figura III.12. Resultados de entropía conformacional para algunos DRCs seleccionados: Gal-1, Gal-4N y PPL13. A) Los valores de entropía conformacional alcanzan la convergencia en el tiempo de simulación (50 ó 60 ns, según el caso). B) Análisis alternativo por regresión lineal de los datos transformados.

Galectina	MBS	Ligando	Energías [kcal mol ⁻¹] ^a				OH/	ΔS [cal K ⁻¹ mol ⁻¹]
			ELE ^b	vdW ^c	GB _{solv} ^d	$\Delta G_{unión}$ ^e		
Gal-1	C	β -D-galactopiranosil-(1→	-30,1	-14,7	32,7	-12,1	2	-41,2
	D	4)-N-acetil-D-glucosamina	-23,2	-8,4	20,8	-10,8	2	
	todo	LacNAc	-53,3	-23,2	46,7	-29,7	4	
Gal-2	C	β -D-galactopiranosil-(1→	-29,1	-12,7	32,1	-9,6	2	-8,2
	D	4)- α -D-glucopiranososa	-9,1	-4,4	10,9	-2,7	1	
	todo	α -lactosa	-38,2	-17,1	39,1	-16,2	3	
Gal-3	C	β -D-galactopiranosil-(1→	-4,4	-14,6	8,0	-10,9	2	-96,0
	D	4)-N-acetil-D-glucosamina	-46,5	-6,5	43,7	-9,3	2	
	todo	LacNAc	-50,9	-21,0	45,2	-26,8	4	
Gal-7	C	β -D-galactopiranosil-(1→	-10,5	-13,8	13,3	-11,1	2	-0,1
	D	4)-N-acetil-D-glucosamina	-37,8	-5,9	36,3	-7,4	2	
	todo	LacNAc	-48,8	-19,8	43,8	-24,7	4	
Gal-8N	C	β -D-galactopiranosil-(1→	-6,1	-14,8	9,4	-11,5	1	-30,2
	D	4)- α -D-glucopiranososa	-41,1	-4,7	43,2	-2,4	2	
	todo	α -lactosa	-47,2	-19,5	48,4	-18,2	3	
Gal-9N	Y	β -D-galactopiranosil-(1→	-2,3	-5,0	7,7	0,4	0	-5,4
	Z	4)-N-acetil-D-glucosaminil-(1→	-7,8	-10,8	14,8	-3,7	1	
	A	3)- β -D-galactopiranosil-(1→	1,3	-7,1	6,4	0,6	1	
	B	4)-N-acetil-D-glucosaminil-(1→	-4,3	-10,2	12,5	-1,9	1	
	C	3)- β -D-galactopiranosil-(1→	-19,1	-15,2	20,3	-14,0	2	
	D	4)-N-acetil-D-glucosaminil-(1→	-27,7	-7,6	30,8	-4,5	2	
	E*	3)- β -D-galactopiranosil-(1→	-0,9	-1,0	3,2	1,4	0	
	F*	4)-N-acetil-D-glucosamina	-3,3	-0,2	4,1	0,6	0	
	todo	tetra-LacNAc	-63,9	-57,5	83,9	-37,7	7	
Gal-10	C	manosa	-20,2	-13,9	22,9	-11,2	2	-s

Tabla III.6. Resultados de energía libre y cambio de entropía conformacional de unión para los MBS ocupados en las galectinas estudiadas, usando MM-GBSA y el método de Schlitter respectivamente. ^aPara todos los casos, los desvíos estándar del análisis de todas las estructuras provenientes de la DM de cada complejo fueron menores a 5 kcal mol⁻¹. ^bELE: energía electrostática. ^cvdW: contribución debida a la energía de van der Waals. ^dGB_{solv}: energía libre de solvatación. ^e $\Delta G_{unión}$: cambio de energía libre de unión. /OH: número de grupos hidroxilo del ligando que poseen la capacidad de interactuar con algún residuo de la proteína en el LBG. sLa entropía conformacional del complejo Gal-10:manosa no pudo ser calculado debido a que el ligando no permanece en el LBG el tiempo de simulación suficiente para alcanzar convergencia de la entropía. Se incluye aquí la contribución energética del ligando entero. *Para los MBS E y F, se muestran los resultados como control.

En forma consistente con la tendencia observada en el análisis previo de cambios en RMSF, todos los resultados son menores a cero, es decir, la unión a ligando produce una disminución de la entropía conformacional de la proteína. Los valores van desde cero para Gal-7 hasta casi 100 cal K⁻¹ mol⁻¹, lo que resulta en una contribución de 30 kcal mol⁻¹ a la pérdida de energía libre de unión a 300 K por costo entrópico en la estructura de la proteína. El rango de valores aquí observado sugiere que la entropía conformacional es una contribución importante en la energía libre de unión y, por lo tanto, en la afinidad por el ligando. Luego de publicarse estos resultados teóricos [40], es interesante remarcar que un trabajo experimental demuestra lo aquí predicho al menos para Gal-3 [38].

Finalmente, para proveer un punto de vista comparativo de la termodinámica de unión a carbohidratos de las diferentes galectinas, se realizó un análisis de MM-GBSA para todos los complejos DCR:ligando estudiados en esta Tesis y por otro lado, se introdujo una estrategia/metodología para contemplar todos aquellos otros complejos no caracterizados experimentalmente. El fin de ambos estudios fue la disección de la selectividad de las galectinas por un grupo de β -galactósidos biológicamente relevantes. Desglosaré ambos enfoques en apartados separados, dado el diferente tenor y relevancia de ambos análisis.

Termodinámica de los LBG

Los resultados (**Tabla III.6**) muestran que la mayoría de los disacáridos poseen *grossa modo* una energía libre de unión negativa entre 20-30 kcal mol⁻¹, salvo para Gal-2 y la unión a manosa por parte de Gal-10 que presentan una unión más débil al sitio de unión D. Usualmente la contribución electrostática es más del doble que la de vdW para este tipo de ligandos, pero esta componente energética está compensada por el costo de solvatación. De forma interesante, en cada DRC, cada monosacárido contribuye a la interacción en forma significativamente diferente, poniendo de manifiesto la variabilidad que puede ser el motivo de regulación de la unión particular de cada galectina a cada ligando. Observando de cerca cada MBS, los análisis muestran una variación sustancial con valores yendo desde -30 a 0 kcal mol⁻¹. Sin embargo, algunas tendencias interesantes pueden remarcarse. La contribución de vdW al sitio C es siempre parecida, entre -11 y -15 kcal mol⁻¹, posiblemente debido a la interacción de apilamiento del ligando con el Trp conservado. Por otro lado, las interacciones electrostáticas son muy variables. Para el sitio D, la contribución de vdW es significativamente menor (alrededor de 6 kcal mol⁻¹); otra vez, la contribución electrostática muestra alta variabilidad a lo largo del análisis de todos los DRC. Finalmente, analizando los datos para el complejo de Gal-9N, se puede evidenciar que el MBS que más contribuye a la unión del ligando entero es el MBS C y en menor medida, el D y el Z, mientras que los otros presentan una interacción despreciable o hasta incluso desfavorable. Cabe notar que a pesar de que no pudimos definir en ningún complejo los MBS E y F (pues estos se encuentran casi completamente expuestos al solvente; ver sección III.2), se realizaron los cálculos de MM-GBSA para estas regiones para el complejo Gal-9N:tetra-LacNAc para tener un control negativo de unión del método y una idea de la dispersión de los resultados. Los valores de energía

son muy bajos en comparación con los anteriormente analizados, en donde realmente existe interacción proteína-ligando. De este modo, estos resultados particulares indican, por un lado, que los dos últimos monosacáridos (hacia el extremo reductor) del ligando para Gal-9N están expuestos al solvente y por otro lado, que la energía obtenida para los otros complejos galectina:ligando proveen resultados consistentes con el error aceptado para el método.

Considerando todas estas tendencias en conjunto, los resultados termodinámicos mostraron que además de las interacciones de vdW conservadas en el MBS C (debido a la presencia del Trp), cada DRC muestra diferentes contribuciones y energías libres totales de unión gracias a la topología de cada MBS en particular y, por lo tanto, posiblemente muestren diferente afinidad por estos ligandos *in vivo*. Algunos resultados experimentales de energía libre de unión que se han reportado para estos complejos se detallan a continuación: Gal-1 de bazo bovino: LacNAc, $-5,5 \text{ kcal mol}^{-1}$ (a 298 K); Gal-1 de corazón bovino: LacNAc, $-5,7 \text{ kcal mol}^{-1}$ (a 300 K); Gal-3 recombinante de ratón: LacNAc, $-5,9 \text{ kcal mol}^{-1}$ (a 300 K); Gal-7 humana recombinante: LacNAc, $-4,4 \text{ kcal mol}^{-1}$ (a 300 K); y Gal-8N humana recombinante fusionada con GST:lactosa, $-5,6 \text{ kcal mol}^{-1}$ (a 298 K) [41-43]. Todos los resultados obtenidos mediante MM-GBSA son mayores (en valor absoluto) que los experimentales, consistente con los problemas ya conocidos de la metodología, los cuales tienen una tendencia a sobreestimar la energía libre de unión [44]. Por esa razón, los resultados computacionales deben ser tomados en cuenta solo en forma comparativa (sección II.1.5.2). De acuerdo al análisis de las uniones de hidrógeno, resulta claro que en cada MBS no existen más de dos de estas uniones entre el ligando y la proteína por MBS. Esto es debido al hecho de que cada monosacárido se une a cada galectina exponiendo algunos de sus grupos hidroxilo al interior de la proteína, mientras que otros apuntan al solvente. Finalmente, es notable que mientras los valores de las energías libre de unión son todos bastante similares, el cambio de entropía evidenciado por cada DRC es realmente variable, dando una idea del impacto diferencial que puede llegar a tener en la movilidad de la proteína la unión de uno u otro ligando.

Selectividad por ligandos

Validación del método. Con el fin de determinar efectivamente si la capacidad de unirse selectivamente a diferentes ligandos está codificada en la topología del LBG y que el ajuste fino y diferencial de ocurrencia y posición de aminoácidos es lo que hace a

una galectina diferente de la otra, se muestran a continuación los resultados del *glycoarray in silico* antes descrito (ver sección III.2) (**Figura III.13**).

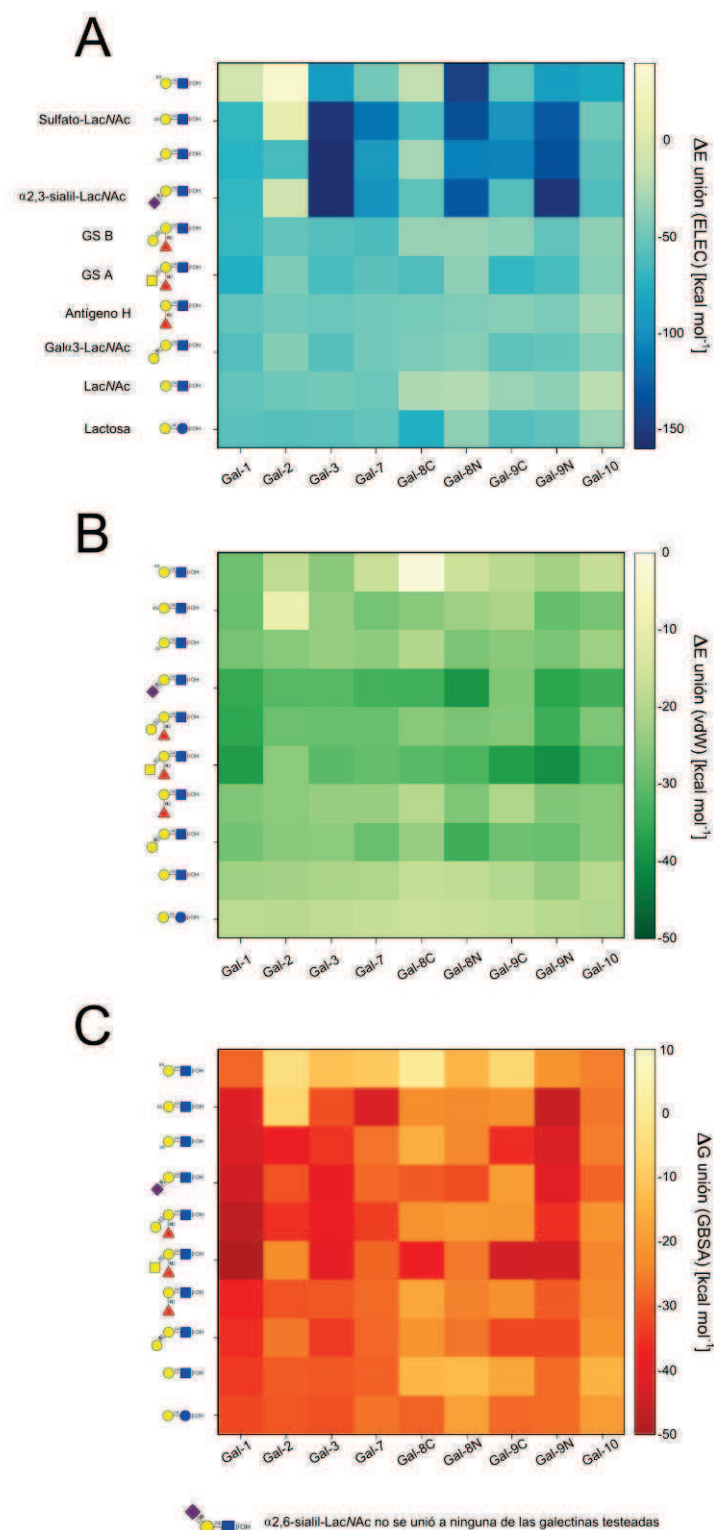


Figura III.13. Resultados del glyco-array *in silico* basado en los cálculos de energía libre de unión con MM-GBSA. Heat-map para los valores de cambio de A) energía electrostática, B) energía de van der Waals y C) energía libre de unión, para todas las combinaciones posibles entre algunas galectinas humanas y los ligandos de la Figura III.4.

En la **Figura III.13**, se pueden apreciar tres gráficos diferentes, en donde se ordena a modo de *array*, con escalas de colores para la energía de interacción Elec (A), vdW (B) y libre de unión (C), el valor para cada galectina con cada ligando ensayado.

Para poder validar la comparación de “todos contra todos”, es importante saber las limitaciones y alcances de este tipo de *screening*. Teniendo en cuenta el problema asociado a la metodología de MM-GBSA para el cálculo de los valores de energía libre de unión (discutidos en el apartado anterior de esta sección y en la sección II.1.3.2), y a partir de los datos de bibliografía disponibles, se procedió a una validación de las tendencias actualmente descritas y estudiadas. Por ejemplo, en el trabajo de Stowell y col [45], se trabajó con un set un poco más variado de ligandos y ensayos experimentales con el fin de comparar la selectividad entre Gal-1, Gal-2 y Gal-3. Nuestros resultados computacionales predicen acertadamente 20/27 tendencias. Efectivamente, nuestro protocolo reproduce totalmente los resultados de afinidad selectiva frente a α 2,3 y α 2,6-sialil-LacNAc y frente a los tres ligandos sulfatados (3S-, 4S- y 6S-LacNAc) para estas tres galectinas. En el trabajo de Brewer [42], se comparan Gal-1 bovina, Gal-3 murina y Gal-7 humana a diferentes temperaturas y con diferentes ligandos. De este estudio, se pueden destacar algunas tendencias que se reflejan también en nuestro array *in silico*, teniendo en cuenta que el estudio no se realiza sobre galectinas de la misma especie y que una sustitución en una posición estratégica del LBG puede cambiar drásticamente la tendencia esperada para las galectinas humanas. Se ha logrado predecir mayor afinidad de Gal-1 que Gal-7 frente a lactosa, similar afinidad por LacNAc para las tres galectinas, que ninguna une α 2,6-sialil-LacNAc, y que Gal-1 y Gal-3 tienen afinidad por α 2,3-sialil-LacNAc similar a la que tienen por LacNAc y que resulta mayor que la afinidad que Gal-7 tiene por este azúcar sialilado. En el trabajo de Ideo y col. [43] se estudia la unión de Gal-8N a algunos derivados sialilados y sulfatados, demostrando una mayor afinidad de Gal-8N (y también de Gal-3) por α 2,3-sialil-LacNAc y por 3S-LacNAc (dato experimental disponible sólo para Gal-3). Nuestro estudio logra reproducir también estas tendencias. En conclusión, utilizando la bibliografía hasta la actualidad disponible fue posible verificar gran parte de los resultados previos sobre la selectividad, indicando que la metodología aquí empleada permitirá una predicción adecuada para las otras galectinas que no han sido aún analizadas.

Tendencias generales. Una vez validada la metodología, es posible comenzar a analizar las tendencias generales de unión de todas las moléculas con todas. Ante todo, es necesario hacer notar la ausencia del ligando α 2,6-sialil-LacNAc en la **Figura III.13**. Dado que en ninguno de los casos fue posible evitar la salida del LBG del ligando durante la termalización, se puede concluir que este ligando no posee afinidad significativa por ninguna de las galectinas aquí estudiadas. Este resultado cobra relevancia, pues es bien sabido que las galectinas son incapaces de unir a derivados sacarídicos que posean el ácido neuramínico en posición 6 del molde de LacNAc, modulando así, la funcionalidad de estas lectinas [46, 47]. Justamente, la presencia de este monosacárido en dicha posición es clave en otras propiedades biológicas de funcionalización con glicanos, incluso no sólo para el ambiente de las galectinas (revisadas en [46]) sino que sorprendentemente, poco han sido exploradas aún en la actualidad más allá del mundo de las galectinas. Estructuralmente, es muy fácil también de poner en evidenciar por qué esta unión no se puede dar en ningún caso. Las galectinas poseen su estructura “tallada” para poder unir el disacárido LacNAc, lo que no solo impone un cierto encastre de este ligando en el LBG, sino que también restringe las posibilidades de los sustituyentes que podrían decorarla. Como muestra la **Figura III.14**, es físicamente imposible permitir que la porción de glicano en posición 6 del disacárido LacNAc interactúe con los MBS sin que se perturbe la unión del molde de LacNAc en los MBS D y C.

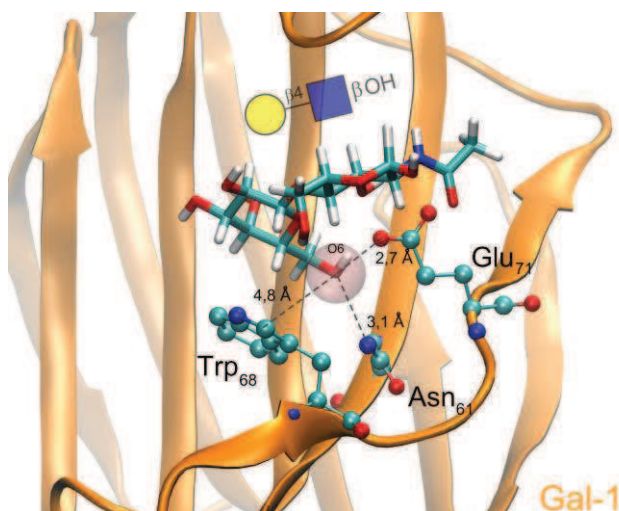


Figura III.14. Espacio disponible para los sustituyentes en posición 6 del molde de LacNAc. Gal-1 y LacNAc en una estructura promedio extraída de las DM ensayadas. El espacio entre el ligando y las cadenas laterales de la proteína es limitado, siendo imposible, por ejemplo, acomodar un sustituyente de ácido neuramínico (como en α 2,6-sialil-LacNAc) sin perturbar completamente la estructura secundaria de la proteína.

Ahora, analizando en forma global la **Figura III.13C** es posible identificar a simple vista ciertas tendencias globales y algunas singularidades. Por ejemplo, es notable la baja afinidad que posee Gal-10 por todos los derivados de LacNAc. Basándonos en la escasa bibliografía que hay sobre la afinidad de Gal-10, es un resultado esperado. Sólo se encuentra una estructura de Gal-10 en complejo con manosa, un monosacárido distinto de todos los azúcares aquí testeados [35]. Por otro lado, si examinamos las afinidades particulares de cada ligando por todas las galectinas, resulta evidente también que 6S-LacNAc es el ligando que interactúa menos fuertemente con los distintos DRC. Analizando esta tendencia con lo observado para el ligando α 2,6-sialil-LacNAc, este resultado aporta aún más evidencia de que la posición 6 en el molde de LacNAc es muy sensible y determinante en la unión (o no) de β -galactósidos por parte de todas las galectinas. En este caso en particular, si bien el sulfato puede ubicarse en el espacio que queda disponible (ver **Figura III.14**), la presencia del residuo conservado Glu71 (*numeración de LGALS1*) hace que esta interacción sea muy desfavorable electrostáticamente. Una excepción se observó en casos como en Gal-8N, en donde en su lugar, debido a un corrimiento de secuencia, esta posición la ocupa la Arg88 (*numeración de LGALS8*) que apunta al solvente la mayoría del tiempo de la DM, lo que deja claramente el espacio para que pueda acomodarse el grupo sulfato en ese entorno. Las singularidades más evidentes son los complejos que presentan mayor afinidad (Gal-1:GSA, Gal-7 y Gal-9N por 4S-LacNAc y Gal-9C y -9N por GSA) y menor afinidad (Gal-8C:6S-LacNAc y todas las galectinas frente al α 2,6-sialil-LacNAc).

Para los ligandos más sencillos de Lactosa y LacNAc se puede observar que todas las galectinas son medianamente afines salvo unas interacciones mas débiles para las galectinas tándem Gal-8N, -8C y -9C (y el caso de Gal-10 antes analizado). Cuando se analiza en particular el cambio de energía libre de unión, se encuentra que para estas galectinas la posición del ligando es muy distinta de aquellas en donde éste valor de energía libre es bien negativo y la unión se podría calificar como “más canónica” o estándar. A pesar de que la estructura secundaria no se ve afectada significativamente, e incluso la cadena lateral de un residuo crítico como es el Trp del MBS D no difiere mucho uno de otro (RMSD_{TRP} relativo a Gal-9N: Gal-8N = 1,21 Å; Gal-9C = 1,08 Å; Gal-8C 1,47 Å), el ligando sufre un posicionamiento desfavorable a la unión en las tres proteínas, siendo más pronunciado para Gal-8N, donde se obtiene el valor de $\Delta G_{\text{unión}}$ menos negativo (RMSD_{LacNAc} relativo a Gal-9N: Gal-8N = 4,45 Å; Gal-

9C = 3,31 Å; Gal-8C = 2,00 Å) (**Figura III.15**). Lo interesante de este análisis es que si miramos aún más en detalle la estructura y la secuencia particular de esta galectina, podemos arribar a alguna explicación. En las cercanías a la subunidad de Glc/GlcNAc, existe en el MBS D una secuencia consenso que determina la afinidad típica de las galectinas por estos azúcares. Si alineamos las estructuras aquí comparadas, veremos claramente que Gal-8N posee la mayoría de los aminoácidos adecuados, excepto una Arg (**Figura III.16**). La presencia de una Ile en lugar de una Arg clave para la interacción con el O2 (en lactosa)/NH del grupo amida (en LacNAc) y el O3 del monosacárido Glc/GlcNAc permitiría justificar puntualmente el porqué de la tendencia. Uno de los trabajos fundamentales en Gal-1 indica que la mutación de esta Arg bloquea completamente la unión a glicanos [33]. Sin embargo, en un trabajo específico de Gal-8, los autores indican que sólo encontraron una leve disminución en la capacidad de unión a Lactosa en presencia de Gal-8 (completa) y las mutantes de esta Arg en el extremo C-terminal y en la Ile del extremo N-terminal [48].

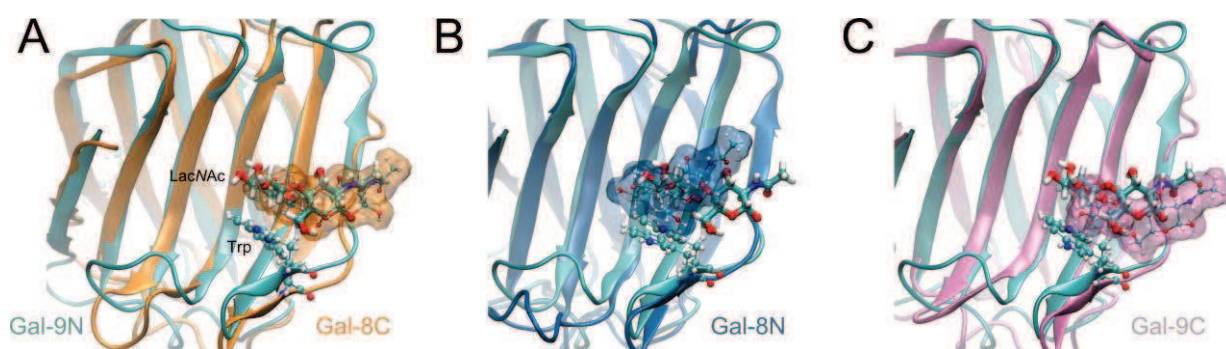


Figura III.15. Comparación de la posición del ligando en las distintas galectinas. En verde se observa la estructura de Gal-9N junto al ligando LacNAc, utilizado como patrón estándar de unión y por consiguiente, comparación con las otras galectinas en tandem. A) Gal-8C en naranja, B) Gal-8N en celeste y C) Gal-9C en rosa. En todas las estructuras, se marca el Trp conservado para referencia espacial dentro del LBG, la molécula de LacNAc del patrón (Gal-9N:LacNAc) en esferas y bastones grandes y la molécula de LacNAc de la galectina en cuestión en esferas y bastones pequeños y encerrada en una nube de superficie de vdW, siguiendo el mismo código de color.

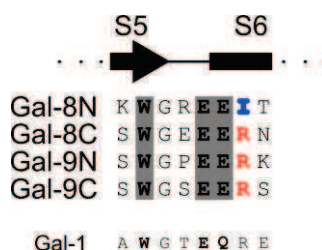


Figura III.16. Alineamiento estructural de las galectinas en tandem aquí estudiadas, mostrando la una de las regiones de interés a la unión a LacNAc y la misma región para Gal-1 a modo de comparación fuera de las galectinas en tandem. Para un alineamiento completo, revisar la **Figura III.6**.

Es notable la poca variabilidad en los valores de afinidad para todas las galectinas frente a los ligandos Gal α 3-LacNAc y el antígeno H, más evidente en el último que en primero. Esto se puede justificar en base a la comparación de estos valores con los observados para Lactosa y LacNAc en donde sucede más o menos lo mismo, y teniendo en cuenta la arquitectura particular los ligandos en cuestión: los nuevos monosacáridos que se agregan en los glicanos Gal α 3-LacNAc y antígeno H, se encuentran unidos en posiciones 3 y 2 a la subunidad de Gal, respectivamente. Mirando de cerca la estructura de estos ligandos en complejo con las galectinas, resulta claro que estos monosacáridos están orientados en dirección al solvente (aquellos unidos en posición 2 mucho más que aquellos en posición 3) y, por lo tanto, no se ve reflejada su presencia por parte de una interacción puntual con el LBG de la proteína (**Figura III.17**). Al calcular la SASA (sección II.1.5.2) por monosacárido en ambos ligandos encontramos el mismo resultado. Tanto para el monosacárido Gal α 3- en el ligando Gal α 3-LacNAc como para el residuo de fucosa en el antígeno H se encontraron valores muy elevados de exposición al solvente en comparación, por ejemplo, al valor para el monosacárido de galactosa del molde de LacNAc, casi completamente enterrado en el LBG (**Figura III.17**). Si uno observa con más atención, la diferencia es notable: para el antígeno H, la presencia de fucosa en posición 2 prácticamente no cambia el $\Delta G_{\text{unión}}$ con respecto a LacNAc (**Figura III.13C**). Sin embargo, aunque puede llegar a ser despreciable, algunas diferencias si comienzan a ser notorias para el ligando Gal α 3-LacNAc, en particular para las galectinas en tándem, donde la presencia de este monosacárido pareciera aumentar un poco más la afinidad por este trisacárido que por LacNAc. Es más, si esto es cierto, debería observarse que la afinidad de las galectinas por el grupo sanguíneo B, una combinación estructural entre el antígeno H y el glicano Gal α 3-LacNAc, debería seguir el mismo patrón de afinidades que el antígeno H y, en líneas generales, es lo que se observa también experimentalmente [45].

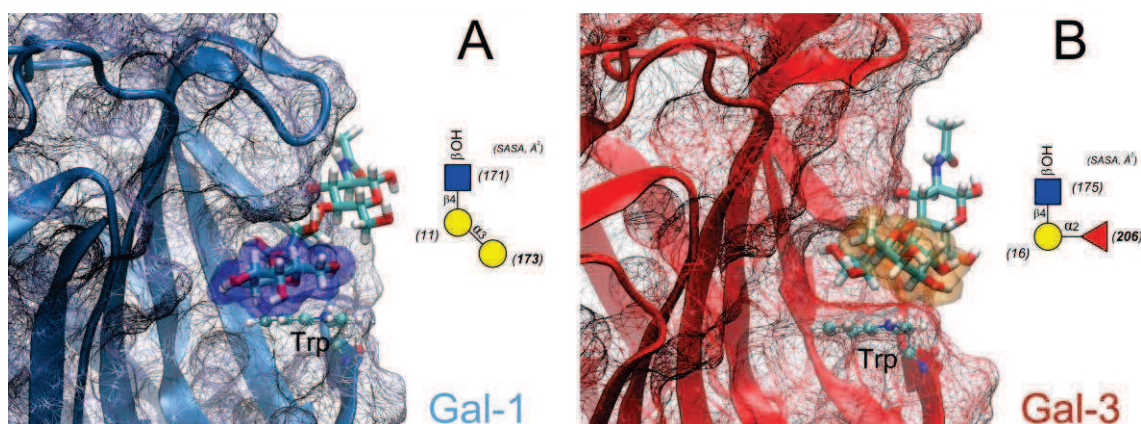


Figura III.17. Los sustituyentes en posición 2 y 3 apuntan al solvente. Se utilizaron las estructuras de A) Gal-1:Gal α 3-LacNAc y B) Gal-3:antígeno H para demostrar la orientación de estos sustituyentes en relación al LBG y el solvente. Encerrado en una nube de vdW se encuentra la estructura de la galectina y también el monosacárido en posición 2 ó 3 a la estructura de LacNAc, según corresponda. Al lado del esquema de cada ligando, se muestra entre paréntesis, la SASA (en Å²) para cada monosacárido en cuestión.

Cuando pasamos al estudio de la comparación de los ligandos cargados (sialilados y sulfatados) el efecto debido a la presencia de estos sustituyentes es evidente en la energía de unión. Resulta muy claro al analizar la **Figura III.13A**, en donde se observa que en la parte superior del gráfico, donde se encuentran estos ligandos, la variabilidad en energías electrostáticas es muy grande, marcando así parte de la fuerza impulsora a la selectividad y afinidad por estos ligandos en particular. A pesar de que la magnitud de las fuerzas dispersivas (**Figura III.13B**) es mucho menor que la de las electrostáticas, se puede observar que son más importantes (más negativas) para los tetrasacáridos de los grupos sanguíneos y para el α 2,3-sialil-LacNAc, mientras que para los residuos sulfatados no muestran diferencias significativas con respecto a lo observado para Lac/LacNAc (salvo casos puntuales como Gal-2:4S-LacNAc y Gal-8:6S-LacNAc, en donde claramente el aporte a la unión de parte de las fuerzas no polares es casi nulo). Si bien el sustituyente SO₃⁻ puede ocupar la posición 6, como antes hemos discutido, el resultado final es una energía libre de unión la mayoría de las veces menos negativa que para el ligando canónico de LacNAc. Por otro lado, no se observa mucha variabilidad de estos mismos valores de energía libre cuando miramos el ligando 3S-LacNAc, en donde podemos justificar esta tendencia por lo anteriormente dicho sobre la exposición al solvente de la posición 3. Sólo se muestra una singularidad de un valor muy negativo en la energía electrostática para el complejo que 3S-LacNAc forma con Gal-3 que no se propaga al valor final de energía libre de unión. Esto es debido a la presencia particular de una Arg apuntando

en dirección al grupo sulfato en la región del MBS B, estabilizando la unión del ligando a la proteína mediante una fuerte interacción carga-carga que no es posible en otras galectinas (por ejemplo, Gal-1 y Gal-2 presentan un residuo de Val en dicha posición) (**Figura III.18**). Notablemente, esta Arg se encuentra también en Gal-7, -8N y -9 (N y C) donde se observan valores de energía electrostática de unión más negativos que los observados para el complejo de dichas galectinas con LacNAc.

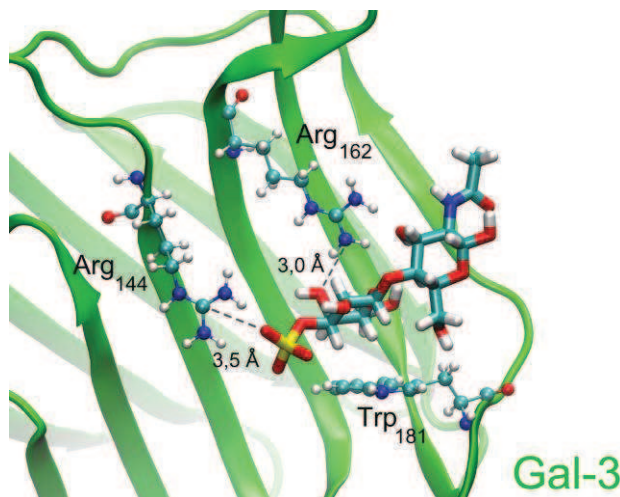


Figura III.18. 3S-LacNAc en Gal-3. La presencia singular de una Arg en el MBS B (Arg144) permite justificar la estabilización electrostática particular de este complejo.

Finalmente, para los complejos formados con 4S-LacNAc es posible observar una mayor variabilidad tanto a nivel electrostático como de balance final en la energía libre de unión. Resultan interesantes los valores más negativos para Gal-7 y Gal-9N (y en menor medida Gal-1) y los valores menos negativos para Gal-2 y los DRC de las otras galectinas en tándem estudiadas (Gal-8N, -8C y -9C). Los valores muy negativos para Gal-7 y Gal-9N se justifican al observar una fuerte interacción entre el grupo negativo sulfato y un grupo de aminoácidos variados que incluyen Arg, Asn e His (**Figura III.19**). A pesar de que esta región topológica de análisis es bastante similar para Gal-2 y las galectinas en tándem, dado que por defecto estas son un poco menos afines por LacNAc, parece ser que la presencia de un grupo sulfato en posición 4 del ligando, si bien no rompe la interacción por completo (en el tiempo simulado), ésta no parece ser tan favorable como en los casos de Gal-7 y -9N pues efectivamente el ligando se encuentra en estas galectinas en tándem un poco más alejado, en dirección al solvente, de los mismos aminoácidos citados en la **Figura III.19**.

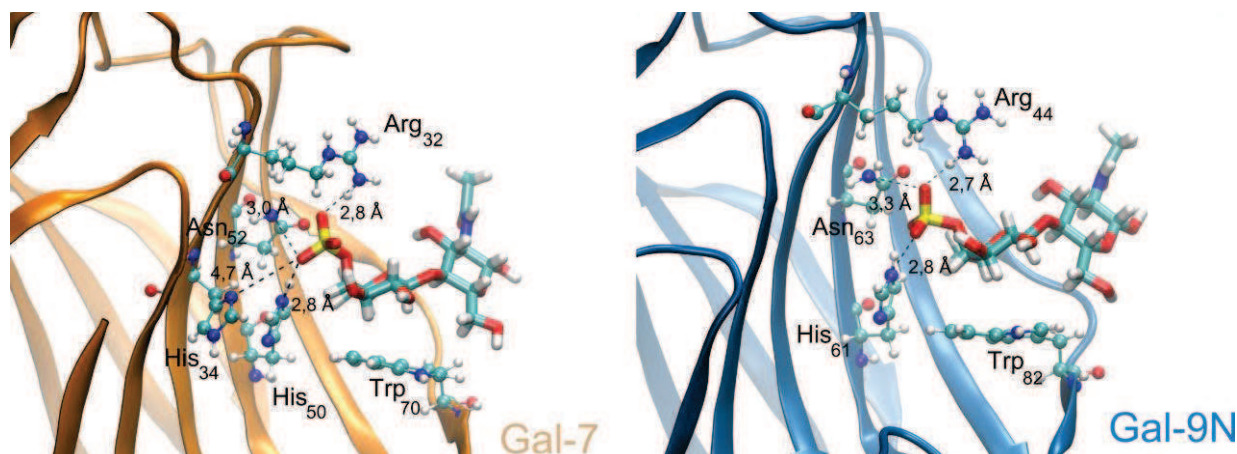


Figura III.19. 4S-LacNAc en A) Gal-7 y B) Gal-9N. Disposición de los aminoácidos claves que participan en la estabilización del grupo sulfato en estas galectinas.

III.4 Discusión y conclusiones

En el capítulo presente, se ha provisto el primer análisis comparativo de estructura, dinámica y propiedades de unión de ligandos de los diferentes miembros de la familia de galectinas humanas. Los resultados proveen explicaciones racionales para las diferentes especificidades por diferentes glicanos, actividad íntimamente relacionada con las funciones de las galectinas. La característica funcional clave que define cada DRC de galectina humana está relacionada con su especificidad por unión a carbohidratos, y su habilidad por multimerizar e interconectar de este modo dos sitios diferentes de unión gracias a la arquitectura de tipo tándem. Nuestra hipótesis central en esta etapa del trabajo se basó en la afirmación de que la especificidad de unión a glicanos y la correspondiente afinidad son resultado de las diferentes arquitecturas, topologías y dinámicas de cada LBG, las cuales a su vez influyen en las propiedades bioquímicas de cada DRC. Para elaborar esta idea, realizamos un análisis comparativo exhaustivo de la estructura y dinámica de cada galectina, usando diversas técnicas de simulación computacional. Como se describió en la introducción al capítulo, los resultados fueron presentados en el contexto de la clasificación de las galectinas de acuerdo a la organización del DRC según prototipo, tándem y tipo quimera [20] y también según la clasificación filogenética [21] más moderna (pero aún poco aceptada en la literatura) que divide a la familia de las galectinas humanas en dos tipos de DRCs: F3 y F4, los cuales a su vez se pueden redistribuir en tres grupos mono-F3, mono-F4 y bi-DRC.

El primer resultado interesante, aunque no sorprendente, que se desprende de este análisis confirma que a pesar de tener un porcentaje promedio de identidad de secuencia de solo el 30%, la estructura de todos los DRCs es muy similar con un plegamiento altamente conservado y con las únicas diferencias estructurales localizadas en regiones específicas de los loops. Sin embargo, un resultado más sutil es el alto grado de similitud con un RMSD en promedio menor a los 2 Å, valores que rondan las desviaciones esperadas en experimentos de cristalografía de rayos-X con proteínas idénticas en distintos ambientes. Estas diferencias, a pesar de que permiten cierta agrupación de los DRCs en diferentes grupos estructurales, muestran casi nula correlación con la secuencia o la clasificación según organización de dominios. Por ende, uno no espera mucha relación entre el mismo grupo (ya sea en una u otra clasificación) y la función que cumple la galectina en cuestión. También, las diferencias en la longitud y las características aminoacídicas en los loops puede ser crucial para determinar la especificidad por un ligando en particular, o incluso en interacciones con otras proteínas. De hecho, la arquitectura de un glicano en particular puede ser realmente compleja y pequeñas diferencias en estas regiones, donde hemos encontrado las principales diferencias estructurales miembro a miembro, pueden ser suficientes, por ejemplo, porque algunas galectinas pueden unir glicanos sialilados y sulfatados, mientras que otras no o al menos con menor afinidad. Esto se pudo verificar a través del análisis posterior del glyco array *in silico* desarrollado.

La segunda conclusión se obtiene del análisis de la dinámica molecular de las galectinas. De manera similar a lo observado con el análisis estructural, el estudio superficial de la dinámica de cada DRC muestra un patrón global de movimientos muy similares para todos, con el principal desplazamiento localizado en las regiones de los loops que conectan las hebras beta de la estructura secundaria. Resulta interesante notar sin embargo que cada loop y su movimiento particular son característica particular de cada DRC, y sólo algunas similitudes menores fueron identificadas comunes a toda la familia. Un resultado relevante de este análisis es aquel que muestra fluctuaciones mayores y más homogéneas para las galectinas del tipo bi-DRC en comparación con las del tipo mono-DRC (**Figura III.9**). Es muy difícil determinar si estos patrones distintivos son una causa o una consecuencia de la organización del DRC. Sin embargo, dado que la organización por dominios no guarda relación filogenética, en vista de estos resultados, es razonable especular que patrones dinámicos diferenciales entre mono- y bi-DRCs sean consecuencia de la organización

por dominio. Las galectinas mono-DRC muestran generalmente equilibrios de dimerización. Por ende, una movilidad elevada podría interferir en el proceso de dimerización. Otro tipo posible de rol funcional de este tipo de dinámica observado es el impacto del movimiento de la proteína sobre la modulación de la afinidad por ligandos, a través del cambio de entropía conformacional del DRC. Como se muestra en nuestros resultados (**Tabla III.6**), la pérdida energética debida a esta entropía puede llegar a ser de hasta 30 kcal mol⁻¹ (a 300 K), pudiendo ser en determinados casos razón suficiente para unir o no un determinado glicanos. En este contexto dinámico, un patrón de movimientos menos restrictivo como el de los bi-DRCs puede llegar a permitir mayor espectro de afinidades que el ofrecido por la dinámica de las galectinas mono-DRCs, con características dinámicas más reguladas para conectar afinidad y dimerización en la misma estructura y dinámica, tal como se ha observado para Gal-1 [49, 50]. Finalmente, alta o baja movilidad/entropía conformacional podrían también explicar porqué algunas galectinas unen glicanos lineales (como poli-N-acetilactosaminas), mientras que otros tienen alta afinidad por glicanos ramificados, dado que la rigidez estructural es menos importante para los primeros que para los segundos [15, 34, 51].

En su conjunto, estos resultados muestran que desde un punto de vista global, todos los DRCs humanos son muy similares tanto en estructura como en dinámica, con algunas singularidades que merecen especial atención. En consecuencia, la pregunta clave que aún persiste es cuál/es es/son las diferencias relevantes que determinan la necesidad para tantas galectinas humanas. Claramente, esta respuesta escapa de las técnicas computacionales y necesita ser abarcada desde un punto de vista sistémico desde la genómica y la experimentación con modelos animales.

Finalmente, se ofrece un estudio comparativo detallado de cada LBG y los residuos relevantes que forman parte de este surco de unión. Las contribuciones a la energía de unión de cada monosacárido, las cuales no se pueden obtener mediante métodos experimentales, y su comparación muestran que las diferencias claves entre DRCs son consecuencia de la topología de cada MBS. Se ha mostrado que muchos residuos a lo largo del LBG en los MBS Z, B, C y D están altamente conservados, como ser el Trp 2-S6, determinante en la elevada y similar contribución de vdW a la energía de unión en el MBS C a lo largo de la familia de galectinas. Sin embargo, para cada DRC existen sustituciones únicas que carecen de relación filogenética o de organización según dominios que resultan en claras diferencias en las contribuciones

energéticas a cada MBS, las cuales podrían explicar la afinidad y selectividad por diferentes glicanos, demostrando así la variedad de funciones biológicas relacionadas con la unión a polisacáridos. Evidencia parcial de este hecho son los casos particulares de análisis de Gal-10 y hGRPC, las cuales poseen más sustituciones de residuos conservados en el LBG que podrían explicar por qué Gal-10 une manosa en lugar de lactosa y por qué hGRPC no une ningún glicano testeado hasta la fecha. Es más, la selectividad diferencial por un conjunto de ligandos en la mayoría de las galectinas aquí estudiadas ha sido puesta a prueba en el desarrollo de la estrategia del *glyco array in silico*. Con el desarrollo y puesta a punto de una metodología computacional para poner de manifiesto las diferencias de unión a glicanos fuimos capaces de lograr rescatar una imagen global de la selectividad de esta familia particular de lectinas. La mayoría de los glycan arrays experimentales hasta ahora desarrollados han mostrado determinadas limitaciones, no sólo para caracterizar la selectividad en galectinas, sino en la mayoría de las proteínas de unión a glicanos [42]. Muchas de estas limitaciones se deben a la dificultad de sintetizar una gran cantidad de glicanos. Además, la forma en la que cada glicano se presenta en el array puede ser clave para permitir una presentación a la lectina de la forma lo más natural posible, sin incorporar artefactos estructurales y dinámicos del glicano adherido a la superficie de prueba que impidan la correcta determinación de la unión [52, 53]. Por otro lado, testear experimentalmente 15 proteínas (como el caso de las galectinas) frente a todos los ligandos disponibles puede ser un desafío de expresión y purificación, sin mencionar la inversión económica que conlleva (la mayoría de las galectinas aquí estudiadas no están disponibles para su correcta expresión, por lo que un objetivo como ese escapa las escalas de tiempo de un proyecto acotado como lo es un doctorado). Por ello, un tratamiento computacional, como el llevado a cabo en esta Tesis, es capaz de sobrellevar estas desventajas en forma considerable. “Sintetizar” todos los ligandos es algo que se hace medianamente fácil con la ayuda del paquete GLYCAM. “Obtener” todas las galectinas fue el objetivo de esta Tesis y se ha logrado, salvo para el extraño C-terminal de Gal-12. “Enfrentar” todas las proteínas con todos los ligandos se hizo manualmente, pero en nuestro grupo se está trabajando para automatizarlo. “Leer” el *glycoarray* es algo casi automático, con la ayuda de los cálculos de MM-GBSA implementados en AMBER y algún programa particular creado para automatizar la lectura y graficado de los resultados de una forma fácil de lectura y visualización. Con este enfoque, pudimos encontrar nueva evidencia energética y estructural del porqué es imposible la unión de las galectinas a

los ligandos α 2,6-sialilados y una menor afinidad por los 6-sulfatados de LacNAc, teniendo esto un impacto clave en la función de galectinas: si existe un entorno en donde estos glicanos predominan, seguramente el rol de galectinas sería nulo o no sería de esperar que sea mediado por la unión a estos azúcares (sino tal vez por interacciones proteína-proteína, pobremente descriptas aún en la bibliografía de galectinas). Por otra parte, hemos logrado resaltar por este medio la baja afinidad de Gal-10 por todos estos β -galactósidos, y complementar así la información antes encontrada en la misma dirección. Finalmente, las singularidades que aparecen en este tipo de screenings pueden ser clave para poder encontrar justificación estructural rápida y directa al seleccionar una pareja galectina:ligando determinada en el rol de determinado efecto biológico, sustentado con evidencia experimental. O también puede ser de gran uso en el diseño racional de inhibidores. Si se desea profundizar una interacción de un ligando en una galectina por sobre otra, es posible detectar con celeridad las posiciones más críticas en un ligando molde determinado para colocar un sustituyente cargado positiva o negativamente o de determinado volumen o las características electrostáticas adecuadas para ponderar una unión por sobre otra.

III.5 Referencias

1. Lopez-Lucendo, M.F., Solis, D., Andre, S., Hirabayashi, J., Kasai, K., Kaltner, H., Gabius, H.J., and Romero, A. **(2004)** Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *Journal of Molecular Biology* 343, 957-70
2. Cho, M. and Cummings, R.D. **(1995)** Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. I. Physical and chemical characterization. *Journal of Biological Chemistry* 270, 5198-206
3. Liao, D.I., Kapadia, G., Ahmed, H., Vasta, G.R., and Herzberg, O. **(1994)** Structure of S-lectin, a developmentally regulated vertebrate beta-galactoside-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* 91, 1428-32
4. Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S.H., and Rini, J.M. **(1998)** X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1 Å resolution. *J Biol Chem* 273, 13047-52
5. Nagae, M., Nishi, N., Nakamura-Tsuruta, S., Hirabayashi, J., Wakatsuki, S., and Kato, R. **(2008)** Structural Analysis of the Human Galectin-9 N-terminal Carbohydrate Recognition Domain Reveals Unexpected Properties that Differ from the Mouse Orthologue. *Journal of Molecular Biology* 375, 119-35
6. Nagae, M., Nishi, N., Murata, T., Usui, T., Nakamura, T., Wakatsuki, S., and Kato, R. **(2009)** Structural analysis of the recognition mechanism of poly-N-acetyllactosamine by the human galectin-9 N-terminal carbohydrate recognition domain. *Glycobiology* 19, 112-17
7. Yoshida, H., Teraoka, M., Nishi, N., Nakakita, S.-i., Nakamura, T., Hirashima, M., and Kamitori, S. **(2010)** X-ray Structures of Human Galectin-9 C-terminal Domain in Complexes with a Biantennary Oligosaccharide and Sialyllactose. *Journal of Biological Chemistry* 285, 36969-76

8. Zhou, D., Ge, H., Sun, J., Gao, Y., Teng, M., and Niu, L. **(2008)** Crystal structure of the C-terminal conserved domain of human GRP, a galectin-related protein, reveals a function mode different from those of galectins. *Proteins: Structure, Function, and Bioinformatics* 71, 1582-88
9. Wälti, M.A., Thore, S., Aebi, M., and Künzler, M. **(2008)** Crystal structure of the putative carbohydrate recognition domain of human galectin-related protein. *Proteins: Structure, Function, and Bioinformatics* 72, 804-08
10. Kato-Murayama, M., Murayama, K., Terada, T., Shirouzu, M., and Yokoyama, S. **(2006)** Riken Structural Genomics PROTEOMICS INITIATIVE (RSGI)
11. Tomizawa, T., Kigawa, T., Saito, K., Koshiba, S., Inoue, M., and Yokoyama, S. **(2005)** Solution Structure of the C-Terminal Gal-Bind Lectin Domain from Human Galectin-4. *Structural Genomics PROTEOMICS INITIATIVE (RSGI)*
12. Kishishita, S., Nishino, A., Murayama, K., Terada, T., Shirouzu, M., and Yokoyama, S. **(2007)** Crystal structure of N-terminal domain of human galectin-8 with D-lactose. *Riken Structural Genomics PROTEOMICS INITIATIVE (Rsgi)*
13. Kishishita, S., Nishino, A., Murayama, K., Terada, T., Shirouzu, M., and Yokoyama, S. **(2007)** Crystal structure of N-terminal domain of human galectin-8. *Riken Structural Genomics PROTEOMICS INITIATIVE (Rsgi)*
14. Tomizawa, T., Koshiba, S., Inoue, M., Kigawa, T., and Yokoyama, S. **(2007)** Solution Structure Of The C-Terminal Gal-Bind Lectin Protein From Human Galectin-8. *Riken, Structural Genomics PROTEOMICS INITIATIVE (RSGI)*
15. Lobsanov, Y.E.A. **(1993)** X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9 Å resolution. *J Biol Chem* 268, 27034-38
16. Leonidas, D.D., Vatzaki, E.H., Vorum, H., Celis, J.E., Madsen, P., and Acharya, K.R. **(1998)** Structural Basis for the Recognition of Carbohydrates by Human Galectin-7. *Biochemistry* 37, 13930-40
17. Leonidas, D.D., Elbert, B.L., Zhou, Z., Leffler, H., Ackerman, S.J., and Acharya, K.R. **(1995)** Crystal structure of human Charcot-Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins. *Structure* 3, 1379-93
18. Visegrády, B., Than, N.G., Kilár, F., Sümegi, B., Than, G.N., and Bohn, H. **(2001)** Homology modelling and molecular dynamics studies of human placental tissue protein 13 (galectin-13). *Protein Engineering* 14, 875-80
19. Than, N.G., Pick, E., Bellyei, S., Szigeti, A., Burger, O., Berente, Z., Janaky, T., Boronkai, A., Kliman, H., Meiri, H., Bohn, H., Than, G.N., and Sumegi, B. **(2004)** Functional analyses of placental protein 13/galectin-13. *European Journal of Biochemistry* 271, 1065-78
20. Hirabayashi, J. and Kasai, K.-i. **(1993)** The family of metazoan metal-independent β -galactoside-binding lectins: structure, function and molecular evolution. *Glycobiology* 3, 297-304
21. Houzelstein, D., Gonçalves, I.R., Fadden, A.J., Sidhu, S.S., Cooper, D.N.W., Drickamer, K., Leffler, H., and Poirier, F. **(2004)** Phylogenetic analysis of the vertebrate galectin family. *Mol Biol Evol* 21, 1177-87
22. Lee, E.H., Hsin, J., Sotomayor, M., Comellas, G., and Schulten, K. **(2009)** Discovery Through the Computational Microscope. *Structure* 17, 1295-306
23. Leach, A.R. **(2001)** *Molecular Modelling: Principles and Applications*. Pearson Education EMA
24. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. **(2006)** Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function and Genetics* 65, 712-25
25. Woods, R.J., Dwek, R.A., Edge, C.J., and Fraser-Reid, B. **(1995)** Molecular Mechanical and Molecular Dynamic Simulations of Glycoproteins and Oligosaccharides. 1. GLYCAM_93 Parameter Development. *The Journal of Physical Chemistry* 99, 3832-46
26. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L., and Woods, R.J. **(2008)** GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *Journal of Computational Chemistry* 29, 622-55
27. Leffler, H., Carlsson, S., Hedlund, M., Qian, Y., and Poirier, F. **(2002)** Introduction to galectins. *Glycoconjugate Journal* 19, 433-40

28. *Essentials of Glycobiology*. (2009) Editado por Ajit Varki, Richard D Cummings, Jeffrey D Esko, Hudson H Freeze, Pamela Stanley, Carolyn R Bertozzi, Gerald W Hart, and Marilynn E Etzler. Cold Spring Harbor Laboratory Press
29. Černý, V. (1985) Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications* 45, 41-51
30. Kirkpatrick, S., Gelatt, C.D., and Vecchi, M.P. (1983) Optimization by Simulated Annealing. *Science* 220, 671-80
31. Moens, L., Vanfleteren, J., Van De Peer, Y., Peeters, K., Kapp, O., Czeluzniak, J., Goodman, M., Blaxter, M., and Vinogradov, S. (1996) Globins in nonvertebrate species: Dispersal by horizontal gene transfer and evolution of the structure-function relationships. *Molecular Biology and Evolution* 13, 324-33
32. Scheeff, E.D. and Bourne, P.E. (2005) Structural Evolution of the Protein Kinase-Like Superfamily. *PLoS Comput Biol* 1, e49
33. Hirabayashi, J. and Kasai, K. (1991) Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. *Journal of Biological Chemistry* 266, 23648-53
34. Leffler, H. and Barondes, S.H. (1986) Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. *Journal of Biological Chemistry* 261, 10119-26
35. Swaminathan, G.J., Leonidas, D.D., Savage, M.P., Ackerman, S.J., and Acharya, K.R. (1999) Selective Recognition of Mannose by the Human Eosinophil Charcot-Leyden Crystal Protein (Galectin-10): A Crystallographic Study at 1.8 Å Resolution. *Biochemistry* 38, 13837-43
36. Ford, M.G., Weimar, T., Köhli, T., and Woods, R.J. (2003) Molecular Dynamics Simulations of Galectin-1-oligosaccharides Complexes Reveal the Molecular Basis of Ligand Diversity. *PROTEINS: Structure, Function and Genetics* 53, 229-40
37. Nesmelova, I.V., Ermakova, E., Daragan, V.A., Pang, M., Menéndez, M., Lagartera, L., Solís, D., Baum, L.G., and Mayo, K.H. (2010) Lactose Binding to Galectin-1 Modulates Structural Dynamics, Increases Conformational Entropy, and Occurs with Apparent Negative Cooperativity. *Journal of Molecular Biology* 397, 1209-30
38. Diehl, C., Engström, O., Delaine, T., Håkansson, M., Genheden, S., Modig, K., Leffler, H., Ryde, U., Nilsson, U.J., and Akke, M. (2010) Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3. *Journal of the American Chemical Society* 132, 14577-89
39. Ermakova, E., Miller, M.C., Nesmelova, I.V., López-Merino, L., Berbís, M.A., Nesmelov, Y., Tkachev, Y.V., Lagartera, L., Daragan, V.A., André, S., Cañada, F.J., Jiménez-Barbero, J., Solís, D., Gabius, H.-J., and Mayo, K.H. (2013) Lactose binding to human galectin-7 (p53-induced gene 1) induces long-range effects through the protein resulting in increased dimer stability and evidence for positive cooperativity. *Glycobiology* 23, 508-23
40. Guardia, C.M.A., Gauto, D.F., Di Lella, S., Rabinovich, G.A., Martí, M.A., and Estrin, D.o.A. (2011) An Integrated Computational Analysis of the Structure, Dynamics, and Ligand Binding Interactions of the Human Galectin Network. *Journal of Chemical Information and Modeling* 51, 1918-30
41. Schwarz, F.P., Ahmed, H., Bianchet, M.A., Amzel, L.M., and Vasta, G.R. (1998) Thermodynamics of Bovine Spleen Galectin-1 Binding to Disaccharides: Correlation with Structure and Its Effect on Oligomerization at the Denaturation Temperature†. *Biochemistry* 37, 5867-77
42. Brewer, C.F. (2005) Thermodynamic binding studies of galectin-1, -3 and -7. *Glycoconjugate Journal* 19, 459-65
43. Ideo, H., Seko, A., Ishizuka, I., and Yamashita, K. (2003) The N-terminal carbohydrate recognition domain of galectin-8 recognizes specific glycosphingolipids with high affinity. *Glycobiology* 13, 713-23
44. Hou, T., Wang, J., Li, Y., and Wang, W. (2010) Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* 51, 69-82

45. Stowell, S.R., Arthur, C.M., Mehta, P., Slanina, K.A., Blixt, O., Leffler, H., Smith, D.F., and Cummings, R.D. **(2008)** Galectin-1, -2, and -3 Exhibit Differential Recognition of Sialylated Glycans and Blood Group Antigens. *Journal of Biological Chemistry* 283, 10109-23
46. Zhuo, Y. and Bellis, S.L. **(2011)** Emerging Role of α 2,6-Sialic Acid as a Negative Regulator of Galectin Binding and Function. *Journal of Biological Chemistry* 286, 5935-41
47. Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W.E.G., Yagi, F., and Kasai, K.-i. **(2002)** Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1572, 232-54
48. Levy, Y., Auslender, S., Eisenstein, M., Vidavski, R.R., Ronen, D., Bershadsky, A.D., and Zick, Y. **(2006)** It depends on the hinge: a structure-functional analysis of galectin-8, a tandem-repeat type lectin. *Glycobiology* 16, 463-76
49. Di Lella, S., Martí, M.A., Croci, D.O., Guardia, C.M.A., Díaz-Ricci, J.C., Rabinovich, G.A., Caramelo, J.J., and Estrin, D.o.A. **(2010)** Linking the Structure and Thermal Stability of β -Galactoside-Binding Protein Galectin-1 to Ligand Binding and Dimerization Equilibria. *Biochemistry* 49, 7652-58
50. Stowell, S.R., Cho, M., Feasley, C.L., Arthur, C.M., Song, X., Colucci, J.K., Karmakar, S., Mehta, P., Dias-Baruffi, M., McEver, R.P., and Cummings, R.D. **(2009)** Ligand reduces galectin-1 sensitivity to oxidative inactivation by enhancing dimer formation. *J Biol Chem* 284, 4989-99
51. Rabinovich, G.A. and Toscano, M. **(2009)** Turning "sweet" on immunity: Galectin-glycan interactions in immune tolerance and inflammation. *Nature Rev Immunol* 9, 338-52
52. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M.E., Alvarez, R., Bryan, M.C., Fazio, F., Calarese, D., Stevens, J., Razi, N., Stevens, D.J., Skehel, J.J., van Die, I., Burton, D.R., Wilson, I.A., Cummings, R., Bovin, N., Wong, C.-H., and Paulson, J.C. **(2004)** Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* 101, 17033-38
53. Leppänen, A., Stowell, S., Blixt, O., and Cummings, R.D. **(2005)** Dimeric Galectin-1 Binds with High Affinity to α 2,3-Sialylated and Non-sialylated Terminal N-Acetylglucosamine Units on Surface-bound Extended Glycans. *Journal of Biological Chemistry* 280, 5549-62



ROL DE LAS CISTEINAS EN LA FUNCION Y DINAMICA DE GAL-1 Y SU FAMILIA

IV.1 Introducción

Desde su descubrimiento, se ha establecido que la mayoría de las galectinas requieren un ambiente reductor para poder llevar a cabo su función [1]. La relevancia de los procesos de oxidación en la estructura y función de ciertas galectinas ha sido demostrada por ensayos de caracterización bioquímica [2-4], y la contribución de los residuos de cisteína en estos procesos a la actividad como lectina se ha estudiado por ensayos de mutagénesis dirigida [5, 6] y modificación química [7-9]. Sin embargo, a pesar de la evidencia disponible demostrando la importancia de la oxidación de Gal-1 en su función, un consenso firme en la importancia de cada cisteína y las bases moleculares para comprender el mecanismo de oxidación no ha sido alcanzado.

Uno de los objetivos de esta Tesis es poder lograr un entendimiento de los determinantes atómicos y moleculares que caracterizan a la familia de galectinas humanas, y en el Capítulo III se ha puesto en evidencia la aparente importancia que tienen las cisteínas en este contexto. En particular, en la **Tabla III.3** se pusieron de manifiesto dos características claves y llamativas en la comparación de ciertos datos estructurales relevantes: por un lado, el caso singular de Gal-1 con 6 residuos de cisteína en su relativamente corta secuencia y, por otra parte, la gran variabilidad en la aparición de este residuo particular a lo largo de toda la familia de galectinas humanas. Con la idea de poder determinar si existe algún motivo estructural, dinámico y funcional particular del porqué/para qué de estas dos características sobresalientes provenientes de esa primera inspección estructural, es que nos adentramos en este capítulo al estudio mucho más detallado del rol de las cisteínas en la relación estructura-función en las galectinas.

La característica bioquímica particular del elevado número de cisteínas en la secuencia de Gal-1 implica una gran sensibilidad a sufrir cambios por oxidación, proceso electroquímico que induce la pérdida de actividad de lectina [10]. Es por ello que, si bien Gal-1 ha sido extensamente utilizada como modelo de unión a glicanos y multimerización, resulta igualmente interesante como modelo para explorar otros comportamientos de la familia de galectinas, como ser su actividad redox. Es pertinente recalcar que la primera estructura de rayos X reportada para Gal-1 humana [11] reveló no sólo la distribución espacial de las cisteínas sino también modificaciones en estos residuos como ser la formación de ácido sulfénico y disulfuros mixtos con 2-mercaptoetanol (ME). Luego, los enlaces disulfuro intracelulares presentes en la estructura de Gal-1 oxidada (aún no obtenida en forma directa) se han caracterizado mediante espectrometría de masa [10, 12]. Mientras que la forma reducida de esta lectina es crítica para llevar a cabo su función pro-apoptótica e inmunoregulatoria, se ha postulado recientemente una nueva función como factor de crecimiento axonal durante la regeneración nerviosa en nervios periféricos para la forma oxidada de Gal-1 [12]. Por lo tanto, estas fluctuaciones entre los diferentes estados redox de Gal-1 podrían llegar a controlar cierto rango determinado de funciones biológicas claramente dependientes del estado de oxidación de las cisteínas.

Es sabido que la inactivación de Gal-1 por aire es un proceso muy lento y que puede ser catalizado por trazas de metales pesados como Cu. Sin embargo, la cinética del proceso de oxidación no ha sido caracterizada. Esto se debe, en parte, a que las proteínas expuestas al aire se encuentran sometidas a la exposición de varios oxidantes que actúan bajo mecanismos aún poco estudiados o comprendidos. Una excepción se puede hacer para el comportamiento frente al peróxido de hidrogeno (H_2O_2), una de las especies reactivas de oxígeno (ROS, del inglés *reactive oxygen species*) más importantes [13]. Sumados a los ya estudiados y bastante comprendidos roles de esta molécula como microbicida, de molécula de envejecimiento y estrés oxidativo, se suma una función crítica como molécula de señalización [14, 15]. Por lo tanto, además de estudiar el comportamiento al aire, hemos incorporado el H_2O_2 a nuestro estudio para explorar la posibilidad de que sea un candidato capaz de oxidar Gal-1 *in vivo*.

Usando una aproximación multidisciplinaria al problema de la oxidación, mediante la realización de experimentos *in silico* e *in vitro*, hemos investigado los determinantes moleculares y estructurales responsables de la inactivación redox-

dependiente de Gal-1 e intentando interpretar estos resultados en el contexto de la comparación de la familia de galectinas humanas.

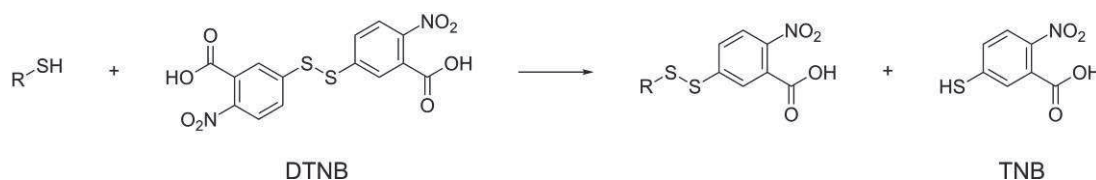
IV.2 Métodos experimentales y computacionales

Expresión y purificación de Gal-1 y las mutantes CXS. La purificación de todas las proteínas recombinantes siguió el protocolo explicado en la sección II.2.2. Adicionalmente, para prevenir la formación enlaces disulfuro mixtos entre las cisteínas y el ME utilizado a lo largo de toda la purificación, antes de cada análisis, el ME fue removido de la solución mediante una incubación de la proteína en PBS con ditioneitol (DTT) 10 mM en hielo durante 30 min. Para separar la proteína completamente libre y reducida del exceso de DTT y ME, las muestras se desalaron con columnas NAP-5 (GE Healthcare), se purgaron con Ar (g) en un vial cerrado y la proteína en solución se mantuvo en hielo hasta su uso. Para generar las mutantes de cada cisteína por serina (CXS) en Gal-1, se utilizó el método de PCR inversa [16] para generar los plásmidos para transformar las bacterias. Para las dobles mutantes C2S-C16S y C2S-C88S se realizó otra ronda de mutación simple partiendo del plásmido que contiene la secuencia de la mutante C2S. El *primer* sentido contiene un error adrede que cambia la cisteína a mutarse por una serina (**Tabla IV.1**). Estos *primers* se usaron en combinación con los *primers* antisentido que comienzan al inicio de los *primers* sentido. El inserto y el vector fueron amplificados en el mismo paso con la polimerasa KOD Hot Start (Novagen) y el producto de PCR fue ligado con la ligasa T4 DNA (Promega). Las mutaciones fueron chequeadas por secuenciación completa del inserto de ADN.

Mutante	Dirección	Primer
C2S	sentido	5'-ATATGGCTT CT GGTCTGG-3'
	antisentido	5'-GTATATCTCCTTCTTAAAGTTAAAC-3'
C16S	sentido	5'-CTGGAGAGT CC CTTCGAGTG-3'
	antisentido	5'-GTTTGAGATTCAGGTTGCTGG-3'
C42S	sentido	5'-CAACCTGT CC CTGCACTTC-3'
	antisentido	5'-TTGCTGTCTTTGCCCAGGTTTC-3'
C60S	sentido	5'-CCATCGTGT CC AACAGCAAG-3'
	antisentido	5'-TGTTGGCGTCGCCGTG-3'
C88S	sentido	5'-CAGAGGTGT CC ATCACCTTC-3'
	antisentido	5'-CAACACTTCCAGGCTGGAAG-3'
C130S	sentido	5'-CAAGATCAAAT CT GTGGCCTTTG-3'
	antisentido	5'-AAGTCACCGTCAGCTGC-3'

Tabla IV.1 Primers utilizados para generar las mutantes CXS. Los codones mutados están resaltados en negrita.

Cuantificación de oxidantes, proteína y tioles. La concentración de H_2O_2 se midió a 240 nm ($\epsilon_{240} = 43,6 \text{ M}^{-1} \text{ cm}^{-1}$). La concentración de proteína libre y reducida se midió también espectrofotométricamente usando una absorptividad molar a 280 nm de $8480 \text{ M}^{-1} \text{ cm}^{-1}$ para Gal-1 y sus mutantes, según la estimación a partir de su secuencia primaria (<http://www.expasy.ch/tools/protparam.html>). La concentración de tioles se determinó mediante el uso del ácido 5,5'-ditiobis-(2-nitrobenzónico) (DTNB), luego de incubar las muestras de Gal-1 con un exceso de DTNB en PBS por 30 min en la oscuridad a temperatura ambiente. Se utilizó una absorptividad molar a 412 nm de $14150 \text{ M}^{-1} \text{ cm}^{-1}$ [17] para cuantificar la concentración del anión 3-tio-6-nitrobenzoato (TNB) (**Esquema IV.1**), usando como blanco la absorbancia de una solución de DTNB y Gal-1 a la misma concentración que el ensayo y a la misma longitud de onda de la determinación.



Esquema IV.1 Reacción de Ellman, para cuantificar tioles reducidos a partir de la absorbancia del anión 3-tio-6-nitrobenzoato a 412 nm.

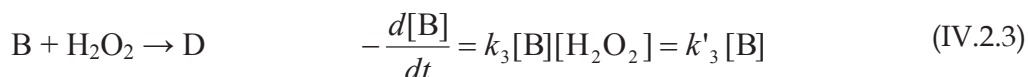
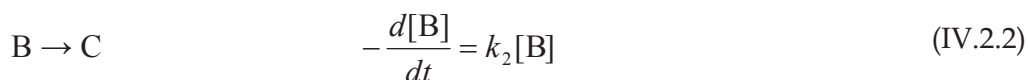
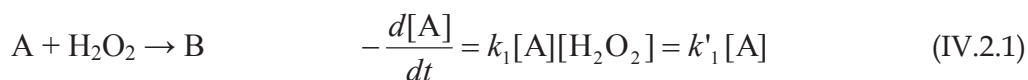
Mediciones espectrofotométricas. Se registraron espectros de DC (sección II.2.4.3) en el UV lejano y cercano (promedios de al menos 8 escaneos) restando los blancos correspondientes a los buffers adecuados (paso óptico = 1 mm). Se realizaron las deconvoluciones de los espectros del UV lejano usando Dichro Web (<http://dichroweb.cryst.bbk.ac.uk>) con el análisis CONTIN y la base de referencia SP175. Para los espectros de fluorescencia (sección II.2.4.2) se utilizó una longitud de onda de excitación de 280 nm (295 nm para los experimentos de la cinética de cambio conformacional) (filtro de paso de banda de 1 nm) y los espectros de emisión fueron registrados en el rango entre 300 y 420 nm (paso de banda de 3 nm). Se promediaron al menos 6 escaneos antes de proceder a la utilización del espectro. Los efectos de dilución se corrigieron, siendo siempre menores al 10%.

Cinética de reacción entre Gal-1 y H_2O_2 . Las constantes de velocidad de oxidación por H_2O_2 de Gal-1 y las mutantes CXS se determinaron usando condiciones de pseudo-primer orden como antes descripto [18]. Las reacciones se llevaron a cabo a concentraciones crecientes de peróxido y usando la misma concentración de proteína

(67 μM). A diferentes tiempos de reacción (0, 2, 5, 8, 15, 35, 60, 90 y 120 min) se extrajeron alícuotas de la mezcla de reacción, se mezclaron con catalasa (100 U mL^{-1}) para eliminar el exceso de H_2O_2 y frenar la reacción de oxidación, y se reservaron en hielo hasta la medición del contenido de tioles libres. Las constantes de velocidad observadas de consumo de tiol (k') fueron determinadas por ajuste de los datos a una función biexponencial de la evolución temporal de la concentración de tioles a las diferentes concentraciones de oxidante ensayadas.

Cinética del cambio conformacional de Gal-1 en la oxidación con H_2O_2 .

Llamemos A a Gal-1 en su forma reducida, B a Gal-1 con sus cisteínas oxidadas a la especie reactiva ácido sulfénico, las cuales son capaces de formar un enlace disulfuro con otra cisteína reducida para arrojar la forma oxidada de Gal-1 final (C), cuyo cambio conformacional es el medido mediante la señal de fluorescencia y que, a su vez, es capaz de formar irreversiblemente otras especies sobreoxidadas, en las formas de ácido sulfínico/sulfónico (D). El mecanismo propuesto para esta serie de reacciones es



Dado que todos los experimentos fueron llevados a cabo en condiciones de pseudo-orden en Gal-1 (A-D), la concentración de H_2O_2 en exceso puede agruparse en constantes de pseudo-primer orden (k'_1 y k'_3). Combinando estas tres ecuaciones, se obtiene una expresión para la variación de la concentración de la especie B en función del tiempo:

$$\frac{d[\text{B}]}{dt} = k'_1[\text{A}] - k_2[\text{B}] - k'_3[\text{B}] \quad (\text{IV.2.4})$$

Por otro lado, la expresión integrada de IV.2.1 muestra la concentración de A en función del tiempo según:

$$[A](t) = [A]_0 e^{-k'_1 t} \quad (\text{IV.2.5})$$

Sustituyendo IV.2.5 en IV.2.4 y resolviendo la ecuación diferencial de primer orden resultante, es posible obtener una expresión analítica para la evolución temporal de [B], y con ella, integrando nuevamente, otras para [C] y [D]:

$$[B](t) = \frac{[A]_0 k'_1}{k_2 + k'_3} \left[e^{-k'_1 t} - e^{-(k_2 + k'_3)t} \right] \quad (\text{IV.2.6})$$

$$[C](t) = \frac{[A]_0 k'_1 k_2}{k_2 + k'_3 - k'_1} \left[\frac{1 - e^{-k'_1 t}}{k'_1} + \frac{e^{-(k_2 + k'_3)t} - 1}{k_2 + k'_3} \right] \quad (\text{IV.2.7})$$

$$[D](t) = \frac{[A]_0 k'_1 k'_3}{k_2 + k'_3 - k'_1} \left[\frac{1 - e^{-k'_1 t}}{k'_1} + \frac{e^{-(k_2 + k'_3)t} - 1}{k_2 + k'_3} \right] \quad (\text{IV.2.8})$$

Con estas ecuaciones es posible realizar un ajuste global de la señal de fluorescencia total a la contribución de cada especie, incluyendo la variación de la concentración de H_2O_2 . Este modelo se utilizó para explicar la forma de la señal en función del tiempo y obtener de allí los distintos parámetros cinéticos de las reacciones que sufre Gal-1 en presencia de H_2O_2 , y poder comparar así los diferentes cambios que sufre la proteína durante la oxidación.

Análisis electroforético. Se realizaron diversos análisis de electroforesis en gel de poliacrilamida y SDS (15:1) para estudiar las estructuras de las proteínas reducidas y oxidadas, tanto en condiciones reductoras como oxidantes. Los mismos se revelaron usando tinción de plata o azul de Coomassie.

Unión de lactosa a Gal-1. Se determinaron (a 25 °C) las constantes termodinámicas de unión de lactosa (K_b) en condiciones reductoras (presencia de DTT) y oxidantes (proteína previamente oxidada al aire o con H_2O_2) siguiendo el cambio del espectro de fluorescencia de Gal-1 a diferentes concentraciones de lactosa. Se tituló una solución de Gal-1 (8 mM) con agregados conocidos y paulatinos de una solución stock de lactosa 100 mM. Se registró la intensidad de emisión de fluorescencia a 363 nm y se

ajustó la variación de esta señal en función de la concentración final de lactosa mediante un modelo de unión simple a un sitio.

Experimentos computacionales. Las estructuras iniciales de Gal-1 fueron obtenidas de la PDB: 1GZW para la proteína wild-type con lactosa como ligando (experimento de difracción de rayos X con 1,65 Å de resolución) y 1W6N para la mutante C2S de Gal-1, libre de ligando. Las estructuras que no estaban disponibles en la PDB (forma oxidada de Gal-1 con el enlace disulfuro Cys16-Cys88, con y sin lactosa) o aquellas que requerían algún tipo de edición de la estructura (por ejemplo, la estructura depositada bajo el código 1GZW para Gal-1 reducida posee cisteínas en la forma ácido sulfénico, no deseadas cuando se requiere estudiar la forma totalmente reducida), fueron generadas usando el paquete computacional Amber9. En todos los casos, se removieron los iones y las moléculas de agua cristalográficas y se resolvataron con aguas TIP3P (sección II.2) en una caja octaédrica (10 Å de distancia entre la superficie proteína y el borde de la misma). Se llevaron a cabo 60 ns (80 ns para los modelos creados) de DM a 1 atm y 300 K según los protocolos y procedimientos detallados en la secciones II.1.2, II.1.3 y II.1.5 e implementados en detalle en el capítulo III (ver sección III.2).

De las DM de Gal-1 monomérica reducida y de la mutante C2S se seleccionaron 50 fotos al azar y para cada una de ellas se calculó el valor de pKa del grupo -SH en todas las cisteínas presentes en la estructura, usando el programa PROPKA [19]. Este programa involucra un cálculo empírico basado en la determinación de efectos de solvatación e interacción intra-protéica del grupo ácido/básico bajo estudio. Los resultados se expresaron como ΔpK_a , es decir, la diferencia entre el valor promedio obtenido del análisis de todas las fotos y el valor de referencia interno usado por el programa (9,00 para la cisteína). Se incluye también la desviación estándar de cada determinación. Para las mismas fotos, se calculó el área accesible superficial al solvente (del inglés SASA, *solvent accessible surface area*, sección II.1.5.2) de la cadena lateral y el átomo de azufre (SG) de cada cisteína usando el plugin vmdICE.

Los parámetros termodinámicos correspondientes a la unión de lactosa a los diferentes DRC analizados fueron calculados con dos estrategias: el método MM-GBSA, ampliamente discutido en la sección II.1.3.2, y el método MM/QM-COSMO desarrollado por Anisimov y Cavasotto [20]. Éste último método combina las energías clásicas provenientes de una DM, una aproximación continua del solvente (sección

II.1.2) y su accesibilidad a las distintas estructuras proteicas, junto con una reevaluación de las trayectorias usando un Hamiltoniano semi-empírico para lograr una descripción más acabada de las energías de interacción proteína-ligando, donde las entropías traslacionales y rotacionales se calculan usando la integral de configuraciones proveniente del análisis termodinámico-estadístico y las entropías internas fueron estimadas haciendo uso de las propiedades termodinámicas de un oscilador armónico como aproximación.

Se realizó un alineamiento múltiple de las secuencias de las 15 galectinas disponibles usando el método y programa CLUSTAL-W [21]: LGALS-1 a -12, correspondientes a todas las galectinas desde Gal-1 hasta Gal-12, LGALS-13 (o PP13), LGALS-14 (o PPL13) y HSPC159 (hGRPC), éstos últimos genes correspondientes a proteínas relacionadas con el plegamiento de galectina. Se utilizó la información estructural del plegamiento de Gal-1 (PDBID = 1GZW) para determinar la penalización de gaps locales e incorporar información estructural al alineamiento.

Ensayo de muerte de células T. Células T Jurkat (5×10^5) fueron cultivadas y activadas como se describió anteriormente [22, 23] e incubadas con o sin Gal-1 (3 μ M) o sus variantes mutantes en medio RPMI suplementado con suero fetal bovino 5%, penicilina (100 mU/mL) y estreptomicina (50 μ g/mL) en platos de 24 pocillos en estufa a 37 °C y 5% CO₂. Para generar condiciones reductoras se agregó ME (0,55 mM concentración final) antes de resuspender las células en el medio. Para determinar el efecto de la oxidación sobre la función de las galectinas, Gal-1 y sus mutantes fueron agregadas a un medio con 10 mM de H₂O₂ en RPMI durante 20 min previos al ensayo de apoptosis. El exceso de ROS fue eliminado usando catalasa (100 U/mL). Luego, el medio se completó con suero y antibióticos, y las células se agregaron en cada pocillo de la placa de cultivo. Luego de 14 h de exposición a las diferentes variantes de Gal-1, las células se lavaron con PBS dos veces. La muerte celular se determinó por marcación con anexina V-FITC/ioduro de propidio (IP) (sección II.2.4.5) en buffer de tinción (100 mM HEPES, 1,4 M NaCl, 25 mM CaCl₂). La fluorescencia resultante (FITC y IP) se analizó en un citómetro FACSCanto. La muerte celular se calculó como el porcentaje de células anexina V-positivas en los pocillos tratados con galectina menos el porcentaje de células anexina V-positivas en los pocillos control (ME para los medios reductores, H₂O₂ seguido de tratamiento con catalasa para los medios oxidantes, sin galectina en ambos casos). Los datos fueron expresados como media $\pm$ desvío estándar y los

resultados se compararon de a dos mediante un test *t* de Student, con valores-*p* de 0,05 o menores considerados como significativos (Prism de GraphPad Software).

IV.3 Resultados

IV.3.1 Algunas cisteínas están involucradas en el proceso de oxidación de Gal-1

Una de las características más impactantes de Gal-1 es la elevada proporción de cisteínas en su secuencia (**Tabla III.3**), teniendo en cuenta que este aminoácido en particular es uno de los menos empleados para la biosíntesis proteica [24]. Existen 6 cisteínas por cada DRC de Gal-1, numeradas Cys2, Cys16, Cys42, Cys60, Cys88 y Cys130. Su distribución espacial en una representación tridimensional de la estructura terciaria de Gal-1 se muestra en la **Figura IV.1**. Estos residuos de cisteína exhiben un amplio rango de accesibilidad de solvente, por lo que es de esperar que su reactividad frente a diferentes oxidantes solubles varíe acorde a esta exposición diferencial. Para verificar esta hipótesis, se midieron los tioles libres usando la reacción de Ellman (ver sección IV.2) con el reactivo DTNB. El análisis experimental de los grupos sulfhidrilos expuestos muestra una relación molar de $3,9 \pm 0,6$, sugiriendo la presencia de aproximadamente cuatro moles de tioles libres por mol de proteína. Para identificar cuáles eran las cisteínas expuestas se calcularon dos parámetros claves para alcanzar este propósito empleando simulaciones de DM. Por un lado se calcularon las funciones de distribución radial $g(r)$ (sección II.1.3.2) de moléculas de agua alrededor de los átomos de azufre de las seis cisteínas y, por otro lado, el SASA por residuo y átomo de azufre, mostrando el perfil diferencial de exposición a solvente previamente esperado (**Figura IV.2**). Los resultados se encuentran en concordancia con lo observado previamente por Lopez-Lucendo y col. [11], en donde se emplea un tipo de cálculo diferente. Las Cys2 y Cys130 son las cisteínas más expuestas (mayor valor de SASA) y solvatadas en Gal-1 monomérica, debido a que se encuentran muy cerca de los extremos terminales de la proteína. Las Cys16 y Cys88 muestran un perfil de solvatación muy similar entre ellas, con un menor grado de exposición al solvente que Cys2 y Cys130, mientras que la ausencia de un típico perfil de solvatación (y SASA despreciable) para las Cys42 y Cys60 indican una oclusión casi total de estas cisteínas al solvente. Estos resultados sugieren que la reactividad frente a la oxidación y/o a

cualquier tipo de funcionalización de estas cisteínas es claramente diferente, siendo Cys42 y Cys60 difícilmente accesibles por reactivos solubles en agua, como previamente se sugirió [9]. Por lo tanto, en principio parece razonable asociar la química redox de Gal-1 a la presencia de Cys2, Cys16, Cys88 y Cys130, las cuatro cisteínas expuestas en la estructura de Gal-1.

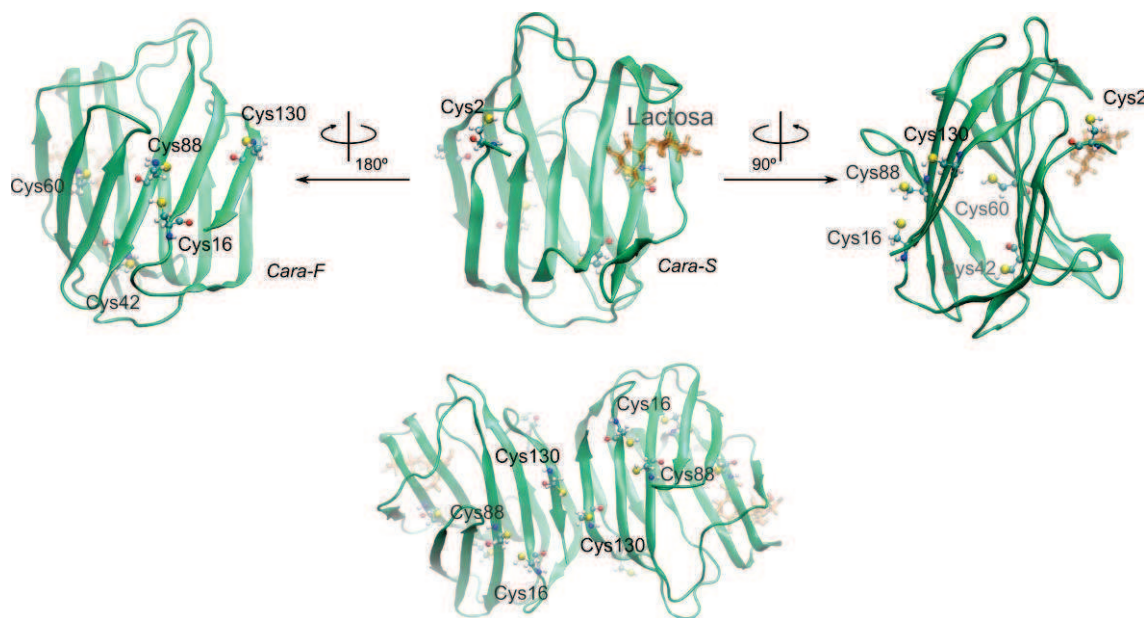


Figura IV.1. Distribución espacial de las seis cisteínas en Gal-1. Existen seis residuos de cisteína en cada monómero de Gal-1, llamados Cys2, Cys16, Cys42, Cys60, Cys88 y Cys130. Mientras que Cys2 y Cys130 están localizados en los extremos N y C terminal respectivamente, las Cys42 y Cys60 están en el interior del sándwich- β del plegamiento de Gal-1. Las Cys16 y Cys88, una muy cerca de la otra, se encuentran en la cara F, diametralmente opuestas al LBG.

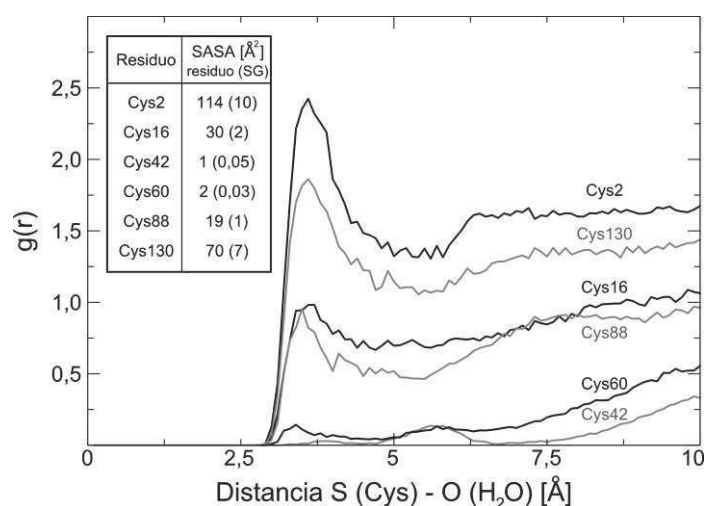


Figura IV.2. Distribución radial de las moléculas de H_2O alrededor del átomo de azufre de cada cisteína en Gal-1. La función de distribución radial $g(r)$ arroja una idea de la probabilidad de encontrar una molécula de agua a la distancia r del átomo de azufre de la cadena lateral de cada cisteína marcada en el gráfico. Es por lo tanto, una medida de la posibilidad de que un oxidante soluble pueda acercarse al átomo reactivo de azufre de una cisteína determinada. Como inserto, se muestran los valores de SASA para el residuo completo y sólo el átomo de azufre de cada cisteína. Con ambos cálculos se observa que, en orden decreciente, la exposición al solvente de las cisteínas en Gal-1 es Cys2 > Cys130 > Cys16 $\approx$ Cys88 $\gg$ Cys60 $\approx$ Cys42.

Usando PROPKA y los resultados provenientes de la DM de Gal-1 wild-type, se calcularon los valores de pKa para las seis cisteínas, y también para la mutante C2S como un control para las mutantes (**Tabla IV.2**).

Aminoácido	ΔpK_a	
	Gal-1 wild-type	Gal-1 C2S
Cys2	$-0,34 \pm 0,82$	-
Cys16	$-1,02 \pm 0,79$	$-0,80 \pm 0,63$
Cys42	$0,66 \pm 0,43$	$0,83 \pm 0,40$
Cys60	$0,84 \pm 0,68$	$0,67 \pm 0,49$
Cys88	$-0,26 \pm 0,27$	$-0,20 \pm 0,56$
Cys130	$-0,43 \pm 0,48$	$-0,52 \pm 0,26$

Tabla IV.2 Estimación del cambio de pKa en las cisteínas de Gal-1 wild-type y su mutante C2S.

Las cisteínas que presentan el cambio más grande en pKa con respecto al valor del aminoácido libre (en solución acuosa) son Cys16, Cys42 y Cys60. Como Cys42 y Cys60 se encuentran enterradas en el interior de la proteína, es esperable una perturbación significativa en el valor de pKa del grupo lateral -SH con respecto al mismo completamente libre en agua. Un valor positivo en esta diferencia confirmó que estas cisteínas se encuentran excluidas del solvente y rodeadas de residuos hidrofóbicos. Por otro lado, Cys16 mostró el mayor cambio (con signo negativo), resultando ser una cisteína más ácida que la cisteína libre. Dado que la oxidación de cisteínas generalmente involucra el ataque nucleofílico de la especie desprotonada (tiolato) al oxidante, el valor bajo de pKa calculado para Cys16 sugiere que, entre todas las cisteínas expuestas, este residuo en particular podría controlar el inicio del proceso de oxidación a pH fisiológico. Sin embargo, es bien conocido que otros factores además del pKa pueden afectar la reactividad de las cisteínas [25], siendo éste uno de los más importantes. Finalmente, a pesar de ser levemente más acídicas, las otras cisteínas expuestas muestran un valor de pKa bastante similar al de la cisteína libre.

IV.3.2 La oxidación y reducción de Gal-1 es un proceso reversible

Cuando se expone una solución de Gal-1 5 μ M en PBS a una oxidación al aire (5 días) o por tratamiento con 30 mM de H₂O₂ durante 30 min, se obtienen espectros de DC en el UV-lejano similares para ambos tratamientos (**Figura IV.3A**), los cuales resultaron ser claramente diferentes del correspondiente espectro para Gal-1 reducida.

Este cambio ocurrió en intensidad y en posición del mínimo de elipticidad, indicando un cambio conformacional hacia un plegamiento alternativo. Los resultados de deconvolución de estos espectros mediante Dichro Web muestran un cambio en contenido de hebra- β de 43% para Gal-1 reducida a 39% luego de la oxidación. Estos cambios se vieron acompañados por leve aumento en la cantidad de loops (de 11,5% a 13%). Resultados similares se han encontrado para esta galectina [26] y para galectinas de otras especies [2-4], confirmando resultados previos y la presencia de un fenómeno de cambio de plegamiento frente al cambio de estado redox.

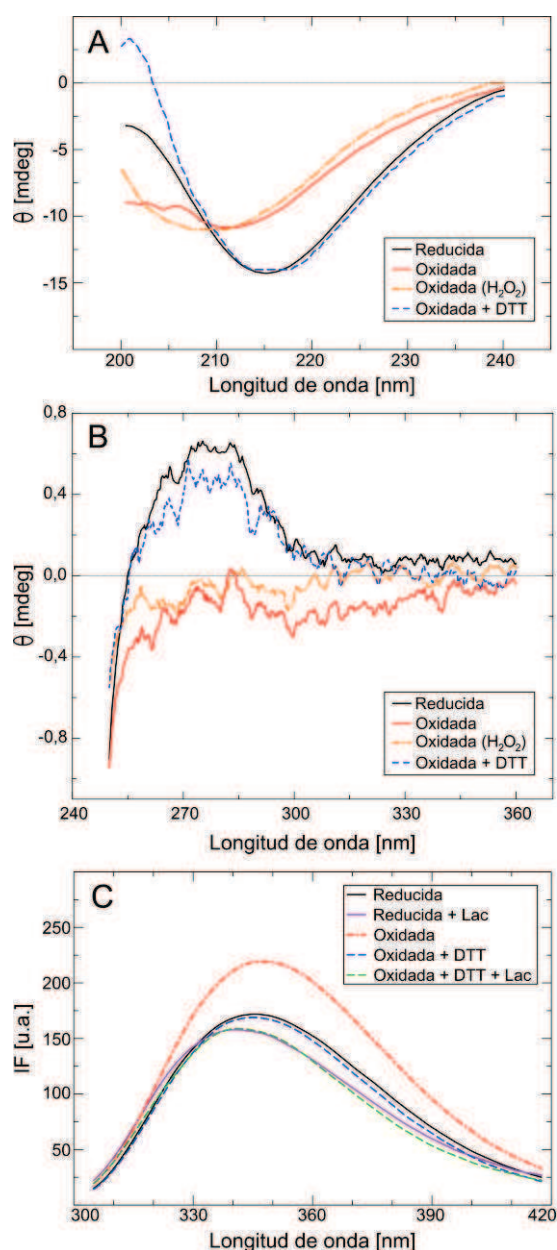


Figura IV.3. Reversibilidad estructural de Gal-1 frente a los cambios redox del medio. Espectros de DC en (A) el UV-lejano y (B) el UV-cercano. (C) Espectros de emisión de fluorescencia de Gal-1 en diferentes situaciones. La referencia a Gal-1 oxidada, corresponde a la oxidación por aire.

Pero lo más interesante fue que cuando la proteína oxidada al aire fue tratada con 2 mM de DTT durante 15 min, se obtuvo un espectro muy similar al de Gal-1 reducida (**Figura IV.3A**), sugiriendo que la oxidación es un proceso capaz de revertirse. Aún más, cuando un exceso de H_2O_2 se agregaba nuevamente a las muestras, se obtenía el espectro de la proteína oxidada. En la misma sintonía, los espectros de DC en el UV-cercano de la proteína oxidada mostraron un achatamiento de las bandas positivas alrededor de los 280 nm, sugiriendo la pérdida de un plegamiento nativo alrededor de los residuos aromáticos (**Figura IV.3B**). Dado que estos últimos experimentos requieren una concentración de proteína mucho más grande (sección II.2.4.3) (70 μ M, en particular), el tratamiento de oxidación indujo agregación de Gal-1. Por esta razón, luego del agregado de DTT a la muestra oxidada, no se pudo lograr la recuperación total de la intensidad de señal del espectro de Gal-1 reducida, aunque la forma es prácticamente la misma, indicando nuevamente que la proteína oxidada que permaneció en solución efectivamente había logrado una reversión a su forma reducida. Finalmente, cuando se realiza un análisis similar utilizando espectroscopía de fluorescencia (**Figura IV.3C**), se encontraron espectros de emisión idénticos cuando la proteína se redujo luego de ser oxidada al aire, incluso en presencia de lactosa. A la vista de los resultados obtenidos, es posible indicar que tanto la estructura secundaria como la terciaria de la forma reducida de Gal-1, así como también la capacidad de unión a lactosa, se pueden recuperar luego de la reducción de la proteína oxidada.

IV.3.3 Las Cys2, 16 y 88 juegan un rol clave

Para lograr una descripción más detallada de la contribución de cada cisteína al proceso de oxidación, se produjeron en forma recombinante las seis mutantes simples de cisteína por serina y un par de dobles mutantes seleccionadas. Estas proteínas mutantes fueron sometidas a los mismos protocolos de reducción y oxidación previamente usados para Gal-1 wild-type. Los cambios conformacionales fueron seguidos mediante el registro de los espectros de DC en el UV-lejano (**Figura IV.4**). De manera similar al comportamiento de Gal-1, la mutación de las cisteínas internas (Cys42 y Cys60) no previno que la proteína alcanzara su conformación oxidada. Un resultado idéntico se observó para la mutante C130S, indicando que esta cisteína expuesta al solvente no es esencial para que se lleven a cabo los cambios conformacionales promovidos por la oxidación. Por el contrario, las proteínas mutadas

en Cys16, Cys88 y sorprendentemente Cys2 no mostraron prácticamente ningún cambio en su espectro de DC luego de la oxidación por aire. Sin embargo, la mutante C2S sufrió cierto cambio conformacional cuando la oxidación se realizó mediante el uso de H_2O_2 . Por otro lado, no se observaron cambios significativos en las dobles mutantes C2SC16S y C2SC88S, debido a la ausencia de las cisteínas más críticas para que se disparen los cambios conformacionales antes mencionados.

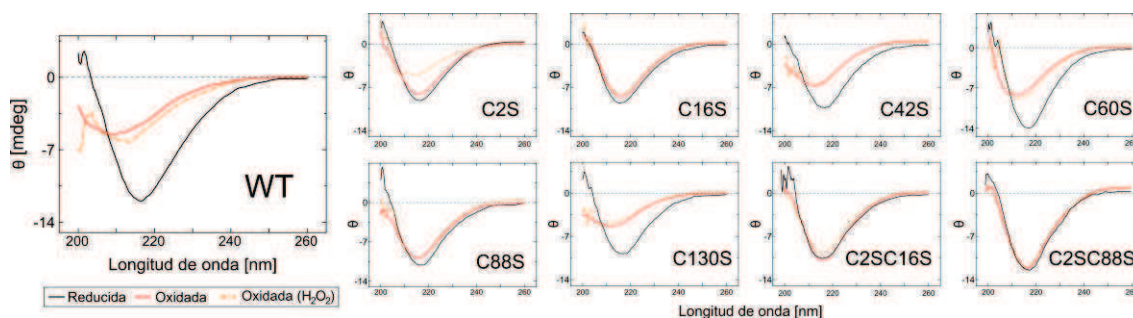


Figura IV.4. Espectros de DC de Gal-1 y sus mutantes CXS en diferentes ambientes redox. Todas las estructuras reducidas de Gal-1 y las mutantes CXS (línea negra continua) son muy similares. Pero cuando se oxidan al aire (línea roja continua) o con H_2O_2 (línea punteada naranja), diferentes espectros se obtienen, dependiendo de la cisteína mutada.

Basándonos en estos resultados, resulta difícil comprender por qué la ausencia de algunas cisteínas en particular previene la transición entre distintas conformaciones, dado que se ha propuesto que el estado oxidado final requiere la presencia de las seis cisteínas para formar tres enlaces disulfuro determinados [10]. Aunque puedan estar ocurriendo cambios que no influyen directamente el espectro de DC, en las condiciones experimentales aquí evaluadas sólo las cisteínas Cys16 y Cys 88 (y Cys2 si la oxidación se lleva a cabo en aire) parecerían ser esenciales para inducir el cambio conformacional que inactiva la proteína, ya que su ausencia en la estructura previene las alteraciones estructurales evidenciadas por DC.

IV.3.4 Oxidación de Gal-1 con H_2O_2

Cuando Gal-1 reducida (recordemos, 4 tioles/proteína) se expone a concentraciones crecientes de H_2O_2 , se observa un perfil de oxidación bifásico en función del tiempo (**Figura IV.5**). Se ajustaron curvas bi-exponenciales a los datos experimentales ($n = 3$) a cada concentración de peróxido estudiada, donde la amplitud de cada caída exponencial coincidió aproximadamente con el consumo de dos tioles. A

partir de la pendiente del gráfico de la constante de velocidad observada (k_{obs}) para la primera exponencial en función de la concentración de H_2O_2 , se pudo recuperar un valor de constante específica de velocidad para el tiol más reactivo en Gal-1 de $1,7 \pm 0,2 \text{ M}^{-1} \text{ s}^{-1}$ a pH 7,4 y 25 °C en PBS (Tabla IV.3). En presencia de lactosa, este valor ha sido levemente superior pero aún dentro del error del método. A pH 6,8 la constante de velocidad fue aproximadamente seis veces menor, lo que es esperable para este tipo de reacciones donde el tiolato es la especie reactiva. Dado que la segunda exponencial corresponde a un proceso que no está completamente resuelto en el tiempo, el valor de la constante obtenida a partir de estos datos experimentales posee mucho error y solo puede ser estimada, del orden de $10^{-3} \text{ M}^{-1} \text{ s}^{-1}$ a pH 7,4 y 25 °C.

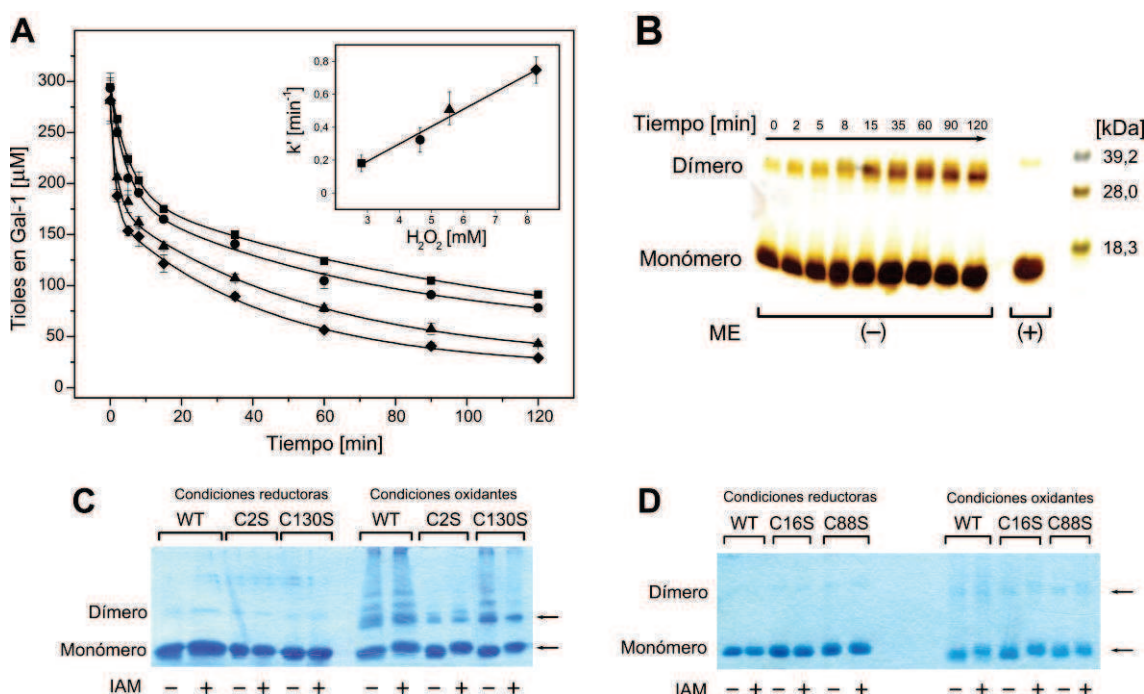


Figura IV.5. Cinética de oxidación de Gal-1 con H_2O_2 . (A) Una solución de Gal-1 (67 μM) se incubó con H_2O_2 2,81 mM (cuadrados), 4,65 mM (círculos), 5,56 mM (triángulos) y 8,28 mM (diamantes) en buffer PBS y a 25 °C. Se tomaron alicuotas del medio de reacción a diferentes tiempos y la reacción se detuvo con el agregado de catalasa, en hielo. Se ajustó una función biexponencial para la determinación de las pseudo-constantes de velocidad de consumo de tioles libres. Las mismas graficadas en función de la concentración de oxidante (inserto) permiten hallar el valor de la constante de velocidad para el tiol más reactivo en Gal-1. (B) Análisis electroforético de las alicuotas removidas de la reacción de Gal-1 con 10 mM de H_2O_2 . Las primeras nueve calles contienen buffer no-reductor con SDS, mientras que la calle once corresponde a la última alicuota ($t = 120 \text{ min}$), con buffer reductor (ME) con SDS. (C) Gal-1, C2S y C130S y (D) Gal-1, C16S y C88S fueron sometidas a oxidación con 10 mM H_2O_2 por 2 h. Se realizaron geles reductores y no reductores luego de cada tratamiento. Iodoacetamida (IAM) se agregó luego de la oxidación para analizar en profundidad el efecto del *crackeo* de las muestras sobre los tioles que no reaccionaron durante la oxidación con peróxido.

Cuando las tres mutantes CXS más relevantes se sometieron a las mismas mediciones, todas las constantes fueron menores que la obtenida para Gal-1 wild-type (**Tabla IV.3**). El valor de la constante de velocidad para la mutante C2S provee el valor real de consumo de tioles para la formación del enlace disulfuro Cys16-Cys88, resultando en un valor de $0,4 \pm 0,1 \text{ M}^{-1} \text{ s}^{-1}$.

Proteína	Condición (T = 25 °C en PBS)	k [M ⁻¹ s ⁻¹]
Gal-1	pH 7,4	1,7 ± 0,2
	pH 7,4 + lactosa 100 mM	2,1 ± 0,2
	pH 6,8	0,3 ± 0,1
C2S		0,4 ± 0,1
C16S	pH 7,4	0,8 ± 0,1
C88S		1,0 ± 0,1

Tabla IV.3 Cinética de consumo de tioles por oxidación con H₂O₂ en Gal-1 y mutantes seleccionadas.

Con respecto a la naturaleza del primer enlace disulfuro formado, dada nuestra evidencia la posibilidad más razonable es el par Cys16-Cys88 debido a las razones anteriormente mencionadas, y además porque cualquier otra reacción de Cys2 con alguna de las cisteínas provenientes de otro monómero de Gal-1 generaría dímeros (puentes intermoleculares). Aunque los dímeros efectivamente se forman durante la reacción de Gal-1 con H₂O₂ (**Figura IV.5B y C**), la cinética de formación de dímeros no coincide con la primera fase de consumo de tioles, la cual posee un tiempo de vida media de 40 s a 10 mM de H₂O₂. Por el contrario, la formación de dímeros no se logró completar en 2 h (**Figura IV.5B**). Además, la banda correspondiente al monómero en aquellas muestras sometidas a oxidación migra más que la misma banda correspondiente a la proteína reducida, debido principalmente a la formación de un puente disulfuro que ocurre en el monómero (**Figura IV.5B**, calles 9 y 11). Esta observación fue corroborada también al analizar la movilidad electroforética de las mutantes C16S y C88S (**Figura IV.5D**). Bajo condiciones no reductoras, ambas mutantes muestran la misma movilidad (la cual es ligeramente menor a la observada para Gal-1 oxidada), apoyando la hipótesis de que las Cys16 y Cys88 son las responsables del enlace disulfuro intramolecular. Cuando realizamos ensayos de oxidación similares con las mutantes C2S y C130S (**Figura IV.5C**), se pudo observar la formación de dímeros no sólo en la proteína wild-type (**Figura IV.5B**) sino también en ambas mutantes. Sorprendentemente, los dímeros fueron reducidos por tratamiento con ME y se formaron inclusive bajo el agregado de iodoacetamida (IAM) luego del

tratamiento con H_2O_2 (para evitar la oxidación de tioles libres durante el craqueo de la proteína, previo a la siembra del gel de electroforesis). Sin embargo, la banda correspondiente al dímero fue más evidente en la calle de la mutante C130S, sugiriendo que Cys2 es la responsable principal para formación de estos dímeros y también para los multímeros de mayor peso molecular.

Para obtener una estimación de la velocidad del cambio conformacional experimentado por Gal-1 durante su oxidación por H_2O_2 , se decidió realizar experimentos siguiendo los cambios en la señal de fluorescencia intrínseca a lo largo de la oxidación. Registramos la intensidad de fluorescencia a 345 nm para una solución 7 μM de Gal-1 en función del tiempo a diferentes concentraciones de H_2O_2 (Figura IV.6).

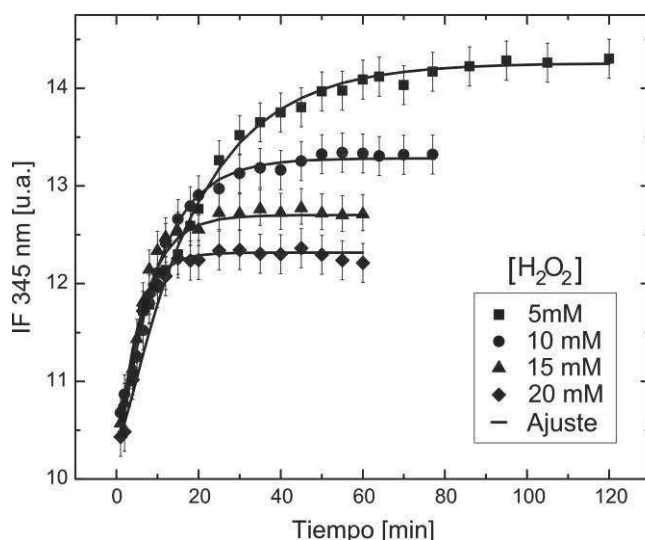


Figura IV.6. Cinéticas del cambio conformacional que sufre Gal-1 por oxidación con H_2O_2 . La intensidad de fluorescencia se registró a 345 nm en función del tiempo y a distintas concentraciones de H_2O_2 (ver gráfico). Se propuso un modelo cinético (sección IV.2) que explicara esta tendencia en forma global y se ajustó a los datos (línea).

Como se destacó anteriormente, la intensidad de fluorescencia aumenta cuando Gal-1 se oxida (Figura IV.3C). Sin embargo, algo más estaba sucediendo al analizar la variación de la fluorescencia con el cambio de concentración de peróxido de hidrógeno: la fluorescencia final disminuía a medida que aumentaba la concentración de oxidante. Se ha demostrado que el cambio conformacional es independiente de la concentración de H_2O_2 una vez que la cisteína alcanza el intermediario de ácido sulfénico ($\text{Cys-SH} + \text{H}_2\text{O}_2 \rightarrow \text{Cys-SOH} + \text{H}_2\text{O}$) [25], por lo que no se esperaba esta tendencia en los espectros de fluorescencia. Como el peróxido de hidrógeno puede inducir no sólo la oxidación del grupo tiol a ácido sulfénico sino también la oxidación del ácido sulfénico

a ácido sulfínico (proceso que se conoce como sobreoxidación del tiol), hemos propuesto un modelo que incluye el proceso de sobreoxidación de la cisteína reactiva, ofreciendo una explicación al comportamiento de la señal y, además, la estimación de la constante de velocidad para el cambio conformacional. Esta reacción compite con la formación del enlace disulfuro, inhibiendo así el cambio conformacional asociado (ver modelo cinético descrito en la sección IV.2). Luego de ajustar el modelo con los datos, además de verificar que el mecanismo de reacción propuesto logró explicar el valor decreciente del *plateau* a medida que aumenta la concentración de peróxido de hidrógeno, se pudieron obtener una constante de velocidad para el cambio conformacional (independiente de H_2O_2) de $0,003 \pm 0,001 \text{ s}^{-1}$ y una constante de velocidad para la sobreoxidación de $0,3 \pm 0,1 \text{ M}^{-1} \text{ s}^{-1}$, comparable con la de consumo de tiol. A bajas concentraciones de H_2O_2 , la velocidad inicial de cambio conformacional resultó más grande que la de sobreoxidación (en 5 mM de H_2O_2 , el cambio conformacional es el doble de rápido). Por otro lado, a concentraciones elevadas de H_2O_2 , la oxidación del intermediario sulfénico puede competir con la reacción del cambio conformacional. Es importante destacar que si bien este es un resultado biofísico razonable, carece de relevancia biológica, dado que estas concentraciones elevadas de H_2O_2 raramente se dan en microambientes fisiológicos, ni siquiera en estados de estrés oxidativo patológicos o desregulados.

IV.3.5 Gal-1 oxidada no une lactosa

Monitoreando la intensidad de fluorescencia intrínseca de Gal-1 a 363 nm (una longitud de onda a la cual la diferencia de intensidad entre la proteína libre y la unida es máxima) en función de la concentración de lactosa fue posible, mediante el ajuste de un modelo sencillo de un sitio de unión, determinar una constante de unión (K_b) de $4,7 \times 10^3 \text{ M}^{-1}$ (**Figura IV.7A**). Cuando se realiza el mismo ensayo con Gal-1 previamente oxidada al aire o con 10 mM de H_2O_2 , el espectro de fluorescencia no se vio modificado a ninguna de las concentraciones de lactosa ensayadas (**Figura IV.7A**). Por lo tanto, la proteína oxidada es incapaz de unir lactosa, independientemente del oxidante utilizado. Para ahondar en las bases moleculares de este efecto, se realizó un ensayo termodinámico computacional del proceso de unión a lactosa, comparando la capacidad de unión de la proteína reducida (Gal-1 wild-type) con la de la proteína oxidada (Gal-1 con un puente disulfuro entre Cys16 y Cys88). Aunque durante los 80

ns de DM la molécula de lactosa se mantuvo en el LBG de Gal-1 oxidada, el $\Delta G_{\text{unión}}$ (calculado con el método MM/QM-COSMO, sección IV.2) aumenta a medida que el tiempo de simulación corre, indicando el inicio del proceso de disociación, hecho consistente con la unión desfavorable de lactosa a Gal-1 con el enlace disulfuro Cys16-Cys88 presente en su estructura [$\Delta G_{\text{unión}}$ Gal-1 reducida = -8,18 kcal mol⁻¹ vs $\Delta G_{\text{unión}}$ Gal-1 oxidada = -4,19 kcal mol⁻¹]. Esta tendencia indica que el azúcar se une con mayor afinidad a la forma reducida, tal como se evidenció en los resultados experimentales. El hecho de que no se pudo observar una disociación completa del ligando de la proteína puede deberse a que la DM requería mayor tiempo de simulación para converger estadísticamente y/o tal vez estos resultados de energía libre de unión, basados en un modelo implícito de solvatación, están sobreestimados y sólo se puede analizar la tendencia comparativa, como se realizó en este apartado y en el capítulo anterior.

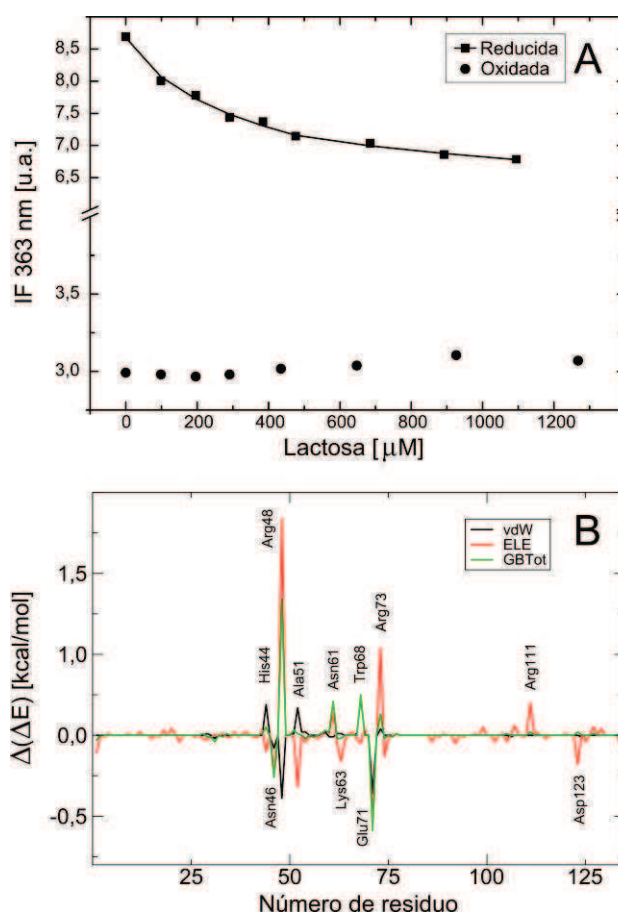


Figura IV.7. Análisis termodinámico de la unión de lactosa a Gal-1. (A) Intensidad de fluorescencia a 363 nm en función de la concentración de lactosa. Se ajustó un modelo simple de unión y de allí se estimó la constante de unión (K_b). No se observó ningún cambio en el espectro de emisión con el agregado de lactosa en el caso en que Gal-1 se encontraba previamente oxidada por H_2O_2 . (B) Diferencias de $\Delta E_{\text{unión}}$ ($\Delta(\Delta E)$) por residuo (MM/GBSA) entre los sistemas Gal-1_{oxidada}:lactosa y Gal-1_{reducida}:lactosa. Se marcan las contribuciones de vdW (negro), electrostáticas (rojo) y el $\Delta G_{\text{unión}}$ (verde) del proceso total.

Utilizando el método de descomposición por aminoácido implementado en la rutina de MM/GBSA, es posible determinar cuáles son los aminoácidos más afectados por la formación del enlace disulfuro. A pesar de que el puente Cys16-Cys88 se encuentra en la cara-*F* de Gal-1, los aminoácidos que más afectan la unión y que resultan sensibles a estos cambios se encuentran efectivamente en el LBG, en la cara-*S* de la proteína (**Figura IV.7B**). Arg48 y Arg73 (y Asn61 y Arg111 en menor medida) fueron los aminoácidos que incrementaron la componente electrostática a favor de la desestabilización de la unión con la oxidación, mientras que Asn46, Ala51, Lys63, Glu71 y Asp123 promovieron el efecto opuesto. Además, His44, Ala51 y Trp68 aumentaron la componente energética de van der Waals (vdW) por ruptura de las interacciones dispersivas con el ligando. Finalmente, Arg48 y Glu71 contribuyeron a la unión en la proteína oxidada a través de la reducción de la contribución energética de vdW, donde la energía total (suma de todas las componentes) indicó una contribución global a la unión en la proteína oxidada. Esta tendencia se vio reflejada a nivel de la distancia entre los aminoácidos y el ligando. Cuando Gal-1 se oxida, las distancias del O3 de la glucosa a Glu71(CD) y Arg48(NH2) aumentaron 0,1 Å y 0,3 Å en los 80 ns de simulación, respectivamente. Por otro lado, Arg48(NH1) se aproximó al O3 del monosacrido galactosa unos 0,1 Å, reflejando así un desplazamiento del ligando significativo del LBG hacia el solvente. Este tipo de análisis permitió la discriminación de la contribución de cada aminoácido a la unión del ligando a Gal-1 luego de la oxidación por formación del puente disulfuro Cys16-Cys88.

IV.3.6 Gal-1 oxidada no modifica la viabilidad de células T

Se ha reportado que Gal-1 es capaz de alterar selectivamente la viabilidad de células T activadas bajo condiciones reductoras [22, 27]. Este efecto involucra la unión a determinados ligandos glicosilados en la superficie de las células T (particularmente CD45 y CD43) y una presencia predominante de la lectina en su forma dimérica. Por ende, hemos estudiado también el impacto de la oxidación de Gal-1 en la estructura y función en el contexto celular, usando un ensayo de apoptosis (ver sección II.2.4.5). Exponiendo células T activadas (línea celular Jurkat) a diferentes concentraciones de Gal-1 en condiciones reductoras (en presencia de 0,55 mM de ME), se ha determinado una concentración óptima de 3 µM de Gal-1 para trabajar en los ensayos posteriores (**Figura IV.8**). Ésta concentración permitió analizar niveles óptimos de apoptosis que

arrojen diferencias significativas en relación a la muerte basal por ausencia de tratamiento en el ensayo control. Este valor es muy cercano al reportado con anterioridad [22].

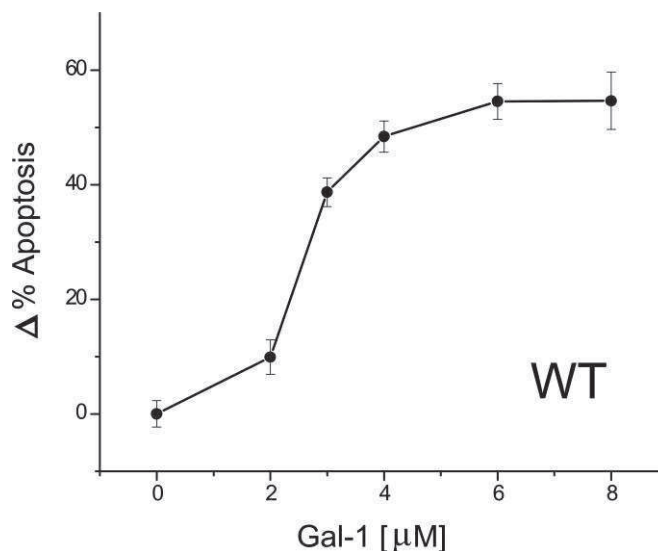


Figura IV.8. Curva dosis-respuesta para el efecto apoptótico de Gal-1 sobre células T. Puesta a punto de la concentración óptima elegida (3 μM) para los ensayos posteriores.

En contraste con los efectos observados bajo condiciones reductoras, encontramos que la oxidación de Gal-1 wild-type y las mutantes C42S, C60S y C130S perdieron gradualmente su actividad apoptótica (**Figura IV.9**). Por el contrario, C2S, C16S, C88S y las doble mutantes C2SC16S y C2SC88S no indujeron cambios significativos en la viabilidad de las células T independientemente del ambiente redox a las que se las sometió. Estos resultados confirman, a nivel *in vivo*, la relevancia de las cisteínas Cys2, Cys16 y Cys88 en la inactivación de Gal-1 y refuerzan la hipótesis de que sólo algunas cisteínas son esenciales en la regulación de la actividad de Gal-1. De forma interesante, las doble mutantes se sobrepusieron exitosamente al desafío oxidativo y fueron capaces de disparar el mismo grado de apoptosis que la proteína wild-type reducida. Así, hemos logrado el diseño de dos mutantes que son sustancialmente más resistentes a la inactivación oxidativa que las mutantes simples, teniendo este resultado una relevancia terapéutica importante, pues la oxidación ocurre en ambientes inflamatorios en donde se requiera el correcto accionar de esta proteína [28, 29].

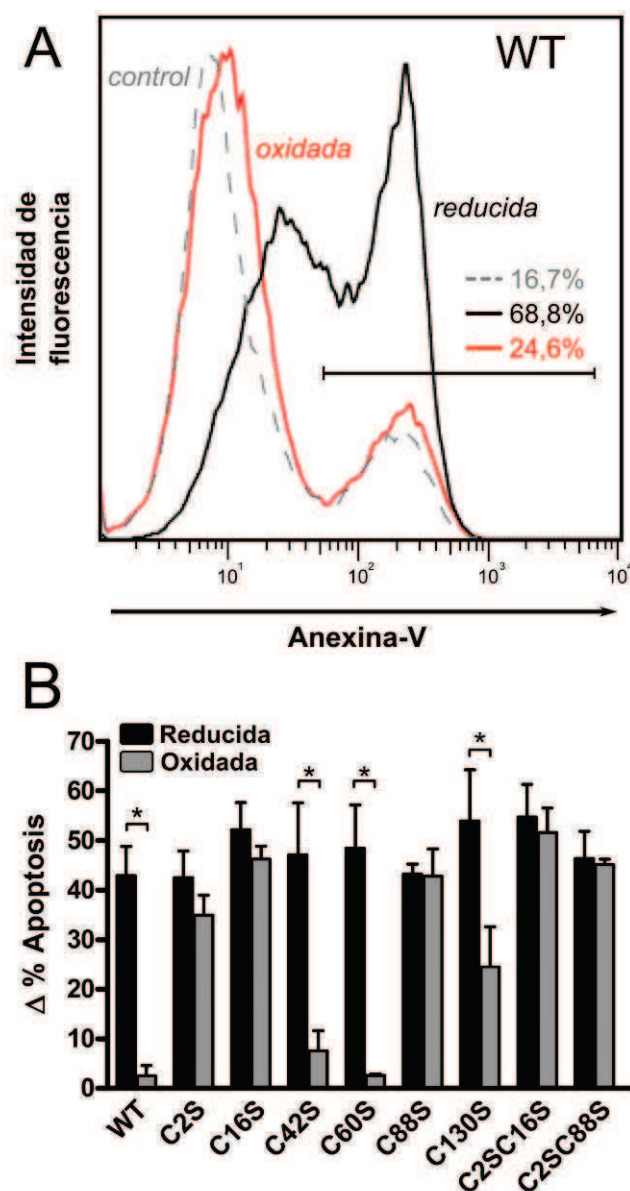


Figura IV.9. Efecto de Gal-1 y sus mutantes CXS en la muerte de células T. (A) Histograma representativo para Gal-1 wild-type, donde se muestra el perfil de fluorescencia (escala logarítmica) de células T marcadas con anexina V-FITC (control, línea gris rayada; Gal-1 reducida, línea negra; Gal-1 oxidada, línea roja). (B) ΔApoptosis (%) observado para cada Gal-1 recombinante ensayada (wild-type, y las mutantes simples CXS y dos dobles mutantes). En negro se muestra la proteína reducida y en gris, la proteína previamente oxidada. Puesta a punto de la concentración óptima elegida (3 μ M) para los ensayos posteriores. Los resultados mostrados son representativos de un mínimo de tres experimentos independientes (media $\pm$ SD; * $p < 0,05$).

IV.3.7 Gal-1, cisteínas y sus hermanas

Finalmente, para lograr una imagen más integrada de estos resultados, hemos evaluado la relevancia de la presencia de estas cisteínas en los otros miembros de la familia de galectinas humanas. La hipótesis es que, dado que la estructura

tridimensional es extremadamente similar entre los miembros de la familia (Capítulo III), aquellas galectinas que presenten las cisteínas más relevantes (según los resultados de este Capítulo) en la posición y ambiente adecuado, es posible que exhiban el mismo perfil y comportamiento estructural y funcional que el demostrado para Gal-1. Para ello, realizamos un alineamiento de secuencia con el programa Clustal-W, incorporando información estructural para que los gaps coincidan con los loops (Figura IV.10).

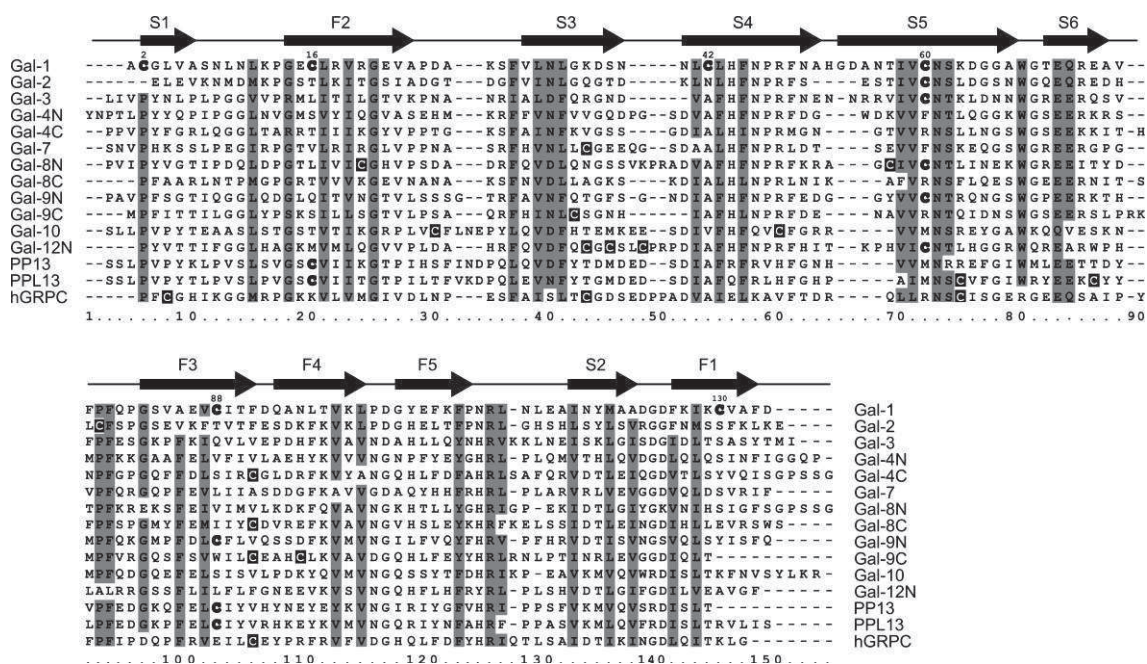


Figura IV.10. Alineamiento múltiple de todas las galectinas humanas. Los esquemas correspondientes a la estructura secundaria y la numeración en el borde superior corresponden a Gal-1. Las láminas- β están denotadas como flechas y los loops que las conectan como líneas. Las cisteínas compartidas con Gal-1 están resaltadas en negrita, mientras que las otras cisteínas se resaltaron con letra blanca y fondo negro. Los residuos conservados (~80%) se marcaron en gris.

Este alineamiento revela que (i) la cisteína más conservada es Cys60; (ii) la cisteína más reemplazada es Cys42 (mayoritariamente por Ala); (iii) los únicos miembros de la familia de galectinas que contienen las Cys16 y Cys88 son PPL13 y PPL13. Aquellas que poseen Cys88 solamente son Gal-3 y Gal-9N. Así, la ocurrencia de Cys60 (cisteína poco relevante para el proceso de oxidación) en esta familia puede jugar un rol en la estabilidad de plegado o estructural, ya que está presente y conservada en casi todos los miembros. Lo mismo puede concluirse para Cys42 pero en dirección opuesta: su rol en la oxidación no parece ser crucial ya que es la más sustituida en toda la familia.

IV.4 Discusión y conclusiones

Aunque la oxidación se creía parte de un proceso de inactivación responsable de la disminución o eliminación completa de la función de galectinas, estudios recientes sugieren que Gal-1 oxidada puede presentar funciones alternativas independientes del reconocimiento de glicanos, como ser la regeneración de nervios periféricos [26]. Durante la oxidación, se ha propuesto que cada subunidad de Gal-1 forma tres puentes disulfuro intramoleculares que desencadenan un cambio conformacional profundo, afectando la dimerización y la unión a glicanos [26, 30]. La estabilidad frente a condiciones oxidantes aumenta considerablemente con las mutaciones de las cisteínas (por serina, generalmente), mientras que la especificidad y afinidad por los ligandos naturales de Gal-1 no se ve afectada [6, 31]. En este contexto, se ha demostrado que la unión a ligandos protege a Gal-1 de la oxidación [30]. Stowell y col. encontraron que la unión a ligandos específicos controla la sensibilidad de Gal-1 a oxidarse, modificando la posición del equilibrio de dimerización hacia mayor componente dimérica, estructura cuaternaria que parecería ser más refractiva a la oxidación. En apoyo a estos resultados, una forma mutante de Gal-1 que exhibe un impedimento hacia la dimerización mostró elevada sensibilidad hacia la oxidación, y falló además en la correcta inducción de apoptosis, actividad biológica comúnmente desarrollada por Gal-1 reducida [30]. Teniendo en cuenta estos resultados previos, y usando una combinación de experimentos *in vitro* e *in silico*, hemos profundizado en el estudio de los mecanismos moleculares que llevan a la oxido-reducción de Gal-1. Hemos podido establecer un orden jerárquico de reactividades e importancia estructural y funcional de cada cisteína y caracterizar la cinética de oxidación con H_2O_2 , arrojando una visión atomística del proceso con la idea de integrar toda la información estructural disponible. El primer resultado sorprendente fue el alto grado de reversibilidad del proceso reducción-oxidación. La estructura secundaria seguida por DC y por espectros de fluorescencia reflejaron el potencial de Gal-1 para moverse en forma muy plástica entre dos estructuras, un balance dictaminado por el potencial redox del ambiente (**Figura IV.3**). Dado que sólo cuatro de los seis tioles presentes en Gal-1 están expuestos al solvente, hemos postulado que los residuos responsables de disparar estos cambios conformacionales se encuentran entre estas cuatro cisteínas particulares. Encontramos que sólo Cys2, Cys16 y Cys88 son importantes para este proceso, dado que la mutante C130S muestra un comportamiento similar al observado

para la proteína wild-type frente a la oxidación (**Figura IV.4**). Es más, dada su proximidad y la particular acidez de uno de estos residuos (**Tabla IV.2**), Cys16 y Cys88 son buenos candidatos para formar entre ellos un enlace disulfuro. Estos resultados están apoyados por otras evidencias experimentales [10]. Ahora bien, la formación de dos o tres enlaces disulfuro involucrando la totalidad de las cisteínas en Gal-1 ha sido reportado (Cys2-Cys16 y Cys42-Cys60 en [32] y Cys2-Cys130, Cys16-Cys88 y Cys42-Cys60 en [10, 26]), y son resultados que no se han podido evidenciar con nuestro *set up* experimental. En nuestros ensayos, la evidencia de un cambio conformacional por oxidación cuando se mutan Cys42, Cys60 o Cys130 indica una relevancia nula de estas cisteínas en el proceso global de oxidación (**Figura IV.4**).

El rol crítico de Cys2 en la oxidación y estabilidad estructural de Gal-1 fue establecido inicialmente por el trabajo pionero de Hirabayashi y Kasai [6]. Nuestros resultados apoyan la idea de que la presencia de este residuo en particular es esencial para la estabilidad conformacional frente a la oxidación *in vitro*. Cuando se reemplaza Cys2 por un residuo serina, se espera algún tipo de perturbación de las interacciones de este aminoácido con los que lo circundan, afectando así, por ejemplo, la reactividad de las otras cisteínas y explicando el cambio de estabilidad frente a la oxidación al aire, que es lo que se evidencia (**Figura IV.4**). Sin embargo, las simulaciones de DM y los cálculos de pKa (**Tabla IV.2**) muestran que esta mutación no alteró significativamente la estructura y energética de las cisteínas reactivas. Dado que la mutante C2S fue efectivamente oxidada en presencia de H₂O₂, justificamos esta discrepancia con el oxidante actuante en el protocolo de oxidación al aire debido a una diferencia en la concentración de oxidantes en solución (menos oxidantes al aire que las concentraciones empleadas de agua oxigenada). Por supuesto, otras posibilidades deben ser exploradas para definitivamente dar con una explicación concreta de las diferencias entre Gal-1 wild-type y la mutante C2S.

La cuantificación del consumo de tioles en función del tiempo por la presencia de H₂O₂ fue utilizado para seguir la cinética de oxidación de Gal-1 (**Figura IV.5**). Gal-1 reacciona con H₂O₂ siguiendo una cinética bifásica, de pseudo-primer orden en oxidante, mostrando una constante de velocidad específica de $1,7 \pm 0,2 \text{ M}^{-1} \text{ s}^{-1}$, un valor muy similar a la constante de oxidación de cisteína libre en solución ($2,9 \pm 0,1 \text{ M}^{-1} \text{ s}^{-1}$ a pH 7,4) [33] o de cisteínas en otras proteínas como albúmina sérica humana ($2,26 \text{ M}^{-1} \text{ s}^{-1}$ a pH 7,4) [18] o tioredoxina ($1,05 \text{ M}^{-1} \text{ s}^{-1}$) [34]. Este valor es realmente bajo en comparación con las constantes registradas para las proteínas especializadas en la

reducción de hidroperóxidos (mediante óxido-reducción de tioles), hecho asociado a su plegamiento particular que les otorga esta habilidad catalítica increíblemente rápida ($10^5 - 10^8 \text{ M}^{-1} \text{ s}^{-1}$). En presencia de lactosa, el valor de la constante de velocidad no se vio afectado significativamente. En este contexto, no hemos podido observar las diferencias que sugieren los estudios previos [30] donde la unión a ligando induce una capacidad protectora frente a la oxidación. Esto se debió simplemente a una cuestión de la puesta a punto experimental, en donde para lograr señales medibles y de buena calidad Señal/Ruido de la absorbancia del TNB, debimos trabajar con concentraciones de Gal-1 mucho mayores que el valor de la constante de dimerización ($K_d 7 \mu\text{M}$) [35]. Por lo tanto, no se pudieron observar cambios en presencia de lactosa. Por otro lado, la reversibilidad del proceso de oxidación sugiere que la oxidación puede llegar a representar un factor transiente que limite la actividad biológica de la proteína en microambientes inflamatorios, donde el riesgo de oxidación es alto, pero que su función puede recuperarse cuando el estímulo oxidativo se elimina o atenúa. Así, el efecto biológico observado con la forma oxidada de la proteína puede ser debido también en parte, al balance entre estas dos formas redox de Gal-1.

La baja (en comparación) constante de velocidad medida con H_2O_2 sugiere que la oxidación de Gal-1 *in vivo* puede estar asociada con alguna actividad enzimática catalítica o estrictamente controlada por la célula y su entorno. Para que sean biológicamente relevantes en un circuito regulatorio complejo, estos efectos estructurales y funcionales deben ocurrir mucho más rápido que los observados aquí *in vitro* con H_2O_2 o aire. Así, la oxidación de Gal-1 podría estar provocada por estos u otros oxidantes en forma indirecta, a través de reacciones de intercambio tiol-disulfuro, con proteínas más reactivas (proteínas que censan peróxidos). Interacciones de este tipo han sido demostradas y estudiadas primeramente en el factor transcripcional Yap, en donde la respuesta frente a peróxido de hidrógeno es mediada por la glutatión peroxidasa 3 y luego confirmadas en otros factores [36, 37].

La cinética de oxidación y el análisis electroforético (**Figura IV.5**) han mostrado la formación dependiente del tiempo de enlaces disulfuro intra e intermoleculares. La banda de menor peso molecular migró desde su posición inicial a comienzos de la cinética de oxidación (**Figura IV.5B**), indicando la formación de un enlace disulfuro intramolecular en el monómero de Gal-1. Al mismo tiempo, aquellas especies con un peso molecular característico de dímero aparecieron durante la oxidación y fueron eliminados en condiciones reductoras, confirmando así la presencia de enlaces

disulfuro intermoleculares entre los diferentes monómeros de Gal-1. La formación de estos dímeros fue más lenta que la primera fase de oxidación, la cual se completa en menos de 10 min (en 10 mM de H_2O_2 y las condiciones de Gal-1, temperatura y pH arriba indicadas).

La cinética del cambio conformacional provocado por oxidación con H_2O_2 y seguida por experimentos de fluorescencia revelaron que la señal final de intensidad de fluorescencia disminuía a medida que aumenta la concentración de oxidante (Figura IV.6). Hemos propuesto un modelo para explicar este fenómeno, el cual requiere la consideración de una reacción que depende de la concentración de peróxido que compita con el cambio conformacional que es independiente de ella: la sobreoxidación de cisteínas. Usando este modelo, se obtuvo una constante de velocidad de cambio conformacional de $0,003 \text{ s}^{-1}$, el cual a altas concentraciones de H_2O_2 compite con la sobreoxidación (formación de ácidos sulfínicos o sulfónicos que no reversion a cisteína) de las cisteínas que dirigen el cambio conformacional. Todos estos resultados se encuentran resumidos e integrados en el esquema de la Figura IV.11. Estos datos se convierten en relevantes más allá de la prueba *in vitro* cuando se considera una regulación estricta de la formación de enlaces disulfuro en Gal-1 para generar modificaciones estructurales sincronizadas con los requerimientos del microambiente tisular.

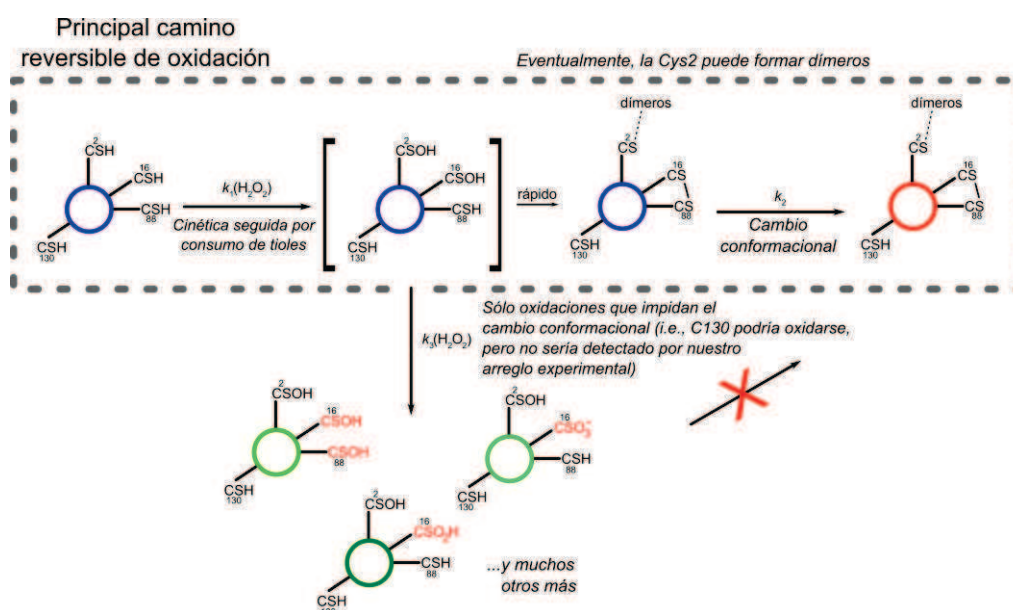


Figura IV.11. Esquema que resume el comportamiento redox de Gal-1 frente a H_2O_2 observado y estudiado en el presente capítulo.

Finalmente, y como principal motivación del estudio profundizado en Gal-1, se realizó una comparación estructural y de secuencia del resto de las galectinas en busca de similitudes en la cantidad y posición de los residuos de cisteína responsables de todo este comportamiento redox. Mediante dicho análisis (**Figura IV.10**) hemos podido detectar que sólo Gal-1, PP13 y PPL13 poseen las cisteínas más relevantes en el proceso de oxidación, resultado de nuestros estudios. Es posible, por lo tanto, que PP13 y PPL13 exhiban un comportamiento redox similar al observado para Gal-1, ya que poseen las cisteínas claves en la misma ubicación topológica. Curiosamente, Gal-1, así como también PP13 y PPL13, han sido propuestas como proteínas claves en el mantenimiento del embarazo, sugiriendo un rol cooperativo o redundante para estas lectinas en el fenómeno de placentación [38]. Además, el estado redox de PP13 y PPL13 es clave en el correcto funcionamiento y regulación de este proceso [39]. Por otro lado, se puede especular que Gal-3 y Gal-9N podrían haberse vuelto insensibles a la oxidación como consecuencia de una presión evolutiva para eliminar Cys16 de su secuencia. O tal vez la regulación redox gobernada por la formación del switch Cys16-Cys88 no sea necesaria para estas galectinas en particular.

IV.5 Referencias

1. *Animal Lectins: A Functional View*. (2008) Editado por Gerardo R. Vasta and Hafiz Ahmed. CRC Press
2. Pande, A.H., Sumati, R.K., and Hajela, K. (2003) Oxidation of Goat Hepatic Galectin-1 Induces Change in Secondary Structure. *Protein and Peptide Letters* 10, 265-75
3. Shahwan, M., Al-Qirim, M.T., Zaidi, S.M.K.R., and Banu, N. (2004) Physicochemical Properties and Amino Acid Sequence of Sheep Brain Galectin-1. *Biochemistry (Moscow)* 69, 506-12
4. Ashraf, G., Banu, N., Ahmad, A., Singh, L., and Kumar, R. (2011) Purification, Characterization, Sequencing and Biological Chemistry of Galectin-1 Purified from Capra hircus (goat) Heart. *The Protein Journal* 30, 39-51
5. Abbott, W.M. and Feizi, T. (1991) Soluble 14-kDa beta-galactoside-specific bovine lectin. Evidence from mutagenesis and proteolysis that almost the complete polypeptide chain is necessary for integrity of the carbohydrate recognition domain. *Journal of Biological Chemistry* 266, 5552-57
6. Hirabayashi, J. and Kasai, K. (1991) Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. *Journal of Biological Chemistry* 266, 23648-53
7. Oda, Y. and Kasai, K.-I. (1983) Purification and characterization of β -galactoside-binding lectin from chick embryonic skin. *Biochimica et Biophysica Acta (BBA) - General Subjects* 761, 237-45
8. Hirabayashi, J., Kawasaki, H., Suzuki, K., and Kasai, K.-i. (1987) Complete Amino Acid Sequence of 14 kDa β -Galactoside-Binding Lectin of Chick Embryo. *Journal of Biochemistry* 101, 775-87
9. Whitney, P.L., Powell, J.T., and Sanford, G.L. (1986) Oxidation and chemical modification of lung beta-galactoside-specific lectin. *Biochemical Journal* 238, 683-89

10. Tracey, B.M., Feizi, T., Abbott, W.M., Carruthers, R.A., Green, B.N., and Lawson, A.M. (1992) Subunit molecular mass assignment of 14,654 Da to the soluble beta-galactoside-binding lectin from bovine heart muscle and demonstration of intramolecular disulfide bonding associated with oxidative inactivation. *Journal of Biological Chemistry* 267, 10342-47
11. Lopez-Lucendo, M.F., Solis, D., Andre, S., Hirabayashi, J., Kasai, K., Kaltner, H., Gabius, H.J., and Romero, A. (2004) Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *Journal of Molecular Biology* 343, 957-70
12. Inagaki, Y., Sohma, Y., Horie, H., Nozawa, R., and Kadoya, T. (2000) Oxidized galectin-1 promotes axonal regeneration in peripheral nerves but does not possess lectin properties. *European Journal of Biochemistry* 267, 2955-64
13. Stone, J.R. and Yang, S. (2006) Hydrogen Peroxide: A Signaling Messenger. *Antioxidants & Redox Signaling* 8, 243-70
14. Veal, E. and Day, A. (2011) Hydrogen peroxide as a signaling molecule. *Antioxid Redox Sign.* 15, 147-51
15. Brigelius-Flohé, R. and Flohé, L. (2011) Basic principles and emerging concepts in the redox control of transcription factors. *Antioxid Redox Sign* 15, 2335 - 81
16. Clackson, T., Detlef, G., and Jones, P.T. (1991) *PCR, a Practical Approach*. IRL Press at Oxford University Press, Oxford.
17. Riddles, P.W., Blakeley, R.L., and Zerner, B. (1979) Ellman's reagent: 5,5'-dithiobis(2-nitrobenzoic acid) – a reexamination. *Analytical Biochemistry* 94, 75-81
18. Carballal, S., Radi, R., Kirk, M.C., Barnes, S., Freeman, B.A., and Alvarez, B. (2003) Sulfenic Acid Formation in Human Serum Albumin by Hydrogen Peroxide and Peroxynitrite†. *Biochemistry* 42, 9906-14
19. Bas, D.C., Rogers, D.M., and Jensen, J.H. (2008) Very fast prediction and rationalization of pKa values for protein–ligand complexes. *Proteins: Structure, Function, and Bioinformatics* 73, 765-83
20. Anisimov, V.M. and Cavasotto, C.N. (2011) Quantum mechanical binding free energy calculation for phosphopeptide inhibitors of the Lck SH2 domain. *Journal of Computational Chemistry* 32, 2254-63
21. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-80
22. Toscano, M., Bianco, G.A., Ilarregui, J.M., Croci, D.O., Correale, J., Hernandez, J.D., Zwirner, N.W., Poirier, F., Riley, E.M., Baum, L.G., and Rabinovich, G.A. (2007) Differential glycosylation of TH1, TH2 and TH-17 effector cells selectively regulates susceptibility to cell death. *Nature Immunology* 8, 825-34
23. Lange, F., Brandt, B., Tiedge, M., Jonas, L., Jeschke, U., Pöhland, R., and Walzel, H. (2009) Galectin-1 induced activation of the mitochondrial apoptotic pathway: evidence for a connection between death-receptor and mitochondrial pathways in human Jurkat T lymphocytes. *Histochemistry and Cell Biology* 132, 211-23
24. Pe'er, I., Felder, C.E., Man, O., Silman, I., Sussman, J.L., and Beckmann, J.S. (2004) Proteomic signatures: Amino acid and oligopeptide compositions differentiate among phyla. *Proteins: Structure, Function, and Bioinformatics* 54, 20-40
25. Ferrer-Sueta, G., Manta, B., Botti, H., Radi, R., Trujillo, M., and Denicola, A. (2011) Factors Affecting Protein Thiol Reactivity and Specificity in Peroxide Reduction. *Chemical Research in Toxicology* 24, 434-50
26. Kadoya, T. and Horie, H. (2005) Structural and Functional Studies of Galectin-1: A Novel Axonal Regeneration-Promoting Activity for Oxidized Galectin-1. *Current Drug Targets* 6, 375-83
27. Perillo, N.L., Pace, K.E., Seilhamer, J.J., and Baum, L.G. (1995) Apoptosis of T cells mediated by galectin-1. *Nature* 378, 736-39
28. Rabinovich, G.A. and Ilarregui, J.M. (2009) Conveying glycan information into T-cell homeostatic programs: a challenging role for galectin-1 in inflammatory and tumor microenvironments. *Immunological Reviews* 230, 144-59
29. Fuchs, J., Zollner, T.M., Kaufmann, R., and Podda, M. (2001) Redox-modulated pathways in inflammatory skin diseases. *Free Radical Biology and Medicine* 30, 337-53

30. Stowell, S.R., Cho, M., Feasley, C.L., Arthur, C.M., Song, X., Colucci, J.K., Karmakar, S., Mehta, P., Dias-Baruffi, M., McEver, R.P., and Cummings, R.D. **(2009)** Ligand reduces galectin-1 sensitivity to oxidative inactivation by enhancing dimer formation. *J Biol Chem* 284, 4989-99
31. Nishi, N., Abe, A., Iwaki, J., Yoshida, H., Itoh, A., Shoji, H., Kamitori, S., Hirabayashi, J., and Nakamura, T. **(2008)** Functional and structural bases of a cysteine-less mutant as a long-lasting substitute for galectin-1. *Glycobiology* 18, 1065-73
32. Yamaoka, K., Ingendoh, A., Tsubuki, S., Nagai, Y., and Sanai, Y. **(1996)** Structural and Functional Characterization of a Novel Tumor-Derived Rat Galectin-1 Having Transforming Growth Factor (TGF) Activity: The Relationship between Intramolecular Disulfide Bridges and TGF Activity. *The Journal of Biochemistry* 119, 878-86
33. Winterbourn, C.C. and Metodiewa, D. **(1999)** Reactivity of biologically important thiol compounds with superoxide and hydrogen peroxide. *Free Radical Biology and Medicine* 27, 322-28
34. Goldman, R., Stoyanovsky, D.A., Day, B.W., and Kagan, V.E. **(1995)** Reduction of Phenoxyl Radicals by Thioredoxin Results in Selective Oxidation of Its SH-Groups to Disulfides. An Antioxidant Function of Thioredoxin. *Biochemistry* 34, 4765-72
35. Cho, M. and Cummings, R.D. **(1996)** Characterization of monomeric forms of galectin-1 generated by site-directed mutagenesis. *Biochemistry* 35, 13081-88
36. Delaunay, A., Pflieger, D., Barrault, M.-B., Vinh, J., and Toledano, M.B. **(2002)** A Thiol Peroxidase Is an H₂O₂ Receptor and Redox-Transducer in Gene Activation. *Cell* 111, 471-81
37. Flohé, L. and Ursini, F. **(2008)** Peroxidase: A term of many meanings. *Antioxid Redox Sign* 10, 1485-90
38. Than, N.G., Romero, R., Erez, O., Weckle, A., Tarca, A.L., Hotra, J., Abbas, A., Han, Y.M., Kim, S.-S., Kusanovic, J.P., Gotsch, F., Hou, Z., Santolaya-Forgas, J., Benirschke, K., Papp, Z., Grossman, L.I., Goodman, M., and Wildman, D.E. **(2008)** Emergence of hormonal and redox regulation of galectin-1 in placental mammals: Implication in maternal-fetal immune tolerance. *Proceedings of the National Academy of Sciences* 105, 15819-24
39. Than, N.G., Pick, E., Bellyei, S., Szigeti, A., Burger, O., Berente, Z., Janaky, T., Boronkai, A., Kliman, H., Meiri, H., Bohn, H., Than, G.N., and Sumegi, B. **(2004)** Functional analyses of placental protein 13/galectin-13. *European Journal of Biochemistry* 271, 1065-78



EXPLORANDO LA ESTRUCTURA CUATERNARIA DE LAS GALECTINAS: ¿EXISTEN LOS LATTICES?

V.1 Introducción

Aunque las galectinas poseen todas un plegamiento similar, es sabido que su asociación cuaternaria es muy variada, principalmente cuando se comparan todas las galectinas entre los cordados [1]. Al sumergirnos especialmente en el estudio de las galectinas humanas, nos encontramos con escasez de estudios y resultados en dicho tema [2]. Sólo es conocido el estudio con cierto detalle de la dimerización de Gal-1, -2 y -7 y la oligomerización de Gal-3 en determinados trabajos analizando algunas de estas galectinas juntas o por separado [3]. Se sabe que Gal-1 existe en equilibrio entre la forma monomérica y su homodímero [4] con una constante de dimerización (*in vitro*) de 7 μM [5] y existen actualmente mutantes de Gal-1 para las cuales predomina la forma monomérica [5], o el dímero se ha transformado en permanente mediante el agregado de un péptido *linker* entre cada DRC [6]. También existe una forma naturalmente procesada en el extremo *N*-terminal que es incapaz de dimerizar y que posee una actividad biológica diferente en cuanto a su capacidad de inducción de apoptosis en ciertas líneas celulares [7]. Estudios cristalográficos revelan que Gal-2 también puede existir como un homodímero [8] muy similar al observado para Gal-1 [9] pero ningún estudio más se ha efectuado para caracterizar el equilibrio de dimerización y compararlo con el resto de las galectinas. Para el caso de Gal-3, el único miembro de las galectinas tipo quimera, se ha estudiado la estructura de la proteína murina intacta a través de estudios de RMN [10] mostrando que la secuencia peptídica al comienzo del extremo *N*-terminal (de alrededor de 120 aminoácidos ricos en prolina, glicina y tirosina) se encuentra desplegada y que ocasionalmente interactúa con el DRC, el cual, aislado de su extremo *N*-terminal, es monomérico [11] a diferencia de

Gal-1, -2 y -7. El DRC por sí solo es capaz de unir lactosa pero no posee actividad de hemaglutinación ni de unión cooperativa como sí muestra la Gal-3 completa [12]. Estos estudios sugieren que el extremo N-terminal entonces es importante para la agregación de la proteína. Pero los resultados sobre el estado de oligomerización son realmente variados. Existen reportes de dímeros de Gal-3 determinados por microscopía electrónica [13]. Algunos ensayos de filtración en gel no son capaces de encontrar multímeros [12, 14]. En la misma dirección, experimentos de velocidad de sedimentación y medidas de equilibrio muestran que la proteína existe predominantemente como monómero [3]. Y para sumar a la discusión, el trabajo tal vez más aceptado en la bibliografía indicaría que Gal-3 interactúa con una estequiometría de pentámero con los glicanos con los que se ha ensayado la precipitación de la proteína [15]. Finalmente, para Gal-7 contamos con una estructura proveniente de estudios cristalográficos [16] en donde se muestra que la dimerización de esta lectina es muy diferente al resto de las galectinas prototipo: un parche hidrofóbico en la cara-*F*, opuesta al LBG, es el responsable de mediar la dimerización, comprometiendo unos 1500 Å² de superficie proteica [17]. Recientemente, un estudio de Ermakova y col. [18], ha profundizado un poco más en este mecanismo particular de dimerización y se ha comenzado a especular en las consecuencias del mismo. Por otro lado, aún no se ha determinado a ciencia cierta el rol en la oligomerización de los distintos DRCs en las galectinas tipo tándem, en donde se especula que dada la variación de afinidad y selectividad de cada DRC dentro de estas galectinas, puede ser el causante de la diferente afinidad por diferentes receptores [19, 20].

Al fin de cuentas, ¿por qué resulta importante la forma y las características de esta oligomerización para comprender la función de las galectinas? Y aquí es donde también se continúa con la especulación en la bibliografía y se pone de manifiesto la necesidad de estudios más concluyentes sobre el rol de la oligomerización en estas lectinas en particular. Los trabajos pioneros en materia del *cross-linking* entre lectinas y glicanos se los debemos al gran aporte de Fred C. Brewer [21-30], quien trabajó en sistemas particulares, la gran mayoría bacterianos, en donde la formación de arreglos muy regulares en la membrana de las bacterias era evidente a la luz de experimentos de diversas microscopías electrónicas. Hoy en día se tiene cierta certeza de que la formación de complejos multivalente de galectinas solubles con receptores glicosilados en la membrana plasmática de las células eucariotas ayuda a organizar diferentes arreglos proteicos en la superficie de la célula [31] (**Figura V.1**).

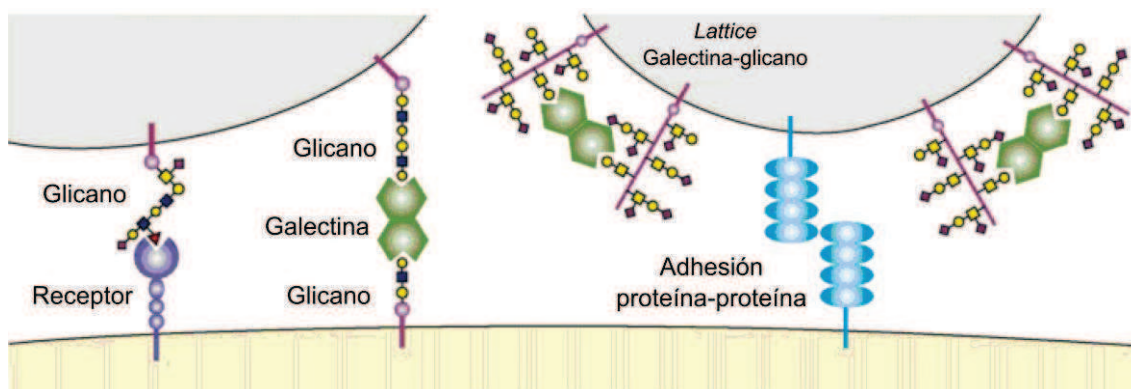


Figura V.1. Modelos de interacción glicano-receptor en procesos de adhesión celular. El modelo más a la izquierda involucra glicoproteínas en una célula interactuando con un receptor (como puede ser una molécula de selectina) en otra célula. Otros estudios muestran que las galectinas unen glicanos entre dos células, pero nueva evidencia sugiere que pueden favorecer la adhesión celular al formar redes o lattices con diferentes mucinas de la superficie celular. Tomado y adaptado de [32].

Las consecuencias de la formación de estos arreglos (ampliamente denominados en la literatura como *lattices*, quizás como consecuencia de los trabajos antes mencionados de Fred Brewer, a pesar de que es un término que indica “regularidad” y se desconoce si es el caso de los arreglos macromoleculares galectina-glicoproteínas) es realmente variada: desde organizar dominios en la membrana plasmática como ser balsas lipídicas (del inglés, *lipid rafts*), pasando por un tráfico adecuado de glicoproteínas en células polarizadas, hasta la regulación de la actividad y capacidad de señalización de algunas glicoproteínas de membrana, como ser ciertos receptores de factores de crecimiento, a través de la modulación del tiempo de residencia del receptor en membrana, inhibiendo la endocitosis y controlando la recirculación del mismo en membrana [31]. Sin embargo, todas estas actividades claves en el abanico de funciones de las galectinas están respaldadas por escasa evidencia experimental, y en donde no se focaliza en el rol de la capacidad de oligomerizar de estas lectinas en el esquema bioquímico completo que se propone en la mayoría de las revisiones del tema [31, 33, 34].

En este sentido, quizás una de las evidencias más robustas de la existencia de los *lattices* y del rol que cumple la oligomerización de las lectinas en la formación de los mismos se exploró en un par de trabajos importantes. El más importante es el trabajo de Nieminen y col. [35] en donde se demuestra la oligomerización de Gal-3 en neutrófilos y células endoteliales usando microscopía de fluorescencia basada en la transferencia de energía de resonancia de Föster (FRET). Los autores de éste trabajo proponen un trío de modelos posibles de interacción Gal-3-glicanos, incluida la

controversial formación de *lattices* (**Figura V.2**). El otro trabajo, mucho más “artificial” que el de Nieminen, es el ideado por Bilardi y col. [36], cuya estrategia se basó en la introducción de diferentes glicoconjugados fluorescentes en la membrana de una línea celular, apuntando hacia el exterior celular, y en la observación de FRET por parte de esas sondas glicoingenierizadas en presencia o ausencia de Gal-1. A pesar de la importancia colosal del asunto, en la actualidad no se cuenta con más evidencia directa de la existencia de estos *lattices* que la aquí citada, motivando así nuestro interés por aportar más evidencia y un mayor entendimiento al rol de la oligomerización de galectinas y la formación de estos agregados en la célula.

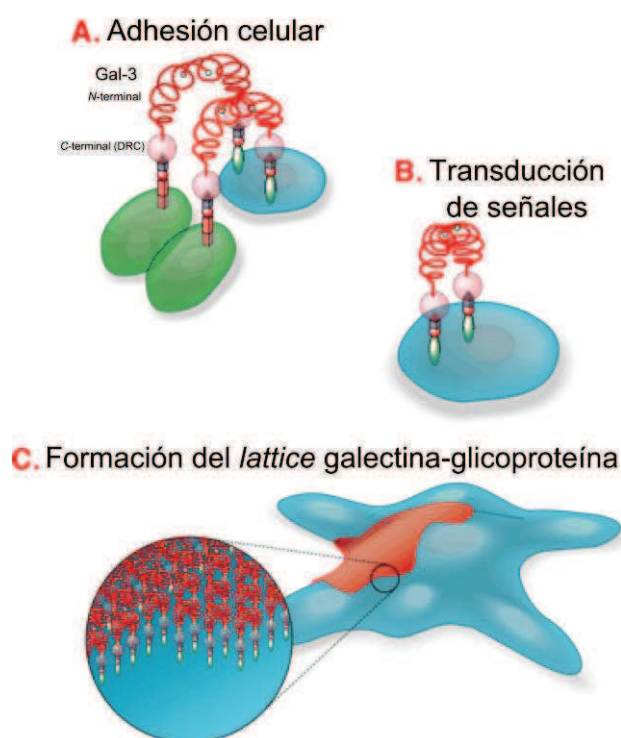


Figura V.2. Oligomerización de Gal-3. Luego de la unión del glicano a través del DRC, el dominio *N*-terminal de Gal-3 oligomeriza, y por lo tanto, entrecruza los glicanos en la superficie celular. Esta oligomerización de Gal-3 se puede observar como (A) parte del proceso de adhesión celular, (B) autoagregado de receptores que transducen señales y (C) formación de *lattices*, todos fenómenos observados por los autores (tomado y adaptado de [35]).

Con la reciente publicación del excepcional trabajo de Croci y col. [37] como parte de los objetivos generales del Laboratorio de Inmunopatología a cargo del Dr. Rabinovich, se ha marcado también un punto de partida para el entendimiento del rol de Gal-1 en el proceso de neovascularización o angiogénesis [38, 39]. El trabajo explora y demuestra que la interacción entre Gal-1 y los glicanos que decoran algunos dominios particulares del receptor de transmembrana tipo 2 de VEGF (VEGFR2, del

inglés, *vascular endothelial growth factor receptor 2*) es la responsable de desencadenar la activación de determinadas señales de crecimiento vascular en células endoteliales en ausencia del ligando natural del receptor, el VEGF, tanto *in vitro* como *in vivo* (**Figura V.3**). El VEGFR2 es uno de los receptores de VEGF, de una familia de tres subtipos. Consiste en una proteína con siete dominios tipo IgG extracelulares, un pequeño dominio transmembrana y un dominio C-terminal con actividad de tirosina-quinasa. (**Figura V.3A**). La unión del ligando natural VEGF induce la dimerización del receptor a través de determinados contactos entre los dominios D2, D4 y D7 del receptor y el ligando (**Figura V.3B y C**) y posterior autofosforilación del dominio intracelular.

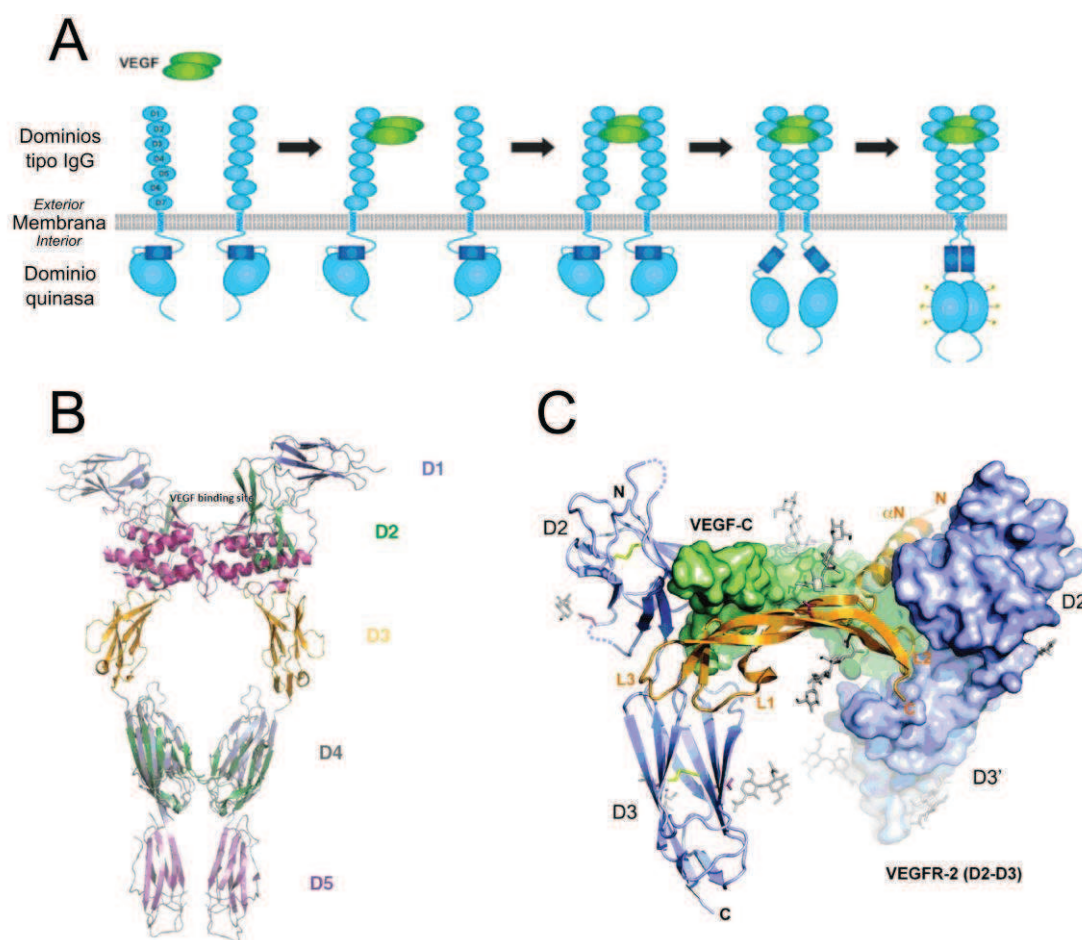


Figura V.3. VEGFR2. (A) Representación esquemática del mecanismo de activación de receptores de VEGF. Un dímero covalente de VEGF (verde) se une a VEGFR (azul), desencadenando la dimerización de dos monómeros de receptores. Interacciones adicionales entre los dominios 4 y 7 de tipo IgG, los dominios transmembranales y los dominios juxtamembranales (azul oscuro) estabilizan el complejo VEGF-VEGFR, resultando en la activación de la quinasa y la autofosforilación, necesaria para la señalización río abajo; (B) Estructura de los cinco primeros dominios (D1 a D5) del dímero de KIT (stem cell factor receptor) el cual tiene motivos estructurales idénticos a los de los receptores de VEGF; (C) Detalle estructural particular de los dominios D2 y D3 de VEGFR2 en complejo con VEGF-C. Observar los residuos de N-glycanos (principalmente LacNAc) los cuales se postulan como sitios de unión de Gal-1 en el trabajo de Croci y col. [37] (PDBID 2X1X).

El estudio de éste receptor interactuando con Gal-1 en diversidad de ensayos y modelos celulares por parte del grupo de trabajo del Dr. Rabinovich fueron el punto de partida para poder explorar, entonces, la formación de los *lattices* entre Gal-1 y VEGFR2 en la membrana de las células endoteliales y poder así aportar al entendimiento de la función de las galectinas a través del análisis de un requerimiento particular de la estructura cuaternaria de estas lectinas. Por lo tanto, en este último capítulo, se exploró la formación de los *lattices* de Gal-1:VEGFR2 mediante ensayos *in vitro* en células endoteliales, utilizando diferentes microscopías y proteínas recombinantes y tratando, nuevamente, de correlacionar estos resultados con las características particulares a nivel atómico-molecular de las galectinas ensayadas, mediante el análisis de la interfase de dimerización de Gal-1 y un grupo de mutantes desarrolladas *in silico*.

V.2 Métodos experimentales y computacionales

Expresión y purificación de Gal-1 y la mutante monomérica mGal-1. La purificación de ambas proteínas recombinantes siguió el protocolo detalladamente explicado en la sección II.2.2. La variante monomérica mGal-1 consiste en una doble mutante: C2S y V5D, siendo ésta última la principal responsable de la desestabilización de las interacciones entre monómeros (pues se encuentra en la interfase de dimerización), produciendo así una mutante de Gal-1 con una constante de homodimerización mucho menor que la observada para la proteína wild-type (ver sección V.3). El plásmido conteniendo el gen mutado fue suministrado por gentileza del Dr. Jun Hirabayashi.

Líneas y cultivo celulares. Se emplearon las células HUVEC provenientes de cultivo primario y como control de no interacción de los anticuerpos, la línea celular B16, según lo descrito en la sección II.2.3.

Ensayo de unión de lectinas. Células HUVEC fueron desprendidas del plato de cultivo por medio de solución no enzimática (Sigma). Luego, las células fueron lavadas con buffer de 'lectinas' (NaCl 150 mM 10mM HEPES pH 7,4) 1% BSA. Posteriormente se incubaron durante 1 h en 100 μ L de buffer 'lectinas' conteniendo las diferentes lectinas conjugadas con biotina: L-PHA (2 μ g/mL); LEL (1 μ g/mL); SNA (2 μ g/mL);

MAL-II (10 µg/mL) y Gal-1 (10 µg/mL). Posteriormente las células se lavaron y fueron incubadas con estreptavidina-PE 1:100 durante 30 min a 25 °C. Luego se lavó el exceso de estreptavidina y se fijaron las células con 1% PFA y se guardaron a 4 °C hasta su análisis por citometría de flujo. Las muestras fueron analizadas en un citómetro de flujo FACSaria utilizando el software FlowJo.

Microscopía de inmunofluorescencia. Se emplearon células B16 (control de inespecífico), células HUVEC sin tratar (control) y tratadas con Gal-1 (se realizaron estudios con gradientes de concentración y ensayos a concentración única de 40 µg/mL, tanto de Gal-1 como de mGal-1, según el caso), cultivadas según se indicó en la sección II.2.3 en discos de vidrio estériles (1 cm de diámetro) y cubiertos con colágeno (sección II.2.3). Previo al tratamiento con Gal-1 se cambió el medio de cultivo a un medio mínimo (medio F-12K + 20 % SFB), sin antibióticos ni cóctel de crecimiento durante 10 h. Luego, se agregó Gal-1 recombinante durante 1 h a 37 °C 5 % CO₂ con la concentración adecuada de lectina y un posterior lavado con PBS 1 % SFB. Las células se fijaron con solución 2 % PFA/sacarosa por 20 min a 25 °C. Posteriormente, se bloquearon por 1 h en 5 % BSA en PBS, se incubaron con un anticuerpo monoclonal anti - VEGFR2 (2 µg/ml; 55B11) en 1 % BSA en PBS durante toda la noche a 4 °C. Luego de tres lavados con PBS, las células se incubaron 1 h en presencia del anticuerpo secundario anti - IgG de conejo acoplado a Alexa-Fluor 488 (1:1000). Finalmente las células se lavaron exhaustivamente y se procedió al montaje de las muestras con medio ProLong® Gold Antifade en presencia de DAPI (300 nM). Los preparados se guardaron a -20 °C en oscuridad. Las imágenes se tomaron utilizando el microscopio DeltaVision Olympus IX70 Deconvolution RT. Los análisis cuantitativos y el procesamiento de las imágenes se realizaron con los programas Metamorph® y ImageJ.

Microscopía de transmisión electrónica de secciones celulares marcadas con inmuno-oro (inmunocitoquímica). Células HUVEC cultivadas según se indicó en la sección II.2.3 fueron estimuladas con 40 µg/mL Gal-1 durante 1 h en estufa de cultivo (37 °C y 5 % CO₂), se lavaron con PBS para eliminar el excedente de Gal-1 unido inespecíficamente y se fijaron en 8 % PFA en PBS pH 7,4 e incubadas a 25 °C durante 1h. Luego, las células fijadas fueron separadas del plato de cultivo utilizando un rastrillo para células y las mismas se colocaron en un tubo para centrifuga. El pellet se lavó tres veces con PBS y se resuspendió en gelatina 5 % en PBS fundida. Se centrifugó

inmediatamente y se enfrió en baño de hielo hasta solidificación de la suspensión. Se cortaron cubos conteniendo el pellet de células y se los criopreservó, sumergiéndolos en una solución al 15 % polivinilpirrolidona y 2 M sacarosa durante 16 h. A partir de estos cubos se cortaron finas secciones (70-100 nm de espesor) a -130 °C (corriente continua de nitrógeno líquido) en un ultramicrotomo Leica UCT (Leica). Las secciones se levantaron suavemente en una gota de una solución con metilcelulosa/sacarosa e inmediatamente se depositaron sobre los grids de cobre que sostienen la muestra dentro del microscopio electrónico (sección II.2.4.4). En dichos grids, la muestra se incubó con anti-VEGFR2 (1:10) y luego con proteína A de conejo conjugada con nanopartículas de Au de 10 nm de diámetro. Finalmente, los grids se embebieron en 2 % de metilcelulosa y 0,3 % acetato de uranilo para contraste de membranas. El exceso de fluido se retiró con un papel de filtro y los grids se visualizaron en un microscopio electrónico de transmisión Tecnai T-12 (sección II.2.4.4).

Microscopía de transmisión crío-electrónica de la célula completa. Se emplearon células HUVEC tratadas con Gal-1 como se explicó anteriormente, y cultivadas según se indicó en la sección II.2.3 pero esta vez sobre grids de Au (sección II.2.4.4), previamente esterilizados con radiación UV y recubiertos con colágeno (sección II.2.3). Luego del tratamiento con Gal-1, las células se fijaron con 8 % PFA en PBS a 25 °C durante 30 min y se enjuagaron tres veces con PBS. A estos grids se les realizó el mismo procedimiento de marcación de inmunofluorescencia arriba mencionado, para verificar que las células muestran los mismos fenómenos de agregación en estas condiciones particulares de cultivo e ,idealmente, poder lograr una correlación entre los resultados de microscopía de fluorescencia y electrónica sobre la misma muestra. Luego de tomadas las imágenes de estas células al microscopio confocal, los grids fueron secados y congelados en etano líquido mantenido a -180 °C con nitrógeno líquido en un equipo Vitrobot® (*vitrification robot*, FEI Company) utilizado para lograr el congelamiento ultra-rápido de muestras acuosas y lograr que el agua de las mismas se vitrifique y permita la transmisión de una imagen de electrones. Los grids así congelados se mantuvieron sumergidos en nitrógeno líquido hasta poder cargar un *cryo-holder* Gatan, modelo 626 (Gatan, Pleasanton, CA) el cual inserta los grids congelados con la muestra en el microscopio de transmisión Polara G2 (FEI Company), operado a temperaturas criogénicas de no más de -180 °C durante toda la toma de imágenes. Típicamente, se registraron imágenes de la célula completa a bajas

dosis de electrones (1-2 electrones por Å²), operando a voltajes de 300 kV y logrando una magnificación efectiva de unos 30.000X.

Experimentos computacionales. Las estructuras iniciales de Gal-1 fueron obtenidas de la PDB: 1GZW para la proteína wild-type cristalizada como dímero (experimento de difracción de rayos X con 1,65 Å de resolución). Para las estructuras restantes ensayadas que no estaban disponibles en la PDB (mutantes simples, dobles o triples en la región de dimerización, a partir de la mutación *in silico* de Gal-1 wild-type, tanto en su forma de monómero como de dímero) fueron generadas usando el paquete computacional Amber9. En todos los casos, se removieron los iones y las moléculas de agua cristalográficas y se re-solvataron con aguas TIP3P (sección II.3) en una caja octaédrica (10 Å de distancia entre la superficie proteína y el borde de la misma). Se llevaron a cabo simulaciones de DM de 60 ns a 1 atm y 300 K según los protocolos y procedimientos detallados en la secciones II.1.2, II.1.3 y II.1.5 e implementados en detalle en el capítulo III (ver sección III.2). Los parámetros termodinámicos correspondientes al proceso de dimerización de los diferentes DRC analizados fueron calculados con el método MM/GBSA, ampliamente discutido en la sección II.1.5.2.

V.3 Resultados

V.3.1 Análisis de la interfase de dimerización de Gal-1

Como punto de partida del análisis de la oligomerización de las galectinas proto-tipo, comenzamos con un estudio detallado de la región proteica que permite la dimerización de Gal-1. En la **Figura V.4** se muestran las únicas estructuras de galectinas como oligómeros (**Figura V.4A**), prestando especial atención a la región de dimerización de Gal-1 en particular (**Figura V.4B**), la cual fue nuestro sistema de estudio. Dada la simetría particular en la que se ensamblan los dos monómeros de Gal-1, es posible observar un enfrentamiento determinado de la misma cadena de aminoácidos de un monómero con la del otro (**Figura V.4C**), tanto de los extremos N-terminal como de los del C-terminal. De un análisis visual detallado de este enfrentamiento de cadenas laterales fue posible encontrar formas particulares de perturbar tanto las interacciones hidrofílicas como las hidrofóbicas existentes y realizar

mutaciones dirigidas con tal fin. En particular, las formas de hacerlo responden a la sustitución de aminoácidos no polares (disminuir los contactos hidrofóbicos entre los monómeros, típicos en regiones de interacción proteína-proteína) por aminoácidos cargados (para aumentar la repulsión electrostática en la interfase de dimerización). Por ello, siguiendo esta estrategia y mirando detalladamente la región de dimerización para seleccionar los aminoácidos interactuantes y mejor posicionados, se plantearon las siguientes proteínas a estudiar: Gal-1 wild-type (WT, como control), C2S-V5D (la mutante monomérica mGal-1 disponible en el laboratorio para hacer experimentos de mesada), C2S-V5D-S7D, C2S-V5D-V131D, C2S-V5D-F133A, A6D, A6D-V5D, A6D-V131D, A6D-V5D-V131D y los controles de las mutaciones simples C2S, V5D y A6D.

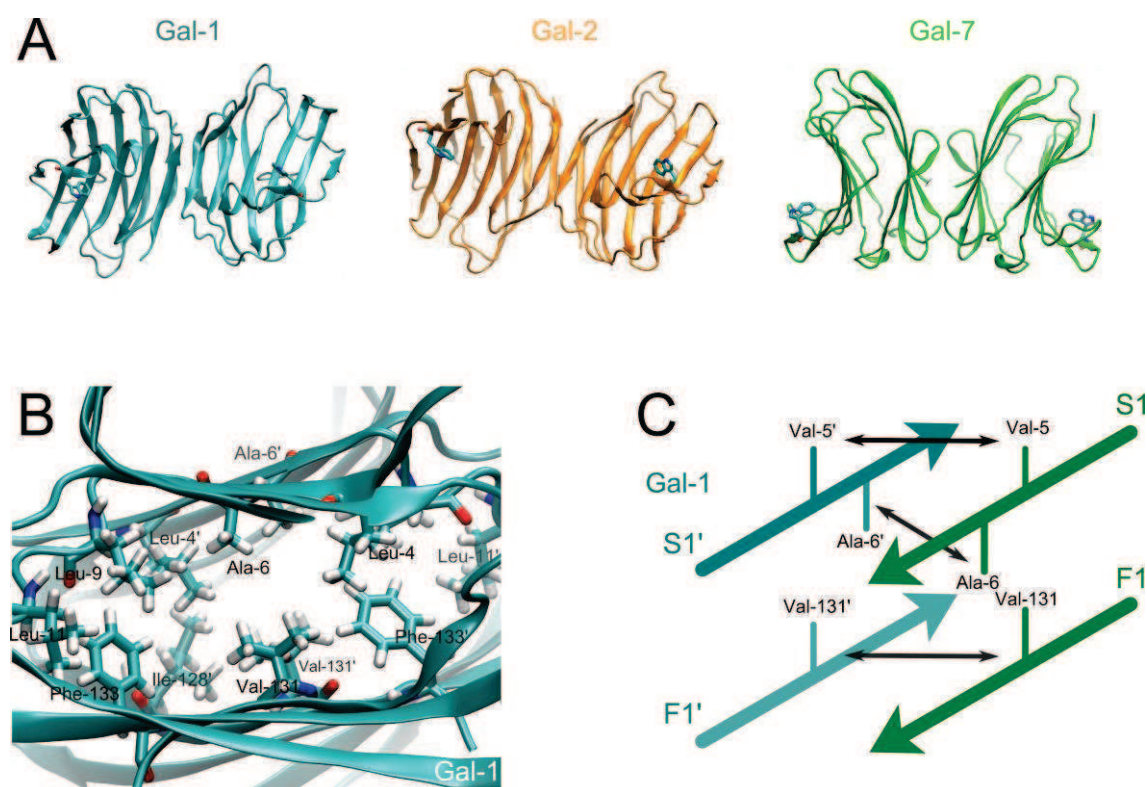


Figura V.4. Oligomerización de galectinas. (A) Estructuras cuaternarias determinadas para Gal-1, -2 y -7. Se marca el residuo de triptófano en cada una de ellas para ubicar el LBG de cada una de ellas; (B) Detalle de los residuos hidrofóbicos (bolsillo hidrofóbico) en la región de dimerización de Gal-1; (C) Representación esquemática de los aminoácidos más importantes para la realización de mutantes que impacten tanto en un monómero como en el otro a la hora de dimerizar. Se utilizó una prima (') para distinguir el mismo aminoácido pero de monómeros diferentes.

Las simulaciones de DM de los monómeros mutantes permitieron diagnosticar si la perturbación introducida por la mutación afecta o compromete el plegamiento de la proteína, más allá de la región de dimerización. Ninguna de las mutantes mostró

perfiles de RMSD en función del tiempo que difieran significativamente con el observado para Gal-1 (**Figura V.5A**). Esto indicaría que todas las mutaciones ensayadas perturban muy poco la estructura del monómero y que son proteínas conformacionalmente estables. Obviamente se podría haber prestado especial atención al hecho posible de que estos pequeños cambios inducidos por las mutaciones afectaran la capacidad de unión a glicanos, pues es en definitiva la función biológica ulterior de la proteína, y simplemente aquí se deseó perturbar el equilibrio de dimerización pero no el de unión a ligandos. Sin embargo, podemos *a priori* especular que dado que estas mutaciones se encuentran realmente lejos de los aminoácidos principalmente involucrados en la unión a azúcares (~ 20 Å), y que el RMSD de todos los monómeros mutados no presentaron grandes diferencias con el plegamiento de Gal-1 wild-type, entonces es esperable que las mutantes conserven su capacidad de unión y especificidad por glicanos.

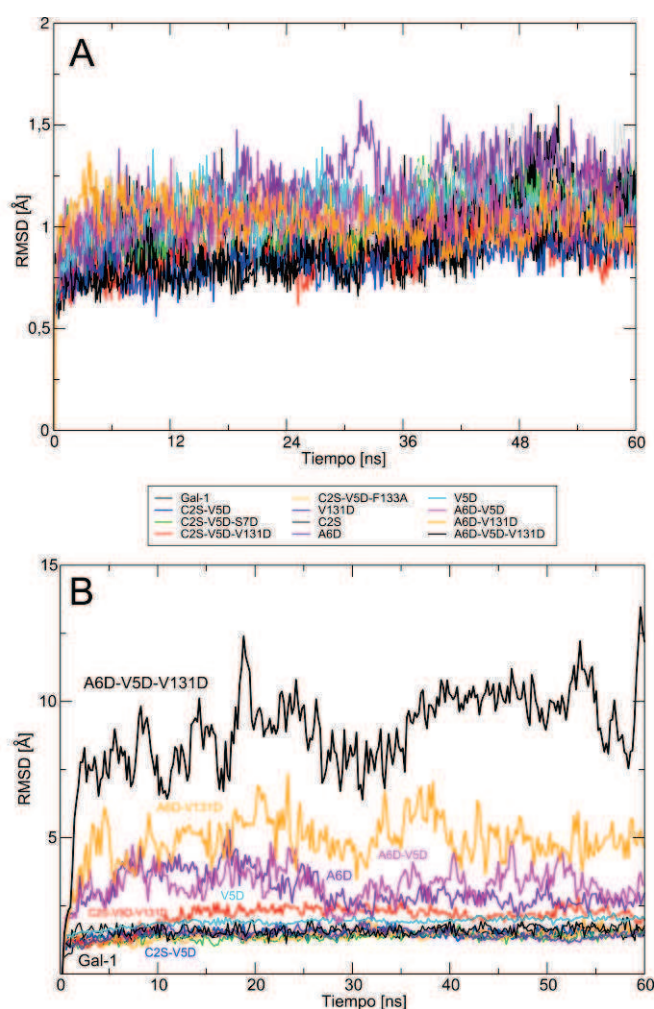


Figura V.5. RMSD en función del tiempo de simulación, (A) para los monómeros y (B) para los dímeros de Gal-1 ensayados.

Los resultados de RMSD en función del tiempo de simulación para los dímeros mutados se muestran en la **Figura V.5B**. Los mismos muestran claramente que las mutantes A6D-V5D-V131D, A6D-V131D, A6D-V5D y A6D son capaces de perturbar la estructura del dímero lo suficiente como para separarlo en monómeros, objetivo principal de la introducción de dichas mutaciones. Por los resultados, la mutación A6D logró mostrar diferencias significativas respecto de la dinámica del dímero de Gal-1. Cuando se analizó el efecto del agregado de una segunda mutación a la estructura de la mutante A6D, se observó que la mutación V5D no afecta significativamente la estabilidad del dímero, salvo en algunos momentos al final del tiempo de simulación analizado. Sin embargo, la mutación V131D logró una mayor perturbación de la estructura, permitiendo valores de RMSD (respecto de Gal-1) por encima de los 4 Å de diferencia. Lo interesante fue la incorporación de las dos últimas mutaciones analizadas en una mutante triple. La estructura de la mutante A6D-V5D-V131D muestra las mayores desviaciones respecto de la estructura canónica del dímero de Gal-1, alcanzando valores de 10 Å de RMSD. Esto se justifica por el hecho de haber incorporado seis cargas negativas en cadenas laterales móviles en la interfase de dimerización (3 por cada monómero) reemplazando seis aminoácidos compactos y no polares. Sin lugar a dudas la nueva situación de interacciones intermoleculares en la región de dimerización resultó más favorable a la separación del dímero y, por ello, la estructura del mismo se vio tan alterada a lo largo de la simulación de DM. Es notable observar que los efectos inducidos por las mutaciones V5D y V131D no son aparentemente aditivos pues las mutantes A6D y A6D-V5D se comportan de manera muy similar, la mutante A6D-V131D aumenta su RMSD unos angstroms, pero la triple mutante triplica su valor de RMSD respecto la mutante A6D (ver el párrafo siguiente para un análisis energético más profundo). Por otro lado, la actual mutante disponible en bibliografía, C2S-V5D, se muestra como una proteína mucho más estable frente a las mutaciones que las propuestas en este estudio. Sólo se lograron perturbaciones significativas al introducir un nuevo residuo de ácido aspártico (C2S-V5D-V131D), tal como se evidencia por los valores similares de RMSD con las mutantes A6D y A6D-V5D e incluso así, el efecto no es tan apreciable como en la otra triple mutante ensayada. Finalmente, observar que la mutación del único residuo no polar y aromático (F133) por una alanina no logró perturbar el equilibrio de dimerización de manera significativa. Por lo que este residuo aromático no debe aportar ningún tipo de

interacción de tipo apilamiento con los restantes aminoácidos del bolsillo hidrofóbico de esta interfase.

Para corroborar aún más el efecto, se realizaron cálculos de la estimación del cambio de energía libre de dimerización ($\Delta G_{\text{dim}} = G_{\text{dímero}} - 2G_{\text{monómero}}$) de cada sistema, utilizando el método de MM/GBSA (sección II.1.5.2). Los resultados se resumen en la **Tabla V.I**. Resulta evidente que salvo para la mutante C2S, todas las otras mutantes presentan valores de $\Delta G_{\text{dimerización}}$ mayores que el calculado para Gal-1 wild-type. Por lo tanto, en una u otra manera, todas las mutaciones afectan la dimerización, facilitando el desplazamiento del equilibrio hacia la formación de monómero. La mutante C2S y las mutaciones dobles y triples que las contienen fueron incluidas en el análisis porque es una herencia de la mutante que se cuenta en el laboratorio para los experimentos. Con estos nuevos resultados demostramos que la presencia de Ser en lugar de Cys2 no perturbará nunca el equilibrio de dimerización (aunque si sabemos que Cys2 posee un rol en la oxidación de la proteína (sección IV.3), motivo por el cual, el plásmido contiene esta mutación en particular, además de V5D).

Proteína	ΔG_{dim}	ΔE_{elec}	ΔE_{vdW}
Gal-1	-36.9 ± 5.3	-200 ± 34	-64.7 ± 4.1
C2S	-36.4 ± 6.1	-249 ± 55	-63.1 ± 4.6
V5D	-26.4 ± 6.8	-153 ± 57	-63.6 ± 5.0
V131D	-27.6 ± 9.0	-140 ± 71	-63.0 ± 4.7
C2S-V5D	-24.2 ± 5.8	-68 ± 53	-60.8 ± 4.0
C2S-V5D-S7D	-28.9 ± 6.8	102 ± 44	-67.1 ± 5.1
C2S-V5D-V131D	-19.6 ± 7.3	173 ± 56	-55.8 ± 5.5
C2S-V5D-F133A	-23.8 ± 7.8	-151 ± 67	-66.4 ± 8.6
A6D	-14.8 ± 8.1	-168 ± 59	-58.2 ± 7.5
A6D-V5D	-17.3 ± 9.1	-88 ± 74	-43.4 ± 6.3
A6D-V131D	-9.1 ± 7.9	-71 ± 61	-28.6 ± 6.6
A6D-V5D-V131D	-7.1 ± 6.1	20 ± 166	-37.1 ± 11.6

Tabla V.I. Resumen de los cálculos de cambio de energía libre de dimerización para todas las proteínas ensayadas.

Todas las energías están en kcal mol⁻¹ y los errores se informan como un DS.

En sintonía con lo que se observó en los perfiles de RMSD, las mutantes C2S-V5D-V131D, A6D, A6D-V131D y A6D-V5D-V131D son las que presentan los ΔG_{dim} menos negativos de todos. Los valores del cambio de energía de van der Waals (E_{vdW}) en todas las mutantes aumentan respecto de Gal-1, pues en casi todos los casos (salvo para las mutaciones de Cys) se están mutando residuos no polares por residuos polares. En consecuencia, también se vio afectada la energía electrostática (E_{elec}). En particular, como las mutaciones se pensaron para que dos cargas negativas de los residuos de aspartato de monómeros diferentes se repelieran, los cambios de E_{elec} serán

menos negativos que para la proteína wild-type. Efectivamente es lo que sucede en todas las mutantes que involucran mutaciones por Asp (D). Inclusive, este cambio puede llegar a ser positivo como en las perturbaciones más grandes (triples mutantes). Notar que si bien el $\Delta G_{\text{dimerización}}$ es suma de muchas contribuciones, incluso algunas calculadas con mucho error como ser la solvatación y los cambios de entropía, el cambio de E_{elec} es mucho más robusto: al valor de ΔE_{elec} de la mutante A6D le podemos sumar una desestabilización de $+80 \text{ kcal mol}^{-1}$ debido a la mutación V5D (diferencia entre el ΔE_{elec} de las mutantes A6D-V5D y A6D) y un valor de $+100 \text{ kcal mol}^{-1}$ debido a la mutación V131D (diferencia entre el ΔE_{elec} de las mutantes A6D-V131D y A6D). Esto arrojaría un valor de 12 kcal mol^{-1} para el ΔE_{elec} de la triple mutante A6D-V5D-V131D, muy cercano al valor de 20 kcal mol^{-1} calculado para dicha mutante en particular. De todas formas, el resultado final que puede compararse con experimentos y otras mediciones, como el RMSD, es el $\Delta G_{\text{dimerización}}$, el cual resulta de una contribución de diferentes términos, no sólo de la componente electrostática de la interacción entre monómeros.

V.3.2 Del modelado molecular a las células: ensayo de unión a lectinas

Como primer paso en los ensayos *in vivo*, se decidió evaluar la presencia de glicanos en la superficie celular utilizando un ensayo de unión de un set de lectinas vegetales biotiniladas capaces de reconocer diferentes estructuras sacarídicas principalmente en *N*-glicanos, el cual ofrece un detalle de los glico-epitopes presentes en la superficie de las células endoteliales, el tipo celular elegido como modelo para evaluar el rol de la estructura cuaternaria de Gal-1. Las diferentes estructuras reconocidas por dichas lectinas se muestran en la **Figura V.6A**.

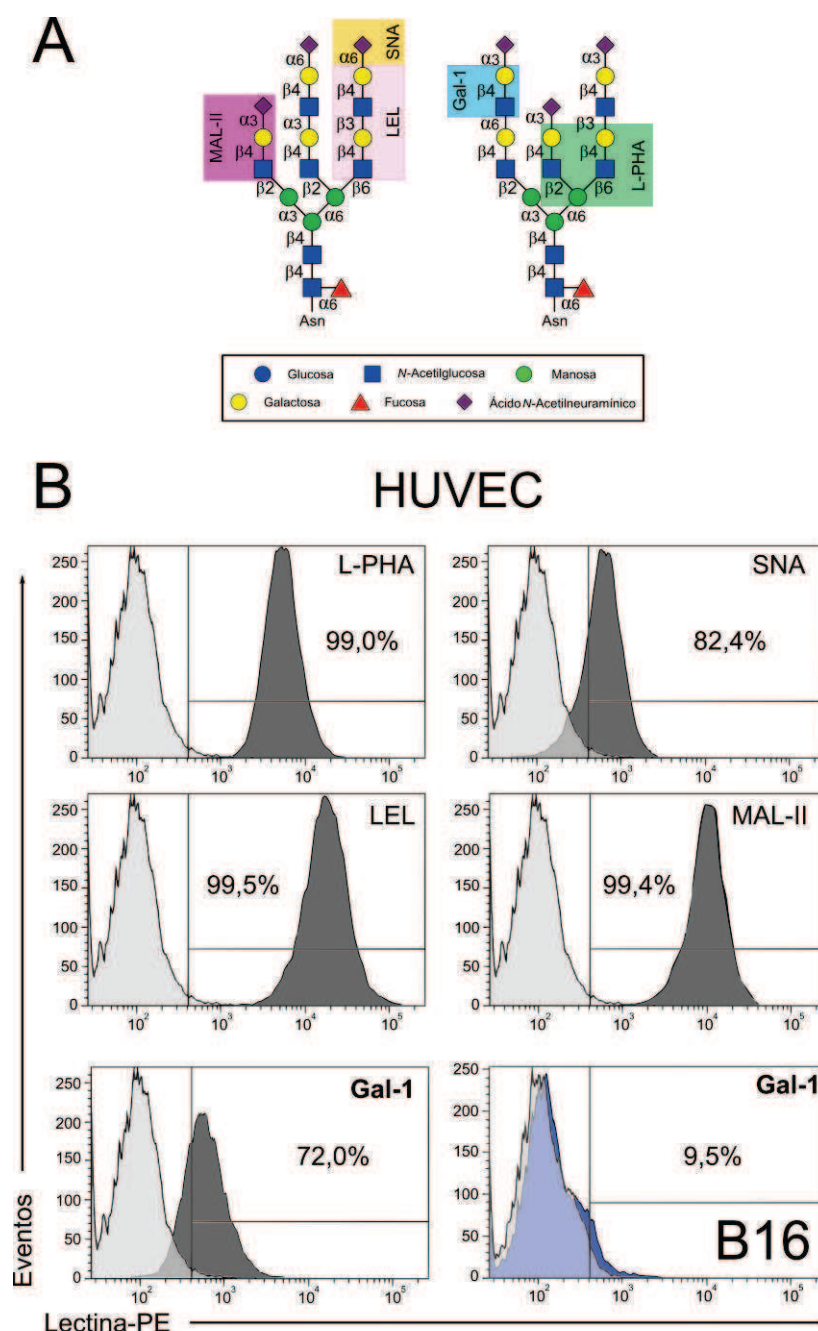


Figura V.6. Perfil de glicosilación y unión a lectinas en células HUVEC. (A) Sitios de unión para las diferentes lectinas L-PHA, SNA, LEL y MAL II en *N*-glicanos, indicando además la porción de glicano más importante para la unión de Gal1 (LacNAc). (B) Repertorio de glicanos en células HUVEC y B16 (sólo se muestra para Gal-1) bajo condiciones de reposo determinado por citometría de flujo a través de la unión selectiva de lectinas biotiniladas L-PHA, LEL, SNA, MAL II, PNA y HPA (histogramas oscuros) o sólo estreptavidina-PE como control (histogramas claros). Los datos son representativos de 2 experimentos independientes.

Tal como fue mencionado en la introducción (sección I.2.4), Gal-1 reconoce residuos individuales de LacNAc y unidades repetitivas presentes en poli-LacNAc [40]. El análisis de la glicosilación de células HUVEC en condiciones basales mostró una afinidad elevada para la lectina L-fitoheemaglutinina (L-PHA) (**Figura V.6B**), capaz de

reconocer ramificaciones en *N*-glicanos generadas por la enzima *N*-acetilglucosaminiltransferasa 5 (GnT5), y una alta reactividad de la lectina *Lycopersicon esculentum* (LEL) (**Figura V.6B**) que reconoce residuos poli-LacNAc en *N*- y en *O*-glicanos. Teniendo en cuenta la participación protagónica de los residuos de ácido siálico en la regulación de la unión de Gal-1 a diversos tipos celulares ([41] y resultados de esta Tesis, sección III.3), se evaluó también la presencia de ácido siálico en posición α 2-6 y α 2-3 en terminales de LacNAc, los cuales restringen o permiten la unión de Gal-1, respectivamente. El análisis reveló niveles bajos de ácido siálico en posición α 2-6, evidenciados por la asociación de la lectina *Sambucus nigra* (SNA) (**Figura V.6B**), en contraste con los altos niveles de ácido siálico α 2-3 evaluados a través de la unión de la lectina *Maackia amurensis* (MAL II) (**Figura V.6B**). Estos resultados demuestran que la expresión reducida del ácido siálico en posición α 2-6 no representa una pérdida global de ácido siálico en la membrana de células endoteliales, sino más bien una regulación diferencial de la biosíntesis de los glicanos. En conjunto, estos datos indican que las células endoteliales poseen un perfil de glicosilación permisivo para la unión de Gal-1, con alto número de ramificaciones en *N*-glicanos, abundancia de residuos terminales de poli-LacNAc y presencia limitada de ácido siálico en posición α 2-6, lo que se demuestra con la capacidad de unión de Gal-1 a éstas células (**Figura V.6B**). Por otro lado, el ensayo simple de la unión de Gal-1 a células tumorales de melanoma, como ser las de la línea celular B16, demostró que la capacidad de asociación de Gal-1 al perfil de glicosilación particular de éstas células (altos niveles de ácido siálico) fue prácticamente nulo (**Figura V.6B**), demostrando así que estos efectos son también dependientes de las células bajo las mismas condiciones basales de crecimiento.

V.3.3 Gal-1 induce la agregación del VEGFR2

A partir del trabajo de Croci y col. [37], quedó fuertemente establecido que VEGFR2 es un glicoreceptor en las células HUVEC para Gal-1, demostrado por múltiples enfoques moleculares y bioquímicos. Por lo tanto, se ha utilizado éste receptor como caso de estudio para explorar el efecto de la estructura cuaternaria de galectinas. Como primer experimento se incubaron células endoteliales HUVEC con diferentes concentraciones de Gal-1 wild-type para lograr obtener una curva de dosis-respuesta y comenzar a explorar los efectos de Gal-1 sobre éste receptor en la membrana celular. De estos experimentos resultó adecuado utilizar una concentración

de Gal-1 de 40 μM para el resto de los ensayos. Los resultados se muestran en la **Figura V.7**. El efecto evidenciado más importante de la presencia de Gal-1 sobre VEGFR2 fue el de la agregación del receptor en los contactos célula-célula. Dicho efecto fue claramente dependiente de la concentración de Gal-1 y anulado con el agregado de lactosa al medio de cultivo, demostrando el requerimiento de la unión a los *N*-glicanos presentes en el receptor.

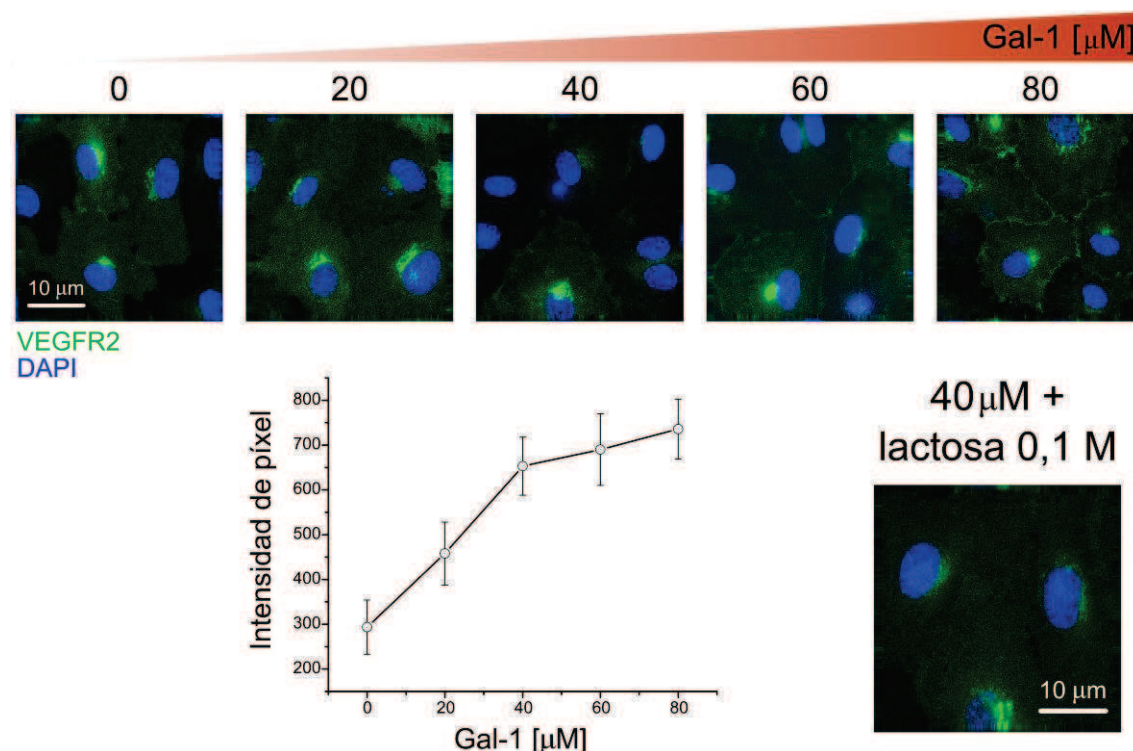


Figura V.7. Agregación de VEGFR2 en células HUVEC en función de la concentración de Gal-1. El efecto más pronunciado se encuentra a partir de una concentración de Gal-1 mayor a 40 μM , siendo el efecto más evidente la agregación del receptor en los contactos célula-célula. Para la cuantificación, se tuvieron en cuenta 30 contactos célula-célula diferentes, se cuantificó la intensidad de fluorescencia a lo largo de una línea que atravesaba el contacto entre células y se registró el valor máximo de intensidad. Los resultados se informan como la media $\pm$ un desvío estándar.

Cuando se analizó en detalle la población celular estimulada con Gal-1, fue posible evidenciar además, que en algunas células aisladas era importante la segregación de VEGFR2 en microdominios del soma celular (**Figura V.8**). Estos dominios se encontraron distribuidos en diferentes partes de las células, pero realmente limitados a una zona muy pequeña de la membrana celular. Se calculó el coeficiente de variación porcentual de la distribución de la intensidad de fluorescencia del VEGFR2 tanto para células sin tratar (control) como aquellas tratadas con Gal-1 y se observó una diferencia del 30 % (**Figura V.8**), indicando una mayor heterogeneidad del VEGFR2 en las células

tratadas, medición a favor de la hipótesis de que el receptor se agrega en presencia de Gal-1. Asombrosamente, mGal-1, la proteína mutante incapaz de dimerizar, no logró, a la misma concentración de 40 μM , inducir ninguno de estos efectos antes citados (Figura V.9).

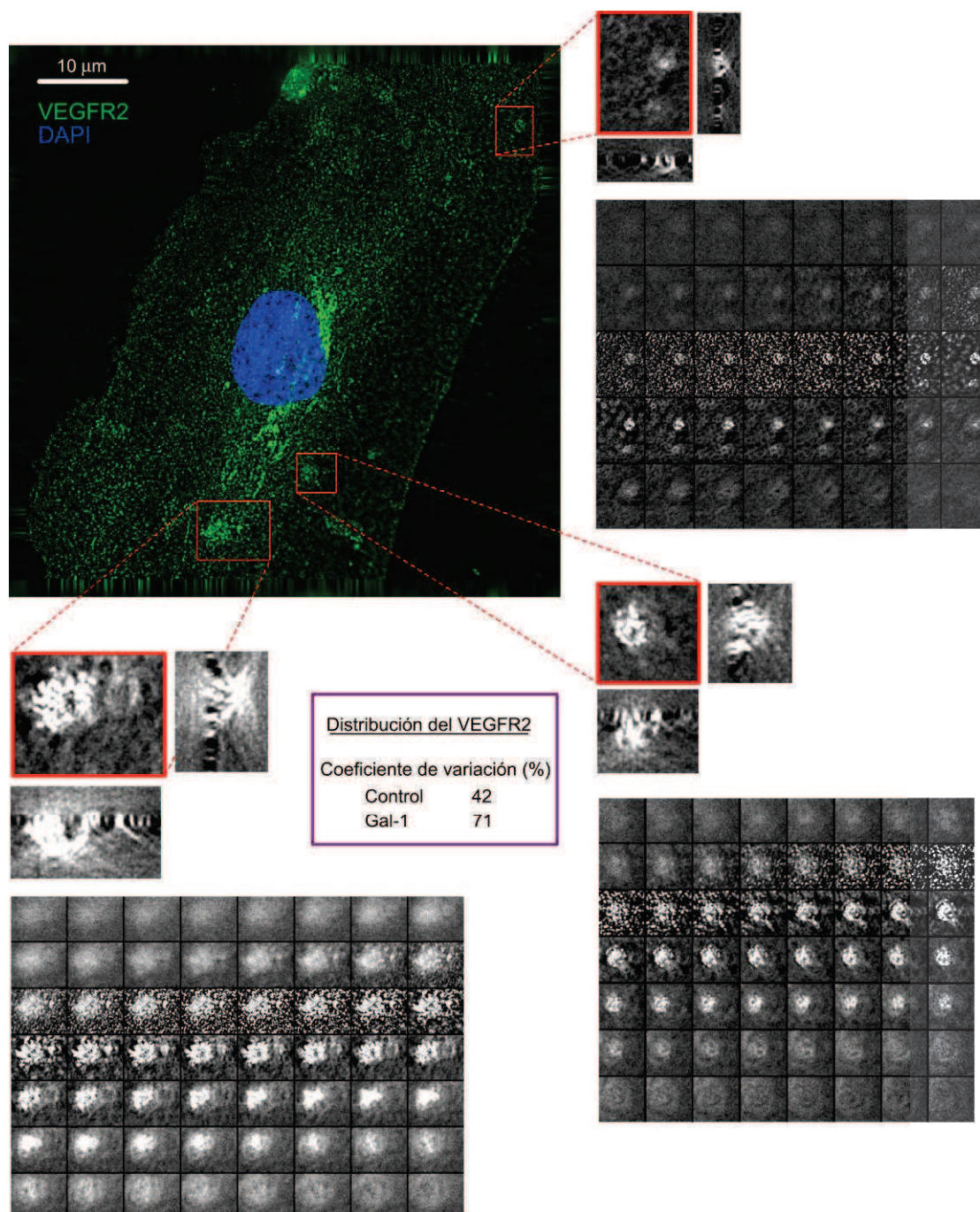


Figura V.8. Agregación de VEGFR2 en células HUVEC en microdominios de la superficie celular. La presencia de agregados de VEGFR2 en determinadas regiones de la superficie celular se hizo evidente al analizar en detalle cada célula. Se presentan las proyecciones en el plano de cada región de interés seleccionada, así como también un panel con las imágenes a cada valor del eje z (profundidad). De esta manera, se demostró que estos agregados se tratan de dominios en la superficie celular exclusivamente. En el recuadro central se ofrecen los resultados de los coeficientes de variación porcentuales de la intensidad de pixel tomadas en 30 regiones al azar ($100 \mu\text{m}^2$) tanto en células control como en tratadas con Gal-1.

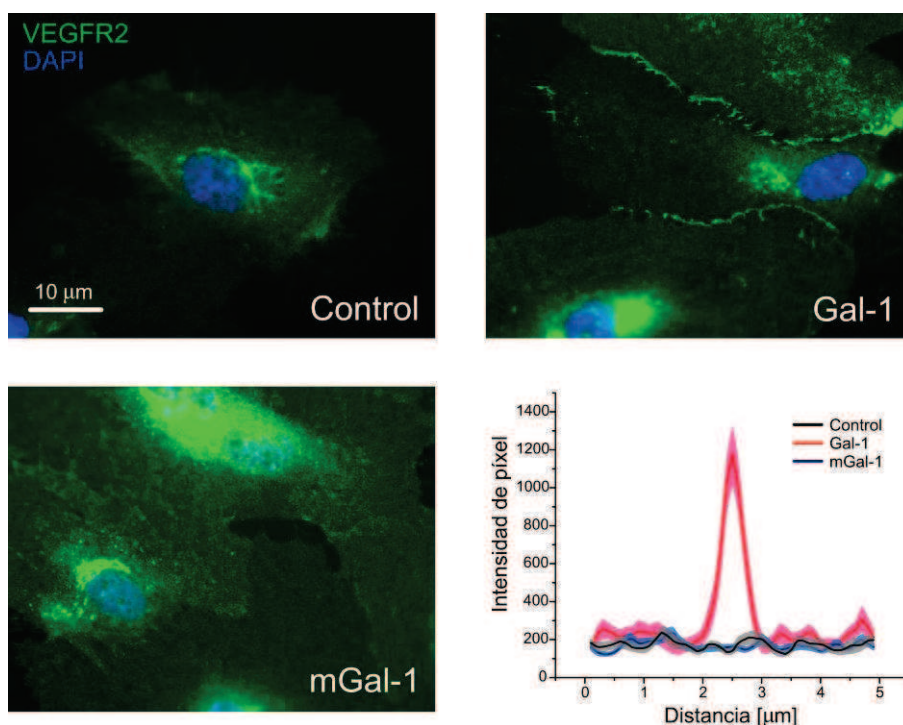


Figura V.9. La agregación de VEGFR2 no ocurre con mGal-1, la mutante de Gal-1 que no dimeriza. Se midió la intensidad de cada pixel a lo largo de una línea que atravesaba cada contacto célula-célula, tal como se explicó para la Figura V.7. Los valores de intensidad se registraron como la media $\pm$ un desvío estándar ($n = 30$), señalado como una sombra respecto de la línea que marca el valor medio.

Para corroborar la agregación de VEGFR2 en presencia de Gal-1 mediante otra metodología, y aumentando el zoom, se realizaron ensayos de inmunocitoquímica sobre las células estimuladas, usando anti-VEGFR2 y proteína-A conjugada con nanopartículas de Au sobre cortes nanométricos del cultivo. La proteína-A reconoce la fracción Fc de los anticuerpos y, por lo tanto, indica la presencia del mismo en una región particular de la célula. En la **Figura V.10** se pueden observar tres secciones representativas de cortes al azar de cultivo HUVEC adecuadamente tratado (ver sección V.2), donde las nanopartículas se encuentran muy distribuidas en el aparato de Golgi (**Figura V.10A**), pero al encontrarlas en los contactos célula-célula (**Figura V.10B**) o en la superficie celular de una célula aislada (**Figura V.10C**), fue notable confirmar la segregación de VEGFR2 en pequeños microdominios con una elevada resolución microscópica nunca antes lograda.

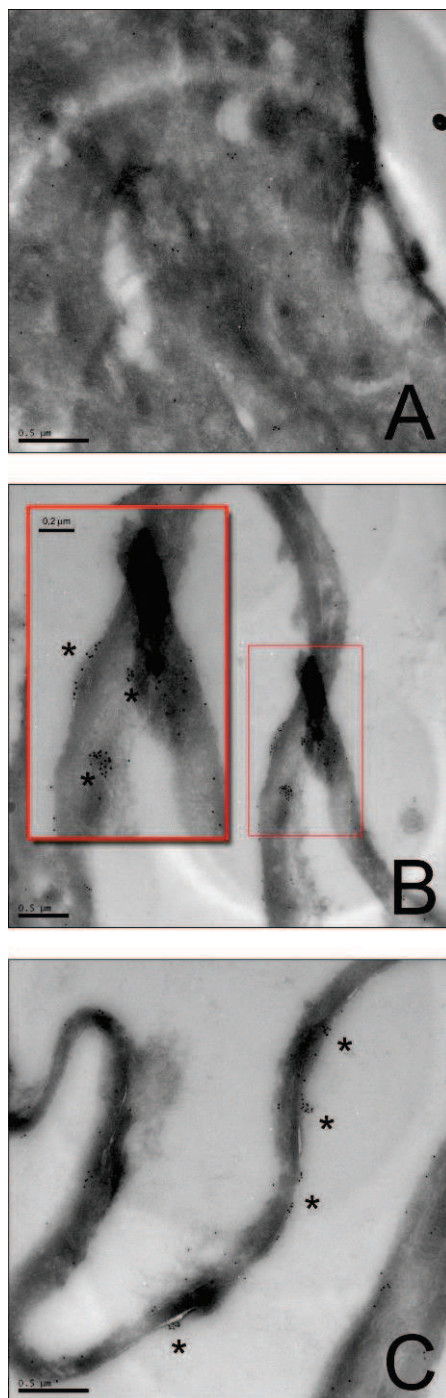


Figura V.10. Inmunocitoquímicas de cortes de HUVEC estimuladas con Gal-1. (A) Aparato de Golgi; (B) contacto célula-célula; (C) superficie de una célula aislada. Tamaño de nanopartícula de Au = 10 nm. Espesor del corte = 70-100 nm. Se agregaron asteriscos en las imágenes para guiar al lector hacia aquellas regiones en donde se observaron los microdominios de VEGFR2 agregados por la presencia de Gal-1. Barra: 0,5 μm.

Finalmente, se decidió recurrir al mayor aumento disponible en la actualidad, llevando estos mismos ensayos a un microscopio electrónico criogénico. Con éste tipo de equipos es posible obtener imágenes de resolución nanométrica de complejos moleculares *in situ*, en la membrana de la célula. Como la muestra debe ser depositada

sobre grids especiales, y el efecto que se deseaba observar requería la presencia de la célula intacta sobre estos grids, se requirió como paso previo comprobar que las células HUVEC cultivadas sobre grids de Au (los grids de Cu pueden ser tóxicos para las células) y estimuladas con Gal-1 mostraban el mismo fenómeno de agregación de VEGFR2 bajo estas nuevas condiciones de cultivo. Los resultados, a través de microscopía confocal de fluorescencia, se muestran en la **Figura V.11**. Fue posible observar que las células fueron capaces de crecer y ser estimuladas de la misma manera que en un plato de cultivo. Si bien no fue posible evidenciar de forma clara el efecto de agregación de VEGFR2 en las regiones de contacto célula-célula, en algunas células sí se encontraron los mismos patrones de segregación en microdominios del soma celular (**Figura V.11**).

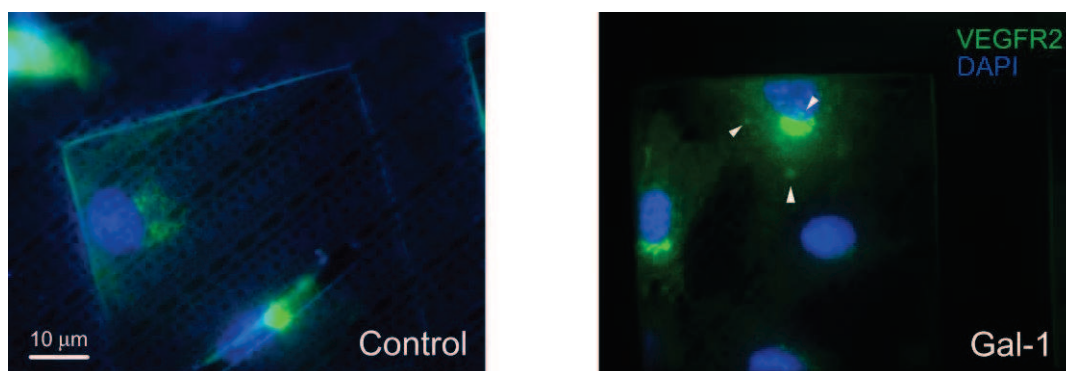


Figura V.11. Experimentos de inmunofluorescencia de HUVEC cultivadas en grids de Au. (A) Control; (B) Células estimuladas con Gal-1; se agregaron puntas de flecha en las imágenes para guiar al lector hacia aquellas regiones en donde se observaron los microdominios de VEGFR2 agregados por la presencia de Gal-1. Observar además los bordes metálicos de los grids y el film de carbono perforado que soporta las células.

El experimento final, y el más importante de todos, era congelar estas muestras fijadas en etano líquido (usando el equipo Vitrobot®, sección V.II). Esto se logró de manera satisfactoria en la mayoría de los casos, aunque en otros no fue posible, dado que requiere mucha práctica y manualidad. Lamentablemente, bajo las condiciones hasta aquí ensayadas no se pudieron obtener imágenes de buena calidad de estos microarreglos, debido al espesor de la célula endotelial. Tampoco se encontraron muchas células en contacto entre sí como para poder evaluar lo que sucede en la interfase célula-célula. Sólo se observaron algunos microfilamentos de actina (polímero de actina F de unos 10 nm de espesor, componente fundamental del citoesqueleto de todas las células eucariotas) y caveolas (del latín “pequeñas cuevas”, balsas lipídicas o invaginaciones de la membrana de unos 100 nm de diámetro) (**Figura V.12**). Dado que

estos estudios se llevaron a cabo durante una estadía de tres meses en los laboratorios del Dr. Subramaniam (NCI/NIH, Bethesda, Estados Unidos) gracias a una beca Fulbright-Bunge y Born, la continuidad de pruebas quedará supeditada a próximos viajes de colaboración con el grupo de microscopía electrónica de alta resolución y biología celular del NIH.

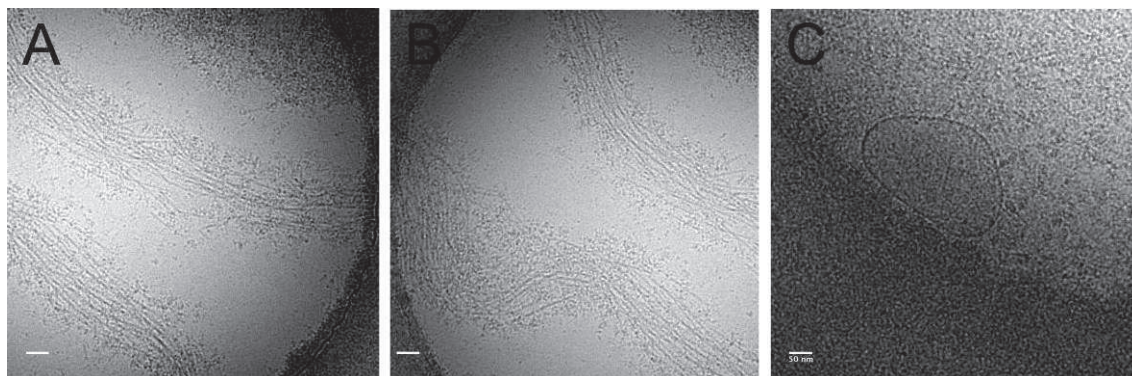


Figura V.12. Experimentos de microscopía crío-electrónica de células HUVEC cultivadas en grids de Au. A) y B) Muestran filamentos de actina y C) una típica estructura membranosa con un tamaño consistente con una caveola. Barra: 50 nm.

V.4 Discusión y conclusiones

En el capítulo presente, se exploró la capacidad de Gal-1 para inducir la agregación de un receptor en particular, el VEGFR2, en células endoteliales como modelo de estudio para un objetivo mucho más general: demostrar la formación de arreglos galectinas-glicoproteínas, actualmente denominados *lattices*.

Para ello, se comenzó por analizar el papel de determinados aminoácidos de los extremos terminales de Gal-1 en el proceso de dimerización de la proteína, requerimiento fundamental para lograr el entrecruzamiento de moléculas. Se ha descrito el bolsillo hidrofóbico en el cual yacen los aminoácidos no polares causantes de la fuerza impulsora de la dimerización, así como también se han puesto a prueba diversas mutantes que interrumpen dicho equilibrio de oligomerización. En base a nuestros resultados, las mutantes de Gal-1 C2S-V5D-V131D, A6D, A6D-V131D y A6D-V5D-V131D se presentaron como las mejores candidatas *in silico* a generar mutantes de Gal-1 incapaces de dimerizar (**Tabla V.1**). Como novedad al conocimiento actual, la consideración de la mutación V131D y el pasaje a la familia de mutantes a partir de A6D en lugar de la disponible en bibliografía (mGal-1, mutante C2S-V5D de Gal-1), son los resultados más importantes de dicho análisis computacional. Sumado a la

información generada en el Capítulo IV de este trabajo de Tesis, en el que asignamos un rol a Cys2 más relacionado con la oxidación que con la dimerización de la proteína, se puede decir que estamos ante una generación nueva de mutantes de Gal-1 que no dimerizan. Estas mutantes han sido diseñadas mucho más racionalmente a comparación de la actual mutante disponible, la cual los mismos autores de su descubrimiento han mencionado que conserva cierta capacidad de aglutinar células en suspensión, actividad únicamente posible si la proteína puede oligomerizar [42]. Por lo tanto, más allá del estudio básico de la relación estructura-función de dimerización, se ha podido ofrecer una nueva posibilidad de mutantes, que aguardan por ser producidas y estudiadas *in vitro* para comprobar los resultados hasta aquí obtenidos.

Si bien el objetivo más ambicioso del presente capítulo era la demostración mediante técnicas de alta resolución estructural, como la microscopía crío-electrónica, de la existencia de un arreglo supramolecular Gal-1-VEGFR2, lamentablemente las técnicas y ensayos realizados no permitieron arribar en forma completa con la verificación de dicho objetivo. Sin embargo, los resultados fueron más que alentadores. Se ha demostrado, en repetidas ocasiones durante el capítulo, la interacción del receptor 2 de VEGF con Gal-1 (**Figuras V.6 a 11**) y la agregación del mismo (**Figuras V.7 a 11**) *in situ* en la membrana de la célula mediante técnicas de microscopía. La evidencia por microscopía de fluorescencia confocal permitió la localización de estos arreglos moleculares en la membrana de la célula y también en los contactos célula-célula. Dado que en todos estos experimentos hay ausencia de ligando natural del receptor (VEGF), es inequívoco pensar que Gal-1 es la causante de este efecto a través de la unión de los *N*-glicanos que decoran el receptor (**Figura V.7**, ensayo en presencia de lactosa y resultados de [37]), aunque existe una batería de experimentos adicionales posibles para arribar a dicha conclusión de forma más certera (inhibir síntesis de determinados glicanos con siRNA, utilizar un tratamiento de glicosidasa PNGasa F, ensayar sobre células KO del gen *Mgat5*, testear estos efectos con una Gal-1 mutante incapaz de unir glicanos para descartar interacciones proteína-proteína). Más allá de haber trabajado en un sistema determinado, estos resultados sobre agregación de receptores mediados por lectinas aportan al *pool* de información disponible sobre el rol de lectinas y glicanos en general. Mucho se ha estudiado sobre el rol de las lectinas y la matriz extracelular y la migración [43], pero poco se sabe sobre la regulación que imponen estas rutas de comunicación celular a través de glicanos y lectinas. Este es, quizás, el comienzo de un nuevo camino a recorrer entre los paradigmas que la

glicobiología puede aportar al entendimiento global del funcionamiento de la comunicación celular.

Todo este trabajo se realizó casi totalmente durante la estadía del autor (tres meses) en los laboratorios del Dr. Sriram Subramaniam, del National Institute of Cancer (NIC) de los National Institutes of Health (NIH), Bethesda (MD), Estados Unidos, gracias a una beca de finalización doctoral Fulbright-Bunge y Born. El laboratorio del Dr. Subramaniam es uno de los laboratorios de más alto prestigio mundial en cuanto al manejo de técnicas de alta resolución al servicio de la determinación de estructuras biológicas. Si bien al establecer un proyecto acotado para la estadía se decidieron las estrategias aquí resumidas, en base a estos resultados se han planteado nuevas alternativas para esquivar los inconvenientes presentados y poder, al fin de cuentas, lograr una descripción nunca antes lograda de los *lattices*. Entre ellas, las posibilidades de atacar el problema incluyen la realización de cortes de las células estimuladas con Gal-1 de muestras embebidas en resinas plásticas (a diferencia del enfoque criogénico) y el empleo de otras técnicas como la tomografía FIB-SEM (del inglés, *focused ion beam scanning electron microscopy*), donde la muestra es escaneada luego de que un haz de iones ha sacado una fina capa de la superficie de la misma, repitiendo el proceso hasta lograr un volumen de muestra y poder reconstruir así una imagen tridimensional del sistema con resoluciones de hasta 5 nm. Todas estas técnicas están disponibles en el laboratorio donde se realizó la estadía y con el cual se ha establecido una colaboración. Sólo será cuestión de continuar en la dirección adecuada para poder cumplir con el objetivo principal.

V.5 Referencias

1. *Animal Lectins: A Functional View*. (2008) Editado por Gerardo R. Vasta and Hafiz Ahmed. CRC Press
2. Fred Brewer, C. (2002) Binding and cross-linking properties of galectins. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1572, 255-62
3. Morris, S., Ahmad, N., André, S., Kaltner, H., Gabius, H.-J., Brenowitz, M., and Brewer, F. (2004) Quaternary solution structures of galectins-1, -3, and -7. *Glycobiology* 14, 293-300
4. Cho, M. and Cummings, R.D. (1995) Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. I. Physical and chemical characterization. *Journal of Biological Chemistry* 270, 5198-206
5. Cho, M. and Cummings, R.D. (1996) Characterization of monomeric forms of galectin-1 generated by site-directed mutagenesis. *Biochemistry* 35, 13081-88
6. Bättig, P., Saudan, P., Gunde, T., and Bachmann, M.F. (2004) Enhanced apoptotic activity of a structurally optimized form of galectin-1. *Molecular Immunology* 41, 9-18
7. Miura, T., Takahashi, M., Horie, H., Kurushima, H., Tsuchimoto, D., Sakumi, K., and Nakabeppu, Y. (2004) Galectin-1[beta], a natural monomeric form of galectin-1 lacking its six

amino-terminal residues promotes axonal regeneration but not cell death. *Cell Death Differ* 11, 1076-83

8. Lobsanov, Y.D., Gitt, M.A., Leffler, H., Barondes, S.H., and Rini, J.M. **(1993)** X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-Å resolution. *Journal of Biological Chemistry* 268, 27034-38

9. Lopez-Lucendo, M.F., Solis, D., Andre, S., Hirabayashi, J., Kasai, K., Kaltner, H., Gabius, H.J., and Romero, A. **(2004)** Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *Journal of Molecular Biology* 343, 957-70

10. Birdsall, B., Feeney, J., Burdett, I.D.J., Bawumia, S., Barboni, E.A.M., and Hughes, R.C. **(2001)** NMR Solution Studies of Hamster Galectin-3 and Electron Microscopic Visualization of Surface-Adsorbed Complexes: Evidence for Interactions between the N- and C-Terminal Domains†. *Biochemistry* 40, 4859-66

11. Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S.H., and Rini, J.M. **(1998)** X-ray Crystal Structure of the Human Galectin-3 Carbohydrate Recognition Domain at 2.1-Å... Resolution. *Journal of Biological Chemistry* 273, 13047-52

12. Hsu, D.K., Zuberi, R.I., and Liu, F.T. **(1992)** Biochemical and biophysical characterization of human recombinant IgE-binding protein, an S-type animal lectin. *Journal of Biological Chemistry* 267, 14167-74

13. Ochieng, J., Platt, D., Tait, L., Hogan, V., Raz, T., Carmi, P., and Raz, A. **(1993)** Structure-function relationship of a recombinant human galactoside-binding protein. *Biochemistry* 32, 4455-60

14. Massa, S.M., Cooper, D.N.W., Leffler, H., and Barondes, S.H. **(1993)** L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity. *Biochemistry* 32, 260-67

15. Ahmad, N., Gabius, H., Andre, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F., and Brewer, C.F. **(2003)** Galectin-3 Precipitates as a Pentamer with Synthetic Multivalent Carbohydrates and Forms Heterogeneous Cross-linked Complexes. *Journal of Biological Chemistry*

16. Leonidas, D.D., Vatzaki, E.H., Vorum, H., Celis, J.E., Madsen, P., and Acharya, K.R. **(1998)** Structural Basis for the Recognition of Carbohydrates by Human Galectin-7. *Biochemistry* 37, 13930-40

17. Rini, J.M. and Lobsanov, Y.D. **(1999)** New animal lectin structures. *Current Opinion in Structural Biology* 9, 578-84

18. Ermakova, E., Miller, M.C., Nesmelova, I.V., López-Merino, L., Berbís, M.A., Nesmelov, Y., Tkachev, Y.V., Lagartera, L., Daragan, V.A., André, S., Cañada, F.J., Jiménez-Barbero, J., Solís, D., Gabius, H.-J., and Mayo, K.H. **(2013)** Lactose binding to human galectin-7 (p53-induced gene 1) induces long-range effects through the protein resulting in increased dimer stability and evidence for positive cooperativity. *Glycobiology* 23, 508-23

19. Stowell, S.R., Arthur, C.M., Slanina, K.A., Horton, J.R., Smith, D.F., and Cummings, R.D. **(2008)** Dimeric Galectin-8 Induces Phosphatidylserine Exposure in Leukocytes through Polylactosamine Recognition by the C-terminal Domain. *Journal of Biological Chemistry* 283, 20547-59

20. Earl, L.A., Bi, S., and Baum, L.G. **(2011)** Galectin multimerization and lattice formation are regulated by linker region structure. *Glycobiology* 21, 6-12

21. Gupta, D., Bhattacharyya, L., Fant, J., Macaluso, F., Sabesan, S., and Brewer, C. **(1994)** Observation of Unique Cross-Linked Lattices between Multiantennary Carbohydrates and Soybean Lectin. Presence of Pseudo-2-Fold Axes of Symmetry in Complex Type Carbohydrates. *Biochemistry* 33, 5614-22

22. Bhattacharyya, L., Khan, M.I., Fant, J., and Brewer, C.F. **(1989)** Formation of highly ordered cross-linked lattices between asparagine-linked oligosaccharides and lectins observed by electron microscopy. *Journal of Biological Chemistry* 264, 11543-45

23. Bhattacharyya, L. and Brewer, C.F. **(1992)** Formation of homogeneous carbohydrate-lectin cross-linked precipitates from mixtures of D-galactose/N-acetyl-D-galactosamine-specific lectins and multiantennary galactosyl carbohydrates. *European Journal of Biochemistry* 208, 179-85

24. Islam Khan, M., Mandal, D.K., and Brewer, C.F. (1991) Interactions of concanavalin A with glycoproteins. A quantitative precipitation study of concanavalin A with the soybean agglutinin. *Carbohydrate Research* 213, 69-77
25. Bhattacharyya, L., Khan, M.I., and Brewer, C.F. (1988) Interactions of concanavalin A with asparagine-linked glycopeptides: formation of homogeneous cross-linked lattices in mixed precipitation systems. *Biochemistry* 27, 8762-67
26. Bhattacharyya, L., Ceccarini, C., Lorenzoni, P., and Brewer, C.F. (1987) Concanavalin A interactions with asparagine-linked glycopeptides. Bivalency of high mannose and bisected hybrid type glycopeptides. *Journal of Biological Chemistry* 262, 1288-93
27. Brewer, C.F. (2001) Lectin Cross-Linking Interactions with Multivalent Carbohydrates. In *The Molecular Immunology of Complex Carbohydrates – 2* (Wu, A., ed), pp. 17-25, Springer US
28. Dam, T.K., Gerken, T.A., and Brewer, C.F. (2009) Thermodynamics of Multivalent Carbohydrate–Lectin Cross-Linking Interactions: Importance of Entropy in the Bind and Jump Mechanism. *Biochemistry* 48, 3822-27
29. Cheng, W., Bullitt, E., Bhattacharyya, L., Brewer, C.F., and Makowski, L. (1998) Electron Microscopy and X-ray Diffraction Studies of Lotus tetragonolobus A Isolectin Cross-linked with a Divalent Lewisx Oligosaccharide, an Oncofetal Antigen. *Journal of Biological Chemistry* 273, 35016-22
30. Olsen, L.R., Dessen, A., Gupta, D., Sabesan, S., Sacchettini, J.C., and Brewer, C.F. (1997) X-ray Crystallographic Studies of Unique Cross-Linked Lattices between Four Isomeric Biantennary Oligosaccharides and Soybean Agglutinin†,‡. *Biochemistry* 36, 15073-80
31. Garner, O.B. and Baum, L.G. (2008) Galectin-glycan lattices regulate cell-surface glycoprotein organization and signalling. *Biochem. Soc. Trans.* 36, 1472-77
32. Taylor, M.E. and Drickamer, K. (2007) Paradigms for glycan-binding receptors in cell adhesion. *Current Opinion in Cell Biology* 19, 572-77
33. Brewer, C.F., Miceli, M.C., and Baum, L.G. (2002) Clusters, bundles, arrays and lattices: novel mechanisms for lectin–saccharide-mediated cellular interactions. *Current Opinion in Structural Biology* 12, 616-23
34. Rabinovich, G.A., Toscano, M.A., Jackson, S.S., and Vasta, G.R. (2007) Functions of cell surface galectin-glycoprotein lattices. *Current Opinion in Structural Biology* 17, 513-20
35. Nieminen, J., Kuno, A., Hirabayashi, J., and Sato, S. (2007) Visualization of Galectin-3 Oligomerization on the Surface of Neutrophils and Endothelial Cells Using Fluorescence Resonance Energy Transfer. *Journal of Biological Chemistry* 282, 1374-83
36. Belardi, B., O'Donoghue, G.P., Smith, A.W., Groves, J.T., and Bertozzi, C.R. (2012) Investigating Cell Surface Galectin-Mediated Cross-Linking on Glycoengineered Cells. *Journal of the American Chemical Society* 134, 9549-52
37. Croci, Diego O., Cerliani, Juan P., Dalotto-Moreno, T., Méndez-Huergo, Santiago P., Mascanfroni, Ivan D., Dergan-Dylon, S., Toscano, Marta A., Caramelo, Julio J., García-Vallejo, Juan J., Ouyang, J., Mesri, Enrique A., Junttila, Melissa R., Bais, C., Shipp, Margaret A., Salatino, M., and Rabinovich, Gabriel A. (2014) Glycosylation-Dependent Lectin-Receptor Interactions Preserve Angiogenesis in Anti-VEGF Refractory Tumors. *Cell* 156, 744-58
38. Folkman, J. (2006) Angiogenesis. *Annual Review of Medicine* 57, 1-18
39. Carmeliet, P. (2005) Angiogenesis in life, disease and medicine. *Nature* 438, 932-36
40. Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W.E.G., Yagi, F., and Kasai, K.-i. (2002) Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1572, 232-54
41. Toscano, M., Bianco, G.A., Ilarregui, J.M., Croci, D.O., Correale, J., Hernandez, J.D., Zwirner, N.W., Poirier, F., Riley, E.M., Baum, L.G., and Rabinovich, G.A. (2007) Differential glycosylation of TH1, TH2 and TH-17 effector cells selectively regulates susceptibility to cell death. *Nature Immunology* 8, 825-34
42. Stowell, S.R., Cho, M., Feasley, C.L., Arthur, C.M., Song, X., Colucci, J.K., Karmakar, S., Mehta, P., Dias-Baruffi, M., McEver, R.P., and Cummings, R.D. (2009) Ligand reduces galectin-1 sensitivity to oxidative inactivation by enhancing dimer formation. *J Biol Chem* 284, 4989-99

-
43. Elola, M.T., Wolfenstein-Todel, C., Troncoso, M.F., Vasta, G.R., and Rabinovich, G.A. **(2007)** Galectins: matricellular glycan-binding proteins linking cell adhesion, migration and survival. *Cell Mol Life Sci* 64, 1679-700



CONCLUSIONES GENERALES

VI.1 Conclusiones generales

La **biología estructural** ofrece un abanico increíble de metodologías al servicio de la resolución de los más variados problemas con los que nos encontramos día a día en nuestra mesada de trabajo. Desde las herramientas de la química computacional y la termodinámica estadística, pasando por las que nos ofrece la biología molecular, siguiendo por las principales técnicas espectroscópicas, desde las más sencillas como la de absorbancia hasta las más complejas como la de RMN, y llegando hasta las microscopías de última generación, todo está disponible para la biología estructural. Si indagamos en cada uno de nuestros sistemas científicos particulares de interés, seguramente encontraremos, en no más de dos experimentos de distancia, uno que involucre alguna de estas técnicas para poder entender lo que sucede en nuestro problema de estudio. Es por ello que en ésta época, en donde las posibilidades de colaborar e interrelacionarse entre varios grupos con diferentes capacidades aseguran el éxito de la comprensión de un sistema y su funcionamiento, esquivar un enfoque multidisciplinario para poner a prueba las hipótesis de cualquier trabajo hubiera sido un gran error.

En este trabajo de Tesis se estudiaron a las galectinas humanas como una familia, utilizando en algunos capítulos a Gal-1 como miembro modelo y tratando de extrapolar los resultados a la comparación con el resto de los miembros de la familia proteica. En particular, se analizaron los determinantes moleculares más importantes que determinan un grupo de funciones bioquímicas de las galectinas: la unión a

glicanos específicos, la susceptibilidad frente a la oxidación y la capacidad de oligomerización.

En el Capítulo III se examinaron detalladamente los componentes estructurales que llevan a las galectinas a unirse en forma selectiva, aunque con similar afinidad, a un grupo de β -galactósidos biológicamente relevantes. Las simulaciones computacionales realizadas no sólo permitieron arrojar información clave sobre la estructura y la dinámica de cada miembro de la familia de galectinas, sino que además facilitaron la predicción y el entendimiento de la especificidad por ligandos mediante una metodología totalmente válida y extensible al estudio de otras familias de lectinas. Se ha demostrado que las galectinas poseen poca identidad de secuencia pero un plegamiento altamente conservado, en donde las principales diferencias se encontraron en las composiciones, longitudes y comportamientos dinámicos de los loops que interconectan las hebras β de cada estructura. Se ha sistematizado el estudio de la especificidad, primero estudiando los complejos experimentales galectina-glicanos disponibles y luego desarrollando una metodología basada en DM y cálculos de energía libre de interacción. Estos resultados permitieron rescatar las pocas tendencias experimentales disponibles y sirvieron además como puntapié inicial para la disección del rol de la unión a determinados azúcares, galectina por galectina, encontrándose similitudes globales como la baja afinidad por residuos α 2,6-sialidados y diferencias muy sutiles como entre los residuos sulfatados de LacNAc. En conclusión, el requerimiento de tantas galectinas humanas no parece estar asociado a la diversidad de plegamientos sino más bien a la especificidad de cada una de ellas por los glicanos a los cuales se puede unir, así como también a estudios que escapan a los objetivos de esta Tesis, como ser diferencias en los patrones de expresión, tanto a nivel tejido como a nivel sistémico.

En el Capítulo IV, dada la observación proveniente del capítulo anterior que pone en evidencia la diversidad en el patrón de aparición de residuos de cisteína entre las distintas galectinas, se decidió profundizar en éste fenómeno estructural en busca de diferencias o similitudes que pudieran ayudar a la comprensión de la reactividad redox de las galectinas. Mediante el uso de simulaciones de DM y una batería amplia de experimentos *in vitro*, se logró determinar el rol de cada cisteína en el proceso de oxidación de Gal-1 y extrapolar este comportamiento al resto de las galectinas

humanas. Se caracterizó la cinética de esta oxidación con H_2O_2 y se puso en contexto con las velocidades de reacción típicas con este oxidante, para poder dar interpretación al fenómeno de inactivación de la función de apoptosis de Gal-1 causada por la oxidación. En el proceso, se presentaron dos mutantes de Gal-1 capaces de resistir la oxidación y cumplir con la función de apoptosis en un ensayo *in vitro* con células T activadas con los mismos niveles de eficacia que la proteína wild-type.

Finalmente, y no por último menor, en el Capítulo V de resultados se presentaron una serie de estudios computacionales y experimentales cuyos objetivos focalizaban con el requerimiento de la oligomerización de las galectinas, además de la unión a glicanos, para poder llevar a cabo una determinada función. Se emplearon técnicas de simulación de DM para la caracterización de la región de dimerización en Gal-1. Estos estudios permitieron la proposición teórica de una nueva generación de mutantes que impiden éste equilibrio de oligomerización de forma más efectiva que la mutante actualmente disponible en la bibliografía. Por otro lado, se aprendieron y usaron técnicas de cultivo celular para poner a prueba, *in situ* en la superficie de las células, la formación de arreglos supramoleculares de Gal-1 y el receptor glicosilado VEGFR2, utilizando diferentes estrategias microscópicas, cada una con una resolución molecular cada vez mayor. A pesar de que los resultados obtenidos de las microscopías electrónicas criogénicas no fueron los esperados, los resultados arrojan evidencia fuerte sobre la capacidad de Gal-1 para segregar en determinados microdominios las moléculas de VEGFR2, tanto en el cuerpo de células endoteliales como en las regiones de contacto célula-célula. Y que esta capacidad se ve determinada por la capacidad de unión selectiva a glicanos presentes en la célula y de dimerizar a las concentraciones de Gal-1 ensayadas. En conjunto con los resultados del grupo, estos ensayos vienen a dar identidad estructural al paradigma de la glicobiología de los *lattices* como mecanismo principal de acción y activación de la función de un receptor, en ausencia de ligando natural y mediante mecanismos no-clásicos, dependientes de lectinas y glicanos.

Por lo tanto el presente trabajo de Tesis ofrece ese viaje desde el átomo a la célula, tal como se propuso en el título del mismo: a) **a nivel atómico-molecular** se encuentra una visión global y acabada de la estructura y dinámica de las galectinas humanas, así como también una interpretación más precisa de la interacción entre Gal-1 y N-glicanos complejos, oxidantes y ella misma en el proceso de dimerización; b)

desde un punto de vista celular se confirman los estudios estructurales trabajando *in situ* con la célula diana de todas las interacciones estudiadas previamente *in silico* y/o *in vitro*.

En resumen, la comunión entre los temas y las herramientas utilizadas en esta Tesis para poner a prueba la hipótesis general de que existe una conexión íntima entre estructura, dinámica y función, ha permitido el hallazgo de importantes tendencias y comportamientos para la comprensión de la función de las galectinas humanas como una familia, y que pueden ser ejemplo para un estudio particular y más profundo de una galectina de interés o de otras familias de proteínas, incluso sin el requerimiento previo de que se traten de lectinas. Porque al fin de cuentas, la biología estructural nos ofrece siempre un nuevo escalón para trepar la escalera de nuestros sistemas, en donde el techo es nuestra imaginación.

VI.2 Perspectivas

Claro está que, incluso al final de este trabajo de Tesis, no deja de sorprender que una proteína tan pequeña y tan promiscua en el organismo pueda ser portadora de tantos determinantes moleculares para llevar a cabo su abanico de funciones. Y esto se afirma con el conocimiento real de que han quedado fuera muchas otras propiedades sin analizar, por cuestiones de tiempo y enfoque. La estructura de Gal-1 oxidada (que es capaz de tener funciones totalmente diferentes a las de la proteína reducida), la síntesis de inhibidores selectivos, la capacidad de unión a ligandos a diferentes pH (dada la presencia de residuos de His en el LBG sensibles a estos cambios), las interacciones proteína-proteína (independientes de la unión a glicanos), el rol de la estructura y la función dentro de la célula, mecanismos de secreción no clásicos, el exótico caso de estudio de Gal-12 (aquella galectina que más diverge en secuencia y estructura de la familia), entre otros desafíos seguramente aún sin descubrir, serán motivos de nuevos descubrimientos en el área. En ese sentido, queda mucho por pensar, por investigar y por hacer. Espero que el esfuerzo grupal que representó este trabajo de Tesis sea útil en esa lista de cuestiones por investigar y sino, al menos, como fuente de inspiración para otros temas relacionados.

Como consecuencia de los trabajos de la oxidación de Gal-1, en el grupo se ha puesto en marcha el estudio de la capacidad de inactivación por oxidación de Gal-1 y las mutantes propuestas en este trabajo de Tesis en modelos murinos *in vivo*. Definitivamente, las perspectivas de este trabajo podrán resolver inquietudes del área de la galectinómica como ser la regulación de la respuesta inmune innata y adaptativa en microambientes inflamatorios oxidantes mediante galectinas endógenas y sus interacciones con glicanos.

Los logros parciales respecto de la exploración de la hipótesis de la formación de *lattices* galectinas-glicoproteínas nos han potenciado en la búsqueda de herramientas complementarias para arribar a la descripción molecular de estas estructuras de manera directa y definitiva. En ese sentido, algunas opciones se han discutido en la sección V.4 y otras aparecerán con el correr de los experimentos. Pero sin lugar a dudas, un trabajo cuyas respuestas generan nuevas preguntas, es el terreno más adecuado y alentador posible para continuar explorando las hipótesis que sacien nuestra sed de curiosidad y motivación para el desarrollo de terapias cuyo éxito dependan de la modulación de este tipo de estructuras.



PUBLICACIONES

Publicaciones que surgieron de resultados de este trabajo de Tesis

- *Structural basis of redox-dependent modulation of galectin-1 dynamics and function.* (2014) **C.M. Guardia**, J.J. Caramelo, M. Trujillo, S. Mendez-Huergo, R. Radi, D.A. Estrin y G.A. Rabinovich. *Glycobiology*. 24(5), 428-41. DOI: 10.1093/glycob/cwu008.
- *An integrated computational analysis of structure, dynamics and ligand binding interaction of the human galectin network.* (2011) **C.M. Guardia**, D.F. Gauto, S. Di Lella, G.A. Rabinovich, M.A. Martí y D.A. Estrin. *J. Chem. Inf. Model.* 51 (8), 1918-30. DOI: 10.1021/ci200180h. **Tapa de revista.**

Publicaciones como co-autor durante el transcurso del doctorado

- *Thiol redox biochemistry: Insights from computer simulations.* (2014) A. Zeida, **C.M. Guardia**, P. Lichtig, L.L. Perissinotti, L. De Felipe, A. Turjanski, R. Radi, M. Trujillo y D.A. Estrin. *Biophysical Reviews*. DOI: 10.1007/s12551-013-0127-x.
- *When galectins recognize glycans: from biochemistry to physiology and back again.* (2011) S. Di Lella, V. Sundblad, J.P. Cerliani, **C.M. Guardia**, D.A. Estrin, G.R. Vasta y G.A. Rabinovich. *Biochemistry* 50 (37), 7842-57. DOI: 10.1021/bi201121m.
- *Linking Structure and Thermal Stability of the β -galactosidase binding protein galectin-1 to ligand binding and dimerization equilibria.* (2010) S. Di Lella, M.A. Martí, D. Croci, **C.M. Guardia**, J.C. Diaz Ricci, G.A. Rabinovich, J.J. Caramelo y D.A. Estrin. *Biochemistry* 49(35), 7652-58. DOI: 10.1021/bi100356g.
- *Carbohydrate-binding proteins: Dissecting ligand structures through solvent environment occupancy.* (2009) D.F. Gauto, S. Di Lella, **C.M. Guardia**, D.A. Estrin y M.A. Martí. *J. Phys. Chem. B* 113(25): 8717-24. DOI: 10.1021/jp901196n.

VIII

AGRADECIMIENTOS

Quisiera agradecer en primer lugar a las entidades que han financiado de una u otra forma el trabajo y mi formación doctoral. Agradezco al CONICET, a la ANPCYT y a la querida FCEyN-UBA. A las fuentes de financiamiento internacionales (CeBEM y Fundación Fulbright) y a la Fundación Bunge y Born les agradezco por el apoyo a la realización de mis experiencias en el extranjero.

Me siento increíblemente agradecido a la vida por haberme permitido el honor de ser “criado” por los dos padres científicos más extraordinarios del mundo. No sólo porque cada uno a su manera me mostró el mundo de la biología estructural, sino porque estamos hablando de personas sensacionales. **Darío**, con su humildad, generosidad y pasión por los desafíos, me ha enseñado a trabajar con respeto y total libertad, inyectándome energía día a día, no sólo en el grupo sino también en las clases que hemos compartido. Un ejemplo de padre y maestro. Siempre recordaré mis primeros pasos en tu grupo, y por ello te agradezco un montón. Y **Gaby**, que llegó a mi primera esfera de coordinación de la mano de Santiago Di Lella, me abrió las puertas de su mundo, sin juzgarme, sin preguntarme nada. Sólo bastaba con tener ganas de trabajar y aprender. En el camino descubrí a uno de los científicos más grandes de la actualidad, que no tiene horarios ni excusas para hacer lo que le gusta por lo que quiere y que siempre estuvo para lo que necesité, tanto a nivel laboral como personal. Me enseñó lo que es la pasión por lo que uno hace y la sensación de estar en familia, invitándome a conocer a los suyos en el IBYME y descubriéndolos míos al final de esta etapa. Por el ejemplo, por la dirección, por el impulso hacia la superación constante y sostenida, y por la amistad sincera y desinteresada, muchas gracias a mis directores de Tesis.

En el proceso de decidirme por el tipo de Tesis que quería desarrollar, **Julio Candy** ha sido clave y quería agradecerle por todas esas horas de clase de QBIIB y todas esas otras de trabajo (y diversión) en la FIL. Te aprecio como docente, como persona, como científico y has sido para mí una inspiración y un soporte en el camino, tanto que te considero mi tercer director, a pesar de que los papeles no lo reflejen. ¡Gracias por todo!

Quisiera también agradecerle al Jurado de esta Tesis, que se ha tomado el tiempo y la dedicación para corregir y criticar nuestro trabajo y enriquecerlo para aportar aún más a mi formación y al conocimiento generado a lo largo de estas páginas.

A mis **compañeros del GMM**, que me han soportado con mi humor (y malhumor) y que de cada uno de ellos me llevo lo mejor. En especial, a **Diego G.**, el hermano más grande que nunca tuve y que la vida me regaló. Sos un amigazo, un ejemplo y un soporte constante. Desde los primeros tiempos en la “cocinita”, las largas charlas de discusiones científicas y no tan científicas, hasta la bienvenida que me diste a tu familia (comenzando por Cynthia y regalándole al mundo al potro de Bauti) y todos esos consejos, sueños y miedos que compartimos en estos últimos años, fueron lo mejor que pasó durante mi estadía en el GMM. Aunque también hacerme amigo íntimo de mi ex alumna **Lucía** también ocupa un lugar en estos agradecimientos. Siempre feliz y con energía para sacar los problemas adelante, y organizando todo para la felicidad del grupo. Una amiga que veo crecer día a día y me llena de orgullo y alegría. A la legión del interior (**JuanPa, Vicky, Clau, Gaby Jara y Ariel P.**) y al miembro único de la legión del exterior (**Ari Z.**), les quería agradecer también por llenarme de todo un vocabulario nuevo y por la amistad y afecto que me han brindado todos estos años. Son unos genios y son geniales, ¡les agradezco por venir a revolucionar el GMM! A los que me recibieron allá por el 2006: **Damián Bikiel y Nano**, han sido los creadores del monstruo Charly, así que, háganse cargo. Jamás olvidaré mi primer *input* de Gaussian con Damián ni mi primer QM/MM con Nano. A **Marce, Santi, Lula, Leo, Adrián T., Lau y Pablo, Pau, Ithu, Vero, Eze, Estefi y Damián Scherlis** por haberme recibido, por seguir construyendo un grupo día a día cada vez mejor y para aquellos que emigraron, por haberlo dejado en condiciones para seguir haciendo cada vez mejor ciencia y mejores científicos. Al peronista amigo de **Adrián Roitberg**, que me bancó un trimestre en su grupo en el extranjero y que siempre está presente en el grupo local y que sin su conexión íntima con AMBER, la parte

computacional de este trabajo (y miles de otros en el resto del mundo) no hubiera sido posible. Al genio de **Lucas** que siempre atendió mis dudas computacionales y alguna que otra discusión administrativa y amorosa de esas que la vida te pone enfrente y que son más difíciles que poner a funcionar el cluster después de un pico de tensión. A los que se han ido incorporando y ahora tienen la posta: **Juan A., Diego H., Facto, Nico F., Fer, Uri, Yami, Francisco, Juan R., Pablo L., Leandro G., Juan B., Lean Radusky, Elías.**

A mis **amigos del IBYME**, los ami-biólogos del grupo de Gaby que me ayudaron a darle sentido a las dinámicas moleculares y permitieron que mis inquietudes explotaran en forma exponencial. En especial, agradecerle al genio de **Panza** (Diego C.) que me enseñó todo lo que sé de expresión y purificación de Gal-1 y en algún momento espero publicar un *Cell* como él. Al **Pelado** (Juan P.C.) por acoplarse después en el entrenamiento de la purificación y hacer que me den más ganas de ir al labo. A **Santi M.H.** por acompañarme a trabajar con células y recordarme que la paciencia con vos existe! A mi *segundo* corazoncito en el grupo, **Sebas D.D.**, por ser tan perseverante y puro y por compartir tantas charlas de ciencia, evolución, parásitos y sexología berreta. A **Tommy** y **Vero** por enseñarme a hacer un PAGE y un Western Blot. A **Rosa**, la mejor bioterista y la mina más copada, con quién aprendí a genotipificar (ratones) y fenotipificar (humanos). A **JuanK** por ser compinche y compartir con nosotros a Gaby. A los pilares femeninos del grupo: **Marta** y **Marian**. A los más nuevitos (y no tanto) pero igualmente queridos de **Ada, Nicolino, Sebas M., Lucho, Flor, Anabela, Marcos, Pablo** y **Claudia**. Y finalmente, como parte de este grupo, pero en Exactas, gracias a **Diego L., Daniel C., Lucas, Felipe, Nacho** y **Geraldine**.

Luego de tres visitas a Uruguay, aprendí muchísimo sobre cinética de oxidación y sobre otros valores científicos y humanos en el grupo CEINBIO de la UdelaR a cargo de **Rafa** (Radi). A él, a **Madia, Martín** y **Silvina**, mi eterno agradecimiento por lo que han contribuido en mi formación, muchas veces con un valor agregado que claramente excede el título de colaboradores.

A los Departamentos de Inorgánica, Analítica y Química Física y de Química Biológica de la FCEyN, donde he saciado mis necesidades vocacionales de docente (lejos de considerarlo un mero trabajo). Allí conocí excelentes JTP y Profesores que me enseñaron desde la misma vereda y excelentes alumnos que me impulsaron, cuatrimestre a cuatrimestre, a superarme. Gracias a mis casi ocho años ininterrumpidos de ayudante conocí a quienes hoy son grandes amigos, que sabiendo de qué se trata

este estilo de vida, me acompañaron en las buenas y en las malas, festejando y sufriendo conmigo. **Paula Monserrat C.** (la luz de mis ojos), **Yami A.**, **Yami J.** y **Days**, hoy mi familia de Facebook, dan testimonio de ello. Las quiero como son y por lo que son y espero que podamos pasar muchos años y reactivos nuevos más, todos juntos, en familia. **Analía**, **Nani**, **Betty**, **Euge**, **María Florencia B.** y **Olga**, docentes mías y hoy compañeras de aventuras y ejemplos de perseverancia y amor por la familia y su trabajo. **Virginia**, con quien hemos compartido tantas noches oscuras de trabajo y tecitos, de pasillos y risas por el entrepiso del DQIAQF, de congresos alocados y anécdotas imposibles de borrar. Sos mi madre científica, de fierro y pecho resistente a cualquier bala, de esas mujeres que te ayudan y enseñan a resistir y a luchar haciendo uso del mejor ejercicio posible: el ejemplo. Te quiero y le agradezco a la Tesis por haberte conocido.

A **Ceci Sp**, que te conozco desde casi el comienzo de la carrera pero que te amo desde la explosión del Universo, te quería agradecer especialmente por escucharme, por aconsejarme, por criticarme, por odiarme, por discutirme, por apachucharme, por arrastrarme a Francés, por cocinar en mis cumpleaños, por estar...Ah, y por corregir esta Tesis como ninguna otra persona. Te adoro.

A **Guiy**, que no recuerdo el momento exacto en que nos conocimos, pero que sí recuerdo que fue amor a primera vista. Te debo muchísimas cosas en mi vida (en especial, mi futura sonrisa Colgate y el cuerpo esculpido como el de Lorenzo Lamas), principalmente gracias a tus consejos, que no faltaron en ningún momento durante todo este parto y a esa energía positiva constante que te impulsa hacia delante y arriba. Hoy miro hacia atrás y hemos pasado cosas lindas y cosas feas, nos hemos ausentado y nos hemos asfixiado, pero mi vida no sería la misma sin que vos estuvieras a mi lado compartiéndola. Te admiro y te quiero como a pocos.

A mis amigos nacidos de la facu y que logramos mantenernos unidos a pesar de los cambios por los que la vida nos lleva a cada uno (por orden alfabético, así no se me ofenden, a cada uno lo quiero por igual): Akemi, Cagno, Cante, Ceco, Funes, Gi, Iñaqui, Javi, Jota, Kuky, Lau, Lean, Leti, Lu, Lucía C., Lucía F., Maru, Nico, Romi, Sil, Sol, Vani y Walter.

A los amigos que me dieron los lugares de trabajo, los cursos y congresos, que hicieron más ameno el proceso de aprendizaje y que con compartir sus experiencias, enriquecieron la mía: Migue, Ana Larralde, Elcho, Caro, Sole, Ceci Sorbello, Reche, Ana Tufo, Néstor F., Beatriz Aráoz, Joe Hodak, Noelia L.Z., Kary y Juan y su progeñie,

Paula Vena, Rodrigo Corti, Diana C., Florenza, Guille B., Andrea Z., Flor Molinas, Nacho F. y esposo, Laura G., Magda, Sabri, Tomás F., Nacho S., Cristian D., Juan P., Ceci V., Meli V., Silvina M., Vanesa M.V., Nico de F., Gi Miana, Andrés R., Bombicino, Ramiro Orsi, Soledad L., Carlos L., Lucas L., Máximo L.M., Lara y Lucas, Carito Domaica, Sherman, Merce F., Marina R., Juan Carlos Calvo, Anita Fiszbein, Rob Pozner, Cintia Berón, Nilda D., Julián F., Gaya, Lau y Aye, Lucas Franchina y especialmente a José Silveira, Cin Dominguez e Ileana D'Allura.

A las secretarias del DQIAQF, que sin ellas todo esto hubiera sido imposible (facturas, pasajes, viajes, informes, licencias, concursos, llaves, reservas ¡puff!): Alejandra, Elizabeth, Lili, Claudia y Cintia.

A mis amigos de Villasoles, con los que aprendí no sólo la LSA, sino también el respeto por la comunidad sorda, la tolerancia y la integración y las ganas de luchar por ellas desde otro espacio. Gracias a Mariana, Pao, Yami, las Natalias, las Vanesas, Javi, Rosana, Paula, Loreana, July y Yani y a todos los profes por ayudarme todos los sábados a crecer como persona y aportar al cable a tierra que resultaban nuestros encuentros.

A la banda de Always Club, que permitieron que toda la estantería donde reposa mi cerebro no se caiga a pedazos y contribuyeron a cambiar mi humor a través de la actividad física: JuanI, Cata, Vale, May, Flor, Sil, Dami, Pablo P, Helene y Georgina.

Un especial agradecimiento a Noelia Canabal por bancarme con su violín en los encuentros mágicos de música que teníamos esporádicamente y que permitieron alimentar no sólo mi alma dura y llena de ecuaciones y protocolos sino también nuestra más profunda y sincera amistad.

A los locos de BET en Facebook, que hicieron de compañía y sinergia (y permitieron también un poco de procrastinación) durante muchas horas-silla-culo enfrente de la computadora.

I would like to sincerely thank Dr. Subramaniam and all his great team for giving me the opportunity to work at NIH. Without his guidance and selfless commitment, I would never have achieved the Fulbright scholarship and learned all you taught me about cell biology and electron microscopy. Special thanks to Lisa, Mario, Lesley and Giovanna, who were the ones that worked with me on the bench during my stay abroad. Thank you very much.

Finalmente, y quizás lo más importante, quería agradecerle (además de dedicarles esta Tesis) a mis padres y mi familia inmediata (mi hermana, mis tías abuelas y mis tíos políticos), porque gracias a ellos es que me encuentro donde estoy. Porque cuando todo lo demás falla, mi familia siempre está, sin margen de error. Fueron mi primer paso, mi primera escuela, mi primera fórmula, mi primer título y lo seguirán siendo para siempre. Gracias por el más puro amor y el mejor de los ejemplos y todos esos sacrificios que hicieron para que hoy pueda cumplir un sueño más en la vida que me dieron. Ojala se sientan tan orgullosos de mis logros como yo me siento de formar parte de la familia que construimos.

Fe de erratas

Pág. 7

- Donde dice: “biosínTesis”, debe decir: “biosíntesis”.
- Donde dice: “sínTesis”, debe decir: “síntesis”.
- Donde dice: “nucléotido dador”, debe decir: “azúcar-nucléotido dador”.
- Donde dice: “(proteómica)”, debe decir: “(proteoma)”.
- Donde dice: “(glicómica)”, debe decir: “(glicoma)”.

Pág. 8

- Donde dice: “ácido siálico”, debe decir: “ácidos siálicos”.
- Donde dice: “glicosilación en mamíferos”, debe decir: “glicosilación de proteínas en mamíferos”.
- El apéndice de la **Figura I.3** debe ser:

Figura I.3. Biosíntesis y diversidad estructural de N-glicanos. Se esquematizan algunos ejemplos particulares de N-glicanos de alta manosa, híbridos y complejos. A su vez se detallan las distintas enzimas que participan en la síntesis. Las flechas rojas indican posibles zonas de extensión. Adviértase que en un N-glicano determinado no se encuentran todos los productos de las enzimas detalladas, ya que algunas compiten por un mismo sustrato. Dol-P, dolicol-fosfato; GnT1, 2, 3 y 5, GlcNAc transferasas; FuT, fucosil transferasas; GalT, galactosil transferasas; ST6Gal, α 2,6-sialil transferasas; ST3Gal, α 2,3-sialil transferasas [24].

Pág. 9

- Donde dice: “(fertilización, embriogénesis y reconocimiento de patógenos)”, debe decir: “(fertilización, embriogénesis, reconocimiento de patógenos, etc)”.

Pág. 10

- **Tabla I.1.** Fila correspondiente al año 1919. Donde dice: “primerea”, debe decir: “primera”.

Pág. 15

- Donde dice: “previos a su secreción”, debe decir: “previos a su excreción”.

Pág. 76

- Donde dice: “van der Walls”, debe decir: “van der Waals”.