



Tesis Doctoral

Estudios estructurales y de interacción de miméticos de Condroitín Sulfato con los factores de crecimiento Midkina y Pleiotrofina

Memoria presentada para aspirar al grado
de Doctora en Química por

Myriam Torres Rico

Sevilla, 2021

Dr. Pedro Manuel Nieto Mesa, Investigador Científico del CSIC y Dr. Jose Luis de Paz Carrera, Científico Titular del CSIC

CERTIFICAN

Que el trabajo titulado *Estudios estructurales y de interacción de miméticos de Condroitín sulfato con los factores de crecimiento Midkina y Pleiotrofina*, se ha llevado a cabo bajo mi dirección en los laboratorios del Instituto de Investigaciones Químicas (IIQ, CSIC-US) del Centro de Investigaciones Científicas Isla de la Cartuja (icCartuja) Sevilla, constituyendo la memoria presentada por Myriam Torres Rico para optar al título de Doctora en Química.

Sevilla, Septiembre 2021

El director

El codirector

Dr. Pedro Manuel Nieto Mesa

Dr. Jose Luis de Paz Carrera

El trabajo de investigación recogido en la presente memoria se ha llevado a cabo fundamentalmente en el Instituto de Investigaciones Químicas (IIQ), que forma parte del Centro de Investigaciones Científicas Isla de la Cartuja, centro mixto del Consejo Superior de Investigaciones Científicas (CSIC) y la Universidad de Sevilla (US). El doctorado ha sido financiado mediante una beca predoctoral del Ministerio de Educación y Ciencia BES-2016-078322 durante 2017-2021.

Dicha beca predoctoral también ha financiado una estancia breve en Universidad de Barcelona, bajo la supervisión de Carme Rovira.

*“Y se abrió un telón desafiando el final,
y en esa brecha de luz, vuelve a bailar...”*

Vetusta Morla

Agradecimientos

En primer lugar, quiero expresar mi agradecimiento a mi director de Tesis Pedro M. Nieto por darme la oportunidad de llevar a cabo esta etapa tan enriquecedora y de tanto aprendizaje, además de resolverme las incontables dudas sobre la RMN. A mi codirector Jose Luis de Paz por su entera disponibilidad siempre que he necesitado cualquier tipo de ayuda, aunque no fuese su campo, y facilitarme tanto esta última etapa de la tesis. También a Javier Rojo por su trato siempre tan cercano conmigo y a Maribel por tutorizar mi tesis y su disponibilidad en todo momento.

Agradecer a Carme Rovira y a todo su grupo por hacerme sentir tan cómoda, en especial a Joan y a Alba, por su ayuda y paciencia en mis frustraciones computacionales. Conseguísteis sin duda, no sólo que me empezase a gustar la química teórica, sino que me enamorase aún más de Barcelona.

A mis compañeros de laboratorio por cuidarme y tratarme tan bien en todo momento: a Pedro y a Antonio, mi primera toma de contacto en Sevilla, habéis sido como dos hermanos para mí, gracias por hacerme tan fácil los comienzos, por haberme acompañado todo el camino (y haberos comido la mitad de mis mudanzas). Sin duda me llevo un montón de buenos momentos con vosotros. A Jorge, por las buenas conversaciones y porque siempre he sabido que he podido contar contigo para lo que fuese. A Noelia...me falta vida para agradecerte que no me hayan echado ya del programa de doctorado por olvidarme del papeleo, además de tu preocupación y tus buenos consejos cuando más se necesitaban. A Adrián por tu ternura y amabilidad con la que siempre me has tratado, te deseo mucha suerte en lo que te queda de tesis. A Javi por su alegría y vitalidad tan necesaria en el tramo final de esta etapa. Y finalmente a M^a José, por hacerme los comienzos computacionales tan fáciles, fuiste clave para empezar a dar los primeros pasitos por este mundillo tan complejo.

Al resto de mi familia Cartujana, los grandes responsables de haber hecho inolvidables estos casi 5 años por esta tierra y que espero que sean muchos más allá donde acabemos: a mi titi Laura, por tu sabiduría, tu personalidad implacable y por haber estado ahí siempre que lo he necesitado. A Crubs, mi mayor descubrimiento cartujano, *moltes gràcies per tantes i tantes rialles*, ojalá una etapa *post*-doctoral en la misma ciudad yendo a monólogos y quedadas por Madrit sin plantones. A mi bebesita Elena, gracias por aportarme tanta calma y paz en todo momento, y estar codo con codo conmigo en todas mis locuras. A mis golditos favoritos, a Alberto por tu exquisito sentido del humor que tan buenos momentos nos has hecho pasar y a mi *caprachosa* Tania, por nuestros piques (siempre desde el cariño), tus consejos, y sobre todo por ser la primera en proponerme un viernes random de Eureka que desencadenó todos los buenos recuerdos que me llevo en la maleta. ¡¡ Espero con ansias vuestra boda!! A Inma, por toda tu ayuda, preocupación y no dejar que me rindiese cuando toqué fondo en la tesis, eres maravillosa. A Carlos, por tus ánimos, confianza y lo agusto que siempre me hiciste sentir. A Patri, por tener siempre una palabra de consuelo y dar un achuchón cuando se necesitaba (cuando se podían dar).

A Carmen, por las buenas risas echadas en tu azotea cuando el confinamiento no permitía mucho más. A Práxedes por sus “enseñanzas” bailando sevillanas. Y por último, no menos importante, a mi Rossmery, mi compañero de sevillanas, qué buenos ratos me has hecho pasar y cuánta cultura musical y televisiva me has descubierto!

No me olvido tampoco del resto de compañeros de Cartuja, entre muchos otros, Noel Irene, Laura, Jon, Diego, Mari, Alejandro, Christian, Pilar, Anabel, ...con los que he compartido buenos momentos y siempre han tenido unas palabras de ánimo durante la tesis. También merecen un especial reconocimiento todo el grupo de informática del IIQ por haber solucionado mis innumerables peleas con el ordenador. A los conserjes por su energía mañanera y rescatar todas las veces que he perdido mi tarjeta de acceso. A Javi López, Marga y Fran por su trato siempre tan agradable y su entera disponibilidad siempre que se necesita.

Tampoco puedo dejarme atrás al resto de la gente que he conocido durante esta etapa y con los que he compartido momentos magníficos: a Fran, mi sucesor computacional y con el que me río de todas las desgracias de nuestros cálculos teóricos. A Lola y mi ex compañera perruna Noa, siempre era un gusto llegar a casa y pasar un rato con vosotras. A mi vecinillo Fran, por tu gran amabilidad y disponibilidad para todo. A Ana, por ser la que sostenía mi paciencia durante los duros meses de confinamiento y la alegría en esos monótonos días. A mis entrenadores de Boxeo, por animarme los únicos momentos de desfogue durante el encierro de escritura y hacer el cardio tan llevadero y divertido.

Dejando un lado a mi familia andaluza, no puedo no agradecer a los que siempre han estado y siguen estando ahí año tras año, mi familia madrileña: a mi gente del Fortuny/Chamberí, más de media vida al pie del cañón y seguís siendo mis imprescindibles siempre que subo a Madrid, sobre todo mis señorinas, Rut y Cy. Mis Quimicuchos, tanto mis niñas como mis niños, es un placer ver como vamos creciendo tanto profesional como personalmente en nuestra carrera científica y que sigamos estando ahí los unos para los otros a pesar de la distancia. A mis chicas de Torrejón, en especial a mi vecinilla Sil, gracias por tantos buenos recuerdos juntas, por hacer mi estancia en Barcelona inolvidable y estar siempre tan presente a pesar de los miles de km que nos separan. A Víctor, mi compañero de batallas, siempre riéndonos de los problemas con unos buenos vinos por Malasaña. Ojalá volver a compartir mesa contigo en Lilly.

Y agradecer en último lugar, pero a la vez el más importante, a mi familia. Que sin saber muy bien de que iba lo que estaba haciendo, siempre han intentado entenderme; brindándome el máximo apoyo y haciéndome sentir que podía con todo, dándome el empujón que necesitaba para salir a flote. La educación y los valores que me habéis aportado durante toda mi vida, me han hecho ser la mujer que soy a día de hoy y no puedo estar mas agradecida.

Resumen:

El Condroitín Sulfato es el glicosaminoglicano (GAG) más abundante en la matriz extracelular, cuyo patrón de sulfatación da lugar a diferentes estructuras que le confiere cierta afinidad con diversas proteínas, como factores de crecimiento, desencadenando distintos roles biológicos. El trabajo presentado en esta Tesis doctoral ha sido principalmente el estudio de interacción entre miméticos de Condroitín Sulfato (CS) con los factores de crecimiento de unión a Heparina, las citoquinas Midkina (MK) y Pleiotrofina (PTN). Dependiendo del código de sulfatación del CS se encuentran diferentes tipos, siendo el derivado de tipo E con unidades de Glc y GalNAc, el que establece uniones fuertes con este grupo de proteínas y con el que se ha trabajado en esta Tesis doctoral. Por lo tanto, los diferentes miméticos estudiados se engloban en dos series: tres ligandos tetrasacáridos y tres pentasacáridos de CS-E con diferentes patrones de sustitución, con los que posteriormente se desarrollarán los estudios de interacción con estas dos citoquinas.

En primer lugar, se ha determinado la estructura de los seis miméticos por RMN utilizando principalmente la técnica 2D-NOESY, complementada con 2D-COSY, 2D-TOCSY y 2D-HSQC. Estos datos experimentales acerca de las conformaciones de los anillos y las orientaciones de los enlaces interglicosídicos han sido contrastadas paralelamente con cálculos computacionales de dinámica molecular, observándose que el patrón de sustitución en el ligando dificultaba su reproducibilidad teórica.

La segunda parte de este proyecto consiste en estudiar la afinidad de estos miméticos con ambas proteínas observando si se establece dicha unión y que partes del carbohidrato y de la citoquina se encuentran involucradas. Mediante los experimentos de STD (*Saturation Transfer Diffusion*) y NOESY transferido se demostró que existía interacción deslocalizada por todo el esqueleto del oligosacárido, además de observarse que la estructura de los ligandos en su forma libre y enlazada no variaba. Paralelamente, la técnica de Polarización de Fluorescencia (FP) avaló esta interacción demostrando además que la presencia de sistemas aromáticos en el esqueleto del carbohidrato la aumentaba. Una vez observado el comportamiento del ligando en su forma complejada, nos planteamos estudiar dicha interacción desde la proteína. Para ello el experimento ¹⁵N-HSQC nos permitió conocer los aminoácidos más involucrados en esta interacción, realizando el experimento con el ligando más sencillo unido a la PTN.

Por otro lado, se llevaron a cabo estudios de interacción teóricos para profundizar en el conocimiento de esta unión tanto con la MK como con la PTN. El *docking* molecular nos permitió ampliar las zonas de afinidad simulando otros bolsillos de interacción confirmando por otra parte las diferentes orientaciones que pueden adquirir los ligandos hacia estas citoquinas. Una vez conocida la estructura molecular de los seis miméticos en la primera parte del trabajo y la zona de mayor afinidad con la proteína, se amplían los cálculos teóricos al estudio del complejo para reproducir su estructura de nuevo mediante simulaciones.

Por último, las aproximaciones MM/PBSA y MM/GBSA (*Molecular Mechanics with Poisson-Boltzmann (PB) or Generalized Born (GB) and Surface Area Solvation*) permiten determinar los cambios de energía libre de unión de nuestros miméticos a estas proteínas conociendo además la participación de los aminoácidos a partir de simulaciones mediante modelos de solvatación.

Abreviaturas:

AMBER	<i>Assisted Model Building with Energy Refinement</i>
CESGA	<i>Centro de Supercomputación de Galicia</i>
CG	<i>sp³ C aliphatic (GLYCAM06 Parameters)</i>
CHARMM	<i>Chemistry at HARwar Macromolecular Mechanics</i>
COSY	<i>¹H-¹H COrrrelation SpectroscopY</i>
CUDA	<i>Compute Unified Device Architecture</i>
CS	<i>Chondroitin Sulfate / Condroitín Sulfato o Sulfato de Condroitina</i>
DS	<i>Dermatan Sulfate / Dermatan Sulfato o Sulfato de Dermatano</i>
FID	<i>Free Induction Decay</i>
FP	<i>Fluorescence Polarization</i>
GAGs	<i>Glycosaminoglycans</i>
GLIDE	<i>Grid-Based Ligand Docking with Energetics</i>
GROMACS	<i>GRoningen MAchine for Chemical Simulations</i>
HP	<i>Heparin / Heparina</i>
HS	<i>Heparan Sulfate / Heparan Sulfato o Sulfato de Heparano</i>
HSQC	<i>¹H-¹³C Heteronuclear Single Quantum Coherence</i>
IC ₅₀	<i>half maximal Inhibitory Concentration</i>
IDR	<i>Intrinsically Disordered Region</i>
ITC	<i>Isothermal titration Calorimetry</i>
KS	<i>Keratan Sulfate / Keratan Sulfato o Sulfato de Keratano</i>
MD	<i>Molecular Dynamics</i>
MD-TAR	<i>Time-Average Restrained (Molecular Dynamics)</i>
MM-PBSA	<i>Molecular Mechanics with Poisson–Boltzmann and Surface Area</i>
MM-GBSA	<i>Molecular Mechanics with Generalized Born and Surface Area</i>
MK	<i>Midkine / Midkina</i>
NMR / RMN	<i>Nuclear Magnetic Resonance / Resonancia Magnética Nuclear</i>
NOESY	<i>Nuclear Overhauser Effect SpectroscopY</i>

NOE	<i>Nuclear Overhauser Effect</i>
PANIC	<i>Peak Amplitude Normalization for Improved Cross-relaxation</i>
PDB	<i>Protein Data Bank</i>
PME	<i>Particle Mesh Ewald</i>
PTN	<i>Pleiotrophin / Pleiotrofina</i>
SD	<i>Stochastic Dynamics</i>
SPR	<i>Surface Plasmon Resonance</i>
STD	<i>Saturation Transfer Difference</i>
TOCSY	<i>¹H-¹H T_{OT}al Correlation SpectroscopY</i>
tr-NOESY	<i>transferred-NOESY/ NOESY transferido</i>
VMD	<i>Visual Molecular Dynamics</i>

Capítulo 1: Introducción y objetivos

1.1	<i>Carbohidratos</i>	2
1.2	<i>Glicosaminoglicanos</i>	4
1.2.1	<i>Estructura química y características de los carbohidratos</i>	4
1.2.2	<i>Biosíntesis de GAGs</i>	7
1.2.3	<i>Estructura e importancia biológica de GAGs</i>	9
1.2.3.1	<i>Estructura del Sulfato de Condroitina</i>	10
1.2.3.2	<i>Funcionalidad biológica del Condroitín sulfato</i>	11
1.3	<i>Factores de crecimiento</i>	12
1.3.1	<i>Midkina</i>	13
1.3.1.1	<i>Estructura de la MK</i>	14
1.3.2	<i>Pleiotrofina</i>	17
1.3.2.1	<i>Estructura de la PTN</i>	17
1.3.3	<i>Receptores candidatos de la MK y la PTN</i>	19
1.4	<i>Objetivos</i>	22

Capítulo 2: Metodología

2.1	<i>Introducción</i>	30
2.2	<i>Métodos teóricos</i>	31
2.2.1	<i>Química Computacional</i>	31
2.2.2	<i>Métodos de campos de fuerza</i>	31
2.2.3	<i>Cargas Parciales</i>	34
2.2.4	<i>Simulaciones Moleculares</i>	35
2.2.5	<i>Dinámicas moleculares promediadas en el tiempo (MD-TAR)</i>	37
2.2.6	<i>Cálculos de Interacción</i>	38
2.2.6.1	<i>Docking (Acoplamiento Molecular o Screening Virtual)</i>	38
2.2.6.2	<i>Métodos de MM/PBSA y MM/GBSA</i>	41
2.3	<i>Métodos experimentales</i>	42
2.3.1	<i>Resonancia Magnética Nuclear (RMN)</i>	43
2.3.1.1	<i>Experimento COSY y TOCSY</i>	44
2.3.1.2	<i>Experimento NOESY</i>	45
2.3.1.3	<i>Experimentos de ligando observado o de transferencia de magnetización</i>	48
2.3.1.4	<i>Experimento de Diferencia de transferencia de saturación (STD)</i>	50
2.3.1.5	<i>Experimentos heteronucleares</i>	52
2.3.2	<i>Polarización de Fluorescencia (FP)</i>	56

<u>Capítulo 3: Estudio estructural de GAGs mediante RMN y cálculos de Dinámica Molecular</u>	
3.1 Antecedentes	66
3.2 Pentasacáridos	67
3.2.1 Pentasacárido 1	67
3.2.2 Pentasacárido 2	71
3.2.3 Pentasacárido 3	77
3.3 Tetrasacáridos	81
3.3.1 Tetrasacárido 4	81
3.3.2 Tetrasacárido 5	84
3.3.3 Tetrasacárido 6	90
3.4 Análisis experimental mediante RMN	96
3.5 Metodología	102
3.5.1 Preparación de inputs y simulaciones	102
3.5.2 Método de análisis de los resultados teóricos	103
3.5.3 Método de análisis de los resultados experimentales	103
<u>Capítulo 4: Estudios de estructura e interacción experimental y teórica de los miméticos de CS-E con PTN</u>	
4.1 Introducción	109
4.2 Estructura de la Pleiotrofina	110
4.3 Estudio experimental.....	113
4.3.1 Experimento de NOESY transferido.....	114
4.3.1.1 Comparativa entre los complejos 1-PTN y 4-PTN	114
4.3.1.2 Comparativa entre los complejos 2-PTN y 5-PTN	118
4.3.1.3 Comparativa entre los complejos 3-PTN y 6-PTN	121
4.3.2 Experimento de STD (Saturation Transfer Difference)	123
4.3.3 Experimento de ^{15}N - ^1H -HSQC	125
4.3.4 Experimento de Polarización de Fluorescencia (FP)	130
4.4 Estudio teórico	132
4.4.1 Docking molecular	132
4.4.2 Dinámica Molecular.....	142
4.4.2.1 Comparativa entre los complejos 1-PTN y 4-PTN	143
4.4.2.2 Comparativa entre los complejos 2-PTN y 5-PTN	152
4.4.2.3 Comparativa entre los complejos 3-PTN y 6-PTN	161
4.4 Métodos MM-PBSA/GBSA	171
4.5 Metodología	179

4.5.1. Preparación de inputs para las dinámicas moleculares.....	179
4.5.2. Análisis de los resultados teóricos	181
4.5.2.1 Análisis estructural.....	181
4.5.2.1 Estudio de interacción	181
4.5.3. Análisis de los resultados experimentales	182
<u>Capítulo 5: Estudios de estructura e interacción experimental y teórica de los miméticos de CS-E con MK</u>	
5.1 Introducción	189
5.2 Estructura de la Midkina	190
5.3 Estudio experimental.....	193
5.3.1 Experimento de NOESY transferido.....	193
5.3.2 Experimento de STD (Saturation Transfer Difference)	194
5.3.3 Experimento de Polarización de Fluorescencia.....	196
5.4 Estudio teórico	197
5.4.1 Docking molecular	197
5.4.2 Dinámica Molecular.....	204
5.4.2.1 Complejo pentasacárido 1-MK.....	205
5.4.2.2 Complejo pentasacárido 2-MK.....	210
5.4.2.3 Complejo pentasacárido 3-MK.....	215
5.4.3 Métodos MM-PBSA/MM-GBSA	222
5.5 Metodología	227
6.1 Conclusiones.....	229

Capítulo 1

Introducción y objetivos

1.1 Carbohidratos

La glicobiología es el estudio de la biosíntesis, estructura, biología y evolución de los carbohidratos (denominados también glicanos o hidratos de carbono) que son un grupo de moléculas, compuestas mayormente de hidrógeno, carbono y oxígeno, distribuidas ampliamente en la naturaleza, considerándose la principal fuente de energía y pudiendo ser reconocidas por otras biomoléculas como las proteínas.^{[1],[2],[3],[4]}

Los carbohidratos son los productos naturales más abundantes desempeñando papeles esenciales en los seres vivos como el de proporcionar energía en nuestro cuerpo para su uso inmediato (glucosa), o conservación (almidón y glucógeno). Además, actúan como bloques de construcción estructural (denominados *building blocks*) de crustáceos e insectos y se encuentran también en las paredes vegetales y celulares de ciertos hongos (quitina y celulosa respectivamente). También constituyen los componentes fundamentales de todas las superficies celulares donde están involucrados procesos de reconocimiento celular esenciales para la vida.^{[5],[6]}

Una característica importante de la química de los carbohidratos es la funcionalización y modificación del anillo de azúcar, implicando la desoxigenación, epimerización, introducción de dobles enlaces y otros grupos funcionales, y modificación del grupo hidroxilo. Todas estas modificaciones están presentes en la naturaleza en multitud de sustancias con actividad biológica y relevancia farmacológica.^[7]

Los glicanos se presentan en forma de estructuras simples o complejas asociadas a otras biomoléculas, formando los denominados glicoconjugados. Su proporción de azúcar puede variar mucho en función de su tamaño total. La superficie de todo tipo de células en la naturaleza está cubierta en su parte periférica por el glicocálix, rico mayormente en carbohidratos. Según su unión a diferentes biomoléculas podemos encontrar varios tipos de glicoconjugados: glicoproteínas, donde la proteína se encarga de transportar uno o más glicanos unidos covalentemente al esqueleto polipeptídico generalmente a través de enlaces de N u O. Un N-glicano es una cadena de azúcar unida covalentemente a un residuo de asparagina de una cadena de polipéptidos que comúnmente involucra a un residuo de N-acetilglucosamina (GlcNAc) en eucariotas. Un O-glucano (oligosacárido unido a O) se une frecuentemente al polipéptido a través del grupo hidroxilo de un residuo de serina o treonina.^{[3],[8],[6]}

Un proteoglicano presenta una o más cadenas de glicosaminoglicanos (GAGs), polímeros con unidades repetitivas de disacárido sulfatado, unidos por O a una *core protein*, que es una proteína central en cuyo extremo contiene un residuo de xilosa unido al grupo hidroxilo de un residuo de serina. Algunos proteoglicanos pueden transportar cadenas GAGs y diferentes O- y N-glucanos.^{[9],[10]} Un glicosilfosfatidilinositol es otro tipo de glicoconjugado que contiene un glicano entre un fosfatidilinositol y una fosfoetanolamina, que está unida mediante enlace amida con el terminal carboxilo de una proteína. Esta estructura generalmente constituye el único anclaje a la membrana de la bicapa lipídica para tales proteínas. Por otra parte, los glicoesfingolípidos (a menudo

denominados glicolípidos) pueden ser neutros o aniónicos y contienen glucosa o galactosa unida al grupo hidroxilo primario terminal de la ceramida de la fracción lipídica. Cabe señalar que estos tipos de glicoconjugados representan solo las clases más comunes de azúcares reportados en las células eucariotas (Figura 1.1).^[4]

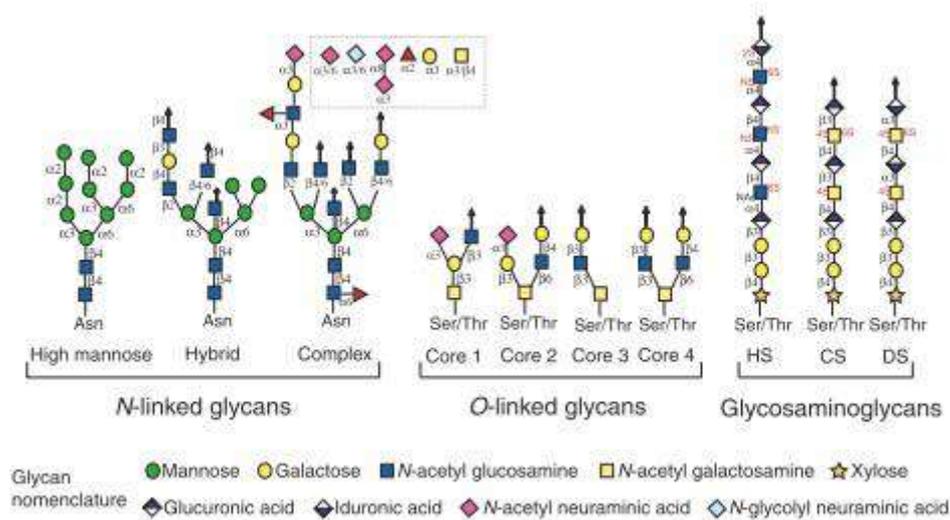


Figura 1.1. Diversidad química de Glicanos.^[4]

Existen evidencias acerca del papel de los glicanos en el crecimiento y desarrollo celular, crecimiento tumoral y metástasis, anticoagulación, reconocimiento/respuesta inmune, comunicación célula-célula y patogénesis microbiana. Al estar presentes en la matriz extracelular, los azúcares se encuentran en un entorno de muchas proteínas, como factores de crecimiento, citocinas, receptores inmunes y enzimas entre otros.

Debido a las interacciones entre ambas biomoléculas, los glicanos presentan numerosas funciones biológicas modulando por tanto la actividad de la proteína en la interfaz extracelular. La glicosilación es un fenómeno postraduccional que se produce en el retículo endoplásmico (RE) y el aparato de Golgi, en donde una misma cadena polipeptídica puede modificar su función eficazmente alterando su patrón estructural. Existen dos tipos de glicosilación dependiendo de las modificaciones que tengan lugar en la proteína. La primera de ellas es la N-glicosilación, si se produce la unión de un residuo de GlcNAc al átomo de nitrógeno de la cadena lateral de un aminoácido de asparagina (Asn) a través de un enlace β -1N. El segundo caso, sería la O-glicosilación que tiene lugar entre residuos de GalNAc y GlcNAc con aminoácidos que contienen grupos hidroxilo como Ser o Thr (ver Figura 1.2).^[8]

La diversidad de las glicoproteínas deriva de si hay uno o más sitios de glicosilación y de la comprensión de sus interacciones bioquímicas, ya que los glicanos pueden unirse a múltiples sitios de las proteínas dando origen al campo de la glicómica. Esta ciencia se centra en estudiar el glicoma, definido como el conjunto de las estructuras de los carbohidratos en todas sus formas. Aunque su diversidad es mucho más compleja, este estudio se realiza de forma análoga a la genómica (estudio del ADN), proteómica y lipidómica (estudio funcional y estructural de proteínas y lípidos respectivamente).

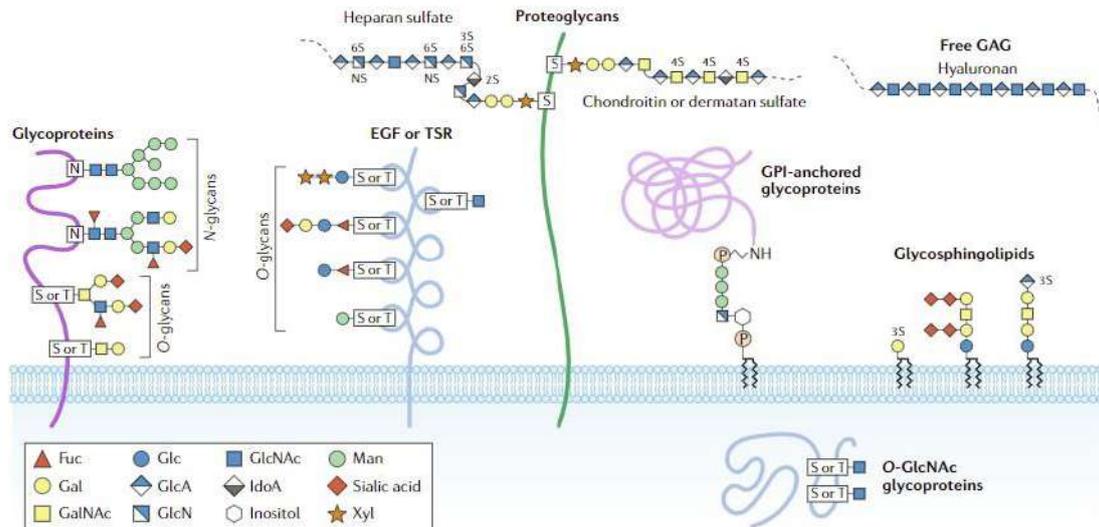


Figura 1.2. Tipos principales de glicosilación en humanos.^[8]

1.2 Glicosaminoglicanos

1.2.1 Estructura química y características de los carbohidratos

Para explicar la estructura de los glicosaminoglicanos y poder entender sus propiedades, comenzaremos describiendo los carbohidratos desde su forma más simple, los monosacáridos. Se definen como polihidroxialdehídos o polihidroxicetonas, compuestos por grupos hidroxilos que se repiten varias veces a lo largo del esqueleto carbonado y un grupo carbonilo en el extremo final que puede ser un aldehído o una cetona. Estos dos tipos de grupos carbonilo dan lugar a las aldosas y cetosas respectivamente y la parte terminal del glicano donde se encuentran estos grupos carbonilos se denomina extremo reductor. Los azúcares monosacáridos unidos entre sí mediante enlaces glicosídicos generan oligosacáridos y polisacáridos lineales o ramificados. La forma cerrada de un anillo monosacáridico genera un centro anomérico quiral en el C-1 para aldosas, y en el C-2 para cetosas.

Existen nueve monosacáridos comunes que se encuentran en los glicoconjugados de vertebrados. Una vez incorporados a proteínas/lípidos, estos “*building blocks*” pueden modificarse aún más para aumentar su diversidad. Los grupos hidroxilos de los diferentes monosacáridos pueden sufrir modificaciones tales como fosforilación, sulfatación, metilación, O-acetilación o acilación. Aunque los grupos amino están comúnmente N-acetilados, pueden presentarse también como N-sulfatados o permanecer sin sustituir. Los grupos carboxilo están ocasionalmente sujetos a lactonización con grupos hidroxilo o incluso a lactamización con grupos amino cercanos.^{[10],[11]}

La unión mediante el enlace glicosídico de un monosacárido a otro residuo tiene lugar a través de un grupo hidroxilo, pero también puede generarse con hidroxilácidos como la serina (Ser), treonina (Thr) y tirosina (Tyr) para formar las O-glicoproteínas anteriormente mencionadas.

El enlace glicosídico genera un centro anómico que puede dar lugar a dos estereoisómeros dependiendo del tipo de enlace (α o β). La unión de dos monosacáridos genera un disacárido, que a su vez puede constituir un punto de ramificación por participación en varios enlaces glicosídicos, formando oligosacáridos (entre 3 y 10 unidades) o cadenas más largas denominadas polisacáridos. La forma acíclica de los monosacáridos con cuatro o más átomos de carbono puede ciclarse por una reacción intramolecular hemiacetalica y generar también dos formas diastereoméricas α y β . Aldosas con 4 o más átomos y cetosas con 5 o más átomos de carbono pueden formar estructuras de furanosa. Similarmente, aldosas con 5 o más átomos y cetosas con 6 o más átomos de carbono pueden formar estructuras de piranosa.

Los carbohidratos son quirales y ópticamente activos, presentándose mayormente en la naturaleza como D-azúcares (siendo D-gliceraldehído y L-gliceraldehído el monosacárido referente para distinguir ambas series de enantiómeros), mientras que los aminoácidos que constituyen las proteínas pertenecen a la serie L en su mayoría.^[12]

En la Figura 1.3 puede verse la proyección de los azúcares en su forma de piranosa introducida por Haworth, donde se muestra la nomenclatura de los mismos en función de la posición del grupo lateral.

Sugar	O1(α)	O1(β)	O2	O3	O4	C6	Schematic Diagram
Pentoses							
Ribose	D	U	D	D	D	-	
Arabinose	D	U	U	D	D	-	
Xylose	D	U	D	U	D	-	
Lyxose	D	U	U	U	D	-	
Hexoses							
Allose	D	U	D	D	D	U	
Altrose	D	U	U	D	D	U	
Glucose	D	U	D	U	D	U	
Mannose	D	U	U	U	D	U	
Gulose	D	U	D	D	U	U	
Idose	D	U	U	D	U	U	
Galactose	D	U	D	U	U	U	
Talose	D	U	U	U	U	U	

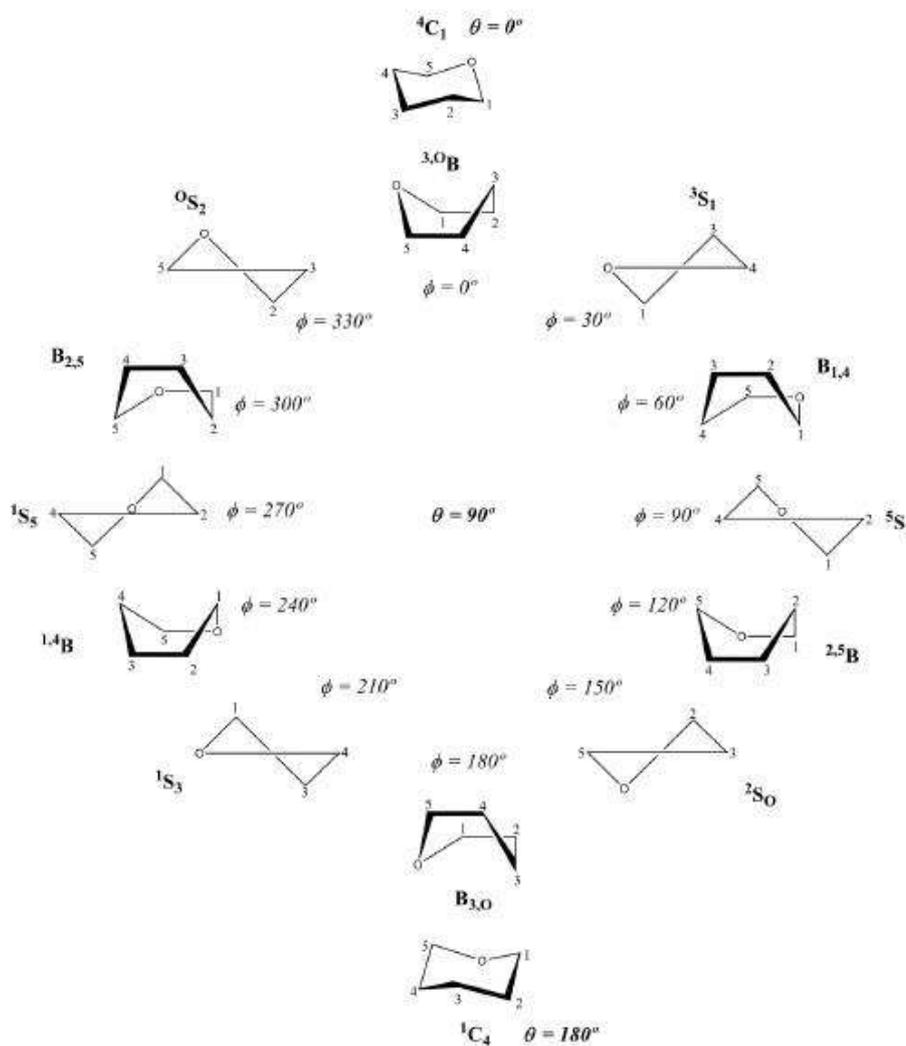


Figura 1.3. Tabla de formulación de los grupos laterales en la proyección de Haworth (arriba). Representación de los parámetros de puckering para un anillo de piranosa (abajo).^{[13],[14]}

En la Figura 1.3 también podemos observar las diferentes conformaciones existentes para los anillos de piranosa. La primera de ellas se denomina “conformación silla”, donde 4C_1 y 1C_4 son imágenes especulares entre sí. En el primer caso, el carbono 4 se encuentra arriba y el carbono 1 está debajo del plano formado por los átomos C2, C3, C5 y O5; al contrario que ocurre con la conformación 1C_4 , estando C1 arriba y C4 abajo del plano.

La segunda conformación es la denominada tipo “bote”, representadas como 1,4B , 2,5B , 3,0B (siendo O, el átomo de oxígeno), $B_{1,4}$, $B_{2,5}$, $B_{3,0}$; y por último las llamadas “bote torcido” constituidas por 1S_3 , 5S_1 , 2S_0 , 3S_1 , 1S_5 y 0S_2 . En la conformación tipo silla, los oxígenos exocíclicos de los anómeros α y β no son químicamente ni geoméricamente equivalentes, sino que uno de ellos está orientado de manera paralela (ecuatorial) y el otro perpendicular (axial) al plano del anillo. El efecto estereoelectrónico que tiene lugar cuando el sustituyente adyacente a un heteroátomo se encuentra preferentemente en posición axial en vez de ecuatorial se denomina efecto anomérico. Fue observado en los anillos de piranosa, considerando el sistema R – Y – C – Z, donde Z es un átomo electronegativo, C es el carbono anomérico e Y tiene un par de electrones solitario.

Factores electrónicos favorecen la disposición *gauche* de R y Z a lo largo del enlace Y – C.^[12]

Por tanto, el efecto anomérico favorece la estabilidad del anómero α frente al β . La explicación más simple considera la disminución de la repulsión entre los dipolos de los oxígenos endo- y exocíclicos cuando ambos sustituyentes electronegativos se encuentran en disposición axial (α -anómero), frente a la disposición ecuatorial (β -anómero).^[15] Otra explicación posible para el efecto anomérico considera la interacción estabilizante, denominada hiperconjugación, que ocurre entre el par de electrones libres situados en el orbital no enlazante del átomo endocíclico del anillo de azúcar y el orbital molecular antienlazante (σ^*) del enlace adyacente C₁-O₁, en el caso del anómero α . Por otro lado, el efecto exoanomérico favorece la conformación *gauche* entre el enlace C₁-O₅ y el enlace O₁-R (Figura 1.4). La hiperconjugación que explica el efecto exo-anomérico se da entre el par de electrones libres del orbital no enlazante (np) del oxígeno glicosídico O₁ y el orbital antienlazante σ^* del enlace O₅-C₁.^{[16],[15]}

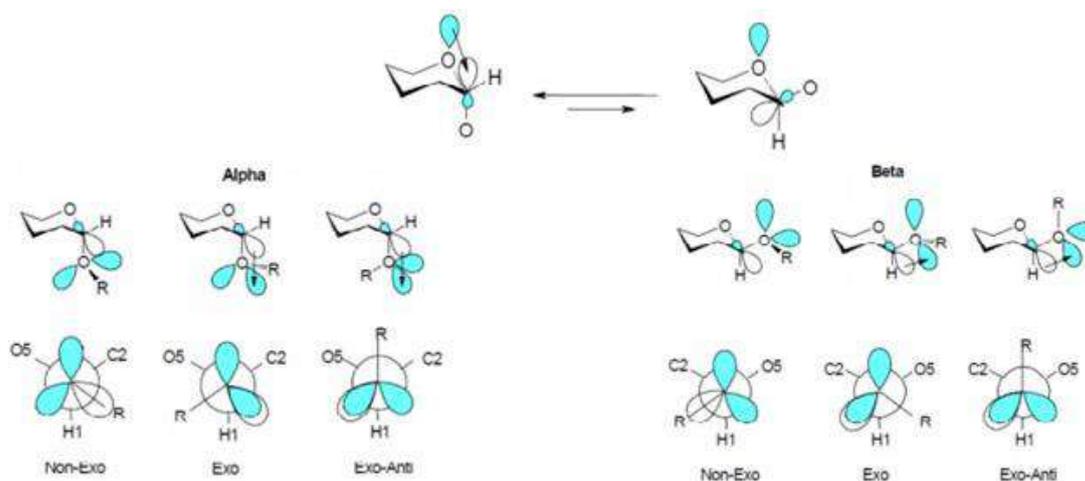


Figura 1.4. Representación esquemática de los efectos anomérico (arriba) y exoanomérico (abajo).

1.2.2 Biosíntesis de GAGs

Como se ha visto antes, la glicosilación es un proceso bioquímico que modifica proteínas tras su síntesis por los ribosomas, mediante la incorporación de carbohidratos. Un rasgo característico de las cadenas de azúcar en su unión a proteínas y lípidos es la variedad estructural generada por las diferentes combinaciones de unidades monosacáridicas originando una amplia variedad de oligosacáridos.^[17] Debido a esta diversidad estructural, y propiedades como el peso molecular, la estereoquímica de los enlaces interglicosídicos entre unidades de repetición de los sacáridos y la composición de dichas unidades monosacáridicas, estas moléculas pueden contener una gran cantidad de información biológica, demostrándose de hecho que muchas de ellas tienen funciones biológicas específicas. Estas actividades funcionales generalmente están relacionadas con las características estructurales de los carbohidratos, como la ausencia o presencia de

grupos sulfato o fosfato. Se cree que el sulfato de condroitina (CS) contiene patrones de sulfatación que sirven de anclaje a moléculas más complejas como son los factores de crecimiento, citoquinas y otras proteínas. Diversos estudios han demostrado que secuencias más largas que un tetrasacárido permiten un mejor ensamblaje de estos complejos, mejorando así el reconocimiento y la actividad neuronal.^[18]

La ruta biosintética de los GAGs tiene lugar en el aparato de Golgi, pero antes tiene lugar en el citoplasma, el proceso de sulfatación. En esta etapa, se absorbe sulfato inorgánico procedente del medio extracelular y se activa a continuación para que pueda reaccionar con el glicano aceptor. Estas formas activadas son el nucleósido difosfato (UDP) y el 3'-fosfoadenosin 5'-fosfosulfato (PAPS).^[19]

Se conocen principalmente seis clases distintas de GAGs y su biosíntesis sigue una ruta sintética diferente para generar las diferentes estructuras polisacarídicas. En el caso del sulfato de condroitina (CS), el sulfato de dermatano (DS), el sulfato de heparano (HS) y la heparina (Hp), la biosíntesis comienza generando una secuencia tetrasacarídica común, Xyl-Gal-Gal-GlcA unidos entre sí por enlaces β (1,3 y 1,4) y enlazados por el átomo de oxígeno de un aminoácido de serina (Ser) de la proteína núcleo (ver Figura 1.5).

A continuación, estos residuos se trasladan al aparato de Golgi a través de transportadores específicos, en donde unas enzimas específicas se encargarán del proceso de adición de residuos monosacarídicos y de sulfatación postraduccionales proporcionando la carga negativa necesaria para que tengan lugar las funciones biológicas de los GAGs. Estas enzimas son las glicosiltransferasas, encargadas de añadir los azúcares-UDP a la cadena polisacarídica, y las sulfotransferasas (ST) que utilizan PAPS para sulfatar en posiciones definidas los residuos de glicano. Este proceso modula señales extracelulares como las interacciones célula-célula y célula-matriz extracelular (ECM).^[18] En el retículo endoplasmático tiene lugar la unión entre el primer residuo de xilosa y la serina de la proteína, mientras que los siguientes pasos de elongación suceden en el aparato de Golgi.

Una vez formado el tetrasacárido unión, el quinto monosacárido decretará el tipo de GAG que se va a sintetizar. Las β -glicosiltransferasas, que unirán residuos de GalNAc, determinarán CS/DS y las α -glicosiltransferasas llevarán residuos de GlcNAc dictando la síntesis de HS/Hep (Figura 1.5). Cabe destacar que, para la obtención de DS y Hep, se produce una epimerización en los residuos de ácidos glucurónicos presentes en CS y HS, contribuyendo significativamente a la diversidad de los GAGs. La biosíntesis de CS y DS se completa mediante la O-sulfatación llevada a cabo por las enzimas PAPS, que transfieren el sulfato a las posiciones específicas en residuos de N-acetilgalactosamina, glucurónico e idurónico.^{[20],[21]}

Para el caso del sulfato de keratano (KS), la síntesis transcurre de manera similar a los anteriores GAGs pero mediante una ruta diferente. Primeramente, se activa el azúcar-UDP en el citosol, continuando por el retículo endoplasmático hasta llegar al aparato de Golgi. Una vez allí, se pueden generar tres clases de KS (tipo I, II y III), dependiendo de su complejación con la proteína. Si la unión tiene lugar por el N del residuo de Asn o por el O de Ser/Thr, la secuencia que se obtiene es de tipo I y II, respectivamente; si por el

contrario el enlace es a través de un residuo de Ser y una unidad de manosa, se originará una secuencia de tipo III. La polimerización de KS se lleva a cabo con glicosiltransferasas que añaden alternativamente GlcNAc y Gal, y la O-sulfatación específica de la posición 6 se realiza mediante sulfotransferasas. Por último, a diferencia de los anteriores, el ácido hialurónico (HA) es un glicosaminoglicano cuya biosíntesis tiene lugar de manera distinta. Se produce en la membrana plasmática, carece de linker, ya que no presenta unión covalente a una proteína central, y no está sulfatado.^[19]

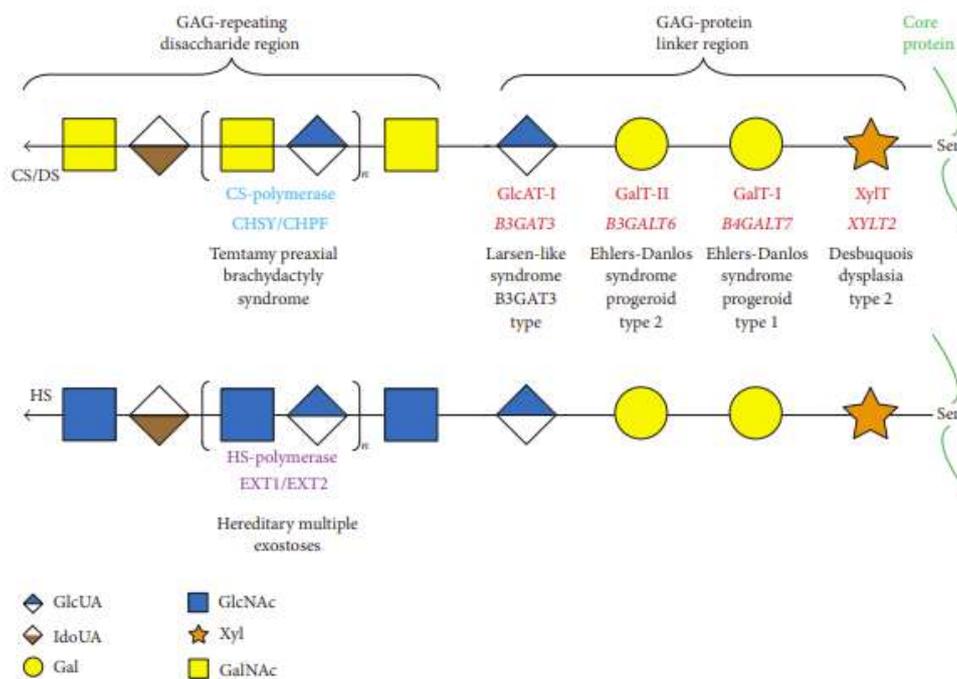


Figura 1.5. Biosíntesis de proteoglicanos de GAGs.^[22]

1.2.3 Estructura e importancia biológica de GAGs

Los GAGs son polisacáridos lineales formados por unidades disacáridicas, que alternan un ácido urónico y una hexosamina, repetidas a lo largo de la cadena, con grupos O-sulfato. Su unión a proteínas da lugar a macromoléculas que se denominan proteoglicanos. El carbohidrato en las glicoproteínas vistas anteriormente puede representar entre el 1 y el 80 de su peso total, pero normalmente es menor al resto de proteína. Por el contrario, los proteoglicanos pueden contener un 95% o más de carbohidratos y sus propiedades son más parecidas a las de los polisacáridos que a las de las proteínas.^[19]

La mayoría de GAGs están altamente sulfatados, uniéndose covalentemente a un núcleo de proteína a través de un linker que enlaza oligosacáridos. Sus unidades repetitivas de disacáridos cargados negativamente se unen covalentemente a proteínas centrales (*core proteins*) proporcionándoles una gran variación estructural y una enorme versatilidad funcional, siendo partícipes en numerosos procesos biológicos. Estos polisacáridos alternan diferentes unidades de azúcares, véase hexosamina (D-glucosamina o D-

galactosamina) y un ácido hexourónico (D-glucurónico o L-idurónico) o hexosa (D-galactosa). Los GAGs más comunes se pueden dividir en dos tipos: glucosaminoglicanos, véase sulfato de queratano (KS, del inglés Keratan sulfate), heparina (Hp, del inglés Heparin), ácido hialurónico (HA, del inglés Hyaluronic acid) y sulfato de heparano (HS, del inglés Heparan Sulfate); y sulfato de dermatano (DS, dermatan sulfate) y sulfato de condroitina (CS, del inglés Chondroitin Sulfate) como galactosaminoglicanos.^[23]

1.2.3.1 Estructura del Sulfato de Condroitina

En esta Tesis nos centraremos en el sulfato de condroitina (CS) siendo uno de los más importantes polisacáridos pertenecientes a la familia de los GAGs. El CS está formado por una cadena de disacáridos de β -D galactosamina N-acetilada y ácido β -D glucurónico unidos entre sí por enlaces β -1,4 y β -1,3. Dependiendo del código de sulfatación que presenten, se obtienen distintos tipos de CS (Figura 1.6).^[24]

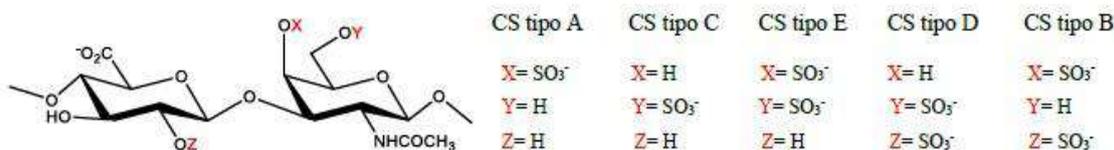


Figura 1.6. Variantes de CS en función de su grado de sulfatación.

Como hemos dicho anteriormente, la unidad básica del CS es un disacárido compuesto por GalNAc y GlcA, unidos entre sí según la secuencia GlcA β (1 \rightarrow 3) GalNAc β (1 \rightarrow 4).^{[25],[26]} La sulfatación es una de las modificaciones más comunes y, tal y como se observa en la Figura 7, los diferentes tipos de CS vienen dados por la disposición de los grupos sulfatos en los oxígenos situados en los carbonos 4 y 6 de las unidades de GalNAc, y en el carbono 2 en el caso de GlcA. Cada cadena de CS está unida a un residuo de serina de la *core protein* a través de una secuencia tetrasacáridica de unión Xyl-Gal-Gal-GlcA (Figura 1.7).

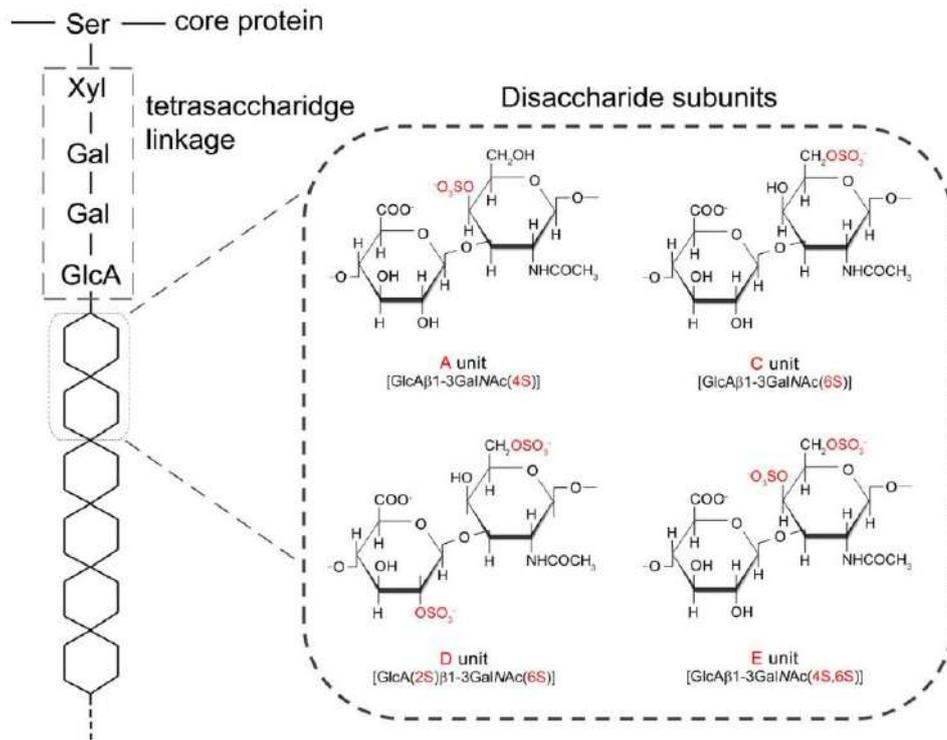


Figura 1.7. Diagrama esquemático de las estructuras CS-GAG.^[25]

La estructura macromolecular y la carga polianiónica proporcionada por los grupos sulfato y carboxilo permiten que los GAGs absorban grandes cantidades de agua y formen soluciones con alta viscosidad y elasticidad. Estas propiedades les permiten desempeñar funciones como lubricantes de articulaciones y cartílagos, amortiguadores y elementos de soporte en el tejido conectivo. Se encuentran ampliamente diseminados en la matriz extracelular y en la superficie celular, participando en el desarrollo embrionario, como receptores de factores de crecimiento y en la regulación de la diferenciación celular^[7] y del crecimiento e invasión tumoral.^{[19],[3]} Se ha demostrado que el sulfato de condroitina y el sulfato de dermatano tienen actividades anticoagulantes (al igual que la Heparina), antilipémicas, antiangiogénicas y antitumorales.^[27] Y, además, en el caso del primero se sabe que participa en la regulación del ensamblaje y mantenimiento de la matriz extracelular (ECM) y en la proliferación celular a través de la interacción con factores de crecimiento que incluyen: factores de crecimiento de fibroblastos, de hepatocitos o de neuritas (Pleiotrofina y Midkina).^{[28], [25]} Estas dos últimas proteínas van a ser nuestro objeto de estudio en esta Tesis.

1.2.3.2 Funcionalidad biológica del Condroitín sulfato

El CS es el GAG más abundante en la matriz del sistema nervioso central (CNS). La sulfatación es una de las principales modificaciones en su estructura y este proceso requiere que se produzca en una determinada posición ya que posteriormente influirá, en mayor o menor grado, en la interacción con las proteínas. Los diferentes patrones de sulfatación dan lugar a diferentes estructuras (ver Figuras 1.6 y 1.7) confiriéndole cierta afinidad por diversas moléculas de señalización con las que interactúa de manera

específica, proporcionándoles diferentes roles biológicos. Estas moléculas se encuentran en la matriz extracelular y pueden ser factores de crecimiento, receptores y moléculas guía.

Como la mayoría de los GAGs, el CS podemos encontrarlo en la matriz extracelular y en la superficie celular unido a una proteína central formando el condroitín sulfato proteoglicano (CSPG). Por otro lado, también aparece en los gránulos intracelulares de ciertas células, como los mastocitos.^{[23],[29]} El CS es un importante componente del cartílago, su resistencia a la compresión viene generada por los grupos sulfato cargados que van a generar una repulsión electrostática proporcionándole esta cualidad. Además, se tiene constancia de que varios CSPG regulan la fibrillogénesis del colágeno y participan en la función biomecánica de los tendones y ligamentos. Estos proteoglicanos también son fundamentales para la formación, desarrollo y mantenimiento de la morfología y función del cerebro^{[30],[29]} y los últimos estudios recientes apuntan también a un papel importante en trastornos cardiovasculares.^[31]

Como ya se ha mencionado, la interacción de los GAGs con una gran variedad de proteínas les confiere una gran importancia en numerosos procesos fisiológicos como la migración, el crecimiento, la diferenciación, la orientación y el desarrollo celular. La especificidad y selectividad de estas interacciones no se limita solo a secuencias proteínicas, sino que dependen también de los oligosacáridos y su patrón de sulfatación. Dentro de las familias de proteínas que interaccionan con el CS encontramos los factores de crecimiento (GF), que son biomoléculas que favorecen la proliferación, el crecimiento, la migración y la diferenciación celular. Además, intervienen en la regulación del metabolismo, la reparación tisular y el mantenimiento de la homeostasis tisular en organismos adultos.^{[32],[29]} Dentro de esta familia, encontramos los factores neurotróficos, que promueven la supervivencia, crecimiento y diferenciación de las células neuronales; y los factores de crecimiento de fibroblastos (FGF), que ejercen efectos pleiotrópicos en diferentes tejidos.^[33]

Midkina (MK) y Pleiotrofina (PTN) son otros ejemplos de citoquinas de unión a heparina que constituyen la familia NEGF (*neurite growth promoting factor*), localizadas en las fibras gliales radiales del cerebro embrionario, a lo largo de las cuales migran las células madre neuronales. MK y PTN median en la adhesión y migración de células neuronales y promueven el crecimiento de neuritas al interactuar con la Hep de la superficie celular.^{[34],[29]} Ambas están involucradas en enfermedades neurodegenerativas como el Alzheimer.^[35] Por lo tanto, conocer la interacción CS-Midkina/Pleiotrofina resulta un objetivo potencial de estudio.

1.3 Factores de crecimiento

Como ya hemos mencionado, los factores de crecimiento son moléculas de señalización extracelular que juegan un papel fundamental en la diferenciación, proliferación, supervivencia, adhesión, migración y otras actividades de respuesta celular. Son proteínas

de bajo peso molecular que se secretan e incluyen hormonas y citoquinas, que se encargan de la comunicación intercelular, y la construcción y remodelación de tejidos,^[36] incluyendo la carcinogénesis.^[35] Los factores de crecimiento incluyen familias de proteínas funcionales como son los factores de crecimiento de los fibroblastos (FGF), y de neuritas (MK y PTN). Estas proteínas son reconocidas por GAGs, caracterizándose por su unión a CS/DS. Su participación en la regulación y desarrollo de enfermedades las convierte en un importante foco de estudio.

Los factores de crecimiento de neuritas están presentes en todos los vertebrados y están compuestos mayoritariamente de aminoácidos básicos y cisteína. Su estructura presenta dos dominios de lámina- β , denominados N-terminal y C-terminal unidos por tres puentes disulfuro en el caso del dominio N y otros dos en el dominio C. Para el caso de la MK, su principal zona de unión a GAGs se produce a través del dominio C, y promueve el crecimiento, supervivencia, migración y expresión génica de las células, entre otras funciones. En cambio, la PTN es una proteína con actividad motora de neuritas que también promueve la actividad de fibroblastos.

Por otro lado, diversos estudios han demostrado que estas proteínas se comportan como factores angiogénicos derivados de tumores que pueden direccionar ribozimas para que supriman la invasión y metástasis del melanoma y el coriocarcinoma.^[37] En términos generales, ambos factores de crecimiento (MK y PTN) se unen fuertemente a sindecanos (un tipo de proteoglicanos de heparán sulfato) y la importancia biológica de ambas moléculas radica en su inusual expresión en determinadas enfermedades.

1.3.1 Midkina

La MK es un factor de crecimiento de unión a heparina con un peso molecular de 13 kDa, cuya expresión se encuentra con mayor fuerza en los primeros episodios embrionarios^[38] y es controlada por el ácido retinoico. Dicha expresión es más intensa en la remodelación de los tejidos nerviosos, epiteliales y los tejidos del mesodermo, siendo más débil o ausente en la mayoría de los tejidos adultos.^[39] La expresión controlada de la MK durante la embriogénesis y sus características estructurales únicas la dotan de gran importancia. La dimerización de sus dominios mejora su actividad, y sus inhibidores son útiles en el tratamiento de cáncer,^[40] esclerosis múltiple, y artritis reumatoide.^[39] La MK mejora el tratamiento de lesiones de isquemia cerebral,^[41] y es promotora de plasminógenos para la fibrinólisis,^[42] degeneración retiniana^[43] e insuficiencia cardíaca^[44]. Pero el principal rol de la MK se encuentra en el sistema nervioso, mejorando la presencia de neuritas de las neuronas embrionarias y promoviendo la diferenciación neuronal. Estudios recientes han revelado la importancia médica de esta nueva familia de factores de crecimiento, debido a que, por ejemplo, la expresión de ARNm de MK se ve incrementada en varios carcinomas humanos.^[45]

Estudios de RMN han revelado que, como se ha mencionado antes, MK presenta dos dominios estabilizados por puentes disulfuro; el primero de ellos se denomina N-terminal y está situado entre los residuos de MK (15-52), y el segundo C-terminal ubicado entre

los residuos (62-104). Este dominio C terminal es el sitio de reconocimiento de carbohidratos, generalmente heparina. Además, se conoce que la dimerización de la MK mejora sus actividades biológicas y los GAGs fomentan la formación del dímero con una forma reticular debido a la acción de una transglutaminasa. Otra singularidad de la MK, que ha sido motivo de interés desde el punto de vista estructural, es la estabilidad que presenta al calor y a los ácidos, sin que afecte significativamente a su actividad potenciadora de plasminógeno y neuritas.^[46]

Estudios recientes también han demostrado que proteoglicanos de CS son receptores de MK y PTN. Estas proteínas se unen con más alta afinidad al azúcar que a la porción proteica del proteoglicano.^[47] Usando heparina y derivados de CS parcialmente sulfatados como modelo, se concluyó que las unidades sulfatadas están implicadas y son necesarias para la completa unión a la MK; en concreto, la unidad N-acetilgalactosamina 4,6-disulfatada del CS de tipo E presenta una unión tan fuerte como la heparina.^[48] Ésta inhibe a la MK; por tanto, es probable que el reconocimiento de GAGs en proteoglicanos sea esencial para la actividad de esta citoquina. Sin embargo, aún sigue siendo objeto de estudio la estructura de los azúcares que interaccionan con MK.

1.3.1.1 Estructura de la MK

Los dos dominios, MK (1-59) y MK (62-104) están sujetos por puentes disulfuro con una homología ligeramente similar a la trombospondina de tipo I (TSR). Presentan tres láminas- β antiparalelas; en el caso del primer dominio, situado más cerca del extremo N-terminal encontramos: β_{N1} (Trp20-Pro25), con una protuberancia en el residuo de Pro22, β_{N2} (Val32-Cys39) y β_{N3} (Gln42-Val50), conectados entre ellos por dos *loops* (Figura 1.8A).

Para las láminas β más próximas al extremo C-terminal tenemos: β_{C1} (Tyr64-Asp73), con una protuberancia en (Asn68-Trp69-Gly70); β_{C2} (Thr78-Leu85) y β_{C3} (Thr97-Thr101) (Figura 9B). Se observa homología entre las secuencias β_{N2} (Arg35–Glu36–Gly37–Thr38), β_{C2} (Arg81–Gln82–Gly83–Thr84), β_{N3} (Ile46–Arg47, Pro51–Cys52) y β_{C3} (Ile98–Arg99, Pro103–Cys104), respectivamente.^[42]

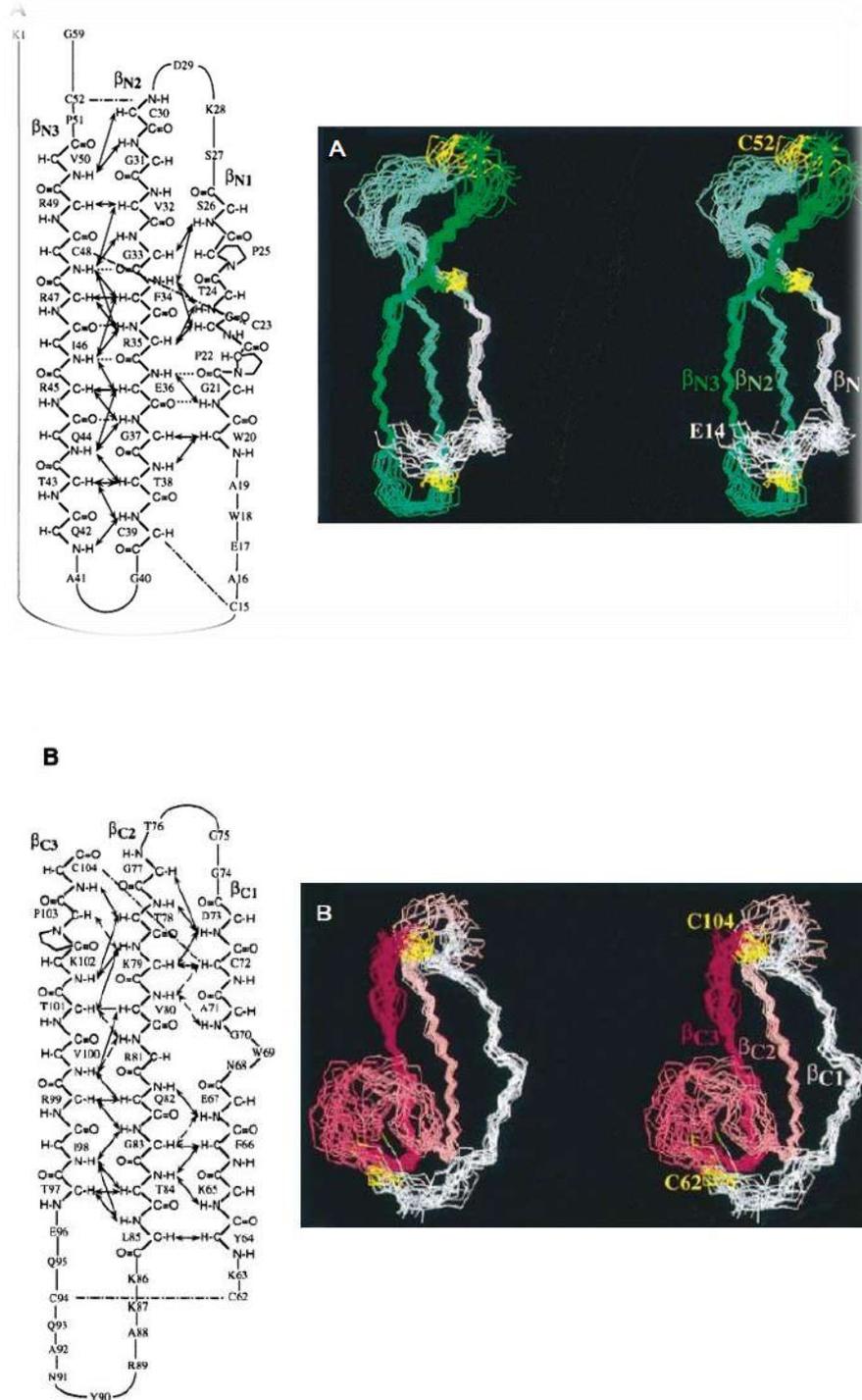


Figura 1.8. Estructura secundaria de (A) MK (1-59) y (B) MK (62-104).^[42]

La diferencia más remarcable entre la estructura secundaria de MK (1-59) y MK (62-104) es la longitud del *loop* de la horquilla- β secundaria que conecta la segunda y tercera lámina- β . Existen clusters hidrofóbicos tanto para MK (1-59) como para MK (62-104); este último, que se encuentra ubicado más cerca del extremo C-terminal, está compuesto por dos zonas de reconocimiento de carbohidratos (*cluster-1* y *cluster-2*) que contienen compuestos por aminoácidos básicos. Diversos estudios de afinidad han demostrado que para generar un fuerte anclaje a la MK es especialmente importante que en el *cluster-1* tenga lugar el reconocimiento de unidades de sulfato de heparina o sulfato de condroitina.

Los aminoácidos en este cluster están localizados en dos láminas β , Lys79 y Arg81 en β_{C2} y Lys102 en β_{C3} .^[35] El aminoácido Trp69 situado en β_{C1} queda enfrentado a Arg81 restringiendo la orientación de la cadena lateral del mismo (Figura 1.9D). El *cluster-2* se encuentra en el *loop* flexible formado por los aminoácidos Lys86, Lys87 y Arg89.^[49]

Varios estudios de RMN con heparina-12mer mostraron que los residuos de Ala88, Arg89, Tyr90, Lys79 y Arg81 presentaban cambios en sus desplazamientos químicos al unirse con MK (62-104) lo que se traduce en ser los principales enclaves de unión a heparina. Por el contrario, aunque MK (1-59) tiene bastantes residuos básicos, no se acoplan a los grupos sulfatos de la heparina y como consecuencia su unión a ella será débil (Figura 1.9A).^[42]

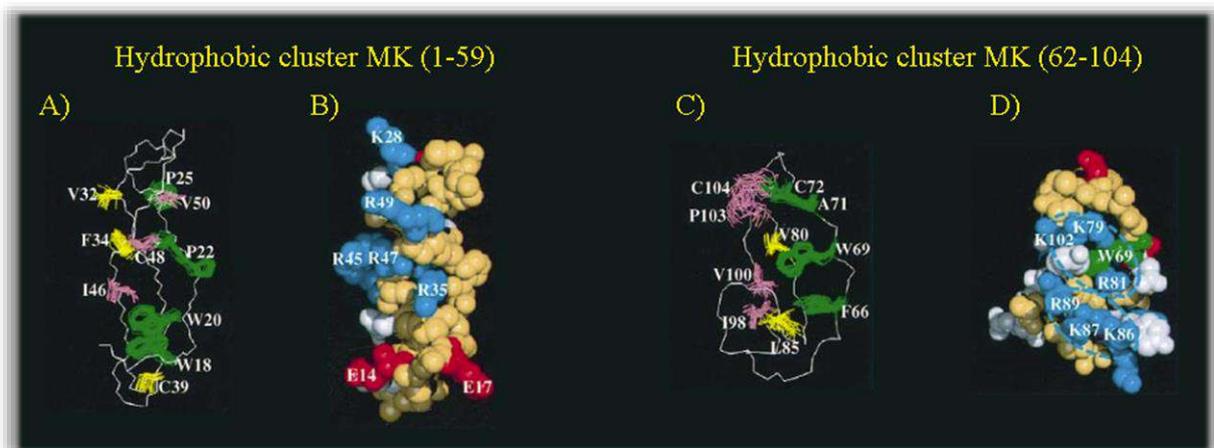


Figura 1.9. Clusters hidrofóbicos de MK (1-59) (A) y MK (62-104) (C), mostrando la superposición de 20 estructuras de las cadenas laterales para residuos hidrofóbicos. Mediante el modelo de esferas se destacan los residuos básicos (azul), ácidos(rojo), hidrofóbicos (beige) y polares (blanco) en la superficie de la proteína (B y D). En el caso del cluster MK (62-104), Trp69 se muestra en verde y los grupos básicos aparecen rodeados por una línea discontinua en azul (D).^[42]

Estudios posteriores con una de las midquinas presentes en el pez zebra (Mdkb) y un pentasacárido sintético de heparina (fondaparinux) sugirieron dos nuevos sitios de reconocimiento de carbohidratos en la proteína (*cluster-3* y *cluster-4*) situados en el dominio N-terminal (Figura 1.10).^[50]

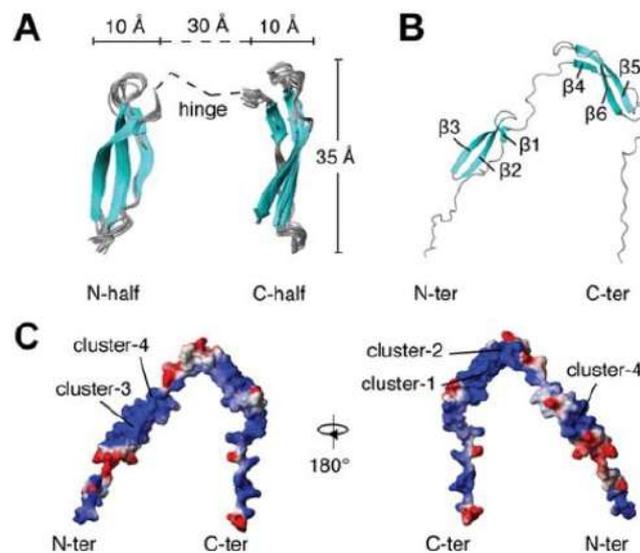


Figura 1.10. Representación de la estructura de la Midkina. A) Superposición de las 10 mejores estructuras secundarias de MKb excluyendo los extremos terminales de N y C. B) Representación de una de las estructuras de MKb. C) Representación electrostática de los cuatro clusters.^[50]

1.3.2 Pleiotrofina

La PTN es otro miembro de la familia de los factores de crecimiento de neuritas (NEGF) que, al igual que MK, tiene un importante rol en el desarrollo neuronal, angiogénesis y regeneración de tejidos.^{[51],[52],[53]} Además, diversos estudios han demostrado que la expresión de esta proteína se encuentra en mayor o menor grado en las células tumorales dependiendo del grado del tumor^{[54],[55],[56]}. La expresión de PTN también puede llegar a ser beneficiosa, ayudando a mantener células madres hematopoyéticas, desencadenando reparación tisular en episodios de isquemia^[57] y proporcionando protección sobre tejidos neuronales, convirtiéndose en un objetivo terapéutico en enfermedades neurodegenerativas como el Parkinson.^{[58],[59]} Dependiendo de la actividad que vaya a desempeñar, la PTN interacciona, del mismo modo que la MK, con el sindecano de tipo N o el dominio proteoglicano CS presente en los receptores de tipo proteína tirosina fosfatasa PTP (proteoglicano de CS).^{[60],[61]} La unión con el primer receptor es crucial durante el desarrollo neuronal; en cambio con el segundo receptor, su asociación con estructuras de CS sobresulfatadas (concretamente de tipo E) facilita la proliferación y el crecimiento celular.

1.3.2.1 Estructura de la PTN

Comparte una homología del 50% con respecto a la MK, pero con una mayor cantidad de aminoácidos que le confiere un peso molecular de 18 kDa. Diversos estudios a diferentes pH mostraron similitudes entre estos dos factores de crecimiento, ya que PTN también presenta dos dominios, similares a la trombospondina de tipo I (TSR), N y C-terminal. Este último dominio C terminal presenta mayor afinidad por heparina, pero su cola terminal, a pesar de contener una gran cantidad de aminoácidos básicos, no contribuye

significativamente a la interacción.^[62] Teniendo en cuenta que la unión con GAGs está mediada por interacciones electrostáticas, los grupos de aminoácidos básicos se convierten en potenciales sitios de unión. Ambos dominios contienen cadenas de residuos de Lys, pero el extremo-C presenta dos *clusters* de aminoácidos básicos en comparación al extremo-N terminal. La composición de cada uno de ellos muestra una equivalencia a los *clusters* presentes en la MK, el primero de ellos compuesto por los residuos Lys68, Lys91 y Lys92 y el segundo *cluster*, ubicado en el lado opuesto de la lámina- β , formado por Lys84, Arg86 y Lys107 (correspondientes a los *clusters* de la MK II y I, respectivamente). Para el extremo-N terminal, la distribución de los aminoácidos en el *cluster* es más escasa, encontrándose situados también en un lado de la lámina- β , Arg35, Arg39 y Lys49, mostrando equivalencia con el *cluster* III identificado en la MK del pez cebra, pero con un aminoácido más.^[42] Cabe destacar, que, aunque los aminoácidos se encuentren ubicados en lados opuestos de las láminas- β , la curvatura que presenta la proteína les hace posicionarse muy cerca los unos de los otros.

Debido a la ausencia de NOEs entre ambos dominios, se cree que no presentan una orientación definida, mostrando la segunda y tercera hebra una mayor estructura de lámina- β con respecto a la primera debido a la distribución de sus enlaces disulfuros. Pero, a pesar de esto, sí se pudieron observar picos cruzados entre el extremo-C terminal y los residuos 58-66 situados en el segmento de la bisagra que conecta ambos dominios. Esto crea una curvatura acercando los aminoácidos básicos Lys60 y Lys61 con el *cluster* I del extremo terminal C (Figura 1.11).

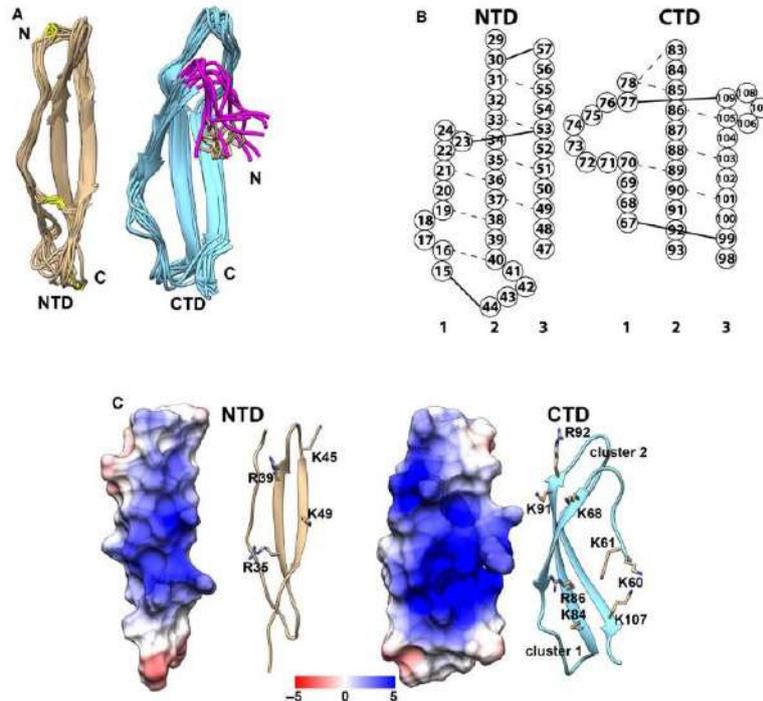


Figura 1.11. Representación de la estructura de la Pleiotrofina. A) Superposición de las 10 mejores estructuras más bajas en energía. De color amarillo los puentes disulfuro, en magenta la zona de la bisagra. B) Descripción esquemática de las láminas- β de la PTN con la correspondiente numeración de sus residuos. Las líneas discontinuas muestran los NOE observados. Las líneas continuas muestran los puentes disulfuro C) Mapa de potencial de los dominios N y C terminal.^[63]

Diversos estudios de RMN confirmaron que los sitios de unión fuerte para un proteoglicano de CS se encuentran en el extremo C-terminal y la zona de la bisagra, con valores de Kd cuatro veces más pequeños con respecto a los de NTD (dominio TSR N-terminal).^[63] Probablemente durante la unión al GAG, la zona de la bisagra se pliega a lo largo del CTD, permitiendo que los residuos Lys60 y Lys61 se integren en el cluster I de este dominio.

Por otro lado, a pesar de que el extremo final del dominio CTD (dominio TSR C-terminal) sea prescindible para unirse a heparina, es necesario para la unión al receptor proteína tirosina fosfatasa (PTP ζ , ver apartado 1.3.3) que se requiere para desencadenar una de las actividades de la PTN. Estudios posteriores de RMN han demostrado que este dominio tiene cierta relevancia para mantener interacciones estables con CS de tipo A^[64] y además puede asociarse a otros receptores como las integrinas, demostrando que el complejo juega un papel crucial en los mecanismos de transformación y migración celular.^{[64],[65]}

Los dominios de PTN presentan diferentes afinidades por los diferentes tipos de CS. Véase el caso del CS de tipo C, que se expresa en niveles más altos en el sistema nervioso central durante los primeros episodios de desarrollo neural,^[66] y muestra una mayor afinidad por esta proteína, mejorando significativamente la formación del complejo con el receptor. Otro ejemplo sería el CS de tipo E, presente en algunos tumores y células metastásicas, cuyos proteoglicanos van a servir como receptores de alta afinidad para PTN, explicando la presencia de la misma en las células cancerosas.^{[63],[67]}

1.3.3 Receptores candidatos de la MK y la PTN

Para que puedan llevar a cabo sus diferentes funciones, estos factores de crecimiento deben anclarse previamente a receptores transmembrana presentes en la matriz extracelular. Estos receptores son moléculas de señalización que regulan numerosos procesos celulares y pueden ser de diferentes tipos en función de las moléculas que reconozcan y/o señalicen.

El receptor-tipo proteína tirosina fosfatasa (PTP ζ) es una proteína transmembrana que se expresa como proteoglicano CS y que contiene una región transmembrana y una de tirosina fosfatasa intracelular. Se sabe que ambos factores de crecimiento, MK y PTN, comparten un sitio de unión común con este receptor transmembrana. Diversos estudios con mutantes de MK permitieron establecer que cuando este receptor se unía a la mitad del extremo C-terminal del cluster I de la MK y tenía lugar una mutación en el aminoácido Arg78, la afinidad del enlace disminuía, a diferencia de las mutaciones de los residuos Lys83 y Lys84 en el cluster II, donde no se vio afectada la unión. Estos hallazgos sugirieron que los aminoácidos básicos en el cluster I de la MK y la PTN son cruciales para conseguir una alta afinidad con PTP ζ y como resultado transducir señales en neuronas. La unión entre PTP ζ y estos factores de crecimiento se ve fuertemente inhibida por heparina, y moderadamente por heparán y condroitín sulfato. En el caso de este último, los sulfatos de condroitina de tipo D y E interaccionan más intensamente que los de B y C, mientras que el de tipo A no ejerció ningún efecto.^{[68],[69]}

Tanto la MK como la PTN presentan dos sitios de unión con este receptor: uno de alta afinidad en la porción donde se encuentra el GAG del proteoglicano ($K_d = 0.58$ nM y $K_d = 0.25$ nM, respectivamente) y otro con menor afinidad situada en la región proteica ($K_d = 3$ nM para ambas).^{[70],[68]} Después de la unión con estos factores de crecimiento, los receptores transmembrana se activan y se desencadena la fosforilación de las tirosinas de la región citoplasmática que da lugar al inicio del sistema de señalización, recibiendo dicha señal dos quinasas tipo MAP y PI3.^[71]

Se cree que la PTP ζ se dimeriza, provocando la inactivación del dominio fosfatasa y el aumento del fosfato de tirosina intracelular. Esto a su vez puede estimular la actividad de otro receptor tipo proteína tirosina, en este caso quinasa, denominada ALK, que fue descubierto debido a su participación en el linfoma anaplásico de células grandes. Muchos tipos de células cancerosas y células normales que expresan ALK presentan proliferación y supervivencia después de la estimulación con MK o PTN, indicando que éste es un receptor funcional importante para estas proteínas, aunque hay aún estudios que defienden que esta enzima no es un receptor directo de la PTN, sino que su actividad está regulada por la actividad de la fosfatasa de PTP ζ como se ha indicado anteriormente.^[72]

Estudios posteriores dilucidaron que existían otras familias de receptores transmembrana distintas de los proteoglicanos que reconocían MK y PTN, como por ejemplo lipoproteínas de baja densidad (LDL) como el LRP. Éste, además de internalizar ligandos en el lisosoma para su degradación y metabolismo también actúa como un receptor endocítico de señales de la membrana celular para activar las vías de señalización intracelulares en la migración y la proliferación celular (ver Figura 1.12). Estudios de afinidad con MK en embriones de ratón mostraron que esta proteína se une con una alta afinidad al LRP ($K_d = 3.5$ nM).^[73]

Entre otros receptores que también se han descubierto que presentan interacción con estos factores de crecimiento, pero en menor grado, están las integrinas, familia de glicoproteínas que forman receptores heterodiméricos (subunidades α y β) para la adhesión célula-sustrato y célula-célula, desempeñando un papel importante en la morfogénesis, inflamación, metástasis de cáncer y autorrenovación de células madre.

También encontramos: neuroglicanos de tipo C (NGC), que son proteoglicanos de sulfato de condroitina transmembrana; los sindecanos, que además de tener un dominio intracelular, presentan otro extracelular compuesto por proteoglicanos de sulfato de heparina; y los *nocht* que son vías de señalización celular altamente conservadas, que desempeñan un papel importante en la comunicación célula-célula.^{[73],[72]}

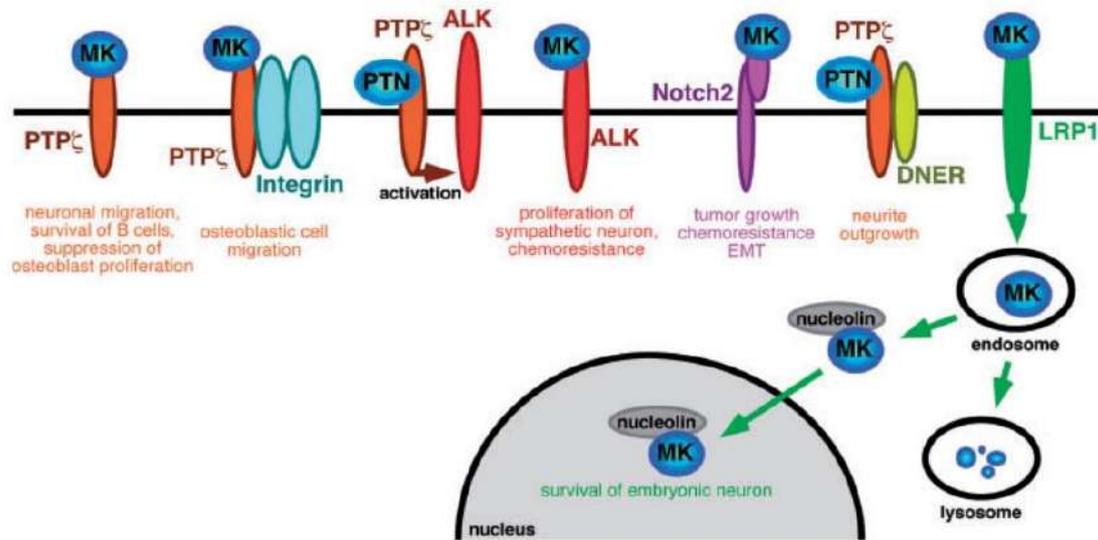


Figura 1.12. Candidatos receptores y proceso de señalización para MDK y PTN. [73]

1.4 Objetivos

Como se ha descrito anteriormente, cuando los GAGs se unen a determinadas proteínas como lectinas y factores de crecimiento, sus complejos pueden desarrollar funciones biológicas importantes, así como participar en el proceso de ciertas enfermedades.

El sitio de reconocimiento de estas proteínas por los GAGs ha sido ampliamente estudiado, demostrándose que pueden presentar más de un sitio de unión (zona principal y secundaria), y que sus superficies electrostáticas son capaces de acomodar los diferentes grupos sulfatos que se encuentran en los oligosacáridos GAG.^{[74],[75],[76]} Las diferencias de tamaño y los distintos patrones de sulfatación que presentan los GAGs van a desencadenar que se orienten de una forma u otra, en uno o ambos sitios de reconocimiento.

En nuestro grupo se han llevado a cabo estudios estructurales de diferentes oligosacáridos de Hp, HS, CS/DS y estudios su interacción con Langerina (Lectina de tipo C),^{[77],[78]} FGF-1, FGF-2^{[79],[80]} y MK^[81] (factores de crecimiento), para demostrar la influencia que tiene la estructura del GAG en su complejación con la proteína. En estos estudios se emplearon métodos teóricos y experimentales, pudiendo determinarse la estructura de los distintos oligosacáridos GAG mediante dinámicas moleculares de cada uno de estos ligandos en su forma libre, comparándose con los resultados experimentales obtenidos con técnicas de Resonancia Magnética Nuclear (RMN). Por otro lado, la estructura del complejo GAG-proteína se estudió mediante *docking* como método teórico y experimentos de NOESY transferido y STD como técnicas de RMN por observación del ligando.

El objetivo general de la presente Tesis Doctoral es el estudio estructural y de interacción entre diversos derivados GAGs y los factores de crecimiento MK y PTN. En concreto, se estudiaron una serie de oligosacáridos relacionados con la estructura del sulfato de condroitina tipo E, ya que se sabe que dicho patrón de sulfatación es capaz de interaccionar con ambos factores de crecimiento (Figura 1.13). Estos oligosacáridos, sintetizados en nuestro grupo de investigación, poseen unidades de glucosa en lugar de los correspondientes ácidos glucurónicos del CS y pueden considerarse miméticos de los productos naturales. Además de los tetra- y pentasacáridos completamente desprotegidos **1** y **4**, también se consideraron en nuestro estudio los intermedios sintéticos **2-3** y **5-6** que poseen algunos o todos los grupos hidroxilos protegidos con sustituyentes hidrofóbicos, ya que resultados previos de nuestro grupo sugerían que este tipo de derivados también puede unirse a MK/PTN.^{[81][82]}

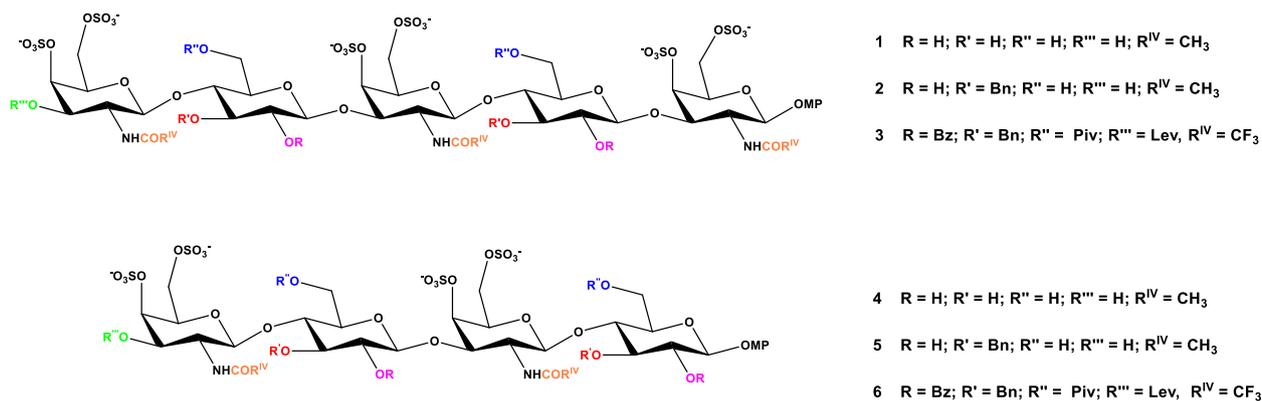


Figura 1.13. *Miméticos de CS con diferentes patrones de sustitución. En el caso de los compuestos completamente desprotegidos (1 y 4), los sustituyentes R, R', R'' y R''' corresponden a átomos de H mientras que R^{IV} es un grupo CH₃. Para los ligandos dibencilados (2 y 5), el único cambio en la estructura de los azúcares consiste en la presencia de un anillo aromático bencílico en el grupo R'. Por último, en los compuestos persustituídos (3 y 6), aumenta la presencia de grupos hidrofóbicos encontrando ahora, además de los grupos bencilo (Bn) en R', grupos benzoilo (Bz) en R; residuos de Pivaloilo (Piv), Levulino (Lev) y CF₃ en R'', R''' y R^{IV} respectivamente.*

El trabajo de esta Tesis se centra en determinar e intentar simular estructuras tridimensionales de cada uno de los compuestos objeto de estudio, además de estudiar los modos de unión y sitios de interacción proteína-ligando. Para ello nos planteamos los siguientes objetivos específicos:

Primer objetivo:

1. Determinar la estructura tridimensional de los ligandos libres **1-6** mediante experimentos de RMN y simulaciones de dinámica molecular. Los datos de RMN proporcionarán valores experimentales para las constantes de acoplamiento y las distancias interprotónicas que se compararán con los valores teóricos obtenidos durante las dinámicas moleculares de los ligandos libres.

Segundo objetivo:

2. Estudio de interacción de los ligandos **1-6** con PTN y MDK mediante ensayos de competición por polarización de fluorescencia (FP).

Tercer objetivo:

3. Determinar las potenciales estructuras de los complejos entre PTN y MDK y los compuestos **1-6** y **1-3**, respectivamente, empleando tres conformaciones de cada proteína representativas de las posibles orientaciones relativas de los dominios

TSR. Para esta tarea se plantea el uso de cálculos de docking con Autodock-vina y Autodock-vina Split. Así se pretende analizar la posible influencia teórica de la proteína en la conformación, orientación y posición relativa del ligando.

Cuarto objetivo:

4. Realización de dinámicas moleculares de los complejos más plausibles, obtenidos en el tercer objetivo: PTN y ligandos **1-6** y MDK y **1-3**, usando tres posibles orientaciones relativas de los dominios estructurados.
5. Análisis de las estructuras experimentales de los ligandos enlazados obtenidas por experimentos de RMN de observación del ligando (NOESY transferido y STD-NMR) y su comparación con las estructuras teóricas de las trayectorias tanto libres como con restricciones promediadas en el tiempo.
6. Estudio del comportamiento conformacional de los ligandos en los distintos de complejos y la influencia en éste de la longitud y de los patrones de sustitución. Análisis de la influencia de la conformación de la proteína en la interacción con el carbohidrato.

Quinto objetivo:

7. Realizar cálculos de Energía libre de interacción empleando métodos MMPBSA y MMGBSA sobre las dinámicas moleculares de cada uno de los complejos obtenidos en el punto anterior. Comparación con los resultados experimentales de STD-NMR y FP.
8. Comparación de la valoración experimental ^{15}N -HSQC entre el compuesto **4** y PTN y los resultados teóricos obtenidos por simulación MMPBSA de ese complejo.

Bibliografía Capítulo 1

- [1] A. Varki in *Glycobiology* **2017**, *27*, 3–49.
- [2] Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321–327.
- [3] A. A. Kulkarni, A. A. Weiss, S. Iyer, *Medical Research Reviews*, **2010**, *30*, 327–393.
- [4] R. Raman, S. Raguram, G. Venkataraman, J. C. Paulson, R. Sasisekharan, *Nat. Methods* **2005**, *2*, 817–824.
- [5] T. Rademacher, R. Parekh, R. Dwek, *Annu. Rev. Biochem.* **1988**, *57*, 785–838.
- [6] G. J. Boons, P. Wu, *Glycobiology* **2016**, *26*, 788.
- [7] R. S. Haltiwanger, J. B. Lowe, *Annu. Rev. Biochem.* **2004**, *73*, 491–537.
- [8] C. Reily, T. J. Stewart, M. B. Renfrow, J. Novak, *Nat. Rev. Nephrol.* **2019**, *15*, 346–366.
- [9] R. D. Cummings, J. M. Pierce, *Chem. Biol.* **2014**, *21*, 1–15.
- [10] A. Varki, H. H. Freeze, A. E. Manzi, *Curr. Protoc. Protein Sci.* **2009**, 1–10.
- [11] J. Guo, X. S. Ye, *Molecules* **2010**, *15*, 7235–7265.
- [12] O. Konur in *Marine Glycobiology*, CRC Press; **2016**, pp:7-21.
- [13] V. S. Raghavendra Rao, P. K. Qasba, P. V. Balaji, R. Chandrasekaraman in *Conformation of Carbohydrates*, CRC Press, **1998**, pp:1-72.
- [14] J. C. Muñoz-García, F. Corzana, J. L. De Paz, J. Angulo, P. M. Nieto, *Glycobiology*, **2013**, *23*, 1220–1229.
- [15] E. Juaristi, Y. Bandala in *Advances in Heterocyclic Chemistry*, Elsevier Inc., **2012**, *105*, pp:189-222.
- [16] G. A. Jeffrey, J. A. Pople, J. S. Binkley, S. Vishveshwara, *J. Am. Chem. Soc.* **1978**, *100*, 373–379.
- [17] K. Honke, *Med. Res. Rev.* **2002**, *22*, 637–654.
- [18] M. Rawat, C. I. Gama, J. B. Matson, L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* **2008**, *130*, 2959–2961.
- [19] F. Sasarman, C. Maftai, P. M. Campeau, C. Brunel-Guitton, G. A. Mitchell, P. Allard, *J. Inherit. Metab. Dis.* **2016**, *39*, 173–188.
- [20] S. Mizumoto, S. Ikegawa, K. Sugahara, *J. Biol. Chem.* **2013**, *288*, 10953–10961.
- [21] A. Malmström, B. Bartolini, M. A. Thelin, B. Pacheco, M. Maccarana, *J. Histochem. Cytochem.* **2012**, *60*, 916–925.
- [22] S. Mizumoto, S. Yamada, K. Sugahara, *Biomed Res. Int.* **2015**, 861752.
- [23] S. Yamada, K. Sugahara, S. Özbek, *Commun. Integr. Biol.* **2011**, *4*, 150–158.

- [24] D. Soares Da Costa, R. L. Reis, I. Pashkuleva, *Annu. Rev. Biomed. Eng.* **2017**, *19*, 1–26.
- [25] J. C. F. Kwok, P. Warren, J. W. Fawcett, *Int. J. Biochem. Cell Biol.* **2012**, *44*, 582–586.
- [26] R. V. Iozzo, A. D. Murdoch, *FASEB J.* **1996**, *10*, 598–614.
- [27] A. Almond, J. K. Sheehan, *Glycobiology* **2000**, *10*, 329–338.
- [28] T. Mikami, H. Kitagawa, *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830*, 4719–4733.
- [29] L. Djerbal, H. Lortat-Jacob, J. Kwok, *Glycoconj. J.* **2017**, *34*, 363–376.
- [30] T. Laabs, D. Carulli, H. M. Geller, J. W. Fawcett, *Curr. Opin. Neurobiol.* **2005**, *15*, 116–120.
- [31] J. Halper, *Adv. Exp. Med. Biol.* **2014**, *802*, 77–94.
- [32] J. Kreuger, D. Spillmann, J. P. Li, U. Lindahl, *J. Cell Biol.* **2006**, *174*, 323–327.
- [33] V. P. Eswarakumar, I. Lax, J. Schlessinger, *Cytokine Growth Factor Rev.* **2005**, *16*, 139–149.
- [34] K. Matsumoto, A. Wanaka, K. Takatsuji, H. Muramatsu, T. Muramatsu, M. Tohyama, *Dev. Brain Res.* **1994**, *79*, 229–241.
- [35] K. Kadomatsu, T. Muramatsu, *Cancer Lett.* **2004**, *204*, 127–143.
- [36] Y. Yoshida, S. Ikematsu, T. Moritoyo, M. Goto, J. ichiro Tsutsui, S. Sakuma, M. Osame, T. Muramatsu, *Brain Res.* **2001**, *894*, 46–55.
- [37] Merenmies, J., Rauvala, H., *J. Biol. Chem.*, **1990**, *28*, 16721–16724.
- [38] K. Kadomatsu, M. Tomomura, T. Muramatsu, *Biochem. Biophys. Res. Commun.* **1988**, *151*, 1312–1318.
- [39] K. Maruyama, H. Muramatsu, N. Ishiguro, T. Muramatsu, *Arthritis Rheum.* **2004**, *50*, 1420–1429.
- [40] Y. Takei, K. Kadomatsu, T. Muramatsu, S. Matsuo, H. Itoh, K. Nakazawa, S. Kubota, *Cancer Res.* **2001**, *61*, 8486–8491.
- [41] J. Wang, H. Takeuchi, Y. Sonobe, S. Jin, T. Mizuno, S. Miyakawa, M. Fujiwara, Y. Nakamura, T. Kato, H. Muramatsu, T. Muramatsu, A. Suzumura, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 3915–3920.
- [42] W. Iwasaki, *EMBO J.* **1997**, *16*, 6936–6946.
- [43] K. Unoki, N. Ohba, H. Arimura, H. Muramatsu, T. Muramatsu, *Investig. Ophthalmol. Vis. Sci.* **1994**, *35*, 4063–4068.
- [44] M. Horiba, K. Kadomatsu, K. Yasui, J. K. Lee, H. Takenaka, A. Sumida, K. Kamiya, S. Chen, S. Sakuma, T. Muramatsu, I. Kodama, *Circulation* **2006**, *114*, 1713–1720.
- [45] Aridome, K., Tsutsui, J., Takao, S., Kadomatsu, K., Ozawa, M., Aikou, T., Muramatsu, T. *J. Cancer Res.* **86**, 655–661, (1995).

- [46] S. Kojima, T. Inui, H. Muramatsu, Y. Suzuki, K. Kadomatsu, M. Yoshizawa, S. Hirose, T. Kimura, S. Sakakibara, T. Muramatsu, *J. Biol. Chem.* **1997**, *272*, 9410–9416.
- [47] M. Qi, S. Ikematsu, N. Maeda, K. Ichihara-Tanaka, S. Sakuma, M. Noda, T. Muramatsu, K. Kadomatsu, *J. Biol. Chem.* **2001**, *276*, 15868–15875.
- [48] C. Ueoka, N. Kaneda, I. Okazaki, S. Nadanaka, T. Muramatsu, K. Sugahara, *J. Biol. Chem.* **2000**, *275*, 37407–37413.
- [49] T. Muramatsu, *Curr. Pharm. Des.* **2011**, *17*, 410–423.
- [50] J. Lim, S. Yao, M. Graf, C. Winkler, D. Yang, *Biochem. J.* **2013**, *451*, 407–415.
- [51] H. A. Himburg, J. R. Harris, T. Ito, P. Daher, J. L. Russell, M. Quarmyne, P. L. Doan, K. Helms, M. Nakamura, E. Fixsen, G. Herradon, T. Reya, N. J. Chao, S. Harroch, J. P. Chute, *Cell Rep.* **2012**, *2*, 964–975.
- [52] H. J. Yeh, Y. Y. He, J. Xu, C. Y. Hsu, T. F. Deuel, *J. Neurosci.* **1998**, *18*, 3699–3707.
- [53] Y. S. Li, P. G. Milner, A. K. Chauhan, M. A. Watson, R. M. Hoffman, C. M. Kodner, J. Milbrandt, T. F. Deuel, *Science*, **1990**, *250*, 1690–1694.
- [54] R. Mentlein, J. Held-Feindt, *J. Neurochem.* **2002**, *83*, 747–753.
- [55] R. Jäger, K. Noll, K. Havemann, K. H. Pflüger, C. Knabbe, H. Rauvala, G. Zugmaier, *Int. J. Cancer* **1997**, *73*, 537–543.
- [56] M. Hatziapostolou, J. Delbe, P. Katsoris, C. Polytarchou, J. Courty, E. Papadimitriou, *Prostate* **2005**, *65*, 151–158.
- [57] G. Herradón, C. Pérez-García, *Br. J. Pharmacol.* **2014**, *171*, 837–848.
- [58] I. R. Taravini, M. Chertoff, E. G. Cafferata, J. Courty, M. G. Murer, F. J. Pitossi, O. S. Gershanik, *Mol. Neurodegener.* **2011**, *6*, 1–12.
- [59] D. M. Marchionini, E. Lehrmann, Y. Chu, B. He, C. E. Sortwell, K. G. Becker, W. J. Freed, J. H. Kordower, T. J. Collier, *Brain Res.* **2007**, *1147*, 77–88.
- [60] A. Kinnunen, T. Kinnunen, M. Kaksonen, R. Nolo, P. Panula, H. Rauvala, *Eur. J. Neurosci.* **1998**, *10*, 635–648.
- [61] K. Meng, A. Rodriguez-Peña, T. Dimitrov, W. Chen, M. Yamin, M. Noda, T. F. Deuel, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 2603–2608.
- [62] E. Raulo, S. Tumova, I. Pavlov, M. Pekkanen, A. Hienola, E. Klankki, N. Kalkkinen, T. Taira, I. Kilpeläinen, H. Rauvala, *J. Biol. Chem.* **2005**, *280*, 41576–41583.
- [63] E. Ryan, D. Shen, X. Wang, *FEBS J.* **2016**, *283*, 1488–1503.
- [64] C. Mikelis, E. Sfaelou, M. Koutsoumpa, N. Kieffer, E. Papadimitriou, *FASEB J.* **2009**, *23*, 1459–1469.

- [65] C. Mikelis, M. Lamprou, M. Koutsioumpa, A. G. Koutsioubas, Z. Spyrinti, A. A. Zompra, N. Spiliopoulos, A. A. Vradis, P. Katsoris, G. A. Spyroulias, P. Cordopatis, J. Courty, E. Papadimitriou, *J. Cell. Biochem.* **2011**, *112*, 1532–1543.
- [66] U. Rauch, P. Gao, A. Janetzko, A. Flaccus, L. Hilgenberg, H. Tekotte, R. K. Margolis, R. U. Margolis, *J. Biol. Chem.* **1991**, *266*, 14785–14801.
- [67] Y. Ito, M. Watanabe, T. Nishizawa, T. Omachi, T. Kobayashi, S. Kasama, O. Habuchi, J. Nakayama, *Acta Histochem. Cytochem.* **2007**, *40*, 53–59.
- [68] N. Maeda, K. Ichihara-Tanaka, T. Kimura, K. Kadomatsu, T. Muramatsu, M. Noda, *J. Biol. Chem.* **1999**, *274*, 12474–12479.
- [69] N. Maeda, T. Nishiwaki, T. Shintani, H. Hamanaka, M. Noda, *J. Biol. Chem.* **1996**, *271*, 21446–21452.
- [70] E. Nakamura, K. Kadomatsu, S. Yuasa, H. Muramatsu, T. Mamiya, T. Nabeshima, Q. W. Fan, K. Ishiguro, T. Igakura, S. Matsubara, T. Kaname, M. Horiba, H. Saito, T. Muramatsu, *Genes to Cells* **1998**, *3*, 811–812.
- [71] T. Muramatsu, *J. Biochem.* **2002**, *132*, 359–371.
- [72] C. Xu, S. Zhu, M. Wu, W. Han, Y. Yu, *Biol. Pharm. Bull.* **2014**, *37*, 511–520.
- [73] K. Kadomatsu, S. Kishida, S. Tsubota, *J. Biochem.* **2013**, *153*, 511–521.
- [74] S. D. Vallet, O. Clerc, S. Ricard-Blum, *J. Histochem. Cytochem.* **2021**, *69*, 93–104.
- [75] A. Sarkar, W. Yu, U. R. Desai, A. D. Mackerell, P. D. Mosier, *Glycobiology* **2016**, *26*, 1041–1047.
- [76] D. Shi, A. Sheng, L. Chi, *Front. Mol. Biosci.* **2021**, *8*, 1–15.
- [77] J. C. Muñoz-García, E. Chabrol, R. R. Vivès, A. Thomas, J. L. De Paz, J. Rojo, A. Imberty, F. Fieschi, P. M. Nieto, J. Angulo, *J. Am. Chem. Soc.* **2015**, *137*, 4100–4110.
- [78] M. José García-Jiménez, F. Corzana, J. L. De Paz, P. M. Nieto, *Nat. Prod. Commun.* **2019**, *14*, 1–6.
- [79] M. J. García-Jiménez, S. Gil-Caballero, Á. Canales, J. Jiménez-Barbero, J. L. de Paz, P. M. Nieto, *Int. J. Mol. Sci.* **2017**, *18*, 4–6.
- [80] J. C. Muñoz-García, M. J. García-Jiménez, P. Carrero, A. Canales, J. Jiménez-Barbero, M. Martín-Lomas, A. Imberty, J. L. de Paz, J. Angulo, H. Lortat-Jacob, P. M. Nieto, *Glycobiology* **2014**, *24*, 1004–1009.
- [81] C. Solera, G. Macchione, S. Maza, M. M. Kayser, F. Corzana, J. L. De Paz, P. M. Nieto, *Chem. A Eur. J.* **2016**, *22*, 2356–2369.
- [82] A. M. J. Garcia-jimenez, S. Gil-caballero, F. Corzana, F. Juarez-vicente, J. R. Miles, K. Sakamoto, K. Kadomatsu, M. Dominguez-garcía, J. L. De Paz, P. M. Nieto, *Chem. Eur. J.*, **2021**, *27*, 12395–12409.

Capítulo 2

Metodología

2.1 Introducción

Como se ha explicado en el capítulo anterior, los carbohidratos en sí mismos son inmensamente importantes para casi todos los aspectos de la vida (fotosíntesis, como fuente de energía...) y además se sabe que su unión a macromoléculas, en nuestro caso proteínas, juega un papel importante en la regulación de actividades biológicas y diversos procesos fisiológicos que tienen lugar en las células. En este último contexto, la capacidad de modelar adecuadamente los carbohidratos y sus complejos con proteínas es de gran importancia para este área. Por ello, este trabajo se va a centrar en el estudio estructural de una serie de oligosacáridos, tanto en su forma libre como asociada, siendo esto crucial para comprender las bases moleculares de la formación del complejo. La importancia de conocer la interacción molecular que tiene lugar ha generado un amplio abanico de técnicas biofísicas que permiten elucidar el complejo proteína-ligando, véase FP (*Fluorescence Polarization*), SPR (*Surface Plasmon Resonance*) e ITC (*Isothermal Titration Calorimetry*), entre otras. Para conocer la estructura atómica, tanto del ligando como del complejo, son imprescindibles las técnicas de Resonancia Magnética Nuclear (RMN), de difracción de rayos X y recientemente la microscopía electrónica de alta resolución.^[1] Pero debido a que los carbohidratos son difíciles de cristalizar en su forma libre, se han tenido que construir modelos 3D por mecánica y dinámica molecular para poder predecir su estructura tridimensional.

La modelización de carbohidratos ha sido hasta hace relativamente poco tiempo un campo poco explorado. Fenómenos estereoelectrónicos como los efectos anomérico y exoanomérico no son fácilmente modelables con campos de fuerzas, esencialmente empíricos. La reparametrización con Glycam de los campos de fuerzas AMBER y la mejora de los parámetros específicos de GROMACS y CHARMM, han mejorado considerablemente la calidad de los cálculos de dinámica molecular.^[2] Al mismo tiempo se han beneficiado de los avances más recientes como los métodos PME (*Particle mesh Ewald*) y el empleo de caja periódica de disolvente.^[3] La combinación de ambos factores ha permitido que sea posible reproducir estructuras 3D de oligosacáridos mediante dinámica molecular.

Para el estudio de complejos intermoleculares carbohidrato-proteína se ha recurrido a métodos de menor calidad y mayor rendimiento. Generalmente se utiliza docking que permite generar un gran número de potenciales estructuras del complejo en un tiempo relativamente corto. Se suele emplear para generar las posibles estructuras del complejo que luego han de ser valoradas.

Recientemente, han resurgido cálculos basados en antiguos métodos como Poisson-Boltzman y Generalized Born, véase MMPBSA (*Molecular Mechanics Poisson-Boltzmann Surface Area*) y MMGBSA (*Molecular Mechanics Generalized Born Surface Area*), que combinados con MD (*Molecular Dynamics*) o SD (*Stochastic Dynamics*) permiten evaluar la energía libre de asociación por estimación de la entropía.^[4]

2.2 Métodos teóricos

2.2.1 Química Computacional

En el contexto de esta tesis, el objetivo principal es estudiar estructuralmente biomoléculas flexibles (como los GAGs) mediante cálculos teóricos, generando estructuras tridimensionales que justifiquen las propiedades experimentales.

En general, el marco más preciso para este estudio es la mecánica cuántica (QM) que se basa en dos aproximaciones: HF SFC (*Hartree-Fock Self Consistent Field*) o DFT (Teoría del Funcional Densidad), pero el principal inconveniente que presenta es el elevado coste computacional que supone.^{[5],[6]} Continuando en la escala de precisión, se encuentran los métodos semiempíricos, para los que se recurre a parametrizaciones experimentales evitando cálculos intensivos. El siguiente nivel son los métodos basados en campos de fuerzas. Estos calculan las energías empleando ecuaciones de un oscilador armónico que han de ser previamente parametrizadas, y solo son válidas en situaciones cercanas al equilibrio térmico. De estos derivan los cálculos de dinámica molecular en los que a partir del campo de fuerzas se resuelven las ecuaciones de Newton del movimiento permitiendo la interconversión entre energía cinética y potencial, simulando el comportamiento de las moléculas a lo largo del tiempo.

Los cálculos en esta tesis están basados en “campos de fuerzas”, siendo éstos un conjunto de ecuaciones con sus correspondientes parámetros. Para reducir significativamente estos cálculos se usan una serie de aproximaciones, permitiendo abordar la mayoría de los sistemas moleculares. La más importante de todas es la aproximación de *Born-Oppenheimer*, que afirma que los movimientos nucleares y electrónicos son independientes entre sí, y que, debido a la menor masa que presentan los electrones, permite calcular la energía de la molécula en función de sus coordenadas nucleares.^[7] Estos cambios de energía son considerados como movimientos en una superficie multidimensional que puede denominarse superficie de energía potencial (PES), representando las conformaciones moleculares en función de su estructura y energía.^[5]

2.2.2 Métodos de campos de fuerza

A través de ecuaciones empíricas derivadas de la mecánica clásica podemos calcular funciones de energía potencial, denominándose el conjunto de éstas, campos de fuerza. Los campos de fuerza relacionan la geometría con la energía y sus variaciones en la zona próxima al equilibrio y contienen la descripción de los cambios en la energía en función de las posiciones relativas de los átomos.

El potencial de campo de fuerza viene determinado por una combinación de términos independientes (ecuación 1):

$$E_{Pot} = \sum_{enlaces} V_{enl} + \sum_{\text{ángulos}} V_{ang} + \sum_{diedros} V_{died} + \sum_{NB} V_{NB} + \sum_{elect} V_{elec} \dots \quad (1)$$

Ecuación 1. *Expresión matemática general para un campo de fuerzas.*

Aunque estas ecuaciones se han desarrollado para una gran variedad de sistemas, como moléculas pequeñas orgánicas, proteínas, ácidos nucleicos y lípidos, los recursos computacionales para realizar cálculos en cadenas o secuencias oligosacáridicas han sido más escasos debido a su naturaleza química. Los efectos anomérico y exoanomérico son de naturaleza estereoelectrónica y por lo tanto difíciles de modelar basándonos en ecuaciones simples.

Por su naturaleza entre moléculas orgánicas y biopolímeros, los campos de fuerzas para carbohidratos se pueden clasificar en dos grupos: aquellos que se desarrollan con simulaciones en un contexto biológico más amplio (CHARMM, GLYCAM, GROMOS, OPLS, etc.)^[8] y los derivados de moléculas orgánicas (TRIPOS, MM3, MM4, etc.)^[9] Los primeros usan una clasificación por residuos en el biopolímero para asignar los parámetros, mientras que en los segundos son los tipos de átomos los que determinan la asignación (Figura 2.1).

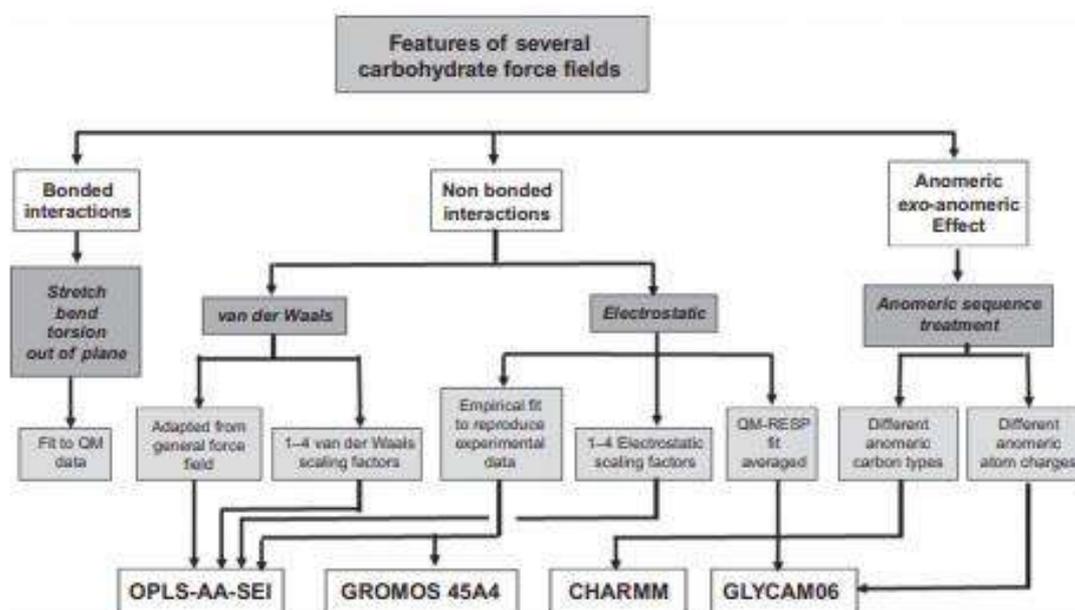


Figura 2.1. *Protocolo de parametrización entre los campos de fuerzas para carbohidratos.*^[10]

Estos campos de fuerza se han empleado regularmente para simulaciones de dinámica molecular de complejos proteína-ligando, pero a pesar de las ventajas que presentaban, no estaban diseñados para la obtención de energías libres o entalpías en el acoplamiento de complejos carbohidrato-proteína.

La presencia de grupos sulfato y sulfamato en los GAGs ha supuesto una importante dificultad a la hora de su modelado. Las aplicaciones más tempranas se centraron en

dilucidar las preferencias conformacionales de aquellos que no estaban sulfatados. La aparición de parámetros para O-sulfonato y N-sulfamato^[11] compatibles con los campos de fuerzas AMBER y CHARMM permitió una comprensión más amplia acerca de la torsión, flexibilidad, interacciones intra- e intermoleculares y los efectos anoméricos y exoanoméricos que gobiernan sobre estos carbohidratos.^[12] Además, a día de hoy, el avance en el desarrollo de los campos de fuerzas permite simulaciones de microsegundos de secuencias largas de GAGs para estudiar así su interacción con proteínas en presencia de disolvente. Pero la expectativa de molecular las propiedades únicas de los GAGs, como una conformación inusual o un dinamismo conformacional con consecuencias biológicas interesantes, ha convertido la parametrización en un objetivo relevante

El campo de fuerzas AMBER, que inicialmente se parametrizó para su uso en proteínas, y posteriormente se extendió a ácidos nucleicos, pertenece a los campos de fuerzas específicos para biomoléculas. Esta primera versión de un campo de fuerzas específico para carbohidratos, obtenido por parametrización, fue ampliado empleando para ello información estructural detallada procedente de cálculos DFT/ab initio (B3LYP/6-31G*) en modelos de carbohidratos (parámetros de Senderovitz).^{[13],[14],[15]} Ambas versiones se encuentran en la versión AMBER* de Macromodel que contiene estos parámetros ampliados y validados por cálculos de energía libre sobre glicanos simples y disacáridos. Sin embargo, esta versión usa campos de fuerzas basados en tipos de átomos y no en residuos.

Debido a la cada vez mayor relevancia de sistemas biológicos que involucran glicoproteínas o glicopéptidos, fue necesario optimizar campos de fuerzas de biomoléculas adicionales para el manejo de carbohidratos. Dentro de los campos de fuerzas desarrollados en este ámbito, y compatible con AMBER, se encuentra GLYCAM, que admite una amplia biblioteca de secuencias de glicanos incluyendo cadenas de GAG insaturadas en el extremo no reductor. A lo largo de los años, versiones actualizadas de la mano de parametrizaciones basadas en el método Hartree-Fock (HF), con un conjunto de bases 6-31G*, permitieron dinámicas moleculares en el estudio de complejos proteína-glicano.^[16] Debido a que el agua ejercía un papel esencial en el equilibrio de las interacciones moleculares de los diferentes rotámeros, para una dinámica molecular es necesario emplear las condiciones de caja periódica, lo que implica usar el método PME para los cálculos de las interacciones electrostáticas.^[17] Esto puso de manifiesto la necesidad de elaborar campos de fuerzas nuevos, diseñados desde el principio para ser usados en condiciones de caja periódica y PME.^[18] En el caso de AMBER, se desarrolló GLYCAM06 y versiones siguientes, que desembocaron en mejores resultados, similares a los observados experimentalmente, incluyendo una caja de agua explícita TIP3P.^[19] Cabe destacar que el empleo de este modelo de agua está basado en diversos estudios que corroboraron que este tipo de solvatación proporcionaba una estructura uniforme y una dinámica molecular más rápida.^[20] Además, se observó que para simulaciones de complejos proteína-ligando, cuando tenía lugar esta unión, las moléculas de agua que se encontraban al lado de los residuos de unión, desaparecían para que dicho acercamiento fuese crucial durante la complejación.^[21]

En general, el uso de GLYCAM06 está ampliamente extendido para modelar carbohidratos, glicoproteínas y glicolípidos, así como complejos proteína-carbohidrato. Además, este campo de fuerzas puede usarse en paquetes de simulación diferentes a AMBER, convirtiendo los archivos con el uso de las herramientas adecuadas. Para facilitar esta transferencia de parámetros, todas las secuencias atómicas tienen un conjunto definido en términos de torsión (sin términos genéricos), y los parámetros *parm94* (su versión actualizada *parm99*^[22]) para términos de van der Waals.^{[23],[24],[25]} Más recientemente, la versión GLYCAM06 se ha actualizado con un conjunto de parámetros más amplio para muchos derivados de carbohidratos que se encuentran comúnmente en sistemas biomoleculares complejos, como el ácido N-acetilneuramínico (ácido siálico), N-acetilglucosamina, N-acetilgalactosamina, N-acetilmanosamina y ácidos glucurónico y galacturónico. Además, esta versión es el único campo de fuerza de carbohidratos actual que contiene parámetros de enlaces N-glicosídicos,^[26] cuya combinación con *parm99sb* de AMBER, que contiene parametrización de proteínas, permite simular glicopéptidos, glicolípidos y glicoproteínas,^[27] describiendo además las propiedades fisicoquímicas de éstos. Una característica distintiva de este campo de fuerzas en comparación con otras versiones de GLYCAM es el uso del mismo tipo de átomos de carbono (CG) para los centros anoméricos α - y β -glucósido, facilitando la interconversión del anillo si procede o el equilibrio entre conformaciones con sustituyentes ecuatoriales y axiales.^[2]

2.2.3 Cargas Parciales

Mediante el campo de fuerzas GLYCAM06 se pueden simular dinámicas moleculares en cajas de agua, generando un conjunto de conformaciones de carbohidratos y reproduciendo además sus energías de rotación si se combina con un set de cargas apropiado. Estas cargas parciales se calcularon inicialmente para cada residuo basándose en el ajuste del promedio del potencial electrostático molecular de la mecánica cuántica (ESP).^[28] Dicha combinación es capaz de reproducir energías de interacción entre moléculas polares al incluir enlaces de hidrógeno y así eludir la necesidad de potenciales de interacción adicionales. A pesar de que esta característica puede ser muy útil en algunas situaciones, en simulaciones donde el carbohidrato se encuentra solvatado, las conformaciones y la dinámica son muy sensibles al equilibrio de fuerzas entre las interacciones agua-azúcar y azúcar-azúcar, involucrando además enlaces de hidrógeno entre los grupos hidroxilo. Por ello, el protocolo actualmente empleado es un refinamiento del método anterior, utilizando restricciones sobre este potencial electrostático (RESP).^{[29],[24]} Este permite usar un conjunto de cargas parciales de magnitudes reducidas que predicen con mayor precisión la fuerza de las interacciones intermoleculares de corto alcance sin que se vean afectados los momentos dipolares.^[30] En cuanto al modelo *ab initio* apropiado para la derivación de las cargas RESP,^[31] se ha demostrado que el conjunto de bases 6-31G*^[32] es capaz de proporcionar una excelente reproducción de las propiedades intermoleculares en fase condensada, como entalpías y densidades líquidas, así como energías libres de solvatación.^[29] Además, este modelo está bien equilibrado con los modelos de agua comúnmente disponibles, véase el TIP3P entre otros.^[33]

2.2.4 Simulaciones Moleculares

La Dinámica Molecular agrupa un conjunto de técnicas de simulación que predicen el comportamiento de átomos o moléculas a lo largo del tiempo, mediante las leyes del movimiento de los átomos clásicas o Newtonianas, permitiendo calcular la evolución de las coordenadas atómicas en el tiempo (trayectoria) de una molécula o sistema a una temperatura determinada. Se utiliza para explorar la relación entre estructura, movimiento y función.

Para obtener resultados significativos y poder explorar el espacio conformacional, es necesario realizar dinámicas de una duración lo suficientemente larga para que los movimientos objeto de estudio puedan ser muestreados. Es importante señalar que el tiempo está severamente limitado por la potencia informática disponible, aunque los avances actuales en la tecnología informática y los algoritmos del software nos permiten muestrear el espacio de sistemas biomoleculares en el orden de cientos de nanosegundos (ns). Este marco de tiempo suele ser adecuado para la mayoría de los movimientos internos de los carbohidratos. El procedimiento común para ejecutar las dinámicas moleculares estables en sistemas de glicosaminoglicanos con AMBER^[22] se muestra en la Figura 2.2.

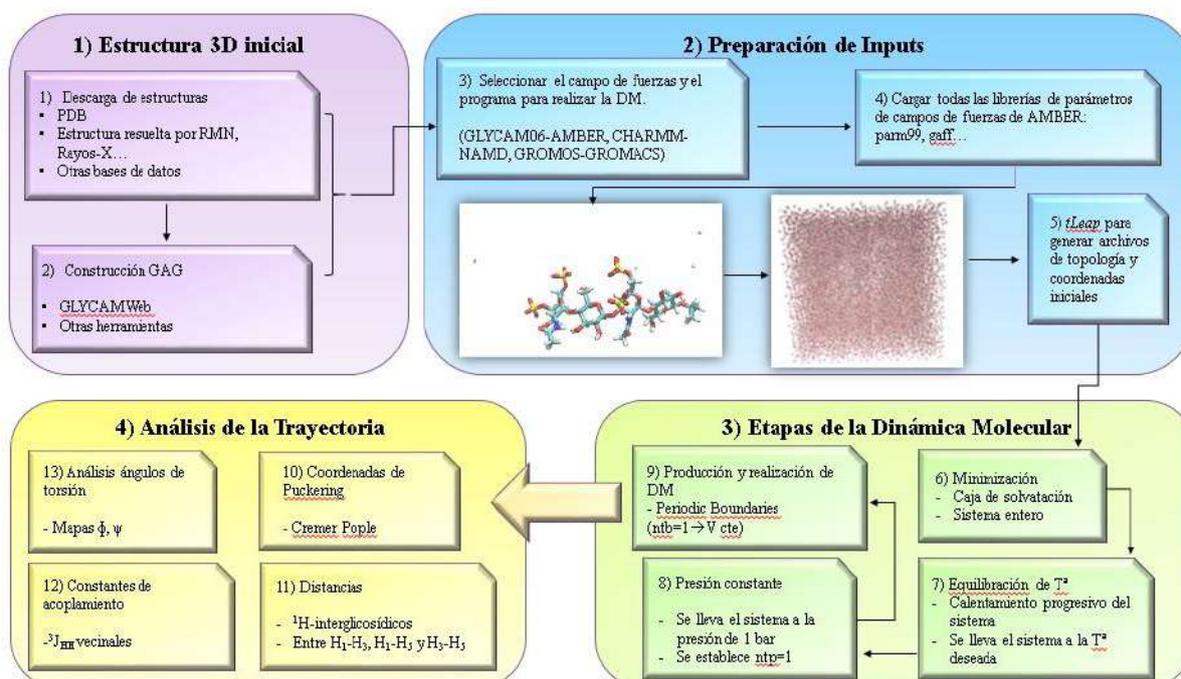


Figura 2.2. Esquema gráfico de los pasos para realizar una Dinámica Molecular y su análisis..

Las simulaciones moleculares constan de varias etapas previas, que se centran principalmente en la preparación de los archivos de entrada requeridos en el cálculo. En el caso de las proteínas, se parten de estructuras que se han obtenido previamente mediante RMN o rayos X, que pueden encontrarse en bases de datos estructurales como Protein Data Bank (PDB). En el caso de los ligandos, las estructuras iniciales pueden

construirse a partir de los residuos implementados en la interfaz gráfica Maestro de Schrödinger Suite.

Una vez construido y generado el modelo estructural inicial, hay que tener en cuenta el entorno que lo rodea, ya que la imitación rigurosa de las condiciones realistas del disolvente durante la simulación es clave para producir resultados comparables y con poder predictivo con respecto a los datos experimentales. Los cálculos se llevan a cabo a pH 7, empleando programas como PROPKA para tener en cuenta los estados de protonación de proteínas.^[34] Un ejemplo de aminoácidos que se deben tener en cuenta en nuestro caso es la histidina (Figura 2.3) que, dependiendo de la carga neta y la localización de sus protones, puede adoptar tres estados distintos: HIP (carga +1, con los nitrógenos δ y ϵ protonados), HID (neutro, con el nitrógeno- δ protonado) e HIE (neutro, nitrógeno- ϵ protonado).^{[35],[36]}

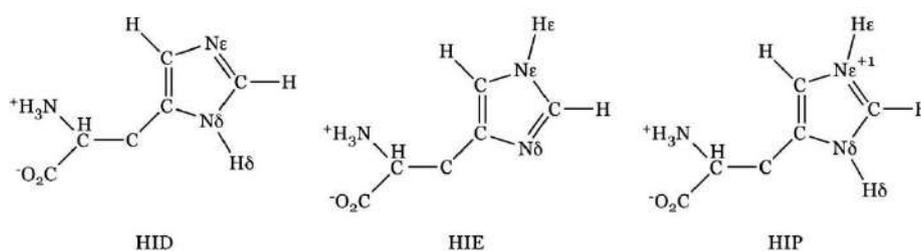


Figura 2.3. Posibles estados de protonación de la Histidina.^[35]

A continuación, el sistema receptor-ligando o molécula de estudio se sumerge en agua explícita fomentando así las interacciones a través de enlaces de hidrógeno, y equilibrando además la cantidad de iones de sal añadidos, asegurando que la carga final del sistema sea neutra. Una vez preparado este modelo estructural inicial se debe abordar la parametrización del sistema. Dependiendo si se trabaja en presencia de proteína o no, los campos de fuerzas generalizados que se emplean son diferentes. Desafortunadamente, debido a la complejidad estructural de muchos compuestos, estos campos de fuerza pueden no contener todos los parámetros que describan en su totalidad a la molécula de estudio; en este caso, por lo tanto, deberían derivatizarse utilizando *ab initio*.^[37]

A continuación, se realizan dos etapas de minimización, restringiéndose primero tanto la molécula de estudio como los contraiones, seguido de una relajación de las moléculas de agua en función de las iteraciones deseadas. A este proceso le siguen varias etapas de estabilización del sistema, incrementando la temperatura hasta llegar a 300 K, a una presión constante de 1 bar. Una vez que el sistema se encuentra en completo equilibrio, tiene lugar la etapa más importante de la simulación, que es la etapa de producción donde se produce la simulación real obteniendo los archivos de coordenadas finales para cada *time-step* y que se utilizarán para los posteriores análisis teóricos. En este punto, conviene revisar los archivos *mdout* generados, ya que de ellos se pueden extraer los datos de energía potencial, cinética y total controlando que no haya ninguna alteración durante la simulación. Una opción para hacer esto es usar el software *gnuplot* que, mediante las representaciones gráficas de estos datos de energía, permite observar la presencia de fluctuaciones en cada una de las etapas si las hubiere.^[38]

2.2.5 Dinámicas moleculares promediadas en el tiempo (MD-TAR)

En caso de que exista un equilibrio conformacional entre varias estructuras discretas que no pueda ser simulado por la dinámica molecular tradicional, es frecuente recurrir a la MD con restricciones promediadas en el tiempo.^[39] Para ello se utilizan restricciones externas generalmente procedentes de los experimentos de RMN, normalmente distancias H-H procedentes de datos de NOE o acoplamiento $^3J_{HH}$, que se introducen en la MD en forma de energía potencial adicional (V_{rest}).

$$V_{Total} = V_{teórica} + V_{rest} \quad (2.1)$$

Este método permite que el sistema salte barreras de potencial sin caer en conformaciones virtuales.^{[40],[41]} Es posible incluir disolvente explícito, con las mismas condiciones que en las dinámicas sin restricciones. Las MD-TAR se van a caracterizar por dos parámetros fundamentalmente: uno de ellos es la constante de fuerza K_{dr} que determina la restricción de fuerza adicional (ecuación 2.2):^[42]

$$V_{rest} = K_{dr} \cdot (\langle R_{ij} \rangle - R_{exp})^2 \quad (2.2)$$

Para aquellas simulaciones cuyas restricciones proceden de datos experimentales originados a partir de mediciones de NOE, se puede tener en cuenta el parámetro correspondiente al tiempo de desintegración (τ_{dr}) determinando la duración del experimento para promediar las distancias durante la simulación.^[42] Por ejemplo, en los experimentos de NOE, la distancia entre dos protones está relacionada con la velocidad de relajación dipolar entre los respectivos núcleos. Tales interacciones descienden con la sexta potencia de la distancia y dado que el valor medido se determina en un tiempo lo suficientemente largo como para que ocurra la interconversión conformacional, el valor experimental que podemos extraer es (ecuación 2.3):

$$r_{NOE} = \langle r(t)^{-6} \rangle^{-1/6} \quad (2.3)$$

donde $r(t)$ es el conjunto de distancias muestreadas por la estructura durante el período en el que el experimento de RMN se está realizando. Sin embargo, si se da un equilibrio conformacional en el que uno de los conformeros presenta dos protones a una distancia en la que se observaría NOE pero el otro no, la distancia promedio puro dirigiría a una conformación virtual inexistente. Mediante la técnica “*time averaged restraints*” el programa almacenaría en una variable temporal durante cuánto tiempo una distancia determinada está siendo sobrepasada, aumentando la energía potencial del sistema paralelamente hasta que provoque el cambio conformacional correspondiente controlado por el campo de fuerzas.

Por tanto, estas dos ecuaciones principales derivan a la expresión matemática final que definen las dinámicas con restricciones promediadas en el tiempo (ecuación 2.4):^[41]

$$(\langle R_{ij}^{-6} \rangle^{-1/6}) = \frac{\sum_0^t R(t)^{-6}^{-1/6}}{N} \quad (2.4)$$

2.2.6 Cálculos de Interacción

Actualmente se encuentran disponibles muchos enfoques para estudiar las interacciones proteína- ligando, bien mediante acoplamiento molecular (*docking*), con el fin de predecir la estructura del complejo resultante y la potencial actividad biológica de un ligando dado, o mediante MM-GB/PBSA (*Molecular Mechanics with Poisson-Boltzmann or Generalized Born and Surface Area Solvation*).^[43] Esto requiere una representación adecuada de las estructuras y propiedades moleculares, algoritmos de búsqueda para muestrear de manera eficiente el espacio de unión, localización y análisis de geometrías que presenten interacciones favorables, y su evaluación y clasificación. En general, las investigaciones computacionales aplican simulaciones de acoplamiento molecular o de dinámica molecular para cuantificar y explorar la interacción entre ligando y receptor. En principio es necesario resolver dos aspectos: primero la búsqueda del sitio y modo de la interacción, para lo cual se suele recurrir a métodos de alto rendimiento de optimización de variables, normalmente basados en algoritmos Monte-Carlo^{[44],[45]}; y segundo, la cuantificación de la energía de interacción. Para el primer problema existen métodos capaces de generar un gran número de conformeros de moléculas pequeñas y evaluar su interacción con un receptor. Para el segundo, se recurre a evaluar la energía de interacción de las “poses” seleccionadas, normalmente mediante cálculos de energía simple (electrostática, van der Waals, tensión del ligando) parametrizados para ser usados en este contexto. Recientemente han surgido aplicaciones de métodos antes empleados para el cálculo de energías de solvatación, basados en los métodos de Poisson Boltzmann y Born Generalizado, que permiten estimar con mayor precisión la energía libre de enlace ($\Delta G_{\text{binding}}$), incluyendo aspectos como la entropía y la solvatación.^[46]

2.2.6.1 Docking (Acoplamiento Molecular o Screening Virtual)

El acoplamiento molecular o *docking* se ha convertido en una poderosa herramienta para explorar y comprender las interacciones ligando-receptor a nivel molecular. Predice los modos de formación de estos complejos, proponiendo una orientación y conformación del ligando en la cavidad del receptor o incluso, en algunos casos, llegando a identificar posibles sitios de unión.^[47] El proceso de acoplamiento no es sencillo, ya que esta unión está influenciada por diversos factores entrópicos y entálpicos. La movilidad tanto del ligando como del receptor, el efecto de la proteína sobre la distribución de cargas del ligando y las interacciones con las moléculas de agua que lo rodean dificultan la descripción cuantitativa de la interacción.^[48] En concreto, el acoplamiento ligando-proteína ocupa un lugar bastante importante en el *docking* por sus aplicaciones en el área de la medicina. Este tipo de aproximaciones son bastante frecuentes cuando se dispone de una estructura resuelta del receptor o bien se tiene un modelo de la misma construido por homología. Esta técnica genera un conjunto de conformaciones del complejo a partir de un muestreo tanto de la conformación como de la orientación del ligando en su unión

al sitio activo del receptor, haciendo una clasificación de sus mejores “poses” en base a su estabilidad y energía de afinidad ponderada.

Debido a que el acoplamiento molecular juega un papel fundamental en el diseño y descubrimiento de fármacos se han desarrollado diferentes metodologías como potentes filtros computacionales, que se suelen utilizar antes del cribado experimental para reducir la mano de obra y los costes. Básicamente todos estos softwares contienen dos componentes comunes: una función de puntuación (*docking score*), que permite calcular el mínimo global de energía, debiendo coincidir con el mínimo global de energía libre del sistema proteína-ligando; y un algoritmo de búsqueda que muestrea exhaustivamente todos los posibles modos de unión con el receptor. Para predecir estas energías libres de enlace existen varios tipos de funciones de puntuación, que se pueden englobar en tres grupos: empíricas, basadas en campos de fuerza o basadas en el conocimiento. Por otro lado, el chequeo conformacional del ligando mediante los algoritmos de búsqueda puede llevarse a cabo mediante: métodos sistemáticos (construcción incremental, bases de datos, búsqueda conformacional), métodos aleatorios o estocásticos (Monte Carlo, algoritmos genéticos, búsqueda tabú) y métodos de simulación (dinámicas moleculares o de minimización energética).

El contrapunto de estas técnicas es que las predicciones de estas “poses” de unión observadas no es completamente fiable, por estar basadas en aproximaciones. Actualmente se siguen dedicando considerables esfuerzos en el desarrollo de métodos computacionales más precisos,^{[49],[50]} cuyas diferencias radican en estos algoritmos de búsqueda, el manejo del ligando y flexibilidad de la proteína, las funciones de puntuación y el tiempo de cálculo requerido para evaluar la afinidad del sistema (Tabla 2.1).

Programas Informáticos	Método de búsqueda conformacional	Funciones de Puntuación
Dock ^[51]	Sistemático y Simulación	Basadas en campos de fuerza
AutoDock ^[52]	Aleatorio/Estocástico y Simulación	Basadas en campos de fuerza
GOLD ^[53]	Aleatorio/Estocástico	Basadas en campos de fuerza
FlexX ^[54]	Sistemático	Empíricas
Glide (Schrodinger) ^[55]	Sistemático y Simulación	Empíricas

Tabla 2.1. Herramientas más comunes para el docking.

La limitación más significativa de este proceso es el uso de un receptor rígido reduciendo el tamaño del espacio conformacional, suponiendo un esfuerzo computacional menor. Pero, sin duda, el mayor avance fue el empleo de rejillas (*grids*) tridimensionales rectangulares que permitían llevar a cabo búsquedas de diferentes conformaciones del ligando en un espacio rotacional y translacional de seis dimensiones mediante algoritmos

de búsqueda. Mediante este método se discretiza el sitio activo de estudio, o el mismo receptor, del espacio que lo rodea, reduciendo notablemente el tiempo de cálculo, evaluando, como se ha comentado antes, las diferentes poses finales a través de funciones de puntuación. Para el acoplamiento automatizado existen dos categorías: la primera es un método de emparejamiento donde un ligando flexible se ensambla a un cuerpo rígido utilizando un algoritmo de coincidencia geométrica. En la segunda técnica, el ligando explora fuera de la proteína hasta encontrar el sitio activo, variando sus conformaciones y orientaciones. Basándose en este último método, el programa Autodock predice el acoplamiento computacional y el cribado virtual de moléculas pequeñas en receptores macromoleculares (ver Figura 2.4).^[56] Este programa está englobado en los métodos aleatorios o estocásticos y está basado en un campo de fuerzas empírico de energía libre, usando un algoritmo genético Lamarquiano como base para el cálculo. Además este *software* ha evolucionado adaptándose a nuevos problemas y, por ejemplo, la extensión Autodock Vina^[57] ofrece una predicción más rápida y precisa del modo de unión utilizando el algoritmo de *Broyden-Fletcher-Goldfarb-Shanno* (BFGS).^[58]

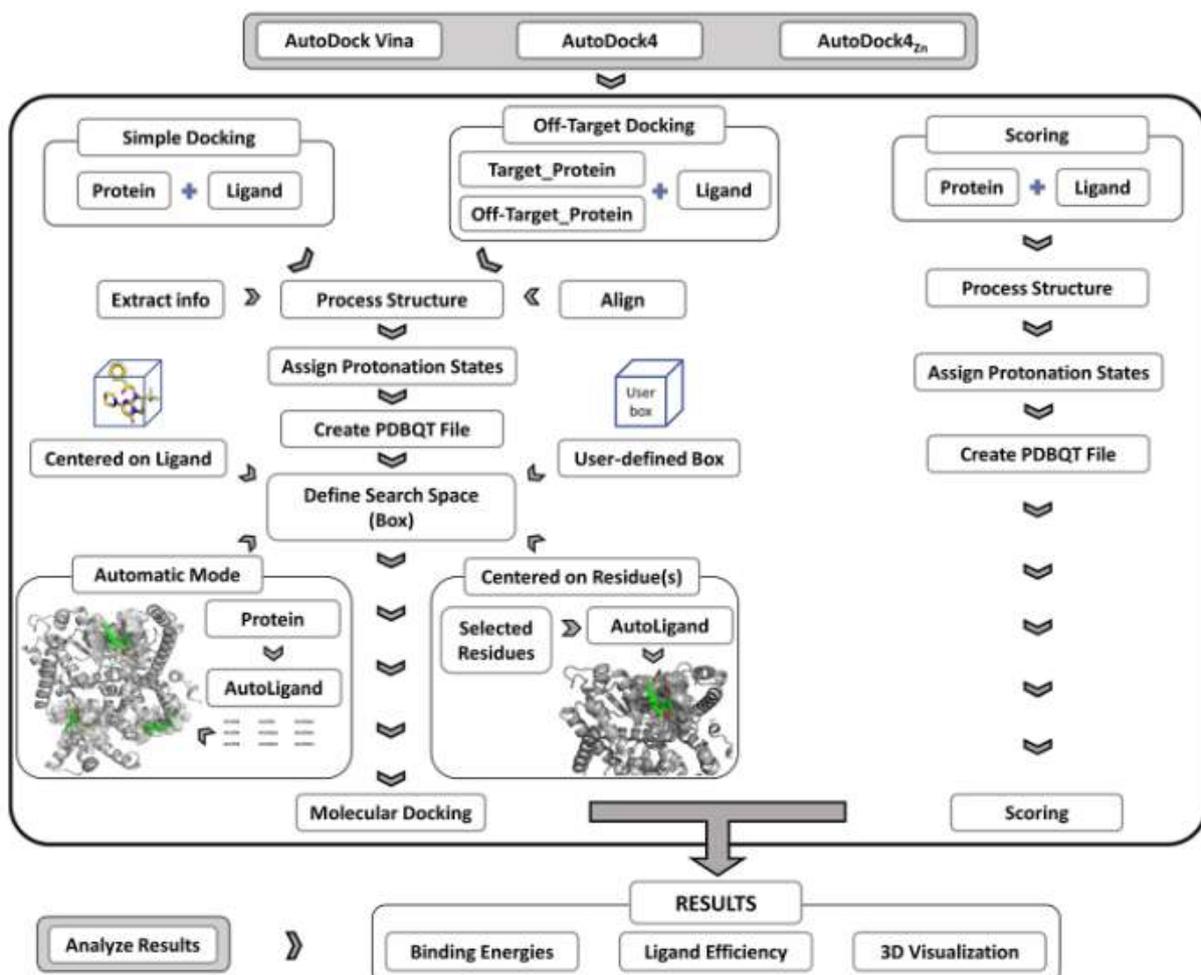


Figura 2.4. Esquema de docking para el programa Autodock.^[59]

2.2.6.2 Métodos de MM/PBSA y MM/GBSA

Desde hace décadas se realizan cálculos de energías libres de unión basados en simulaciones de mecánica molecular mediante dinámicas moleculares o métodos Monte Carlo.^[60] Estas técnicas permiten determinar los cambios de energía libre de unión de pequeños ligandos a macromoléculas biológicas a partir de simulaciones. Así, estos métodos realizan un riguroso tratamiento de todos los grados de libertad del complejo proteína-ligando incluyendo modelos de solvatación continuos, suponiendo un coste computacional considerablemente grande. En un intento de combinar precisión y rapidez, se han desarrollado varias aproximaciones en las que destacan los métodos de MM/PBSA y MM/GBSA (*Molecular Mechanics with Poisson-Boltzmann (PB) or Generalized Born (GB) and Surface Area Solvation*).^[43] Desarrollado por Srinivassan y Kollman en 1998, esta técnica está basada en mecánica estadística, que contiene distintos términos fisicoquímicos que intervienen en el proceso de unión ligando-proteína, esquematizado en la Figura 2.5 mediante un ciclo termodinámico:

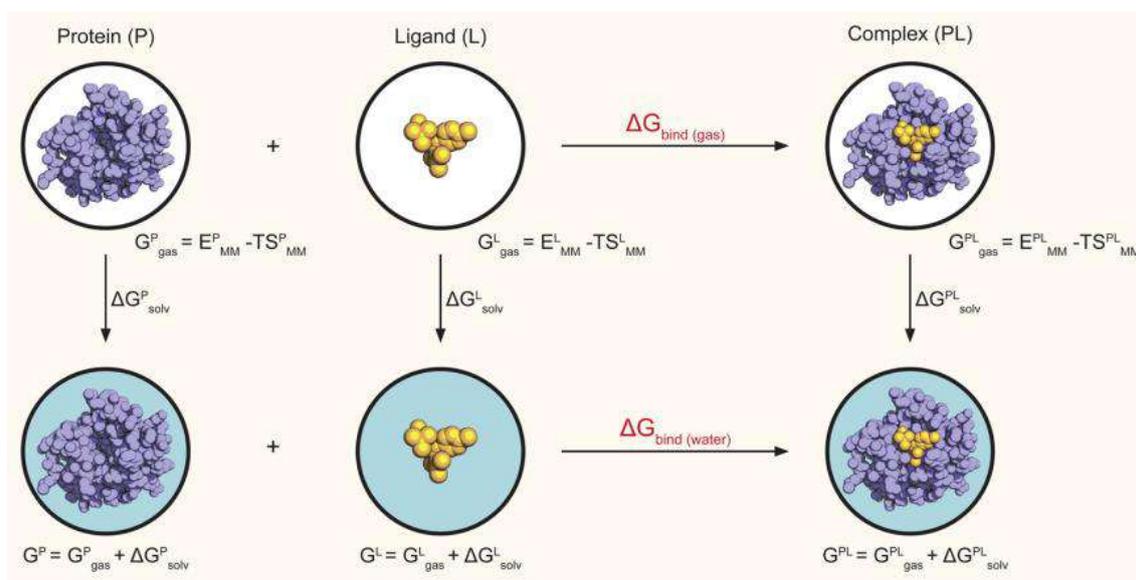


Figura 2.5. Esquema termodinámico para la obtención de la energía libre. Los fondos blanco y azul indican un sistema en vacío y en disolución acuosa, respectivamente. La coloración naranja del ligando establece que está interactuando completamente con el medio.^[61]

Como se observa en el esquema anterior, la unión entre el ligando y el receptor, previamente solvatados por moléculas de agua, dará lugar a un complejo que se encuentra estabilizado por las interacciones intermoleculares no enlazantes. Durante este proceso, se produce un cambio entrópico debido a la reducción de la libertad conformacional del ligando y al desordenamiento de las moléculas de agua que lo rodean (denominado efecto hidrofóbico), que previamente se encontraban ordenadas cuando se encontraban aisladas ambas especies. Basándose en este fundamento termodinámico, la energía libre se puede evaluar a partir de las energías de la mecánica molecular combinadas con un término de energía de solvatación calculado a partir del área de superficie accesible (SASA), una

estimación de la energía electrostática mediante Poisson-Boltzmann y un término entrópico.

La energía libre (ΔG_{bind}) entre el ligando (L) y el receptor (R) para formar el complejo RL es calculado mediante las siguientes ecuaciones matemáticas (ecuaciones 2.5-2.7):

$$\Delta G_{\text{binding}} = \Delta H - T\Delta S \approx \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \quad (2.5)$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{internal}} + \Delta E_{\text{electrostatic}} + \Delta E_{\text{vdw}} \quad (2.6)$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{PB/GB}} + \Delta G_{\text{SA}} \quad (2.7)$$

donde ΔE_{MM} , engloba las energías de unión de la mecánica molecular: $\Delta E_{\text{internal}}$ (enlace, ángulo y diedro), interacciones electrostáticas ($\Delta E_{\text{electrostatic}}$) y de Van der Waals (ΔE_{vdw}). Por otro lado, el término ΔG_{sol} indica la suma de la energía libre de solvatación electrostática ($\Delta G_{\text{PB/GB}}$), y el componente de solvatación ΔG_{SA} (contribución polar y no polar, respectivamente). El cambio de entropía conformacional, $-T\Delta S$, normalmente es calculado por un método normal de análisis a partir de un conjunto de instantáneas conformacionales o *snapshots* tomadas de las simulaciones de dinámica molecular.^{[62],[63]}

El término electrostático supone la contribución más importante, debido a la fuerza existente en las interacciones soluto-solvente y se evalúa a partir de modelos continuos de disolvente: bien a partir de la resolución de la ecuación de Poisson-Boltzmann (ΔG_{PB} , siendo el método MM-PBSA) o mediante un modelo Generalizado de Born (ΔG_{GB} , siendo el método MM-GBSA). Aunque se obtienen resultados similares de ambas técnicas, ésta última requiere menor tiempo de cálculo. Los protocolos para poder aplicar ambos métodos pueden ser dos: i) realizarlos sobre *snapshots* del ligando, proteína y complejo extraídos de una única dinámica molecular de este último, o ii) realizar tres simulaciones independientes de cada especie, extrayendo estas instantáneas de cada una de ellas.^{[64],[65]}

Al ser métodos empíricos, es importante: investigar sistemáticamente ciertos factores como la duración de las simulaciones y la constante dieléctrica del soluto (debe ser la adecuada para poder calcular las energías polares de solvatación); realizar cálculos de entropía precisos; y evaluar la energía libre de unión comparando ambas técnicas, para poder clasificar las afinidades de los ligandos con el receptor.^[63]

2.3 Métodos experimentales

Diversas técnicas biofísicas han sido desarrolladas para caracterizar complejos proteína-ligando, véase experimentos de SPR (*Surface Plasmon Resonance*), que permiten obtener parámetros cinéticos de interacción; FP (*Fluorescence Polarization*), con ensayos de tipo competitivo para estudiar la interacción ligando-proteína utilizando luz polarizada; ITC (*Isothermal Titration Calorimetry*) que permite medir las propiedades termodinámicas en disolución, entre otras.

2.3.1 Resonancia Magnética Nuclear (RMN)

Nuestro grupo de investigación se centra principalmente en el estudio estructural de ligandos glucídicos y sus complejos con proteínas. Las técnicas estructurales de alta resolución más empleadas para ello son la difracción de rayos X y la espectroscopía de RMN. En concreto, esta última es la principal herramienta experimental utilizada en carbohidratos, tanto para la elucidación estructural como para el estudio de las interacciones moleculares. En el caso de los complejos, su extraordinaria sensibilidad a las distintas escalas de afinidad y/o tiempo, al poder emplear diferentes magnitudes derivadas de RMN, la han convertido en la principal herramienta para su estudio. Entre las diferentes técnicas y métodos destacan: la determinación de afinidad tanto alta como baja; la caracterización estructural y sus interacciones; la identificación de epítomos de unión en la proteína y ligando, etc. Además, al realizarlos en disolución, los experimentos de RMN se llevan a cabo en condiciones próximas a las fisiológicas.

Una característica de la RMN es la capacidad de observación de señales individuales correspondientes a un mismo tipo de núcleo en entornos químicos diferentes. Las propiedades de las señales de RMN se pueden simplificar en desplazamiento químico, multiplicidad o acoplamiento escalar con otros núcleos y relajación, velocidad y mecanismos de extinción de señales. Gran parte de las aplicaciones de la RMN en química orgánica, biológica u organometálica hacen uso de estas propiedades.

Otra de las características de la RMN es la propiedad de transferir magnetización entre distintos núcleos por diferentes mecanismos, y poder detectar el origen y el destino de la magnetización. Estos mecanismos son tres: el primero, el acoplamiento escalar que se produce a través de los electrones de enlace y puede ser tanto homonuclear como heteronuclear; el acoplamiento dipolar, a través del espacio; y el intercambio químico en el que es el propio núcleo el que se transfiere de un entorno a otro. Basados en estas potenciales características se pueden realizar experimentos multidimensionales para estudiar y detectar la magnetización transferida entre núcleos.^[66]

Los experimentos monodimensionales ¹H-RMN contienen información del número de protones que generan cada señal y con cuantos átomos vecinos se acoplan generando la multiplicidad de cada señal. El resto de la información tradicionalmente se ha obtenido mediante complejos procesos de asignación basados en el conocimiento previo del desplazamiento químico de cada agrupación de átomos. Sin embargo, con los experimentos bidimensionales es posible correlacionar los átomos de las moléculas mediante sus acoplamientos escalares (a través de los electrones de los enlaces) COSY, TOCSY y HSQC, o a través del espacio, NOESY y ROESY.^[67]

Un experimento multidimensional de RMN consta de varias etapas separadas convenientemente por pulsos de radiofrecuencia: preparación, evolución o marcaje de frecuencias después de la excitación de los núcleos durante un tiempo variable t_1 , mezcla de magnetización en un intervalo de tiempo τ y detección convencional de la FID (Free Induction Decay)^[68] en un tiempo t_2 (Figura 2.6):

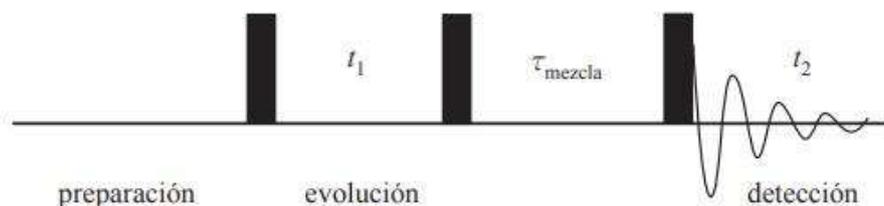


Figura 2.6. Esquema general de una secuencia de pulsos para un experimento de RMN-2D. Éste se repite un determinado número de veces con distintos incrementos de t_1 generando una señal bidimensional (t_1 , t_2).^[68]

Por cada tiempo variable t_1 se registra una fid t_2 que se almacena para su procesamiento independiente. La información que se obtiene de los experimentos COSY (*CORrelation Spectroscopy*) se refiere a los protones que están directamente acoplados entre sí; de los TOCSY (*TOTAL Correlation Spectroscopy*) se obtiene información de aquellos protones que constituyen un sistema de spin, es decir, todos aquellos protones acoplados a través de otros;^[69] los NOESY (*Nuclear Overhauser Enhancement Spectroscopy*) y análogos, relacionan a aquellos protones con proximidad espacial. La combinación de ambos tipos de experimentos permite determinar la secuencia primaria de los oligosacáridos y, adecuadamente cuantificados, los experimentos de acoplamiento dipolar permiten establecer distancias interprotónicas. Otro tipo de experimento es el HSQC, (*¹H Detected, Single Quantum Correlation*) que correlaciona ^1H con el heteroátomo al que está unido ^{13}C o ^{15}N . En el contexto de esta tesis, se han usado los HSQC registrados durante la síntesis de los oligosacáridos estudiados, además del experimento marcado con ^{15}N para estudiar la interacción proteína-ligando.

2.3.1.1 Experimento COSY y TOCSY

El experimento COSY más simple consta de dos pulsos de 90° separados por un tiempo variable t_1 realizándose posteriormente la adquisición de la FID durante t_2 . Esta técnica proporciona información sobre el acoplamiento entre dos núcleos de ^1H - ^1H con una J_{HH} directa (Figura 2.7). Su aplicación más extendida es sobre el núcleo ^1H pero también puede realizarse sobre otros átomos.^{[70],[71]} El experimento TOCSY constituye una variante del experimento COSY en el que, después del marcaje de frecuencias, se ha introducido un tiempo de mezcla variable consistente en un bloqueo de espines o spin lock. Durante este periodo el sistema opera bajo acoplamiento fuerte transmitiendo la magnetización a lo largo del sistema de espines en función de la duración del bloqueo y la magnitud de la constante de acoplamiento. El periodo del bloqueo de espines es una secuencia multipulso que se optimiza en función de la aplicación y el campo, siendo los más empleados el MLEV-32 o 64 y el dipsi2.^{[69],[72]}

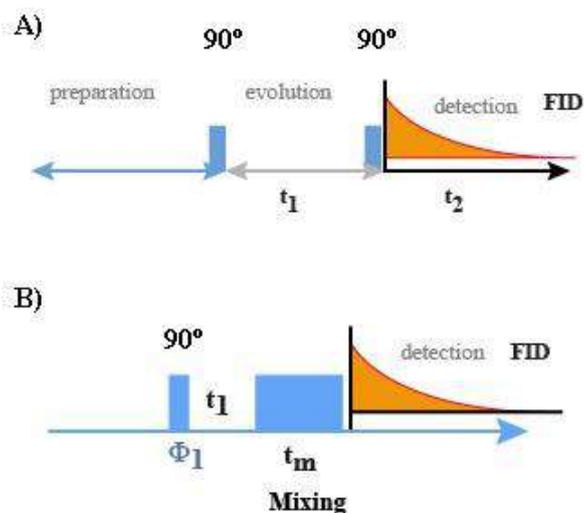


Figura 2.7. A) Secuencia de pulsos genérico para un experimento COSY. B) Secuencia de pulsos para un experimento TOCSY.

2.3.1.2 Experimento NOESY

El experimento NOESY básico consta de tres pulsos de 90° . El primero de ellos crea una magnetización transversal alineada sobre uno de los ejes del sistema de coordenadas, provocando que aquellos vectores correspondientes a núcleos independientes, moviéndose a velocidades diferentes, se separen formando ángulos distintos con el plano transversal durante un tiempo variable t_1 . Este es el conocido como marcaje de frecuencias y está en el origen de la espectroscopía multidimensional. El segundo pulso que marcará el final del tiempo de evolución genera de nuevo una magnetización longitudinal diferente para cada núcleo en función de su frecuencia durante t_1 . En este punto comienza el llamado tiempo de mezcla, durante el cual los núcleos intercambian magnetización longitudinal por relajación dipolar a través del espacio, que finaliza con el tercer pulso de 90° , que convierte la magnetización longitudinal en transversal observable (Figura 2.8).^{[73],[74],[75]}

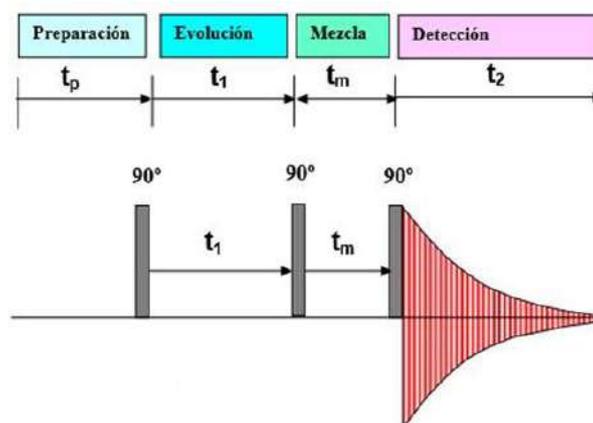


Figura 2.8. Secuencia de pulsos para un experimento de NOESY.

El Efecto Nuclear Overhauser (NOE) se define como el cambio que experimenta un núcleo (S) en la intensidad de su señal de RMN al perturbar otro (I) próximo en el espacio con el que esté acoplado mediante interacción dipolo - dipolo. Si esta intensidad aumenta con la perturbación entonces el NOE se define como positivo ($I > I_0$), y en cambio será negativo cuando ésta disminuya ($I < I_0$) (Ecuación 2.8).^[73]

$$\eta_I\{S\} = \frac{I - I_0}{I_0} \quad (2.8)$$

Para facilitar la comprensión del fundamento del NOE, se expone un diagrama de energía para un sistema de dos espines homonucleares, I y S, acoplados dipolarmente en una molécula rígida que gira isotrópicamente en una disolución (Figura 2.9):

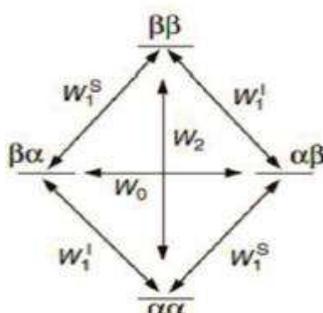


Figura 2.9. Niveles de energía para un sistema de dos espines acoplados. Se distinguen 6 tipos de transiciones: cuatro de estas corresponden a las de cuanto simple (W_1) que involucran saltos de un único spin y corresponden a las transiciones que dan lugar a las líneas del espectro, por ejemplo $\alpha\alpha \rightarrow \alpha\beta$; y dos transiciones flip-flop $\alpha\beta \rightarrow \beta\alpha$ y $\alpha\alpha \rightarrow \beta\beta$ que involucran ambos espines S e I siendo de cero cuanto si $\Delta M=0$ (W_0) o doble cuanto (W_2) si $\Delta M=2$, ambos con mecanismos de relajación cruzada.^[76]

Como se puede observar en la imagen de arriba, el NOE resulta del equilibrio entre distintas vías de relajación en competencia entre sí. Por un lado, se encuentran las vías de relajación cruzada W_0 y W_2 que son responsables del signo del NOE. El proceso de relajación W_2 (de menor energía), que produce NOEs homonucleares positivos es predominante en ligandos de bajo peso molecular que se mueven rápidamente en disolución y están caracterizados por tiempos de correlación cortos. En el caso opuesto, moléculas de peso molecular grande, con tiempos de correlación largos, dan lugar a NOEs negativos siendo predominante el mecanismo de relajación W_0 . Cada vía de relajación cruzada se diferencia en su cinética y extensión. La relajación de doble (W_2) cuanto crece más rápidamente y puede llegar a ser el doble de la de cero cuanto (W_0). Además de estos mecanismos de relajación cruzada responsables del crecimiento del NOE, reestableciendo las diferencias de población de equilibrio del núcleo no saturado (I) se encuentran los mecanismos de relajación cuántica simple (W_1), que son los que originan las señales de los protones involucrados (las observables).

De estas consideraciones y de la Figura 2.9 se puede deducir la expresión de la ecuación del NOE del estado estacionario (o ecuación de Solomon 2.9).^[76]

$$\eta_I\{S\} = \frac{\gamma_S}{\gamma_I} \left[\frac{W_2 - W_0}{W_0 + 2W_1 + W_2} \right] = \frac{\gamma_S}{\gamma_I} \left[\frac{\sigma_{IS}}{\rho_{IS}} \right] \quad (2.9)$$

Las probabilidades de las transiciones dependen a su vez de la función de densidad espectral, $J(\omega)$, que corresponde a la transformada de Fourier de la función de correlación, cuya transición va a depender por lo tanto del tamaño molecular. Así, las moléculas pequeñas que, con tiempos de correlación cortos, relajan muy lentamente pudiendo muestrear una amplia gama de frecuencias, presentan una función de densidad espectral amplia, y viceversa (Figura 2.10).

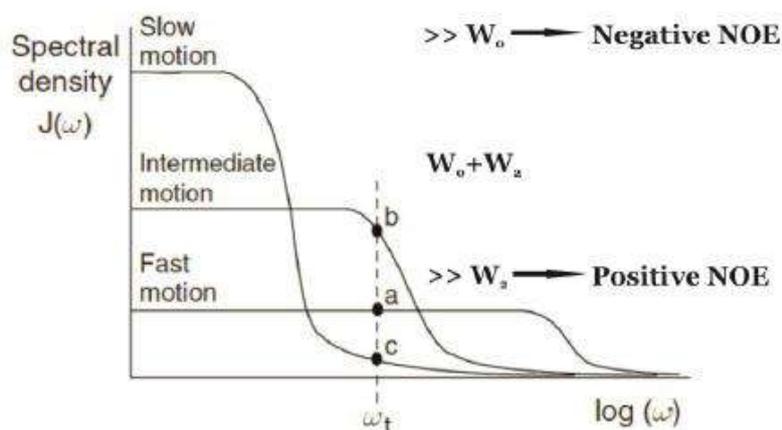


Figura 2.10. Evolución de la función de densidad espectral ($J(\omega)$) con respecto a una función de frecuencia (ω), siendo ω_t la frecuencia de transición de spin en escala logarítmica.^[76]

➤ Determinación de la distancia interprotónica

Para obtener los datos cinéticos, se recurre al NOESY (NOE transitorio) en el que se invierte el spin en lugar de saturarlo. El resultado es el mismo salvo que la intensidad es diferente (la mitad) y que es posible obtener cinéticas de crecimiento en función del tiempo de mezcla. La cinética del crecimiento del NOE en experimentos NOESY se muestra en la Figura 2.11.

El resultado es la acumulación de una determinada cantidad de NOE durante un período de tiempo que acaba disipándose a medida que comienza a operar la relajación y el NOE se difunde a través de otros contactos con otros núcleos.

Su versión monodimensional, el 1D-NOESY, es muy útil para la elucidación estructural. Se puede asumir que la velocidad inicial de crecimiento del NOE entre dos protones se debe sólo a la velocidad de relajación cruzada entre ambos, σ_{NOE} . Esta se puede calcular a partir de la curva de crecimiento del pico de cruce dividido por el pico diagonal al mismo tiempo de mezcla. La velocidad de relajación cruzada es proporcional a la función espectral y a r^{-6} (ver Figura 2.11):

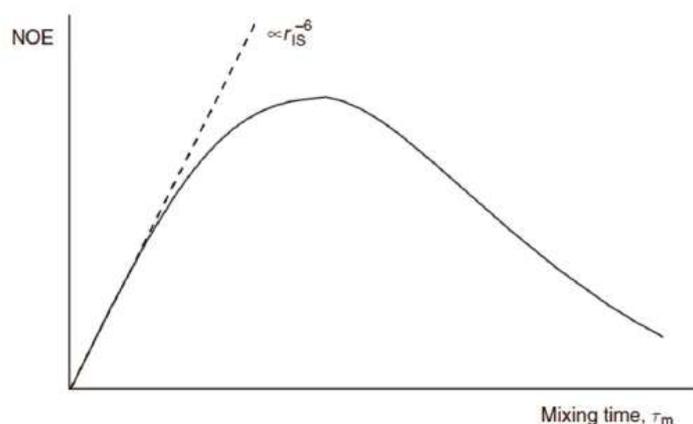


Figura 2.11. Evolución de la curva de crecimiento de un experimento de NOE transitorio.^[76]

Por lo tanto, en esta aproximación el valor del NOE depende linealmente del tiempo de mezcla (τ_m) hasta que, la relajación de espín no perturbado predomina sobre la relajación cruzada (σ_{IS}) en tiempos de mezcla largos, y se pierde esta linealidad. De acuerdo con esto, la intensidad del NOE (η_{IS}) entre I y S (siendo éste el spin invertido) en un tiempo de mezcla τ_m puede ser calculado mediante la siguiente ecuación 3.0:

$$\eta_{IS} = \sigma_{IS} \tau_m \quad (3.0)$$

Además, la velocidad de relajación cruzada es proporcional a la distancia internuclear (r_{IS}^{-6}) según la ecuación 3.1.

$$\sigma_{IS} = k r_{IS}^{-6} \quad (3.1)$$

siendo,

$$k = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2 \gamma^4}{10} \left(\frac{6\tau_c}{1 + 4\omega^2 \tau_c^2} - \tau_c \right) \quad (3.2)$$

La constante de proporcionalidad k contiene el tiempo de correlación τ_c y las constantes físicas γ , μ_0 y ω son valores conocidos para un determinado experimento. Por lo tanto, si se conoce una distancia nuclear (r_{I2S})⁻⁶ entre los núcleos I y S se puede establecer un NOE de referencia, calculándose por comparación directa la distancia internuclear desconocida.^[77]

$$\frac{\eta_{I1S}}{\eta_{I2S}} = \frac{r_{I1S}^{-6}}{r_{I2S}^{-6}} \quad (3.3)$$

2.3.1.3 Experimentos de ligando observado o de transferencia de magnetización

La RMN es particularmente sensible a procesos dinámicos como son la formación y disociación de complejos entre diferentes moléculas. Ello ha llevado al desarrollo de una serie de experimentos para el estudio de complejos entre moléculas grandes y pequeñas,

denominados de ligando observado. Si la cinética de asociación-disociación del complejo es suficientemente rápida, es posible, en una situación en la que la proporción ligando-receptor está desplazada hacia el ligando, observar ciertas propiedades del ligando enlazado en los espectros de RMN. Estos experimentos tienen la gran ventaja de simplificar el proceso de asignación de las señales del complejo, porque en este caso aparecen donde las del ligando libre ya que dependen de las proporciones de poblaciones en el equilibrio. Alternativamente, también es posible hacer uso de la alta capacidad de relajación de las macromoléculas para transferir magnetización entre la macromolécula y el ligando (STD-NMR). También, hay otras técnicas que emplean el disolvente como sonda para la exploración del complejo (waterLOGSY).

➤ ***Experimento de NOESY transferido (tr-NOESY)***

En el caso de estudio de interacciones biomoleculares podemos considerar múltiples especies y diferentes situaciones. Puede darse un equilibrio rápido entre el ligando libre y el asociado, o que el equilibrio esté desplazado hacia un complejo cinéticamente estable. En este último caso, el análisis más frecuente es el que se efectúa sobre la variación de las señales de la proteína durante una valoración con el ligando, obteniendo datos de las zonas de unión. En esta situación el análisis se suele complementar con el estudio paralelo de la estructura tridimensional de la biomolécula receptora. En los experimentos transitorios es posible estudiar la estructura del ligando enlazado a través del análisis de las señales promedio empleando una cantidad subestequiométrica de receptor. Ello se basa en el diferente comportamiento del relajamiento dipolar de las moléculas pequeñas y las macromoléculas. En condiciones adecuadas, la relajación dipolar del ligando enlazado predomina sobre la del libre, aunque esté en una proporción mucho menor. En una situación óptima los efectos NOE del ligando enlazado predominan sobre los del libre. Además, si se comporta como una molécula pequeña los NOE estarán en diferentes regímenes apareciendo con signos opuestos. Recientemente, ha resurgido la técnica de transferencia de saturación haciendo uso de las mejoras en el equipamiento moderno, que rápidamente se ha incorporado a las herramientas para el estudio de complejos biomoleculares.

Aunque bien conocido, el NOESY transferido ha experimentado una explosión de aplicaciones en los últimos años. En especial por su combinación con otros experimentos de ligando observado como el STD, o el waterLOGSY en el estudio de complejos entre macromoléculas y ligandos sencillos. Considerando una molécula pequeña que se une a una macromolécula y que su intercambio entre la forma libre y ligada es rápido en la escala de tiempos del desplazamiento químico, la velocidad de relajación cruzada observada en las señales promedio será próxima a la del complejo, mientras que su desplazamiento químico estará muy próximo al de las señales del ligando libre.^[78] Por consiguiente, un NOESY adquirido en estas condiciones registrará mayoritariamente, si no exclusivamente, la especie unida y, por lo tanto, será negativo, de evolución rápida y mayor intensidad. Si el ligando es una molécula pequeña o mediana de NOE positivo, ésta es una prueba de la formación del complejo entre ligando y receptor. Si hubiera un cambio conformacional debido a la formación del complejo, éste se verá reflejado en el

patrón de NOE.^{[79],[80]} Los NOE transferido intermoleculares entre ligando y proteína se pueden usar para establecer la orientación del ligando unido en el sitio activo del receptor.^[81] En el experimento NOESY transferido debe cumplirse la siguiente desigualdad:

$$|N_{\text{enlazado}} \cdot \sigma_{\text{enlazado}}| \gg |N_{\text{libre}} \cdot \sigma_{\text{libre}}| \quad (3.4)$$

Donde σ es la velocidad de relajación cruzada y N el número de moléculas en los estados libre y unido. Como se ha mencionado antes, la intensidad de los NOE transferidos depende principalmente del exceso del ligando sobre la proteína ($[L]_T/[P]_T$), y por consiguiente, de las fracciones en sus formas libre o asociada. Por lo tanto, si esta relación es demasiado alta, existirá un exceso de ligando disociado en disolución lo que se traduce en NOEs positivos que pueden disminuir o incluso cancelar las señales que indican formación de complejo. A pesar de todas las ventajas y facilidades que nos ofrece esta técnica, también hay que tener en cuenta ciertos inconvenientes. Uno de los principales es la posible difusión de spin, típica en macromoléculas, que genera picos cruzados negativos entre protones que se encuentran conectados a través de un tercer spin, conduciendo a errores en la interpretación de los espectros. No obstante, una adecuada cuantificación de la curva de crecimiento del NOE es suficiente para evitar este artefacto.

2.3.1.4 Experimento de Diferencia de transferencia de saturación (STD)

Otro de los experimentos basados en la observación del ligando es el experimento de diferencia de transferencia de saturación que permite estudiar la interacción en disolución entre una macromolécula y un ligando, mapeando los protones que están involucrados en esta unión. Está basado en la difusión de espín, siendo éste un fenómeno característico en el NOE de moléculas de tiempo de correlación largo, que se encuentran en la zona correspondiente a moléculas grandes de la curva del NOE en función de la movilidad. Dada la eficacia de la transferencia de magnetización de esta zona, una vez transferida al espín receptor se puede continuar transfiriendo a sucesivos núcleos, generando así sucesivos acoplamientos dipolares activos.

A partir de un exceso de molécula pequeña con respecto al receptor (desde 1:10 hasta 1:1000), se llevan a cabo dos experimentos monodimensionales ^1H -RMN: el primero de ellos se utiliza como espectro de referencia; realizado en condiciones de equilibrio térmico donde se establece una frecuencia de irradiación lejos de la zona de interés (“*off-resonance*”). En el segundo, se aplica una radiofrecuencia de baja potencia y larga duración sobre una zona donde sólo existen señales de la proteína, saturándolas y transfiriendo dicha magnetización al ligando mediante difusión de espín a través de NOEs intermoleculares (experimento “*on resonance*”).^[81] La diferencia entre ambos espectros refleja la cantidad de magnetización recibida desde el receptor por el ligando, siendo proporcional a la distancia entre los núcleos de ambos y revelando exclusivamente las señales del ligando que tienen interacción, cuya intensidad será mayor cuánto más cerca se encuentren (ver Figura 2.12).^[83]

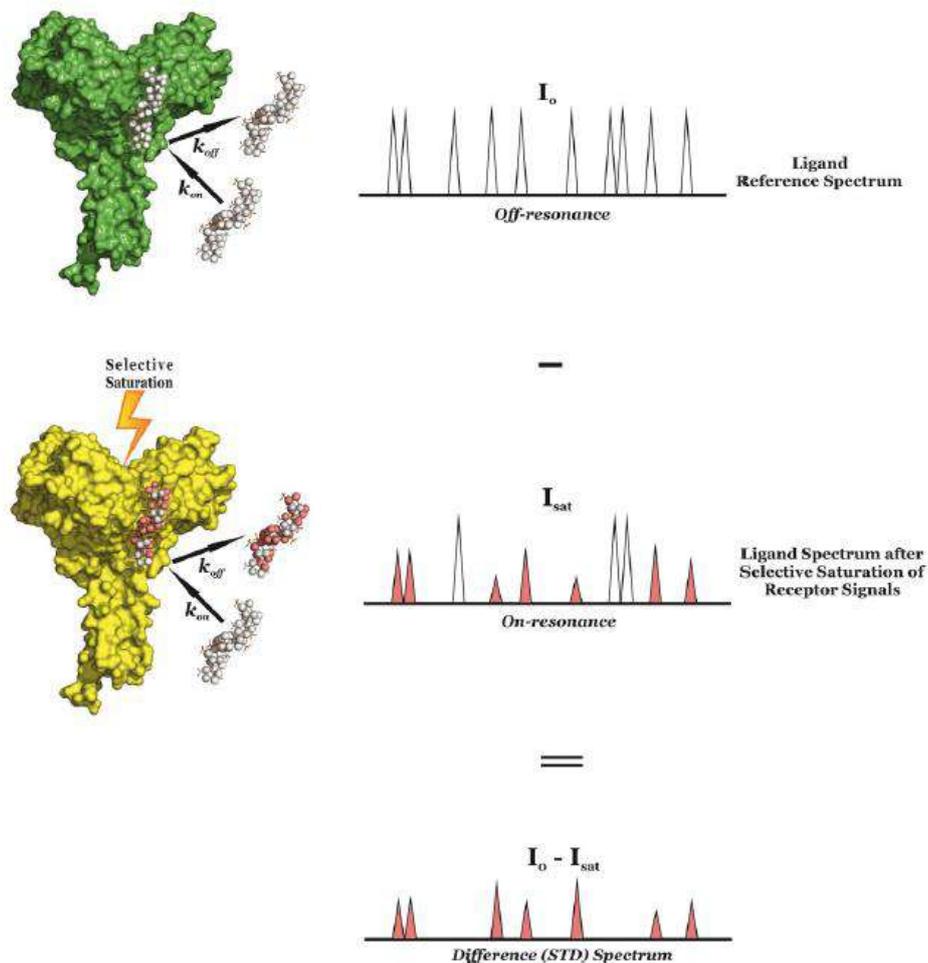


Figura 2.12. Ilustración del experimento de STD.

La distribución de la saturación entre los protones más cercanos y la intensidad de la señal muestra, de una forma cualitativa, la interacción que tiene lugar entre el ligando y la proteína. Pero este resultado puede analizarse de manera semicuantitativa comparando porcentajes relativos de saturación, normalizando todas las intensidades de STD medidas. Para ello, se establece una escala relativa de saturaciones de 0% a 100%, asignando el porcentaje más alto a la señal de mayor transferencia de saturación. Esto genera mapas de epítipo con los que es posible analizar las zonas de mayor interacción del ligando con la proteína.

Los valores de STD están influidos por la relajación pudiendo no ser uniformes en todos los núcleos. Para cancelar las posibles influencias de la relajación en el valor del STD se suele recurrir a calcular la velocidad inicial donde la relajación no ha operado todavía. Ésta se obtiene por el ajuste de la curva de crecimiento del STD a una ecuación exponencial, empleando tiempos de saturación crecientes y obteniendo la velocidad inicial de crecimiento del STD, STD_0 .^[81]

Tal y como se muestra en el esquema de la Figura 2.13, el experimento STD consta de tres bloques principales. En el primero de ellos, el bloque A, se aplica un tren de pulsos selectivos durante un tiempo de saturación (t_{sat}): para el experimento de referencia (*experimento off-resonance*) se irradia muy lejos de las señales de proteína y ligando

mientras que en el experimento *on-resonance* se irradia directamente sobre la proteína sin afectar al ligando. En el bloque B, tras un fuerte pulso de 90° se aplica un “spin lock” para eliminar las señales de fondo de la proteína. Por último, para muestras que contienen una cantidad significativa de agua, se puede añadir un módulo de supresión de agua, watergate, o excitation sculpting que tiene lugar en el bloque C.^[81]

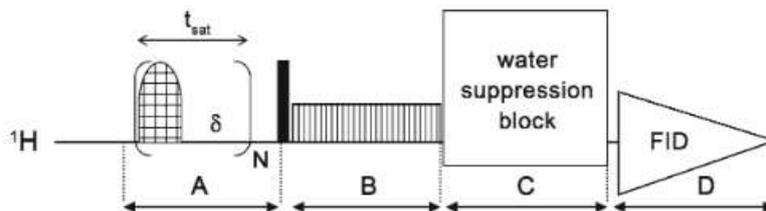


Figura 2.13. Esquema de pulsos para el experimento STD.^[81]

2.3.1.5 Experimentos heteronucleares

➤ Experimento de HSQC

Como se ha explicado anteriormente, la mayoría de los experimentos de observación de proteína para caracterizar su unión con un ligando, implican espectros bidimensionales (2D) de RMN debido a que en los espectros 1D no se pueden identificar señales individuales de ésta. Para solventar esta limitación, es muy útil la espectroscopía heteronuclear 2D, donde se estudian otros núcleos magnéticamente activos como ^{13}C o ^{15}N presentes en estas macromoléculas. Debido a que la cantidad presente en la naturaleza de estos átomos es baja, se suelen recurrir a marcajes isotópicos mediante expresiones en bacterias o levaduras. El experimento heteronuclear 2D más conocido es el denominado HSQC (*Heteronuclear Single-Quantum Correlation*), cuya secuencia de pulsos se muestra en la Figura 2.14.

La formulación de la espectroscopía heteronuclear inversa supuso la proliferación de experimentos de correlación ^1H -heteronúcleo, y la consiguiente extensión de sus aplicaciones en química y bioquímica. Esencialmente, para ^{13}C y ^{31}P en abundancia natural y para ^{15}N en muestras con marcaje isotópico, son experimentos rutinarios en RMN formando parte de secuencias más complejas. Su versión más sencilla consiste en un módulo INEPT, que transfiere la magnetización de ^1H al heteronúcleo, donde tras el marcaje de frecuencias durante t_1 se revierte a ^1H mediante un módulo retroINEPT donde se detecta con o sin desacoplamiento (ver Figura 2.14).^{[84],[85]}

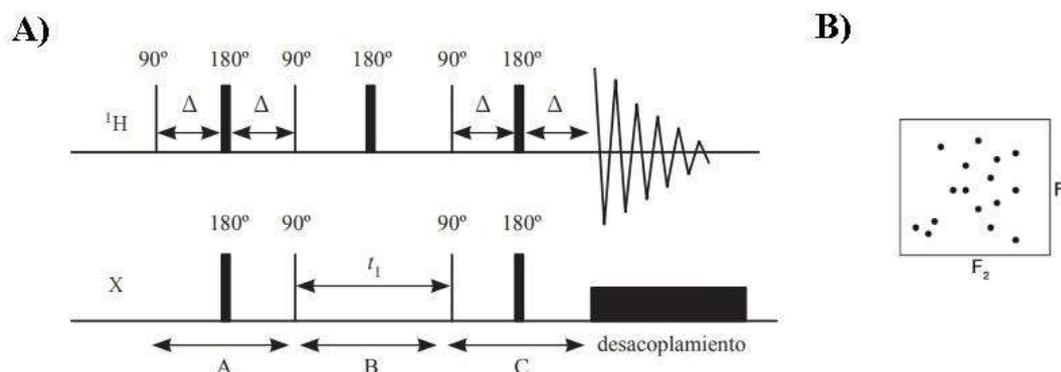


Figura 2.14. A) Secuencia de pulsos para el experimento HSQC entre ^1H y un heteroátomo.^[85] B) Representación del espectro de la correlación entre los desplazamientos químicos de ambos núcleos acoplados directamente a través de una constante de acoplamiento J .^[81]

Esta secuencia básica ha sufrido múltiples mejoras de las cuales destaca, por sus ventajas, la implementación de gradientes de campo para la selección de magnetización. Está basada en la relación de las constantes giromagnéticas de los núcleos permitiendo una extraordinaria reducción del ciclo de fases y una mejor optimización del receptor. Otra mejora esencial ha sido la optimización de la secuencia para obtener una sensibilidad mejorada incluyendo detección eco-antieco con un segundo retroINEPT. Actualmente, en aplicaciones para química se emplea con frecuencia secuencias híbridas HSQC – DEPT, denominadas HSQC editado, en las que la fase de las señales de correlación depende de la multiplicidad del carbono.^{[84],[85]}

➤ ^{15}N -HSQC

Debido a la baja abundancia y escasa sensibilidad en RMN del isótopo ^{15}N , no es posible medir ^{15}N en abundancia natural de forma directa. Sin embargo, gracias a la expresión en sistemas biológicos, ha sido posible y accesible la incorporación de ^{15}N a proteínas y ácidos nucleicos. Los experimentos ^{15}N -HSQC han encontrado una gran cantidad de aplicaciones debido a su posición estratégica en el esqueleto de las proteínas y la sensibilidad de su desplazamiento químico a determinados aspectos estructurales (estructura secundaria, enlaces de hidrógeno, movimientos locales y otros). La RMN basada en ^{15}N es una técnica que nos puede aportar información muy relevante acerca de la secuenciación de péptidos, estudiando su desplazamiento químico. Este experimento implica observar las perturbaciones de los espectros de la proteína tras la adición de cantidades crecientes del ligando, en donde las señales que se vean afectadas corresponderán a núcleos que cambiarán su entorno químico tras unirse a él. Por tanto, estas señales pueden estar ubicadas en el sitio de interacción del ligando o próximo a él, o en cambio están involucradas en un reordenamiento estructural inducido por éste^[86]. Para entender la interacción proteína-ligando y las perturbaciones que tienen lugar, tenemos que considerar el siguiente equilibrio caracterizado por el cociente de la tasa de activación, k_{on} y la de desactivación k_{off} , dando lugar a la K_D (ecuación 3.5):



Durante el experimento, la aparición de los espectros depende de la diferencia de desplazamiento químico (^1H y/o ^{15}N), $\Delta\omega$, entre la proteína libre y unida al ligando en comparación con la tasa de cambio, k_{ex} que viene dada por la ecuación 3.6:

$$k_{\text{ex}} = k_{\text{on}}[\text{L}]_{\text{free}} + k_{\text{off}} \quad (3.6)$$

donde $[\text{L}]_{\text{free}}$ es la concentración de ligando no unido. La formación del complejo puede establecerse en un régimen de intercambio rápido ($k_{\text{ex}} \gg \Delta\omega$), donde la señal de la proteína libre se mueve hacia la señal de la proteína saturada de ligando. Por el contrario, si el intercambio es lento ($k_{\text{ex}} \ll \Delta\omega$), solo las intensidades relativas de las señales de los picos de proteína libre y unida cambiarán, por lo tanto no se observarán perturbaciones de desplazamiento químico directamente, sino solo modulaciones de las intensidades de la señal de RMN^[87]. La perturbación del desplazamiento químico ($\Delta\delta$) observado en el experimento viene dada por la siguiente ecuación 3.7:^{[86],[88]}

$$\Delta\delta = \delta_{\text{obs}} - \delta_{\text{free}} = \frac{(\delta_{\text{bound}} - \delta_{\text{free}})}{2[P]} \left[[P] + [L] + KD - \sqrt{([P] + [L] + KD)^2 - 4[P][L]} \right] \quad (3.7)$$

En la Figura 2.15 se muestra un ejemplo de esta técnica comparando los dos tipos de intercambio que tienen lugar mostrando las diferencias que se aprecian en las señales de los picos:

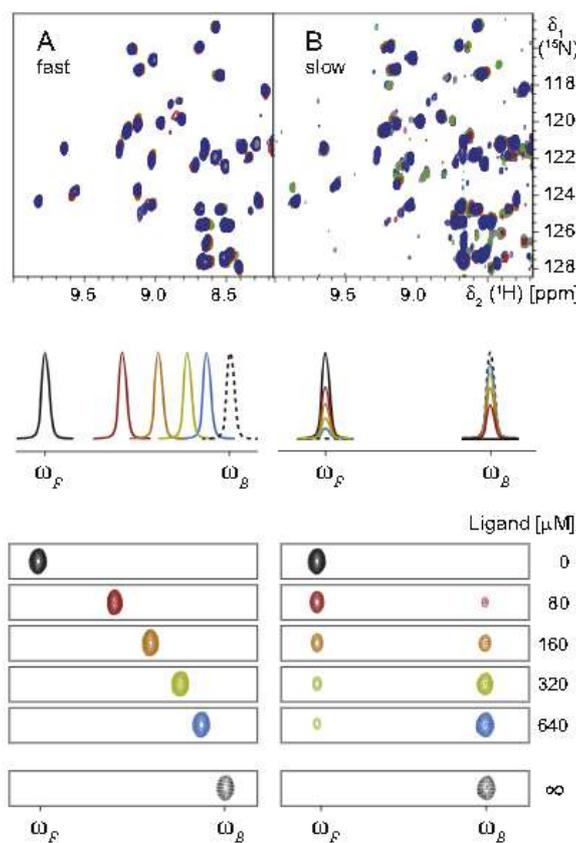


Figura 2.15. Ejemplo de cambios de desplazamiento químico tras el marcaje con ^{15}N y adición creciente de ligando para un intercambio rápido y lento.^[89]

El primero paso del experimento sería la asignación de las resonancias de los diferentes núcleos; en el caso de proteínas mayores a 15 kDa se deben asignar también ^{15}N y/o ^{13}C ; mientras que en las proteínas pequeñas se sigue la estrategia desarrollada por Wüthrich.^[90] Primeramente, se identifican los sistemas de espín de cada aminoácido a partir de sus acoplamientos escalares, mediante técnicas de 2D COSY y TOCSY, o 3D ^1H - ^{15}N TOCSY-HSQC. La identificación a tres enlaces entre los pares de protones NH- $\text{H}\alpha$, $\text{H}\alpha$ - $\text{H}\beta$, $\text{H}\beta$ - $\text{H}\gamma$...de un mismo aminoácido se realiza con el experimento COSY, mientras que con el TOCSY se obtiene las señales de correlación de todas las frecuencias de aquellos protones pertenecientes al mismo sistema de spin.

En cambio, en el experimento 3D ^1H - ^{15}N TOCSY-HSQC están concentradas todas las correlaciones TOCSY de los sistemas de espín que contienen los valores de δ del N – amídico, con un solapamiento menor que en 2D. De esta forma se pueden crear patrones característicos por los desplazamientos químicos conocidos que pertenecen a diferentes sistemas de espines, pudiendo agrupar los 20 aminoácidos en base a este criterio. El siguiente paso consiste en la identificación de cada sistema de espín con un residuo concreto de la secuencia de la proteína mediante los experimentos 2D NOESY y 3D ^1H - ^{15}N NOESY-HSQC (ver Figura 2.16).^{[91],[92]}

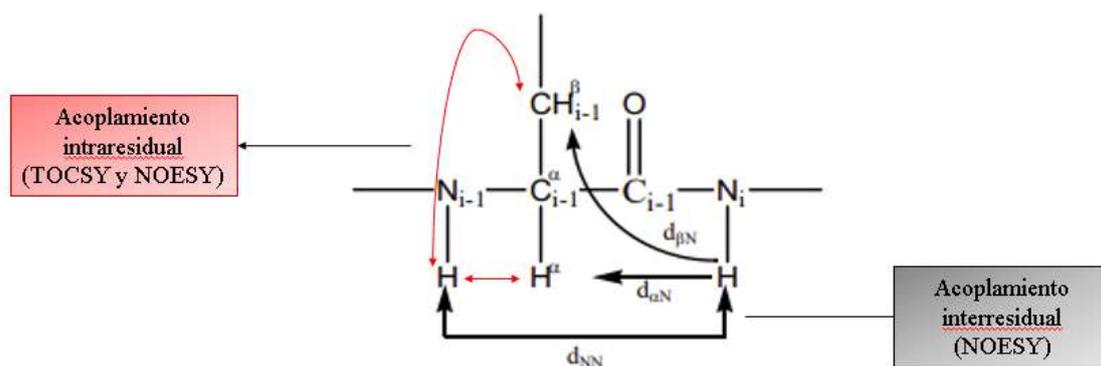


Figura 2.16. Conexiones secuenciales que se pueden encontrar en un fragmento de cadena peptídica.

Estos desplazamientos en los aminoácidos dependen de los residuos vecinos y, además, el acoplamiento $^3J_{\text{NH}}$ a tres enlaces del protón α puede aportar información sobre el ángulo de torsión del enlace peptídico. Dentro de sus limitaciones se encuentra la baja sensibilidad, por su escasa abundancia isotópica (0-36%), y la posible cancelación de algunas de sus señales debido al efecto NOE negativo.^[93] Pero, a pesar de esto, es un experimento clave para el estudio de interacciones proteína-ligando en proteínas marcadas con ^{15}N , observándose al menos una señal por aminoácido. El experimento y la secuencia de pulsos es igual a la que se ha detallado en el apartado anterior, con una primera parte donde mediante el método INEPT se transfiere la magnetización de los protones de amida al núcleo ^{15}N adyacente, una segunda etapa donde tiene lugar la generación de la dimensión F1 durante el incremento de t_1 y el último bloque que contiene un INEPT inverso donde se retransfiere la magnetización de nuevo al ^1H . Finalmente, se detecta la señal en ^1H durante t_2 con desacoplamiento simultáneo.^{[81],[94]}

La mayor parte de las versiones diferentes de ^{15}N HSQC incluyen esquemas de supresión de disolvente. Al ser los protones amida intercambiables con el agua no es posible registrar los experimentos en D_2O , por lo que es necesario usar H_2O y esquemas de supresión de disolvente. En general, los más usados y eficaces son los que se introducen al final de la secuencia antes de la detección, al minimizar el intercambio previo a la detección.

2.3.2 Polarización de Fluorescencia (FP)

La polarización de fluorescencia es una técnica muy versátil, potente y sensible, ampliamente utilizada en estudios sobre el movimiento de moléculas fluorescentes en disolución, en inmunoensayos para proteínas, monitorización terapéutica de fármacos y ensayos de compuestos tóxicos para el medio ambiente. Descrita por primera vez en 1926,^[95] la aplicación se ha ido desarrollando y, a día de hoy, es una técnica muy extendida en el área del descubrimiento de fármacos, centrada en el estudio de interacciones ADN-proteína, ADN-ADN y proteína-proteína. Esta técnica se basa en la irradiación de la muestra con un haz de luz monocromática, que se encuentra polarizada tras atravesar un filtro de polarización vertical. Aquellas moléculas fluorescentes, orientadas adecuadamente sobre este plano de luz absorbida, serán excitadas y la luz emitida, en los planos horizontal y vertical, se recogerá en un detector tras atravesar otro filtro polarizado (Figura 2.17).^{[96][97]} La fórmula matemática que relaciona las intensidades de luz emitida en los planos horizontal y vertical^[98] nos permite obtener los correspondientes valores de polarización de fluorescencia (FP), que indican el grado de polarización de la luz emitida.

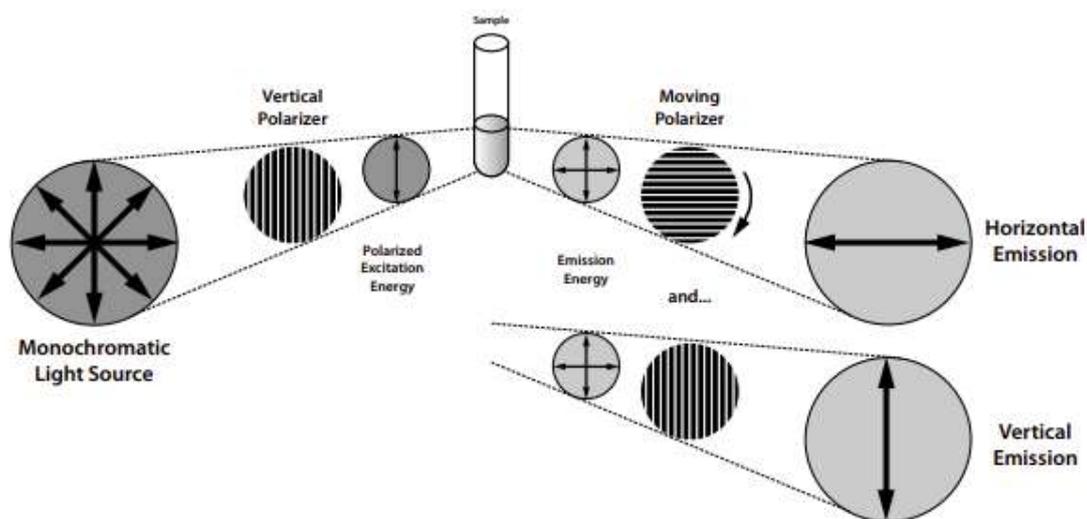


Figura 2.17. Representación esquemática de la detección de la polarización de fluorescencia.^[96]

El valor de polarización de una molécula es proporcional al tiempo de relajación rotacional de la misma, o lo que es lo mismo, el tiempo que tarda en girar un ángulo de un radián. Este valor depende de la viscosidad, la temperatura absoluta y el volumen molecular.^[96] Por lo tanto, si los dos primeros parámetros se mantienen constantes, la polarización está relacionada directamente con el tamaño molecular. Cuando un ligando

fluorescente es pequeño (<1500 Da), al rotar rápidamente en disolución, la luz polarizada con la que se excita se despolarizará. En cambio, cuando esta molécula pequeña se une a otra más grande (>10 kDa), como por ejemplo proteínas, el movimiento de rotación del fluoróforo se vuelve más lento debido a la restricción en su movimiento, generando de este modo luz polarizada (ver Figura 2.18).^[98]

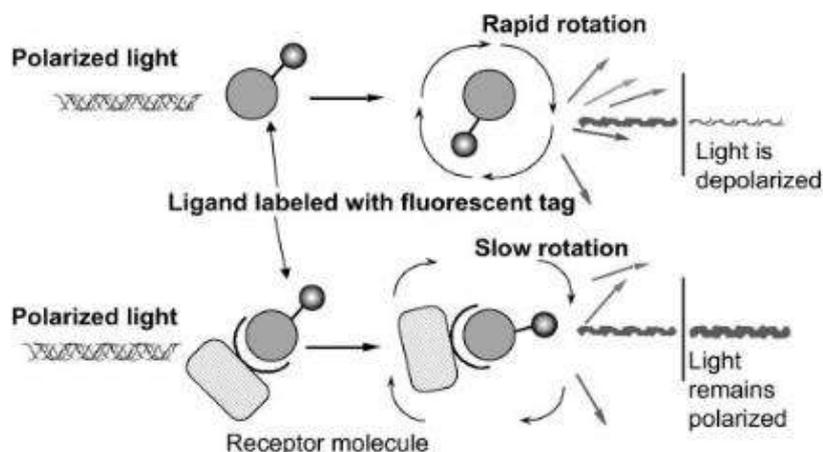


Figura 2.18. Esquema representativo de la polarización de la luz emitida en función del tamaño molecular.^[98]

De este modo, las medidas de FP son ideales para estudiar las interacciones entre una molécula fluorescente de bajo peso molecular, como un oligosacárido, y una proteína. Además, se pueden desarrollar experimentos de competición (ver más abajo) en los que se analiza la capacidad de un ligando no fluorescente para desplazar a la sonda fluorescente del complejo con la proteína. Este tipo de experimentos son los que se han empleado en esta Tesis para calcular las afinidades relativas de nuestros oligosacáridos.

Estos ensayos de tipo competitivo presentan una serie de ventajas: se llevan a cabo en disolución, usando una etiqueta fluorescente única que evita el uso de radioisótopos; no requieren ningún paso de filtración o separación de ligando unido y libre; y se puede trabajar a volúmenes bajos (del orden de 10 μ L) haciéndolo muy útil en cribados de alto rendimiento.^[99] Estos experimentos se realizan utilizando microplacas de titulación que contienen diferentes números de pocillos, desde 96 hasta 1536. Además, existen diferentes tipos de compuestos que pueden usarse como etiquetas fluorescentes, siempre que sus máximos de absorción y emisión coincidan con las longitudes de onda de los filtros polarizados empleados en la técnica.

Para elegir el fluoróforo o marcador fluorescente, algunos parámetros clave son el rendimiento cuántico y el tiempo de vida útil del estado excitado por fluorescencia. Así, los fluoróforos con un tiempo de vida de 4 ns, como la fluoresceína y rodaminas, proporcionan una buena medición de FP entre ligandos pequeños y proteínas de pesos moleculares mayores a 10 kDa. La fluoresceína ha sido la etiqueta fluorescente más comúnmente utilizada para este tipo de experimentos de competición, y en casi todos los lectores de placas se instalan filtros adecuados para sus longitudes de onda de excitación y emisión.^[99]

➤ **Elección de la sonda fluorescente**

En los ensayos de FP, se tienen que tener en cuenta ciertos requisitos para escoger una sonda fluorescente adecuada. En primer lugar, ésta debe presentar una alta afinidad por la proteína que se quiere estudiar. Por lo tanto, es necesario realizar una medición de esta interacción sonda fluorescente-proteína. Para ello se añaden concentraciones crecientes de proteína sobre una concentración constante de sonda y, a través de su representación gráfica, se calcula la correspondiente constante de disociación K_D .^[99] Las curvas son analizadas como isotermas de Langmuir, asumiendo que el sistema alcanza el equilibrio durante la incubación (ver Figura 2.19):

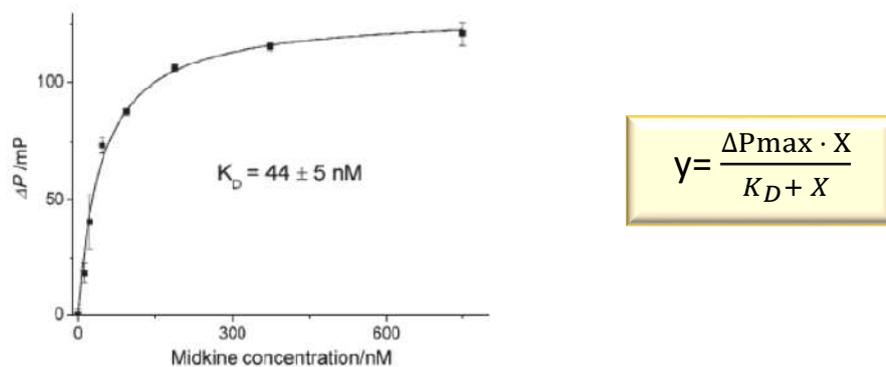
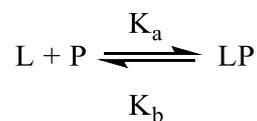


Figura 2.19. Representación gráfica de la variación de polarización (ΔP) frente a la concentración de proteína (Midkina) ajustando la curva a un modelo de interacción 1:1.^[100]

Cada punto en la curva representa la medición de FP de cada pocillo. A concentraciones muy bajas de proteína la sonda fluorescente (ligando) permanece libre mientras que, a concentraciones más altas, todo el ligando se encuentra formando parte del complejo. El valor de la constante termodinámica K_D se corresponde con la concentración de proteína en la que el 50% del ligando está libre y el otro 50% está formando parte del complejo.

La expresión para el cálculo de K_D viene dada por la cinética de una reacción bimolecular reversible entre un ligando (L) y la proteína (P) para dar lugar al complejo (LP):^[101]



Esta formación del complejo puede expresarse en términos de velocidad (ecuación 3.8):

$$\frac{d[LP]}{dt} = K_a[L][P] - K_b[LP] \quad (3.8)$$

Si se sustituye la concentración de ligando libre [L] por la diferencia entre la cantidad de ligando inicial $[L_0]$ y la que se encuentra unido formando el complejo, [LP], tenemos la ecuación 3.9:

$$\frac{d[LP]}{dt} = K_a[P]([L_0] - [LP]) - K_b[LP] \quad (3.9)$$

Por tanto, teniendo en cuenta que cuando se alcanza el equilibrio $d[LP]/dt = 0$ y $K_D = K_b/K_a$, y que la concentración del complejo es proporcional al incremento de polarización (ΔP), se establece (ecuación 4.0):

$$K_a[P](\Delta P_{\max} - \Delta P) - K_b\Delta P = 0 \quad (4.0)$$

Desde esta ecuación se ha deducido la expresión matemática representada en la Figura 2.19 para calcular la K_D del complejo ligando fluorescente-proteína.

➤ **Experimento de competición de FP**

Una vez que tenemos una adecuada sonda fluorescente, capaz de unirse a la proteína en estudio, pueden desarrollarse experimentos de competición para calcular las afinidades relativas de cualquier molécula no fluorescente. En este ensayo competitivo (Figura 2.20) se analiza la capacidad de un ligando (A) para interactuar con una proteína que previamente se ha incubado con una sonda fluorescente (B). La cantidad de complejo existente entre la sonda fluorescente y la proteína se va midiendo al registrar la señal de polarización de pocillos que contienen concentraciones fijas de ambos, pero crecientes de ligando competidor A.^[102] Si se produce el desplazamiento de la sonda por el competidor (A), se observará una disminución en el valor de FP, que corresponde ahora al de una molécula fluorescente de bajo peso molecular (ver Figura 2.20):

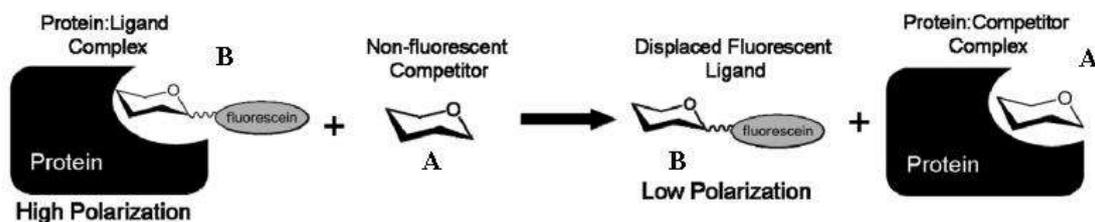


Figura 2.20. Representación esquemática para un ensayo de tipo competitivo. La intensidad de polarización se ve afectada tras el desplazamiento de la sonda fluorescente unida a una proteína por un competidor activo.^[102]

Bibliografía Capítulo 2:

- [1] A. Ciulli, *Methods Mol. Biol.*, **2015**, 1008, 357-388.
- [2] E. Fadda, R. J. Woods, *Drug Discov. Today* **2010**, 15, 596–609.
- [3] A. Iupinov in *Implementation of the Particle Mesh Ewald method on a GPU*. PhD Thesis, Stockholm University, **2016**.
- [4] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, *J. Chem. Theory Comput.* **2012**, 8, 3314–3321.
- [5] J. Chen, K. N. Houk, *J. Chem. Inf. Comput. Sci.* **1998**, 38, 939–939.
- [6] R. A. Kendall, R. J. Harrison, R. J. Littlefield, M. F. Guest in *Reviews in Computational Chemistry*, Wiley-VCH, **1995**, 6, 209–316.
- [7] G. J. Boons, Y. Lee, A. Suzuki, N. Taniguchi, A.G.J. Vorage, *Comprehensive Glycoscience (From Chemistry to Systems Biology)*, Elsevier Science, **2007**, 2–4.
- [8] R. J. Woods, R. A. Dwek, C. J. Edge, B. Fraser-Reid, *J. Phys. Chem.* **1995**, 99, 3832–3846.
- [9] B. L. Foley, M. B. Tessier, R. J. Woods, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, 2, 652–697.
- [10] S. Pérez, I. Tvaroška, in *Advances in Carbohydrate Chemistry and Biochemistry*, Elsevier, **2014**, 71, 9-136.
- [11] H. Bayraktar, E. Akal, O. Sarper, T. Varnali, *J. Mol. Struct. (THEOCHEM)*, **2004**, 683, 121–132.
- [12] H. Verli, J. A. Guimarães, *Carbohydr. Res.* **2004**, 339, 281–290.
- [13] S. W. Homans, *Biochemistry* **1990**, 29, 9110–9118.
- [14] F. A. Momany, J. L. Willett, *Carbohydr. Res.* **2000**, 326, 194–209.
- [15] H. Senderowitz, W. C. Still, *J. Org. Chem.* **1997**, 62, 1427–1438.
- [16] B. Nagarajan, N. V. Sankaranarayanan, U. R. Desai, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2019**, 9, 1–35.
- [17] H. G. Petersen, *Phys. Fluids* **1967**, 10, 3668–3679.
- [18] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, 26, 1668–1688.
- [19] K. N. Kirschner, R. J. Woods, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 10541–10545.
- [20] P. Mark, L. Nilsson, *J. Phys. Chem. A* **2001**, 105, 9954–9960.
- [21] M. Jana, S. Bandyopadhyay, *Phys. Chem. Chem. Phys.* **2012**, 14, 6628–6638.

- [22] D. A. Case, R. M. Betz, D. S. Cerutti, C. I. T.E., T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C.Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, P. A. Kollman in *Amber Reference Manual*, Univ. California, San Fr. **2016**.
- [23] Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C. in *PROTEINS: Structure, Function, and Bioinformatics*, **2006**, 65, 712-725.
- [24] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, 21, 1049–1074.
- [25] C. I. Bayly, K. M. Merz, D. M. Ferguson, W. D. Cornell, T. Fox, J. W. Caldwell, P. A. Kollman, P. Cieplak, I. R. Gould, D. C. Spellmeyer, *J. Am. Chem. Soc.* **1995**, 117, 5179–5197.
- [26] K. Kirschner, A. Yongye, S. Tschampel, *J. Comput. Chem.*, **2008**, 29, 622–655.
- [27] M. B. Tessier, M. L. DeMarco, A. B. Yongye, R. J. Woods, *Mol. Simul.* **2008**, 34, 349–364.
- [28] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, 97, 10269–10280.
- [29] W. D. Cornell, P. Cieplak, C. I. Bayly, P. A. Kollman, *J. Am. Chem. Soc.* **1993**, 115, 9620–9631.
- [30] M. Basma, S. Sundara, D. Çalgan, T. Vernali, R. J. Woods, *J. Comput. Chem.* **2001**, 22, 1125–1137.
- [31] R. J. Woods, R. Chappelle, *J. Mol. Struct. (THEOCHEM)*, **2000**, 527, 149–156.
- [32] F. Y. Dupradeau, C. Cézard, R. Lelong, É. Stanislawiak, J. Pêcher, J. C. Delepine, P. Cieplak, *Nucleic Acids Res.* **2008**, 36, 360–367.
- [33] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, 79, 926–935.
- [34] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, 7, 2284–2295.
- [35] M. O. Kim, S. E. Nichols, Y. Wang, J. A. McCammon, *J. Comput. Aided. Mol. Des.* **2013**, 27, 235–246.
- [36] S. Li, M. Hong, *J. Am. Chem. Soc.* **2011**, 133, 1534–1544.
- [37] J. A. Hadden, J. R. Perilla, *Methods Mol. Biol.* **2018**, 1762, 245–270.
- [38] J. Racine, *J. Appl. Econom.* **2006**, 21, 133–141.
- [39] U. Schmitz, N. B. Ulyanov, A. Kumar, T. L. James. *J.Mol.Biol.* **1993**, 234, 373-389.
- [40] D. A. Pearlman, *J. Biomol. NMR* **1994**, 4, 279–299.
- [41] D. A. Pearlman, P. A. Kollman, *J. Mol. Biol.* **1991**, 220, 457–479.

- [42] i) A. P. Nanzer, W. F. van Gunsteren, A. E. Torda, *J. Biomol. NMR* **1995**, *6*, 313–320. ii) W. R. P. Scott, A. E. Mark, W. F. Gunsteren, *J. Biomol. NMR*, **1998**, *12*, 501–508.
- [43] i) S. Genheden, U. Ryde, *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. ii) N. A. Baker in *Reviews in Computational Chemistry*, Wiley-VCH, **2005**, *21*, pp: 349–379.
- [44] D. S. Goodsell, A. J. Olson, *Proteins Struct. Funct. Bioinforma.* **1990**, *8*, 195–202.
- [45] M. Liu, S. Wang, *J. Comput. Aided. Mol. Des.* **1999**, *13*, 435–451.
- [46] M. J. R. Yunta, *Am. J. Model. Optim.* **2016**, *4*, 74–114.
- [47] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662.
- [48] D. Soares Da Costa, R. L. Reis, I. Pashkuleva, *Annu. Rev. Biomed. Eng.* **2017**, *19*, 1–26.
- [49] M. Kontoyianni, L. M. McClellan, G. S. Sokol, *J. Med. Chem.* **2004**, *47*, 558–565.
- [50] M. D. Cummings, R. L. DesJarlais, A. C. Gibbs, V. Mohan, E. P. Jaeger, *J. Med. Chem.* **2005**, *48*, 962–976.
- [51] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, T. E. Ferrin, *J. Mol. Biol.* **1982**, *161*, 269–288.
- [52] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.*, **2009**, *30*, 2785–2791.
- [53] G. Jones, P. Willet, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.*, **1997**, *267*, 727–748.
- [54] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, *J. Mol. Biol.* **1996**, *261*, 470–489.
- [55] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750–1759.
- [56] S. Cosconati, S. Forli, A. L. Perryman, R. Harris, D. S. Goodsell, A. J. Olson, *Expert Opin. Drug Discov.* **2010**, *5*, 597–607.
- [57] T. F. Vieira, S. F. Sousa, *Appl. Sci.* **2019**, *9*, 4538.
- [58] A. Allouche, *J. Comput. Chem.* **2012**, *32*, 174–182.
- [59] M. S. Valdés-Tresanco, M. E. Valdés-Tresanco, P. A. Valiente, E. Moreno, *Biol. Direct* **2020**, *15*, 1–12.
- [60] J. P. M. Postma, J. C. Berendsen, J. R. Haak, *Faraday Symp. Chem. Soc.*, **1982**, *17*, 55–67.
- [61] M. Aldeghi, M. J. Bodkin, S. Knapp, P. C. Biggin, *J. Chem. Inf. Model.* **2017**, *57*, 2203–2221.
- [62] C. Mulakala, V. N. Viswanadhan, *J. Mol. Graph. Model.* **2013**, *46*, 41–51.

- [63] T. Hou, J. Wang, Y. Li, W. Wang, *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
- [64] Wang, J.M, Hou, T.J, Xu, X.J. *Curr. Comput-Aided Drug Des.*, **2006**, *2*(3), 287-306.
- [65] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham, *Acc. Chem. Res.* **2000**, *33*, 889–897.
- [66] A. Ross, G. Schlotterbeck, F. Dieterle, H. Senn in *Metabonomics and Metabolomics*, Elsevier B.V., **2007**, pp:55-101.
- [67] L. Nieto, J. Jiménez-Barbero in *Encyclopedia of Biophysics*, **2013**, pp-228-232.
- [68] J.Jeener, B. H.Meier, P.Bachmann, R.R. Ernst. *J. Chem. Phys.* **1979**, *71*, 4546.
- [69] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 521–528.
- [70] W. P. Aue, E. Bartholdi, R. R. Ernst, *J. Chem. Phys.* **1976**, *64*, 2229–2246.
- [71] M. D.George, V. Adrian, *J. Chem. Educ.*, **1993**, *70*, 1–16.
- [72] G. Wider in *Progress in Nuclear Magnetic Resonance Spectroscopy*, Elsevier, **1998**, *32*, 193-275.
- [73] R. R. Gil, A. Navarro in *Modern NMR Approaches to the Structure Elucidation of Natural Products: Data Acquisition and Applications to Compound Classes*, Royal Society of Chemistry, **2016**, *2*, pp:1-38.
- [74] D. Neuhaus, M. P. Williamson in *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, **2000**.
- [75] W. S. Price, *Concepts Magn. Reson. Part A* **2009**, *34A*, 60–61.
- [76] T. D. Claridge in *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier, **2016**.
- [77] C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby, N. J. Barron, *Org. Biomol. Chem.* **2011**, *9*, 177–184.
- [78] J. Duus, C. H. Gotfredsen, K. Bock, *Chem. Rev.* **2000**, *100*, 4589–4614.
- [79] G. M. Clore, A. M. Gronenborn, *J. Magn. Reson.* **1982**, *48*, 402–417.
- [80] D. B. Davies, *Spectrochim. Acta Part A Mol. Spectrosc.* **1991**, *47*, 1641.
- [81] M. Goldfl, T. Tarragó, M. Gairí, E. Giralt, *Methods Mol. Biol.*, **2012**, *831*, 233–259.
- [82] J. Angulo Álvarez, P. M. Nieto Mesa, *An. Quím.* **2011**, *107*, 358–366.
- [83] V. H. Pomin, X. Wang, *Molecules* **2018**, *23*(9), 2314.
- [84] G. A. Morris, R. Freeman, *J. Am. Chem. Soc.* **1979**, *101*, 760–762.
- [85] M. Rico, *An. Quím.* **2011**, *107*, 46–57.
- [86] A. L. Skinner, J. S. Laurence, *J. Pharm. Sci.*, **2008**, *97*(11), 4670-4695.
- [87] W. Becker, K. C. Bhattiprolu, N. Gubensäk, K. Zangger, *Chem. Phy. sChem.*,

- 2018**, *19*, 895–906.
- [88] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.
- [89] A. D. Gossert, W. Jahnke, *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *97*, 82–125.
- [90] B. Pedrini, P. Serrano, B. Mohanty, M. Geralt, K. Wüthrich, *Biopolymers* **2013**, *99*, 825–831.
- [91] J. J. Ziarek, F. C. Peterson, B. L. Lytle, B. F. Volkman, *Methods Enzymol.*, **2011**, *493*, 241–275.
- [92] C. W. Akey, D. S. Goldfarb, *J. Cell Biol.* **1989**, *109*, 971–982.
- [93] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* **1980**, *69*, 185–189.
- [94] L. E. Kay, P. Keifer, T. Saarinen, *J. Am. Chem. Soc.* **1992**, *114*, 10663–10665.
- [95] F. Perrin, *J. Phys. le Radium* **1926**, *7*, 390–401.
- [96] J. R. Albani in *Structure and Dynamics Macromolecules: Absorption Fluorescence Studies*, Elsevier, **2004**, pp:193–216.
- [97] W. A. Lea, A. Simeonov, *Expert Opin. Drug Discov.* **2011**, *6*, 17–32.
- [98] K. Kakehi, Y. Oda, M. Kinoshita, *Anal. Biochem.* **2001**, *297*, 111–116.
- [99] N. J. Moerke, *Curr. Protoc. Chem. Biol.* **2009**, *1*, 1–15.
- [100] C. Solera, G. Macchione, S. Maza, M. M. Kayser, F. Corzana, J. L. De Paz, P. M. Nieto, *Chem. - A Eur. J.* **2016**, *22*, 2356–2369.
- [101] P. H. Liang, S. K. Wang, C. H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 11177–11184.
- [102] S. Maza, M. Mar Kayser, G. MacChione, J. López-Prados, J. Angulo, J. L. De Paz, P. M. Nieto, *Org. Biomol. Chem.* **2013**, *11*, 3510–3525.

Capítulo 3

*Estudio estructural de GAGs
mediante cálculos de Dinámica
Molecular y RMN*

3.1 Antecedentes

Como se ha explicado en el **Capítulo 1**, muchos GAGs pueden presentar funciones biológicas específicas dependiendo de la estructura y longitud de sus cadenas. Su actividad funcional depende estrechamente de sus características estructurales, como la secuencia primaria y el mayor o menor grado de sulfatación, que les va a permitir unirse a moléculas como los factores de crecimiento. En general, se ha demostrado que cuanto más larga es la cadena de azúcar la interacción mejora considerablemente, favoreciendo un mejor reconocimiento y actividad neuronal.^[1]

Nuestro grupo de investigación se puso como objetivo sintetizar una serie de miméticos de CS, en los que se sustituían los grupos CO_2^- por grupos CH_2OH , reemplazando por tanto las unidades de GlcA por Glc^[2]. Este cambio facilitaba en buena medida la preparación de estas moléculas. En este trabajo nos centraremos en estudiar una familia de tetrasacáridos y pentasacáridos que presentan una secuencia alterna de β -D glucosa y β -D galactosamina N-acetilada (ver Figura 1.14, **Capítulo 1**).

En cuanto al grado de sulfatación, se prepararon oligosacáridos miméticos de CS de tipo E, que se sabe que interacciona con MK y PTN.^[3] También se consideraron en nuestro estudio intermedios sintéticos con diversos patrones de sustitución de distintos grupos protectores, ya que resultados previos de nuestro grupo indicaban que este tipo de derivados pueden interaccionar fuertemente con nuestras proteínas objetivo MK y PTN.^[4] Los azúcares completamente protegidos (**3** y **6**) presentan, en las unidades de N-acetilgalactosamina un grupo CF_3 sustituyendo el CH_3 del aceto de la posición 2, y en el extremo no reductor un residuo Lev situado en la posición 3, además de grupos Piv en las posiciones 6 de los anillos de glucosa (grupos COR^{IV} , R^{III} y R^{II} , Figura 1.14 del **Capítulo 1**). En primer lugar, se van a discutir los resultados estructurales teóricos obtenidos tanto para la serie pentasacáridica como tetrasacáridica en su forma libre, contrastándose posteriormente estos resultados con los datos experimentales obtenidos por técnicas de RMN.

Los análisis conformacionales teóricos de los miméticos se han llevado a cabo mediante cálculos de mecánica y dinámica molecular utilizando como campo de fuerzas GLYCAM06 dentro del programa AMBER.^[5] Se procedió al estudio, no sólo de los oligosacáridos miméticos finales **1** y **4**, sino también de los compuestos precursores **2**, **3**, **5** y **6**. Trabajos previos en el grupo de investigación habían parametrizado los residuos Bn, Bz, Piv y Lev empleando el servidor web R.E.D.^[6] Estos datos se utilizaron en esta Tesis junto a la parametrización del grupo CF_3 que fue necesario realizar; de esta manera se obtendrían los archivos de topología. Cabe añadir que, para poder llevar a cabo este cálculo, fue necesario construir previamente tanto el ligando como cada residuo en el programa MAESTRO, generando archivos PDB de cada uno de ellos. Con la idea de simplificar los cálculos de las simulaciones, se substituyó el grupo 4-metoxifenilo (MP) por un isopropilo que se encontraba parametrizado en trabajos anteriores del grupo^[7]. Se consideró que el comportamiento estructural de la cadena sacarídica no se vería muy afectado por este grupo protector del extremo reductor. Además, el residuo de isopropilo,

al ser un alcohol secundario en el terminal reductor, imita mejor la estructura del polisacárido natural y, por ello, se trabajó con él en vez de con MP.

A continuación, los archivos de topología y coordenadas creados se utilizaron como *inputs* de nuestra dinámica molecular, que consta de varias etapas de diferente duración con una previa minimización del sistema y la caja de agua (ver sección 3.5). Para aquellos ligandos que mostraron distorsiones en los anillos, se realizó una dinámica molecular más corta con restricciones promediadas en el tiempo (MD-TAR) utilizando las distancias experimentales obtenidas de los espectros NOESY.^[8]

3.2 Pentasacáridos

Comenzaremos describiendo el análisis estructural llevado a cabo en los pentasacáridos, todos ellos miméticos de CS de tipo E con una secuencia alternada de residuos de GalNAc y Glc. Se describirá primero el ligando desprotegido **1**, siguiéndole el derivado **2** parcialmente sustituido con grupos bencilos en la posición 3 de los anillos de Glc y, finalmente, el producto completamente protegido **3**. Para todos ellos se llevaron a cabo estudios cuantitativos de dinámicas moleculares (en algunos casos con restricciones) donde se determinaron aquellas conformaciones energéticamente favorables a lo largo del tiempo, se calcularon las correspondientes constantes de acoplamiento y se obtuvieron valores de las distancias interprotónicas y ángulos diedros que se compararon con los valores σ^{NOE} obtenidos de los estudios experimentales de RMN.

Estos experimentos pudieron llevarse a cabo en todos los ligandos libres, aunque el análisis del pentasacárido persustituido **3** por RMN se vio dificultado debido al fuerte solapamiento de señales que obstaculizó la medida de restricciones experimentales.

3.2.1 Pentasacárido 1

La estructura del mimético de CS representado en la Figura 3.1, consiste en una secuencia tipo E que contiene GalNAc tanto en el extremo reductor como en el no reductor, alternada con residuos de Glc. Muestra 6 grupos SO_3^- y ningún hidroxilo protegido (salvo el anomérico).

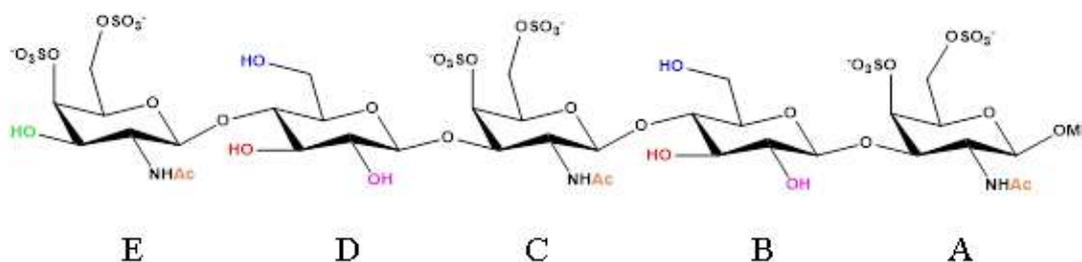


Figura 3.1. Pentasacárido 1 desprotegido de secuencia CS tipo E.

Para el estudio estructural teórico de este ligando libre se llevó a cabo, en primer lugar, una dinámica molecular de 500 ns, donde se analizaron las coordenadas de puckering de los archivos de las trayectorias generadas para esclarecer las conformaciones presentes en cada anillo de azúcar. En la Figura 3.2 se observan las coordenadas de puckering de todos ellos a lo largo de toda la simulación.

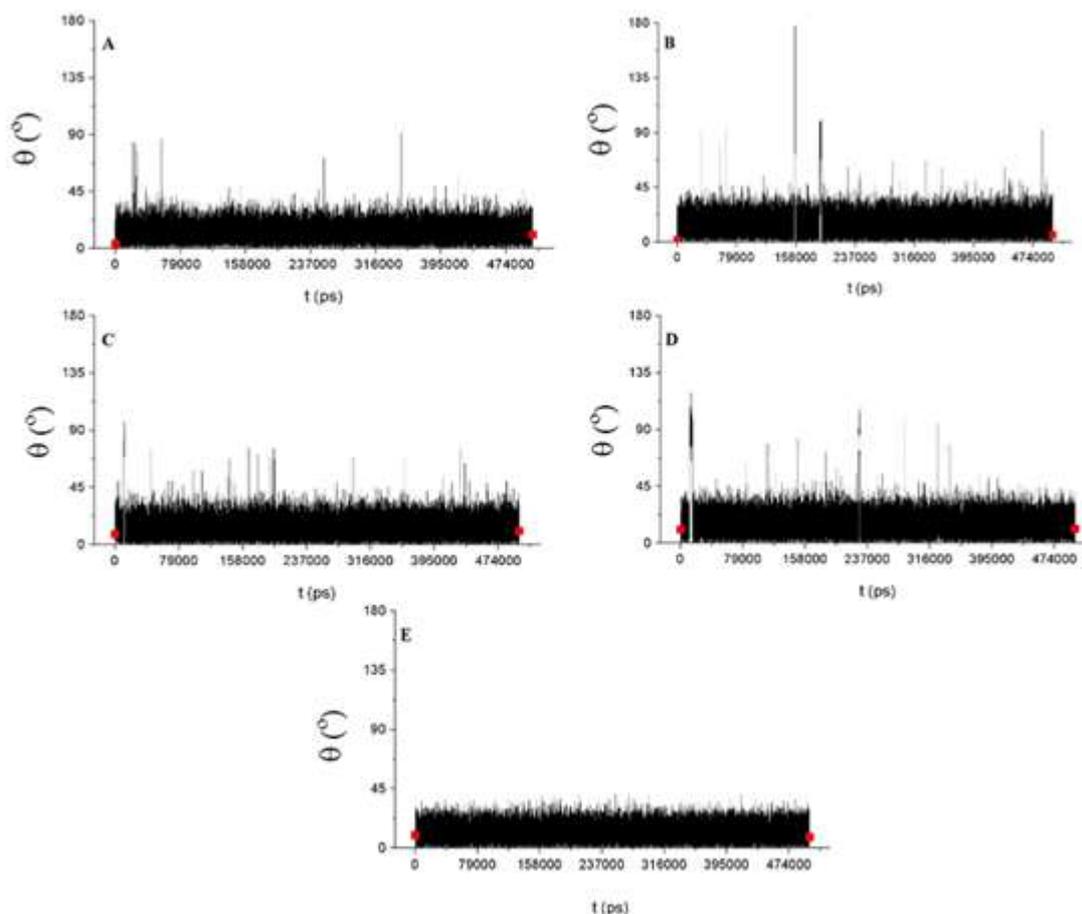


Figura 3.2. Gráficas de Cremer-Pople para el pentasacárido 1.

Siguiendo el modelo de las coordenadas de la esfera de Cremer-Pople, un ángulo θ cercano a 0° queda situado en el polo norte de la esfera de Cremer, indicando una conformación mayoritaria 4C_1 . En base a esto, se establece que durante toda la trayectoria los anillos mantienen esencialmente esta única conformación, tanto los residuos de Glc como los de GalNAc.

A continuación, se compararon los resultados cuantitativos obtenidos de los experimentos de RMN con los calculados durante la trayectoria de la dinámica molecular. Se calcularon las distancias interprotónicas comprendidas entre las parejas de hidrógenos H₁-H₃, H₁-H₅, H₃-H₅ de todos los anillos y las interglucosídicas H₁-H₄ o H₁-H₃ (véase Tabla 3.1).

Las distancias experimentales proceden de los valores σ^{NOE} medidos a 300K. Estos se obtienen de la curva ajustada a una ecuación exponencial de la velocidad inicial del crecimiento del NOE en función del tiempo de mezcla (véase punto 3.5.3).

Cabe destacar que se tuvieron ciertas dificultades para el cálculo de las distancias experimentales debido al solapamiento de las señales en los espectros de RMN, lo que impidió la medida de algunas de ellas, como se explicará más detalladamente en el punto 3.4 de este capítulo.

Distancias	1 RMN 600 MHz	1 MD 500 ns
A H1H3	2.4	2.6
A H1H5	2.8	2.6
A H3H5		2.5
Interg. BH1-AH3	2.2	2.4
B H1H3	2.8	2.8
B H1H5	2.5	2.5
B H3H5		2.7
Interg.CH1-BH4	2.6	2.3
C H1H3	2.9	2.7
C H1H5	2.5	2.6
C H3H5		2.5
Interg.DH1-CH3	2.6	2.5
D H1H3	2.6	2.8
D H1H5	2.8	2.5
D H3H5		2.7
Interg.EH1-DH4		2.3
E H1H3		2.7
E H1H5		2.6
E H3H5		2.5

Tabla 3.1. Distancias experimentales (NOESY, RMN) medidas a 600 MHz y teóricas (MD sin restricciones) para el pentasacárido **1** medidas a 300 K.

En la Tabla 3.1 se puede observar que, tanto las distancias experimentales intraanulares que pudieron medirse como las teóricas calculadas a lo largo de las trayectorias promediadas en $\langle r^{-6} \rangle^{[8]}$, presentan un valor promedio de 2.5 Å corroborando una conformación 4C_1 para los anillos (a excepción del azúcar E debido a la falta de resultados experimentales concluyentes).

Las constantes de acoplamiento ${}^3J_{HH}$ promedio, obtenidas de la trayectoria de la dinámica molecular, han sido calculadas a través de la ecuación de Altona^[9] y se han contrastado con las extraídas experimentalmente de los espectros 1H -RMN (Tabla 3.2). Estos valores coinciden tanto para los anillos de GalNAc A, C, E como para los de Glc B y D. Con los resultados analizados hasta ahora, no se ha observado ningún cambio de conformación en las dinámicas, por lo que se puede corroborar que estos valores están de acuerdo con las conformaciones 4C_1 deducidas.

$^3J_{HH}$ coupling	$^3J_{HH}$ exp. RMN 600 MHz	$^3J_{HH}$ MD 500 ns
A1A2	8.5	8.2
A2A3	11.7	10.6
A3A4	3.0	2.5
A4A5	1.4	1.2
B1B2	7.9	7.3
B2B3	8.6	8.6
B3B4	9.2	8.7
B4B5	9.5	10.2
C1C2	8.1	8.1
C2C3	10.5	10.6
C3C4	2.5	2.5
C4C5		1.2
D1D2	7.7	7.6
D2D3	9.1	9.4
D3D4	9.5	9.2
D4D5	9.3	9.6
E1E2	8.4	8.2
E2E3	10.8	10.4
E3E4	3.0	2.8
E4E5		1.3

Tabla 3.2. $^3J_{HH}$ (en Hz) experimentales medidas a 600 MHz y teóricas (MD sin restricciones) para el pentasacárido **1** medidas a 300 K.

Como se ha explicado en el **Capítulo 1**, los GAGs presentan una disposición lineal y mayormente rígida, debido a la presencia de grupos cargados, ya que la repulsión electrostática entre ellos favorece una orientación *syn* del ángulo ψ de los enlaces interglicosídicos. Si eliminamos estos grupos, la disminución de las cargas favorecería el giro de estos ángulos diedro, aumentando así la flexibilidad del enlace.

Para estudiar esta flexibilidad conformacional de forma teórica, se han representado en la Figura 3.3 los mapas de distribución de población de los ángulos (ϕ , ψ) obtenidos en la dinámica. Para las cuatro uniones interglicosídicas se repite un patrón conformacional en donde, por un lado, el ángulo ψ , que representa la mayor o menor flexibilidad del enlace, fluctúa entre -60° y 60° , aunque unos pocos *frames* en las uniones AB y CD muestran una amplitud hasta 90° . En el caso del ángulo ϕ , se observa en su mayor parte una oscilación entre 0° y 60° , y en mucha menor proporción un submínimo de población que llega a alcanzar un valor de -60° , que no cumpliría el efecto exoanomérico. A pesar de esto, se puede concluir que todas las uniones presentan principalmente una orientación *syn- ψ /exo- ϕ* . Los resultados del pentasacárido **1** están de acuerdo con los obtenidos previamente en nuestro grupo usando oligosacáridos naturales de CS tipo E.^[3]

Este hecho confirma que derivados como **1** pueden considerarse como miméticos estructurales de las secuencias naturales.

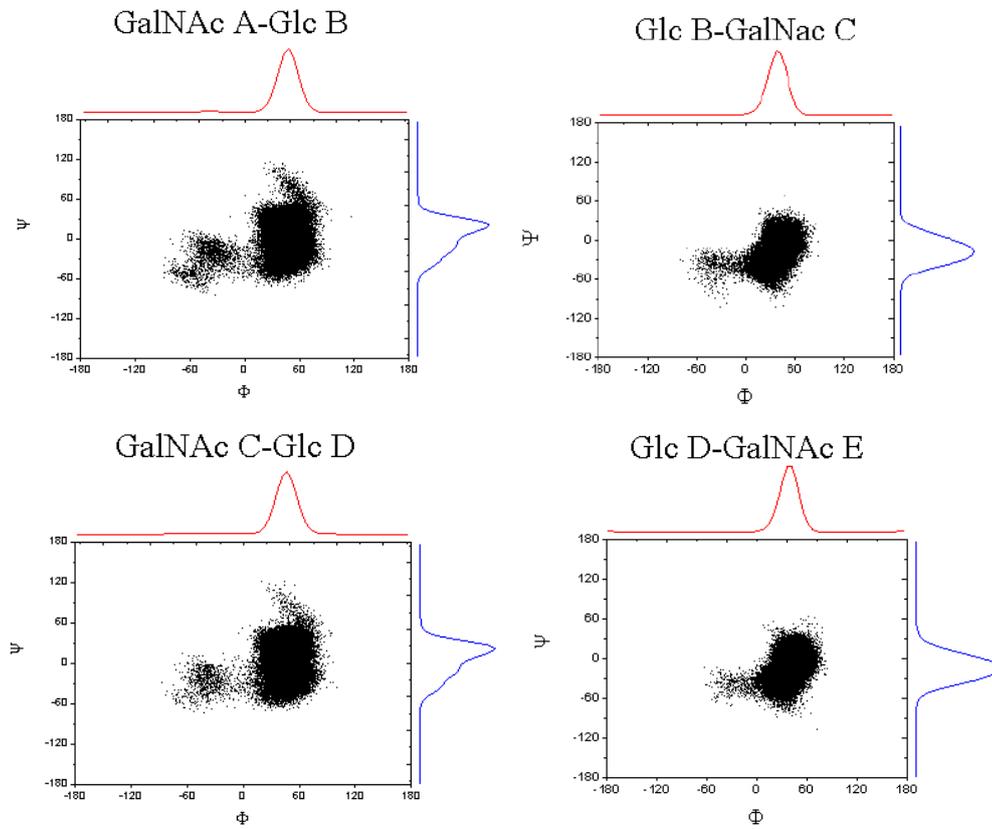


Figura 3.3. Trayectorias (ϕ , ψ) obtenidas de la dinámica molecular libre de 500 ns para el pentasacárido **1**.

3.2.2 Pentasacárido 2

La estructura del mimético de CS parcialmente sustituido representado en la Figura 3.4 es similar al anterior, diferenciándose únicamente en la presencia de grupos bencilos en las posiciones 3 de los residuos de Glc, aumentando así la hidrofobicidad del pentasacárido.

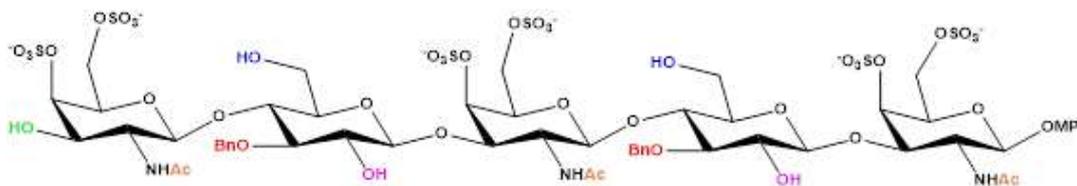


Figura 3.4. Pentasacárido **2** de secuencia CS tipo E con grupos bencilos.

Como este mimético contiene sustituyentes más voluminosos que los que presentaba el compuesto anterior, se decidió hacer una dinámica molecular por partes, dividiendo el cálculo en 50 ns hasta un total de 500 ns, chequeando cada una de ellas.

Cada etapa se lanzó consecutivamente, de manera que las coordenadas finales de una eran el archivo de entrada de la siguiente. Se observó en el cómputo de todos los *frames*, una distorsión del anillo de glucosa D (Figura 3.5):

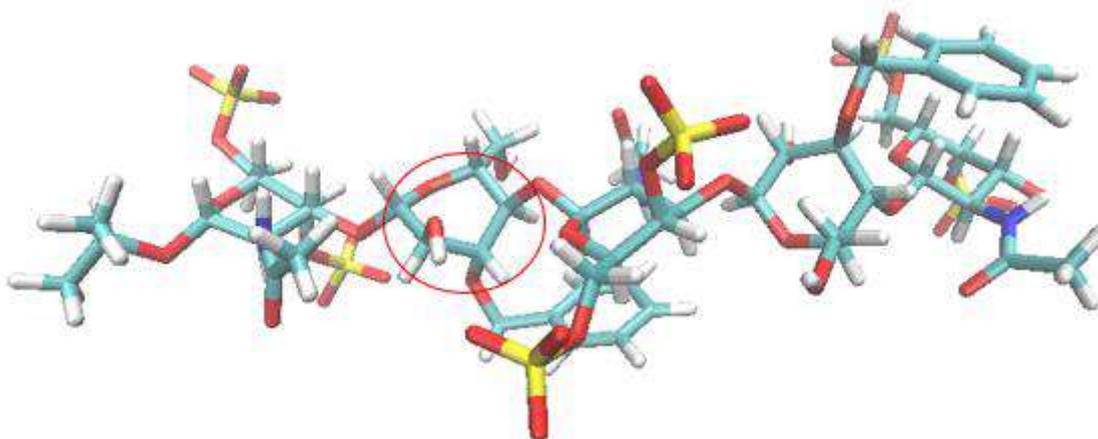
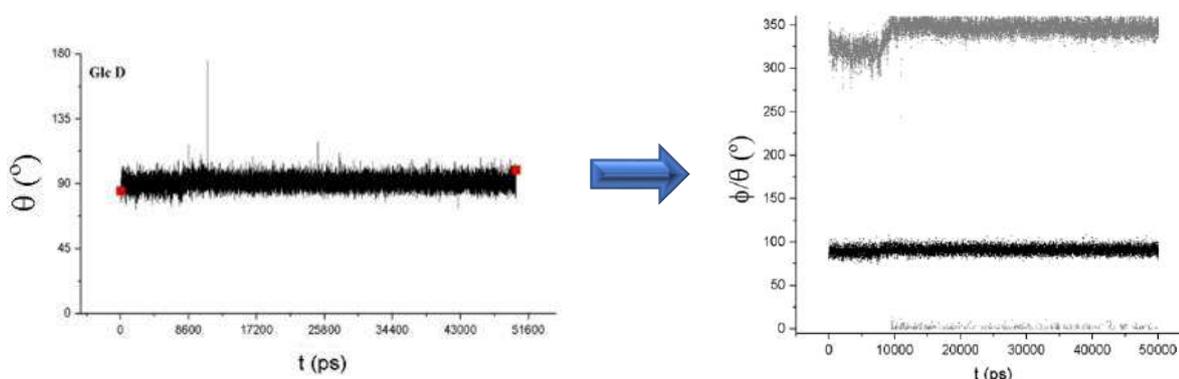


Figura 3.5. Representación de la distorsión observada en el anillo D del pentasacárido 2.

Esta deformidad se observa también en la Figura 6, donde se han representado las coordenadas de puckering para la última etapa de 50 ns en el anillo D. En la Figura 6A esta glucosa presenta un ángulo $\theta = 90^\circ$, situado en el ecuador de la esfera de Cremer-Pople donde se encuentran todas las conformaciones de bote posibles (siguiendo el modelo de la esfera de Cremer comentado en el **Capítulo 1**). Para conocer qué tipo de distorsión presenta el anillo D, se representó su correspondiente ángulo ϕ para ese valor de θ , y tal y como se observa en la Figura 3.6, ϕ presenta valores entre 330° - 350° lo que equivaldría a una conformación tipo bote 0S_2 . Se puede apreciar también, en mucha menor proporción, que ϕ también adquiere un valor de 0° que corresponde a una conformación ${}^{3,0}B$.

En el resto de azúcares se puede observar que las GalNAc A, C y E y la Glc B presentan el mismo comportamiento conformacional con un ángulo $\theta=30^\circ$ próximo a 0° , esperable para este tipo de anillos con conformación de silla 4C_1 (Figura 6B).

A)



B)

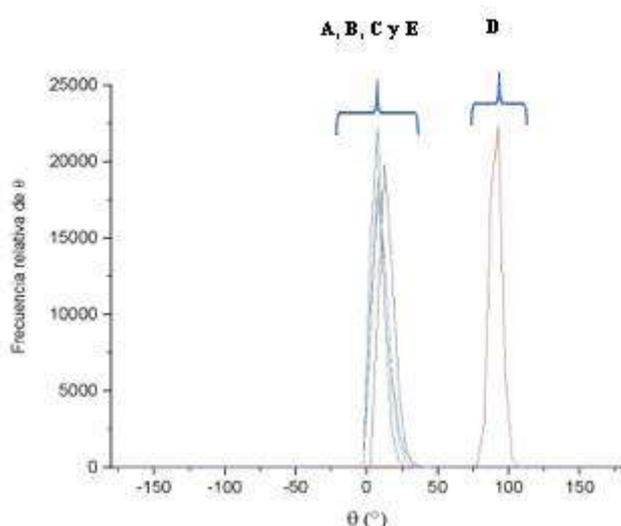


Figura 3.6. (A) A la izquierda, gráfica de Cremer-Popple para el ángulo θ del anillo D y a la derecha superposición de este ángulo con ϕ para determinar el tipo de distorsión del azúcar; (B) gráfica de la representación de la frecuencia relativa de θ para todos los anillos del compuesto 2. En los dos casos son resultados procedentes de la última etapa de 50 ns para una dinámica total de 500 ns sin restricciones.

Debido a esta distorsión, se decidió lanzar una dinámica con restricciones promediadas en el tiempo (MD-TAR), utilizando las distancias experimentales para las parejas de hidrógenos H₁-H₅, H₁-H₃ y H₃-H₅ de este anillo obtenidas por la técnica NOESY de RMN (ver apartado 3.4). Primeramente, se lanzó una simulación con una duración de 8 ns, pero como no se consiguió subsanar dicha distorsión, se probó a aumentar el tiempo de cálculo de la dinámica. Así, se decidió llevar a cabo una simulación de 50 ns, que no suponía demasiado gasto computacional, y finalmente se consiguió corregir la conformación del anillo tal y como se puede observar en la Figura 3.7:

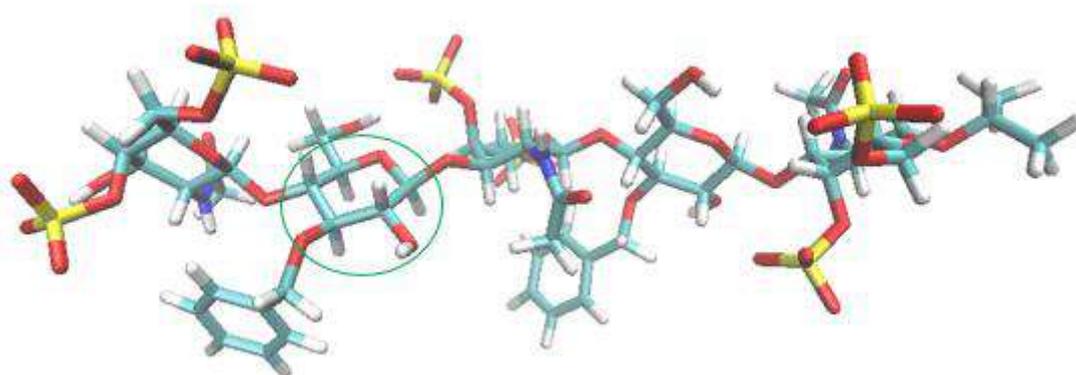


Figura 3.7. Representación de la corrección del anillo D en el pentasacárido 2 (MD-TAR de 50 ns).

En la Figura 3.8 se puede observar que el azúcar D adopta ahora la conformación esperada, presentando valores de θ correspondientes a una conformación 4C_1 .

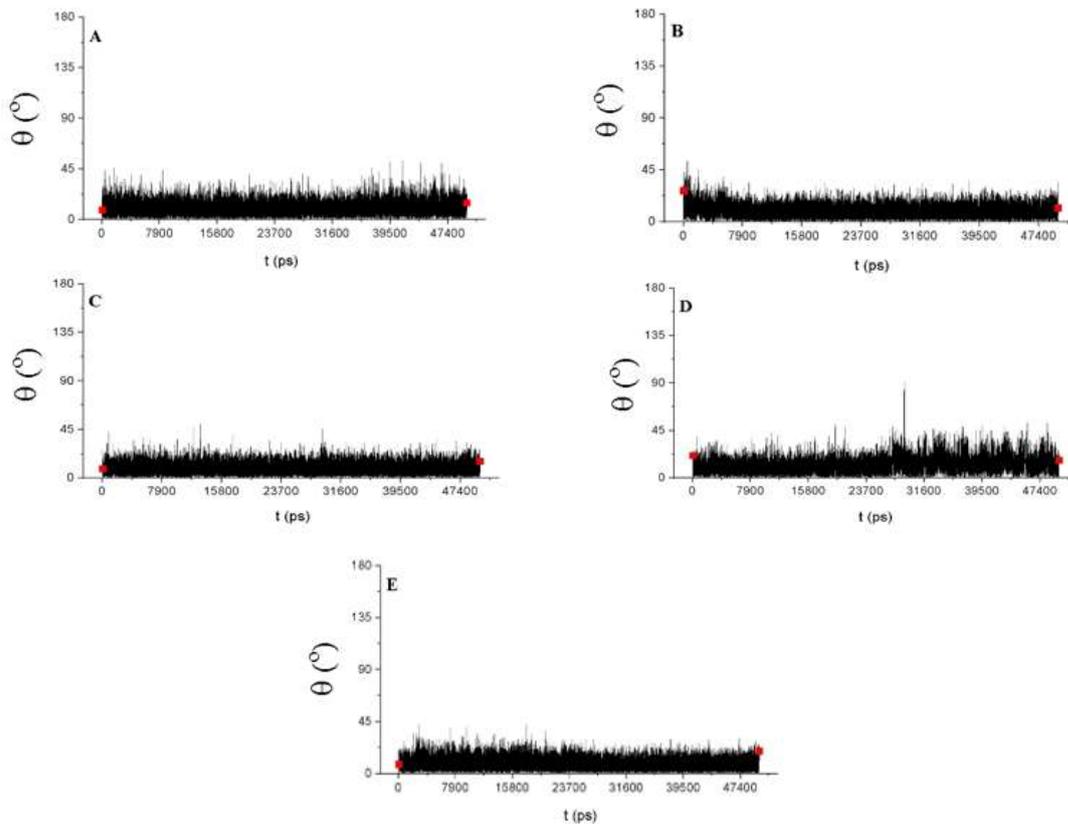
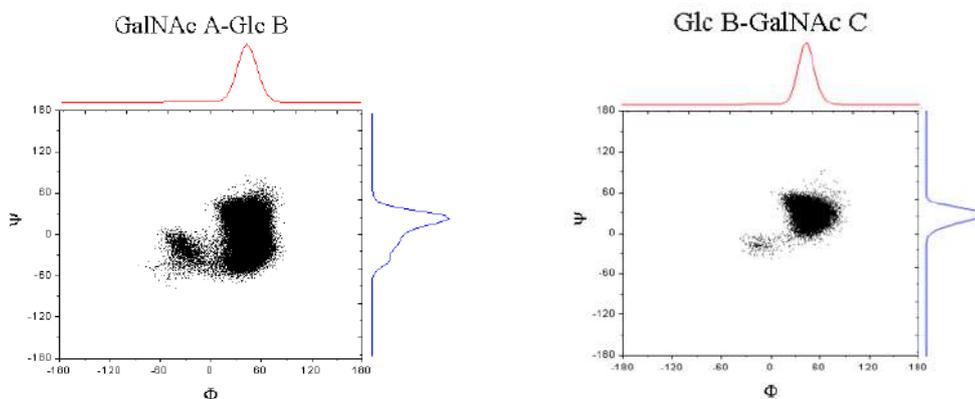


Figura 3.8. Representación de las coordenadas de Cremer-Pople (θ) para los 5 anillos del pentasacárido 2 procedentes de los cálculos de dinámica molecular TAR con restricciones.

Sobre la trayectoria de esta dinámica se midieron los ángulos de todos los enlaces interglucosídicos. Como se puede observar en la Figura 3.9, hay presencia de submínimos para las uniones GalNAc A-Glc B y GalNAc C-Glc C tanto en el ángulo ψ como ϕ . Aunque estas pequeñas poblaciones se extienden en valores de ángulo ϕ cercanos a -60° incumpliendo el efecto exoanomérico, sus proporciones son tan bajas que es prácticamente despreciable. Por lo tanto, estas poblaciones de mínima energía no afectan a la geometría mayoritaria del enlace, que se encuentra en orientación *syn- ψ /exo- ϕ* al igual que los demás.



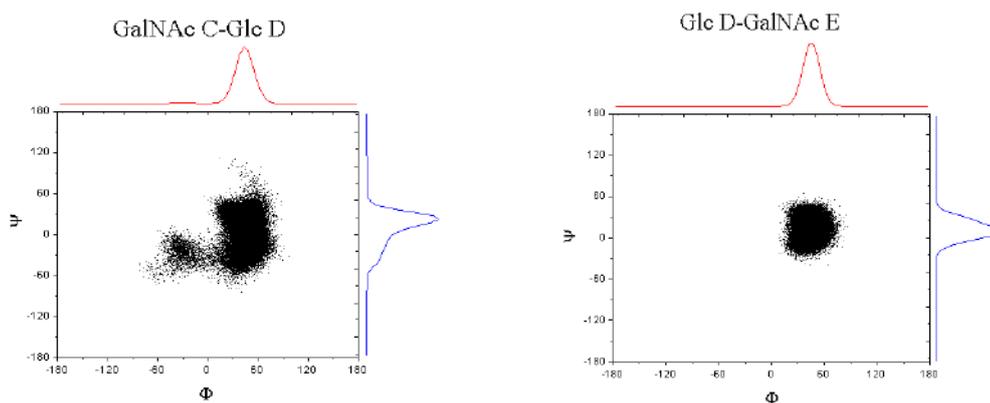


Figura 3.9. Trayectorias (ϕ , ψ) obtenidas de la dinámica molecular (con restricciones) para el pentasacárido 2.

Las distancias interprotónicas intraanulares H1-H3, H1-H5, H3-H5 y las interglicosídicas H1-H4 y H1-H3 teóricas de los anillos se comparan con las obtenidas experimentalmente por RMN en la Tabla 3.3. En ella se observa que, las distancias H1-H3 y H3-H5 obtenidas de la dinámica sin restricciones para la Glc D son inusualmente grandes para lo que se esperaría en un anillo que tiene una conformación 4C_1 . Esto se podría explicar por la distorsión de la silla tras introducir los anillos bencílicos que tienden a situarse en el plano central de la misma como sugiere el análisis de ${}^3J_{HH}$. Tras realizar la MD-TAR con los datos experimentales medidos en los espectros de RMN, conseguimos reestablecer las distancias interprotónicas esperadas en el anillo D, corrigiendo su distorsión y obteniendo valores promedio cercanos a 2.5 Å similares al resto de anillos. En esta ocasión se pudieron medir la mayoría de las distancias experimentales del pentasacárido, por lo que se pudo concluir sin ninguna duda que ambos resultados, teóricos y experimentales, se apoyaban entre sí.

Distancias	2	2	2
	RMN 600 MHz	MD 500 ns	MD-TAR 50 ns
A H1H3	2.6	2.7	2.7
A H1H5	2.7	2.4	2.5
A H3H5	2.6	2.5	2.4
Interg.BH1-AH3	2.4	2.5	2.4
B H1H3	2.7	2.8	2.7
B H1H5	2.4	2.4	2.4
B H3H5	2.5	2.6	2.6
Interg.CH1-BH4	2.6	2.4	2.5
C H1H3	2.8	2.7	2.7
C H1H5	2.6	2.4	2.5
C H3H5		2.6	2.6
Interg.DH1-CH3	2.3	2.2	2.4

Tabla 3.3. Distancias experimentales (NOESY) medidas a 600MHz, y teóricas a 300 K (MD con y sin restricciones) para el pentasacárido 2.

Distancias	2 RMN 600 MHz	2 MD 500 ns	2 MD-TAR 50 ns
D H1H3	2.8	4.2	2.7
D H1H5	2.4	2.6	2.5
D H3H5	2.9	3.9	2.6
Interg.EH1-DH4	2.3	2.3	2.4
E H1H3	2.5	2.6	2.7
E H1H5	2.7	2.5	2.6
E H3H5		2.5	2.6

Tabla 3.3. Continuación.

Con respecto a las constantes de acoplamiento se obtuvieron menos valores experimentales que para el ligando **1** ya que la presencia de los grupos aromáticos aumenta el solapamiento en el espectro de RMN y las mediciones de algunas señales se vieron afectadas.

Del mismo modo que en el análisis anterior, los valores de $^3J_{HH}$ calculados para la dinámica sin restricciones reflejan discrepancias en la glucosa D. Este residuo presenta constantes de acoplamiento calculadas mucho más pequeñas a las esperadas para este tipo de azúcares y a las experimentales. En la MD-TAR, estos valores se corrigen y se obtienen constantes de acoplamiento correspondientes a una conformación 4C_1 . En el resto de los anillos los valores respaldan este mismo tipo de conformación 4C_1 , como se puede observar en la Tabla 3.4.

$^3J_{HH}$ coupling	$^3J_{HH}$ exp. RMN 600 MHz	$^3J_{HH}$ MD 500 ns	$^3J_{HH}$ MD-TAR 50 ns
A1A2	8.9	8.5	8.5
A2A3	11.0	10.4	10.5
A3A4	2.9	2.7	2.7
A4A5		1.0	1.0
B1B2	8.0	7.5	7.5
B2B3	8.9	9.6	9.3
B3B4	9.1	9.1	9.0
B4B5	7.8	9.7	9.8
C1C2	8.2	8.5	8.5
C2C3		10.5	10.4
C3C4	2.7	2.8	2.8
C4C5		1.1	1.1

Tabla 3.4. $^3J_{HH}$ experimentales medidas a 600 MHz, y teóricas a 298 K (MD con y sin restricciones) del pentasacárido **2**.

$^3J_{HH}$ coupling	$^3J_{HH}$ exp. RMN 600 MHz	$^3J_{HH}$ MD 500 ns	$^3J_{HH}$ MD-TAR 50 ns
D1D2	7.8	1.5	7.6
D2D3		3.4	9.4
D3D4	9.2	1.7	9.2
D4D5		6.1	9.7
E1E2	8.3	8.5	8.5
E2E3		10.4	10.5
E3E4	3.0	2.9	3.2
E4E5		1.2	1.4

Tabla 3.4. Continuación.

3.2.3 Pentasacárido 3

El mimético de CS persustituido representado en la Figura 3.10, presenta todos los hidroxilos protegidos por diferentes grupos protectores, véase benzoilos, bencilos, pivaloilos y levulino, lo que le proporciona mayor hidrofobicidad que los pentasacáridos anteriores.

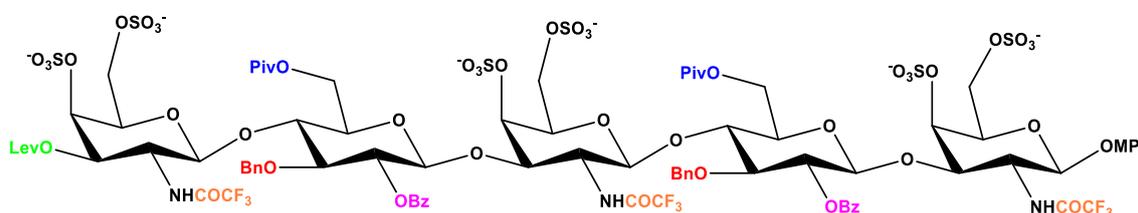


Figura 3.10. Pentasacárido 3 de secuencia CS tipo E persustituido.

Como el pentasacárido 3 presenta en su estructura más grupos protectores, y tras haber observado en el compuesto 2 distorsiones en cada una de las etapas de simulación de 50 ns para un total de 150 ns, se decidió ampliar el tiempo de cálculo a 250 ns. De esta forma se podría ir chequeando de la misma manera que antes, cada etapa de simulación. Los resultados que se obtuvieron indicaron que este pentasacárido, sorprendentemente, no mostraba apenas distorsiones a lo largo de todos los *frames* por lo que, sin necesidad de restricciones ajenas al campo de fuerzas, se procedió a analizar sus trayectorias finales.

A continuación, en la Figura 3.11, se muestran las coordenadas de Cremer-Pople, donde se puede observar que, a lo largo de toda la dinámica, los azúcares presentan la conformación esperada 4C_1 , ya que el ángulo θ se mantiene cercano a 0.

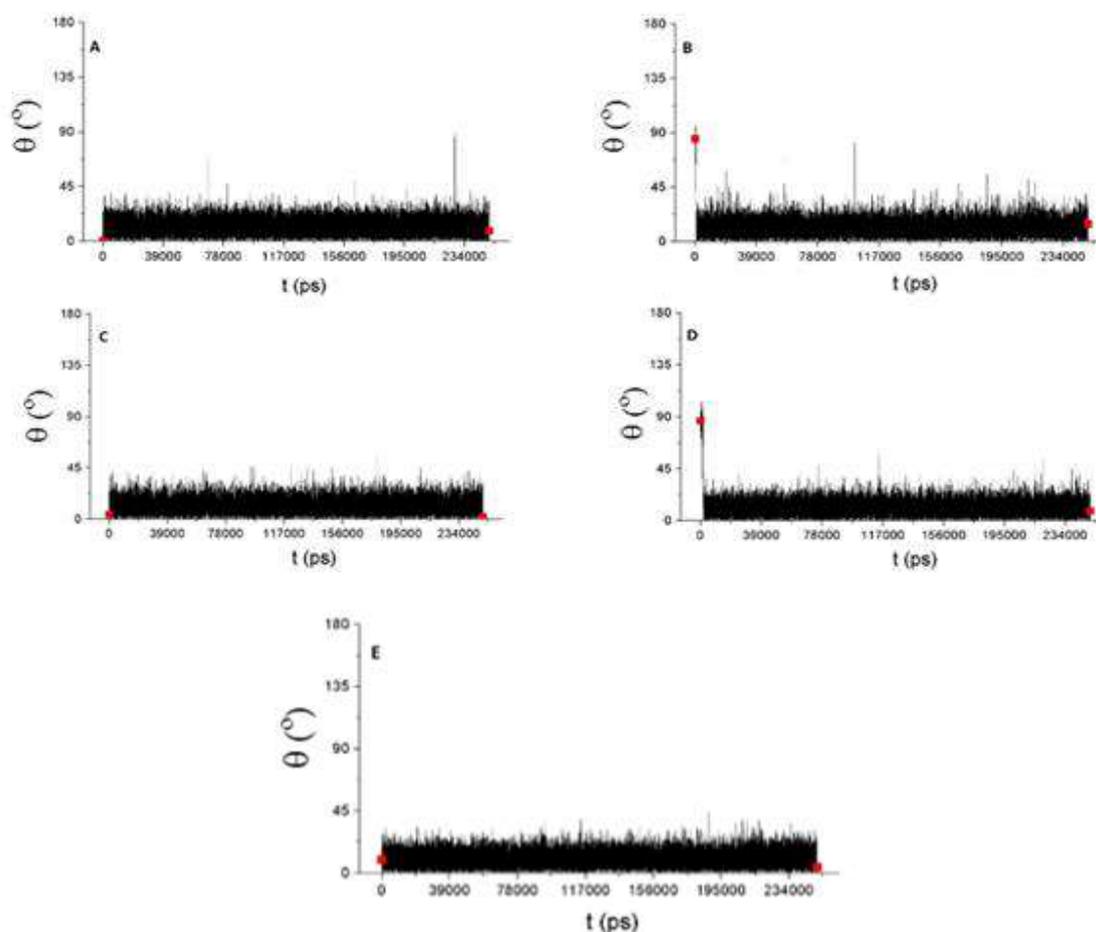


Figura 3.11. Representación de las coordenadas de Cremer-Pople para los 5 anillos del pentasacárido **3** procedentes de los cálculos de dinámica molecular de 250 ns.

Los cálculos de las distancias interprotónicas sólo fueron posibles a nivel teórico. A pesar de que el compuesto pentasacárido pudo disolverse en el buffer acuoso, las señales que se obtuvieron en el espectro de NOESY presentaron un fuerte solapamiento para ser cuantificadas, lo que impidió la determinación de los valores experimentales (ver apartado **3.4** del capítulo). Las distancias teóricas con un valor promedio de 2.5 Å apoyan las conformaciones 4C_1 para todos los residuos de azúcar (Tabla 3.5):

Distancias	3 MD 250 ns
A H1H3	2.6
A H1H5	2.5
A H3H5	2.5
Interg.BH1-AH3	2.4
B H1H3	2.7
B H1H5	2.5
B H3H5	2.5
Interg.CH1-BH4	2.4

Tabla 3.5. Distancias teóricas (MD sin restricciones) para el pentasacárido **3** medidas a 300 K.

Distancias	3 MD 250 ns
C H1H3	2.6
C H1H5	2.6
C H3H5	2.5
Interg.DH1-CH3	2.4
D H1H3	2.7
D H1H5	2.5
D H3H5	2.6
Interg.EH1-DH4	2.4
E H1H3	2.5
E H1H5	2.5
E H3H5	2.5

Tabla 3.5. Continuación.

En cambio, sí fue posible determinar algunos valores experimentales de $^3J_{HH}$, aunque en menor número que en los pentasacáridos anteriores debido al solapamiento que presentan las señales en el espectro de 1H -NMR por la presencia de más sustituyentes en su estructura (Figura 3.12):

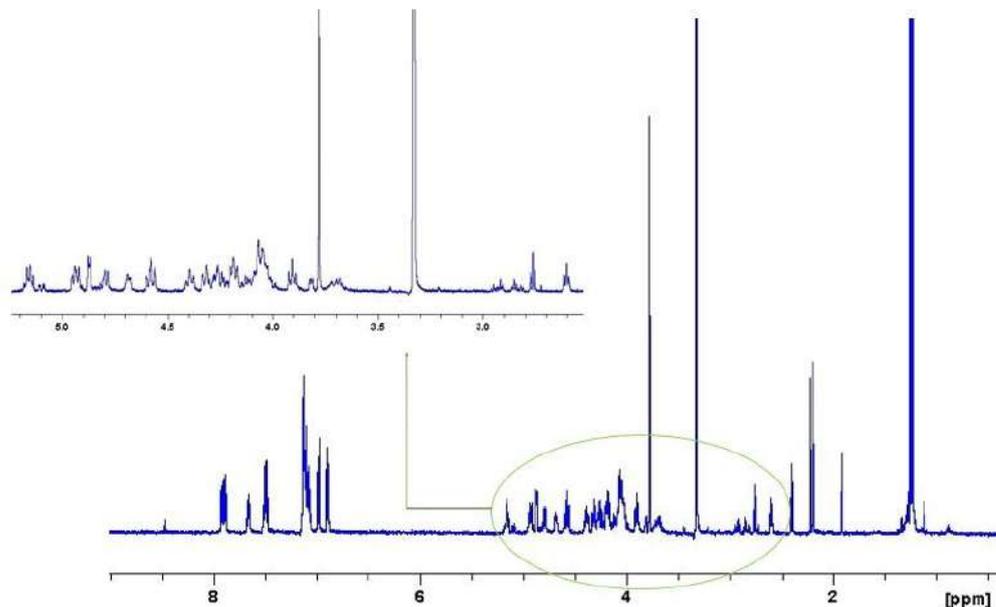


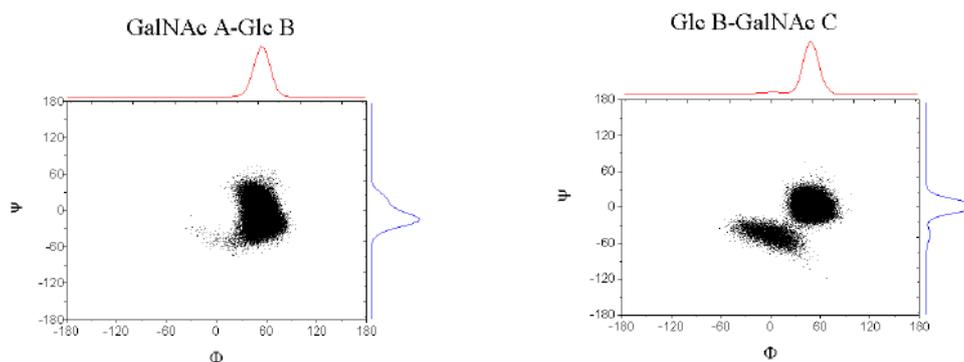
Figura 3.12. Espectro de 1H -NMR para el pentasacárido 3.

Al comparar las $^3J_{HH}$ promedio obtenidas de la trayectoria final de la dinámica con las experimentales, se corroboró experimentalmente la prevalencia de la conformación 4C_1 canónica para los residuos de GalNAc de los anillos A, C y E, y para los de tipo Glc B y D (Tabla 3.6).

$^3J_{HH}$ coupling	$^3J_{HH}$ exp. RMN 600 MHz	$^3J_{HH}$ MD 250 ns
A1A2	8.4	8.5
A2A3	10.7	10.5
A3A4		2.7
A4A5		1.0
B1B2		7.5
B2B3	8.9	9.3
B3B4	9.0	9.0
B4B5		9.8
C1C2	8.0	8.5
C2C3		10.4
C3C4	2.5	2.8
C4C5		1.1
D1D2		7.6
D2D3	8.5	9.4
D3D4		9.2
D4D5	8.0	9.7
E1E2	8.5	8.5
E2E3	10.9	10.5
E3E4		3.2
E4E5		1.4

Tabla 3.6. $^3J_{HH}$ (en Hz) experimentales medidas a 600 MHz y teóricas (MD sin restricciones) para el pentasacárido **3** medidas a 300 K.

Los mapas (ψ , ϕ) obtenidos de las coordenadas de la dinámica muestran el comportamiento conformacional de las uniones interglucosídicas, reflejando todas ellas una conformación tipo *syn- ψ* (Figura 3.13). Cabe destacar que las uniones GalNAc A-Glc B y GalNAc C-Glc D muestran una mayor flexibilidad con valores de ψ entre -60° y 60° . En definitiva, la presencia de un mayor número de grupos protectores en el compuesto **3** no provoca un cambio significativo en la conformación del enlace interglucosídico.



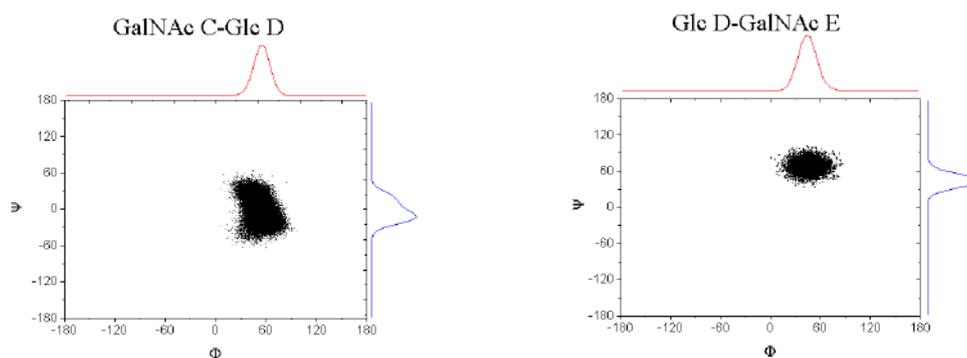


Figura 3.13. Trayectorias (ϕ , ψ) obtenidas de la dinámica molecular para el pentasacárido **3**.

3.3 Tetrasacáridos

Siguiendo el esquema de trabajo de los pentasacáridos, se van a exponer a continuación los resultados obtenidos del estudio estructural de los compuestos tetrasacáridicos. Estos productos son, de nuevo, derivados de CS de tipo E que presentan una secuencia alternada con residuos de GalNAc y Glc, pero con una unidad monosacáridica menos. Se comenzará describiendo el ligando que no presenta grupos protectores y se continuará en orden creciente de sustitución en el azúcar.

De igual manera que los compuestos anteriores, se realizaron primeramente simulaciones moleculares (aplicando restricciones cuando fue necesario), calculando los valores promedio de las constantes de acoplamiento, distancias interprotónicas y ángulos diedros de las trayectorias finales, confrontándolos con los resultados extraídos de los espectros de NOESY y ^1H -RMN.

3.3.1 Tetrasacárido 4

La estructura de este mimético de CS de tipo E, representada en la Figura 3.14, es homóloga a la del pentasacárido **1**, pero presentando una unidad menos de GalNAc en el terminal reductor. El compuesto posee cuatro grupos SO_3^- y ningún grupo protector sobre los hidroxilos (salvo el anomérico).

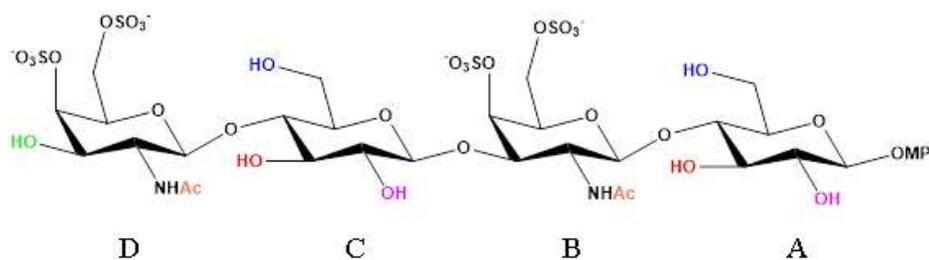


Figura 3.14. Tetrasacárido **4** de secuencia CS tipo E.

Del mismo modo que en el pentasacárido **1**, se realizaron los diferentes cálculos sobre la trayectoria final de la simulación molecular de 500 ns.

El análisis de las coordenadas de plegamiento indica que todos los anillos presentan la conformación canónica de tipo silla 4C_1 con valores esperados de θ entre 0° - 30° para los residuos de GalNAc y Glc (Figura 15):

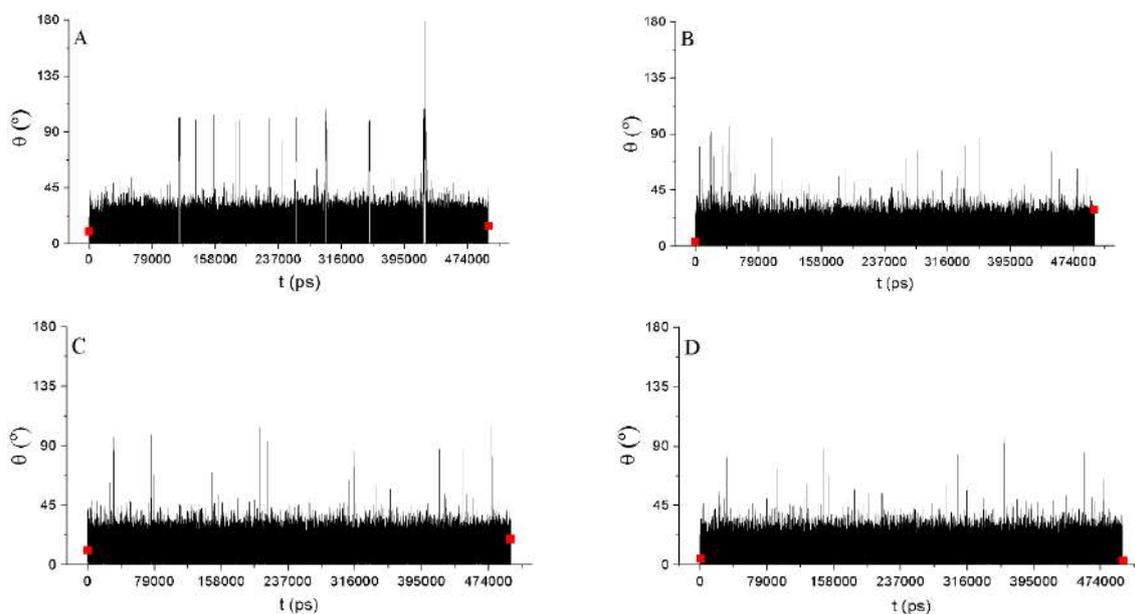


Figura 15. Gráficas de Cremer-Pople para el tetrasacárido **4**.

Estos resultados se corroboraron con los valores de las distancias interprotónicas intraanulares, las cuáles no mostraron distorsión en ninguno de los anillos. Estas distancias estaban de acuerdo con las obtenidas en el experimento de NOESY, con valores promedios de 2.5 \AA y las constantes de acoplamiento obtenidas experimentalmente también respaldaban las ${}^3J_{\text{HH}}$ calculadas en la simulación de MD (Tablas 3.7 y 3.8):

Distancias	4	
	RMN 600 MHz	MD 500 ns
A H1H3	2.8	2.8
A H1H5	2.4	2.5
AH3H5		2.7
Interg.BH1-AH4	2.4	2.4
B H1H3	2.5	2.7
B H1H5	2.9	2.6
B H3H5	2.6	2.5
Interg.CH1-BH3	2.3	2.4
C H1H3	2.8	2.8
C H1H5	2.5	2.5
C H3H5		2.7
Interg.DH1-CH4	2.5	2.8
D H1H3	2.7	2.7
D H1H5	2.6	2.6
D H3H5	2.6	2.5

Tabla 3.7. Distancias experimentales (NOESY) y teóricas (MD sin restricciones) para el tetrasacárido 4 medidas a 300 K.

$^3J_{HH}$ coupling	$^3J_{HH}$	
	exp. RMN 600 MHz	MD 500 ns
A1A2	8.4	7.3
A2A3	10.7	8.6
A3A4		8.6
A4A5		10.2
B1B2		8.2
B2B3	8.9	10.6
B3B4	3.1	2.7
B4B5		2.0
C1C2	8.0	7.3
C2C3		8.6
C3C4	8.4	8.6
C4C5		10.1
D1D2		8.2
D2D3	8.5	10.3
D3D4		2.7
D4D5	2.9	1.9

Tabla 3.8. $^3J_{HH}$ experimentales medidas a 600 MHz y teóricas (MD sin restricciones) para el tetrasacárido 4 medidas a 300 K.

Como se ha comentado anteriormente, la flexibilidad conformacional del enlace interglucosídico puede verse afectada por la presencia de cargas negativas en la molécula debido a las repulsiones electrostáticas entre los grupos, que confieren rigidez a los

enlaces. En base a esto, al presentar los tetrasacáridos dos grupos sulfatos menos que la serie pentasacáridica, se esperaba que existiera una menor rigidez en los ángulos glicosídicos ψ (Figura 3.16), aunque se puede pensar que al presentar éstos un enlace interglicosídico más, el efecto de las cargas podría repartirse en mayor grado confiriendo al azúcar cierta flexibilidad.

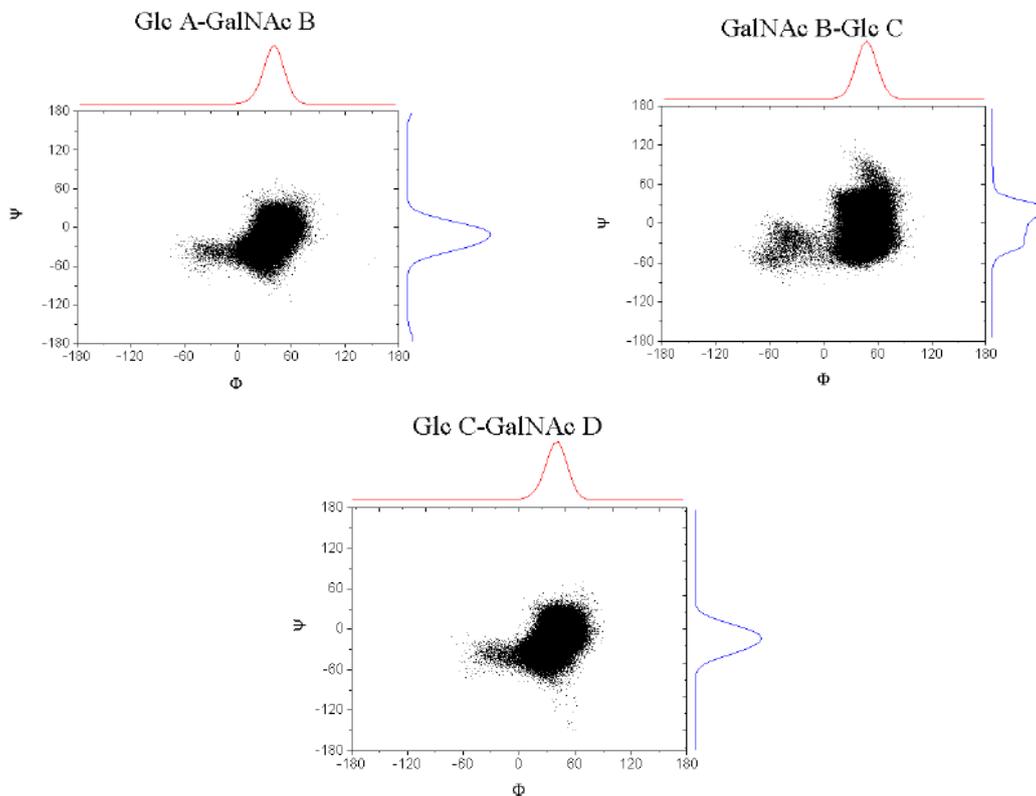


Figura 3.16. Trayectorias (ϕ , ψ) obtenidas de la dinámica molecular para el tetrasacárido **4**.

En los tres mapas de población podemos observar que el ángulo ψ fluctúa entre -60° y 30° para los enlaces AB y CD, indicando cierta flexibilidad, mientras que la unión GalNAc B-Glc C muestra una mayor flexibilidad, llegando a alcanzar ángulos ψ de hasta 60° . Por otro lado, el ángulo ϕ muestra una población mayoritaria entre 0° - 60° con un submínimo, muy poco significativo, en torno a -60° que incumple el efecto exoanomérico. A pesar de estas apreciaciones, el compuesto **4** presenta una geometría tipo *syn- ψ /exo- ϕ* predominante en todos los anillos. Estos resultados coinciden con los obtenidos para el pentasacárido **1**.

3.3.2 Tetrasacárido 5

El siguiente compuesto que estudiamos fue el mimético tetrasacáridico **5**, representado en la Figura 3.17, que está dibencilado en las posiciones 3 de los residuos de glucosa, aumentando así la hidrofobicidad del glicosaminoglicano con respecto al ligando **4**.

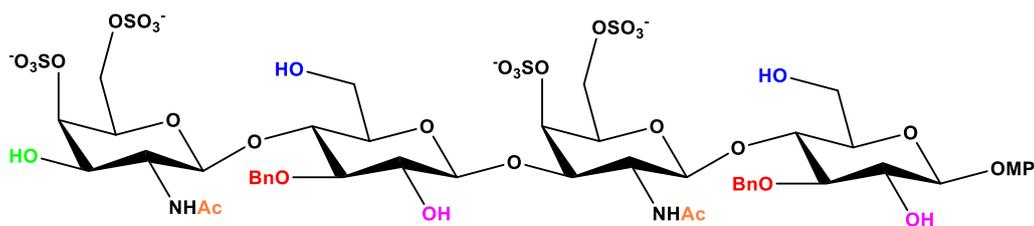


Figura 3.17. Tetrasacárido 5 de secuencia CS tipo E, sustituido con grupos bencilo en la posición 3 de las unidades de Glc.

Para este tetrasacárido se llevó a cabo una dinámica molecular de 250 ns ya que, en el caso de su homólogo (pentasacárido **2**), con una simulación de 150 ns se observaron distorsiones en algunos de los anillos. Por tanto, la idea inicial era alargar el tiempo de cálculo para ver si podían corregirse las distorsiones, sin necesidad de recurrir a una MD-TAR.

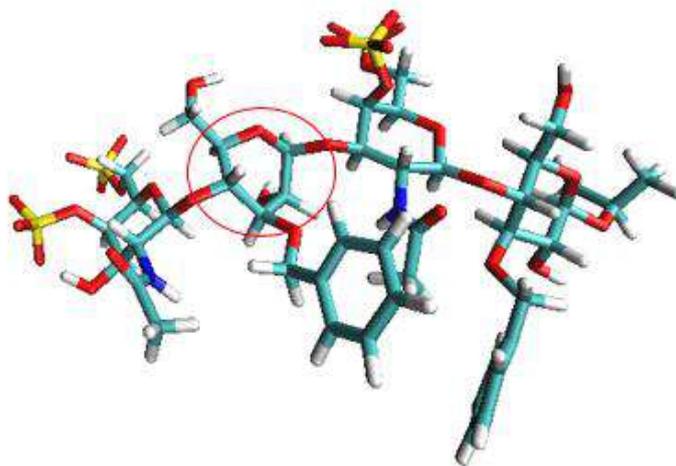


Figura 3.18. Representación de la distorsión en el anillo C del compuesto 5 para la dinámica libre de 250 ns.

Como se puede observar en la Figura 3.18 y en las gráficas de Cremer-Pople de la Figura 3.19A, la glucosa C, a excepción de los primeros *frames*, presenta una distorsión, al igual que ocurría en el compuesto **2**, con valores de ángulo θ de 90° y, en menor medida, de 180° . En esa misma Figura 3.19B, se muestra la superposición de la gráfica del ángulo θ con los datos del ángulo ϕ para el anillo C. En ella puede verse que cuando estamos en el ecuador de la esfera de Cremer-Pople ($\theta = 90^\circ$), el ángulo ϕ adquiere valores cercanos a 0 y entre $330\text{-}350^\circ$, lo que indica conformaciones de bote tipo 3,0B y 0S_2 , respectivamente.

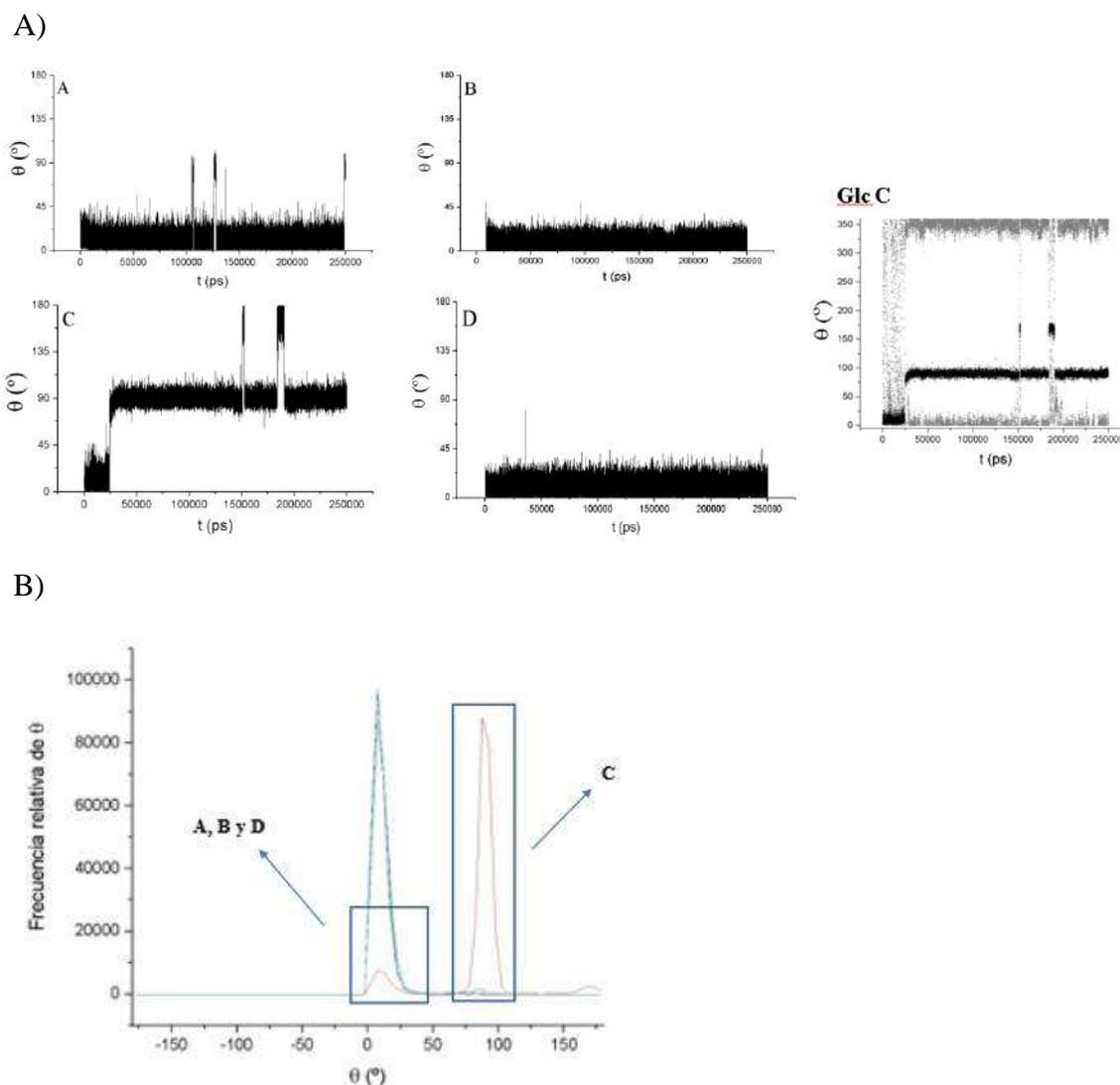


Figura 3.19. (A) Gráficas de Cremer-Popple para todos los anillos del tetrasacárido 5. (B) Gráfica de frecuencia relativa. Todos los resultados proceden de la dinámica molecular de 250 ns sin restricciones.

A excepción del anillo C, los demás residuos presentan una conformación 4C_1 por su valor del ángulo θ entre 0° - 30° . Para intentar corregir la distorsión del anillo C, se procedió a realizar una dinámica TAR utilizando las distancias experimentales medidas en NOESY (Figura 3.20).

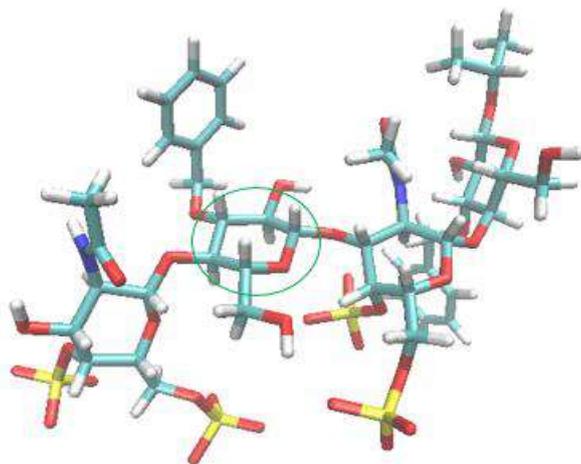


Figura 3.20. Representación de la corrección del anillo C del compuesto 5 en la MD-TAR de 8 ns.

Una duración para la simulación de 8 ns fue suficiente para que se corrigiese la distorsión del anillo. Como se puede observar en la Figura 3.21, todos los azúcares presentaban ahora el mismo valor de θ correspondiente a una conformación de silla 4C_1 .

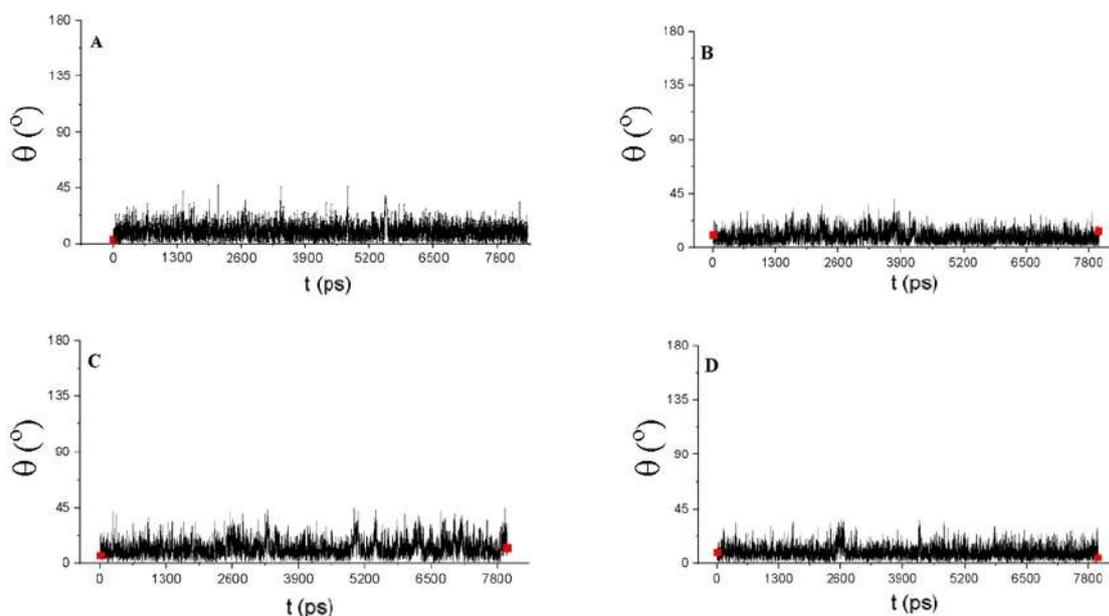
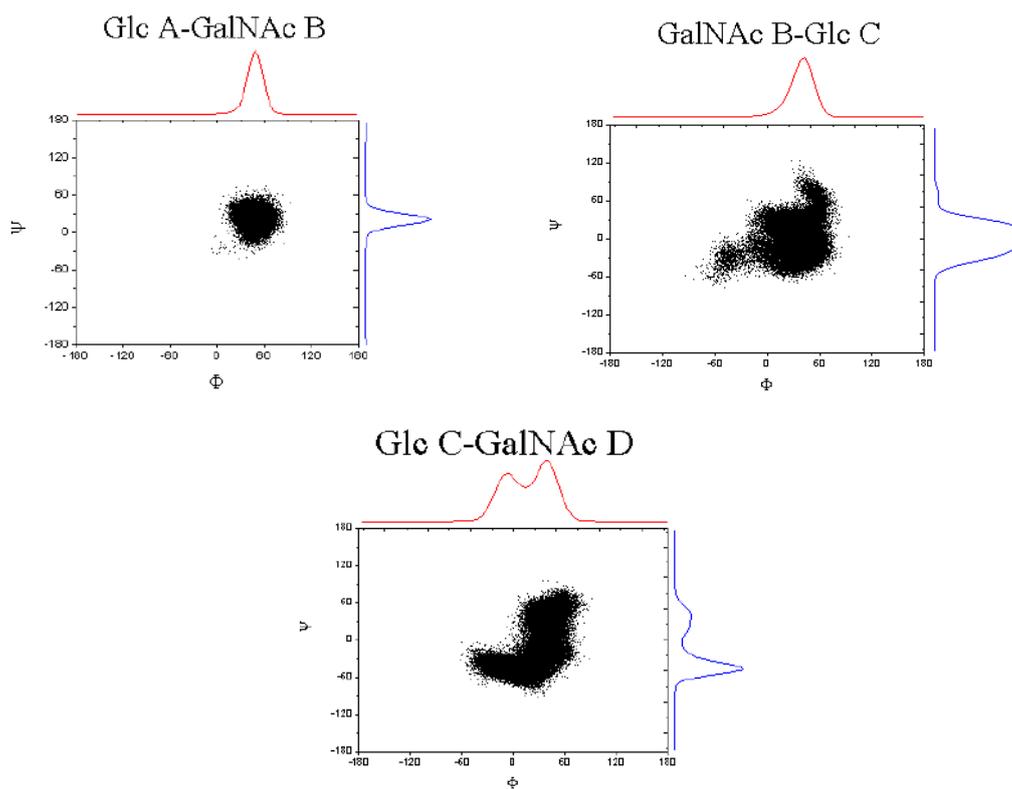


Figura 3.21. Gráficas de Cremer-Popple para el tetrasacárido 5 de la MD-TAR.

Sobre estas mismas coordenadas se realizó el estudio de los ángulos de torsión. En la Figura 3.22 se comparan los resultados obtenidos en la dinámica libre de 250 ns y en la MD-TAR de 8 ns.

A)



B)

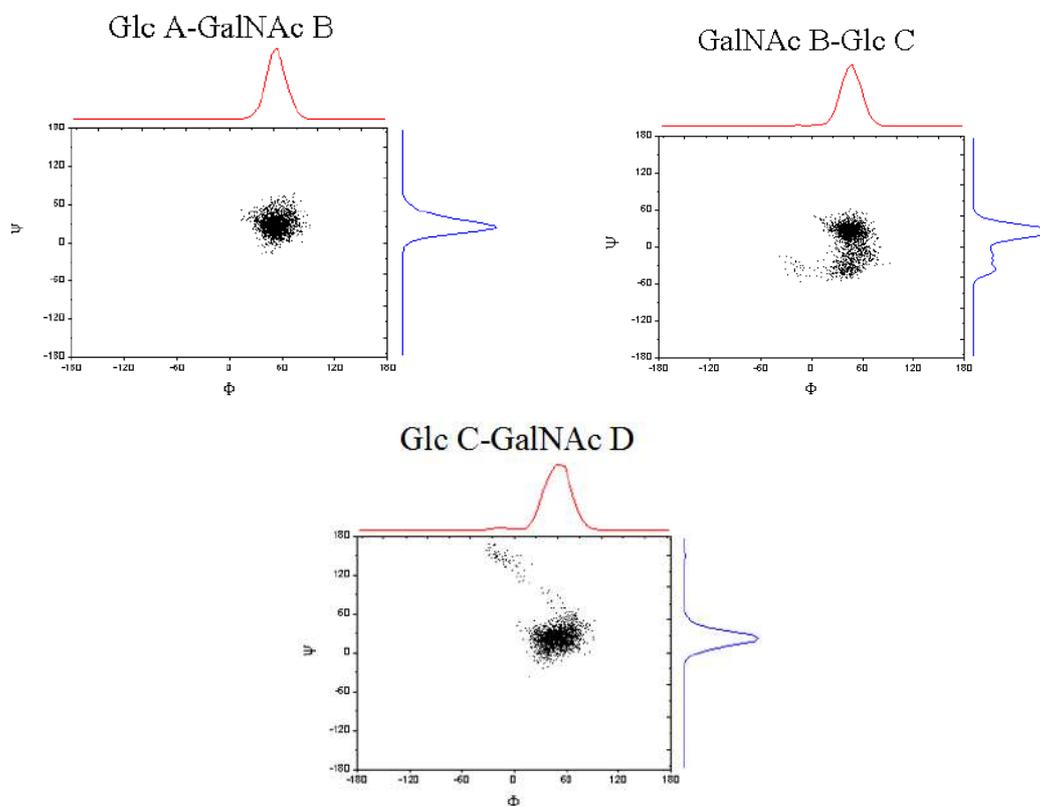


Figura 3.22. (A) Trayectorias (ϕ , ψ) obtenidas de la dinámica molecular libre de 250 ns para los tres enlaces interglucosídicos del tetrasacárido 5. (B) Trayectoria (ϕ , ψ) analizada de la dinámica molecular MD-TAR de 8ns.

Como podemos observar en la Figura 3.22A, la gráfica que corresponde al enlace interglucosídico Glc C-GalNAc D de la dinámica libre presenta una flexibilidad mayor en el ángulo ϕ con respecto al de la MD-TAR. Este enlace presenta en la dinámica libre una población que se extiende desde 60° hasta -60° , lo que indica que en este último tramo no se cumple el efecto exoanomérico. Al aplicar las restricciones experimentales, la población del ángulo ϕ se concentra únicamente entre 0 y 60° , lo que está de acuerdo con el efecto exoanomérico (Figura 3.22B). El resto de uniones glicosídicas siguen el mismo patrón, con un ángulo ψ que oscila entre -60 y 60° . Si comparamos ambas dinámicas, con y sin restricciones, observamos que hay mayor rigidez en los enlaces de esta última. De este estudio se puede concluir que todos los enlaces glicosídicos presentan una conformación *syn- ψ /exo- ϕ* .

Siguiendo con el análisis de ambas dinámicas (con y sin restricciones), se procedieron a calcular las distancias interprotónicas teóricas y se compararon con las medidas experimentalmente en el espectro de NOESY (Tabla 3.9):

Distancias	5	5	5
	RMN 600 MHz	MD 250 ns	MD-TAR 8 ns
A H1H3	3.1	2.7	2.7
A H1H5	2.5	2.5	2.8
A H3H5		2.6	2.8
Interg.BH1-AH4	2.4	2.6	2.7
B H1H3	3.1	2.7	2.7
B H1H5		2.5	2.5
B H3H5		2.5	2.5
Interg.CH1-BH3	2.6	2.3	2.5
C H1H3	2.7	4.0	2.7
C H1H5	2.5	2.6	2.5
C H3H5	2.6	3.9	2.7
C Interg.DH1-CH4	2.2	2.4	2.7
D H1H3	2.7	2.7	2.7
D H1H5	2.5	2.6	2.5
D H3H5		2.5	2.6

Tabla 3.9. Distancias experimentales (NOESY) y teóricas (MD con y sin restricciones) para el tetrasacárido 5 medidas a 300 K.

A la vista de los datos de la Tabla 3.9, las distancias interprotónicas H₁-H₃ y H₃-H₅ del anillo C mostraban, en la dinámica libre, valores más altos a lo esperable para un anillo de glucosa con una conformación ⁴C₁. Utilizando los datos experimentales como restricciones en la MD-TAR, conseguimos la corrección de la conformación del anillo y unas distancias teóricas compatibles con las experimentales, confirmando una conformación ⁴C₁.

Las constantes de acoplamiento teóricas para la simulación sin restricciones, además de respaldar la distorsión del azúcar C con valores de $^3J_{HH}$ muy pequeños, indicaban que el enlace A4-A5 mostraba una constante pequeña, que no corresponde a una glucosa en conformación 4C_1 . Tras realizar la MD-TAR, ambos azúcares se corrigen, obteniendo unos valores de $^3J_{HH}$ de Glc y GalNAc en conformación 4C_1 (Tabla 3.10).

$^3J_{HH}$ coupling	$^3J_{HH}$ exp. RMN 600 MHz	$^3J_{HH}$ MD 250 ns	$^3J_{HH}$ MD-TAR 8 ns
A1A2	8.4	7.0	7.5
A2A3	9.5	9.7	9.8
A3A4		8.9	8.9
A4A5		2.6	8.6
B1B2	8.9	8.4	8.5
B2B3	11.4	10.3	10.5
B3B4	2.9	2.8	2.8
B4B5		1.0	0.9
C1C2	7.8	2.6	7.6
C2C3	8.0	3.6	9.6
C3C4	9.2	2.8	8.8
C4C5		5.2	9.5
D1D2	8.2	8.4	8.4
D2D3		10.4	10.4
D3D4		3.0	2.9
D4D5		1.2	1.2

Tabla 3.10. $^3J_{HH}$ experimentales medidas a 600 MHz, y teóricas a 300 K (MD con y sin restricciones) del tetrasacárido 5.

3.3.3 Tetrasacárido 6

El último tetrasacárido libre objeto de estudio fue el representado en la Figura 3.23, el mimético persustituido **6**, que presenta también todos los hidroxilos protegidos por grupos protectores, como el pentasacárido **3** descrito anteriormente.

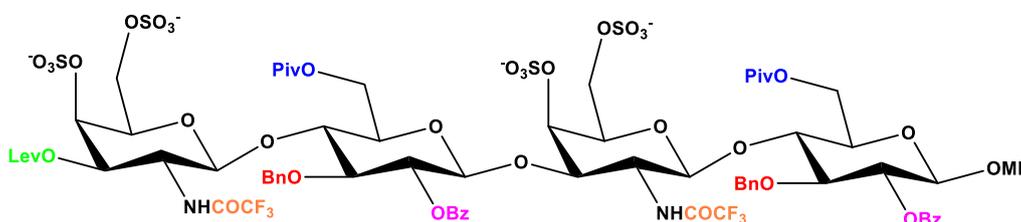


Figura 3.23. Tetrasacárido **6** de secuencia CS de tipo E persustituido.

El alto número de sustituyentes complicó el análisis tanto teórico como experimental para este ligando. A nivel teórico, se siguió el mismo protocolo de cálculo que para su

homólogo pentasacárido (compuesto **3**). Se comenzó realizando una dinámica molecular libre, sin restricciones, con un tiempo de cálculo largo de 250 ns que se fue chequeando desde el principio, y en la que se observaron distorsiones en las glucosas A y C (Figura 3.24).

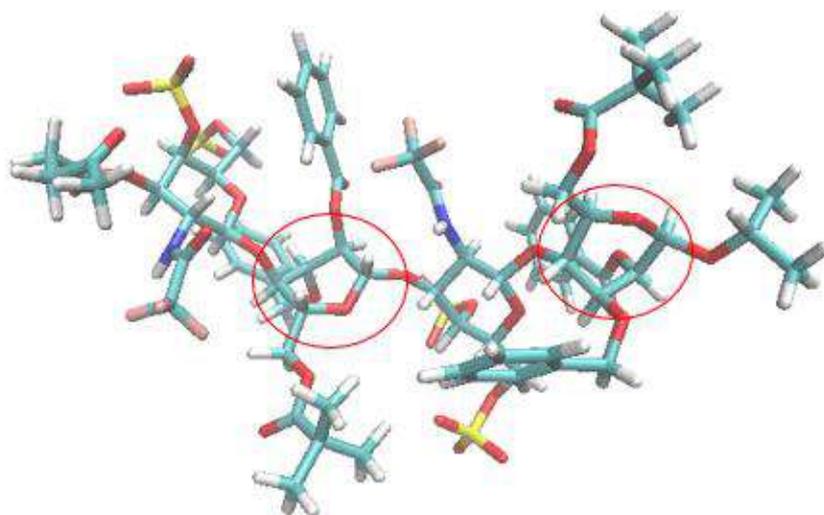


Figura 3.24. Representación de las distorsiones en los anillos A y C del tetrasacárido **6** para la dinámica de 250 ns (sin restricciones).

En la Figura 3.25 se puede observar que, para el anillo A, el ángulo θ pasa de 0° - 30° en los primeros frames, lo que corresponde a una conformación tipo silla 4C_1 , a valores de θ alternados entre 90° y 180° , lo que indicaría una conformación de bote y silla 1C_4 respectivamente. Para conocer más a fondo el tipo de distorsión que está teniendo lugar en A, se representó una vez más el ángulo ϕ superponiéndolo con el gráfico anterior. Así, cuando nos encontramos en el ecuador de la esfera de Cremer-Pople (valores de $\theta = 90^\circ$), los valores de ϕ son cercanos a 330° apuntando a conformaciones tipo bote torcido 0S_2 . En cambio, cuando θ se sitúa en el polo de la esfera indicando un ángulo de 180° , el azúcar presenta valores indeterminados para ϕ . Para el azúcar C la situación es similar, empezando con un ángulo θ cercano a 0° , que apunta a una conformación 4C_1 durante un breve período de tiempo. El resto de la trayectoria permanece en un ángulo θ cercano a 90° , que indica una distorsión del anillo de tipo bote torcido 0S_2 .

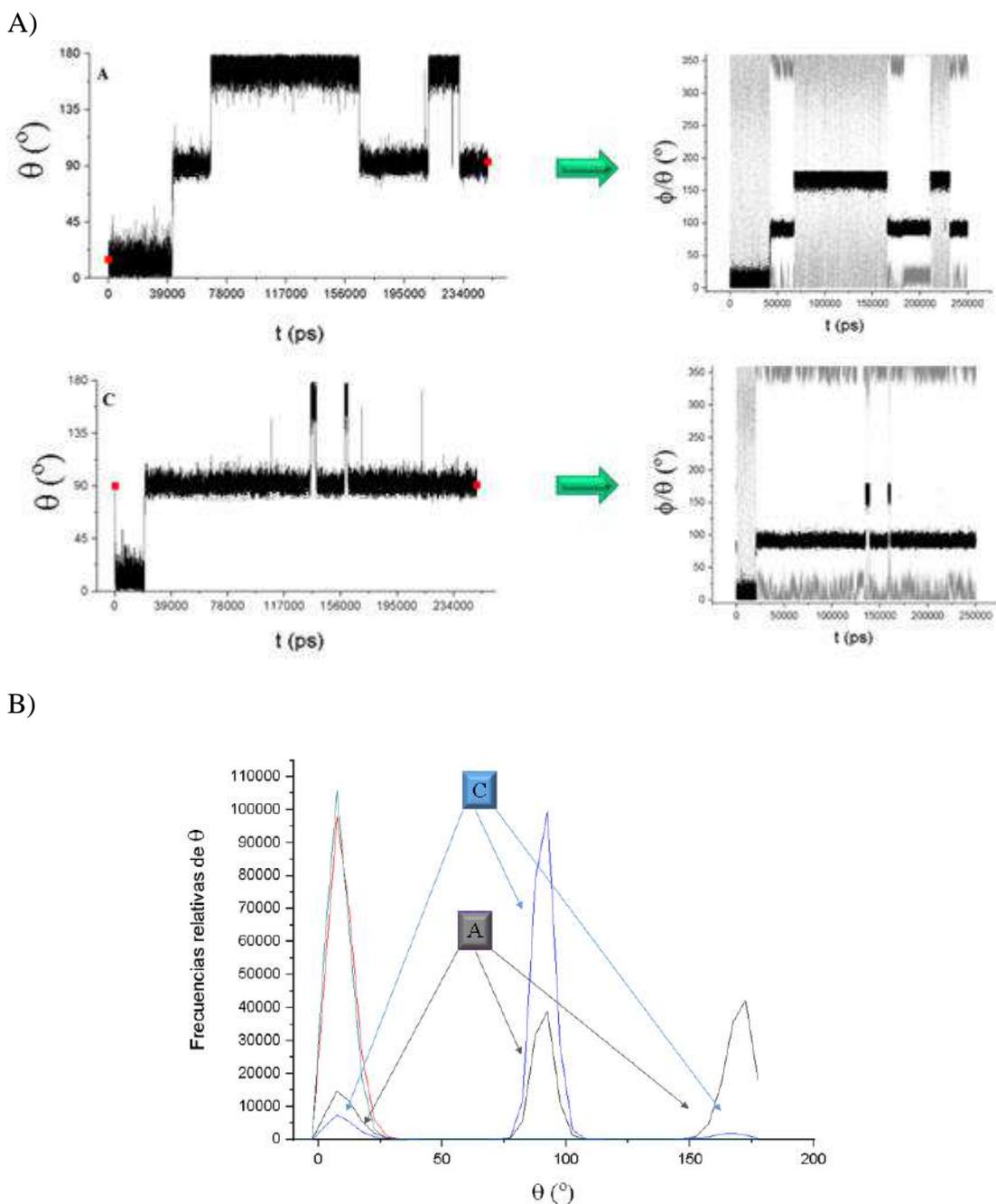


Figura 3.25. (A) Gráficas de Cremer-Pople (ángulo θ en negro y ángulo ϕ en gris) para los anillos A (arriba) y C (abajo) que permiten conocer el tipo de conformación predominante en los residuos del compuesto 6. (B) Gráfica de la frecuencia relativa frente a θ para todos los anillos. Los resultados proceden de la dinámica de 250 ns sin restricciones.

A la vista de estos resultados, se decidió lanzar una dinámica de 8 ns con restricciones en estos dos anillos. Debido a los problemas que se encontraron a nivel experimental (ver punto 3.4), no se pudieron medir distancias interprotónicas a partir de los espectros de NOESY. Por ello, en esta ocasión se usaron como restricciones los valores teóricos de su homólogo 3 (Figura 3.26).

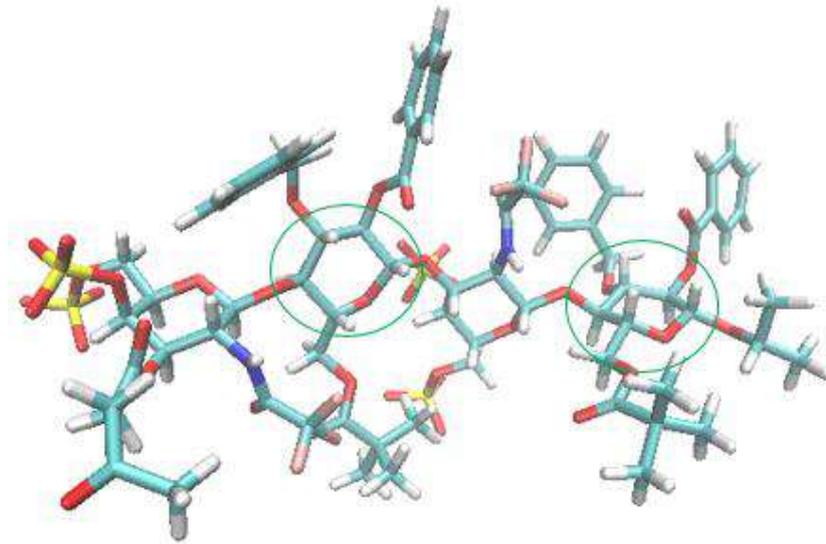


Figura 3.26. Representación de la MD-TAR de 8 ns donde se pueden ver las conformaciones de los anillos A y C en el tetrasacárido 6.

En esta simulación, se pudo apreciar que ambas glucosas corrigieron su distorsión, adquiriendo un valor de ángulo θ cercano a 0° , deduciéndose de estos resultados que todos los anillos adoptan una conformación canónica tipo silla 4C_1 (Figura 3.27).

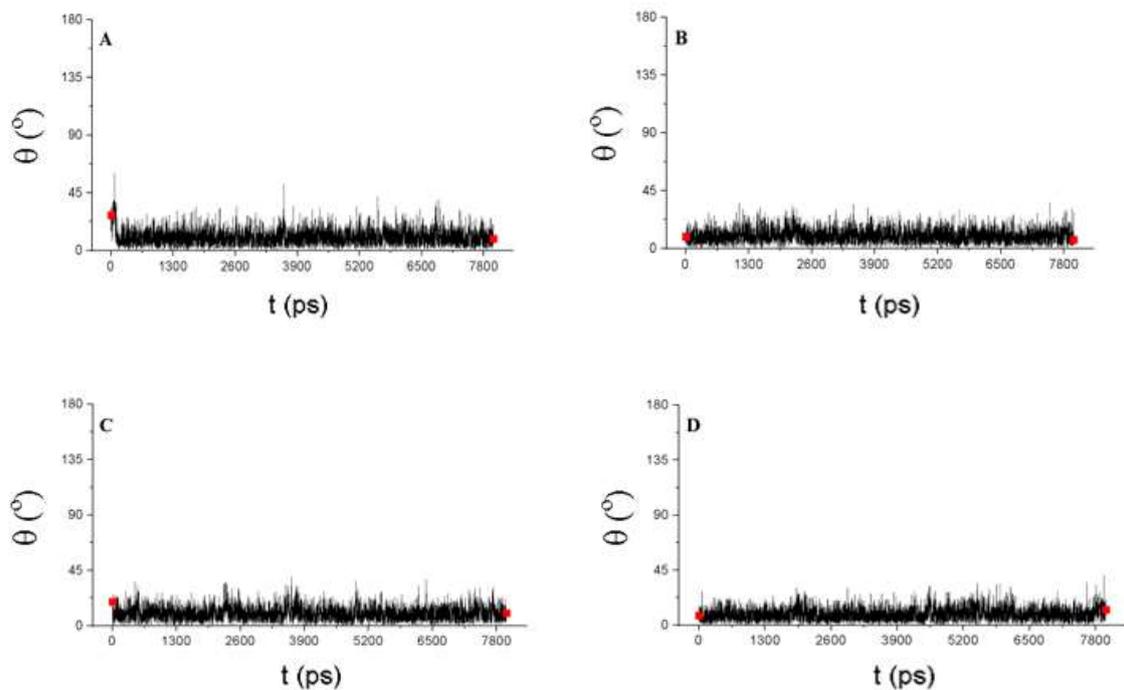
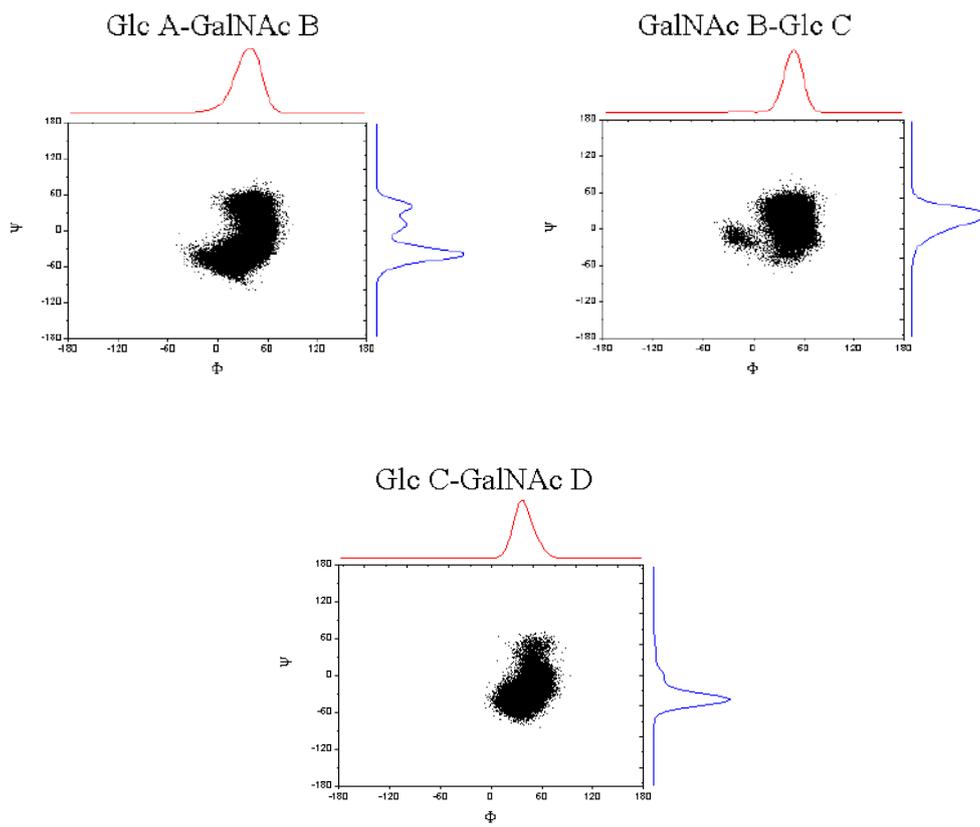


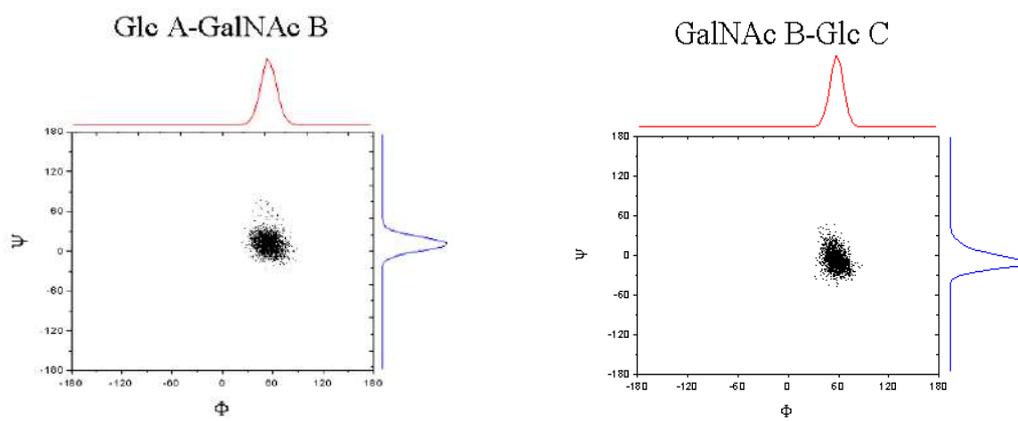
Figura 3.27. Gráficas de Cremer-Popple para todos los anillos del tetrasacárido 6 de la dinámica molecular con restricciones.

Sobre las coordenadas de ambas simulaciones (250 ns y 8 ns, sin y con restricciones, respectivamente) se realizaron los cálculos de los ángulos diedros para poder estudiar cómo afectan a los enlaces interglucosídicos la presencia de tantos grupos protectores. En la Figura 28 se han representado los mapas (ψ, ϕ) correspondientes.

A)



B)



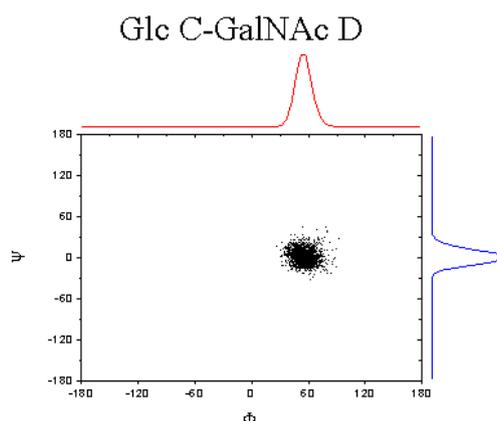


Figura 3.28. (a) Trayectorias (ϕ , ψ) obtenidas de la dinámica molecular (sin restricciones) de 250 ns para los tres enlaces interglucosídicos del compuesto **6** (b) abajo, la trayectoria (ϕ , ψ) analizada de la dinámica molecular de 8 ns (con restricciones) para las tres uniones.

Si comparamos ambas simulaciones (Figuras 3.28.A y 3.28.B), en términos generales, los enlaces interglucosídicos en la dinámica de 250 ns muestran mayor flexibilidad con respecto a los de la MD-TAR debido a la ausencia de restricciones. En todos los casos, las uniones entre los azúcares mantienen una conformación *syn- ψ* , oscilando el valor de ψ entre -60° y 60° para la simulación de 250 ns, y con una población más concentrada en valores cercanos a 0° para la MD-TAR. Además en ambas simulaciones, los enlaces cumplen el efecto exoanomérico.

Como se ha mencionado antes, no se pudieron determinar ni las distancias ni las constantes de acoplamiento experimentales debido a que no fue posible disolver el compuesto en el buffer acuoso utilizado para los experimentos de RMN. Por ello, a continuación se discutirán únicamente los resultados obtenidos teóricamente tanto para las dinámicas de 250 ns como para las TAR de 8 ns (Tablas 3.11 y 3.12).

Distancias	6 MD 250 ns	6 MD-TAR 8 ns
A H1H3	3.9	2.8
A H1H5	3.3	2.5
A H3H5	3.9	2.7
Interg.BH1-AH4	2.4	2.6
B H1H3	2.5	2.7
B H1H5	2.5	2.6
B H3H5	2.4	2.5
Interg.CH1-BH3	2.4	2.4
C H1H3	3.9	2.8
C H1H5	2.6	2.5
C H3H5	4.0	2.6
Interg.DH1-CH4	2.3	2.5

Tabla 3.11. Distancias teóricas (MD con y sin restricciones) para el tetrasacárido **6** medidas a 300 K.

Distancias	6	6
	MD 250 ns	MD-TAR 8 ns
D H1H3	2.5	2.7
D H1H5	2.5	2.5
D H3H5	2.5	2.5

Tabla 3.11. Continuación.

Tal y como se ha visto hasta ahora, las distancias interprotónicas teóricas calculadas en la dinámica sin restricciones confirman las distorsiones en las glucosas A y C, con valores promedios más altos de 2.5 Å. Estas distancias más largas se ven corregidas al incluir las restricciones en la simulación de 8 ns. Para las constantes de acoplamiento, se observa el mismo comportamiento, pudiéndose confirmar únicamente la conformación canónica 4C_1 para los anillos de GalNAc B y D en la dinámica de 250 ns. Tras realizar la TAR, los anillos A y C se corrigen, adquiriendo también la conformación esperada para un residuo de Glc (Tabla 3.12):

${}^3J_{HH}$ coupling	${}^3J_{HH}$ MD 250 ns	${}^3J_{HH}$ MD-TAR 8 ns
A1A2	3.2	7.8
A2A3	3.6	9.6
A3A4	3.6	8.9
A4A5	3.6	9.6
B1B2	8.4	8.2
B2B3	10.6	10.4
B3B4	2.7	2.9
B4B5	1.1	1.1
C1C2	3.6	7.8
C2C3	2.8	9.7
C3C4	3.3	9.3
C4C5	3.7	9.7
D1D2	8.5	8.3
D2D3	10.6	10.5
D3D4	3.1	3.0
D4D5	1.4	1.0

Tabla 3.12. ${}^3J_{HH}$ teóricas a 300 K (MD con y sin restricciones) del tetrasacárido 6.

3.4 Análisis experimental mediante RMN

En paralelo a los cálculos de MD descritos anteriormente, se procedió al estudio de la estructura tridimensional de todos los oligosacáridos mediante técnicas de RMN.

Primeramente, se realizaron experimentos monodimensionales de ^1H -RMN^[2]. En la Figura 3.29 se muestran los espectros correspondientes a la serie pentasacarídica en un buffer 1x PBS en D_2O a pH 7.8.

Cuando se aumenta la hidrofobicidad de los compuestos, al incrementar el número de grupos protectores, se observa un leve ensanchamiento de las señales en el espectro que se debe al incremento en el tiempo de correlación. La utilización de este tampón para la disolución de los ligandos se debe a que el análisis experimental de los complejos con proteína iba a requerir estas condiciones. Por tanto, el medio crítico para facilitar la comparación entre ambas condiciones, que sólo se diferenciarán en la presencia o ausencia de receptor en una concentración muy baja (Figura 3.29).

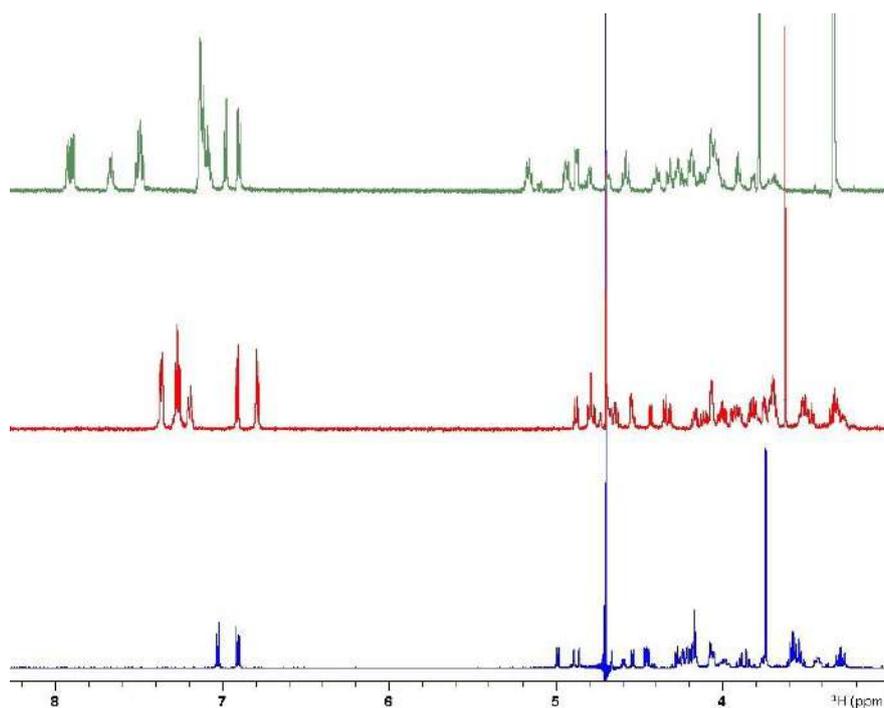


Figura 3.29. Espectros monodimensionales de ^1H -RMN para la serie pentasacarídica. De arriba a abajo, pentasacárido **3** (persustituido), **2** (dibencilado) y **1** (desprotegido) a 1.5 mM y 300 K (600 MHz).

Estos experimentos monodimensionales de ^1H -RMN se llevaron a cabo también para la serie tetrasacarídica (Figura 3.30). En el caso de los ligandos **4** y **5**, el empleo del buffer provocó el ensanchamiento y solapamiento de las señales dificultando su análisis. Por lo tanto, se llevaron a cabo los experimentos disolviendo los compuestos únicamente en agua pura para así obtener unos espectros de mejor resolución.

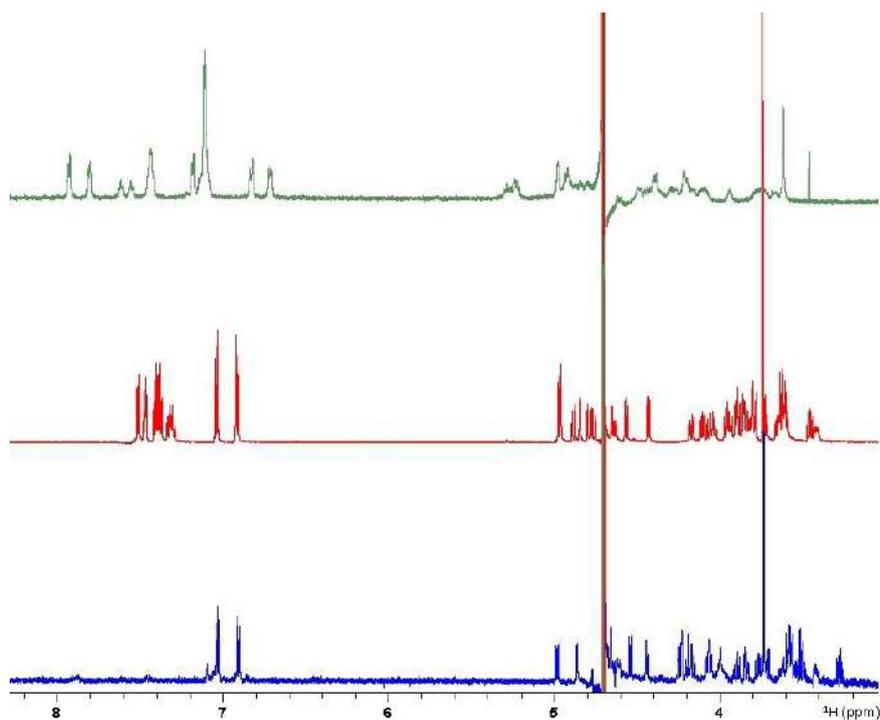


Figura 3.30. Espectros monodimensionales de ^1H -RMN para la serie tetrasacáridica. De arriba a abajo, tetrasacárido **6** (persustituido), **5** (dibencilado) y **4** (desprotegido) a 1.5 mM y 300 K (600 MHz), en 1 x PBS, pH = 7.8.

Como se observa en la Figura 3.30, al igual que en la serie anterior, el gran número de grupos protectores en la molécula **6** provocó un importante ensanchamiento de las señales. Esto puede deberse al aumento del tiempo de correlación por la viscosidad del medio y la agregación del compuesto en medio acuoso, lo que genera moléculas más grandes, con tiempos de correlación más largos que originan señales más anchas. Este hecho impidió el análisis del espectro del tetrasacárido persustituído.

Una vez asignados los ligandos **1-5**, se procedió a calcular las distancias interprotónicas H1-H3, H1-H5, H3-H5 de cada anillo y también las distancias interglucosídicas. Este cálculo se realizó mediante el análisis de los espectros NOESY a distintos tiempos de mezcla, teniendo en cuenta las áreas de cada pico de cruce ^1H - ^1H , las señales de la diagonal y los tiempos de mezcla del experimento.

En algunos casos, no fue posible analizar todos los picos de cruce relevantes debido al solapamiento de las señales, lo que dificultó su integración. Así, por ejemplo, en el espectro NOESY para el ligando **1** (Figura 3.31) no fue posible analizar las señales H3-H5. En cambio, sí pudieron integrarse los picos de cruce H1-H3 y H1-H5, que proporcionaron las correspondientes distancias interprotónicas (Tabla 3.1) que estaban de acuerdo con una conformación $^4\text{C}_1$ para los anillos de Glc y GalNAc. Por otra parte, se analizaron e integraron los picos de cruce interglucosídicos H1B-H3A, H1C-H4B y H1D-H3C que corroboraron la conformación *syn- ψ* para los enlaces glicosídicos.

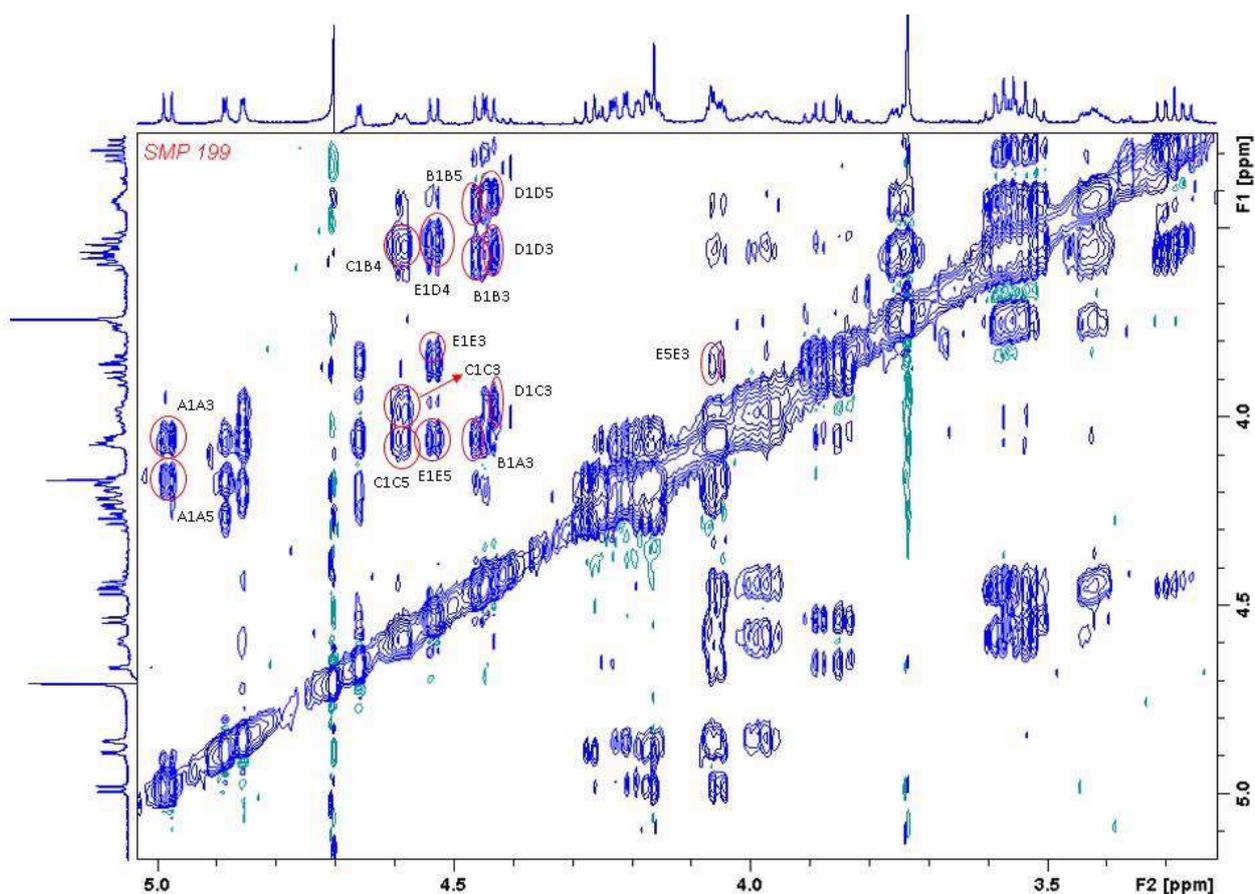


Figura 3.31. Espectro NOESY para el pentasacárido **1** a 1.5 mM, 300 K y 600 MHz.

En el compuesto **2** fue posible el análisis de la mayoría de los picos NOE exceptuando H3C-H5C y H3E-H5E cuya identificación se ve impedida por su proximidad con la diagonal. El análisis de las señales interglucosídicas H1B-H3A, H1C-H4B, H1D-H3C y H4D-H1E corroboraron la conformación *syn-ψ* de los enlaces glucosídicos, lo que está de acuerdo con los resultados teóricos (Figura 3.32 y Tabla 3.3).

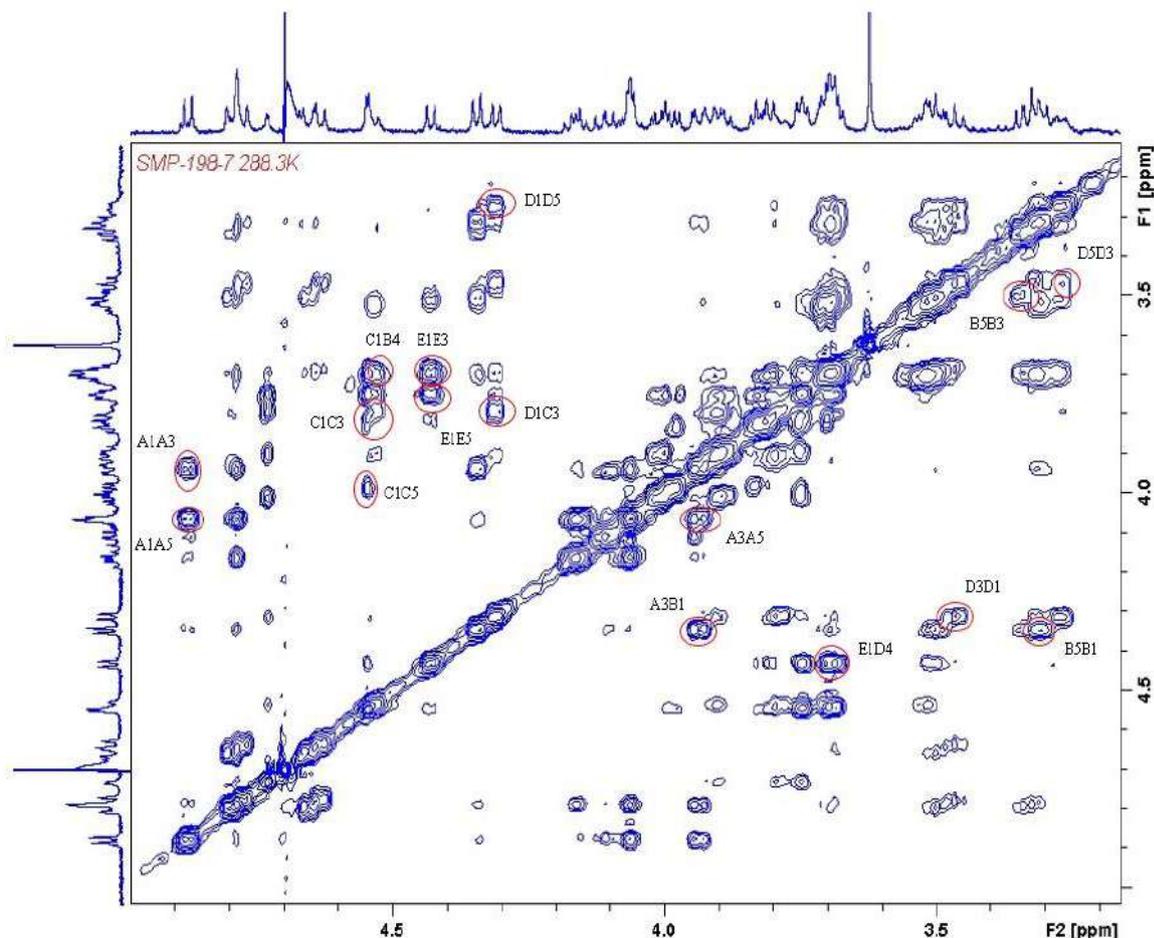


Figura 3.32. Espectro NOESY para el pentasacárido **2** a 1.5 mM, 300 K y 600 MHz.

Como se ha mencionado en el punto 3.2.3, aunque el ligando **3** pudo solubilizarse en el *buffer* PBS, los espectros NOESY que se obtuvieron a los diferentes tiempos de mezcla mostraban señales muy anchas que impidieron su evaluación y, por tanto, no se pudieron generar las curvas de crecimiento correspondientes y calcular las distancias experimentales.

Para los compuestos **4** y **5**, se pudieron identificar e integrar la mayoría de los NOEs significativos exceptuando los H3-H5 de los azúcares A y C para el primero de ellos y de los anillos A, B y D del segundo (Tablas 3.7 y 3.9). Al igual que en los compuestos pentasacarídicos, los picos H1B-H4A, H1C-H3B y H1D-H4C interglucosídicos confirmaban de nuevo la orientación *syn-ψ* de los enlaces (Figura 3.33).

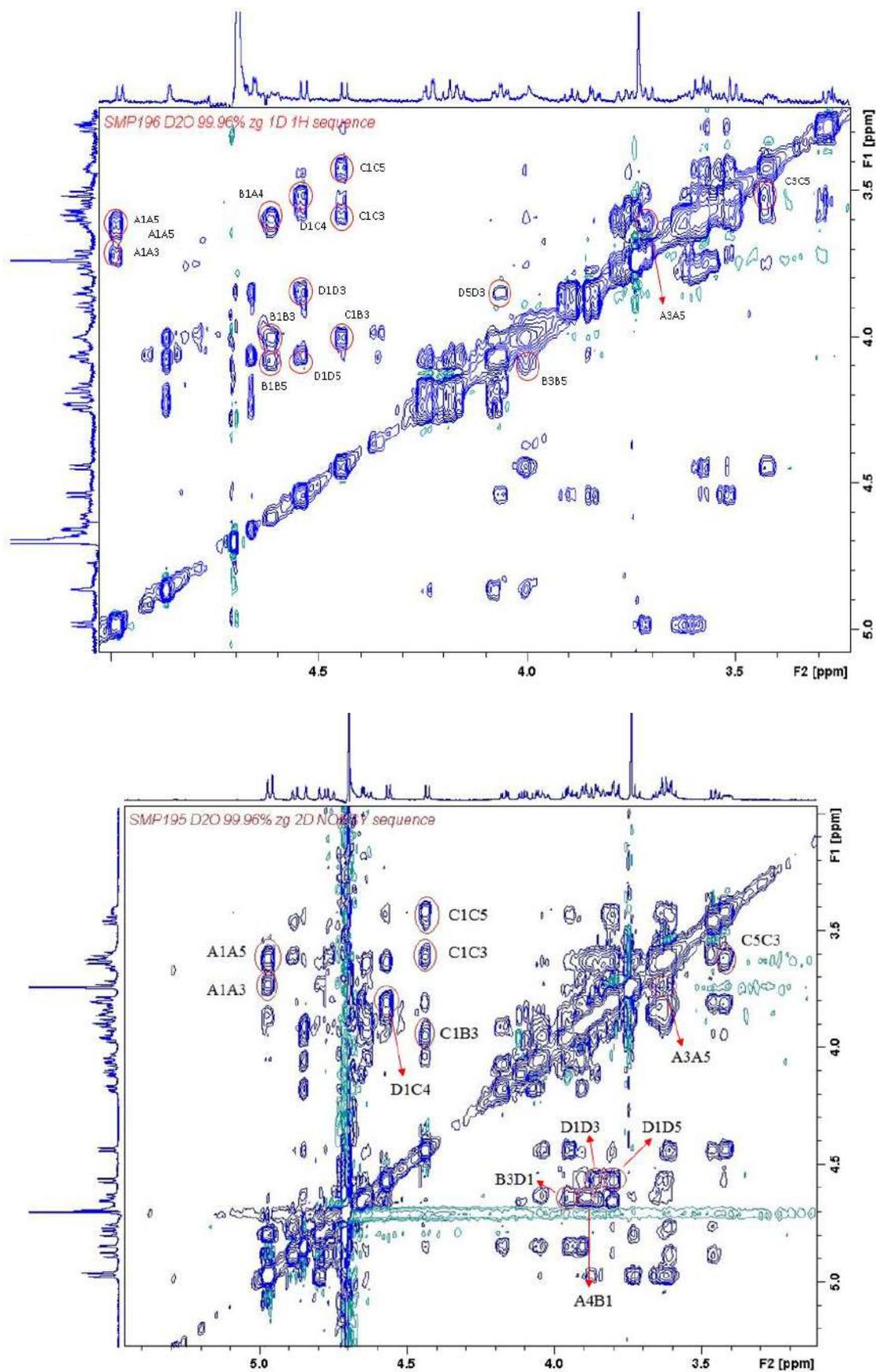


Figura 3.33. Espectro NOESY para el tetrasacárido **4** (arriba) y el tetrasacárido **5** (abajo) a 1.5 mM, 300 K y 600 MHz.

En resumen, los resultados de este capítulo indican que todos los oligosacáridos estudiados **1-6** poseen una estructura tridimensional similar, con todos los anillos en conformación de silla 4C_1 y los enlaces glicosídicos con una disposición *syn-ψ* predominante. Esta estructura no cambia de forma significativa al introducir sustituyentes voluminosos en los miméticos, como ocurre en los derivados **2-3** y **5-6**. Además, esta estructura coincide con la de los oligosacáridos de CS tipo E que contienen unidades de GlcA en lugar de residuos de Glc, por lo que nuestros datos confirman que los compuestos **1-6** pueden considerarse miméticos estructurales de las secuencias naturales.

3.5 Metodología

3.5.1 Preparación de inputs y simulaciones

Las dinámicas moleculares se llevaron a cabo en el cluster Carbohidratos perteneciente al Instituto de Investigaciones Químicas (CicCartuja) de Sevilla. El diseño individual de cada uno de los ligandos y sus diferentes grupos protectores fue llevado a cabo mediante el programa Maestro de la plataforma Schrödinger, estableciendo la correspondiente nomenclatura y numeración de sus átomos y residuos. A continuación, las cargas parciales del residuo CF_3 se generaron mediante el método RESP^[6] derivadas del potencial electrostático molecular MEP en las condiciones del procedimiento desarrollado bajo el campo de fuerzas GLYCAM06^[5]. Su geometría se ha calculado mediante el método Hartree Fock con la variante 6-31G* como conjunto de bases^[10], empleando la plataforma R.E.D 2.0 y R.E.D IV del servidor web R.E.D.

Mediante el módulo de *antechamber* y su extensión *parmchk* en *AmberTools* se crearon los archivos *prep* y *frmod* respectivamente. Estos inputs son necesarios en la ejecución del módulo *tLEAP*, para el cual se empleó la versión AMBER 12, generando los archivos de topología y coordenadas iniciales fundamentales para la simulación molecular. En este procedimiento se utilizó de nuevo el campo de fuerzas GLYCAM06 con el set de parámetros *parm99*. Todos los ligandos fueron neutralizados con iones sodio sumergidos en una caja de agua *TIP3P* como disolvente explícito^[11].

Antes de comenzar las etapas de la simulación, es necesario minimizar el sistema para relajar la energía entre el soluto y el disolvente. Los primeros pasos de minimización y las etapas de equilibrio y presión constante se realizaron con *sander.MPI* mientras que en la etapa de producción de las dinámicas se utilizó *pmemd.MPI* (*Particle Mesh Ewald Method*) en la versión de AMBER12^{[12],[13]}. Durante el primer paso se minimiza únicamente la caja de agua *TIP3P* dejando fijo el ligando; en la segunda etapa, ambos son minimizados. A continuación, se llevaron a cabo diferentes etapas consecutivas de dinámica molecular, comenzando por un aumento de la temperatura del sistema de 0 K a 300 K. Este paso de la simulación tiene una duración de 80 ps y se produce a volumen constante usando el método SHAKE.

Una vez equilibrada la temperatura a 300 K, el siguiente paso con una duración de 200 ps, consiste en alcanzar una densidad de 1g/mL a una presión constante de 1 bar. En este momento, nuestro sistema ya se encuentra en completo equilibrio y puede llevarse a cabo un último paso de mayor duración (250-500 ns total) donde se introdujeron, en los casos que fue necesario, las restricciones experimentales promediadas en el tiempo (MD-TAR) procedentes de los experimentos de NOESY (en estos casos, las duraciones de las dinámicas fueron más cortas, de 8 ns). Cabe destacar que el programa de AMBER *pmemd.MPI* se utilizó tanto en las dinámicas libres como en las que tienen restricciones. Todos los archivos finales de las trayectorias de coordenadas se almacenaron cada 0.5 ps.

3.5.2 Método de análisis de los resultados teóricos

Finalizada la dinámica, se analizaron los *outputs* de coordenadas generados cuyos datos fueron procesados con el módulo *cpptraj* de la versión AMBER12 y visualizados con el programa gráfico VMD.

El estudio de las conformaciones de cada anillo se realizó con el comando *pucker* utilizando el método de Cremer-Pople^[14]. Las constantes de acoplamiento se calcularon mediante un promedio de todos los valores a lo largo de la simulación empleando ecuaciones de Altona parametrizadas para cada enlace^[15]. Se escogió este método ya que tiene en cuenta las electronegatividades de cada sustituyente (Eq.1), pero estas ³J_{HH} teóricas pueden obtenerse también a través de la ecuación de Karplus^[16].

$${}^3J = p1 \cos^2(\phi) + p2 \cos(\phi) + p3 + \sum \lambda_i (p4 + p5 \cos^2(\epsilon_i \phi + p6 |\lambda_i|)) \quad (1)$$

siendo $\lambda_i = (X_\alpha - X_H) + p7 \sum (X_\beta - X_H)$

Para completar el estudio geométrico de nuestros ligandos, se analizaron sobre los frames de la simulación final, usando de nuevo el módulo *cpptraj*, las distancias interprotónicas (H₁-H₃, H₁-H₅, H₃-H₅) e interglucosídicas (H₁-H₄ y H₁-H₃), junto a las torsiones ϕ (H₁'-C₁'-O-C₄/C₃) y ψ (C₁'-O-C₄/C₃-H₄/H₃). Todos estos valores fueron promediados sobre las 125.000.000 estructuras de la última etapa de simulación.

3.5.3 Método de análisis de los resultados experimentales

Los experimentos de RMN han sido realizados en un equipo Bruker con un imán ASCEND III HD 600 equipado con una criosonda QCI S3 H/F-C/N-D 05-Z, previa asignación en un equipo Bruker ASCEND III 400 con sonda BBO 400S1 BBF-H-D-05 Z SP.

Con el fin de que todos los compuestos objeto de estudio pudieran ser comparados se realizaron los experimentos de RMN de los ligandos libres en condiciones similares entre sí e idénticas a la de los compuestos asociados. Los ligandos se prepararon a una concentración 1.5 mM en 250 μ L de buffer 1x PBS en D₂O, empleando tubos de RMN de 3 mm de diámetro en la zona de medida. La presencia de buffer provocó un ensanchamiento de las señales en todos los compuestos. En el caso de los ligandos **4** y **5**, esto dificultó su asignación, por lo que finalmente se asignaron los espectros de los

ligandos libres en D₂O. En los compuestos persustituidos **3** y **6**, el análisis de los espectros de RMN se vio gravemente dificultado por el solapamiento y ensanchamiento de las señales.

Las técnicas que se emplearon para la asignación fueron los experimentos bidimensionales DQF-COSY, TOCSY, NOESY y HSQC. Los experimentos homonucleares se realizaron con versiones sensibles a la fase con selección de gradientes de campo de la biblioteca del fabricante. En los experimentos TOCSY se usó para el periodo de bloqueo de espines la secuencia MLEV17 a 400 MHz y dipsi2 a 600 MHz. Los experimentos NOESY y TOCSY a 600 MHz se registraron con filtro de cero cuanto y con presaturación en la señal del agua.

Los experimentos HSQC realizados a 400 MHz se registraron con secuencias que emplean metodología de eco-antieco-TPPI y gradientes de campo para la selección de las señales en el periodo de etiquetado de frecuencias, con sensibilidad incrementada, y con edición en función de la multiplicidad de la señal (par-impar). A 600 MHz se reemplazaron los pulsos de 180 grados del canal de carbono por pulsos adiabáticos Chirp debido al mayor rango de frecuencias excitadas eficazmente.

Para la cuantificación de distancias interprotónicas, se emplearon experimentos NOESY realizados a 600 MHz, con tiempos de mezcla desde 150-200 ms hasta 800-1000 ms, siguiendo la aproximación de espines aislados (ISPA).^[17] Para calcular la velocidad inicial de crecimiento, los valores del pico de cruce se dividieron entre el pico diagonal, al mismo tiempo de mezcla (t_m). Mediante el ajuste de estos valores a la siguiente ecuación exponencial “Box Lucas” (que describe la evolución de los valores calculados anteriormente frente al tiempo):

$$I_x/I_d(t_m) = a (1 - \exp(b t_m)) \quad (2)$$

se calcula la velocidad inicial de crecimiento (σ_0) como la derivada de dicha ecuación, el producto ab , siguiendo la aproximación *Peak Amplitude Normalization for Improved Cross-relaxation* (PANIC).^{[18][19]} A continuación, empleando la relación de la ecuación 3:

$$d_x = d_{ref} (\sigma_{ref}/\sigma_x)^{1/6} \quad (3)$$

se calculan las distancias experimentales en base a las velocidades de relajación cruzada (σ) para cada pareja de protones.

Referencias Capítulo 3

- [1] M. Rawat, C. I. Gama, J. B. Matson, L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* **2008**, *130*, 2959–2961.
- [2] M. Torres Rico, S. Maza, J. L. de Paz, P. M. Nieto, *Org. Biomol. Chem.* **2021**, *19*, 5312–5326.
- [3] C. Solera, G. Macchione, S. Maza, M. M. Kayser, F. Corzana, J. L. De Paz, P. M. Nieto, *Chem. A Eur. J.* **2016**, *22*, 2356–2369.
- [4] J. L. De Paz, P. M. Nieto, *Org. Biomol. Chem.* **2016**, *14*, 3506–3509.
- [5] K. Kirschner, A. Yongye, S. Tschampel, *J Comput Chem*, **2008**, *29*, 622–655.
- [6] F. Y. Dupradeau, C. Cézard, R. Lelong, É. Stanislawiak, J. Pêcher, J. C. Delepine, P. Cieplak, *Nucleic Acids Res.* **2008**, *36*, 360–367.
- [7] J. C. Muñoz-García, F. Corzana, J. L. De Paz, J. Angulo, P. M. Nieto, *Glycobiology* **2013**, *23*, 1220–1229.
- [8] D. A. Pearlman, P. A. Kollman, *J. Mol. Biol.* **1991**, *220*, 457–479.
- [9] M. Kontoyianni, L. M. McClellan, G. S. Sokol, *J. Med. Chem.* **2004**, *47*, 558–565.
- [10] R. J. Woods, R. Chappelle, *J. Mol. Struct. (THEOCHEM)*, **2000**, *527*, 149–156.
- [11] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [12] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.
- [13] H. G. Petersen, *J. Chem. Phys.* **1995**, *103*, 3668–3679.
- [14] M. Sega, E. Autieri, F. Pederiva, *Mol. Phys.* **2011**, *109*, 141–148.
- [15] A. Navarro-Vázquez, J. C. Carlos, F. J. Sardina, J. Casanueva, E. Díez, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1680–1685.
- [16] M. Karplus, *J. Am. Chem. Soc.* **1963**, *85*, 2870–2871.
- [17] S. Macura, B. T. Farmer, L. R. Brown, *J. Magn. Reson.* **1986**, *70*, 493–499.
- [18] J. C. Muñoz-García, F. Corzana, J. L. De Paz, J. Angulo, P. M. Nieto, *Glycobiology* **2013**, *23*, 1220–1229.
- [19] H. Hu, K. Krishnamurthy, *J. Magn. Reson.* **2006**, *182*, 173–177.

Capítulo 4

*Estudios de estructura e interacción
experimental y teórica de los miméticos
de CS-E con PTN*

4.1 Introducción

Como se ha descrito previamente en el **Capítulo 1**, los GAGs interactúan con una gran variedad de proteínas debido a su amplia diversidad estructural, regulando así la actividad de muchas de ellas, permitiéndoles ejercer diversas funciones biológicas.^{[1],[2]} Los factores de crecimiento son un ejemplo de estas proteínas que, asociadas a la matriz extracelular, desempeñan labores importantes como neuromoduladores con múltiples funciones sobre el desarrollo neuronal.

En este capítulo vamos a centrarnos en el estudio de una de ellas, la Pleiotrofina (PTN), una potente citoquina mitogénica y angiogénica capaz de modular, además, un amplio espectro de procesos fisiológicos relacionados con la proliferación celular. Las diversas actividades de la PTN están ligadas a la capacidad que tiene para unirse a una gran variedad de receptores de membrana siendo éstos, en una gran proporción, proteoglicanos. Un ejemplo son los receptores de tipo proteína tirosina fosfatasa (PTP) que presentan cadenas de Condroitín Sulfato (CS) de tipo C y E principalmente.^{[3],[4]} La PTN es un ávido aglutinante de GAGs y sus interacciones con estos polisacáridos son cruciales para el desarrollo de sus actividades biológicas. El patrón de sulfatación de los GAGs es un determinante crítico en sus interacciones con esta citoquina,^[5] observándose una mayor afinidad por aquellos que se encuentran altamente sulfatados. Sobre estas premisas, y basándonos en estudios de interacción que se han realizado con CS donde se ha demostrado que la afinidad de la PTN por CS de tipo A es menor que en el caso del CS de tipo E,^[6] se consideró interesante realizar el estudio de interacción de nuestros miméticos de CS-E con esta citoquina.

Este estudio de interacción se llevó a cabo, tanto a nivel experimental como teórico, para las dos series de complejos formados con los ligandos **1-6** del **Capítulo 3**, utilizando las estructuras y los resultados obtenidos del análisis estructural de los oligosacáridos libres. De esta forma, se podría determinar si la formación del complejo genera o no distorsiones en el azúcar, e incluso intentar obtener una idea del comportamiento de la PTN cuando se encuentra enlazada. Una vez confirmada la interacción entre esta proteína y los miméticos de CS-E, se intentará reproducir una estructura o un modelo tridimensional del complejo mediante cálculos de dinámica molecular tomando como referencia los resultados expuestos en el **Capítulo 3**, donde la estructura de los seis ligandos mostró conformaciones canónicas 4C_1 en todos sus anillos, con uniones interglucosídicas *syn-ψ/exo-φ*. El análisis conformacional de los oligosacáridos se llevó a cabo de manera similar al procedimiento descrito para su forma libre, aunque con algunas variaciones en la preparación de los inputs, que se detallarán en la sección **4.5** de este capítulo. Estos estudios de interacción complementan los realizados con la citoquina MK en nuestro grupo de investigación, que se han llevado a cabo utilizando otro tipo de derivados GAGs.^[7]

4.2 Estructura de la Pleiotrofina

Tal y como se ha descrito en el **Capítulo 1**, la PTN se compone de dos dominios de repetición de pseudotrombospondina tipo 1 (TSR), conocidos como el dominio N-terminal (NTD) y el dominio C-terminal (CTD), que se encuentran flanqueados por dos terminales no estructurados. Ambas zonas están altamente enriquecidas por aminoácidos básicos, como Arg52 y Lys54 en el NTD, Lys60 y Lys61 en la zona enlazante entre ambos sitios de unión, denominada bisagra, y Lys84, Arg86, Lys107 y Lys68, Lys91 y Arg92 de los grupos 1 y 2, respectivamente, que conforman el CTD.^[5] Se sabe además que, de los dos dominios que presenta, el extremo C-terminal es el que tiene mayor afinidad por azúcares.^[6] Por lo tanto, se consideró interesante conocer, desde el punto de vista teórico, el mecanismo de reconocimiento con los ligandos **1-6**, y cómo se comportan estructuralmente tras esta unión. De las diez estructuras obtenidas por RMN que se conocen de la PTN, se escogieron tres conformeros con diferentes grados de plegamiento, realizando los diferentes estudios teóricos con cada una de esas conformaciones representativas de la proteína, para cada ligando de ambas series, tetra y pentasacáridica. La PTN se puede clasificar como una proteína con regiones intrínsecamente desestructuradas denominadas IDR (*Intrinsically Disordered Region*) pudiendo no ser capaces de tener el núcleo hidrofóbico que constituye el dominio bien organizado, debido a la carencia, o en este caso, a la baja proporción de aminoácidos hidrofóbicos voluminosos.^{[8],[7]} En este caso, la región central entre ambos dominios TSR, denominada *hinge* (o bisagra), es la que se encuentra sin estructura. Las regiones estructuradas lo son gracias a puentes disulfuro entre las láminas beta que constituyen los dos dominios NTD y CTD (ver Figura 4.1).

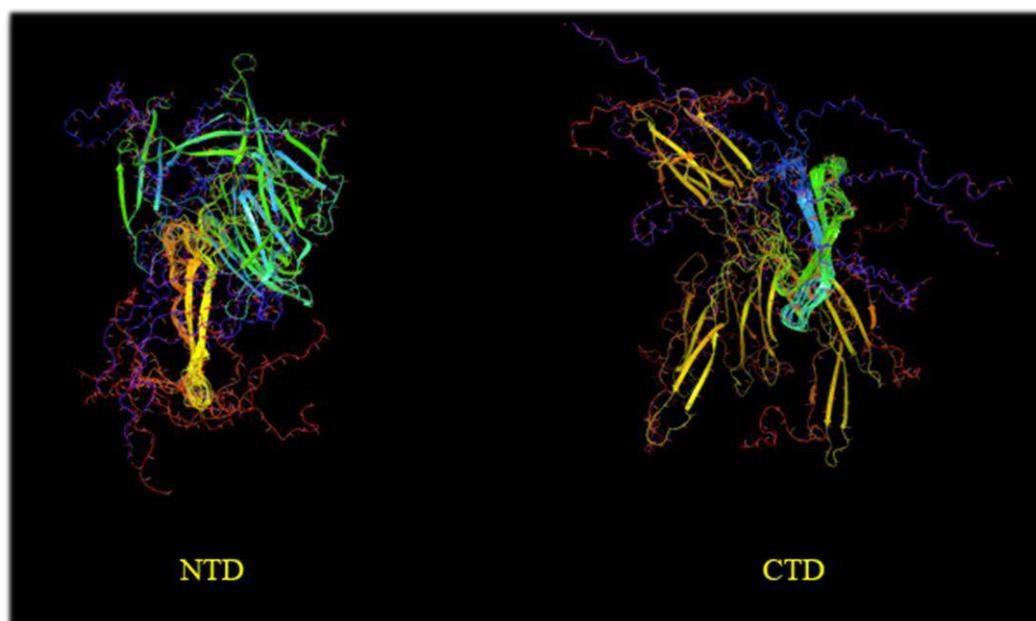
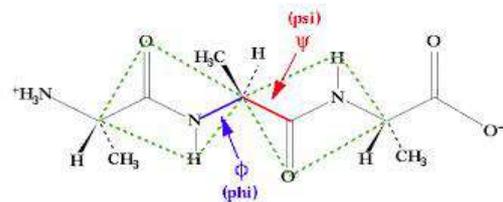
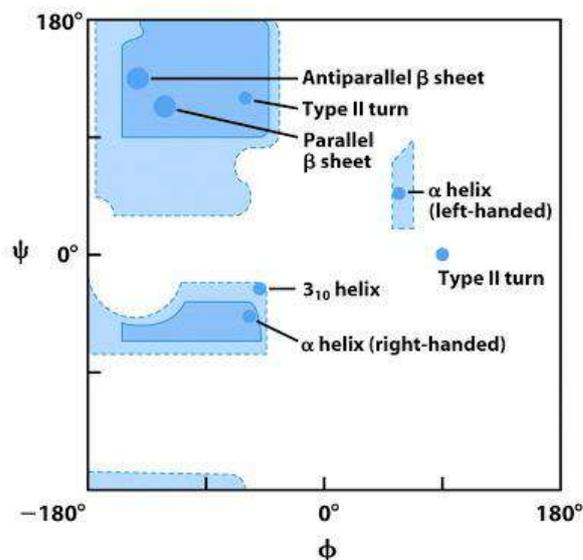


Figura 4.1. Representación tridimensional de la superposición alternativa de los dos dominios de lámina β de la PTN en las 10 estructuras del pdb.

Como se sabe, el enlace peptídico introduce las principales restricciones que limitan la conformación de un polipéptido, contribuyendo de esta manera a la estabilización de proteínas. Los carbonos α de los distintos aminoácidos de una proteína forman parte de dos planos peptídicos. La orientación relativa de estos dos planos está determinada por los ángulos ϕ y ψ (ver Figura 4.2 a). El primero de ellos mide el giro en torno al enlace que une el carbono α de un residuo con el N, mientras que el segundo lo hace con el C carbonílico del mismo residuo. La representación de estos ángulos en un diagrama bidimensional se denomina mapa de Ramachandran. Mediante estos gráficos se representan las regiones energéticamente permitidas para estos ángulos ψ/ϕ del esqueleto de la PTN.^[9] Las zonas más favorables son aquellas que presentan valores negativos de ϕ y una pequeña región con valores positivos del mismo. Además, si dos o más aminoácidos consecutivos adoptan ángulos de ψ y ϕ cercanos a $+130^\circ$ y -140° , respectivamente, aparecen láminas β , que es lo que encontramos en la estructura de la PTN. Sobre los mapas de la Figura 4.2, correspondientes a las tres conformaciones seleccionadas de la proteína, se indican dónde aparecen los residuos de prolina (Pro) y de glicina (Gly). Los primeros aportan rigidez al polipéptido mientras que los residuos de glicina son los responsables de su flexibilidad.^[10]

Como hemos indicado, en la Figura 4.2 se han representado los diferentes mapas para las tres conformaciones escogidas de la PTN, con la finalidad de conocer qué proporción de residuos se encuentran formando parte de su estructura secundaria y/o formando ángulos forzados, lo que nos indicaría que esa conformación no es adecuada. Estas estructuras han sido extraídas de las diez disponibles en el archivo de la base de datos Protein Data Bank (pdb:2N6F), asignadas con los códigos 4, 3 y 8 (PTN4, conformación plegada; PTN3, conformación abierta; y PTN8, conformación bisagra).

a) Diagrama de Ramachandran genérico



d) Conformación bisagra (PTN8)

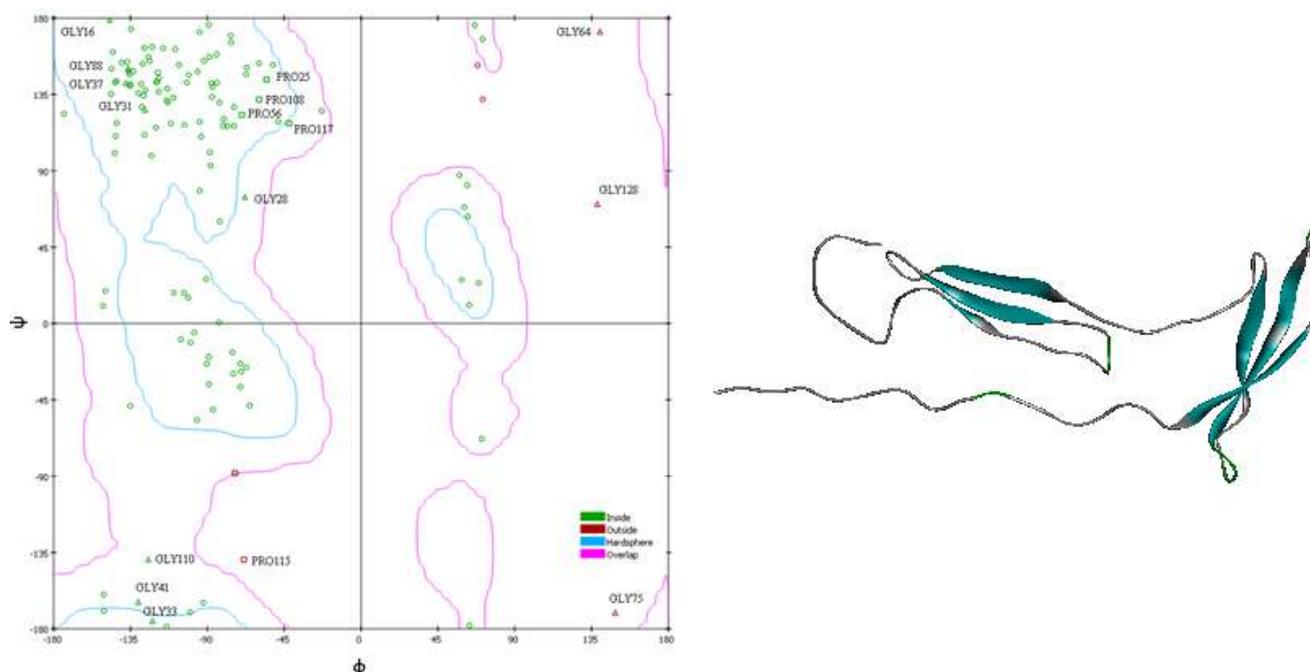


Figura 4.2. A la izquierda, mapas de Ramachandran genérico (a) e individuales (b-d) de cada conformación. A la derecha (b-d), las estructuras tridimensionales de la PTN para los tres conformémeros seleccionados.

Como se puede observar, la mayor parte de los residuos de la proteína se encuentran dentro de las zonas permitidas del diagrama, especialmente en las regiones que corresponden a las zonas de hélice α y lámina β , encontrándose en ésta última una mayor proporción de residuos. Por el contrario, la presencia de aminoácidos en las zonas prohibidas del mapa es minoritaria en el caso de los tres conformémeros, por lo que podemos concluir que estas tres estructuras son adecuadas para utilizarlas como modelo en los cálculos moleculares.

A continuación, se describirán y discutirán los resultados de interacción experimentales obtenidos a partir del empleo de diferentes técnicas de RMN y Polarización de Fluorescencia para cada uno de los complejos formados entre la PTN y los oligosacáridos **1-6**. Seguidamente, estos datos se contrastarán con los métodos de interacción teóricos de *docking*, dinámica molecular y MMPBSA/GBSA, usando las conformaciones de la PTN representadas en la Figura 4.2. Se intentará, además, reproducir una estructura tridimensional de cada complejo de acuerdo con estos resultados, analizándose las diferentes conformaciones que pueda adoptar cada uno de los ligandos.

4.3 Estudio experimental

Diversas líneas de investigación centradas en estudiar las afinidades GAG-PTN sugieren que las interacciones entre CS y estas citoquinas están reguladas por tres parámetros: 1) longitud de la cadena, 2) cantidad de estructuras CS sobresulfatadas (preferiblemente CS-

E y CS-A) y 3) la multivalencia de la interacción. Esto sugiere que la heterogeneidad estructural del CS juega un papel importante en la transducción de las señales de la PTN, y, en concreto, el CS de tipo E favorece, por ejemplo, el crecimiento de neuritas^[3].

Teniendo en cuenta esta información, se consideró interesante estudiar las interacciones entre PTN y los oligosacáridos **1-6**, que contienen un patrón de sulfatación relacionado con el CS-E. Se decidió realizar primero el estudio de interacción utilizando las técnicas más usadas en RMN, véase NOESY transferido y STD (*Saturation Transfer Difference*). Además, a partir del experimento con proteína marcada ¹⁵N-HSQC se puede tener información del área de interacción de la proteína y de su afinidad por este tipo de miméticos de CS-E. Por último, para cuantificar la interacción de todos estos ligandos con la PTN se llevaron a cabo diferentes ensayos de competición a través de la técnica de Polarización de Fluorescencia.

4.3.1 Experimento de NOESY transferido

El análisis de la conformación del enlace glicosídico se realiza mediante experimentos NOESY. Al no existir acoplamiento ¹H-¹H entre los anillos, para el análisis conformacional se debe recurrir a una medida de efecto NOE entre protones de los distintos anillos, complementado por una búsqueda por mecánica molecular de las geometrías conformacionales de mínima energía que se detallarán más adelante. El primer paso del análisis, previamente hecho en el grupo, consistió en la asignación completa de la mayoría de las señales del espectro ¹H-RMN utilizando en algunos casos una combinación de experimentos estándar COSY,^{[11],[12]} TOCSY^{[13],[14]} y HSQC^{[15],[16]}. Sobre estos espectros se midieron las constantes de acoplamiento que, al compararlas con los datos obtenidos en su estado libre, no indicaron cambios significativos. Para la extracción de restricciones experimentales de los complejos (que más adelante se utilizarán en las MD-TAR), se registraron experimentos de NOESY transferido a diferentes tiempos de mezcla en presencia de 20 μM de proteína, en un buffer 1x PBS a pH = 7.8 y a una temperatura de 288 K. De esta forma se obtuvieron, a partir de las velocidades de relajación cruzada σ_{NOE} , las distancias experimentales. Para ello, se construyeron las curvas de crecimiento NOE para los protones de mayor interés, representando el cociente entre el pico de cruce y la diagonal frente al tiempo de mezcla. De su ajuste a una ecuación exponencial y por extrapolación de esta curva a tiempo cero se calculó el σ_{NOE} . Basándonos en la aproximación de spines aislados,^[17] todas las distancias de interés fueron calculadas como $\langle r^{-6} \rangle^{-1/6}$ a partir de sus respectivos σ_{NOE} y una distancia de referencia conocida que se detallará en la sección 4.5 de este capítulo. A continuación, se mostrarán los resultados de cada uno de los complejos comparando las parejas de homólogos de cada serie.

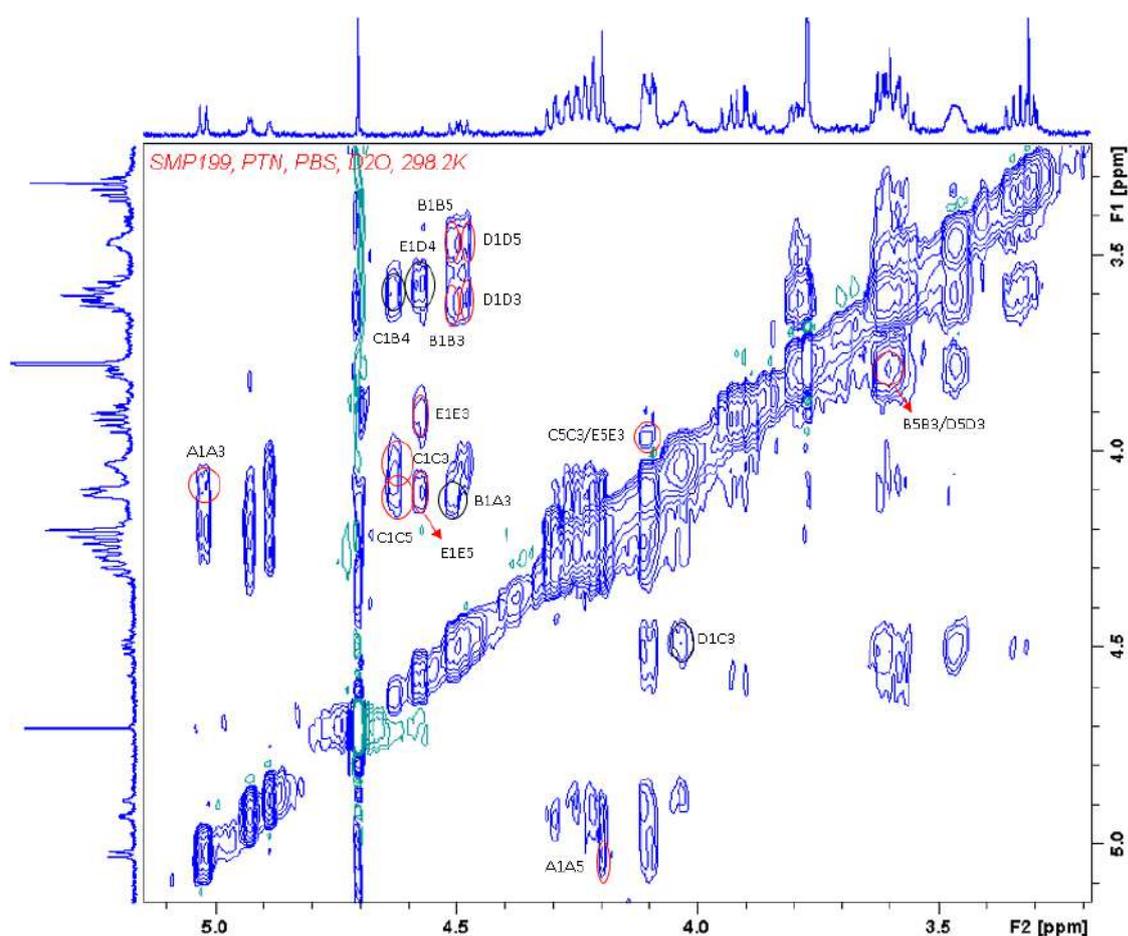
4.3.1.1 Comparativa entre los complejos 1-PTN y 4-PTN

En primer lugar, se van a exponer los análisis de los complejos que presentan el ligando completamente desprotegido tanto de la serie pentasacáridica (**1-PTN**) como de la tetrasacáridica (**4-PTN**). En la Figura 4.3 se han representado los espectros NOESY

transferido obtenidos, observándose que, al ser del mismo signo los picos de cruce y los de la diagonal, no se puede deducir de este hecho que haya tenido lugar la formación del complejo. Sin embargo, el apreciable ensanchamiento de las señales debido al cambio en el tiempo de correlación del ligando sí indica su complejación. En cada uno de los espectros se han señalado en color rojo los protones intrarresiduales y en color azul los interresiduales que se han podido identificar; de su evolución frente al tiempo de mezcla se extraen las distancias interprotónicas experimentales (ver sección 4.5 del capítulo). Como referencia, se considera que el enlace es *syn-ψ/exo-φ* cuando las distancias interglucosídicas (H1'-H4 y H1'-H3 para las uniones GalNAc-Glc y Glc-GalNAc respectivamente) toman valores menores de 2.5 Å, mientras que, para cualquier disposición anti, la distancia debe ser mayor de 3.1 Å.

En ambos espectros de NOESY transferido se pudo confirmar la presencia de los NOE interresiduales H1'-H4 y H1'-H3 que corresponden a una geometría *syn-ψ* de los enlaces, mientras que las conformaciones de tipo silla 4C_1 de los residuos de los carbohidratos fueron corroboradas por la presencia de los NOE intrarresiduales H1-H3, H1-H5 y H3-H5. Estas conclusiones fueron confirmadas paralelamente con el análisis de sus constantes de acoplamiento.

A)



B)

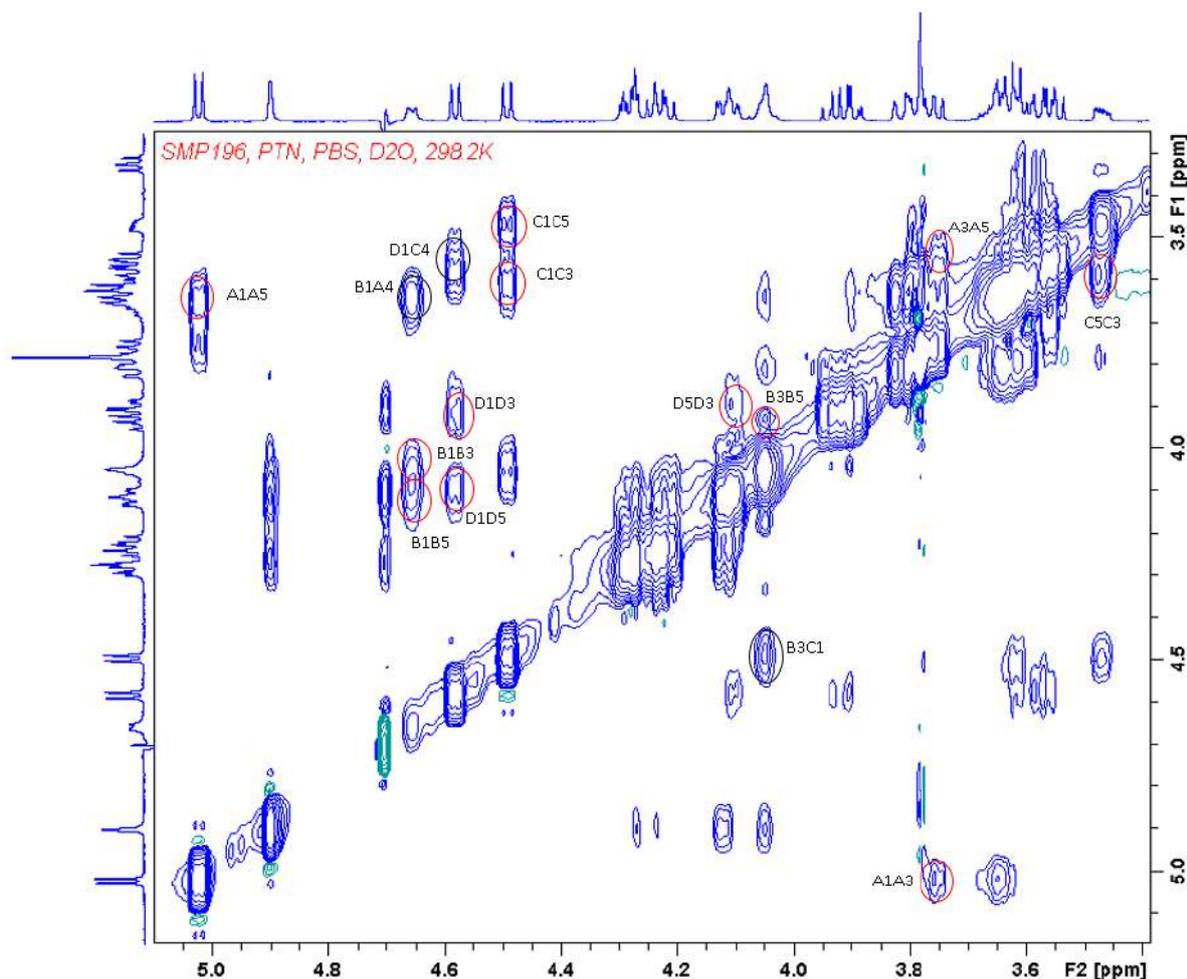


Figura 4.3. Espectros de NOESY transferido del complejo **1-PTN** (A) y del complejo **4-PTN** (B), realizados a 600 MHz y 300K (tiempo de mezcla 500 ms). Se asignan los picos de cruce más relevantes.

Como se ha mencionado antes, a partir del cociente entre el valor de la integral del pico de cruce y el promedio de la diagonal, se pudieron construir las curvas de crecimiento de NOE, extrayéndose los distintos valores de σ_{NOE} , y a partir de estos, las correspondientes distancias experimentales. En las Tablas 4.1 y 4.2 se han representado estos valores que nos indican la presencia de un único conformero *exo- ϕ* y que a su vez podrían estar presentes las dos conformaciones *syn- ψ* muy similares en ambos complejos.

Las distancias medidas de los contactos NOE entre los protones H1-H3, H1-H5 y H3-H5 intrarresiduales mostraron ligeras diferencias entre **1-PTN** y **4-PTN**, siendo para el primero algunas de ellas ligeramente más altas, posiblemente por el solapamiento de los picos de cruce, lo que dificultó su integración. A pesar de ello, se pudo establecer que estos valores indicaban conformaciones ${}^4\text{C}_1$ en todos los anillos, que ya habían descrito las constantes de acoplamiento de los experimentos monodimensionales (Figura A.1 del Anexo).

$^3J_{HH}$	1-PTN	Distancias	1-PTN
A1A2	8.8	A H1H3	2.6
A2A3		A H1H5	2.6
A3A4	2.6	A H3H5	2.4
A4A5		Interg.BH1-AH3	2.2
B1B2	8.3	B H1H3	2.5
B2B3	8.7	B H1H5	2.8
B3B4	9.1	B H3H5	3.1
B4B5	9.7	Interg.CH1-BH4	2.6
C1C2	8.9	C H1H3	2.2
C2C3	10.4	C H1H5	
C3C4		C H3H5	2.6
C4C5		Interg.DH1-CH3	2.6
D1D2	8.3	D H1H3	3.0
D2D3	8.4	D H1H5	2.3
D3D4	9.1	D H3H5	
D4D5	9.1	Interg.EH1-DH4	2.7
E1E2	8.9	E H1H3	3.1
E2E3	11	E H1H5	2.2
E3E4	2.6	E H3H5	
E4E5			

Tabla 4.1. Constantes de acoplamiento y distancias experimentales para el ligando **1** en presencia de PTN medidas a 298K.

$^3J_{HH}$	4-PTN	Distancias	4-PTN
A1A2	7.9	A H1H3	2.4
A2A3	9.3	A H1H5	2.8
A3A4	9.2	A H3H5	
A4A5		Interg.BH1-AH4	2.0
B1B2		B H1H3	2.5
B2B3		B H1H5	2.9
B3B4		B H3H5	2.6
B4B5		Interg.CH1-BH3	2.3
C1C2	7.9	C H1H3	2.8
C2C3	9.7	C H1H5	2.5
C3C4		C H3H5	
C4C5		Interg.DH1-CH4	2.5

Tabla 4.2. Constantes de acoplamiento y distancias experimentales para el ligando **4** en presencia de PTN medidas a 298K.

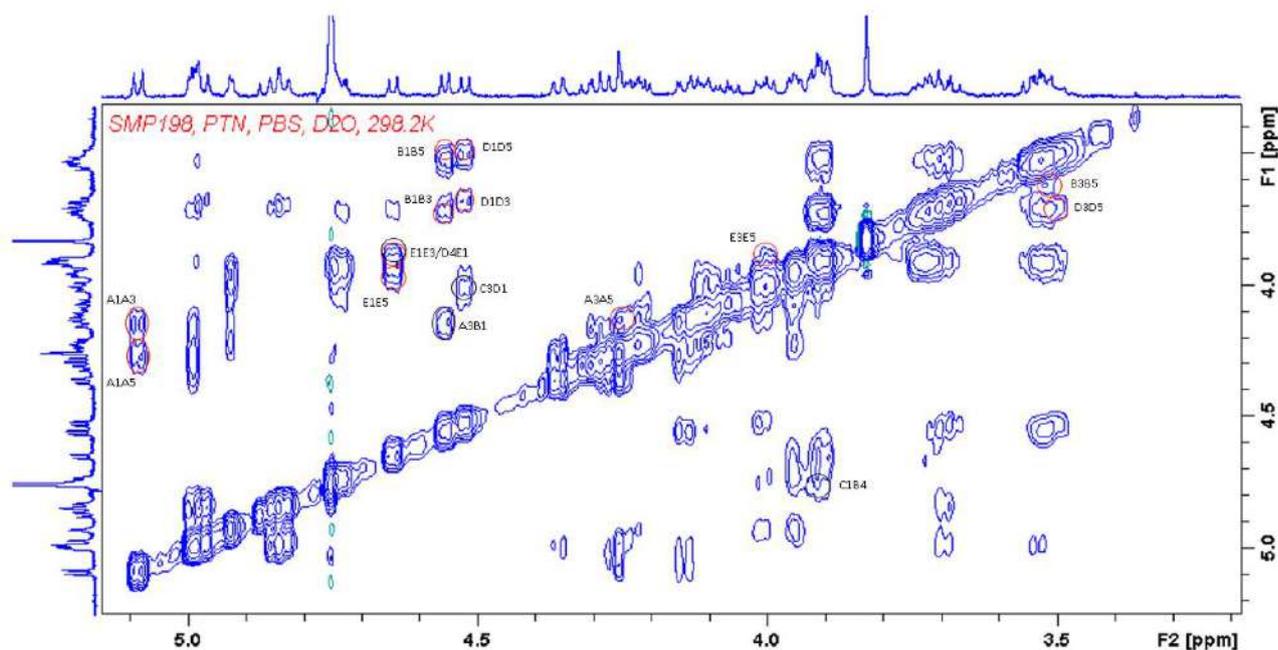
$^3J_{HH}$	4-PTN	Distancias	4-PTN
D1D2	8.5	D H1H3	2.7
D2D3	10.7	D H1H5	2.6
D3D4	2.6	D H3H5	
D4D5			

Tabla 4.2. Continuación.

4.3.1.2 Comparativa entre los complejos 2-PTN y 5-PTN

Para la asignación de los complejos con los ligandos dibencilados **2-PTN** y **5-PTN** se utilizó la misma metodología anterior, aunque en este caso los experimentos de NOESY *transfer* fueron más difíciles de analizar debido a la alta señal de agua que se observaba en los espectros. Esto pudo deberse a varios motivos, como la preparación de la muestra o el tiempo de espera hasta que se midió, ya que nuestros compuestos son higroscópicos. Por este motivo, fue necesario recurrir a una presaturación en la secuencia de pulsos como método de supresión de disolvente. Aunque se consiguió rebajar la intensidad de esta señal, fue inevitable que las señales contiguas a ella se viesan afectadas, impidiendo su medición. La Figura 4.4 muestra las regiones del espectro en las que se han señalado los picos de cruce descriptivos de ambas estructuras: en rojo las correlaciones intrarresiduales y en negro las interresiduales. De nuevo, la observación de los picos de cruce H1-H3, H1-H5 y H3-H5 sustentan la conformación de tipo silla 4C_1 y los picos interresiduales H1'-H3 y H1'-H4 indican la geometría *syn* del enlace interglicosídico.

A)



B)

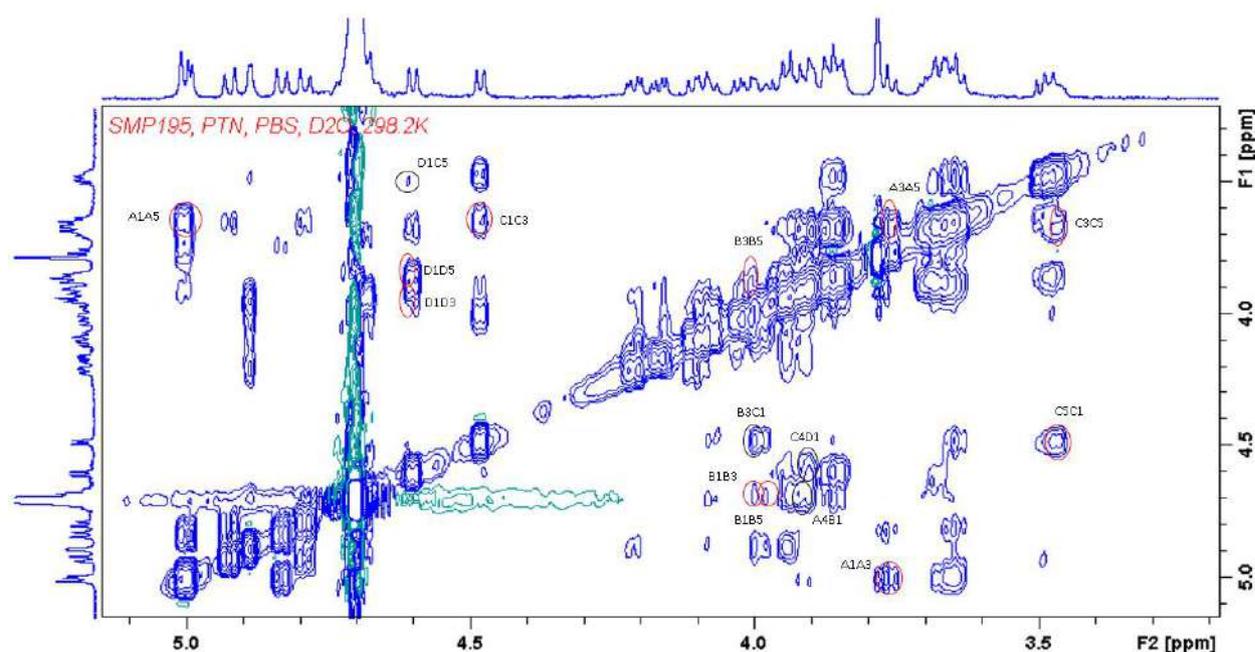


Figura 4.4. Espectros de NOESY transferido de los complejos **2-PTN** (A) y **5-PTN** (B), realizados a 600 MHz y 300K (tiempo de mezcla 500 ms). Se indican los picos de cruce más relevantes con su asignación correspondiente.

El análisis cuantitativo de las distancias interglucosídicas experimentales, a partir de las señales que se pudieron determinar, parece reflejar la presencia de las dos conformaciones tipo *syn-ψ* en los enlaces glucosídicos, aunque no fue posible el cálculo de todas las distancias, como es el caso de B1-A3 y E1-D4 en el caso del complejo **2-PTN**, y B1-A4, D1-C4 en su homólogo **5-PTN** (Tablas 4.3 y 4.4). Además, el rango en el que se encuentran estos valores indica conformaciones *exo-φ*. Debido a la proximidad entre la señal del agua y los protones H1 de las GalNAc C y B de los complejos **2-PTN** y **5-PTN** respectivamente, la presencia de estos NOEs interprotónicos es muy débil, entremezclándose con la diagonal o en algunos casos sin ni siquiera llegar a detectarse. Por otro lado, tampoco fueron posibles las asignaciones de todas las $^3J_{HH}$ experimentales sobre los espectros monodimensionales (Figura A.13 del Anexo), pero aquellas que sí pudieron observarse indicaron que tanto las Glc como las GalNAc se encontraban en la conformación de silla canónica 4C_1 (Tablas 4.3 y 4.4). A pesar de que las constantes de acoplamiento se promedian proporcionalmente a sus poblaciones, a diferencia del NOE que lo hace en función de los tiempos de correlación, la valoración de todas las constantes de acoplamiento del anillo sugiere que los residuos se encuentran en la misma conformación que en el estado libre. Esto se confirma por los NOE entre los protones del anillo en disposición *syn*-diaxial que muestran NOEs correspondientes a una disposición 4C_1 del anillo.

$^3J_{HH}$	2-PTN	Distancias	2-PTN
A1A2	8.6	A H1H3	2.6
A2A3	10.1	A H1H5	2.7
A3A4	2.8	A H3H5	
A4A5		Interg.BH1-AH3	
B1B2	8.1	B H1H3	2.5
B2B3	8.6	B H1H5	2.6
B3B4	9.1	B H3H5	2.3
B4B5	7.2	Interg.CH1-BH4	2.2
C1C2		C H1H3	
C2C3	10	C H1H5	
C3C4	2.4	C H3H5	
C4C5		Interg.DH1-CH3	2.2
D1D2	8.1	D H1H3	2.6
D2D3	6.8	D H1H5	2.6
D3D4	9.5	D H3H5	2.5
D4D5		Interg.EH1-DH4	
E1E2	8.2	E H1H3	2.4
E2E3		E H1H5	2.7
E3E4	2.7	E H3H5	2.4
E4E5			

Tabla 4.3. Constantes de acoplamiento y distancias experimentales para el ligando 2 en presencia de PTN medidas a 298K.

$^3J_{HH}$	5-PTN	Distancias	5-PTN
A1A2	7.9	A H1H3	2.5
A2A3	9.2	A H1H5	2.8
A3A4	9.2	A H3H5	
A4A5	7.9	Interg.BH1-AH4	
B1B2		B H1H3	
B2B3	10.8	B H1H5	
B3B4	2.9	B H3H5	2.6
B4B5	1.4	Interg.CH1-BH3	2.3
C1C2	8	C H1H3	2.7
C2C3	8.3	C H1H5	2.5
C3C4	9.4	C H3H5	2.7
C4C5	9.4	Interg.DH1-CH4	

Tabla 4.4. Constantes de acoplamiento y distancias experimentales para el ligando 5 en presencia de PTN medidas a 298K.

$^3J_{HH}$	5-PTN	Distancias	5-PTN
D1D2	8	D H1H3	2.8
D2D3	10.3	D H1H5	2.3
D3D4		D H3H5	2.5
D4D5			

Tabla 4.4. *Constitución.*

4.3.1.3 Comparativa entre los complejos 3-PTN y 6-PTN

A diferencia de los resultados descritos para los complejos anteriores, para aquellos que presentaban el ligando completamente sustituido (**3** y **6**) resultó más complejo extraer valores experimentales. En el caso del complejo pentasacárido (**3-PTN**), la presencia de las dos cargas negativas adicionales procedentes de los grupos sulfato del anillo de GalNAc del terminal no reductor, hizo posible su disolución en el *buffer* pudiéndose generar el espectro monodimensional de $^1\text{H-RMN}$ (ver Figura A.30 del Anexo), a diferencia del complejo de la serie tetrasacárido (**6-PTN**) donde no fue posible. Como se puede observar en la Tabla 4.5, el solapamiento de las señales impidió determinar algunas de las constantes de acoplamiento y distancias experimentales (Figura 4.5). Con respecto a los protones interresiduales, sólo fue posible identificar una señal aislada que podría confirmar que la unión C1-B4 se encontraba en geometría *syn* (color negro). En el caso de las conformaciones de los anillos sí se pudieron determinar algunos de los picos de cruce más representativos, que sostenían de nuevo una silla de tipo $^4\text{C}_1$.

En cambio, en el caso de su homólogo tetrasacárido **6-PTN**, la ausencia del anillo de GalNAc adicional impidió la solubilización completa en el *buffer* imposibilitando el análisis experimental del complejo.

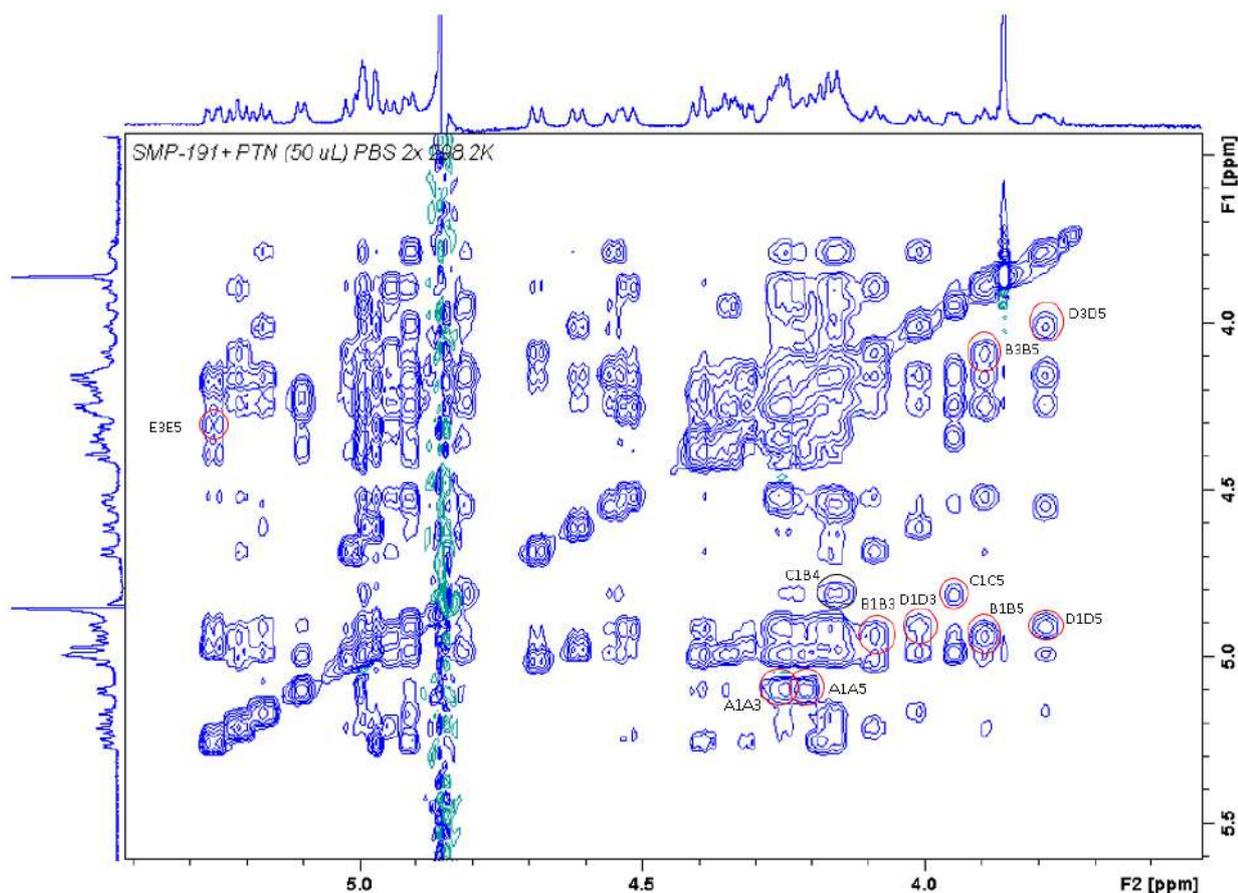


Figura 4.5. Espectro de NOESY transferido del complejo 3-PTN realizado a 600 MHz y 300K (tiempo de mezcla 500 ms). No se pudieron asignar todos los picos de cruce más representativos.

$^3J_{HH}$	3-PTN	Distancias	3-PTN
A1A2	8.0	A H1H3	2.6
A2A3	10.1	A H1H5	2.5
A3A4	2.2	A H3H5	
A4A5		Interg.BH1-AH3	
B1B2	7.9	B H1H3	2.6
B2B3	7.7	B H1H5	2.7
B3B4	9.0	B H3H5	2.6
B4B5	7.6	Interg.CH1-BH4	2.4
C1C2	8.4	C H1H3	
C2C3	10.5	C H1H5	
C3C4		C H3H5	
C4C5		Interg.DH1-CH3	

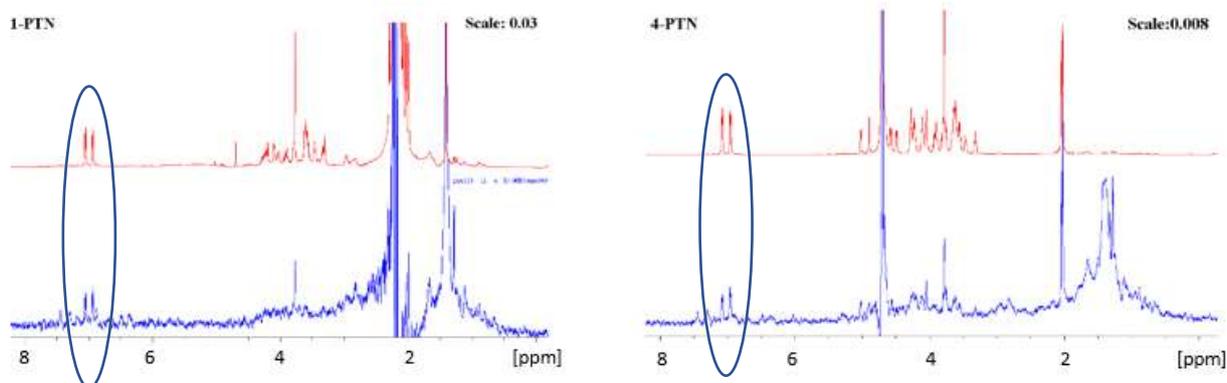
Tabla 4.5. Constantes de acoplamiento y distancias experimentales para el ligando 3 en presencia de PTN medidas a 298K.

$^3J_{HH}$	3-PTN	Distancias	3-PTN
D1D2	7.5	D H1H3	2.2
D2D3	8.3	D H1H5	3.0
D3D4	9.1	D H3H5	2.5
D4D5	8.5	Interg.EH1-DH4	
E1E2		E H1H3	
E2E3	10.7	E H1H5	
E3E4		E H3H5	2.9
E4E5			

Tabla 4.5. Continuación.

4.3.2 Experimento de STD (Saturation Transfer Difference)

Otra de las técnicas de RMN empleadas para determinar si existe o no interacción entre el ligando y la proteína es el STD. Como se ha descrito en el **Capítulo 2**, este experimento consiste en irradiar los protones de la proteína, en este caso de la PTN, transfiriendo dicha magnetización al ligando libre en función del tiempo de saturación, gracias al equilibrio asociación-disociación. La cantidad de irradiación recibida para cada protón del ligando se puede establecer en rangos de porcentaje de saturación que pueden representarse mediante los llamados mapas de epítipo. En la Figura 4.6 se han representado los espectros para los complejos con los ligandos desprotegidos y dibencilados de ambas series. En cambio, para los complejos con los ligandos persustituídos, debido al gran solapamiento de las señales (en el caso de **3-PTN**) y a la insolubilidad del compuesto en el *buffer* (**6-PTN**), no fue posible generar un espectro de STD interpretativo. Como se puede ver en la imagen, se ha representado en color rojo el experimento de ^1H -RMN *off-resonance*, y en azul el espectro de diferencia de STD resultante. Cabe destacar que no fue posible extraer datos de STD en algunos de los protones debido al solapamiento entre algunos de ellos y el bajo nivel de saturación recibido por parte de la proteína.



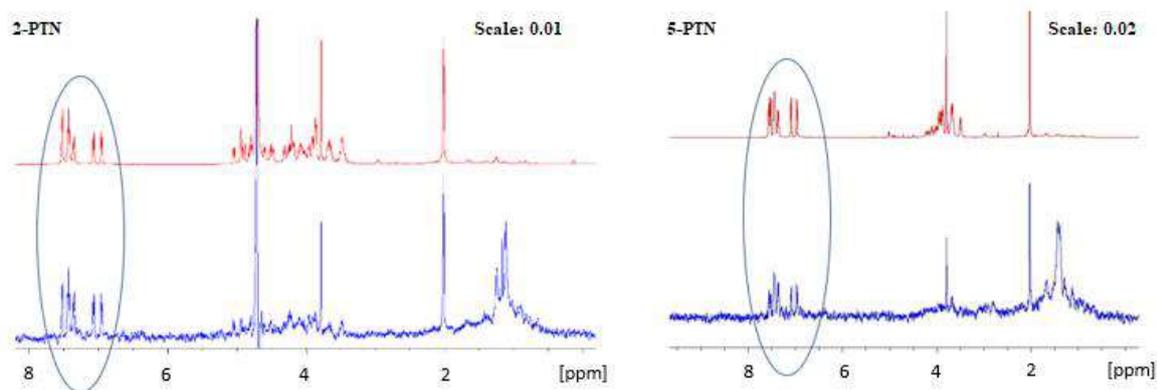
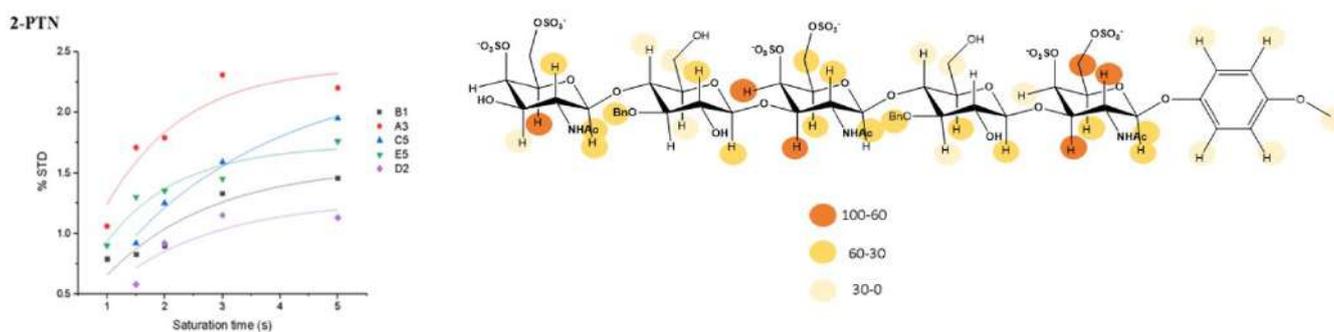


Figura 4.6. Espectros obtenidos de los experimentos de STD. En rojo se representan los espectros de ^1H -RMN y en azul los espectros de ^1H -STD para los ligandos desprotegidos (**1** y **4**) y los dibencilados (**2** y **5**).

A pesar de la baja resolución de las señales, los espectros de la Figura 4.6 confirman la unión de los miméticos de CS-E con la PTN. Para el caso de los complejos **1-PTN** y **5-PTN**, no fue posible extraer valores suficientes para crear las curvas de crecimiento; pero las señales situadas a campo bajo correspondientes a los protones aromáticos del grupo OMP de los extremos reductores y de los grupos bencilos, bastaron para confirmar la interacción, sugiriendo una vez más la interacción de estas citoquinas con grupos hidrofóbicos^[18]. A pesar del solapamiento de los protones anoméricos con los protones CH_2 de los grupos bencilos, en el caso del complejo pentasacárido **2-PTN** sí fue posible extraer valores de STD, al igual que para el complejo tetrasacárido **4-PTN**.



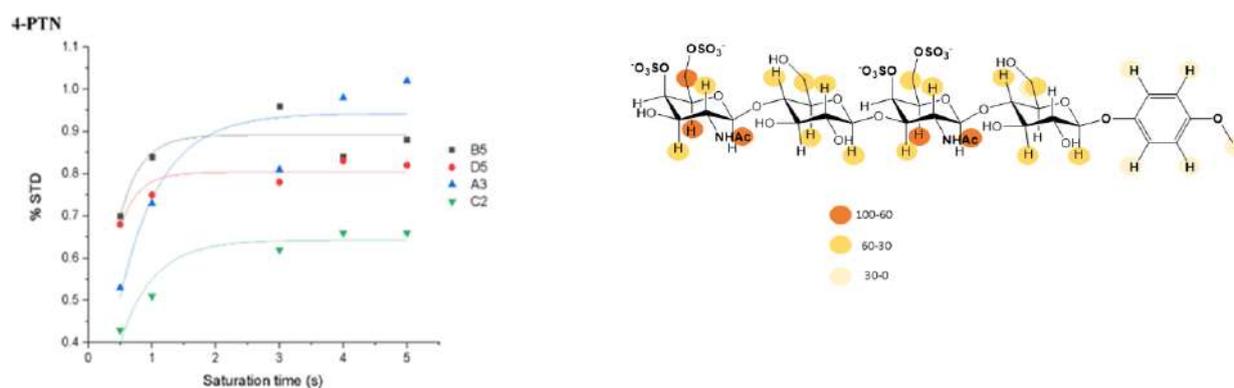


Figura 4.7. A la izquierda, curvas de crecimiento de STD y a la derecha sus correspondientes mapas de epítopo para los complejos pentasacarídico dibencilado (**2-PTN**) y tetrasacarídico completamente desprotegido (**4-PTN**).

En la Figura 4.7 se han representado las curvas de crecimiento de STD para algunos protones de cada anillo de azúcar para los únicos complejos en los que fue posible medir la interacción, con los ligandos **2** y **4**. Para visualizar estos resultados, se representaron los mapas de epítopo de ambos ligandos en los que, mediante una gama de colores, establecemos los puntos que presentan mayor afinidad con la PTN. Como se puede observar en la imagen anterior, la interacción se distribuye por toda la estructura, por lo que no fue posible determinar qué residuo o qué zona es más favorable para que tenga lugar esta unión.

4.3.3 Experimento de ^{15}N - ^1H -HSQC

Hasta ahora se han empleado técnicas de observación del ligando. Éstas han permitido confirmar la formación de complejos transitorios, determinar la estructura del ligando en el complejo, así como determinar sus modos de interacción con el receptor. La posibilidad de disponer de PTN con marcaje isotópico en ^{15}N nos ha permitido abordar la interacción entre PTN y los ligandos desde el punto de vista del receptor.

Tanto PTN como MK se pueden expresar en medios enriquecidos en ^{15}N obteniendo la proteína con marcaje isotópico lo que permite realizar experimentos de correlación ^{15}N HSQC (*Heteronuclear Single Quantum Coherence*).^[19] Este experimento, que se realiza en H_2O , detecta los pares $^1\text{H} - ^{15}\text{N}$ de todos los sistemas ^{15}N -H unidos por un enlace; normalmente se restringe a las amidas por ser las que aportan mayor información. Esta información es muy útil para el estudio de las variaciones de los desplazamientos químicos asociadas a cambios conformacionales, interacciones intermoleculares, variaciones de estructura secundaria, etc ...^[20]

En nuestro caso, debido a que los desplazamientos químicos de estos dos átomos de la proteína son muy sensibles a las alteraciones en el entorno magnético como resultado de la interacción con el ligando, la perturbación dará lugar a cambios en los desplazamientos químicos de su posición original indicando la formación del complejo. Como ya se ha

explicado en el **Capítulo 2**, en función de la cinética relativa a la diferencia de desplazamientos químicos existen dos posibilidades de intercambio: rápido o lento. En el rápido se observa el pico promedio entre el estado libre y el asociado, mientras que en el lento se observan dos picos para cada uno de los estados con intensidades proporcionales a sus poblaciones.^[21]

En nuestro caso, al tratarse de compuesto miméticos de CS-E, nuestro primer objetivo fue determinar si el complejo formado era similar al constituido por los ligandos naturales. Para este experimento se eligió el ligando tetrasacárido **4**, similar a la secuencia natural de CS-E, salvo por el reemplazo del carboxilato de GlcA por un grupo hidroximetilo. Al mismo tiempo, se trataba del ligando más sencillo, frente al pentasacárido. Así, se realizó una valoración con **4** de una muestra de ¹⁵N-PTN, con precauciones para evitar la influencia de cambios en la concentración de *buffer* o proteína en los desplazamientos químicos, ver parte experimental. De la serie de experimentos ¹⁵N-HSQC, se extrae la variación de desplazamiento químico normalizada en función de la posición del residuo individual.

Investigaciones previas de la interacción entre GAG y PTN, empleando la misma técnica, han encontrado que, en general, existe un mayor cambio de desplazamiento químico en el dominio CTD que en el NTD.^[5]

Además, hay estudios donde se han empleado hexasacáridos de CS-A, CS-E, DS y de heparina para analizar la interacción con la PTN encontrando que los residuos básicos de la zona de la bisagra (concretamente Lys60, Gln62, Phe63, Gly64 y Cys67) y los aminoácidos Lys91, Arg92, Asn96, Glu98 y Cys99 en el dominio CTD, mostraron mayores desplazamientos químicos tras la unión.^[6] Por otro lado, los residuos en el dominio NTD y la cola C-terminal apenas mostraron perturbación alguna, pero aquellos que sí muestran algún cambio químico considerable, se sitúan cerca de los aminoácidos básicos del CTD^[6]. En la Figura 4.8 se muestra el espectro bidimensional de ¹⁵N-HSQC donde cada pico de cruce representa una de las parejas N-H de la proteína. Puede observarse que muchos grupos NH se ven afectados por la interacción y presentan variaciones tanto en δ de ¹H como en el ¹⁵N según va variando la concentración de ligando.

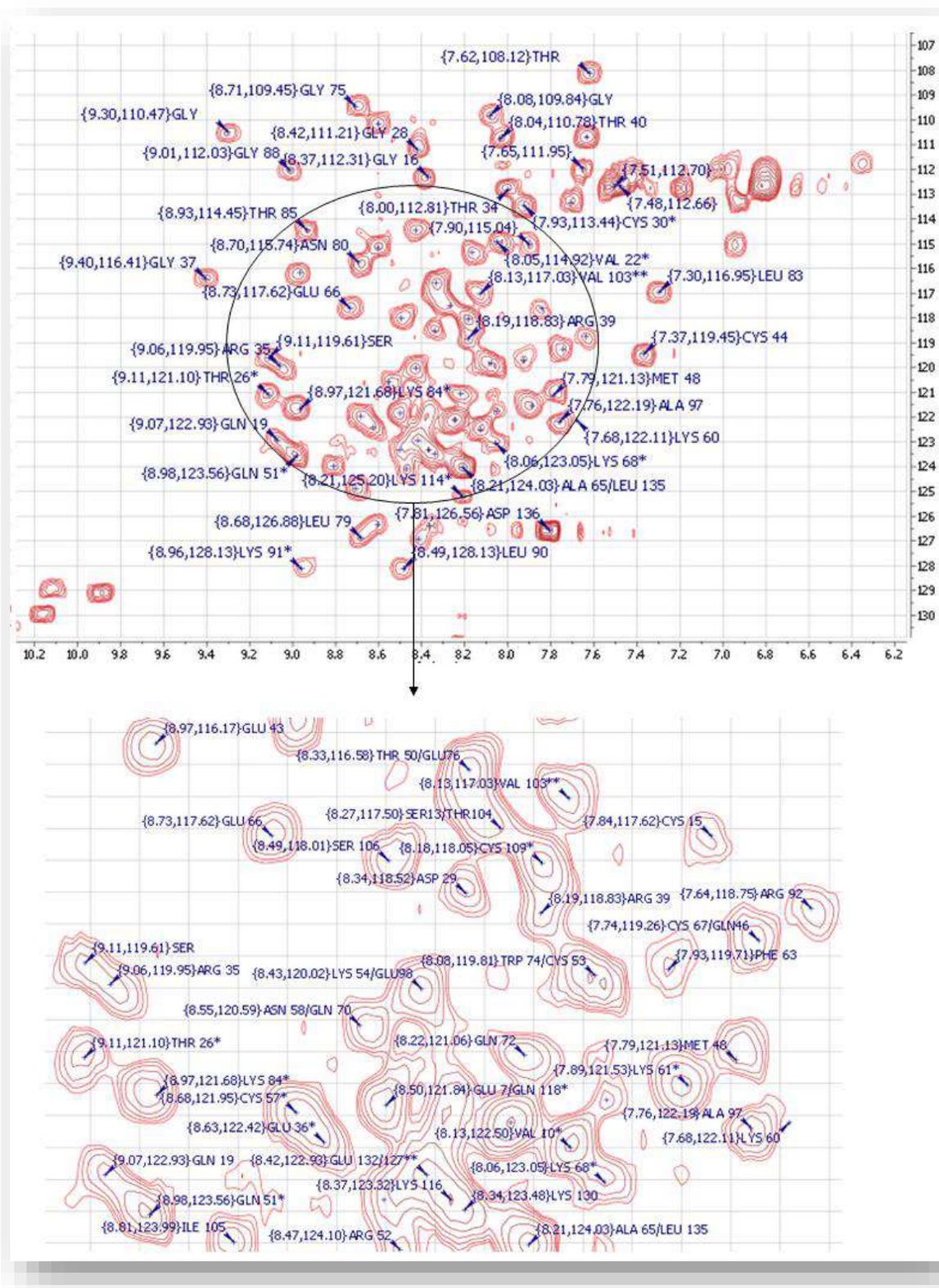


Figura 4.8. Experimento ^{15}N - ^1H -HSQC para el complejo tetrasacárido 4-PTN realizado a 600 MHz y 300K.

Descartando las prolinas, en el espectro bidimensional se esperan señales de correlación heteronuclear correspondientes a las amidas de las cadenas principales, además de los NH₂ adicionales procedentes de los aminoácidos Asn y Gln, y los NH aromáticos de los residuos de Trp. Por otro lado, es posible observar señales de las cadenas laterales de las Lys (NH₃), Arg (NHC(NH₂)₂) e His (NH aromático) aunque estén fuera del rango. De todos ellos, hemos conseguido identificar inequívocamente 44 de los NH amídicos comprendidos desde 9.5 ppm hasta 7 ppm. El principal motivo de esta disminución ha sido el elevado solapamiento accidental de las señales causado por el ensanchamiento de las mismas. Una vez identificados los aminoácidos con mayor afinidad, se ha representado en el diagrama de la Figura 4.9 las perturbaciones de δ observadas para este complejo.

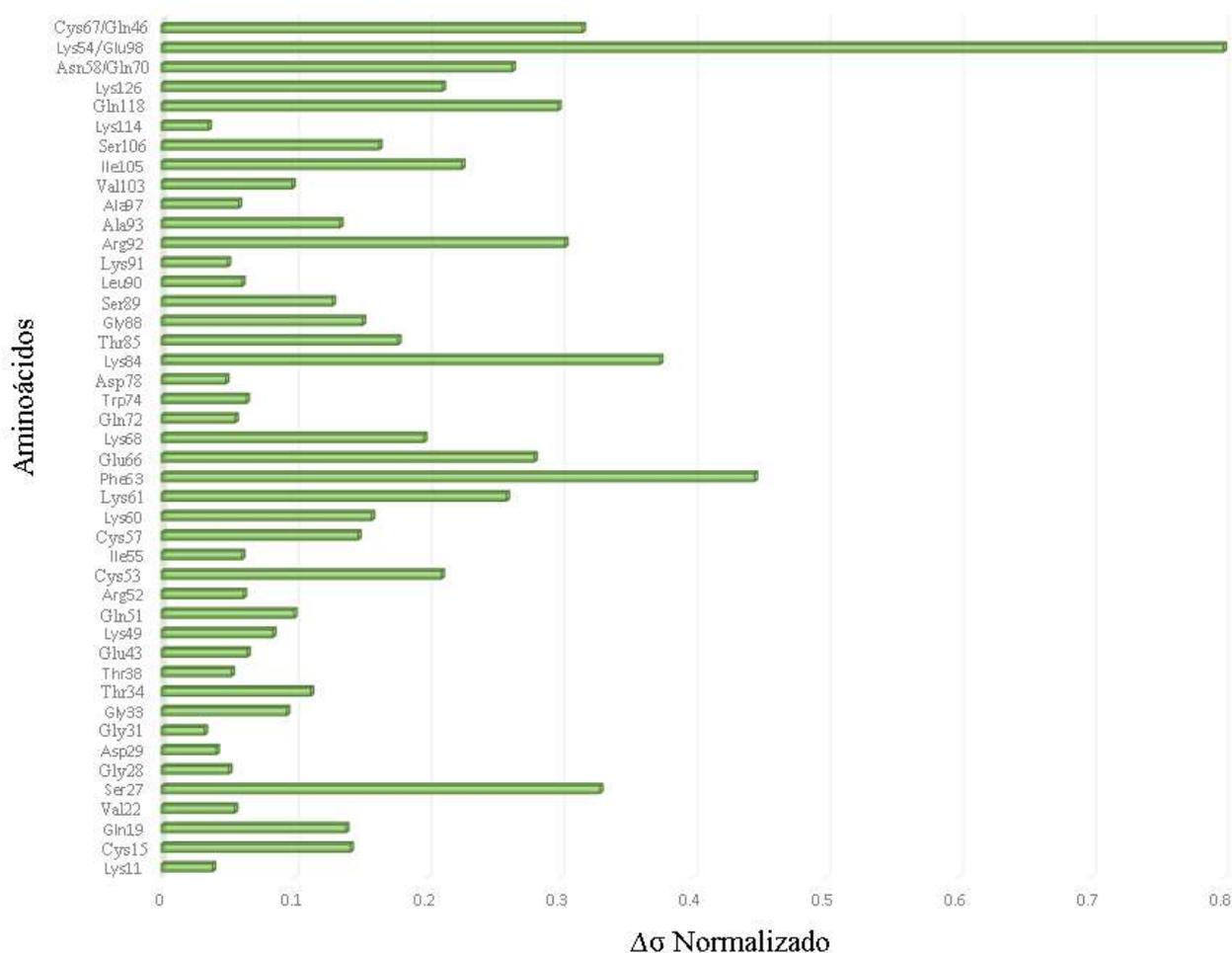


Figura 4.9. Diagrama de interacción del experimento ^{15}N - ^1H -HSQC a 600 MHz y 300K.

De estudios previos se sabe que las regiones de la PTN con mayor interacción por GAGs (Heparina y CS-E) son el dominio NTD, la zona bisagra y el dominio CTD.^{[22],[6],[3],[2]} En el dominio NTD, que comprende del residuo 15 al 57, los más involucrados son los próximos al final, del 51 al 57; en la zona *hinge* o bisagra (residuos del 58 al 66) prácticamente todos ellos; para el dominio CTD (residuos del 67 al 107), aquellos que pertenecen al grupo II (aminoácidos 90-100).^[5] El análisis de nuestros resultados muestra que las áreas de interacción están bastante distribuidas en la secuencia primaria, siendo

participes aminoácidos presentes en los tres sitios de unión principales. Aquellos que se encuentran en el dominio CTD y en la zona *hinge* son los que presentan valores más altos de $\sigma_{\text{normalizado}}$. Dentro de este segundo dominio encontramos una variación mayor ($\sigma > 0.2$ ppm) por parte de los aminoácidos Thr85, Lys84, Ile105, Lys91, Lys68, Arg92 y Gln118. Además, debido a la cercanía entre los picos de cruce, no fue posible asignar inequívocamente las señales de algunos de ellos, como Asn58 o Gln70, Cys67 o Gln46 y Lys54 o Glu98, siendo este último el aminoácido que ha presentado la mayor interacción con un valor de $\sigma = 0.8$. Tanto la zona *hinge* como el dominio NTD, presentan una interacción dominante en los residuos Phe63, Lys61 y Cys53. Además, se observa una interacción en la zona de la cola C terminal por parte del residuo Glc118, sin detectarse una interacción significativa en el otro extremo terminal. De los residuos que comprenden el primer dominio NTD, los aminoácidos que han experimentado desplazamiento químico presentan por lo general valores de $\sigma < 0.1$ ppm. Con la intención de poder discernir entre las parejas de aminoácidos dudosas que se han expuesto, se han representado los modelos tridimensionales extraídos de las coordenadas finales de las dinámicas moleculares que veremos a continuación, señalando en estas estructuras los residuos para intentar identificarlos en función de su proximidad al ligando azúcar. En la Figura 4.10 se observan los modelos con las tres conformaciones de la proteína, aprovechando la flexibilidad que experimenta durante la simulación, intentando reproducir una situación similar a la que nos podríamos encontrar experimentalmente.

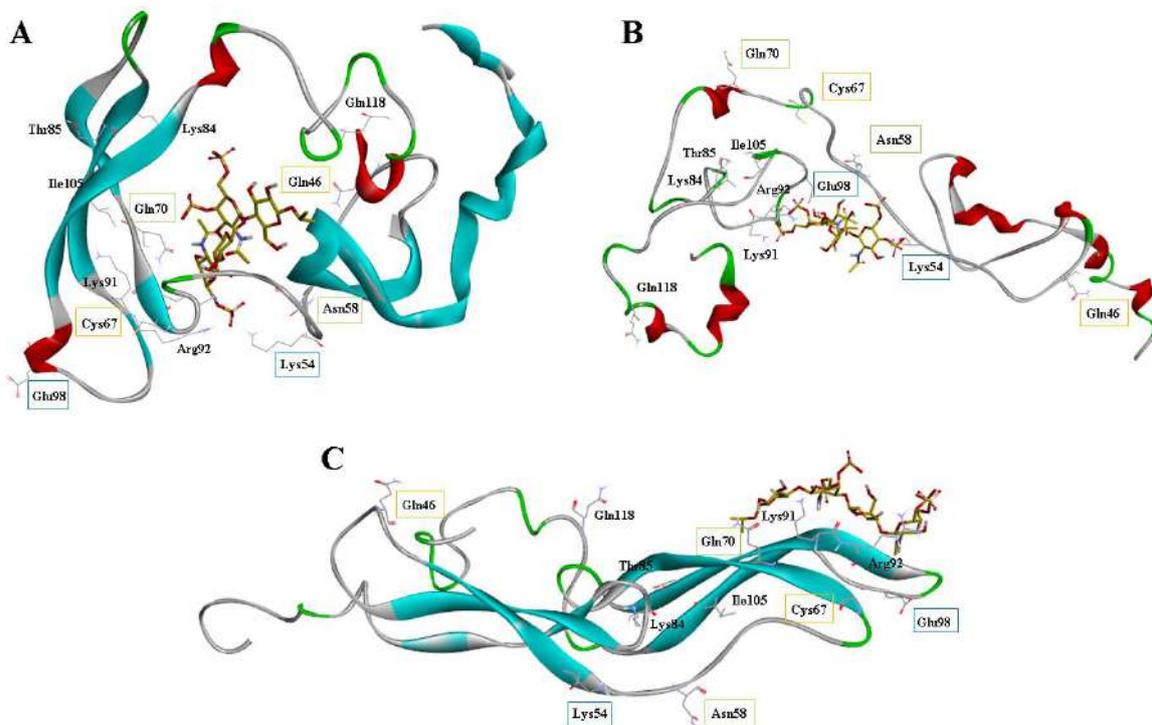


Figura 4.10. Representación tridimensional de los aminoácidos implicados en la interacción del experimento ^{15}N -HSQC en las conformaciones plegada (A), abierta (B) y en forma de horquilla (C) de la PTN para determinar su grado de implicación en la interacción con el ligando. Se han remarcado en diferentes colores las parejas de aminoácidos que no se pudieron asignar experimentalmente.

Como dependiendo de la conformación de la PTN estos residuos pueden estar más o menos próximos al ligando, nos podríamos inclinar por el dominio de la proteína en el que se encuentran, ya que los pertenecientes a las zonas central y dominio CTD son los que han mostrado una afinidad mayor con el ligando (por lo tanto, podrían asignarse a Glu98, Glu70 y Cys67).

4.3.4 Experimento de Polarización de Fluorescencia (FP)

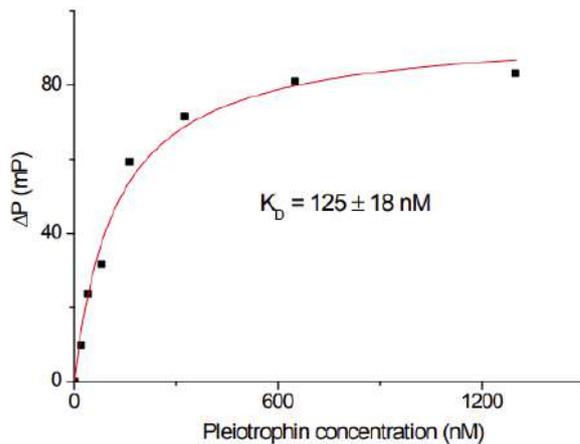


Figura 4.11. Curva de interacción del experimento de unión directa entre la sonda fluorescente y la PTN. ΔP se ha representado frente a la concentración de proteína ajustando la curva a un modelo de un único sitio de interacción.

En última instancia, se van a exponer los resultados experimentales obtenidos a partir de diferentes ensayos de competición llevados a cabo mediante la técnica de Polarización de Fluorescencia, que corroboraron y cuantificaron la interacción de estos miméticos de CS-E con la PTN. Como ya se ha descrito en el **Capítulo 2**, una de las ventajas de este experimento es que nos permite conocer el valor de la afinidad empleando muy poca cantidad de muestra, convirtiéndola así en una técnica eficaz para el cribado de alto rendimiento e ideal para el análisis de interacción entre proteínas y ligandos pequeños o relativamente grandes.

Primeramente, se tiene que realizar un primer ensayo que nos permite conocer si la sonda fluorescente que vamos a emplear tiene una fuerte interacción (idealmente en el rango nanomolar) con la proteína de estudio, en este caso PTN. Trabajos desarrollados con anterioridad en nuestro grupo de investigación permitieron la obtención de un hexasacárido de heparina marcado fluorescentemente.^{[23][24]} Este compuesto se eligió como sonda fluorescente para nuestros ensayos, ya que se tenía constancia de que presentaba una fuerte interacción con otras citoquinas.^[18] Para este experimento se registró la polarización en pocillos de microplaca que contenían concentraciones crecientes de proteína y 10 nM de sonda.

Una vez medida la polarización, se representó la variación de polarización (ΔP), definida como la diferencia entre los valores de FP de cada muestra y el valor de FP residual de la sonda, frente a la concentración de proteína, obteniéndose la correspondiente curva de unión (Figura 4.11). Esta curva se ajustó al modelo de interacción con un único centro activo (ver desarrollo en el **Capítulo 2**), obteniéndose la constante de disociación de la sonda fluorescente ($K_D=125$ nM).^[25]

Una vez que se determinó que la sonda fluorescente presentaba una alta afinidad por PTN y era adecuada para realizar los ensayos de competición descritos en el Capítulo 2, se llevaron a cabo dicho experimentos para calcular las afinidades relativas no sólo de los

ligandos sin proteger **1** y **4**, sino también de los dibencilados y los completamente protegidos (**2-3**, **5-6**), ya que se ha demostrado en estudios anteriores que la presencia de grupos hidrófobos proporciona uniones más fuertes a ciertas citoquinas.^[25] Para ello, se llevaron a cabo medidas de FP de pocillos con concentraciones fijas de proteína y sonda fluorescente, pero crecientes del ligando de interés, obteniéndose las curvas de inhibición representadas en la Figura 4.12. A partir del ajuste de dichas curvas, se pudieron determinar los valores de IC_{50} , definidos como la concentración de ligando que proporciona un 50% de inhibición.

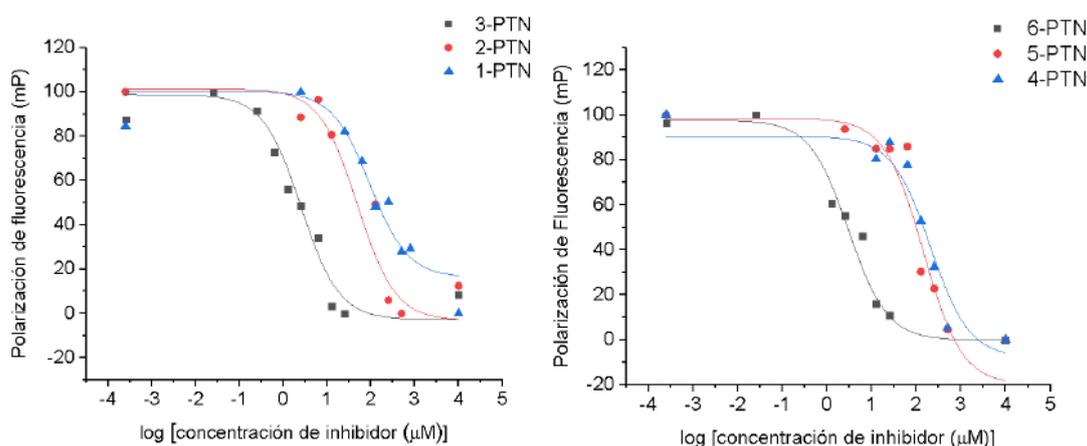


Figura 4.12. Representación de las curvas de inhibición para los seis compuestos con PTN. A la izquierda se muestra la gráfica correspondiente a los complejos pentasacáridicos y a la derecha la de los complejos tetrasacáridicos.

Los puntos de los extremos de las curvas corresponden a los controles que se han empleado para realizar el experimento. Por un lado, el valor mínimo de FP de la curva, que indica el 100% de inhibición, corresponde a un pocillo que contenía únicamente la sonda fluorescente, en ausencia de proteína. El punto máximo de FP, que representa el 0% de inhibición, corresponde a la muestra control que contiene PTN y sonda, sin presencia de ligando competidor. Para el ajuste de las curvas se siguió la ecuación del modelo de competición con un único sitio de reconocimiento, que se detalla en la parte experimental, obteniéndose los valores de IC_{50} que se muestran en la Tabla 4.5:

Pentasacáridos	IC_{50} (μM)	Tetrasacáridos	IC_{50} (μM)
1	194 ± 142	4	278 ± 96
2	84 ± 40	5	180 ± 39
3	3 ± 1	6	6 ± 4

Tabla 4.5. Representación de los valores de IC_{50} para los seis complejos. Los errores representan las desviaciones estándar de las diferentes repeticiones realizadas de cada experimento.

Los resultados de la Tabla 4.5 confirmaron que la presencia de grupos hidrofóbicos mejoraba notablemente la interacción con la proteína. Así, los compuestos desprotegidos **1** y **4** se unen a la PTN débilmente, en el rango micromolar alto, mientras que se obtienen valores de afinidad mayores para los derivados **2** y **5**, con la presencia de grupos bencilos hidrofóbicos. Por último, las afinidades relativas aumentan hasta aproximadamente 50 veces, en el caso de los productos **3** y **6**, que se encuentran totalmente protegidos. Además, los valores de IC₅₀ obtenidos con los pentasacáridos son, en general, más bajos que los obtenidos para los tetras.

4.4 Estudio teórico

En este apartado comienza la descripción del estudio teórico de las interacciones entre PTN y los compuestos miméticos de GAG, complementario del análisis experimental. A diferencia del capítulo anterior, los estudios se realizan sobre complejos carbohidrato-proteína, añadiendo una dificultad adicional.

En primer lugar, se van a exponer los resultados del *docking* molecular y la generación de los complejos bimoleculares. Se discutirán las diferentes orientaciones que adoptan los ligandos azúcar en su estado enlazado y las áreas prioritarias de interacción de la PTN. Una vez determinada la estructura de los complejos, la zona de interacción y los conformeros encontrados, se detallarán los resultados de los cálculos de dinámica molecular. Además, se describirán las estructuras de los ligandos y de la proteína en los correspondientes complejos y sus cambios relativos a las especies libres. En último lugar, sobre las trayectorias finales de cada una de las simulaciones de MD realizadas, se han llevado a cabo los cálculos de interacción de MM-PBSA/GBSA. Estos cálculos permitirán valorar la contribución entrópica a la energía de afinidad y la aportación energética individualizada de los aminoácidos de la PTN.

4.4.1 Docking molecular

Como ya se ha descrito en el **Capítulo 1**, la PTN es una proteína compuesta mayoritariamente por aminoácidos básicos (36 están cargados positivamente y 17 negativamente). Tiene dos dominios unidos por una zona intrínsecamente desordenada y termina en dos regiones sin estructura. Los dos dominios estructurados tienen una longitud aproximadamente de 40 residuos y adoptan el mismo tipo de plegamiento que la TSR de tipo 1, constando de una lámina-β simple de tres cadenas estabilizada por puentes disulfuro que estabilizan su plegamiento. Una de las características más significativas de la PTN es que las caras opuestas de cada dominio tienen potenciales electrostáticos diferentes. Debido a esta topología, los aminoácidos básicos se sitúan en una sola cara convirtiéndola por tanto en un sitio de unión crucial para las interacciones de PTN con muchos de sus receptores.^{[26],[22]}

Como se ha mencionado antes, muchos de estos receptores presentan en sus estructuras largas cadenas de GAGs sulfatadas que presentan interacciones con esta proteína.

Numerosos estudios demuestran que las afinidades PTN-GAG son extremadamente altas ($K_d \sim 1$ nM). Se conoce que el extremo C-terminal, en particular, tiene una potente actividad biológica, donde las interacciones entre los aminoácidos básicos y los aromáticos y los GAGs juegan un papel importante.^[22] Además, diversos estudios de interacciones GAG-proteína por RMN muestran que los residuos encontrados en el grupo II del dominio CTD presentan perturbaciones y experimentan cambios en el desplazamiento de la señal en presencia de estos ligandos oligosacáridos.^[22]

En la estructura de la PTN descrita por Wang^[6], al no detectar NOEs entre ambos dominios TSR de tipo 1, no pudieron descartar que actuasen de manera independiente y que cuanto más larga fuese la cadena oligosacárida, mayor afinidad presentaría por la PTN. Esto se basa en la hipótesis de que los GAGs compuestos por diez o más unidades monosacáridicas, podrían llegar a enlazar ambos dominios simultáneamente.^[22] Otros resultados de investigaciones llevadas a cabo por este investigador, indicaron que en función del tipo de CS considerado el papel y las actividades de los dominios tipo TSR eran diferentes. Pero, a pesar de no haberse encontrado estos contactos NOE, sí que se localizaron picos de cruce definidos entre el dominio CTD y residuos localizados en la zona que los conecta. Específicamente, los protones aromáticos de la Phe63 se ponían en contacto con las cadenas laterales de los aminoácidos Tyr69, Val103 e Ile105, pudiendo tener consecuencias importantes en la unión a GAGs, ya que la curvatura que se genera en esta zona enlazante de los dos dominios posiciona a las Lys60 y Lys61 cerca del grupo 1 del CTD.^[6]

Por otra parte, los resultados experimentales obtenidos en nuestro grupo de investigación han mostrado que la presencia de grupos hidrofóbicos, bencilos o benzoilos, en las cadenas oligosacáridicas aumentan la interacción con estas citoquinas.^{[18],[25]} Por lo tanto, para contrastar estos resultados experimentales con modelos teóricos se llevó a cabo, en primer lugar, *docking* molecular en ambas series de complejos.

Inicialmente, mediante el programa Discovery Studio Visualizer, se realizó una búsqueda global sobre las estructuras depositadas en la base de datos Protein Data Bank para conocer las áreas de potencial afinidad. De las 10 estructuras depositadas se han seleccionado tres, consideradas las más representativas del conjunto. Los resultados de este análisis mostraron: 6 cavidades para la conformación más plegada de la proteína (PTN4) y 3 cavidades en sus formas abierta y horquilla (PTN3 y PTN8, respectivamente) (Figura 4.13).

Para el primer caso (PTN4), el sitio de unión mejor clasificado (Sitio 1) se sitúa entre ambos dominios flanqueado por otros 5 sitios activos más; tres de ellos sobre el CTD y los otros dos sobre los aminoácidos de la cola terminal, conectando con algunos residuos del dominio NTD debido al plegamiento. Para la conformación más abierta (PTN3), el bolsillo principal se encuentra situado en el grupo II del segundo dominio y, en menor proporción, en la zona bisagra. Además, se encontró, en una pequeña cavidad creada entre los aminoácidos Cys109-Thr113, un tercer sitio de interacción. Para el último caso, con la proteína en forma de horquilla, el bolsillo con mayor afinidad abarca desde los residuos

Asn58 hasta la Lys111, es decir, comprende tanto la zona *hinge* central como el dominio C-terminal prácticamente completo.

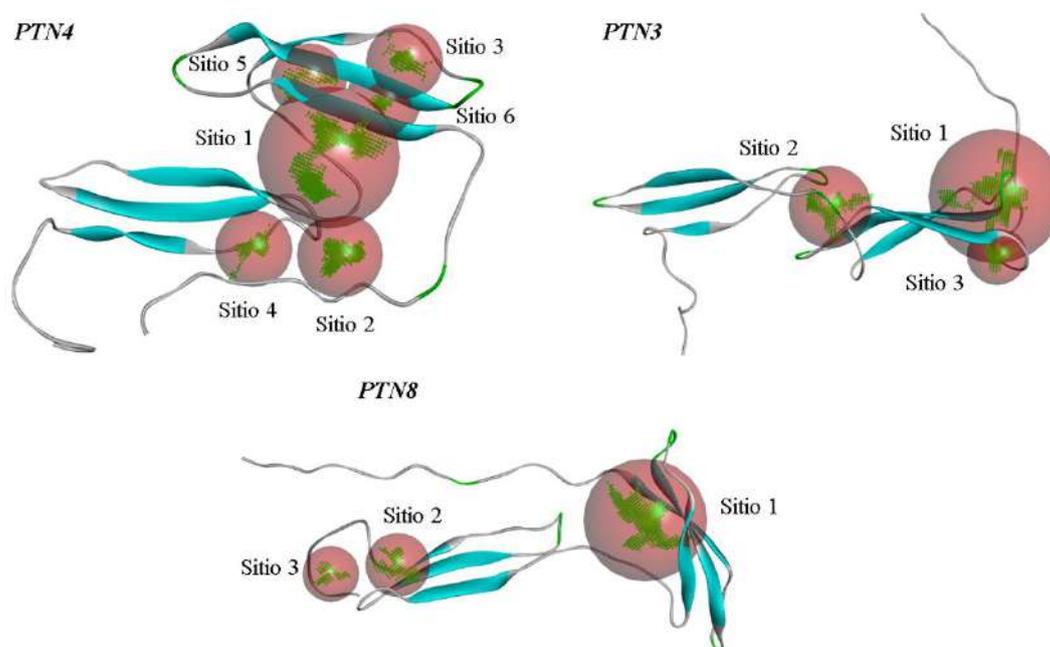


Figura 4.13. Conformaciones de la PTN representando las cavidades con mayor afinidad propuestas por el programa Discovery Studio Visualizer.

Tras estas aproximaciones, se realizaron los estudios de interacción de *docking* molecular empleando los programas Autodock y Autodock Vina sin restringir la localización del ligando a ninguna región de la proteína, todo ello delimitado por una caja o *grid* de dimensiones estándar 62 x 44 x 68 Å. Los resultados de este cálculo generaron nueve complejos para cada estructura pdb (PTN4, PTN3 y PTN8), usando aquellos de mayor afinidad por el receptor en el cálculo sobre las tres estructuras estudiadas en este capítulo. A la hora de analizar estos resultados se han tenido en cuenta los datos del experimento ^{15}N -HSQC, observando ahora la predilección del ligando sobre el sitio activo. Como se observa en la Figura 4.14, para la conformación de la proteína plegada las cadenas de carbohidrato se intercalan en las de la proteína y ésta envuelve a los ligandos entre sus dominios, generando una distribución de los conformeros sobre toda su superficie, quedando repartidos entre ambos dominios, sin observarse una tendencia clara en el área de unión.

Para este sistema, la principal contribución a la afinidad proviene de las interacciones electrostáticas entre los grupos sulfato, desprotonados a pH fisiológico, y las cadenas laterales básicas de Arg, Lys y, en menor medida, His. Éstas se refuerzan con interacciones de Van der Waals, enlaces de hidrógeno e interacciones hidrófobas con las estructuras del esqueleto de carbohidratos. Se observa la participación en la asociación de aminoácidos polares mediante enlaces de hidrógeno como Asn, Gln, flexibles como

Gly o rígidos como Pro.^[27] La coexistencia de residuos rígidos y flexibles sugiere que las regiones en las que se encuentran tienen un papel importante en la función de la PTN.

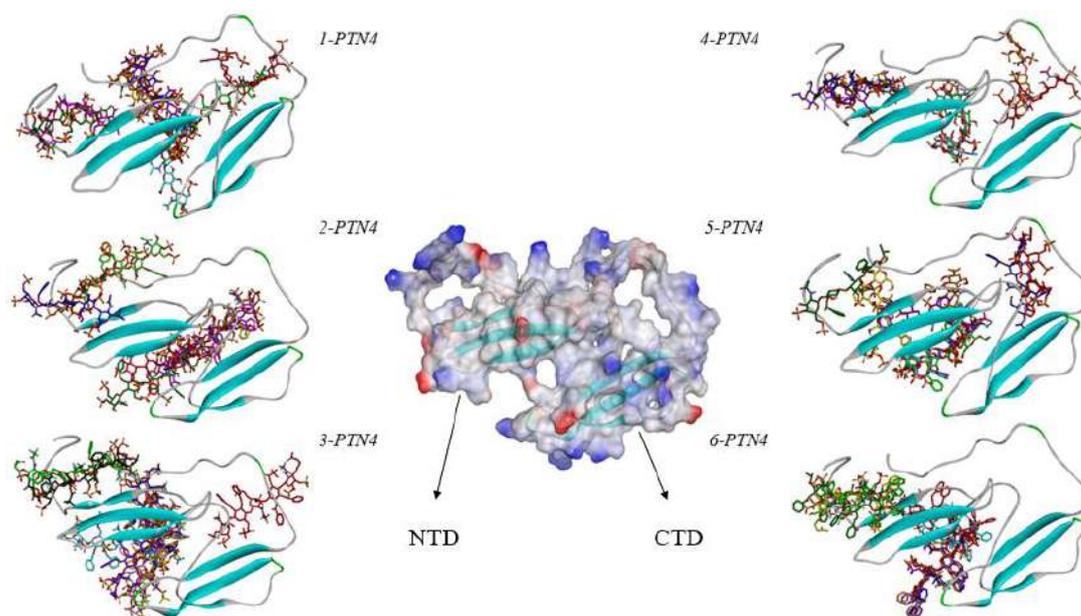


Figura 4.14. Representación tridimensional de los resultados del docking molecular de los seis complejos. Se muestra la superficie de potencial de la PTN en su forma plegada, y la distribución de los 9 conformeros obtenidos del docking para cada ligando.

Para comprender las interacciones resultantes del *docking*, se explican a continuación las interacciones más relevantes siguiendo la clasificación de Discovery Studio Visualizer:^[28]

➤ Clasificadas como “no covalentes” aparecen varios tipos:

1) Las interacciones de enlace de hidrógeno convencionales entre un átomo donador y un átomo aceptor. Los átomos N, O, P y S se consideran átomos donantes clásicos. Para que se considere como unión fuerte, la distancia tiene que ser aproximadamente de 3.4Å.^[29]

2) Las interacciones por puente salino que son relativamente fuertes entre pares de grupos con carga opuesta donde también tiene lugar un enlace de hidrógeno entre ellos.

3) Las interacciones C-H se consideran enlaces de hidrógeno más débiles cuando el donante es un átomo de carbono que se encuentra en un grupo acetileno o adyacente a un oxígeno/nitrógeno.

➤ Clasificadas como electrostáticas se encuentran, además de las de puente salino:

1) Las interacciones π -catión entre los electrones de un sistema π deslocalizado y un residuo cargado positivamente, como suelen ser las cadenas laterales de los aminoácidos Lys y Arg.

2) Las π -anión se establecen de la misma manera, pero con un residuo cargado negativamente como Asp o Glu.

➤ Finalmente, clasificadas como interacciones hidrofóbicas:

1) π - π *stacked* y π - π *T-shaped* donde los anillos π se consideran sistemas planos que interactúan entre sí (sin limitarse únicamente a sistemas aromáticos) y que se comentarán más adelante.

2) Entre grupos alquilo, como las cadenas laterales de los aminoácidos alifáticos.

3) π - σ , que son interacciones débiles entre un hidrógeno y un sistema π ; además como requisito, el hidrógeno que actúa como donante debe estar unido a un carbono no aromático.

4) π -alquil, interacciones entre sistemas aromáticos y grupos alquilo.

En la Figura 4.15 se muestran las principales interacciones que tienen lugar en los resultados del *docking*, procesados a través del programa Discovery Studio, sobre el conformero más estable para cada uno de los ligandos de cada serie y la conformación plegada de la PTN.

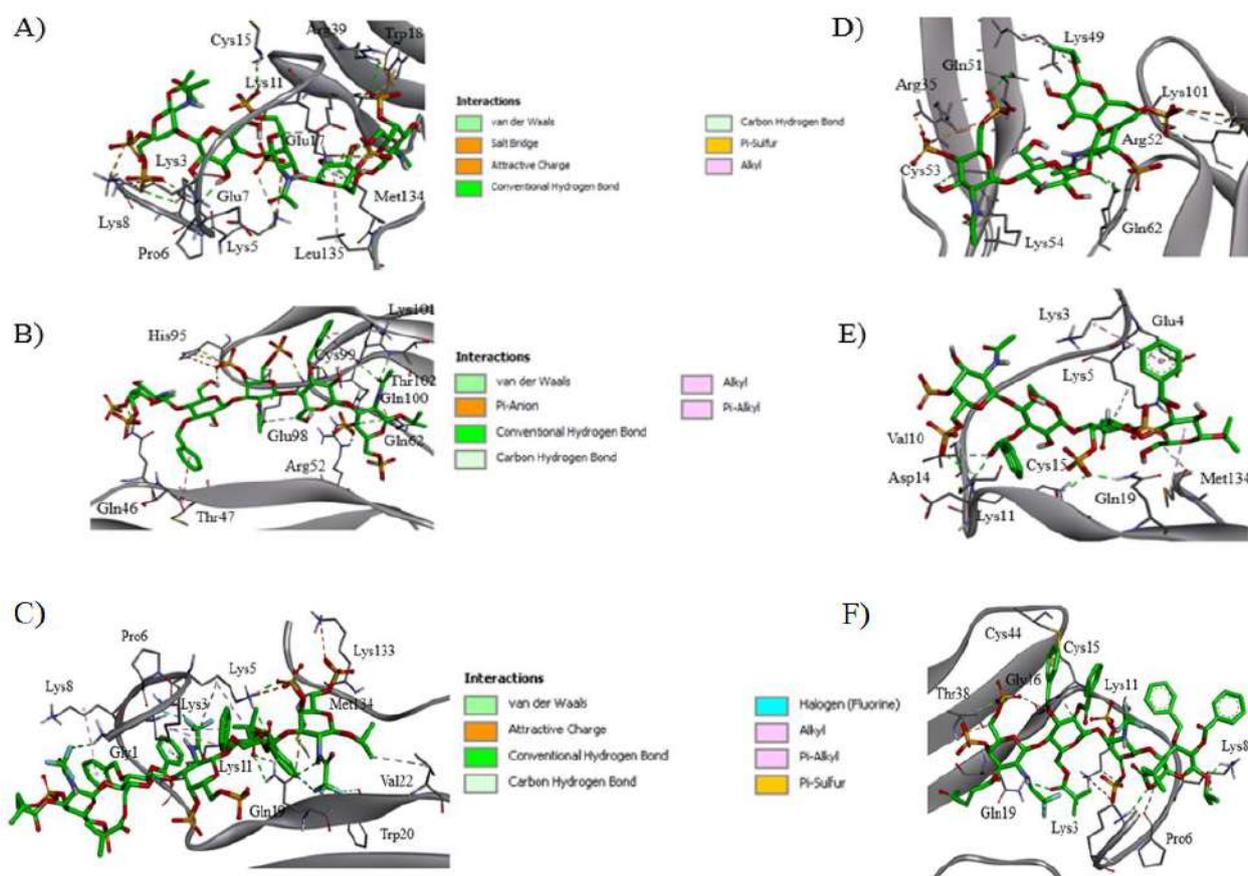


Figura 4.15. A izquierda y derecha se muestran las principales interacciones establecidas entre el conformero más estable de cada ligando y la conformación plegada de la PTN: pentasacárido desprotegido (A), dibencilado (B) y persustituido (C); y tetrasacárido desprotegido (D), dibencilado (E) y persustituido (F). En el centro se señalan los tipos de interacciones más representativas que tienen lugar.

Estos acoplamientos muestran que, en ambas series, existen ciertas interacciones que son comunes en los seis complejos: las establecidas por puentes salinos y atracción de cargas están protagonizadas por lisinas y argininas únicamente, con los oxígenos de los grupos sulfato; los enlaces de hidrógeno convencionales; y las interacciones entre grupos alquílicos.

En las uniones que tienen lugar con los azúcares que contienen grupos protectores en su estructura, aumentan las interacciones con la proteína, estableciéndose nuevos enlaces, ya sea entre sistemas π y grupos alquílicos o residuos cargados, donde los anillos aromáticos pueden pertenecer al residuo de azúcar o de la proteína (Figura 4.16).

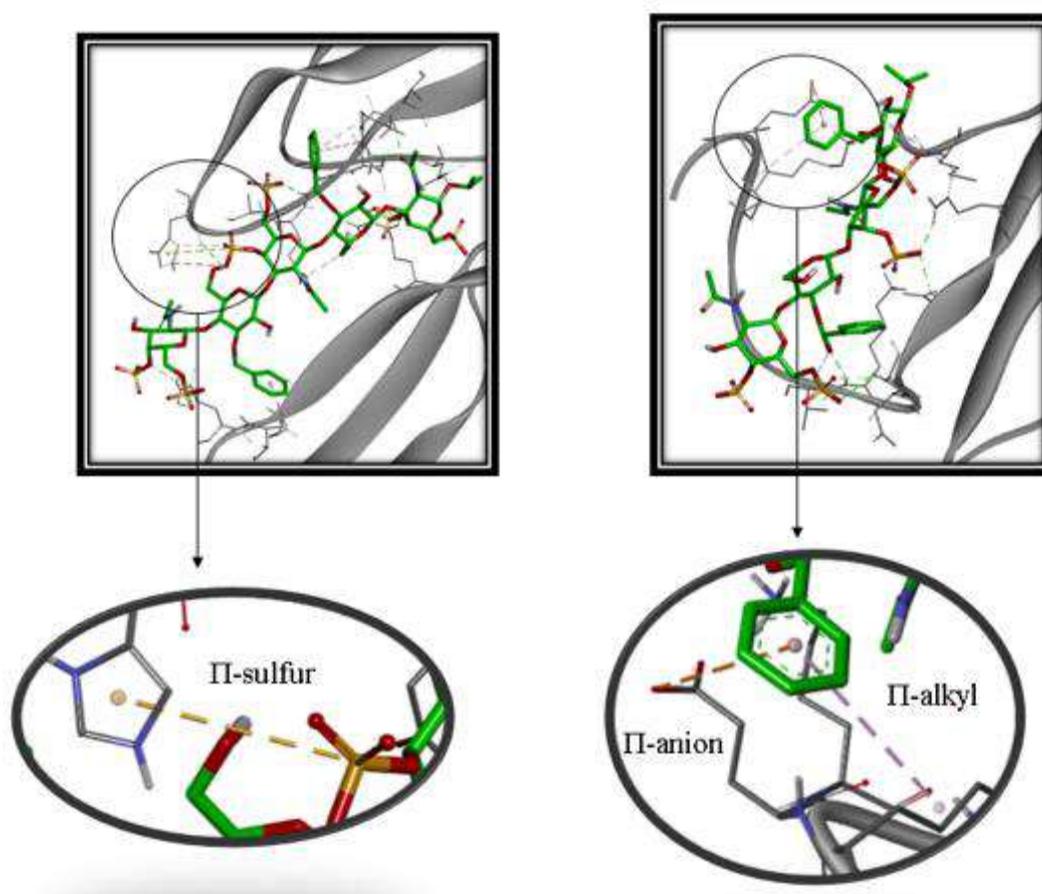


Figura 4.16. Representación tridimensional de los ejemplos de interacciones π con sistemas no aromáticos.

Siguiendo con el análisis de los resultados de *docking*, para el caso en el que la proteína se encuentra en su conformación más abierta, **PTN3**, encontramos un comportamiento diferente de los conformeros dependiendo del ligando. Los complejos dibencilados (**2-PTN3** y **5-PTN3**) y persustituídos (**3-PTN3** y **6-PTN3**), se encuentran distribuidos entre los dos dominios, abarcando prácticamente toda la zona central de la proteína. En cambio, para los compuestos con los grupos hidroxilo sin proteger, la mayoría de los conformeros (complejos **1-PTN3** y **4-PTN3**) muestran de nuevo una predilección hacia un dominio concreto. Si bien cabe destacar que, en el caso de la conformación anterior de la proteína, éstos tendían a unirse sobre el primer dominio de la PTN; en este caso, la interacción se establece preferentemente en el CTD (Figura 4.17).

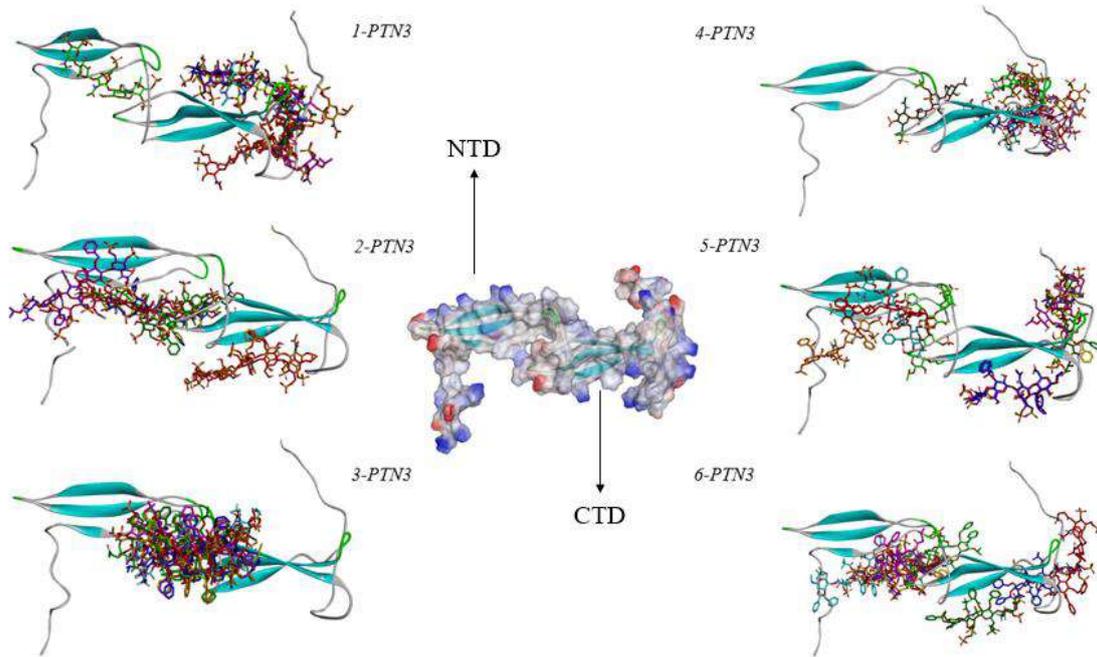


Figura 4.17. Representación tridimensional de los resultados del docking molecular de los seis complejos. Se muestra la superficie de potencial de la PTN en su forma más abierta, y la distribución de los 9 conformeros obtenidos del docking de cada ligando.

En la Figura 4.18 se exponen de nuevo las interacciones por atracción de carga o puente salino entre lisinas y argininas principalmente, junto a las uniones C-H \cdots O, N-H \cdots O e incluso C-O-H \cdots O, y las alquilo-alquilo. Junto a éstas, se identifica un nuevo tipo de unión no convencional π -donor hydrogen que es un tipo de unión establecido entre un hidrógeno aceptor presente en un sistema π y uno dador.

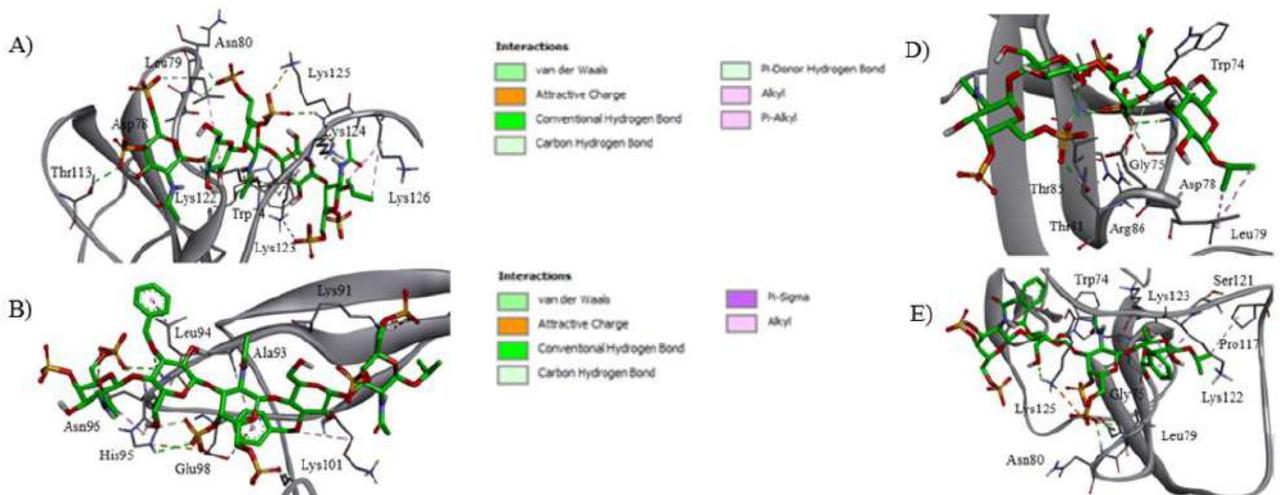




Figura 4.18. A izquierda y derecha se muestran las principales interacciones establecidas entre el conformero más estable de cada ligando y la conformación abierta de la PTN: pentasacárido desprotegido (A), dibencilado (B) y persustituído (C); y tetrasacárido desprotegido (D), dibencilado (E) y persustituído (F). En el centro se señalan los tipos de interacciones más representativas que tienen lugar.

Por otro lado, la introducción de anillos aromáticos en el ligando genera nuevos tipos de interacciones π - π . Éstas se establecen por la interacción de las nubes de anillos aromáticos y pueden tener varias geometrías:

1. *Stacked* o *face to face*, donde la interacción se da entre las dos caras de los anillos aromáticos.
2. *T shaped* o *face to edge*, se produce la interacción del borde de un anillo aromático con la cara del otro.

En la Figura 4.19 se han representado tridimensionalmente un ejemplo para cada situación entre un aminoácido aromático y un grupo bencilo del ligando.

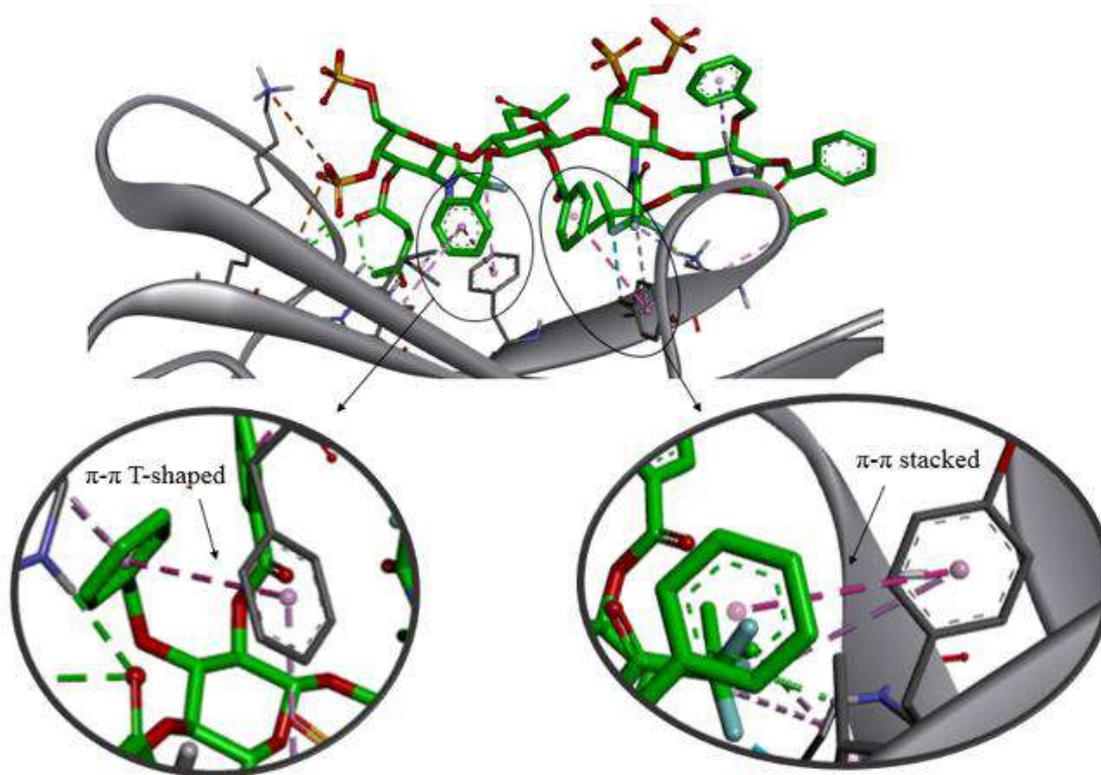


Figura 4.19. Representación tridimensional de los dos ejemplos de interacciones π - π . A la izquierda tenemos la unión π - π T-shaped, y a la derecha π - π stacked para el complejo 6-PTN3.

Siguiendo con el estudio de las interacciones π , se han analizado también aquellas que tienen lugar entre un átomo cargado y el sistema aromático, pudiendo distinguir para este caso entre las π -azufre y las π -anión. La primera de ellas tiene lugar entre un átomo de azufre y una nube π , pudiendo establecerse de frente o de lado con el anillo; mientras que en el segundo tipo de enlace, el átomo que aporta la carga negativa es el aminoácido.

Por otro lado, no debemos olvidar para el caso de los complejos con el ligando completamente protegido las interacciones con halógenos $C-X \cdots B-Y$. El átomo X correspondería a flúor estableciendo interacciones similares a los enlaces de hidrógeno débiles con un átomo de carbono unido a nitrógeno u oxígeno. Concretamente las interacciones flúor con carbono y oxígeno están limitadas a grupos carbonilos tomando como criterio una distancia máxima de 3.7 Å aproximadamente.^[29] Por lo tanto, en este caso, el grupo C-X sería el CF_3 del azúcar que se uniría a los grupos carbonilo de los aminoácidos.

Por último, utilizando la conformación de la PTN en forma de bisagra, la mayoría de los conformeros presentan un comportamiento muy homogéneo, con una preferencia por la zona situada cerca del primer dominio (entre los residuos 15-55). El complejo **5-PTN8** es una excepción, ya que en éste la mayor parte de los ligandos se sitúan mayoritariamente en la región central (Figura 4.20).

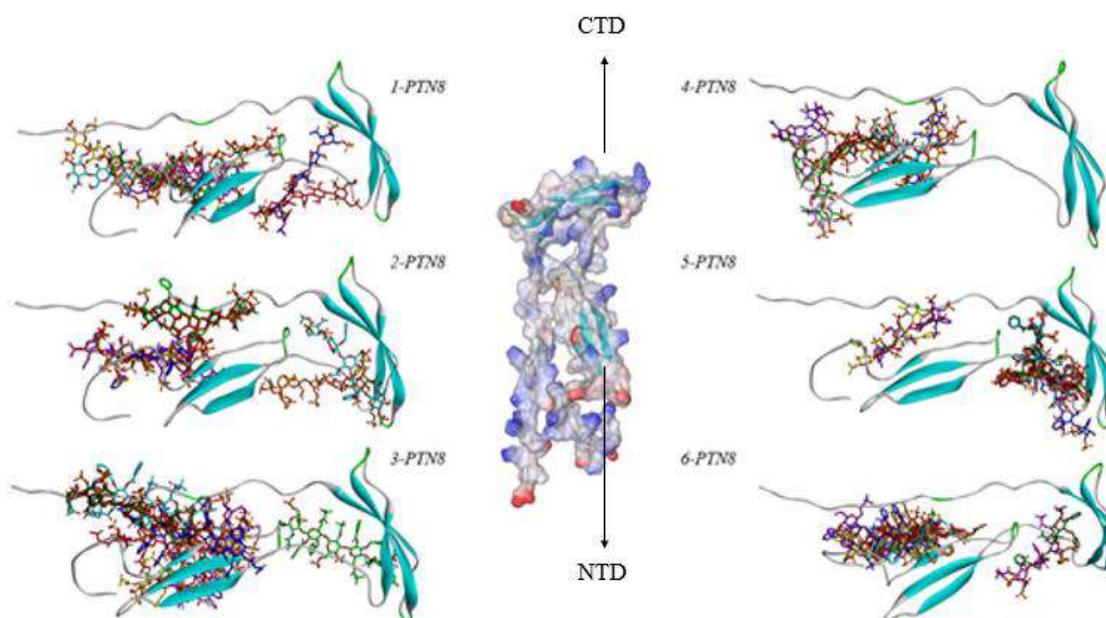


Figura 4.20. A izquierda y derecha se muestran las principales interacciones establecidas entre el conformero más estable de cada ligando y la conformación bisagra de la PTN: pentasacárido desprotegido (A), dibencilado (B) y persustituído (C); y tetrasacárido desprotegido (D), dibencilado (E) y persustituído (F). En el centro se señalan los tipos de interacciones más representativas que tienen lugar.

En esta ocasión aparece una menor variedad de interacciones en el que estén involucrados sistemas aromáticos, pero de nuevo se siguen encontrando para esta conformación enlaces π - π T-shaped, π -catión, π -azufre y π -alquilo, además de las uniones convencionales encontradas en las tres estructuras de la PTN (Figura 4.21). Un dato a añadir es que los enlaces de hidrógeno observados en las tres conformaciones de la PTN ponen en contacto

átomos de las cadenas laterales de la proteína (no implicados en estos puentes) con otros residuos del carbohidrato, estableciendo contactos de Van der Waals en todos los casos.

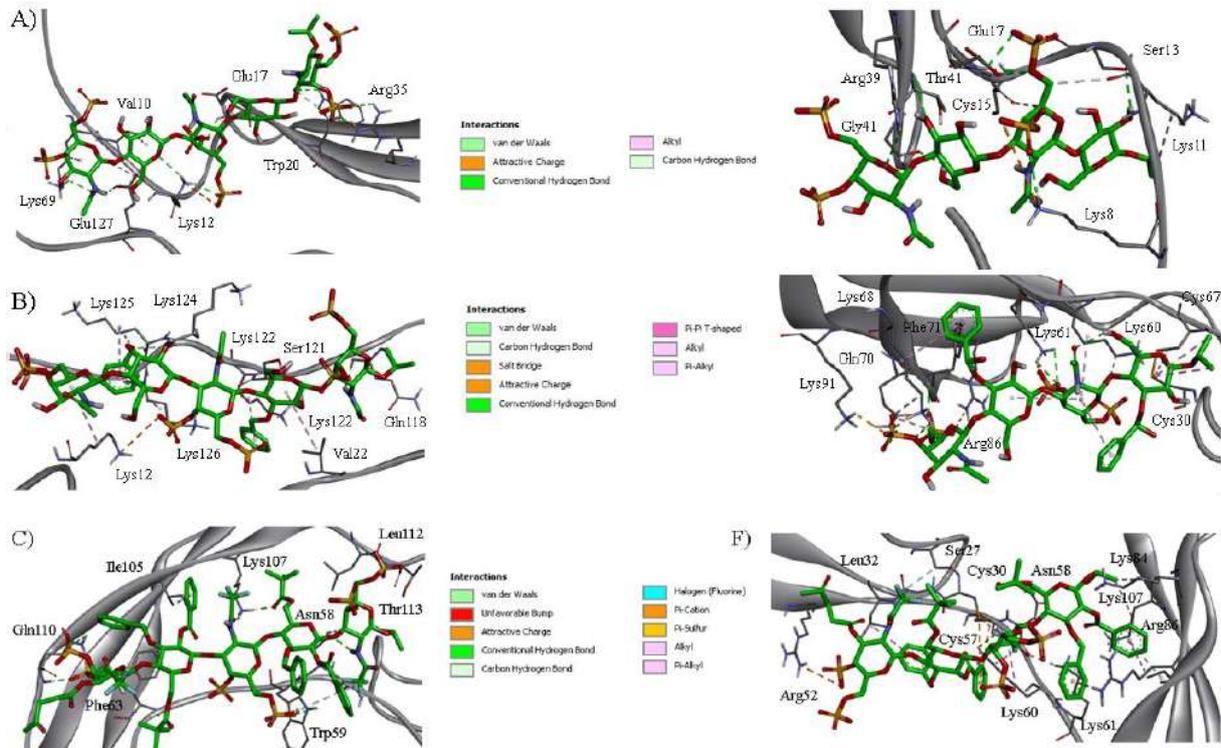


Figura 4.21. A la izquierda y derecha se muestran las principales interacciones establecidas entre el conformero más estable de cada ligando y la conformación horquilla de la PTN: pentasacárido desprotegido (A), dibencilado (B) y persustituído (C); y tetrasacárido desprotegido (D), dibencilado (E) y persustituído (F). En el centro se señalan las interacciones que tienen lugar.

En resumen, los análisis de los seis complejos indican que la interacción PTN-ligando está mediada principalmente por interacciones electrostáticas, siendo los aminoácidos básicos a menudo potenciales sitios de unión a los oligosacáridos. Además, la presencia de grupos hidrofóbicos en la estructura del azúcar incrementa la afinidad con el receptor (de acuerdo con los resultados de polarización de fluorescencia) al establecerse nuevas uniones; en concreto los sistemas aromáticos establecen diversas interacciones π - π . En el Anexo II, se han descrito con más detalle todas estas interacciones observables en los modelos tridimensionales representativos para cada una de las estructuras de la PTN con los conformeros de mayor afinidad.

Cabe añadir que estas estructuras generadas respaldan los resultados de STD expuestos anteriormente, en donde las curvas de crecimiento no mostraban una interacción localizada en el ligando, sino que quedaba distribuida a lo largo de su estructura y por ambas caras.

Como dato importante, en la Figura 4.22 se refleja la variedad de posiciones que puede adoptar el ligando en el complejo. En este caso se representan dos complejos con orientación opuesta respecto al sitio de unión con la proteína. Estos conformeros pueden presentarse enfrentados (paralelos), con orientaciones opuestas (antiparalelos) o perpendiculares. Esta diversidad de posiciones conlleva un amplio abanico de

interacciones por parte del ligando y, debido a que sus energías de afinidad teóricas no difieren demasiado entre las nueve conformaciones seleccionadas, se puede concluir que cualquiera de las disposiciones espaciales del azúcar es igualmente válida para su interacción.

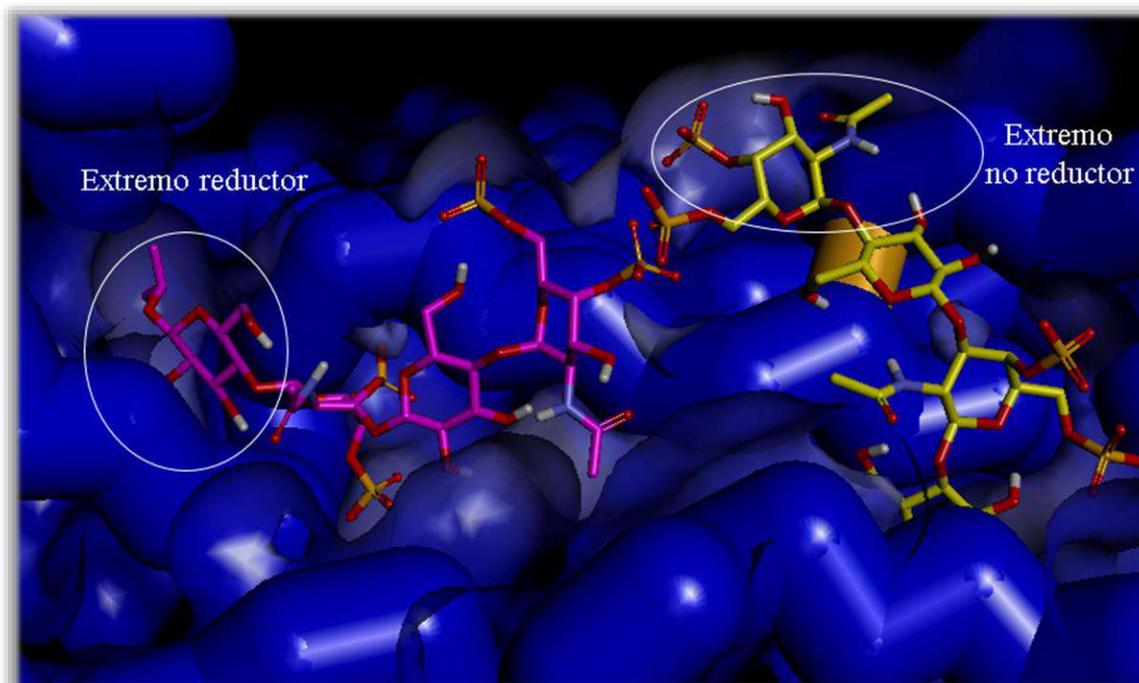


Figura 4.22. Representación tridimensional utilizando como ejemplo dos conformeros obtenidos del docking molecular para el complejo **4-PTN8** mostrándose las diferentes orientaciones del ligando con distintas posibilidades de interacción con la proteína.

4.4.2 Dinámica Molecular

Una vez localizada la zona de mayor interacción entre la PTN y los ligandos **1 – 6** y con el objetivo de mejorar la resolución, se prepararon nuevas estructuras de partida. Éstas se centran en la región comprendida entre el dominio CTD, más concretamente el cluster 1, y la zona *hinge*. Tras la preparación de los modelos de partida se procedió a realizar cálculos con Autodock centrados en esta región.^{[26],[22]} Para ello se emplearon como estructuras de partida las de los carbohidratos libres determinadas en el **Capítulo 3** y tres de las conformaciones representativas de la PTN depositadas en el archivo pdb: PTN4 (estructura plegada), PTN3 (estructura abierta) y PTN8 (estructura bisagra). Para la generación del complejo se utilizó Autodock estableciendo los límites mediante el *grid* englobando los residuos de la zona *hinge* y el *cluster* I del dominio CTD. Mediante un *docking* molecular, se generaron los archivos de partida para realizar los cálculos de dinámica molecular (este procedimiento se detallará en el punto 4.5 de metodología) empleando el primer conformero del docking, ya que es el primero en energía de interacción.

Los archivos *inputs* necesarios para la simulación contienen las coordenadas iniciales del complejo y las parametrizaciones de cada residuo del azúcar. Las topologías y

coordenadas iniciales se crearon mediante el módulo *tLEAP* de AMBER y las dinámicas moleculares se realizaron en disolvente explícito y PME.^[30] El tiempo de cálculo para las dinámicas sin restricciones fue inicialmente de 200 ns, con particiones de 40 ns, y en las MD-TAR de 40 ns con particiones de 8 ns. Todas las etapas fueron analizadas mediante el software *gnuplot*,^[31] representando gráficamente la estabilidad de la simulación a lo largo de toda la trayectoria mediante los archivos de energía creados en cada etapa.

Del mismo modo que en las dinámicas moleculares llevadas a cabo para los ligandos en estado libre, aquellas estructuras que presentaron distorsiones incompatibles con los resultados experimentales se corrigieron empleando cálculos MD-TAR. Para ello, se tomaron como valores de referencia las distancias interprotónicas intraanulares H1 – H3 – H5 e interglucosídicas H1_{n+1} – H3/H4_n obtenidas de los experimentos NOESY transfer. En aquellos casos en los que los resultados seguían discrepando de los valores experimentales se repitieron los cálculos utilizando otro conformero del ligando procedente del *docking*. Se empleó como criterio de elección el siguiente en energía de interacción o aquel que presentase una orientación diferente respecto a la proteína.

La trayectoria de las simulaciones se procesó para obtener la geometría promediada de los complejos. Y de ésta se calcularon valores para comparar con los obtenidos experimentalmente (con valor conformacional, sección 4.3): Análisis de Cremer-Pople de la conformación de cada anillo de azúcar, ángulos de torsión interglucosídicos y distancias interprotónicas.

4.4.2.1 Comparativa entre los complejos 1-PTN y 4-PTN

El análisis teórico dio comienzo por los complejos más simples, con **1** y **4**, aquellos que presentaban los carbohidratos completamente desprotegidos. Se empezó por la conformación más plegada de la PTN (PTN4). Se llevaron a cabo, en primer lugar, dinámicas moleculares libres de restricciones con una duración de 200 ns. Las trayectorias finales mostraron una conformación ⁴C₁ para todos los azúcares en ambos complejos, **1-PTN4** y **4-PTN4**. (Figuras A.2 y A.3 del Anexo). Por otro lado, los mapas de distribución de población de los enlaces glicosídicos son compatibles con un único mínimo absoluto de tipo *syn-ψ/exo-φ*. En este caso, el complejo **4-PTN4** presenta una mayor flexibilidad y los ángulos ψ comprenden valores desde 30° hasta -60°. Del hecho de presentar valores negativos y positivos de ψ podría deducirse que estos enlaces fluctúan entre las dos conformaciones *syn* existentes. En general, las distancias experimentales interglucosídicas, 2.3 Å y 2.6 Å, están de acuerdo con las estructuras predichas. Además, la presencia de dos grupos sulfato menos en el tetrasacárido y, por tanto, la menor repulsión estructural, justificaría esta mayor flexibilidad (Figuras A.4 y A.5 del Anexo).

En la Figura 4.23 se ha representado un fragmento de la dinámica procedente del archivo de trayectoria final, en donde se observa que el ligando se mantiene en el sitio activo de la PTN durante toda la simulación. En lo que respecta a la conformación de la proteína,

mantuvo su plegamiento a lo largo de la trayectoria, aunque en el complejo con el ligando tetrasacárido **4-PTN4**, se observa un despliegue de sus extremos terminales.

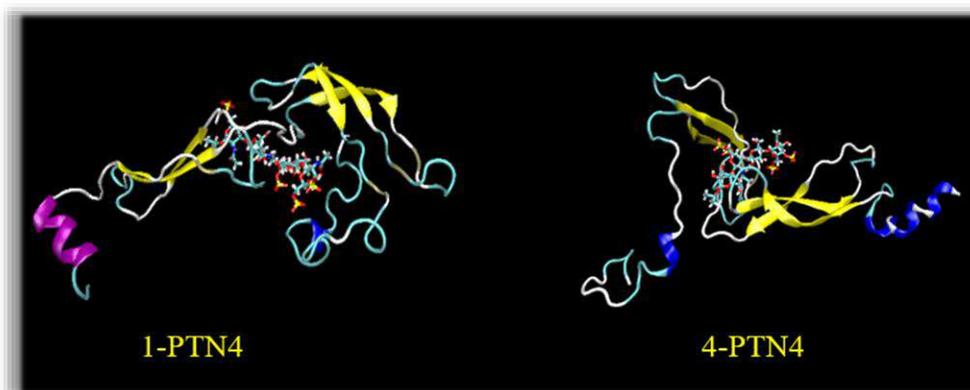


Figura 4.23. A la izquierda, complejo pentasacárido **1-PTN4**, a la derecha complejo tetrasacárido **4-PTN4** tras la simulación de 200 ns.

Siguiendo el mismo protocolo, se realizaron dinámicas moleculares libres partiendo de la estructura lineal. En este caso, se observa que la PTN se pliega sobre el ligando para el complejo con el tetrasacárido (**4-PTN3**), mientras que para el complejo **1-PTN3** sólo se mantuvo plegada en la primera parte de la simulación. En ambos casos el ligando permaneció en el sitio activo de la PTN (Figura 4.24):

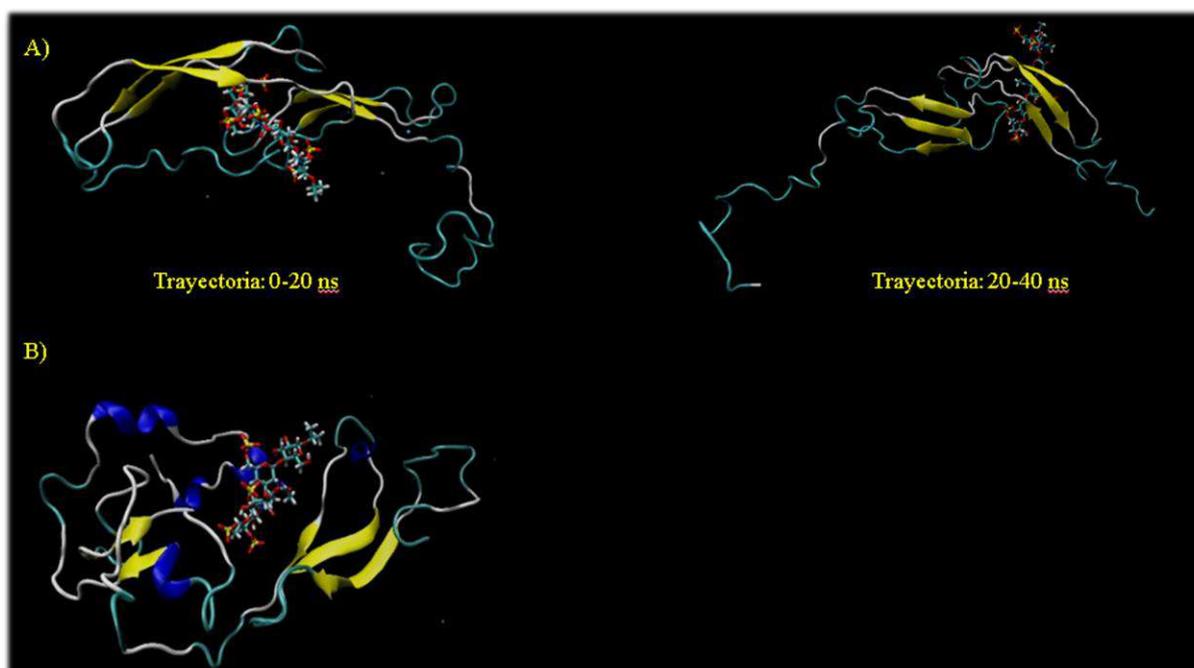


Figura 4.24. A) A la izquierda, primer fragmento de la simulación para el complejo pentasacárido **1-PTN3**, a su derecha, la segunda parte de la dinámica molecular de 200 ns. B) complejo tetrasacárido **4-PTN3** tras la simulación de 200 ns a lo largo de toda la dinámica.

De nuevo, los análisis de ambas trayectorias no mostraron discrepancias estructurales importantes con los datos experimentales. Ambos complejos presentan los ligandos con

conformaciones 4C_1 y enlaces interglicosídicos con geometría *syn-ψ/exo-φ*, aunque para el caso del complejo **4-PTN3** aparece mayor flexibilidad (Figuras A.6-A.8 del Anexo). Cabe añadir que el ángulo ψ para el enlace Glc C-GalNAc D, presenta transiciones esporádicas a conformaciones *anti-ψ* que no son compatibles con los resultados experimentales.

Finalmente, el estudio se completó con la tercera estructura seleccionada de la proteína en forma de horquilla PTN8. La conformación de la proteína permanece estable a lo largo de toda la simulación y el ligando localizado en el sitio de interés (Figura 4.25).

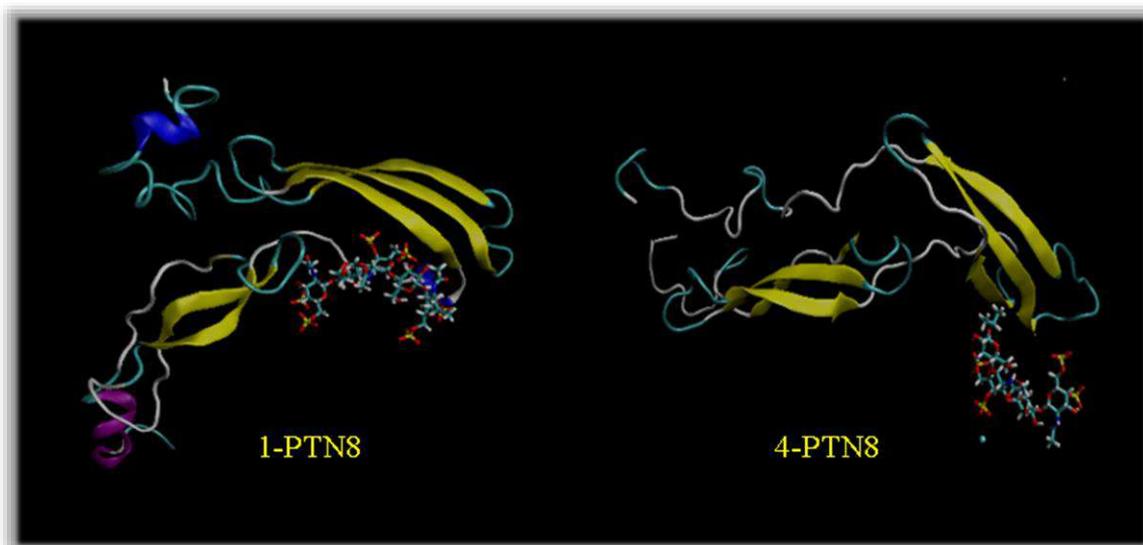


Figura 4.25. En el margen izquierdo, dinámica molecular con restricciones (40 ns) para el complejo pentasacarídico **1-PTN8**, y a su derecha, para el complejo tetrasacarídico **4-PTN8**.

A diferencia de las dinámicas anteriores, los resultados sí que presentaron distorsiones en sus geometrías relativas a los datos experimentales. Se realizaron en ambos casos MD-TAR, pero sólo se consiguió la corrección en el complejo **1-PTN8** tanto para el anillo de Glc D como el ángulo ϕ de éste con la GalNAc C (Figuras A.9, A.10A, A.11 y A.12A del Anexo). En la Figura 4.26 se han representado tridimensionalmente ambas dinámicas mostrando estos resultados.

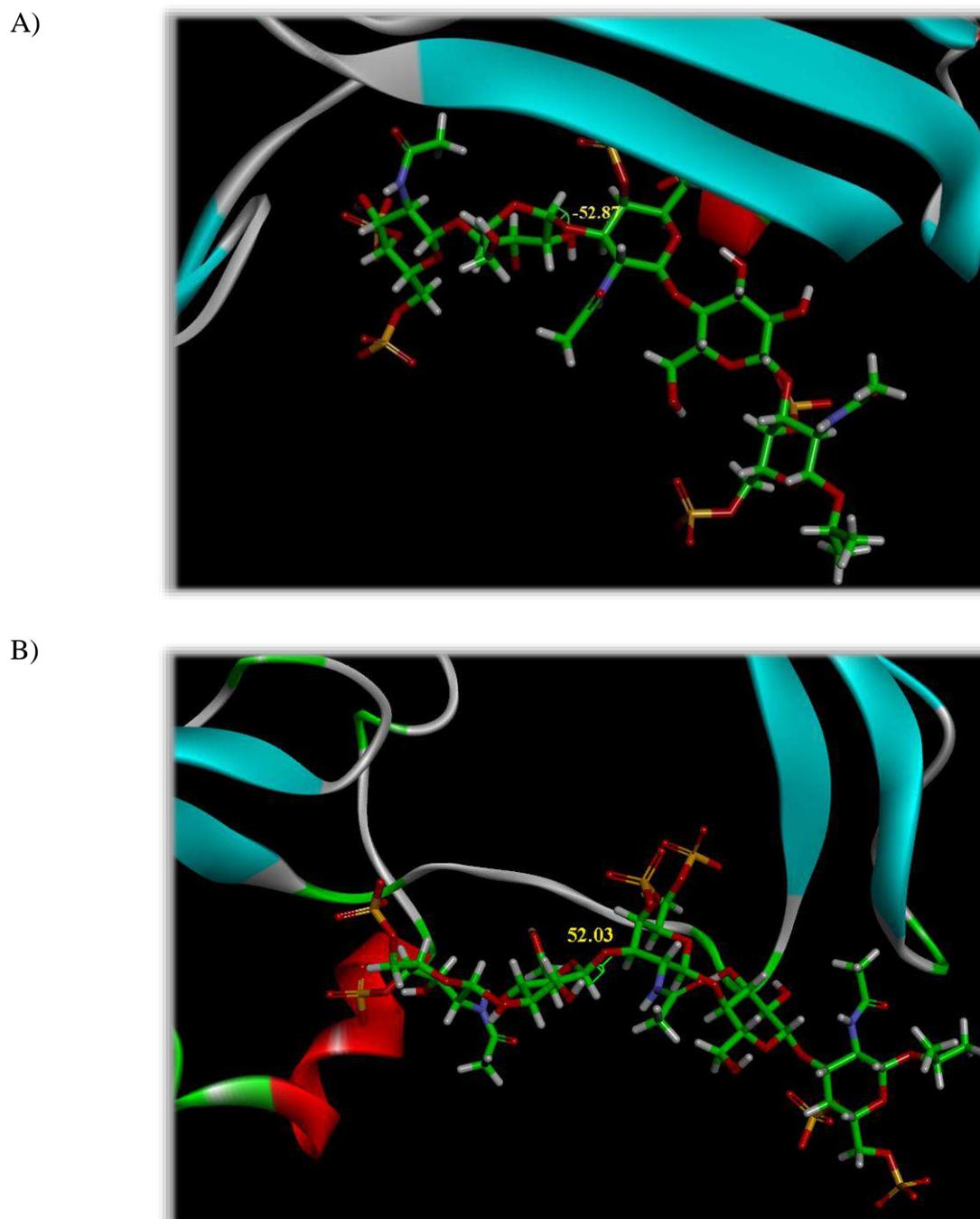
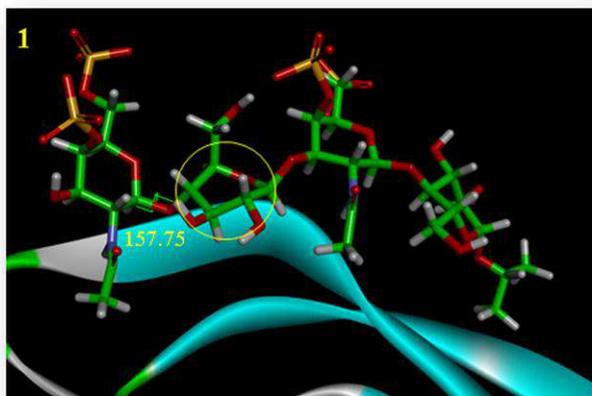


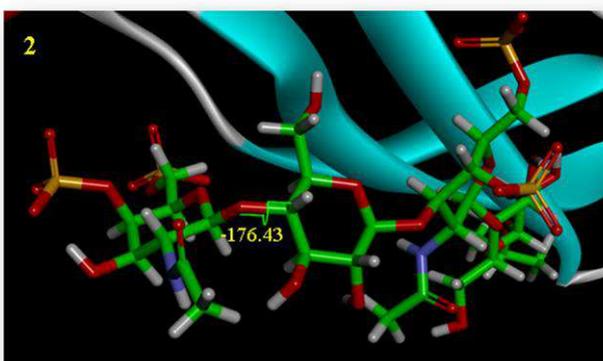
Figura 4.26. A) Distorsión en el anillo de glucosa D del complejo **1-PTN8** y conformación interglucosídica C-D no *exo- ϕ* resultado de la dinámica de 200 ns. B) Corrección de la geometría resultado de la dinámica con restricciones de 40 ns.

Sin embargo, para el complejo **4-PTN8** no se pudieron reproducir los resultados experimentales, ni mediante MD-TAR, ni variando la constante de fuerza de las restricciones. Tampoco se consiguió un resultado óptimo eligiendo el conformero 3 al presentar una orientación contraria al primer ligando y estar situado en otra zona del sitio activo (ver Figura 4.27).



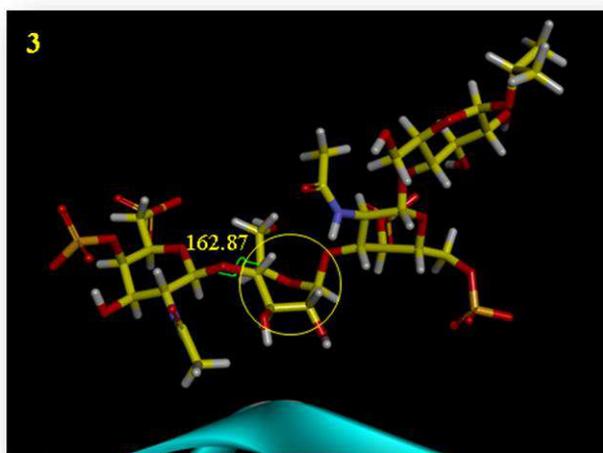
Dinámica molecular libre 200 ns

Confórmero 1



MD-TAR 40 ns

Confórmero 1



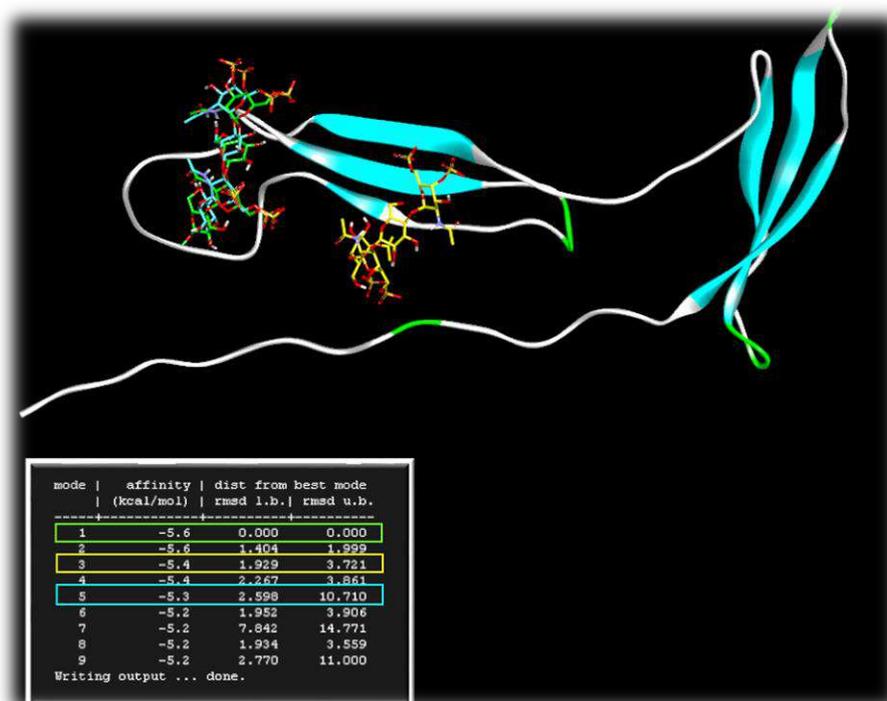
MD-TAR 40 ns

Confórmero 3

Figura 4.27. Representación tridimensional de los resultados de las dinámicas moleculares, tanto libres como con restricciones, para cada uno de los conformeros del complejo **4-PTN8** utilizados en la simulación, mostrando las distorsiones y correcciones que han tenido lugar.

Finalmente se pudo reproducir los resultados experimentales al elegir un nuevo conformero situado en la misma región con orientación contraria (ver Figura 4.28). En este caso, los resultados de la dinámica MD-TAR fueron satisfactorios, mostrando todos los anillos de azúcar conformaciones de silla canónicas 4C_1 , aunque el ángulo *syn-ψ/exo-φ* del enlace Glc C-GalNAc D se mostraba bastante flexible a pesar de las restricciones establecidas (ver Figuras A.10B y A.12B del Anexo).

A)



B)

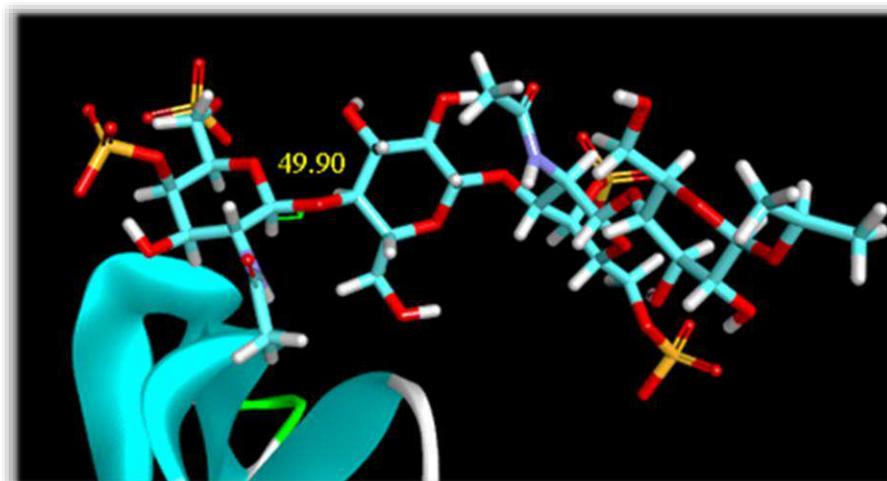


Figura 4.28. A) Representación de los tres ligandos escogidos para las simulaciones llevadas a cabo para el complejo 4-PTN8. En color verde, el primer ligando más estable escogido para la simulación, en amarillo, el segundo más energéticamente favorable y por último en color azul el conformero 5, que dio los resultados más satisfactorios. B) Representación tridimensional al utilizar esta última conformación sin mostrar distorsión alguna y todos los enlaces interglucosídicos en disposición *syn-ψ/exo-φ*.

Sobre los archivos de coordenadas finales de ambos complejos para las tres estructuras de la PTN se calcularon las distancias intrarresiduales H1-H3, H1-H5 y H3-H5, y las interresiduales H1-H3 y H1-H4, además de las constantes de acoplamiento; y se compararon con las medidas experimentales obtenidas en el punto 4.3. Estos valores se

pueden comparar con los obtenidos experimentalmente según la teoría del NOE transferido. Los resultados de distancias interprotónicas indican que la conformación del ligando complejado es la misma que la del estado libre (Capítulo 3), por lo que no es posible determinar si alguno de los tres complejos es predominante. Se asume la formación completa del complejo (o los complejos) por el exceso de ligando. Aunque las constantes de acoplamiento no se deban usar para el análisis de las conformaciones en el complejo, debido a su modo de promediado, en este caso la comparación simultánea de las constantes del mismo anillo indica la presencia de la misma conformación. Por lo tanto, al estar en acuerdo con las distancias intrarresiduales calculadas por NOE transferido se asume que no hay cambio conformacional a nivel de los anillos de los azúcares. En las Tablas 4.7- 4.10 se comparan estos valores, indicando aquellos que denotan la distorsión en la estructura o una disposición anti del enlace interglucosídico.

$^3J_{HH}$ coupling	1-PTN RMN	1-PTN4 MD 200ns	1-PTN3 MD 200ns	1-PTN8 MD 200ns	1-PTN8 MD-TAR 40ns
A1A2	8.8	8.2	8.3	7.9	8.2
A2A3		10.5	10.7	10.4	10.6
A3A4	2.6	2.4	2.2	2.3	2.3
A4A5		1.1	1.2	0.8	1.0
B1B2	8.3	7.8	7.6	7.4	7.5
B2B3	8.7	9.2	8.1	9.3	9.6
B3B4	9.1	8.7	8.5	9.8	9.4
B4B5	9.7	9.6	9.2	9.7	9.7
C1C2	8.9	8.2	7.7	8.1	8.4
C2C3	10.4	10	10.5	10.7	10.8
C3C4		2.3	1.9	1.7	1.7
C4C5		1	0.7	0.7	0.8
D1D2	8.3	7.5	7.5	4.9	7.5
D2D3	8.4	8.4	7.7	9.2	8.8
D3D4	9.1	9.5	8.6	7.8	9
D4D5	9.1	9.7	9.7	3.4	9.7
E1E2	8.9	8.5	8.5	8.2	8.4
E2E3	11	10.3	10.3	10.4	10.5
E3E4	2.6	3.1	3.2	2.5	3.1
E4E5		1.4	1.3	0.9	2.2

Tabla 4.7. Constantes de acoplamiento experimentales y teóricas (a 300 K) para el ligando **I** en presencia de **PTN** calculadas para las 3 conformaciones escogidas de la proteína.

$^3J_{HH}$ coupling	4-PTN RMN	4-PTN4 MD 200ns	4-PTN3 MD 200ns	4-PTN8 MD 200ns Conf.1	4-PTN8 MD 40ns Conf.5
A1A2	7.9	7.7	7.0	7.2	7.6
A2A3	9.3	9.2	9.3	9.1	9.4
A3A4	9.2	8.9	9.9	9.6	9.4
A4A5		9.5	9.7	9.7	9.6
B1B2		8.3	8.5	8.4	8.4
B2B3		10.1	10.8	10.7	10.5
B3B4		2.6	1.7	2.2	2.3
B4B5		0.8	0.9	1.0	1.1
C1C2	7.9	7.7	7.7	5.4	7.8
C2C3	9.7	9.5	9.4	10.0	9.5
C3C4		9.4	9.7	7.7	9.2
C4C5		9.7	9.7	2.7	9.5
D1D2	8.5	8.4	8.4	8.4	8.4
D2D3	10.7	10.4	10.4	10.4	10.4
D3D4	2.6	2.7	2.8	2.7	3.3
D4D5		1.2	1.4	2.0	1.7

Tabla 4.8. Constantes de acoplamiento experimentales y teóricas (a 300 K) para el ligando **4** en presencia de PTN calculadas para las 3 conformaciones escogidas de la proteína.

Distancias	1-PTN RMN	1-PTN4 MD 200ns	1-PTN3 MD 200ns	1-PTN8 MD 200ns	1-PTN8 MD-TAR 40ns
A H1H3	2.6	2.6	2.5	2.7	2.6
A H1H5	2.6	2.6	2.6	2.6	2.6
A H3H5	2.4	2.4	2.4	2.4	2.4
Interg.BH1-AH3	2.2	2.7	2.1	2.5	2.5
B H1H3	2.5	2.6	2.6	2.7	2.6
B H1H5	2.8	2.4	2.5	2.6	2.5
B H3H5	3.1	2.7	2.7	2.5	2.5
Interg.CH1-BH4	2.6	2.4	2.2	2.2	2.1

Tabla 4.9. Distancias experimentales de cada pareja de protones intrarresiduales e interglucosídicos para el ligando **1** en presencia de PTN obtenidas de los NOE calculados y distancias teóricas promediadas de las simulaciones moleculares (a 300 K) para las 3 conformaciones escogidas de la proteína.

Distancias	1-PTN RMN	1-PTN4 MD 200ns	1-PTN3 MD 200ns	1-PTN8 MD 200ns	1-PTN8 MD-TAR 40ns
C H1H3	2.2	2.6	2.7	2.6	2.4
C H1H5		2.6	2.7	2.7	2.5
C H3H5	2.6	2.4	2.3	2.3	2.2
Interg.DH1-CH3	2.6	2.3	2.2	2.8	2.4
D H1H3	3.0	2.9	2.9	2.8	2.7
D H1H5	2.3	2.7	2.6	3.1	2.4
D H3H5		2.5	2.6	3.0	2.6
Interg.EH1-DH4	2.7	2.4	2.3	2.3	2.2
E H1H3	3.1	2.6	2.9	2.6	2.5
E H1H5	2.2	2.4	2.4	2.6	2.6
E H3H5	2.6	2.6	2.6	2.4	2.7

Tabla 4.9. Continuación.

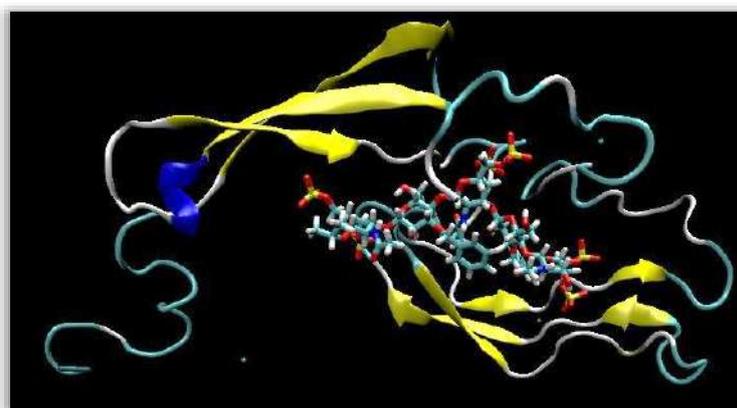
Distancias	4-PTN RMN	4-PTN4 MD 200ns	4-PTN3 MD 200ns	4-PTN8 MD 200ns Conf.1	4-PTN8 MD-TAR 40ns Conf.5
A H1H3	2.4	2.6	2.7	2.8	2.7
A H1H5	2.8	2.5	2.6	2.7	2.6
A H3H5		2.7	2.4	2.4	2.6
Interg.BH1-AH4	2.0	2.3	2.4	2.2	2.3
B H1H3	2.5	2.8	2.4	2.5	2.6
B H1H5	2.9	2.6	2.6	2.7	2.6
B H3H5	2.6	2.4	2.3	2.4	2.4
Interg.CH1-BH3	2.3	2.5	2.2	2.3	2.5
C H1H3	2.8	2.6	2.6	2.8	2.5
C H1H5	2.5	2.7	2.4	4.1	2.6
C H3H5		2.6	2.5	3.6	2.7
Interg.DH1-CH4	2.5	2.2	2.2	2.6	2.4
D H1H3	2.7	2.6	2.5	2.5	2.6
D H1H5	2.6	2.5	2.6	2.6	2.5
D H3H5		2.5	2.6	2.5	2.7

Tabla 4.10. Distancias experimentales de cada pareja de protones intrarresiduales e interglicosídicos para el ligando 4 en presencia de PTN obtenidas de los NOE calculados y distancias teóricas promediadas de las dinámicas moleculares (a 300 K) para las 3 conformaciones escogidas de la proteína.

4.4.2.2 Comparativa entre los complejos 2-PTN y 5-PTN

Para el análisis de los siguientes complejos con el azúcar dibencilado se utilizó la misma metodología anterior, con el objetivo de comparar las diferencias conformacionales que pudieran existir con respecto a los complejos en los que el ligando no presentaba residuos aromáticos. Los resultados teóricos obtenidos de las simulaciones no fueron completamente satisfactorios ya que en ambos compuestos fue necesario llevar a cabo simulaciones MD-TAR para intentar reproducir una estructura acorde a los resultados experimentales. Siguiendo con el mismo protocolo de estudio, se realizaron en primer lugar las simulaciones de los complejos con la conformación más plegada de la PTN, PTN4. Como se puede observar en la Figura 4.29, la visualización de la dinámica molecular libre muestra que la proteína conserva su plegamiento a lo largo de toda la trayectoria. En cambio, al analizar la estructura de los ligandos, se pudieron observar ciertas diferencias.

A)



B)

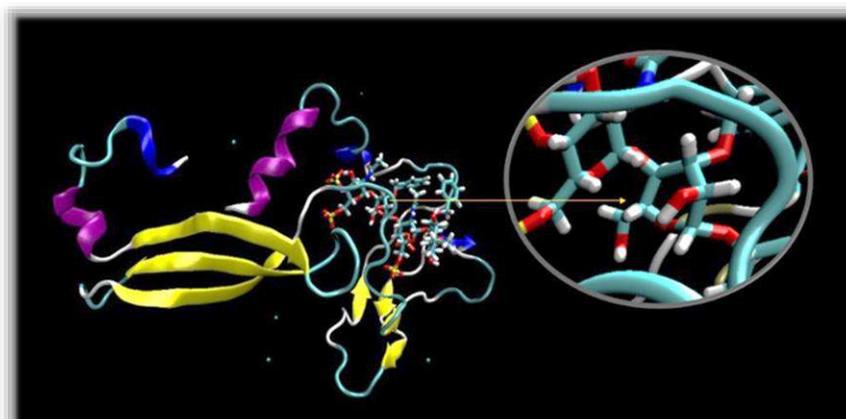


Figura 4.29. Arriba, dinámica molecular sin restricciones (200 ns) para el complejo pentasacarídico 2-PTN4 (A), abajo para el complejo tetrasacarídico 5-PTN4 (B) con una ampliación del anillo distorsionado C.

En el estudio de estos complejos se esperaba encontrar más dificultad en los cálculos por dos motivos: por el aumento del volumen de estos azúcares con respecto a los complejos con el ligando desprotegido; y porque los parámetros para los grupos aromáticos no estaban incluidos en el campo de fuerzas GLYCAM. Aunque el complejo **2-PTN4** presenta un residuo más de azúcar, su geometría no se vio tan afectada como la de su homólogo tetrasacarídico **5-PTN4**, mostrando en la dinámica libre un comportamiento monoconformacional 4C_1 en todos los anillos y enlaces interglicosídicos con una geometría *syn-ψ/exo-φ* principalmente, aunque en la unión Glc B-GalNAc C el ángulo ϕ se aproxima a valores de -60° (Figuras A.14 y A.16 del Anexo). Para el compuesto **5-PTN4**, la Glc C mostró ángulos de θ y ψ próximos a 90° y 330° respectivamente, indicando una conformación tipo bote torcido 0S_2 incompatible con los resultados experimentales (Figura A.15 del Anexo), y los enlaces Glc A-GalNAc B y Glc C-GalNAc D mostraron orientaciones tipo *anti-ψ* como se puede ver en la Figura 4.30.

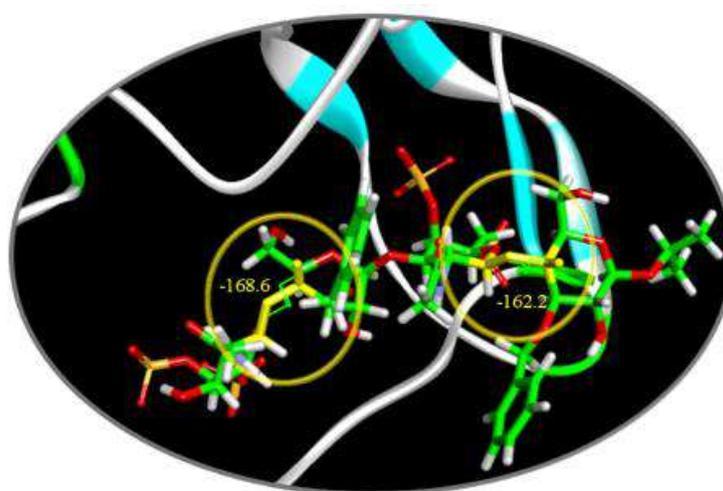


Figura 4.30. Conformaciones interglicosídicas *exo-φ*, *anti-ψ* para el complejo **5-PTN4** resultado de la dinámica de 200 ns.

Para conseguir una estructura que reprodujese los resultados experimentales fue necesario usar el método MD-TAR. Adicionalmente, hubo que aumentar el valor de la constante de fuerza K de 30 a 50 $\text{Kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$. A pesar de corregirse la conformación de la Glc C, la GalNAc B adoptó conformaciones de silla y bote torcido (valores de θ entre $0-30^\circ$ y 90°) mientras que los enlaces interglicosídicos mostraron una geometría *syn-ψ/exo-φ* (A.17 del Anexo).

Siguiendo con el estudio para la forma abierta de la PTN (Figura 4.31), el complejo **2-PTN3** mantenía la conformación 4C_1 en todos sus anillos y una geometría *syn-ψ/exo-φ* en una MD sin restricciones (Figuras A.18 y A.22 del Anexo). Para el complejo **5-PTN3** fue necesario utilizar otra conformación de ligando, al no poder corregir las distorsiones en las glucosas, que se mostraban en una conformación de bote torcido 0S_2 aun estableciendo restricciones (Figuras A.19 y A.20 del Anexo).

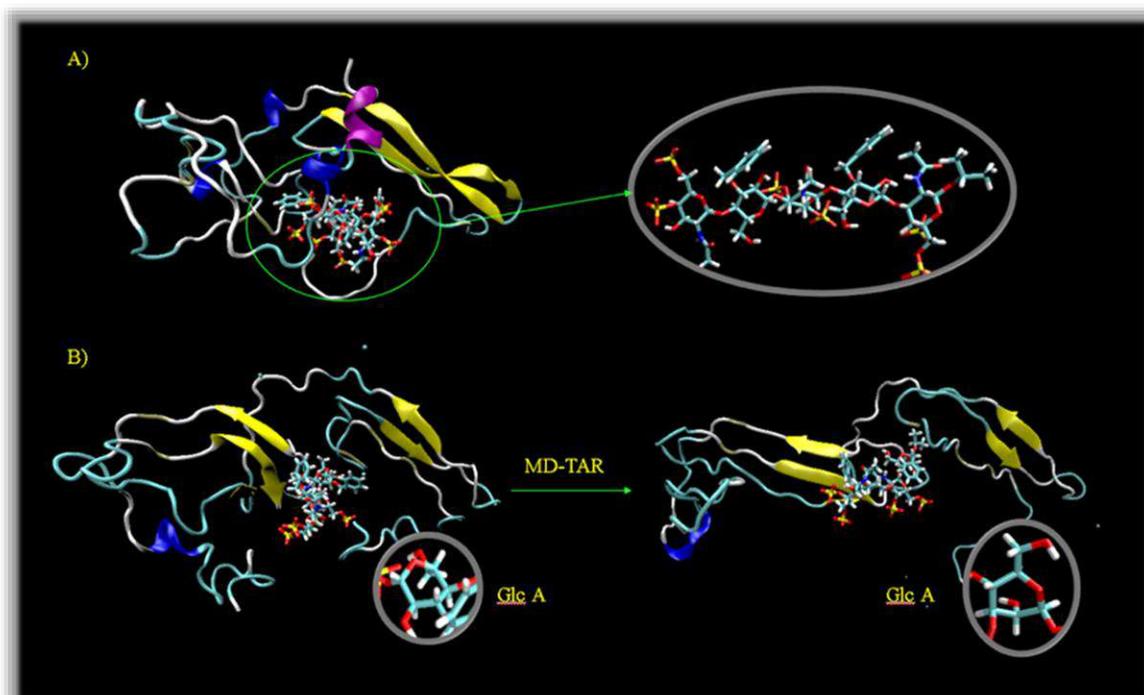


Figura 4.31. (A) Dinámica molecular (200 ns) para el complejo pentasacarídico **2-PTN3**. (B) A la izquierda, dinámica de 200 ns, a la derecha MD-TAR para el complejo tetrasacarídico **5-PTN3** con el primer conformero.

Debido a las discrepancias con los resultados experimentales en el compuesto **5-PTN3** se recurrió a cambiar de conformero del ligando, escogiendo para ello el segundo en energía de interacción, por sus diferencias en orientación y localización con respecto al primero (ver Figura 4.32). La MD-TAR con las mismas restricciones que la anterior corrige en gran parte la distorsión del anillo de Glc A (a excepción de unos pocos *frames*) y la distorsión del mapa (ψ/ϕ) en el enlace GalNAc C-Glc D (ver A.21 y A.23 del Anexo).

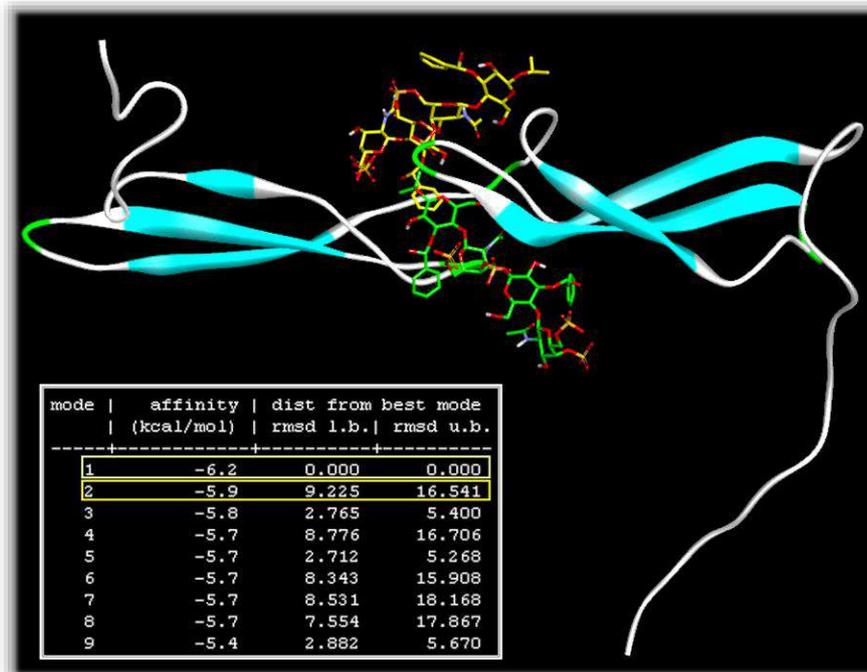


Figura 4.32. Representación de los dos conformeros utilizados para las dinámicas moleculares llevadas a cabo para el complejo 5-PTN3. En color verde, el primer ligando más estable escogido para la simulación, en amarillo, el segundo más energéticamente favorable.

El estudio de los complejos con el ligando dibencilado finaliza con la PTN en forma de bisagra. Hubo que recurrir a restricciones MD-TAR para ambos complejos. De la observación de las estructuras en las trayectorias finales de estas simulaciones (Figura 4.33) se concluye que, aunque la proteína conserva su forma, lo hace en menor grado que en las dinámicas anteriores.

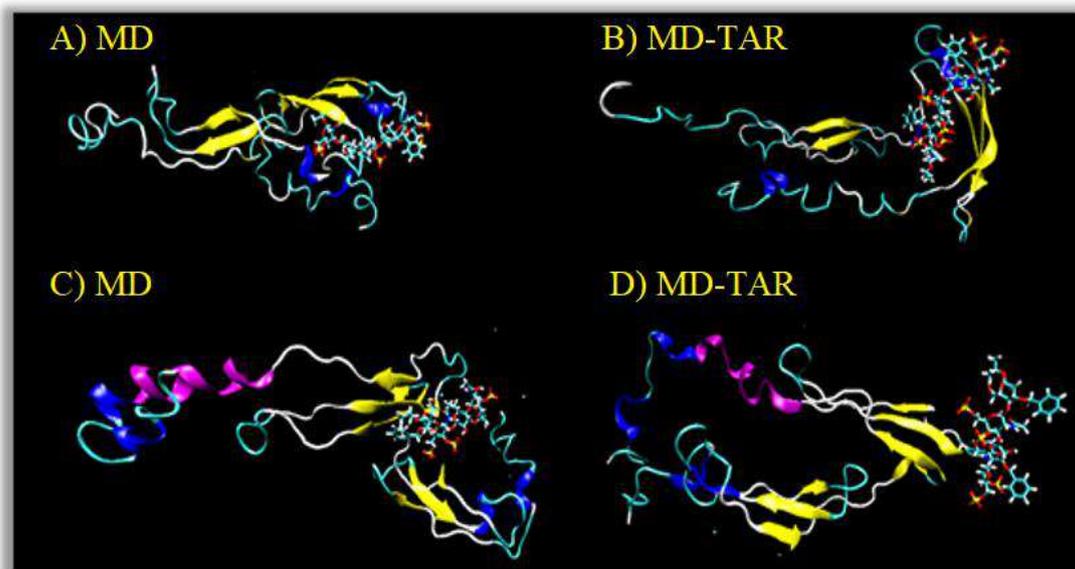


Figura 4.33. Arriba: Dinámicas moleculares libres (200 ns) y con restricciones (40 ns) para el complejo pentasacarídico 2-PTN8 (A y B respectivamente). Abajo: para el complejo tetrasacarídico 5-PTN8, representación tanto de la dinámica molecular de 200 ns como la MD-TAR de 40 ns (C y D respectivamente).

Aunque las coordenadas de *puckering* no mostraron ninguna distorsión en los anillos de azúcar para el complejo pentasacárido **2-PTN8**, se realizó MD-TAR para intentar corregir la conformación *anti-ψ* del enlace Glc B-GalNAc C (Figura 4.34). Aunque el resultado de la MD-TAR corrigió esta orientación y, por lo general, se observaron poblaciones con una geometría *syn-ψ/exo-φ*, ciertos valores del ángulo ϕ se mostraron cercanos a -60° para el enlace GalNAc A-Glc B en contra del efecto exoanomérico (Figuras A.24 y A.26 del Anexo).

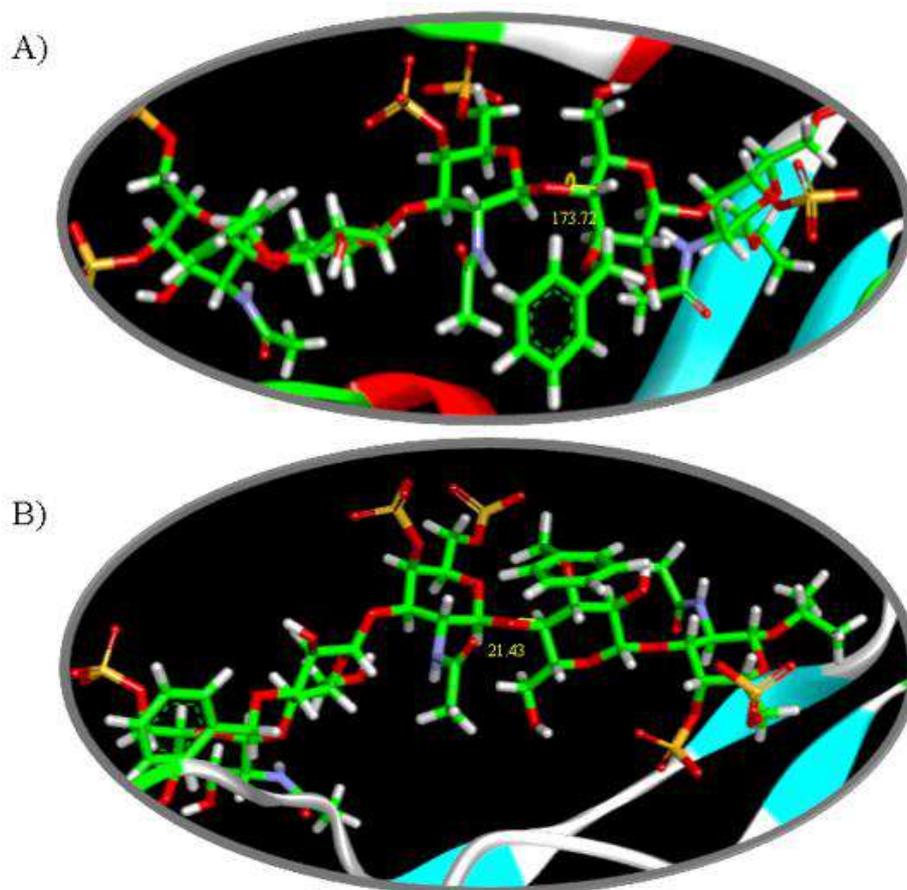
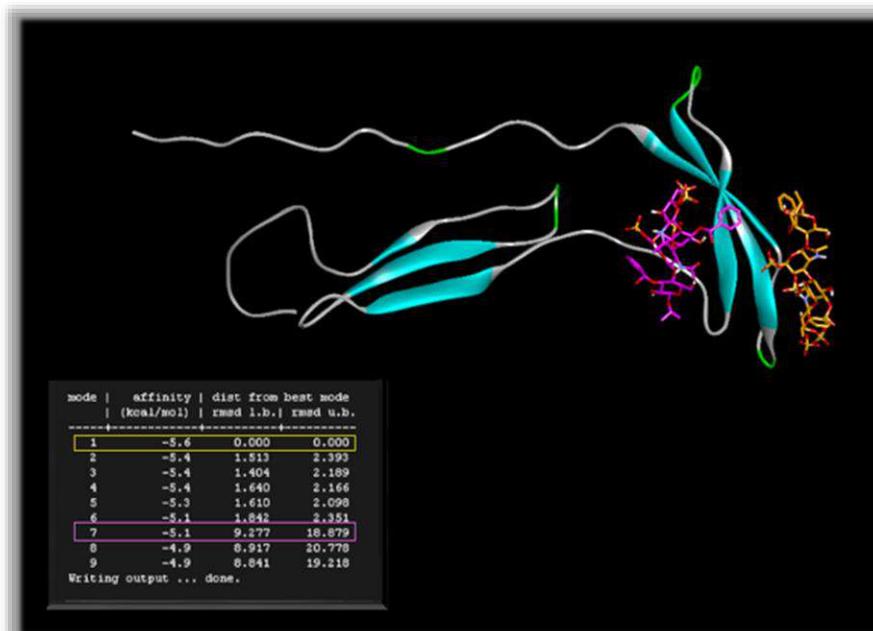


Figura 4.34. (A) Representación tridimensional del enlace interglicosídico en disposición *anti-ψ* en el complejo **2-PTN8** para una dinámica molecular de 200 ns. (B) Corrección del enlace tras aplicar restricciones (MD-TAR 40 ns).

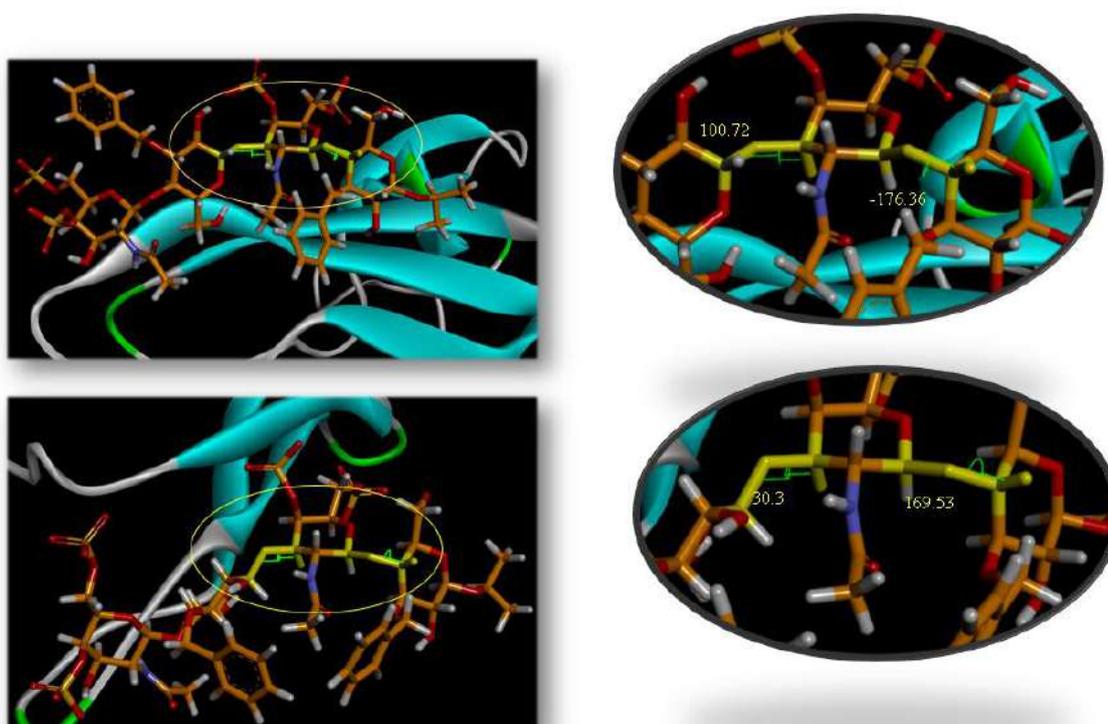
La MD de 200 ns de su homólogo tetrasacárido **5-PTN8** mostró una conformación de bote torcido 0S_2 ($\theta = 90^\circ$) para la Glc A, incompatible con los resultados experimentales, aunque sus enlaces interglicosídicos se encontraron en una orientación *syn-ψ/exo-φ*. Para este caso se repitieron varias simulaciones MD-TAR aumentando el parámetro de la constante de fuerza K de 30 a 50 $\text{Kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ sobre el residuo de Glc A. A pesar de corregir la conformación a una silla 4C_1 , en el enlace interglicosídico Glc A-GalNAc B se encontraron tres conformaciones predominantes de baja energía. Siguiendo el mismo protocolo empleado anteriormente, se usó como alternativa el conformero 7,

consiguiendo resultados compatibles a los datos experimentales. En la Figura 4.35 se muestra un resumen de las estructuras tridimensionales obtenidas de las dinámicas MD-TAR llevadas a cabo para ambos conformeros de azúcar (Figuras A.25, A.27, A.28 y A.29 del Anexo).

A)



B)



C)

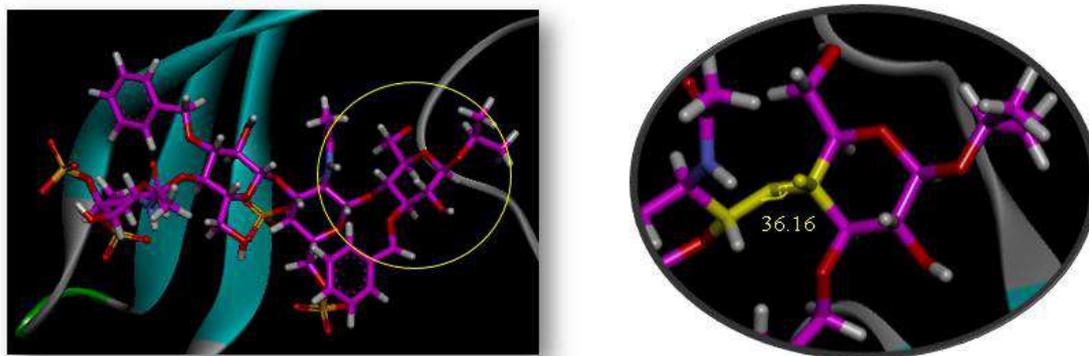


Figura 4.35. A) Representación de los dos conformeros utilizados para las dinámicas moleculares llevadas a cabo para el complejo 5-PTN8. B) Arriba, estructura tridimensional para la dinámica MD-TAR (40 ns) con restricciones en el anillo de glucosa A únicamente. Abajo, repetición de la simulación con restricción además en el enlace AB. C) Repetición MD-TAR utilizando el conformero 7 con corrección de la Glc A y enlace interglucosídico A-B.

A continuación, se exponen los valores de distancias y constantes de acoplamiento teóricas obtenidas de las dinámicas moleculares, contrastándolas con los resultados que se obtuvieron por RMN. A pesar de las repeticiones de los cálculos realizados en esta ocasión, se puede afirmar que se ha conseguido reproducir una estructura tridimensional que apoya los datos experimentales (Tablas 4.11- 4.14).

$^3J_{HH}$ coupling	2-PTN RMN	2-PTN4 MD 200ns	2-PTN3 MD 200ns	2-PTN8 MD 200ns	2-PTN8 MD-TAR 40ns
A1A2	8.6	8.4	8.3	8.5	8.5
A2A3	10.1	10.5	10.8	10.6	10.6
A3A4	2.8	2.2	1.6	2.4	2.1
A4A5		1	0.9	1.2	1
B1B2	8.1	7.8	7.4	7.5	7.7
B2B3	8.6	9.4	9.4	9.8	9.3
B3B4	9.1	8.5	9.1	9.7	7.8
B4B5	7.2	9.5	9.8	9.6	8.9
C1C2		8.4	8.5	8.4	7.3
C2C3	10	10.4	10.6	10.7	10.2
C3C4	2.4	1.7	2.5	2.4	2
C4C5		0.8	1.1	1.2	1.6

Tabla 4.11. Constantes de acoplamiento experimentales y teóricas (a 300 K en MD con y sin restricciones) para el complejo 2-PTN calculadas para las 3 conformaciones escogidas de la proteína.

$^3J_{HH}$ coupling	2-PTN RMN	2-PTN4 MD 200ns	2-PTN3 MD 200ns	2-PTN8 MD 200ns	2-PTN8 MD-TAR 40ns
D1D2	8.1	7.7	7.6	7.6	7.6
D2D3	6.8	8.5	9.9	9.8	9.6
D3D4	9.5	9.8	9.5	9.1	8.9
D4D5		8.8	9.7	9.7	9.5
E1E2	8.2	8.4	8.5	8.5	8.4
E2E3		10.5	10.2	10.4	10.3
E3E4	2.7	2.9	3.1	3.0	3.2
E4E5		1.5	1.4	1.5	1.5

Tabla 4.11. Continuación.

$^3J_{HH}$ coupling	5-PTN RMN	5-PTN4 MD 200ns	5-PTN4 MD-TAR 40ns	5-PTN3 MD 200ns	5-PTN3 MD-TAR 40ns	5-PTN3 MD-TAR 40ns Conf.2	5-PTN8 MD 200ns	5-PTN8 MD-TAR 40ns Conf.7
A1A2	7.9	7.6	7.3	7.3	1.2	7.6	1.8	7.5
A2A3	9.2	9.7	9	8.8	3.5	8.9	2.6	8.4
A3A4	9.2	9.5	8.1	8.3	0.7	7.6	2.2	7.3
A4A5	7.9	7.2	6.7	7.6	6.0	7.1	2.9	7.0
B1B2		8.5	8.4	8.5	7.8	8.4	8.4	8.2
B2B3	10.8	10.6	10.7	10.8	10.8	10.6	10.7	10.5
B3B4	2.9	1.9	2.2	2.5	3.0	2.2	2.5	2.3
B4B5	1.4	0.8	1.1	1	1.9	1.1	1.5	1.2
C1C2	8.0	1.2	7.6	3.8	6.7	7.9	6.9	7.3
C2C3	8.3	1.7	9.8	10.1	8.8	9.4	9.1	9.3
C3C4	9.4	2.5	8.9	9.3	9.3	9.2	9.8	9.7
C4C5	9.4	9.7	9.2	4.9	9.7	9.5	9.6	9.6
D1D2	8.0	8.4	8.4	8.5	8.4	8.5	8.4	8.3
D2D3	10.3	10.4	10.3	10.2	10.2	10.3	10.6	10.4
D3D4		2.8	3	2.6	2.6	2.9	2.8	2.5
D4D5		1.4	1.5	0.6	0.6	1.4	1.7	1.3

Tabla 4.12. Constantes de acoplamiento experimentales y teóricas (a 300 K en MD con y sin restricciones) para el complejo 5-PTN calculadas para las 3 conformaciones escogidas de la proteína.

Distancias	2-PTN RMN	2-PTN4 MD 200ns	2-PTN3 MD 200ns	2-PTN8 MD 200ns	2-PTN8 MD-TAR 40ns
A H1H3	2.6	2.6	2.6	2.6	2.6
A H1H5	2.7	2.6	2.8	2.7	2.7
A H3H5		2.4	2.3	2.5	2.4
Interg.BH1-AH3	2.5	2.3	2.3	2.5	2.2
B H1H3	2.5	2.7	2.7	2.6	2.7
B H1H5	2.6	2.5	2.5	2.6	2.4
B H3H5	2.3	2.7	2.6	2.4	2.8
Interg.CH1-BH4	2.2	2.5	2.6	3.7	2.5
C H1H3		2.6	2.5	2.8	2.8
C H1H5		2.7	2.4	2.7	2.9
C H3H5		2.3	2.4	2.4	2.4
Interg.DH1-CH3	2.2	2.3	2.2	2.6	2.2
D H1H3	2.6	2.6	2.7	2.7	2.7
D H1H5	2.6	2.6	2.6	2.5	2.5
D H3H5	2.5	2.8	2.5	2.6	2.7
Interg.EH1-DH4	2.8	2.4	2.4	2.5	2.6
E H1H3	2.4	2.5	2.6	2.6	2.6
E H1H5	2.7	2.4	2.6	2.5	2.5
E H3H5	2.4	2.5	2.5	2.5	2.6

Tabla 4.13. Distancias experimentales de cada pareja de protones intrarresiduales e interglucosídicos para el complejo **2-PTN** obtenidas de los NOE calculados y distancias teóricas promediadas de las simulaciones moleculares para los conformeros de mínima energía (a 300 K en MD con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

Distancias	5-PTN RMN	5-PTN4 MD 200ns	5-PTN4 MD-TAR 40ns	5-PTN3 MD 200ns	5-PTN3 MD-TAR 40ns	5-PTN3 MD-TAR 40ns Conf.2	5-PTN8 MD 200ns	5-PTN8 MD-TAR 40ns Conf.7
A H1H3	2.5	2.6	2.8	2.8	3.8	2.7	4.1	2.7
A H1H5	2.8	2.5	2.6	2.4	3.2	2.8	2.8	2.5
A H3H5		2.7	2.8	2.8	3.2	2.9	3.5	2.7
Interg.BH1-AH4		3.6	2.5	2.5	2.4	2.4	2.3	2.5
B H1H3		2.6	2.6	2.5	2.6	2.6	2.6	2.7
B H1H5		2.6	2.6	2.6	2.6	2.7	2.7	2.6
B H3H5	2.6	2.3	2.4	2.4	2.4	2.4	2.6	2.4
Interg.CH1-BH3	2.3	2.3	2.4	2.4	2.5	2.4	2.6	2.4

Tabla 4.14. Distancias experimentales de cada pareja de protones intrarresiduales e interglucosídicos para el complejo **5-PTN** obtenidas de los NOE calculados y distancias teóricas promediadas de las simulaciones moleculares para los conformeros de mínima energía (a 300 K en MD con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

Distancias	5-PTN	5-PTN4	5-PTN4	5-PTN3	5-PTN3	5-PTN3	5-PTN8	5-PTN8
	RMN	MD 200ns	MD-TAR 40ns	MD 200ns	MD-TAR 40ns	MD-TAR 40ns Conf.2	MD 200ns	MD-TAR 40ns Conf.7
C H1H3	2.7	4.1	2.6	3.1	2.5	2.6	2.9	2.5
C H1H5	2.5	2.8	2.6	4.1	2.5	2.5	2.6	2.5
C H3H5	2.7	3.3	2.7	3.2	2.7	2.6	2.5	2.6
Interg.DH1-CH4	2.8	3.7	2.4	3.2	2.4	2.4	2.2	2.5
D H1H3	2.8	2.5	2.6	2.4	2.6	2.5	2.4	2.6
D H1H5	2.3	2.5	2.5	2.4	2.6	2.6	2.5	2.5
D H3H5		2.9	2.5	2.5	2.5	2.5	2.5	2.5

Tabla 4.14. Distancias experimentales de cada pareja de protones intrarresiduales e interglucosídicos para el complejo **5-PTN** obtenidas de los NOE calculados y distancias teóricas promediadas de las simulaciones moleculares para los conformeros de mínima energía (a 300 K en MD con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

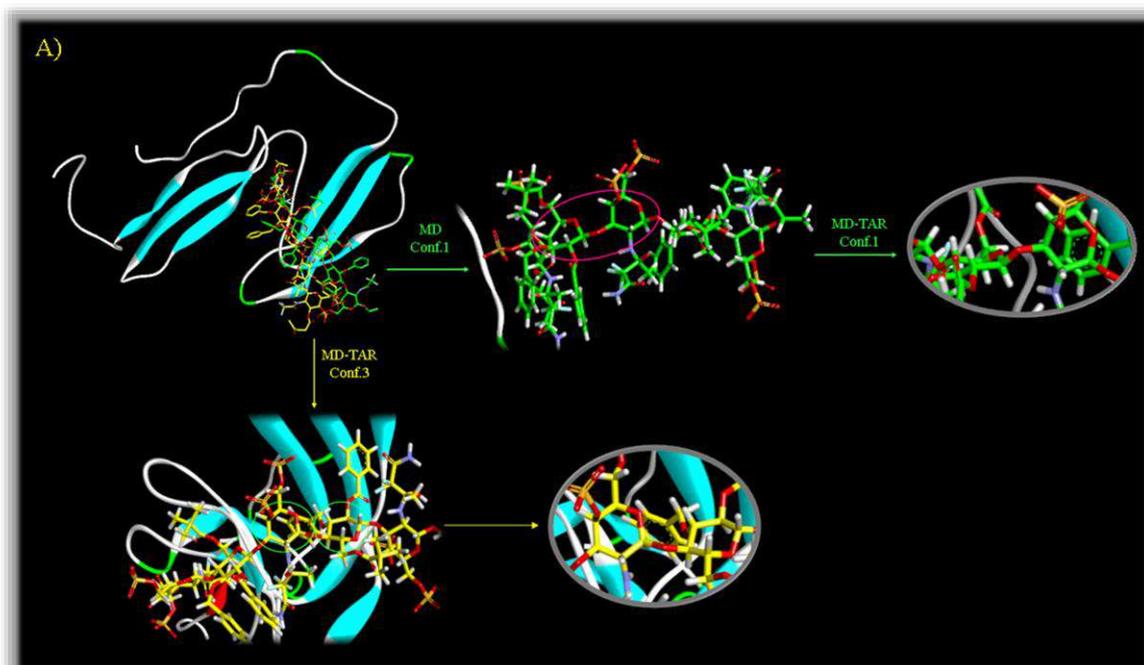
4.4.2.3 Comparativa entre los complejos **3-PTN** y **6-PTN**

En último lugar, han sido analizados los dos complejos con los ligandos completamente protegidos con la misma metodología descrita hasta ahora. A nivel experimental sólo disponíamos de datos del complejo con el mimético pentasacárido **3-PTN** (Figura A.30 del Anexo) ya que su homólogo **6** resultó ser insoluble en el tampón, probablemente debido a la presencia de dos grupos sulfato menos, en comparación con **3**.

En primer lugar, se realizaron dinámicas moleculares sin restricciones para verificar si éstas predecían correctamente la estructura tridimensional de **3**. Debido al volumen de los sustituyentes introducidos en el azúcar, éste podría distorsionarse más fácilmente tras la unión al receptor. Comenzando con el análisis de las trayectorias finales de ambos complejos para la estructura más plegada de la PTN (PTN4), se observaron distorsiones en algunos de los anillos. En el caso del complejo pentasacárido **3-PTN4**, las representaciones de *puckering* mostraron claras deformidades en los anillos C y D, mientras que para su homólogo **6-PTN4**, se encontraron distorsiones en los azúcares A y D, todas ellas incompatibles con los resultados experimentales.

En el caso del complejo **3-PTN4**, las coordenadas de *puckering* indican la presencia de conformaciones localizadas en el ecuador de la esfera, denotando la presencia de conformaciones tipo E, H, B o S (ver Figura A.31 y A.33 del Anexo). Para el complejo **6-PTN4**, la Glc A es la que presenta una distorsión 5S_1 , y el anillo D oscila entre conformaciones de bote ($B_{2,5}$ o ${}^1{}^4B$), bote torcido (1S_5 o 1S_3) y silla 4C_1 (ver Figura A.35 y A.36 del Anexo). En este caso, para ninguno de los dos complejos se predijo la conformación experimental. Con respecto a las geometrías de los enlaces interglucosídicos, ambos compuestos presentaron una geometría *syn-ψ/exo-φ*.

A la vista de estos resultados se realizaron las correspondientes MD-TAR para ambos complejos. Para el compuesto **6-PTN4** se utilizaron como restricciones las distancias que se midieron en el espectro de NOESY del estado libre pertenecientes al ligando **3**, ya que no disponíamos de datos experimentales. (Tablas 5 y 11 del **Capítulo 3**). En la Figura 4.36 se resumen las simulaciones realizadas en ambos casos con las distorsiones y sus respectivas correcciones (Figuras A.32A, A.34A para el complejo **3-PTN4** y A.37, A.39A para el complejo **6-PTN4**). Al no obtener resultados satisfactorios con la primera estructura se decidió cambiar el conformero inicial. Al utilizar el conformero 2 en el caso del complejo **6-PTN4** y el conformero 3 para su homólogo pentasacárido **3**, se pudieron reproducir los datos experimentales empleando MD-TAR.



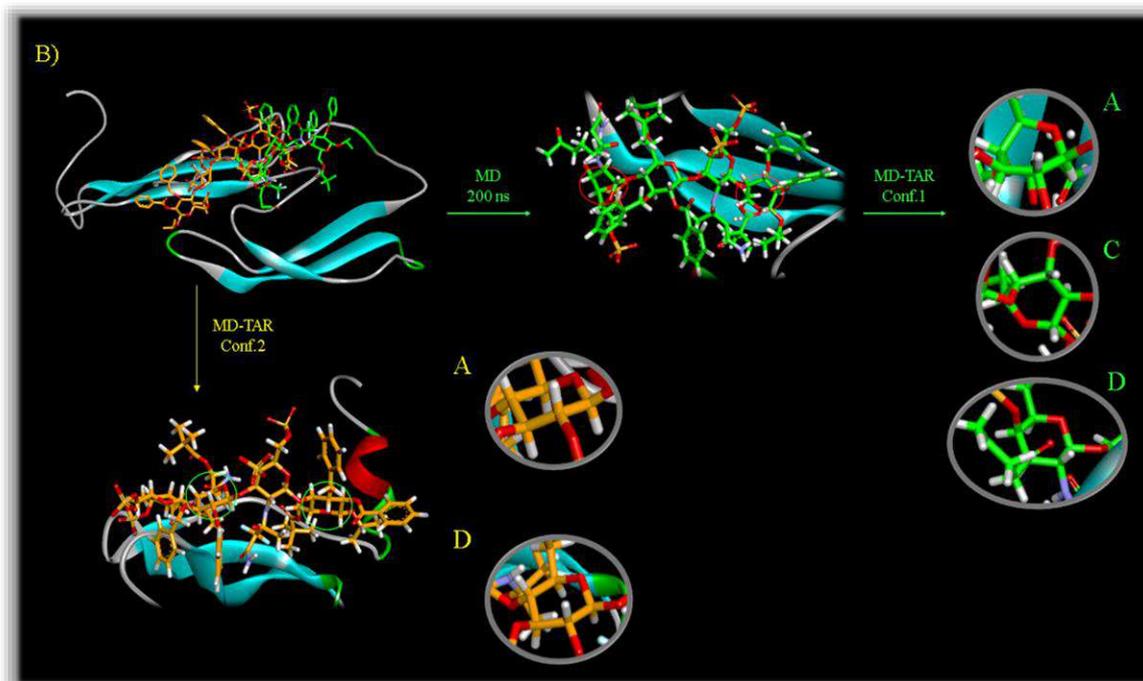


Figura 4.36. A) Representación tridimensional de las distorsiones del complejo **3-PTN4** para una dinámica sin restricciones (200 ns) y las MD-TAR con las restricciones correspondientes de los conformeros 1 y 3. B) Representación tridimensional de las distorsiones del complejo **6-PTN4** para una dinámica de 200 ns, y cuando aplicamos la MD-TAR con las restricciones correspondientes tanto para el primer conformero como para el segundo más favorable en energía.

Las gráficas de Cremer-Pople del compuesto **6-PTN4** mostraron una conformación 4C_1 en todos los anillos y los mapas de distribución predijeron un único mínimo de población con una orientación mayoritaria *syn-ψ/exo-φ* (Figuras A.38 y A.39B del Anexo). En cambio, para el complejo **3-PTN4** los resultados que se obtuvieron con el conformero 3 fueron prácticamente los mismos que con el primero, corrigiéndose el azúcar D mientras que el ángulo θ del azúcar C pasó de fluctuar entre $45-90^\circ$ a mantenerse entre 30° y 45° al igual que el anillo B. Por otro lado, los mapas de torsión sí mostraron la conformación esperada *syn-ψ /exo-φ* en los enlaces interglicosídicos (Figuras A.32B y A.34B del Anexo).

De nuevo, el acuerdo de las estructuras calculadas y los resultados experimentales supuso una ardua tarea en el caso de la conformación abierta de la proteína. Hubo que recurrir al uso de restricciones promediadas en el tiempo y cambios del conformero de partida. En la Figura 4.37 se han representado todas las simulaciones llevadas a cabo, y que se discutirán a continuación.

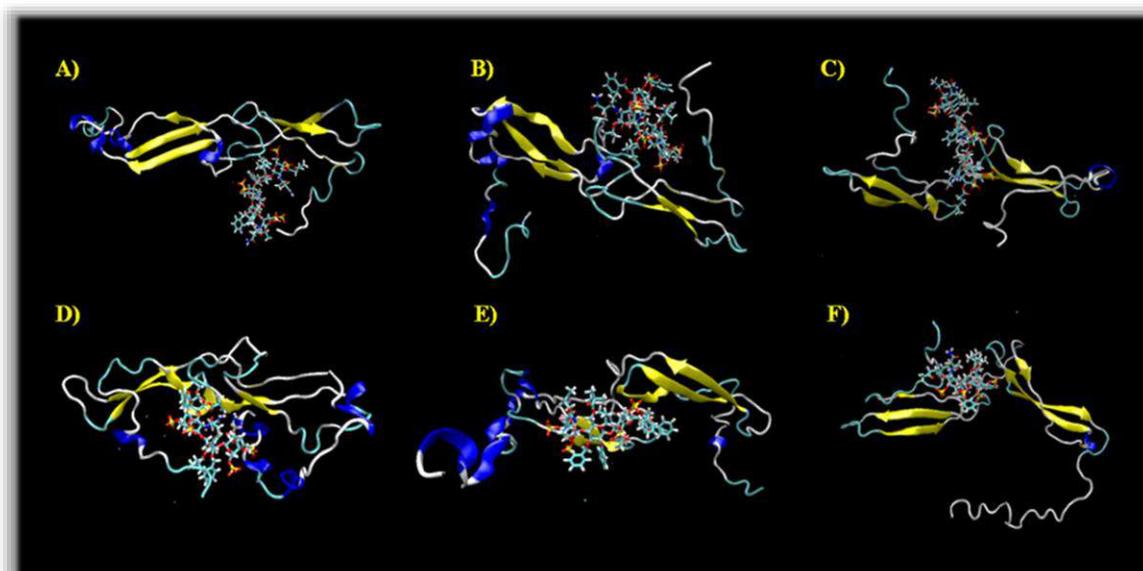
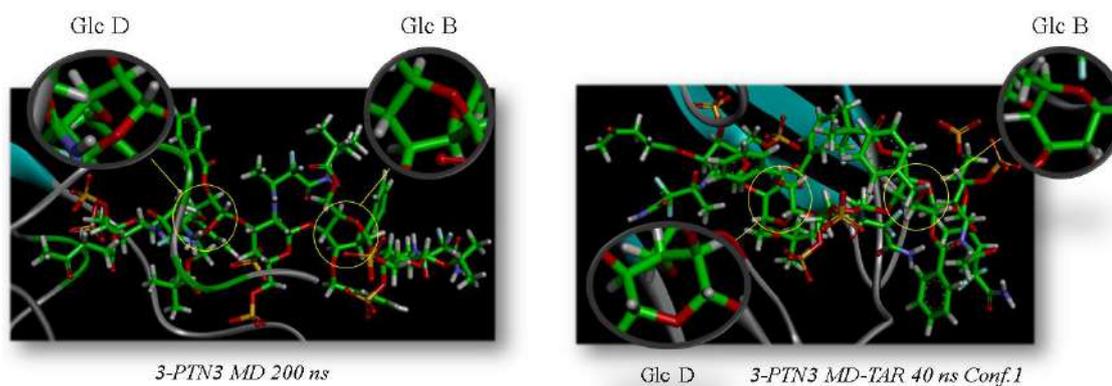


Figura 4.37. Representación tridimensional de las distorsiones de los complejos **3-PTN3** y **6-PTN3** de las simulaciones moleculares llevadas a cabo: A) Estructura final para una dinámica sin restricciones (200 ns) para el complejo pentasacarídico **3-PTN3** utilizando el ligando de mayor afinidad con la proteína. B) Estructura tridimensional para una dinámica MD-TAR de 40 ns con restricciones en los anillos B y D para el mismo conformero. C) Repetición de la dinámica molecular con restricciones utilizando otra conformación de ligando (conf. 5) con MD-TAR. D) Estructura final para una dinámica sin restricciones (200 ns) para el complejo tetrasacarídico **6-PTN3** utilizando el ligando de mayor afinidad con la proteína. E) Estructura tridimensional para la dinámica MD-TAR de 40 ns con restricciones en los anillos A, C y D, corrigiéndose este último. F) Repetición de la simulación MD-TAR cambiando la conformación de ligando (conf.3), utilizando un valor de $K=50 \text{ Kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ para los anillos A y C sin éxito.

En primer lugar, se comentarán los resultados tanto de las dinámicas libres como de las MD-TAR para el mejor conformero de Autodock analizado (Figura 4.38). Para el complejo **3-PTN3** ni las coordenadas de *puckering* ni los enlaces glicosídicos están en acuerdo con los resultados experimentales, mostrando distorsiones de bote torcido 0S_2 y de silla 1C_4 en los anillos B y D, y además con un enlace GalNAc C-Glc D también en desacuerdo con los resultados experimentales (ver Figuras A.40-A.43 del Anexo). Para el complejo tetrasacarídico **6-PTN3**, los cálculos MD-TAR proporcionaron conformaciones razonables para los mapas (ψ , ϕ), con una orientación *syn- ψ /exo- ϕ* para todos los enlaces, pero con distorsiones en los anillos (Figuras A.46, A.47, A.48A, A.48B y A.49A del Anexo).

A)



B)

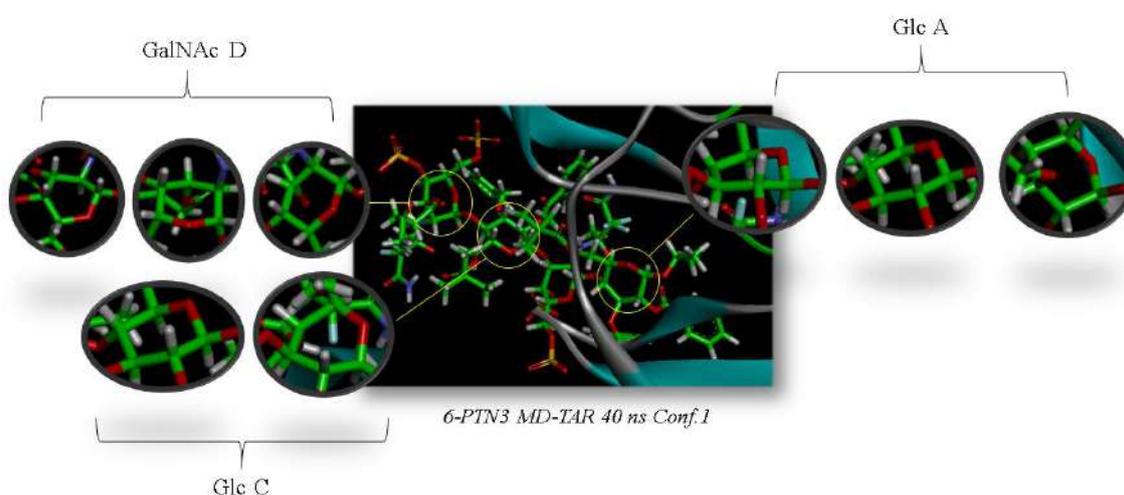


Figura 4.38. A) Representación tridimensional de las distorsiones del complejo **3-PTN3** para una dinámica sin restricciones (200 ns) y para las MD-TAR de 80 ns tanto para el conformero 1 como para el 5 corrigiéndose las distorsiones. B) Representación tridimensional de las distorsiones del complejo **6-PTN3** cuando aplicamos la MD-TAR con las restricciones correspondientes sobre los conformeros 1 y 3.

Siguiendo el mismo procedimiento, se procedió a elegir otra estructura de ligando procedente del *docking*. En la Figura 4.39 se han representado los conformeros escogidos para la simulación: el 5 para el complejo **3-PTN3** y el 3 para su homólogo tetrasacárido **6**. Tras el análisis habitual, sólo se consiguió reproducir una estructura de acuerdo con los resultados experimentales para el complejo **3-PTN3** (Figuras A.44 y A.45 del Anexo). Para el complejo **6-PTN3** no se ha conseguido un resultado satisfactorio (Figuras A.48C y A.49B del Anexo). Los mapas de distribución de población mostraron en ambos casos una geometría *syn-ψ/exo-φ* en todas las uniones interglucosídicas. Tras estos intentos, sólo fuimos capaces de obtener una estructura compatible con los resultados experimentales para el complejo **3-PTN3**, aunque no se descarta que con algunos de los otros conformeros del compuesto **6-PTN3** se puedan alcanzar los mismos resultados, a pesar de no tener información del complejo experimental.

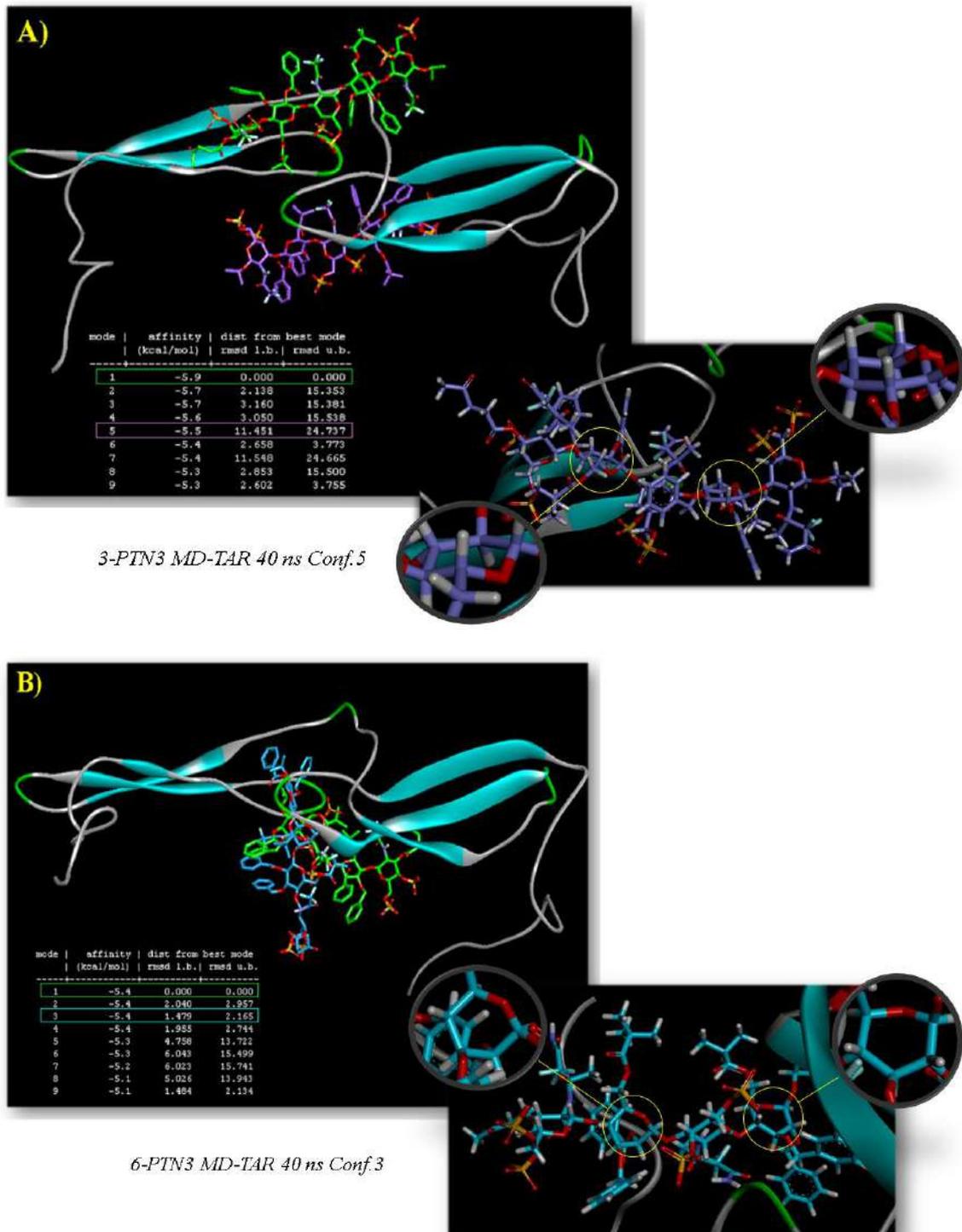


Figura 4.39. A) Para el complejo 3-PTN3 se muestra en color verde, el primer conformero más estable escogido para la simulación y en morado, el conformero 5 que, aunque presentaba una energía de afinidad ligeramente más baja con respecto al primero, disponía de una conformación más adecuada. A su derecha, el resultado de la simulación con la corrección de los azúcares B y D. B) Para el complejo 6-PTN3 se muestra en color verde de nuevo el primer conformero con mayor energía de afinidad y en color azul, el conformero 3 que, en este caso, presentaba la misma afinidad y una conformación más apropiada. A su derecha el resultado de la simulación, sin éxito.

Siguiendo el mismo procedimiento que para los complejos anteriores se analizan ahora los complejos con la conformación bisagra, PTN8. Tras los resultados anteriores se comenzó directamente por las simulaciones MD-TAR. A pesar de introducir restricciones, el estudio del plegamiento de **3-PTN8** mostró una conformación de bote torcido 2S_0 prácticamente a lo largo de toda la trayectoria en los anillos B, D y, en el caso de C, fluctuaciones entre dos conformeros ecuatoriales 1,4B o 1S_3 y 4E o 4H_3 (ver Figuras A.50A y A.51A del Anexo). Sin embargo, para el ligando tetrasacárido **6-PTN8**, los cuatro anillos muestran una conformación mayoritariamente de silla 4C_1 . Por último, se repitieron los cálculos de MD-TAR con el conformero 2 del ligando **3** (ver Figura 4.40), y en este caso el análisis de las trayectorias mostró sillas canónicas 4C_1 estables (Figuras A.52 y A.53 del Anexo).

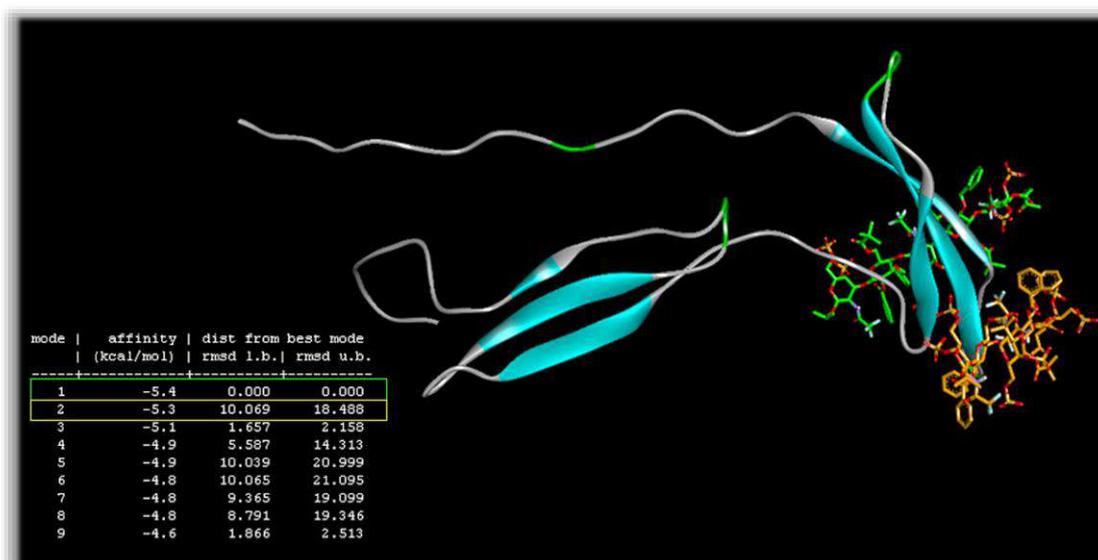


Figura 4.40. Representación de los dos conformeros utilizados para las MD-TAR llevadas a cabo para el complejo **3-PTN8**. En color verde, el primer conformero más estable escogido para la simulación; en naranja, el segundo más energéticamente favorable.

Los mapas de distribución de población de ambos complejos son similares y están de acuerdo con los descritos para **1**, **2**, **4** y **5-PTN**, mostrando en ambos una orientación *syn-ψ/exo-φ* en todos los enlaces (ver Figuras A.51B y A.53 del Anexo). Los análisis de todas las distancias procedentes de las simulaciones más representativas están recogidos en las Tablas 4.15 - 4.18, indicando también las distorsiones que han tenido lugar y las conformaciones anti que han aparecido. Comparando estos valores con los datos experimentales, podemos concluir que, con la excepción del complejo **6-PTN3**, ambos estudios están de acuerdo y confirman conformaciones de silla 4C_1 tanto para los anillos de Glc como de GalNAc, y geometrías de sus enlaces interglucosídicos *syn-ψ/exo-φ*.

$^3J_{HH}$ coupling	3-PTN4 RMN	3-PTN4 MD 200 ns	3-PTN4 MD-TAR 80ns	3-PTN3 MD 200ns	3-PTN3 MD-TAR Conf.1 40ns	3-PTN3 MD-TAR Conf.5 40ns	3-PTN8 MD-TAR Conf.1 40ns	3-PTN8 MD-TAR Conf.2 40ns
A1A2	8.0	8.0	8.0	8.2	8.2	8.2	7.2	7.5
A2A3	10.1	10.6	10.6	10.6	10.6	10.6	10.4	10.4
A3A4	2.2	1.6	1.7	2.0	1.9	2.0	2.2	1.9
A4A5		0.6	0.7	0.6	0.9	0.9	1.7	1.0
B1B2	7.9	6.5	7.0	2.0	3.2	5.4	5.6	6.9
B2B3	7.7	7.4	7.7	2.6	1.1	8.9	3	7.5
B3B4	9.0	8	8.1	1.2	2	8.6	1.3	7
B4B5	7.6	9.4	9.6	9.0	2.9	9.8	9.3	9.5
C1C2	8.4	4.5	7.8	8.4	8.5	7.9	5.3	8.1
C2C3	10.5	10.4	9.6	10.6	10.8	10.6	10.0	10.6
C3C4		1.5	1.5	2.0	2.2	1.9	1.5	1.9
C4C5		2.9	0.7	0.5	1.3	0.8	1.2	0.8
D1D2	7.5	5.7	7.3	2.3	2.1	7.8	1.8	6.9
D2D3	8.3	1.5	9.7	1.7	2.2	10.1	1.9	7.3
D3D4	9.1	1.2	9.5	2.4	1.0	9.0	1.8	7.0
D4D5	8.5	4.8	9.7	9.0	6.7	9.8	9.3	9.6
E1E2		8.3	8.4	8.3	8.2	8.3	8.3	8.5
E2E3	10.7	10.7	10.7	10.8	10.7	10.7	10.7	10.7
E3E4		2.4	2.3	1.6	2.0	2.0	2.3	2.5
E4E5		1.2	1.1	0.8	1.0	0.9	1.3	1.3

Tabla 4.15. Constantes de acoplamiento teóricas calculadas a partir de las MD libres de restricciones de 200 ns y de las dinámicas MD-TAR para el complejo 3-PTN en las 3 conformaciones escogidas de la proteína, con los conformeros elegidos con mayor afinidad.

$^3J_{HH}$ coupling	6-PTN4 MD 200 ns	6-PTN4 MD-TAR 40ns Conf.1	6-PTN4 MD-TAR 40ns Conf.2	6-PTN3 MD 200ns	6-PTN3 MD-TAR 40ns Conf.1	6-PTN3 MD-TAR 40ns Conf.3	6-PTN8 MD-TAR 40ns
A1A2	8.1	6.3	7.4	5.8	8.2	7.4	6.5
A2A3	5.6	5.4	8	7.4	8.2	7.9	8.7
A3A4	1.2	1.2	7	3.0	2.6	6.0	8.6
A4A5	1.1	1.4	9.3	2.9	5.2	8.9	9.6
B1B2	8.5	8.1	8.2	8.4	8.4	8.5	8.2
B2B3	10.8	10.5	10.3	10.8	10.7	10.7	10.6
B3B4	1.9	1.9	2.4	2.2	2.4	2.4	1.8
B4B5	0.4	1.0	1.8	1.7	1.9	1.8	0.8

Tabla 4.16. Constantes de acoplamiento teóricas calculadas a partir de las MD libres de restricciones de 200 ns y de las MD-TAR para el complejo 6-PTN en las 3 conformaciones escogidas de la proteína con los conformeros elegidos de mayor afinidad

$^3J_{HH}$ coupling	6-PTN4 MD 200 ns	6-PTN4 MD-TAR 40ns Conf.1	6-PTN4 MD-TAR 40ns Conf.2	6-PTN3 MD 200ns	6-PTN3 MD-TAR 40ns Conf.1	6-PTN3 MD-TAR 40ns Conf.3	6-PTN8 MD-TAR 40ns
C1C2	8.1	7.2	7.1	6.7	7.8	3.0	7.4
C2C3	9.8	5.9	9.0	1.7	9.7	1.5	8.8
C3C4	7.4	1.2	8.5	1.0	6.1	2.8	7.6
C4C5	8.6	1.3	9.8	4.3	7.3	2.1	9.5
D1D2	4.9	8.0	8.5	1.2	5.8	8.3	8.0
D2D3	10.0	10.7	10.8	1.6	3.4	10.8	10.5
D3D4	1.9	2.3	2.3	2.7	4.1	2.3	2.0
D4D5	2.1	0.8	1.6	5.0	7.1	0.9	0.6

Tabla 4.16. Continuación

Distancias	3-PTN4 RMN	3-PTN4 MD 200 ns	3-PTN4 MD-TAR 40ns Conf.1	3-PTN3 MD-TAR 40ns Conf.1	3-PTN3 MD-TAR 40ns Conf.5	3-PTN8 MD-TAR 40ns Conf.1	3-PTN8 MD-TAR 40ns Conf.2
A H1H3	2.6	2.6	2.5	2.5	2.5	2.7	2.6
A H1H5	2.5	2.6	2.6	2.6	2.6	2.9	2.7
A H3H5		2.4	2.3	2.4	2.4	2.5	2.3
Interg.BH1-AH3		2.6	2.6	2.3	2.6	2.4	2.3
B H1H3	2.6	3.1	3.0	3.7	2.9	4.2	3.0
B H1H5	2.7	2.4	2.3	2.5	2.4	2.8	2.3
B H3H5	2.6	2.9	2.9	4.0	2.6	3.5	2.9
Interg.CH1-BH4	2.4	2.5	2.7	2.3	2.5	2.5	2.4
C H1H3		3.1	3.1	2.4	2.6	3.0	2.6
C H1H5		3.7	3.2	2.6	2.7	3.2	2.7
C H3H5		2.5	2.3	2.5	2.3	2.3	2.4
Interg.DH1-CH3		2.4	2.8	3.0	2.4	2.2	2.5
D H1H3	2.2	3.6	2.8	4.1	2.7	4.0	3.0
D H1H5	3.0	2.4	2.5	2.5	2.4	2.5	2.4
D H3H5	2.5	3.8	2.5	3.8	2.5	3.3	2.8
Interg.EH1-DH4		3.3	2.4	2.4	2.4	2.4	2.3
E H1H3		2.4	2.5	2.5	2.5	2.5	2.4
E H1H5		2.4	2.6	2.6	2.6	2.6	2.5
E H3H5	2.9	2.5	2.4	2.4	2.4	2.5	2.5

Tabla 4.17. Distancias experimentales de cada pareja de protones intrarresiduales e interglucosídicos para el complejo 3-PTN y distancias teóricas promediadas de las simulaciones moleculares (a 300 K en MD de 200 ns y MD-TAR de 40 ns) para las 3 conformaciones escogidas de la proteína y los conformeros elegidos con mayor afinidad.

Distancias	6-PTN4	6PTN4	6-PTN4	6-PTN3	6-PTN3	6-PTN3	6-PTN8
	MD 200ns	MD-TAR 40ns Conf.1	MD 40ns Conf.2	MD 200ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 3	MD-TAR 40ns
A H1H3	3.1	3.3	2.9	3.1	2.8	3.0	2.8
A H1H5	3.3	3.6	2.8	3.5	2.5	2.5	2.6
A H3H5	4.2	4.1	2.9	3.7	3.6	3.1	2.7
Interg.BH1-AH4	2.2	2.3	2.7	2.2	2.5	3.1	2.5
B H1H3	2.4	2.5	2.6	2.4	2.2	2.4	2.5
B H1H5	2.5	2.6	2.5	2.6	2.6	2.5	2.7
B H3H5	2.4	2.4	2.4	2.5	2.6	2.5	2.4
Interg.CH1-BH3	2.3	2.5	2.4	2.9	2.5	2.4	2.6
C H1H3	2.5	3.2	3.0	3.6	2.8	4.0	3.0
C H1H5	2.4	3.5	2.5	2.4	2.8	3.2	2.4
C H3H5	2.9	4.1	2.7	3.9	3.2	4.2	2.8
Interg.DH1-CH4	2.4	2.3	2.4	2.9	2.2	2.9	2.6
D H1H3	3.0	2.6	2.3	4.2	3.5	2.6	2.6
D H1H5	3.3	2.5	2.5	4.1	3.1	2.5	2.5
D H3H5	2.4	2.3	2.4	4.2	4.1	2.4	2.3

Tabla 4.18. Distancias experimentales de cada pareja de protones intrarresiduales e interglucosídicos para el complejo **6-PTN** y distancias teóricas promediadas de las simulaciones moleculares (a 300 K en MD de 200 ns y MD-TAR de 40 ns) para las 3 conformaciones escogidas de la proteína y los conformeros elegidos con mayor afinidad.

En la Tabla 4.19 se recoge toda la información acerca de la metodología seguida para cada uno de los complejos **1-6 PTN**, en las dinámicas moleculares con y sin restricciones que se han llevado a cabo. A la vista de estos datos, los resultados de las dinámicas fueron menos satisfactorios cuanto más sustituido se encontraba el ligando, teniendo que utilizar otro conformero de partida en varias ocasiones. No obstante, en el caso del compuesto **4-PTN8**, a pesar de no presentar sustituyentes el carbohidrato, la obtención de una estructura acorde a los resultados experimentales fue más complicada que en el caso de su homólogo **1-PTN8**. Cabe añadir que, con respecto a la conformación que adopta la PTN en las simulaciones, el patrón de comportamiento ha sido diferente dependiendo de si se establecían o no restricciones. Para las dinámicas moleculares libres, la estructura de la proteína tiende a englobar al ligando entre sus dominios; en cambio, en las MD-TAR se mantenía a lo largo de toda la trayectoria la conformación inicial de la que se parte.

En general, para la serie tetrasacáridica se realizaron más pruebas hasta llegar a una estructura del complejo que apoyase los datos de RMN, pudiendo ser debido este hecho a la mayor flexibilidad que presenta el ligando, al contar con menos repulsión de cargas sobre el esqueleto carbonado. En lo que respecta a la estructura de la proteína, se observaron menos distorsiones en la conformación más plegada (PTN4), a diferencia de las otras dos. Esto puede ser debido a que, cuando tiene lugar la formación del complejo, el ligando se sitúa sobre el centro activo sin que se vea afectado por los dominios

terminales. En cambio, en las conformaciones abierta y en forma de bisagra (PTN3 y PTN8 respectivamente), estos dominios terminales tienden a plegarse, en mayor grado cuando no se establecen restricciones en la simulación, encerrando al ligando y pudiendo generar deformación de los anillos.

Métodos	MD 200 ns	MD-TAR 40 ns Conformero 1	MD-TAR 40 ns (Cambio de conformero)
Complejos Pentasacarídicos			
Conformación PTN4			
1-PTN4	✓		
2-PTN4	✓		
3-PTN4	×	✓	
Conformación PTN3			
1-PTN3	✓		
2-PTN3	✓		
3-PTN3	×	×	Conformero 5 ✓
Conformación PTN8			
1-PTN8	✓	✓	
2-PTN8	×	✓	
3-PTN8	×	×	Conformero 2 ✓
Complejos Tetrasacarídicos			
Conformación PTN4			
4-PTN4	✓		
5-PTN4	×	✓	
6-PTN4	×	×	Conformero 2 ✓
Conformación PTN3			
4-PTN3	✓		
5-PTN3	×	×	Conformero 2 ✓
6-PTN3	×	×	Conformero 3 ×
Conformación PTN8			
4-PTN8	×	×	Conformero 3 × Conformero 5 ✓
5-PTN8	×	×	Conformero 7 ✓
6-PTN8	×	✓	

Tabla 4.19. Resumen de los resultados obtenidos de las dinámicas moleculares con y sin restricciones para ambas series de complejos en las tres conformaciones de la PTN.

4.4 Métodos MM-PBSA/GBSA

El último análisis de interacción que se llevó a cabo para estos complejos fue el de cuantificación de la afinidad con valores de energía libre. Para ello, se realizaron cálculos de MM-PBSA y MM-GBSA (*Molecular Mechanics Energies Combined with the Poisson-Boltzmann or Generalized Born and Surface Area continuum solvation*) en ambas series oligosacáridicas, a partir de los archivos de trayectorias finales generados de las dinámicas moleculares del apartado anterior. Estas dos técnicas permiten calcular la

energía libre de unión relativa proteína-ligando y energías libres absolutas de una molécula en disolución^[32]. En este trabajo se han utilizado ambos métodos para cuantificar la contribución energética en la unión de esta serie de miméticos de CS-E con la PTN. Su análisis nos va a permitir determinar la fortaleza de cada interacción individual determinando cuáles son los aminoácidos más relevantes para la formación del complejo.

En primer lugar, se presentan en la Tabla 4.20 los resultados de los cálculos MM-GBSA para los complejos pentasacáridicos y las tres conformaciones de la PTN (PTN4, PTN3 y PTN8), tanto de las dinámicas moleculares libres como de las MD-TAR.

Cabe destacar que, para realizar estos cálculos de manera homogénea y comparable, se han expuesto los valores obtenidos de las simulaciones correspondientes al primer conformero en energía de afinidad, que ya se encontraba situado en el sitio activo de la PTN.

GB Energies (kcal/mol)	VDW	EEL	E _{GB}	E _{Surf}	ΔG_{gas}	ΔG_{solv}	ΔG_{T}
1-PTN4	-54.81	-6154.84	6107.45	-14.63	-6209.65	6209.81	-116.84
1-PTN3	-15.83	-4663.19	4608.61	-7.34	-4679.03	4601.26	-77.76
1-PTN8	-24.23	-5529.91	5428.71	-10.08	-5554.14	5418.63	-135.51
2-PTN4	-28.37	-6425.57	6306.24	-11.9	-6453.94	6294.34	-159.6
2-PTN3	-44.38	-6798.96	6716.75	-14.72	-6843.34	6702.02	-141.32
2-PTN8	-56.05	-4742.02	4686.85	-13.49	-4798.07	4673.36	-124.71
2-PTN8-TAR	-34.52	-3727.71	3722.09	-7.01	-3762.23	3715.08	-47.15
3-PTN4	-60.55	-5686.06	5601.39	-13.32	-5746.62	5588.07	-158.55
3-PTN4-TAR	-47.55	-5003.6	4985.48	-10.94	-5061.15	4974.53	-66.62
3-PTN3	-46.78	-8076.94	7878.09	-15.89	-8123.73	7862.2	-261.53
3-PTN3-TAR	-49.63	-4204.26	4185.99	-9.15	-4253.89	4176.84	-77.05
3-PTN8	-62.65	-6167.15	6117.68	-14.95	-6229.8	6102.73	-127.07
3-PTN8-TAR	-33.96	-4947.51	4940.6	-7.78	-4981.47	4932.81	-48.66

Tabla 4.20. Energía de unión (kcal/mol) para la serie pentasacáridica con el método MM-GBSA a partir de los resultados de dinámica molecular libres y MD-TAR.

De acuerdo con los datos de FP la inclusión de anillos aromáticos en el ligando genera un aumento en el valor de ΔG_{T} proporcional a su número. Si estudiamos individualmente los diferentes complejos para conocer qué conformación de la proteína es la más favorable en su unión al carbohidrato, los resultados no son unánimes.

En el caso del primer ligando (**1-PTN**), el modo de unión más estable con la PTN se establece cuando ésta se presenta en forma de horquilla (PTN8). Para el complejo dibencilado (**2-PTN**), la energía libre es algo mayor cuando la proteína se encuentra plegada (PTN4) y éste puede situarse entre ambos dominios siendo rodeado por la proteína; mientras que el complejo persustituído (**3-PTN**) muestra preferencia por la

conformación abierta (PTN3). Posiblemente el volumen que le otorgan los grupos protectores a la estructura a lo largo de esta serie induce una interacción más favorable con las estructuras abiertas por el aumento del volumen.

En la Tabla 4.20 también se incluyen simulaciones MD-TAR. Los valores de ΔG_T lógicamente son diferentes debido a la inclusión de la energía de penalización. Como no fue necesario establecer restricciones en todos los cálculos, no se ha podido hacer una comparación más precisa para extraer más conclusiones acerca de estos resultados.

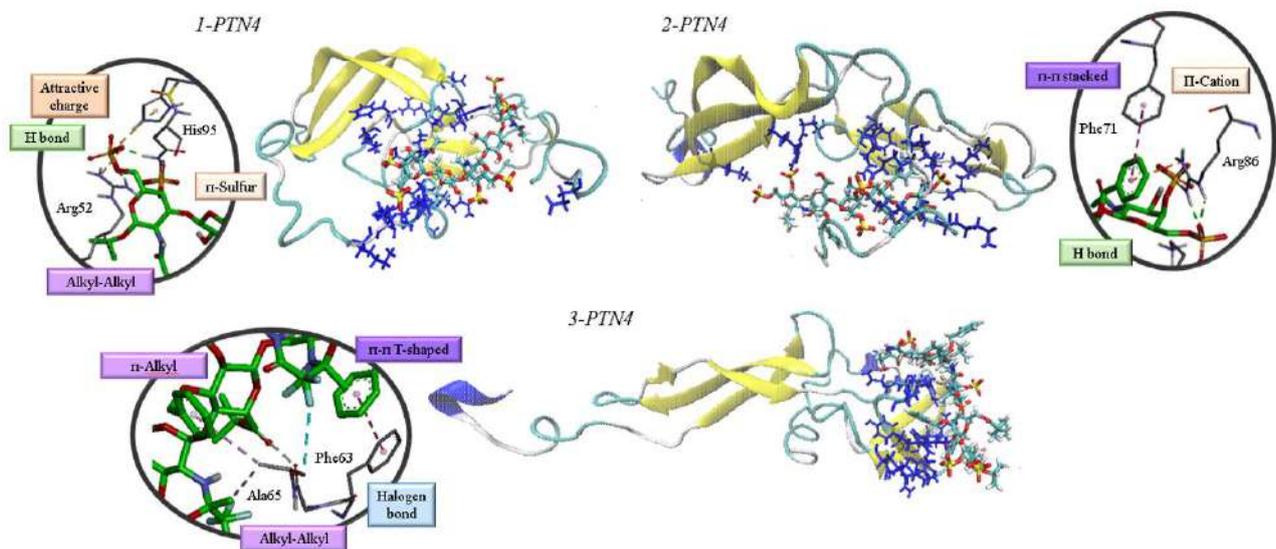
Posteriormente, se han calculado los valores de interacción mediante el método MM-GBSA para los complejos tetrasacarídicos **4-**, **5-** y **6-PTN** (ver Tabla 4.21). Estos han seguido un patrón de comportamiento similar al de los complejos de PTN con **1**, **2** y **3**. En este caso esperábamos obtener una afinidad mayor por la PTN con la serie pentasacarídica como se observó en los datos de polarización. En cambio, los resultados obtenidos no mostraron grandes diferencias entre los tetras y los pentasacáridos. Incluso en algunos casos, como el complejo **5-PTN**, la energía total de interacción fue ligeramente mayor con respecto a su homólogo **2-PTN**. En el caso de los complejos cuyos ligandos presentan anillos aromáticos adicionales en su estructura, (**5-PTN** y **6-PTN**), siguen la misma tendencia que la encontrada en la serie pentasacarídica, con preferencia por las conformaciones de la proteína plegada y abierta, PTN4 y PTN3, respectivamente. Para el complejo tetrasacarídico completamente desprotegido su comportamiento es diferente, mostrando mayor energía de unión cuando la proteína se presenta desplegada.

GB Energies (Kcal/mol)	VDW	EEL	E _{GB}	E _{Surf}	ΔG_{gas}	ΔG_{solv}	ΔG_T
4-PTN4	-30.08	-4389.44	4309.5	-9.88	-4419.53	4299.62	-119.9
4-PTN3	-20.72	-4155.43	4036.53	-9.66	-4176.14	4026.86	-149.28
4-PTN8	-48.5	-2167.64	2165.6	-7.44	-2216.14	2158.16	-57.98
4-PTN8-TAR	-45.87	-2388.47	2384.41	-7.19	-2434.34	2377.22	-57.12
5-PTN4	-48.5	-4548.04	4428.47	-12.27	-4596.54	4416.19	-180.35
5-PTN4-TAR	-30.2	-3526.43	3478.27	-7.27	-3556.64	3471.01	-85.63
5-PTN3	-56.94	-4366.98	4302.98	-12.3	-4423.92	4290.68	-133.24
5-PTN3-TAR	-45.95	-2683.59	2692.24	-8.23	-2729.54	2684.01	-45.53
5-PTN8	-58.45	-4550.06	4475.34	-13.38	-4608.51	4461.96	-146.55
5-PTN8-TAR	-19.24	-2213.37	2207.2	-4.58	-2232.61	2202.61	-29.99
6-PTN4	-64.65	-4735.38	4678.38	-14.53	-4800.03	4663.84	-136.18
6-PTN4-TAR	-39.73	-2944.55	2950.17	-6.50	-2984.27	2943.66	-40.61
6-PTN3	-75.92	-5278.69	5180.76	-15.59	-5354.61	5165.17	-189.44
6-PTN3-TAR	-69.79	-2749.61	2791.42	-9.52	-2819.4	2781.89	-37.51
6-PTN8	-40.8	-5242.49	5155.86	-11.36	-5283.3	5144.5	-138.8
6-PTN8-TAR	-49.69	-2754.56	2769.09	-8.08	-2804.25	2761.02	-43.24

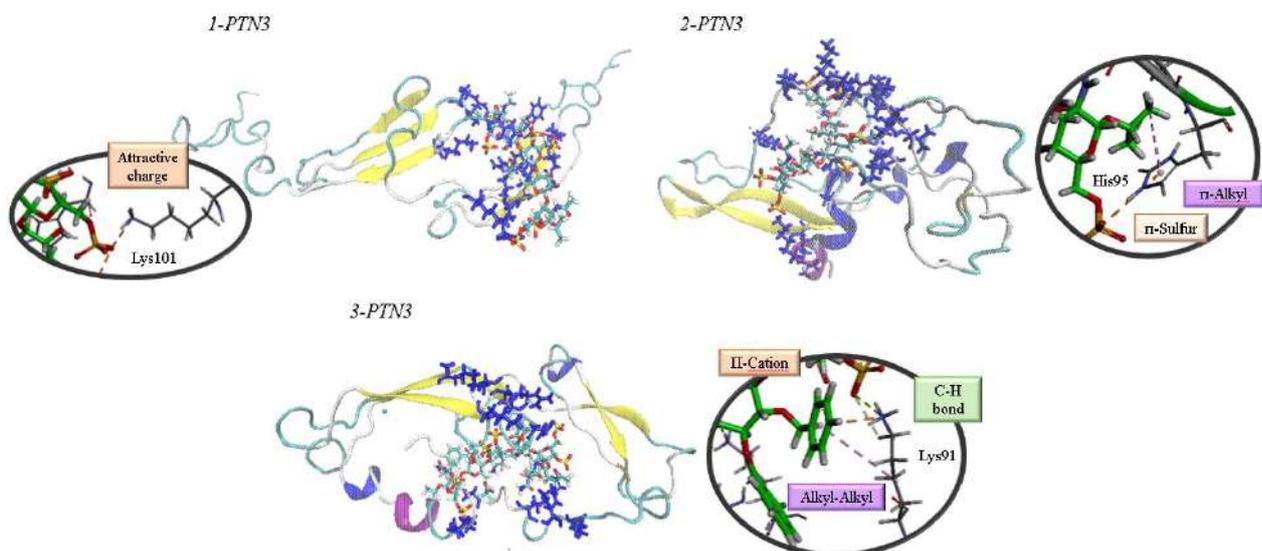
Tabla 4.21. Energía de unión (kcal/mol) para la serie tetrasacarídica con el método MM-GBSA a partir de los resultados de dinámica molecular libres y MD-TAR.

Estos cálculos de MM-GBSA se han combinado con las técnicas de MM-PBSA utilizando las mismas estructuras finales tanto de las dinámicas moleculares como de las MD-TAR de cada complejo. Aunque estas simulaciones fueron llevadas a cabo orientando el sitio activo de la proteína, como se ha mencionado anteriormente, no ha sido ésta la única región que ha mostrado interacción. En primer lugar, se ha calculado la contribución energética individual de cada aminoácido para los complejos con los pentasacáridos **1**, **2** y **3**. Se considera una interacción favorable cuando ΔG_T alcanza valores negativos en energía por debajo de -1 kcal/mol. Los resultados de afinidad por residuo para los complejos pentasacáridicos con las tres conformaciones de la proteína se muestran en el Anexo (Tablas A.T1-T3). En la Figura 4.41, se han representado los aminoácidos con mayor grado de participación en la interacción con los pentasacáridos, señalando en azul los residuos de la proteína con las energías de unión más fuertes.

a) *Complejos pentasacáridicos con la conformación plegada de la proteína.*



b) *Complejos pentasacáridicos con la conformación abierta de la proteína.*



c) Complejos pentasacáridos con la conformación bisagra de la proteína.

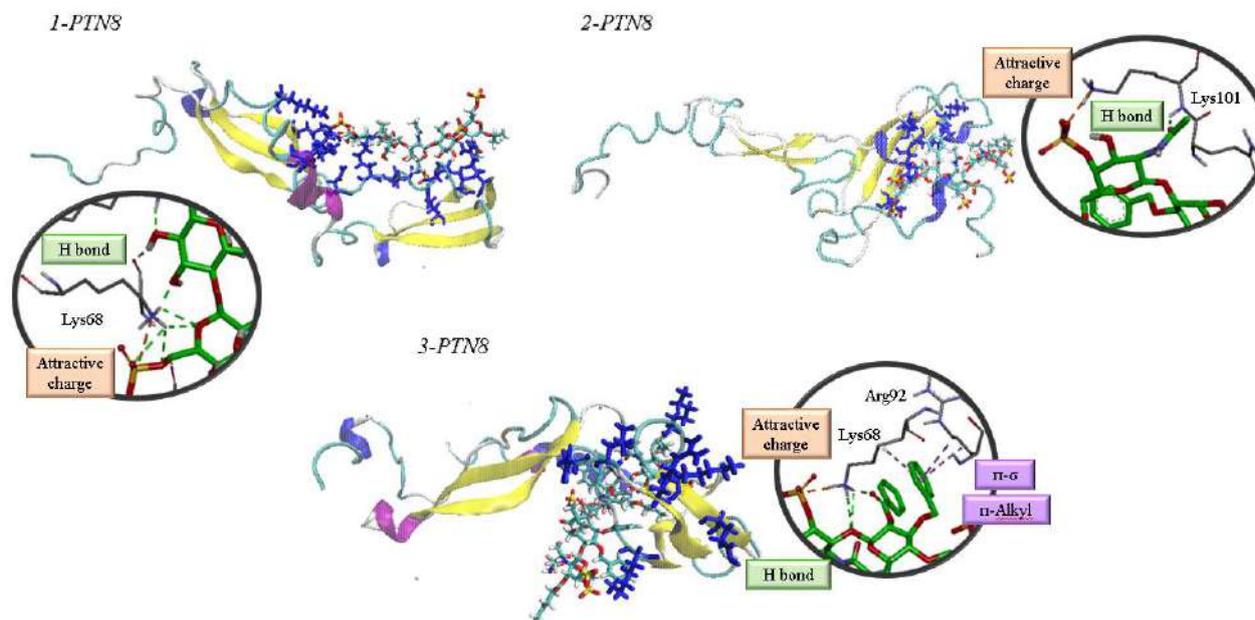


Figura 4.41. Representación tridimensional de los aminoácidos más implicados en la interacción entre los ligandos pentasacáridos y la PTN en sus tres conformaciones, obtenidos mediante la técnica MM-PBSA. Se han ampliado las regiones que muestran ciertas interacciones que tienen lugar con algunos de los aminoácidos de mayor afinidad.

Como era de esperar los resultados de interacción indican que los principales responsables de la formación de estos complejos son los residuos básicos Arg52, Arg92, Lys60, Lys68 y Lys101 mostrando uniones electrostáticas con valores de energía altos, aunque también encontramos fuertes interacciones con His95.

Cabe destacar que, debido al grado de plegamiento que muestra la primera de las conformaciones (PTN4), se han obtenidos valores altos de afinidad con los residuos básicos situados en el extremo final del dominio C-terminal, sin volver a observarse esto en las otras dos disposiciones de la proteína.

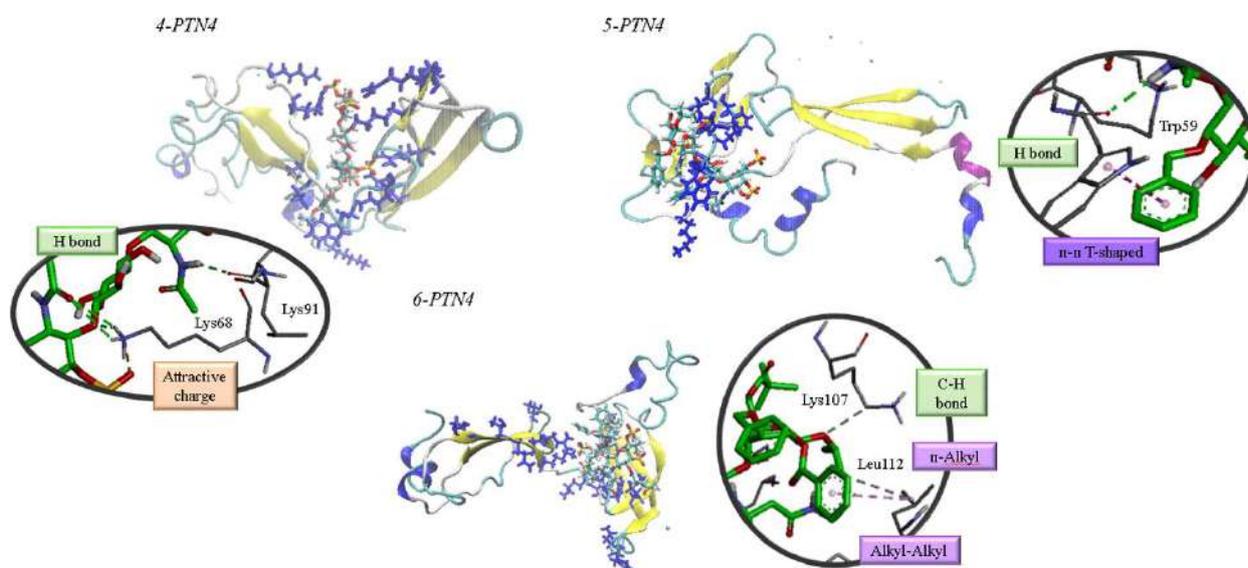
Además, a pesar de que la mayor parte de los aminoácidos participantes se encuentran en este dominio, el alto grado de contribución energética de la Arg52 y la presencia de una fuerte interacción con otros aminoácidos, véase Ser27 o Arg35, presente en las otras disposiciones, nos indica que también existe afinidad por el dominio N-terminal, aunque en mucha menos proporción.

Por otro lado, la presencia de grupos aromáticos en las estructuras de los miméticos de GAGs genera a su vez nuevas interacciones con aminoácidos también aromáticos, como Phe63 y Phe71, creando interacciones tipo π - π . Estos cálculos se han realizado sobre las trayectorias finales de aquellas dinámicas compatibles con los resultados experimentales. Por ello, en algunos casos, como para el complejo completamente sustituido, fue

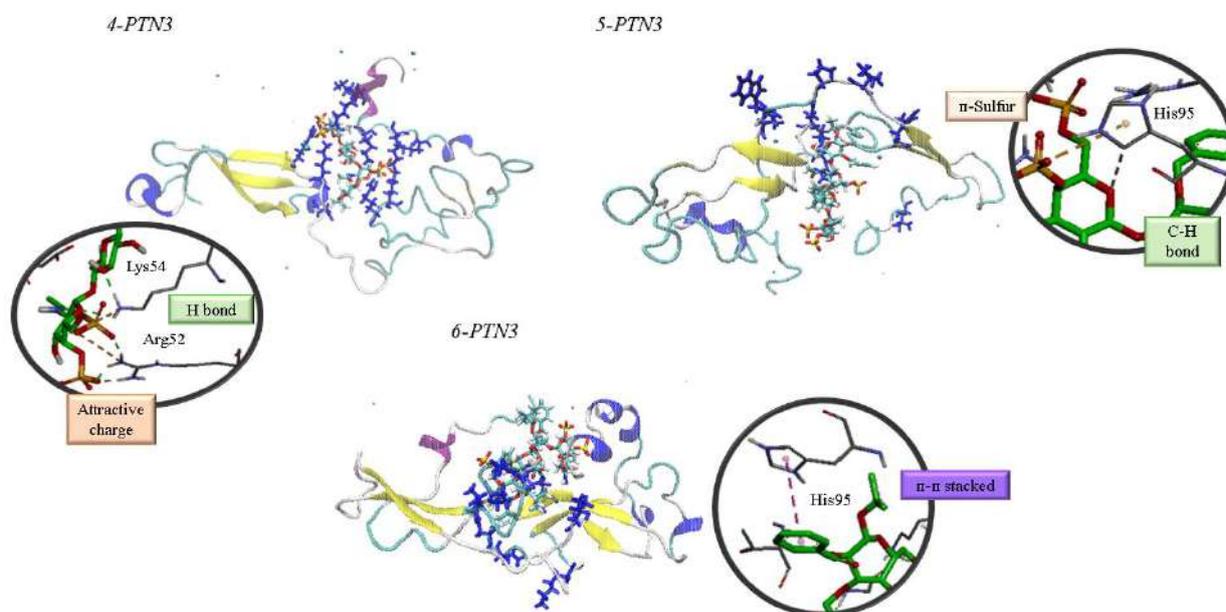
necesario recurrir a diferentes conformeros, con una afinidad menor, mostrando una interacción más baja en los cálculos de MMPBSA (Tablas A.T1-A.T3).

La serie tetrasacáridica **4-**, **5-** y **6-PTN** presenta resultados similares, compatibles con una interacción electrostática con predominio de los aminoácidos Arg y Lys y los grupos sulfato del ligando. Es destacable la participación del extremo C terminal en las conformaciones lineales (PTN3) de la proteína. Sin embargo, el mayor número de interacciones favorables se establecen en los dominios CTD y en la región central con respecto a los extremos terminales y el dominio NTD. Además, la interacción con el extremo inicial de la proteína sólo se observa en el caso del complejo dibencilado en su estructura más abierta, a diferencia de lo que ocurría antes para los pentasacáridos. Por otro lado, para los ligandos dibencilado y persustituído reaparecen las interacciones π - π con los aminoácidos Phe63, Phe71, Tyr69 y Trp59, aunque con afinidades relativamente bajas (Tablas A.T4-A.T6). A continuación, en la Figura 4.42 se muestran de nuevo las estructuras tridimensionales de los complejos para esta serie.

a) *Complejos tetrasacáridicos con la conformación plegada de la proteína*



b) Complejos tetrasacarídicos con la conformación abierta de la proteína



c) Complejos tetrasacarídicos con la conformación bisagra de la proteína.

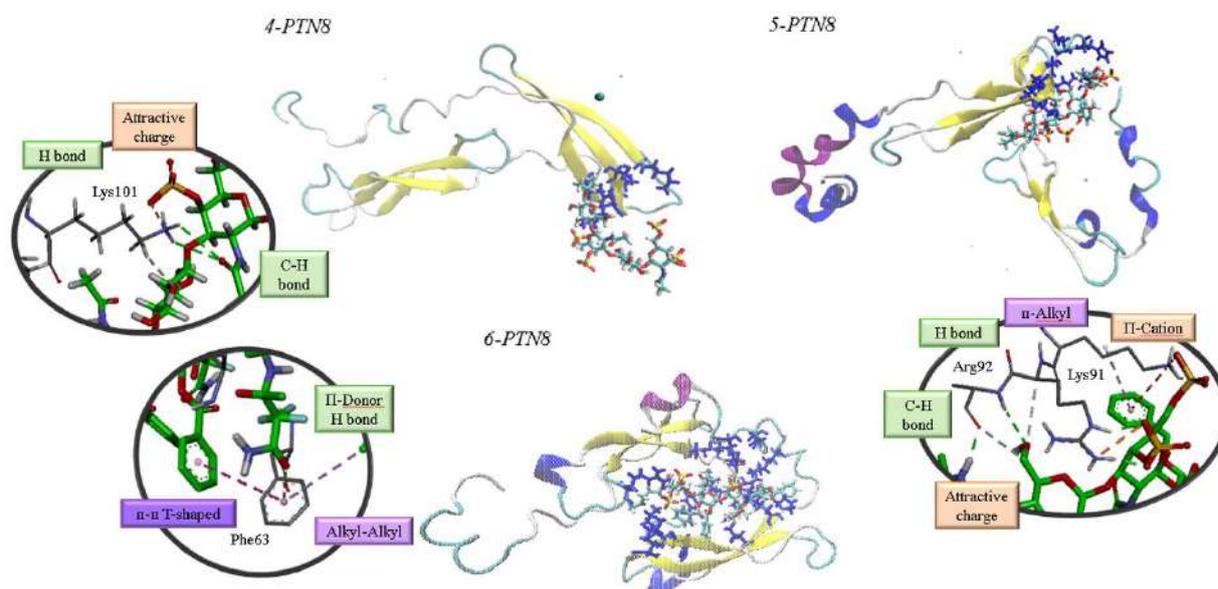
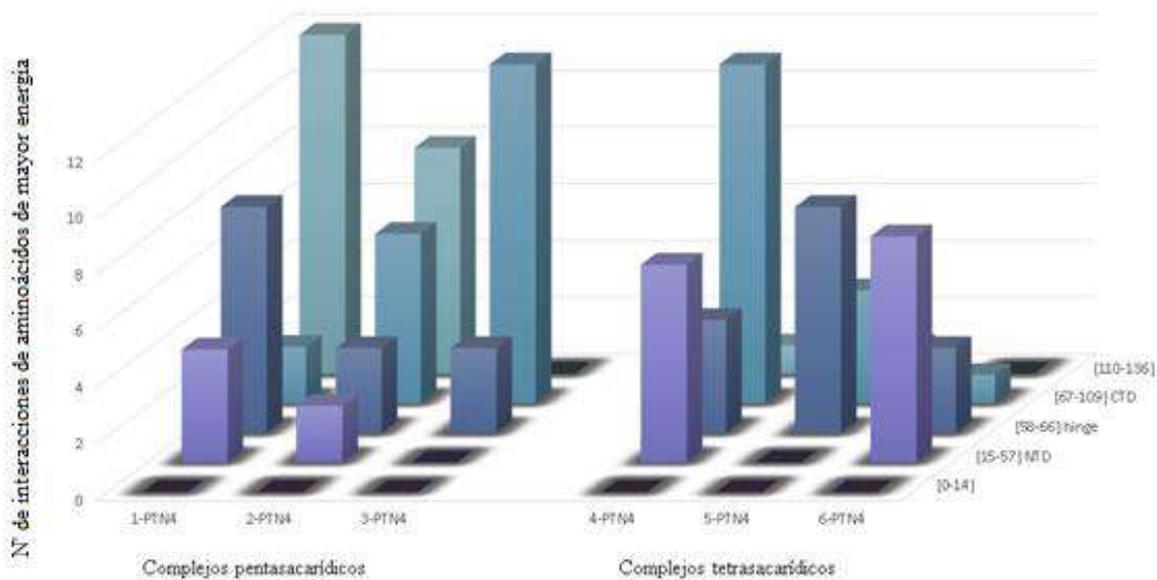


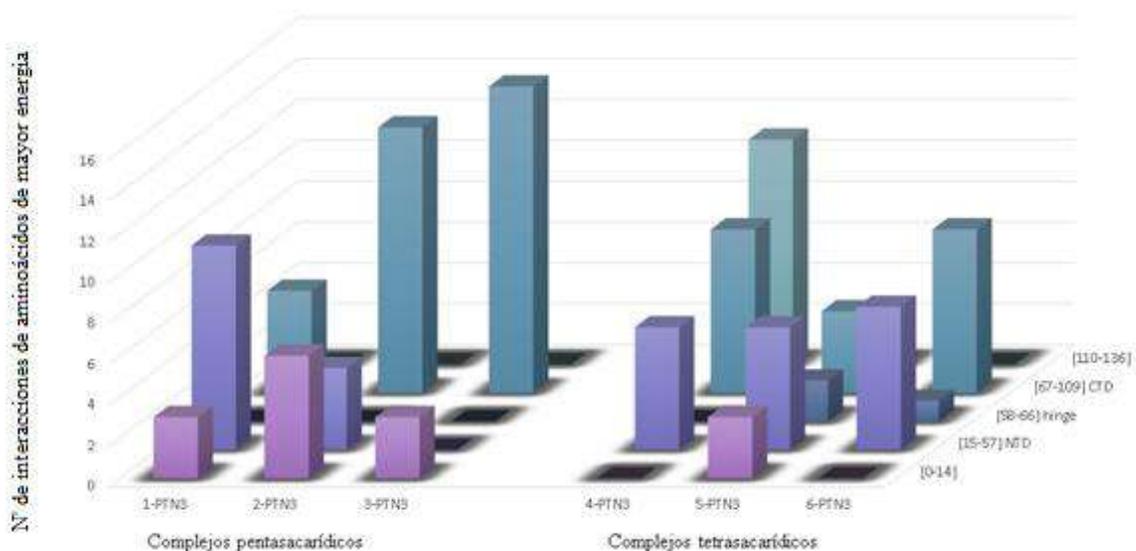
Figura 4.42. Representación tridimensional de los aminoácidos más implicados en la interacción entre los ligandos tetrasacarídicos y la PTN en sus tres conformaciones, obtenidos mediante la técnica MM-PBSA. Se han ampliado las regiones que muestran ciertas interacciones que tienen lugar con algunos de los aminoácidos de mayor afinidad.

Como resumen, en la Figura 4.43 se ha representado gráficamente el número de interacciones favorables en cada zona de la proteína, comparando ambas series para los tres conformeros de la PTN. Salvo algunas excepciones, como el complejo **4-PTN4** que presenta una interacción considerable con el dominio N-terminal de la PTN, se puede concluir que la mayor afinidad se produce en el dominio CTD. Le siguen las zonas central y extremo C-terminal de la proteína que también presentan interacción, exceptuando para la conformación abierta (PTN3) donde apenas se aprecia participación pero, en cambio, algunos residuos del extremo N-terminal sí muestran cierta afinidad.

➤ *Comparativa de ambas series para la conformación plegada de la PTN*



➤ *Comparativa de ambas series para la conformación abierta de la PTN*



➤ **Comparativa de ambas series para la conformación bisagra de la PTN**

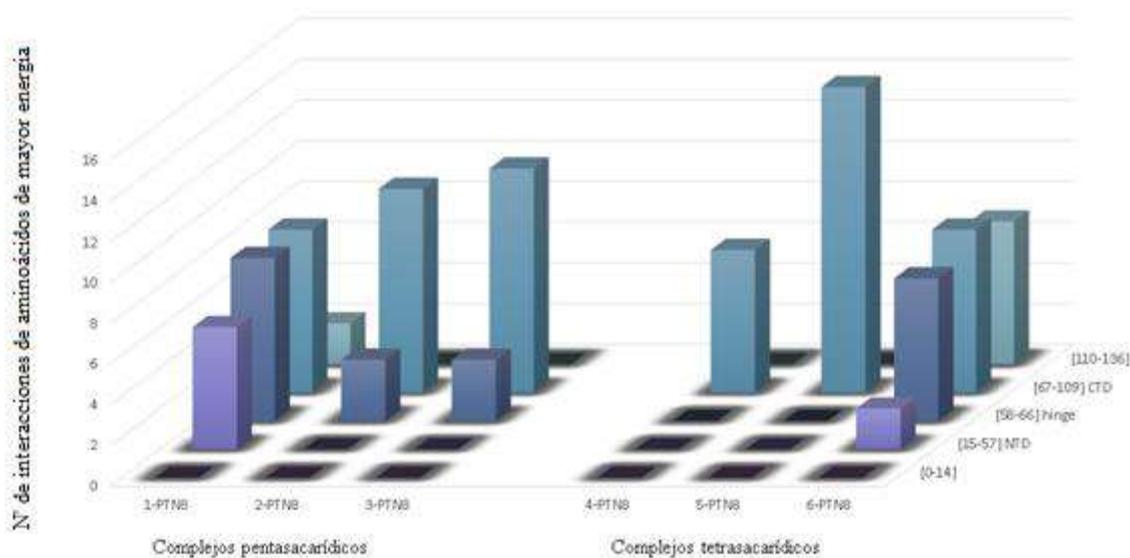


Figura 4.43. Representación gráfica de los resultados de las tablas de MM-PBSA comparando ambas series para cada conformero de la PTN, indicando las zonas de ésta que presentan mayor interacción con el carbohidrato.

4.5 Metodología

4.5.1. Preparación de inputs para las dinámicas moleculares

Las dinámicas moleculares se llevaron a cabo en diferentes *clusters* para rentabilizar al máximo los tiempos de cálculos, aumentando y mejorando los recursos que ya disponíamos. El primero en el que se realizaron estas simulaciones fue en CTE-POWER perteneciente al Centro Nacional de Supercomputación de Barcelona; posteriormente, se utilizó el *cluster* CESGA que se encuentra en el Centro de Supercomputación de Galicia. Para la construcción del complejo se reutilizaron los archivos de geometría iniciales de los ligandos libres descritos en el **Capítulo 3** que fueron creados previamente con la interfaz MAESTRO perteneciente al programa Schrödinger Suite 2012. Con el campo de fuerzas GLYCAM-06 se prepararon todos los residuos pertenecientes al ligando, utilizando para su parametrización los paquetes *pam94* y *parm99* de AMBER. Además, se reutilizaron los archivos de topología *prep* y *frmod* generados en los cálculos de los ligandos libres, cuyo procedimiento ha sido desarrollado en la metodología del **Capítulo 3**.

Por otro lado, las 3 conformaciones que se escogieron de la PTN se obtuvieron a partir de la base de datos Protein Data Bank: la disposición más plegada de la proteína (PTN4); como contrapunto a ésta, la conformación que estaba completamente abierta (PTN3); y,

por último, su estructura más frecuente en forma de horquilla (PTN8), plegada por la zona central que comprende ambos dominios.

Para el montaje del complejo, primeramente, se realizaron los cálculos de *docking* con el fin de obtener una población de las conformaciones y orientaciones más estables del ligando tanto en el sitio activo de la proteína como en toda su estructura, para abarcar todas las posibilidades. Estos estudios se llevaron a cabo en primer lugar con el programa Autodock Tools,^[33] generando los archivos en formato PDBQT tanto del carbohidrato como de la PTN, añadiendo los hidrógenos polares correspondientes. El sitio de acoplamiento en la proteína se definió estableciendo una caja denominada “*grid*” con diferentes dimensiones de X, Y y Z, dependiendo del sitio activo de interés, con un espaciado de 1.0 Å. Para el *docking* molecular en el que se chequearon ambos dominios de la proteína incluida la zona central, se establecieron unas dimensiones de 62 x 44 x 68 Å; en cambio, cuando el cálculo se realizó sobre el sitio activo únicamente, las dimensiones del *grid* se redujeron a 40 x 36 x 38 Å.

Con todos estos datos, se ejecutó el cálculo usando el programa Autodock Vina,^[34] iniciándose varias ejecuciones de conformaciones aleatorias determinadas por el parámetro *exhaustiveness*. A través del script se ejecuta el cálculo de *docking* aprovechando múltiples CPUs, acelerando el tiempo de cálculo y generando las diferentes conformaciones que adquiere nuestro ligando alrededor del sitio de interés de la proteína, junto a su correspondiente tabla de energías de enlace que recoge la afinidad de cada uno de los confórmeros finales en kcal/mol.

Una vez construido el modelo del complejo de estudio, es importante tener en cuenta las condiciones en las que se va a desarrollar la simulación, ya que deben acercarse y reproducir las condiciones experimentales de la forma más realista posible. Para ello, la metodología empleada ha sido el programa PROPKA,^[35] que ha permitido tener en cuenta los estados de protonación de algunos aminoácidos. Uno de estos residuos a tener en cuenta ha sido la histidina, ya que, dependiendo de la carga neta y la localización de algunos de sus protones, puede adoptar tres estados diferentes. El que se ha escogido en nuestro caso ha sido el estado HIP (carga +1, con los nitrógenos δ y ϵ protonados) a pesar de que en nuestro sistema se trabaja a pH 7 y esta forma se encuentra a un pH algo más bajo. El motivo principal se debe a la interacción, ya que una histidina catiónica puede servir como un donante de protón ácido y enlace de hidrógeno general.^[36]

Estas interacciones se analizaron, a continuación, con los programas PyMol y Discovery Studio. Una vez escogido el ligando idóneo, se monta el archivo PDB del complejo que, mediante el módulo *tLEAP* empleando la versión de AMBER 14, genera los inputs de topología y coordenadas iniciales para la simulación. Para ello, se cargaron los sets de parámetros correspondientes a los campos de fuerza GLYCAM06 (para el carbohidrato) y AMBER99SB (para la proteína), neutralizando el sistema con iones cloruro en una caja de agua empleando el campo de fuerzas TIP3P.

Se comenzó con varios ciclos de minimización de los iones y moléculas de agua manteniendo tanto la proteína como el sustrato fijo, seguido de una última etapa de

minimización adicional relajando todo el sistema (ambas etapas constan de 20000 ciclos) para intentar modular las interacciones electrostáticas. A continuación, un calentamiento gradual desde 100K a 300K, en intervalos de 50 ps exceptuando el último (100 ps), y una presurización a 1 bar, lleva al sistema a las condiciones objetivo para alcanzar una fase de equilibrio. Por último, la etapa de producción, a volumen constante, es la de duración más larga. Para aquellas dinámicas en las que no se han utilizado restricciones experimentales, el tiempo de simulación es de 200ns, dividido en bloques de 50 ns para ajustarlo a los recursos del *cluster*; por el contrario, en las MD-TAR su duración es de 40 ns, dividido en bloques de 8 ns en los que se ha ido chequeando. Cabe añadir que, para este último tipo de dinámicas, se han utilizado las distancias experimentales obtenidas de los experimentos NOESY transferido como restricciones de la simulación.

Los cálculos de dinámica molecular fueron similares a los llevados a cabo para los ligandos libres y se realizaron bajo condiciones de límite periódico basados en el método *pmemd* (*Particle Mesh Ewald*). Cabe destacar que en concreto la variante *pmemd.cuda* utiliza GPUs soportando también otro tipo de cálculos como los de Born Generalizado^[37] y fueron establecidos mayormente para todas las etapas salvo las de minimización, donde se empleó el método *pmemd.MPI*.

4.5.2. Análisis de los resultados teóricos

4.5.2.1 Análisis estructural

En primer lugar, se utilizan los programas VMD, PyMOL y Discovery Studio, para la visualización de los resultados, complementándose entre sí para la representación gráfica de estructuras moleculares, ya que admiten diferentes formatos de archivo. Concretamente PyMOL es ampliamente utilizado en este tipo de cálculos para la visualización 3D de acoplamiento y la generación de mapas de densidad electrónica.^[38]

Del mismo modo que en los ligandos libres, los outputs de coordenadas y trayectorias de las etapas de producción de las dinámicas moleculares fueron procesados con el módulo *cpptraj* en la versión AMBER 12 para el análisis estructural. Las conformaciones de cada anillo de azúcar se analizaron con el comando *pucker*, basándose en la metodología de Cremer Pople; los valores de las constantes de acoplamiento se calcularon mediante la ecuación de Altona y, por último, se midieron las distancias interprotónicas intrarresiduales (H₁-H₃, H₁-H₅ y H₃-H₅) e interglucosídicas (H₁-H₄ y H₁-H₃) y los ángulos de torsión ϕ (H₁'-C₁'-OC₄/C₃) y ψ (C₁'-O-C₄/C₃-H₄/H₃).

4.5.2.1 Estudio de interacción

Para cada sistema, la energía libre de unión se calculó mediante las técnicas de MM-PBSA y MM-GBSA en función de la estructura final derivada de la etapa final de producción de las dinámicas moleculares (con y sin restricciones), donde ΔG_T se descompone en términos de: interacciones electrostáticas (ΔE_{elec}), van der Waals (ΔE_{vdw}), de solvatación (ΔG_{sol}), que contiene partes polares y partes no polares, y

finalmente, el cambio de entropía conformacional (-TΔS), que no se consideró en este trabajo debido al alto coste computacional y la baja precisión de la predicción.

Estos cálculos a menudo requieren múltiples archivos de topología que contienen parámetros correspondientes al campo de fuerza y desarrollados en disolvente explícito. Con la extensión de Python *ante-MMPBSA.py* se permite generar fácilmente estos archivos topológicos solvatados y no solvatados individuales para el ligando y la proteína, verificándolos antes de la realización de los correspondientes cálculos de energía. Dentro de los cálculos GB se encuentran una serie de variables como la concentración de sal (saltcon) y la representación de los efectos de la solvatación que puede seguir tres metodologías diferentes. La primera de ellas desarrollada por Hawkins,^[39] GB^{HCT} (en donde *igb*= 1); y las dos últimas implementadas por Onufriev,^[40] GB^{OBC1} y GB^{OBC2} (siendo *igb*= 2 y 5 respectivamente). Este último (*igb*= 5) es el que se ha utilizado para llevar a cabo los cálculos, autoguardándose instantáneas cada 10 ps extraídas de las trayectorias de la dinámica, siendo la energía de enlace final un promedio de todas éstas.

Para generar la descomposición de energía por residuo, tanto en PB como en GB, se deben definir sus variables control como *idecomp*, adquiriendo los valores 1 o 2 si se quiere estimar la contribución de cada residuo a la energía libre total; o los valores 3 y 4 si se hace por pares de residuos específicos. Nosotros hemos utilizado *idecomp*=3 especificando un subconjunto para reducir la cantidad de datos impresos. Cabe añadir que, tanto para los cálculos de MM-GBSA como MM-PBSA, se ha usado la extensión MMPBSA.py de Python todo ello bajo el programa *sander* de AMBER16.^[41]

4.5.3. Análisis de los resultados experimentales

➤ RMN

Los experimentos de RMN han sido realizados en un equipo Bruker usando un imán ASCEND III HD 600 equipado con una criosonda QCI S3 H/F-C/N-D 05-Z, previa asignación en un equipo Bruker ASCEND III 400 con sonda BBO 400S1 BBF-H-D-05 Z SP.

Los experimentos de observación del ligando NOESY transferido se realizaron a 600 MHz y a 300K registrándose con tiempos de mezcla comprendidos entre 150-800 ms. Se utilizó un filtro de cero cuanto antes de la adquisición, y dos bloques de gradientes pulsados flanqueando un pulso de 180° durante el tiempo de mezcla, además de una presaturación en la señal de agua para mejorar la relación señal/ruido. La preparación de los ligandos fue similar al de los experimentos de referencia en ausencia de proteína del **Capítulo 3** (1 x PBS en D₂O 1.5 mM en 300 μL), incluyendo 50 μM de PTN y llevándose a cabo varios ciclos de solubilización-lifilización para asegurar así la eliminación completa de H₂O. El complejo persustituido **3**-PTN no se pudo asignar completamente debido al solapamiento de las señales. En el caso del tetrasacárido **6** no se pudo registrar ningún experimento debido a su completa insolubilidad en el buffer.

Para la obtención de las distancias experimentales se asumió la aproximación de los espines aislados (ISPA).^[42] Para el cálculo de la velocidad inicial de crecimiento se registraron experimentos a tiempo de mezcla variable (100 – 800 ms). De éstos se calcularon los valores del pico de cruce dividido por el pico diagonal, al mismo tiempo de mezcla y mediante el ajuste de estos valores a la ecuación exponencial (2) del punto **3.5.3** del **Capítulo 3** se obtiene la velocidad inicial de crecimiento (σ_0) siguiendo la aproximación PANIC.^{[43][44]} Posteriormente, empleando la relación de la ecuación (3) (mismo punto **3.5.3** del capítulo 3) y una distancia de referencia ya conocida, se calculan para cada par de protones las distancias experimentales a partir de las velocidades de relajación cruzada (σ).^[45]

Los STD-NMR se adquirieron empleando una secuencia de la librería del espectrómetro. En esta, la saturación se realiza mediante un tren de pulsos gaussianos de 50 ms a baja potencia para aumentar su selectividad. Tras el primer pulso se introduce un filtro $T_{1\rho}$ de 10 ms para disminuir las señales residuales de la proteína. Finalmente se aplica un tren 3 9 19 franqueado por gradientes de campo, antes de la adquisición, para la supresión de la señal del agua. Los experimentos *on-resonance* a 0.3 ppm y *off-resonance* a 40-50 ppm se registraron alternativamente y fueron procesados independientemente y posteriormente sustraídos. Las curvas de crecimiento del STD frente al tiempo se ajustaron a la siguiente ecuación monoexponencial:

$$\text{STD}(t_{\text{sat}}) = a(1 - \exp(-bt_{\text{sat}})) \quad (2)$$

de donde se extrajo la velocidad inicial de crecimiento, STD_0 , como el producto ab . Posteriormente los valores se normalizaron asignando un valor del 100% al valor mayor.

Los experimentos ^{15}N -HSQC se han realizado empleando la biblioteca del espectrómetro para una correlación heteronuclear con doble transferencia, mejora de sensibilidad, desacoplamiento durante la adquisición y usando un pulso *flip-back*. Se preparó la muestra en un tubo de 3 mm en 300 μL de buffer 1 x PBS pH=7.8. Las demoras de los periodos de reenfoque INEPT fueron optimizadas para una constante de acoplamiento de 95 Hz. Se realizaron experimentos con 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 3, y 4 equivalentes de carbohidrato.

➤ *Polarización de Fluorescencia*

Las medidas de polarización de fluorescencia fueron llevadas a cabo en placas de microtitulación de 384 pocillos (poliestireno negro, no tratado, Corning), y los datos obtenidos fueron generados utilizando un lector de microplacas usando filtros de polarización con longitudes de onda de excitación y emisión de 485 y 535 nm, respectivamente. Como sonda fluorescente se ha utilizado un hexasacárido de heparina, preparado previamente en nuestro laboratorio, marcado con fluoresceína, disuelto en tampón PBS (10 mM, pH 7.4). La proteína empleada para el estudio, en este caso PTN, fue disuelta utilizando el mismo tampón que contenía en esta ocasión BSA (albúmina de suero bovino) al 1%. Las disoluciones de los compuestos **1-6** se prepararon empleando de nuevo el tampón PBS de concentración 10 mM (sin variar el pH). En el caso de **3** y **6** se añadió un 10% de DMSO en la disolución madre 1 mM para facilitar su solubilización.

Para el ensayo de unión directa de la sonda con la PTN, fueron necesarias preparaciones de disoluciones con concentraciones crecientes de proteína (desde 20 nM hasta 1.3 μ M) y una disolución 20 nM de sonda fluorescente, transfiriéndose a cada pocillo de la microplaca 15 μ L de cada una de ellas, para alcanzar un volumen final de muestra de 30 μ L por pocillo. Se realizaron hasta tres réplicas de cada concentración de proteína y fue necesaria la agitación en la oscuridad durante un tiempo estimado de 5 minutos antes de realizar la lectura. Como referencia se preparó un blanco, que contenía 15 μ L de PTN (650 nM) y 15 μ L de tampón PBS, cuyos valores se restaron a todos los valores de fluorescencia que se fueron midiendo en cada uno de los pocillos. Además, se midió la polarización de una muestra que contenía una disolución de la sonda en PBS con un 0.5% de BSA, que proporcionaba el valor de la polarización residual en ausencia de la proteína. Este valor se restaba a la FP de las disoluciones muestra con proteína, para obtener la diferencia de polarización (ΔP). La representación gráfica de ΔP frente a la concentración de proteína y el correspondiente ajuste de la curva se ha detallado en el **Capítulo 2** y el apartado **4.3.4** del presente capítulo.

Para los experimentos de competición se midió la FP de pocillos que contenían: 20 μ L de la disolución de PTN 162.5 nM y 10 μ L de la sonda 40 nM en presencia de 10 μ L de solución de inhibidor de concentraciones progresivas desde 0.1 μ M a 2 mM. El volumen total de muestra de cada pocillo fue, por tanto, de 40 μ L con una composición final de tampón PBS + BSA al 0.5%, con concentraciones finales de sonda y PTN de 10 nM y 63 nM, respectivamente, mientras que la de inhibidor varió de 500 μ M a 25 nM. Una vez que todos los pocillos están preparados, se agita la microplaca en la oscuridad durante 5 min y se va registrando la polarización de cada uno de ellos para la determinación de los valores de IC_{50} . Estos valores de polarización se representaron frente al logaritmo de la concentración de inhibidor, generando una curva que se ajustó a la ecuación (1) que corresponde a un único sitio de competición:

$$y = A_2 + (A_1 - A_2) / [1 + 10^{(x - \log IC_{50})}] \quad (1)$$

donde IC_{50} corresponde a la concentración de competidor que da como resultado una inhibición del 50%, mientras que A_1 y A_2 son los valores máximos y mínimos de polarización que corresponden a los dos controles incluidos en nuestros experimentos de competición. El primero de ellos contiene PTN y sonda en ausencia de inhibidor, por lo que muestra un valor de polarización máximo, mientras que el segundo control sólo presenta sonda fluorescente, proporcionando el valor de polarización mínimo esperado para una inhibición del 100%.

Bibliografía Capítulo 4

- [1] D. Nikitovic, M. Assouti, M. Sifaki, P. Katonis, K. Krasagakis, N. K. Karamanos, G. N. Tzanakakis, *Int. J. Biochem. Cell Biol.* **2008**, *40*, 72–83.
- [2] S. Mizumoto, D. Fongmoon, K. Sugahara, *Glycoconj. J.* **2013**, *30*, 619–632.
- [3] N. Maeda, N. Fukazawa, T. Hata, *J. Biol. Chem.* **2006**, *281*, 4894–4902.
- [4] C. González-Castillo, D. Ortuño-Sahagún, C. Guzmán-Brambila, M. Pallàs, A. E. Rojas-Mayorquín, *Front. Cell. Neurosci.* **2015**, *8*, 1–7.
- [5] E. Ryan, D. Shen, X. Wang, *FEBS Lett.* **2021**, *595*, 925–941.
- [6] E. Ryan, D. Shen, X. Wang, *FEBS J.* **2016**, *283*, 1488–1503.
- [7] A. M. J. Garcia-jimenez, S. Gil-caballero, F. Corzana, F. Juarez-vicente, J. R. Miles, K. Sakamoto, K. Kadomatsu, M. Dominguez-garcía, J. L. De Paz, P. M. Nieto, **2021**, *27*, 12395-12409.
- [8] R. Van Der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright, M. M. Babu, *Chem. Rev.* **2014**, *114*, 6589–6631.
- [9] S. A. Hollingsworth, P. A. Karplus, *Biomol. Concepts* **2010**, *1*, 271–283.
- [10] H. Choi, H. Park, K. Son, H. M. Kim, Y. Jung, *Chem. Sci.* **2019**, *10*, 10428–10435.
- [11] M. D. George, V. Adrian, *J. Chem. Educ.*, **1993**, *70*, 1–16.
- [12] W. P. Aue, E. Bartholdi, R. R. Ernst, *J. Chem. Phys.* **1976**, *64*, 2229–2246.
- [13] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 521–528.
- [14] I. V. Zhukov, A. S. Kiryutin, F. Ferrage, G. Buntkowsky, A. V. Yurkovskaya, K. L. Ivanov, *J. Phys. Chem. Lett.* **2020**, *11*, 7291–7296.
- [15] P. Kiraly, R. W. Adams, L. Paudel, M. Foroozandeh, J. A. Aguilar, I. Timári, M. J. Cliff, M. Nilsson, P. Sándor, G. Batta, J. P. Waltho, K. E. Kövér, G. A. Morris, *J. Biomol. NMR* **2015**, *62*, 43–52.
- [16] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* **1980**, *69*, 185–189.
- [17] G. M. Clore, A. M. Gronenborn, *J. Magn. Reson.* **1985**, *61*, 158–164.
- [18] C. Solera, G. Macchione, S. Maza, M. M. Kayser, F. Corzana, J. L. De Paz, P. M. Nieto, *Chem. A Eur. J.* **2016**, *22*, 2356–2369.
- [19] V. H. Pomin, X. Wang, *Molecules* **2018**, *23*(9), 2314.
- [20] C. P. Interactions, J. L. Ortega-roldan, M. Blackledge, *Methods Mol. Biol.*, **2018**, *1764*, 73–85.
- [21] W. Becker, K. C. Bhattiprolu, N. Gubensäk, K. Zangger, *Chem. Phys. Chem.*, **2018**, *19*, 895–906.
- [22] X. Wang, *Adv. Clin. Chem.* **2020**, *98*, 51-89.

- [23] S. Maza, M. Mar Kayser, G. MacChione, J. López-Prados, J. Angulo, J. L. De Paz, P. M. Nieto, *Org. Biomol. Chem.* **2013**, *11*, 3510–3525.
- [24] S. Maza, N. Gandía-Aguado, J. L. de Paz, P. M. Nieto, *Bioorganic Med. Chem.* **2018**, *26*, 1076–1085.
- [25] J. L. De Paz, P. M. Nieto, *Org. Biomol. Chem.* **2016**, *14*, 3506–3509.
- [26] C. Mikelis, M. Lamprou, M. Koutsioumpa, A. G. Koutsioubas, Z. Spyranti, A. A. Zompra, N. Spiliopoulos, A. A. Vradis, P. Katsoris, G. A. Spyroulias, P. Cordopatis, J. Courty, E. Papadimitriou, *J. Cell. Biochem.* **2011**, *112*, 1532–1543.
- [27] N. S. Gandhi, R. L. Mancera, *Chem. Biol. Drug Des.* **2008**, *72*, 455–482.
- [28] i) Discovery Studio Visualizer: <https://discover.3ds.com/discovery-studio-visualizer-download>.
- [29] J.-S. Gómez-Jeria, A. Robles-Navarro, G. Kpotin, N. Garrido-Sáez, G.-D. Nelson, *Chem. Res. J.* **2020**, *5*, 32–52.
- [30] H. G. Petersen, *Phys. Fluids* **1967**, *10*, 3668–3679.
- [31] J. Racine, *J. Appl. Econom.* **2006**, *21*, 133–141.
- [32] S. Genheden, U. Ryde, *Expert Opin. Drug Discov.* **2015**, *10*, 449–461.
- [33] S. Cosconati, S. Forli, A. L. Perryman, R. Harris, D. S. Goodsell, A. J. Olson, *Expert Opin. Drug Discov.* **2010**, *5*, 597–607.
- [34] J. Kim, J. Kwon, M. Kim, J. Do, D. Lee, H. Han, *Polym. J.* **2016**, *48*, 829–834.
- [35] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295.
- [36] J. A. Hadden, J. R. Perilla, *Methods Mol. Biol.* **2018**, *1762*, 245–270.
- [37] D. A. Case, D. S. Cerutti, T. E. I. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman in *AMBER Manual*, University of California, San Francisc., **2017**.
- [38] D. Seeliger, B. L. De Groot, *J. Comput. Aided. Mol. Des.* **2010**, *24*, 417–422.
- [39] G. D. Hawkins, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem.* **1996**, *100*, 19824–19839.
- [40] A. Onufriev, D. Bashford, D. A. Case, *Proteins Struct. Funct. Genet.* **2004**, *55*, 383–394.
- [41] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.
- [42] S. Macura, B. T. Farmer, L. R. Brown, *J. Magn. Reson.* **1986**, *70*, 493–499.

- [43] J. C. Muñoz-García, F. Corzana, J. L. De Paz, J. Angulo, P. M. Nieto, *Glycobiology* **2013**, *23*, 1220–1229.
- [44] H. Hu, K. Krishnamurthy, *J. Magn. Reson.* **2006**, *182*, 173–177.
- [45] D. Neuhaus, M. P. Williamson in *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, **2000**.

Capítulo 5

Estudios estructurales y de interacción experimental y teórica de los miméticos pentasacarídicos de CS-E con MK

5.1 Introducción

Siguiendo con el estudio de proteínas que interaccionan con GAGs para regular así una gran variedad de procesos biológicos, en este capítulo nos centraremos en estudiar la interacción con Midkina (MK). Esta citoquina pertenece a la misma familia que la PTN. Ambas constituyen la familia de factores de crecimiento denominada “*neurite growth-promoting factor*”.^{[1],[2],[3],[4]} Se sabe que esta proteína presenta un sitio de reconocimiento de carbohidratos, en concreto de heparina y CS, y que además su dimerización mejora notablemente sus actividades biológicas.^[5] También se ha demostrado que la interacción entre MK y CS está mediada por secuencias de oligosacáridos específicas, con un motivo de sulfatación particular.^[6] Diversos estudios han confirmado que concretamente el CS de tipo E, que presenta una unidad de disacárido con *N*-acetilgalactosamina 4,6-disulfatada, se une a esta citoquina con una afinidad tan fuerte como la heparina.^{[7],[8]} De la misma forma que ocurría en la PTN, la orientación del ligando en el complejo es muy variable, orientándose dentro de la cavidad central de la proteína, constituida por dos dominios estructurados de láminas β unidos por una región intrínsecamente desordenada.^[9] En la Figura 5.1 se ha representado tridimensionalmente la superposición de todas las estructuras de MK resueltas por RMN, primero sobre el NTD (izquierda) y después sobre el CTD (derecha). Debido a la flexibilidad de la MK, al superponer un dominio se desordena el otro y viceversa. Esto está en consonancia con la propuesta como IDR del segmento de unión de ambos dominios.

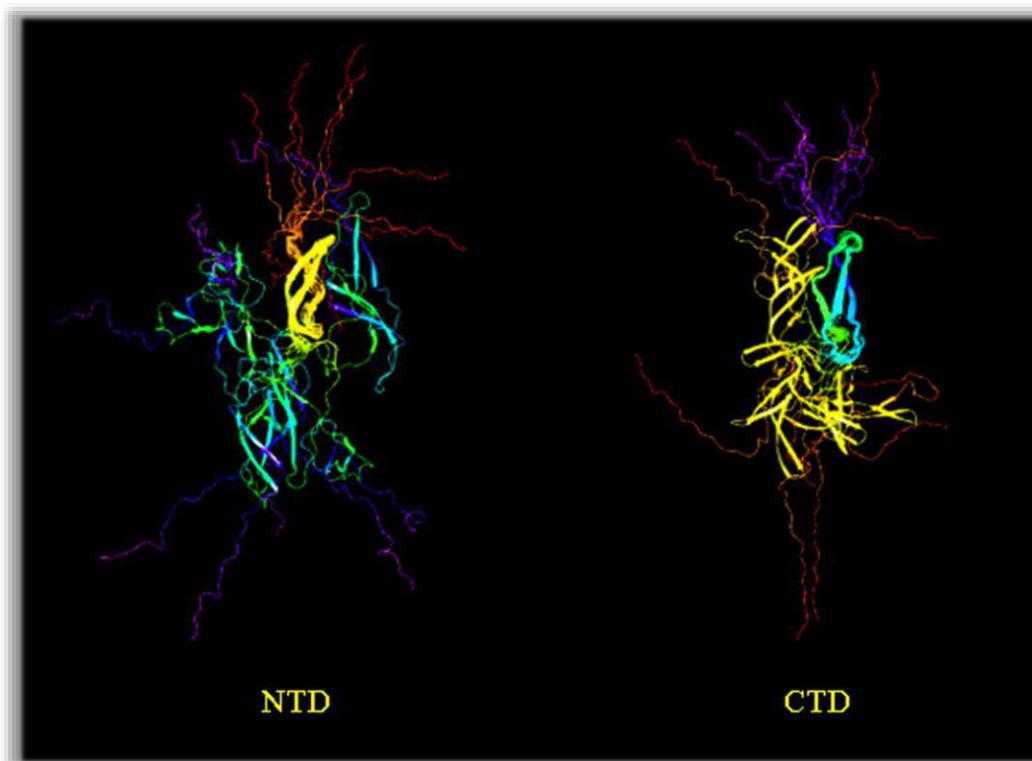


Figura 5.1. Superposición tridimensional de ambos dominios de la MK procedentes de las diez conformaciones resueltas por RMN.

Teniendo en cuenta estas premisas y los estudios previos de interacción realizados en nuestro grupo con otras secuencias oligosacarídicas,^[9] se decidió completar el estudio de la interacción entre los miméticos de CS-E y la MK. Para ello se realizó el estudio experimental por RMN y FP de los complejos, que se completó con un estudio teórico, más breve con respecto al mostrado en el **Capítulo 4**, ya que se utilizó únicamente una de las series, los ligandos pentasacarídicos. De igual manera que se ha llevado a cabo con los complejos con PTN, se escogieron de nuevo para la MK tres conformaciones representativas, para examinar la influencia de la conformación de la MK sobre la estructura de estos ligandos miméticos de CS-E. Para este análisis se calcularán de nuevo las estructuras de los oligosacáridos asociados, en este caso con MK, empleando las mismas técnicas que en el **Capítulo 4**. Los resultados de cálculos de dinámica molecular de los complejos, en primer lugar sin restricciones, permitirán deducir las conformaciones de ligando involucradas, observando los ángulos de torsión, constantes de acoplamiento y las coordenadas de *puckering*. En caso de discrepancias entre los resultados teóricos y los resultados experimentales se recurrieron a cálculos MD-TAR. El análisis se complementa con estudios de *docking* molecular que permiten generar los diferentes conformeros de mayor afinidad por la MK, explorando su orientación hacia ésta; y en último lugar se han realizado cálculos de MM-PBSA/MM-GBSA a partir de estas trayectorias, para conocer las interacciones que tienen lugar y las energías de estas afinidades.

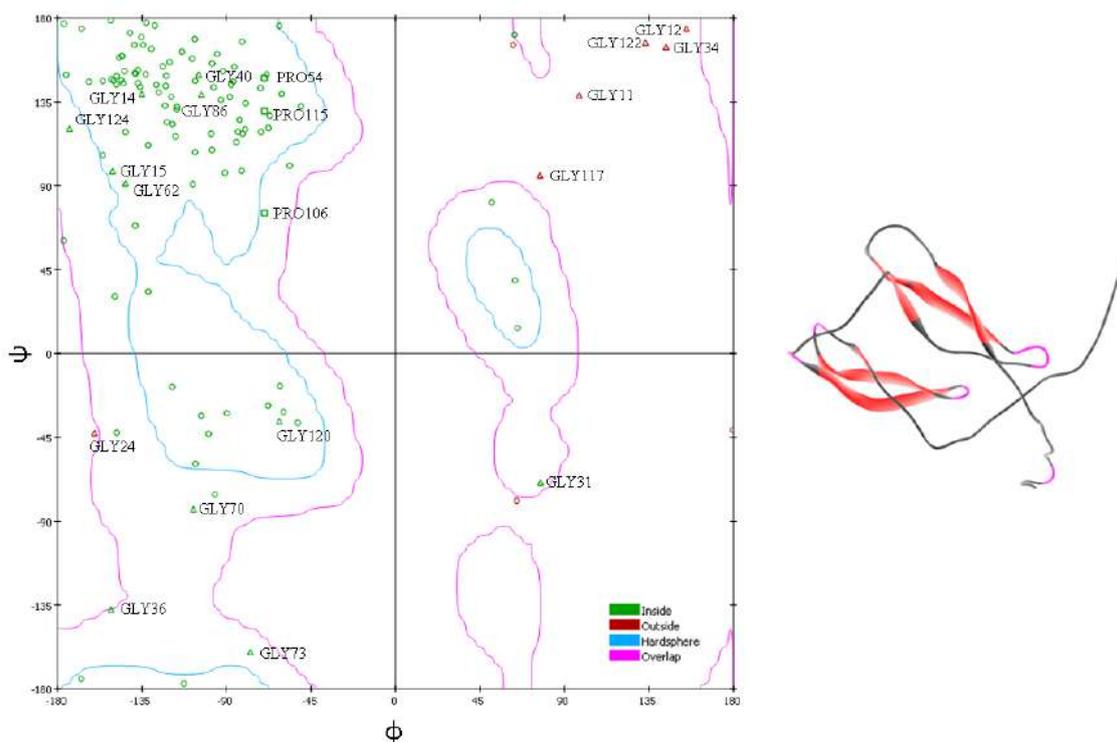
5.2 Estructura de la Midkina

Se han usado las coordenadas depositadas en la base de datos de estructuras de proteínas para la MK de pez cebra (isoforma MK-a: 2LUT).^{[10],[11],[12]} En este caso se han obtenido dos estructuras, una para cada isoforma y, al igual que en la PTN, la MK-a se ha obtenido por RMN y consta de diez estructuras compatibles con las restricciones experimentales. Esta proteína está compuesta por dos dominios de lámina β : N-terminal (Gly15-Arg52) y C-terminal (Gly62-Thr104), conectados entre sí por una zona desordenada central denominada *hinge* (Asn54-Asp62), ambos enriquecidos en aminoácidos básicos y cisteínas.^{[13],[5]} Dado que el sitio de unión más importante para GAG está localizado en el dominio C-terminal (residuos 62-104), la secuencia Leu-Lys-Lys-Ala-Arg-Phe (residuos 85-90) puede jugar un rol importante en esta interacción debido a la flexibilidad encontrada en esta zona de la proteína.^{[13],[14]} El motivo principal por el cual el dominio N-terminal presentaba una interacción más débil frente a GAG era que los aminoácidos básicos de este dominio no se agrupan en *clusters* para interaccionar con los grupos sulfatos de los ligandos de azúcar, además de contener también residuos ácidos que desfavorecían la unión.^{[13],[15]}

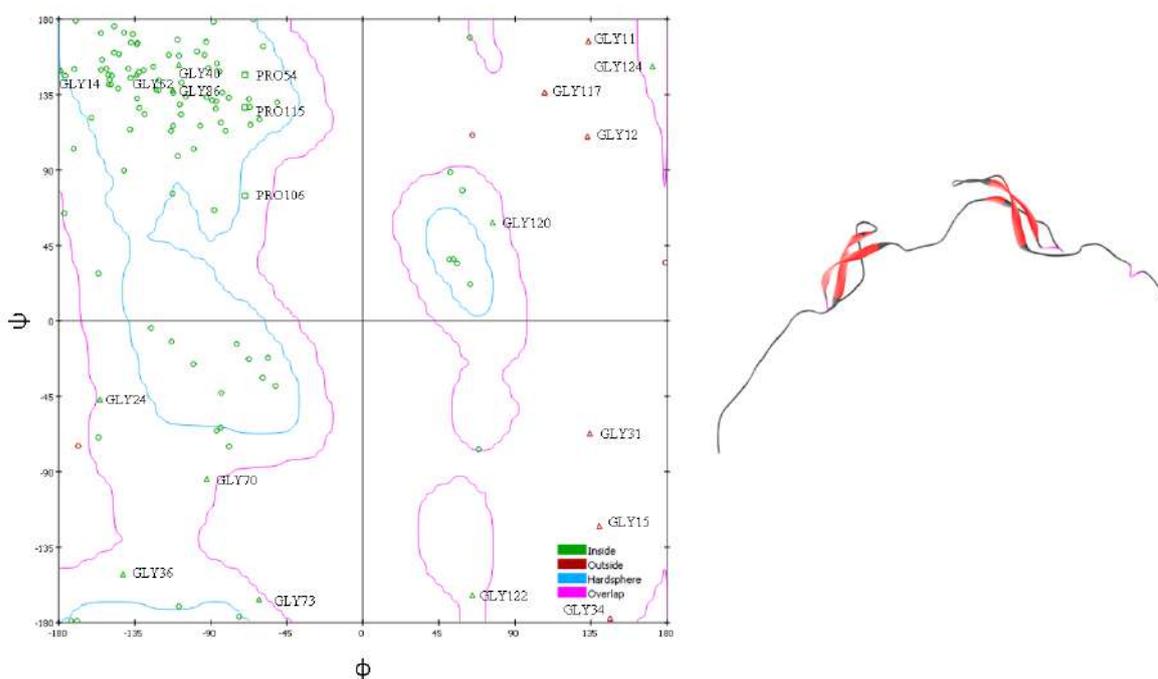
Estudios recientes han determinado que tanto la zona *hinge* como residuos del dominio N-terminal (15-52) también participan activamente en la unión a derivados de heparina y que el CS-E se une simultáneamente a ambos dominios NTD y CTD de la MK.^{[16],[17],[18]} A raíz de estos resultados, y con los conocimientos que teníamos de trabajos anteriores en el grupo, se consideró interesante conocer las preferencias de unión de los miméticos

1 – 3 frente a la estructura completa de MK mediante técnicas de *docking* molecular. Del mismo modo que con la PTN, se han escogido para el estudio tres conformaciones de la MK representativas del espacio conformacional de la proteína y se han representado los mapas de Ramachandran de cada una de ellas para garantizar que los conformeros escogidos son adecuados para llevar a cabo los posteriores cálculos (Figura 5.2).

➤ *Conformación plegada de la midkina (MK4)*



➤ *Conformación lineal de la midkina (MK0)*



➤ *Conformación de bisagra de la midkina (MK2)*

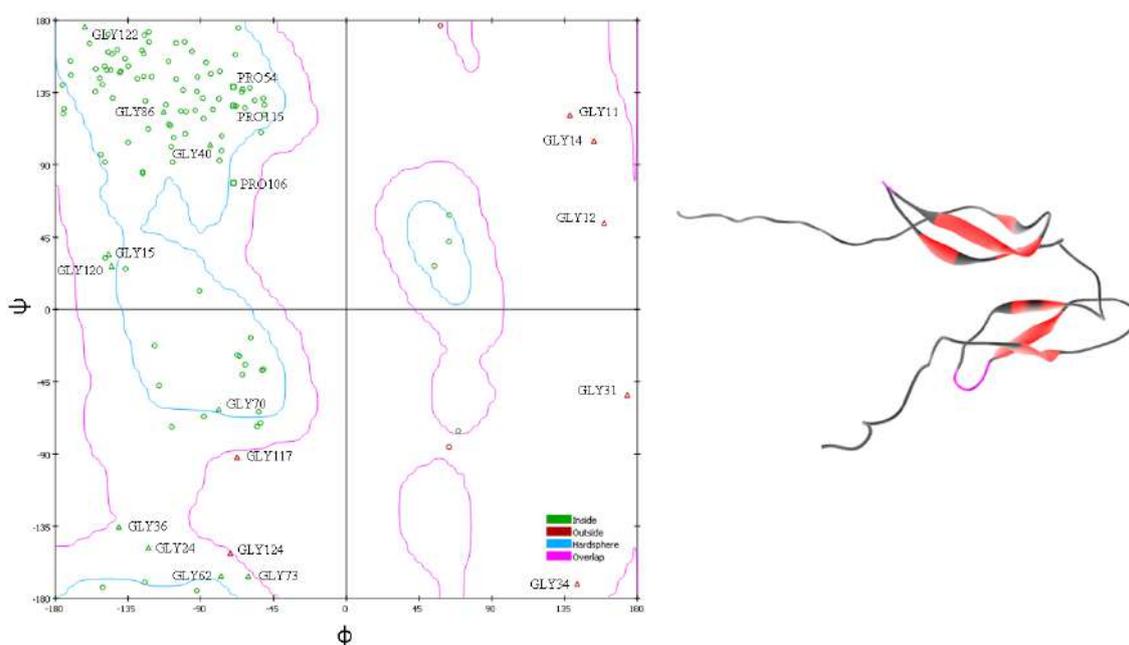


Figura 5.2. A la izquierda, mapas de Ramachandran individuales de cada conformación, y a la derecha las estructuras tridimensionales de la MK empleadas en esta Tesis.

Los diferentes diagramas presentan conformaciones probables, en donde la flexibilidad de esta proteína se refleja en el alto número de glicinas que contiene, situadas la mayoría de éstas en las zonas más favorables, siendo minoritaria la presencia de prolinas que le

otorgan rigidez. Las prolinas suelen encontrarse en puntos de especial importancia estructural debido a que su cadena lateral cíclica induce sobre la estructura un giro inverso que se ha demostrado que puede acomodar el sitio de unión al proporcionar mayor accesibilidad a los aminoácidos adyacentes, mejorando así su unión a GAGs.^[19]

Estos análisis previos se utilizaron como punto de partida para los cálculos de dinámica molecular y para el cálculo de las energías de afinidad para cada mimético, centrándonos en la principal zona de unión a CS, para observar la flexibilidad del ligando al complejarse. Todos estos estudios fueron repetidos para cada uno de los confórmers de la MK representados en la Figura 5.2.

5.3 Estudio experimental

Con la experiencia que tenía el grupo en el estudio de interacción de GAGs con MK, se procedió a caracterizar los complejos por técnicas de RMN de observación del ligando, como las empleadas en el capítulo anterior para PTN, utilizando de nuevo muestras 1.0 mM de ligando, 20 μ M de MK, en buffer 1x PBS a pH = 7.8.

En primer lugar, se expondrán los resultados experimentales de NOESY transferido. De esta manera se comprobará si existe interacción por el cambio en el volumen de los picos de cruce. Si hubiera un cambio conformacional en el mimético inducido por la formación del complejo éste se reflejará en un cambio en el patrón de los picos de cruce NOE. En segundo lugar, se describirán los resultados de los experimentos de STD, que indicarán las regiones del ligando más involucradas en esta asociación.

Finalmente, se cuantificará la afinidad a través de los ensayos de competición por Polarización de fluorescencia, que nos permitirán comparar los resultados de IC₅₀ con los que se obtuvieron en el **Capítulo 4** para PTN. Con estos estudios se podría determinar si estos ligandos presentan una afinidad diferente con respecto a la PTN, y si de nuevo la interacción vuelve a encontrarse repartida por todo el esqueleto del carbohidrato. Además, se podrán comparar los resultados de estos miméticos de CS con los de oligosacáridos de secuencia natural, revelando los cambios que podrían tener lugar al sustituir sus carboxilatos por grupos CH₂OH.^[20]

5.3.1 Experimento de NOESY transferido

En vista de la experiencia que tenía el grupo analizando otras series de GAGs se procedió a la realización de experimentos de NOESY transferido en presencia de MK. Al igual que en el capítulo anterior y en estudios previos sobre otros oligosacáridos la presencia de MK no influye apreciablemente en la estructura de los ligandos enlazados.^{[9],[16]} Del mismo modo que en presencia de la PTN, el espectro del NOESY transferido de **1** con MK (A.54 del Anexo) apoya conformaciones de los anillos ⁴C₁ por la presencia de picos NOE entre los protones H1-H3, H1-H5 y H3-H5 en disposición *syn*-diaxial. Adicionalmente, las uniones interglucosídicas mantienen la geometría *syn-ψ/exo-φ* en ambos tipos de enlace glicosídico. En el caso del complejo con el azúcar dibencilado **2**,

el solapamiento de las señales fue más severo, impidiendo la determinación de algunos de los picos de cruce más característicos. En este caso sólo se pudieron asignar los NOEs entre los protones interglucosídicos E1-D4 y A3-B1 característicos de una geometría *syn* entre esos anillos. Pero el resto de contactos intrarresiduales H1-H3, H1-H5 y H3-H5 apoyaban los resultados de conformaciones tipo silla 4C_1 (ver Figura A.60 del Anexo). Por último, para el complejo con el ligando **3** completamente protegido, la determinación de los picos más relevantes resultó más compleja, debido a la baja resolución del espectro (Figura A.66 del Anexo).

Sólo fue posible determinar una geometría *syn* para el enlace interglucosídico C1-B4; el resto de contactos NOE intrarresiduales que se identificaron fueron compatibles una vez más con conformaciones 4C_1 en los anillos de Glc y GalNAc. En este caso, no se usaron curvas de crecimiento para la estimación precisa de distancias interprotónicas, asumiendo que la conformación enlazada era la misma que en presencia de PTN. Por lo tanto, las restricciones experimentales que se utilizarán en el estudio teórico de MD-TAR serán los datos de las tablas de distancias del **Capítulo 4**.

5.3.2 Experimento de STD (*Saturation Transfer Difference*)

Para estudiar la interacción de estos miméticos de CS-E con la MK se registraron los experimentos de STD para los complejos de los oligosacáridos desprotegidos y dibencilados (**1-MK**, **4-MK** y **2-MK**, **5-MK**, respectivamente) a diferentes tiempos de saturación. Mediante la generación de las curvas de crecimiento se obtendrían los valores de intensidad de saturación independientes de la relajación. La metodología empleada para conseguir esta información ha sido la misma que la descrita en el punto **4.5** del **Capítulo 4**. Una vez más, en los complejos con los ligandos completamente protegidos **3** y **6** no se pudo obtener información sobre su capacidad de interacción. En el caso del complejo con el tetrasacárido **6**, su baja solubilidad en el tampón hizo inviable la medición; para su homólogo pentasacarídico **3**, debido al alto solapamiento y al ensanchamiento de las señales, no se pudieron generar curvas de crecimiento de suficiente calidad para extraer datos cuantitativos.

Los resultados del resto de ligandos, **1**, **2**, **4** y **5**, sugieren un comportamiento desordenado, tanto en la serie pentasacarídica como tetrasacarídica. La transferencia de magnetización se distribuye a lo largo de toda la cadena oligosacarídica, desde todas las direcciones del espacio. Este comportamiento ha sido asignado a la coexistencia de diferentes geometrías en el sitio de unión, véase el análisis teórico posterior. A raíz de los porcentajes de saturación que se obtuvieron de las curvas de crecimiento, se pudo determinar que los complejos con ambas citoquinas, PTN y MK, presentaban un comportamiento similar. En la Figura 5.3 se han representado los mapas de epítopo de los complejos entre **1**, **2**, **4** y **5** y MK. Además, se muestran para cada compuesto las curvas de crecimiento, obtenidas al representar los diferentes porcentajes de STD para cada tiempo de saturación, indicando aquellos protones que reciben más magnetización de la MK. Estos resultados pueden indicar tanto una disposición variable del ligando dentro del complejo como una variación de la estructura de la proteína que lo rodea.

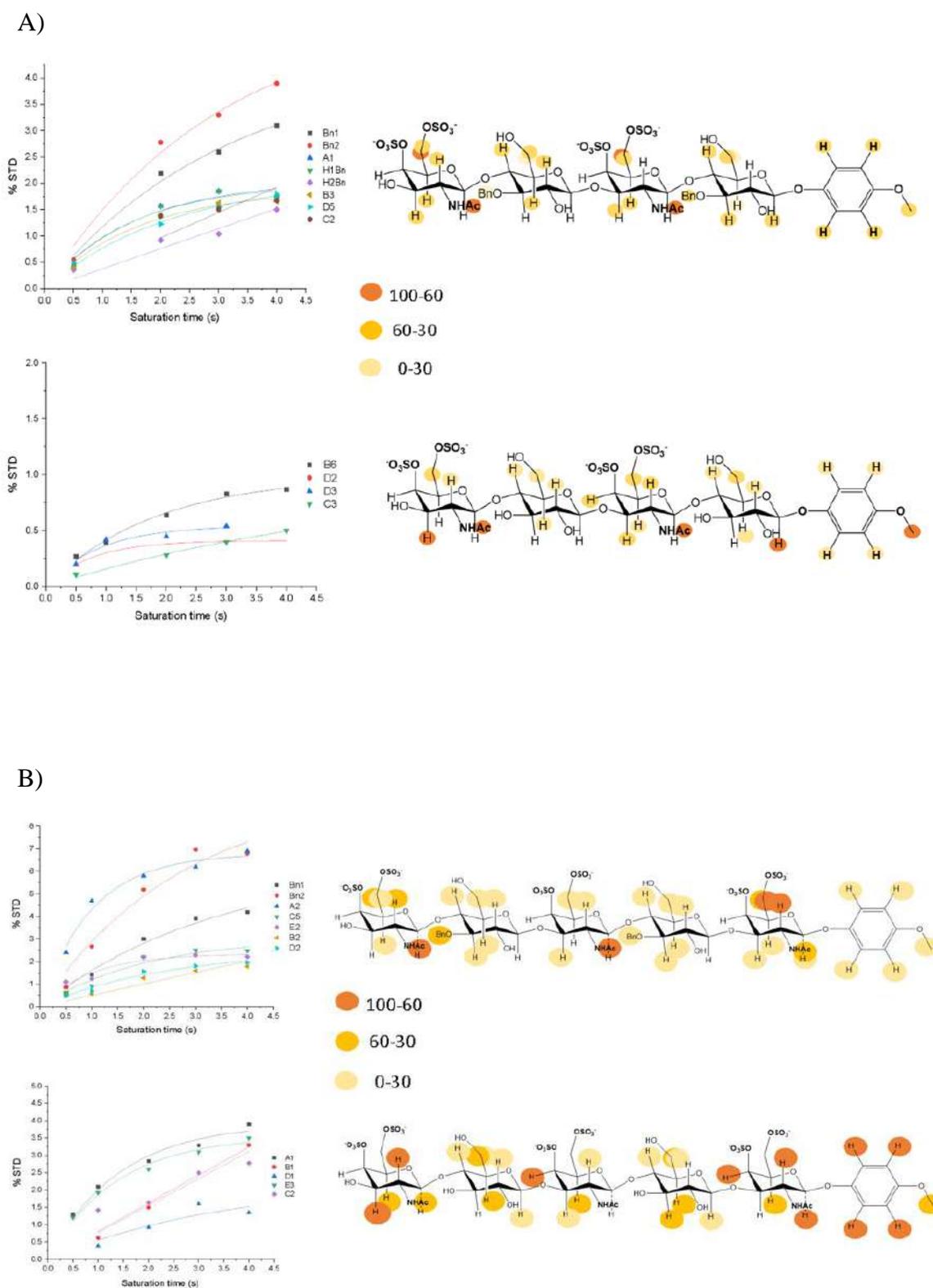


Figura 5.3. Curvas de crecimiento de STD (izquierda) y mapas de epítipo (derecha) para: A) Los complejos tetrasacáridicos (5-MK y 4-MK). B) Los complejos pentasacáridicos (1-MK y 2-MK).

5.3.3 Experimento de Polarización de Fluorescencia

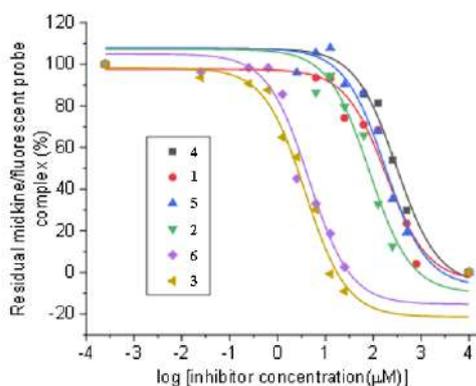


Figura 5.4. Representación de las curvas de inhibición para los seis compuestos con MK.

A continuación, se procedió a estudiar la interacción de los oligosacáridos con la MK de manera cuantitativa, del mismo modo que se analizaron las interacciones con la PTN, usando ensayos de competición de polarización de fluorescencia (ver Figura 5.4). Se decidió utilizar la misma sonda fluorescente, un hexasacárido de heparina,^{[21],[22]} ya que también mostró tener una fuerte interacción con esta citoquina (ver **Capítulo 2**, Figura 2.19, interacción MK/sonda fluorescente). A continuación, se calcularon las afinidades de unión relativas de los oligosacáridos **1-6** midiendo la polarización de fluorescencia de muestras que contenían concentraciones fijas de proteína y sonda, y concentraciones crecientes de ligando. Como se muestra en la Figura 5.4, se representaron los valores de FP normalizados frente a la concentración de oligosacárido. Las curvas de inhibición obtenidas se ajustaron al modelo de competición para un único centro activo, calculando los valores de IC_{50} de cada mimético (Tabla 5.1).^[20]

Los datos indicaron una vez más que a medida que aumentaba la presencia de grupos hidrofóbicos en la estructura del azúcar, la afinidad con la MK mejoraba. Así, los compuestos **1** y **4**, completamente desprotegidos, se unen a esta proteína con una afinidad en el rango alto micromolar, a diferencia de los ligandos persustituídos **3** y **6** que llegaron a mostrar un aumento en la inhibición de hasta 50 veces. Cabe añadir, además, que la interacción es algo mayor para la serie pentasacarídica que para la tetrasacarídica. Estos datos sugieren que compuestos como **3** y **6** pueden ser útiles para modular interacciones CS-proteína y las correspondientes funciones biológicas.

Además, los valores de IC_{50} obtenidos son similares a los mostrados por oligosacáridos de CS-E que contienen unidades de ácido glucurónico en su estructura,^[23] por lo que estos resultados indican que los derivados **1-6** pueden considerarse miméticos de las secuencias naturales.

Pentasacáridos	IC ₅₀ (μ M)	Tetrasacáridos	IC ₅₀ (μ M)
1	130 \pm 80	4	189 \pm 107
2	83 \pm 1	5	154 \pm 35
3	3.8 \pm 0.1	6	3.6 \pm 1.3

Tabla 5.1. Representación de los valores de IC₅₀. Los errores representan las desviaciones estándar de las repeticiones de los ensayos realizados en cada caso.^[20]

Si comparamos estos datos con los que se midieron con la PTN, se puede afirmar que, en general, los oligosacáridos presentan mayor interacción con la MK, aunque la diferencia en afinidades es pequeña. Esta conclusión coincide además con los valores de interacción entre la sonda fluorescente y la proteína: mientras que el valor de K_D de la interacción sonda/MK era 44 nM, la afinidad con la PTN era más débil (K_D = 125 nM).

5.4 Estudio teórico

Como ya vimos en el capítulo anterior, el estudio teórico complementario que se llevó a cabo nos permitió conocer más a fondo cómo tiene lugar esta interacción y de qué manera puede verse afectado el ligando en su estado enlazado, con datos cuantitativos a nivel de estructura, energía de afinidad y energía libre de unión. Cabe añadir que ya se disponía de información experimental por trabajos anteriores del grupo acerca del comportamiento de oligosacáridos derivados de CS en presencia de MK,^{[16],[9]} por lo que en esta ocasión se decidió llevar a cabo el estudio teórico únicamente para la serie pentasacáridica **1**, **2** y **3**. Nuestra idea era conseguir de una forma más simplificada información suficiente que pudiera ser comparada con los resultados obtenidos con la otra citoquina, PTN, y poder definir así los modos de interacción y afinidades con miméticos de CS de tipo E.

5.4.1 Docking molecular

Sabiendo que ambos dominios de lámina β de la MK podían interactuar con los derivados de CS-E, se realizaron en primer lugar estudios de *docking* sobre las estructuras de la serie pentasacáridica, **1**, **2** y **3**. Los análisis se llevaron a cabo utilizando de nuevo el programa AutoDock y Autodock Vina, considerando toda la estructura de la proteína, para conocer las preferencias de los ligandos por los posibles bolsillos de unión, con una caja (*grid*) de dimensiones 70 x 66 x 72 Å. Los resultados para los tres complejos se agruparon en conjuntos de 9 poses que ofrecían la mejor energía de unión para cada una de las conformaciones de la MK. Según la metodología utilizada previamente, se mantuvo la estructura de la proteína rígida y la del ligando flexible durante el cálculo.

- Conformación plegada de la MK (MK4)

Debido a la similitud entre ambas citoquinas, y a los resultados previos de RMN, se asumió que la MK también poseía tres sitios activos con preferencia a interactuar con GAGs (dominios C y N-terminal y el área central que los une).^{[16],[17]} Primero, se describen los resultados del docking molecular de los tres ligandos pentasacarídicos, **1**, **2** y **3**, con la conformación más plegada de esta proteína (MK4). En la Figura 5.5 se puede observar que existe una preferencia de los confórmeros de los tres ligandos oligosacarídicos por el dominio-C terminal y la zona central o *hinge*. Sin embargo, debido al plegamiento de la estructura proteica, también se establecen interacciones con aminoácidos situados en el dominio NTD. Si se comparan estos resultados con los de la simulación llevada a cabo con la conformación plegada de la PTN, ésta última no mostraba una tendencia tan clara en la serie pentasacarídica, y las estructuras se encontraban mucho más deslocalizadas.

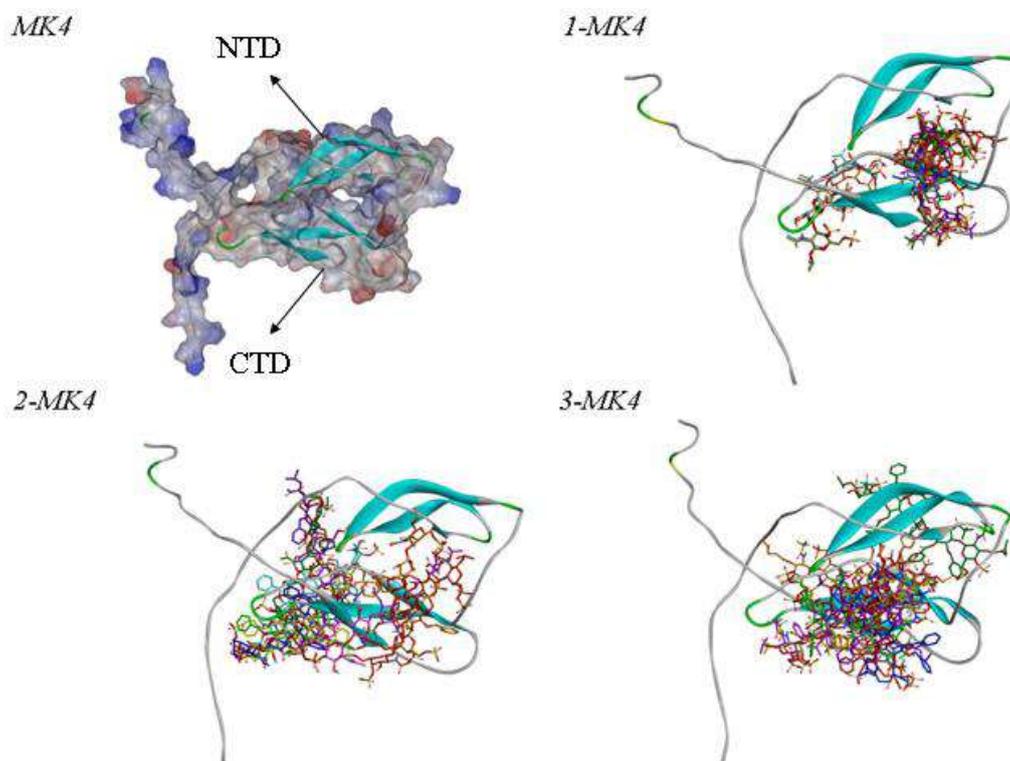


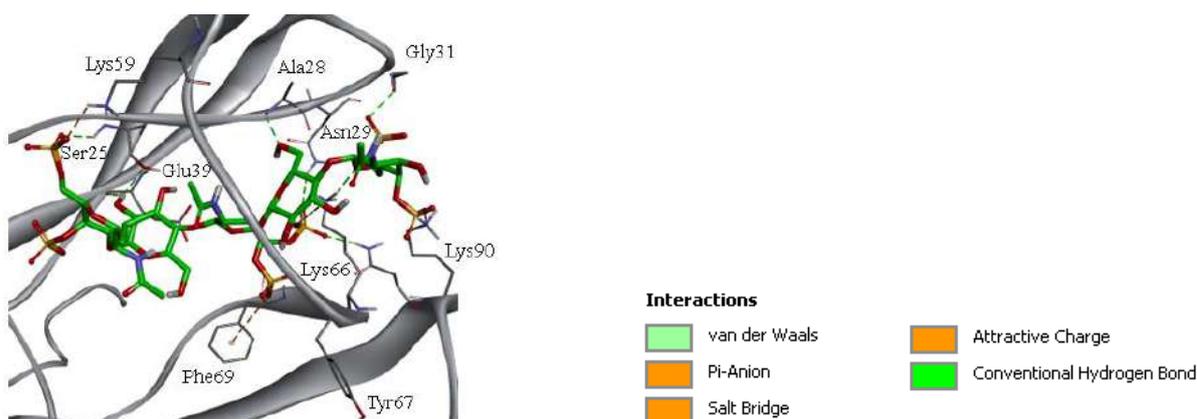
Figura 5.5. Representación de la estructura tridimensional de la MK es su forma plegada (MK4) y sus complejos con **1**, **2** y **3** obtenidos por cálculos de docking molecular (9 conformaciones oligosacarídicas).

En la Figura 5.6 se muestran con más detalle las interacciones más predominantes ligando-proteína de esta serie, representándose el confórmero con mayor energía de afinidad por la MK. Se puede observar que, al igual que ocurría con la PTN, los aminoácidos, véase Ser, Lys y Gly, establecen uniones mediante enlaces de hidrógeno o

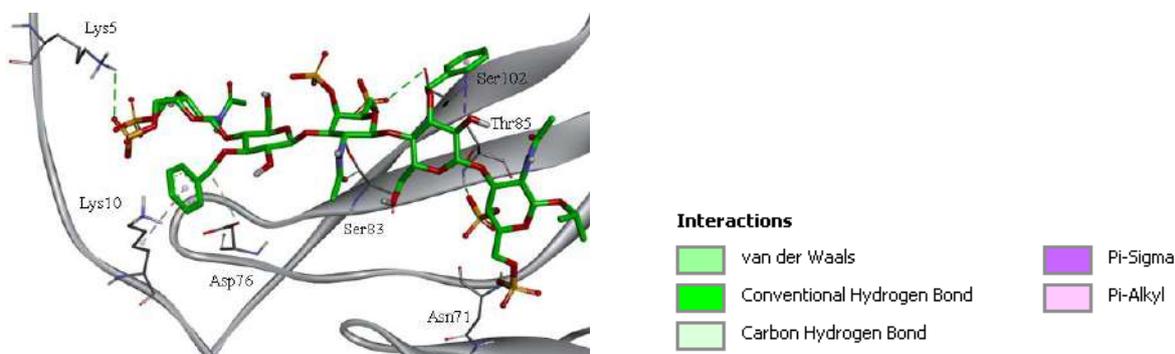
interacciones electrostáticas con los grupos sulfato del oligosacárido. Siguiendo con la clasificación establecida por el programa Discovery Studio Visualizer, en el Anexo II se han descrito con más detalle algunas de las interacciones que tienen lugar. La flexibilidad de la cadena probablemente permita una interacción mayor entre los grupos sulfato aniónicos del carbohidrato y los aminoácidos básicos, dando lugar a las interacciones carga-carga por puente salino. También se observa que los dominios de lámina β contienen aminoácidos tales como Asp y Glu que son capaces de formar enlaces de hidrógeno. A estas interacciones se suman otras, ya descritas en el capítulo anterior, como son las existentes entre un heteroátomo e hidrógeno, $X-H \cdots Y$ ($X=O$ o N ; $Y=O$, N o halógeno), o $C-X \cdots B-Y$ siendo X en nuestro caso un átomo de Flúor con el grupo $C=O$ del aminoácido.^[24]

Como ya se había visto anteriormente, a medida que se añaden grupos aromáticos en la estructura del oligosacárido, aparecen nuevas interacciones favorables. Éstas son responsables del aumento de su afinidad por la citoquina, que se ha observado en las medidas de FP (véase tabla 5.1). Para esta conformación de proteína encontramos interacciones con los sistemas aromáticos, véase, π -alquil, π -anión π -catión y π - σ .

A)



B)



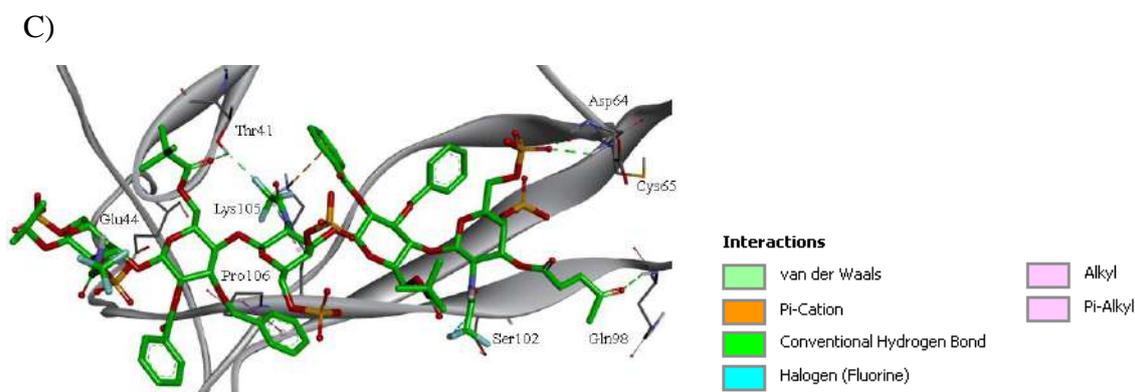


Figura 5.6. Imágenes de las mejores poses de docking obtenidas para los complejos con el ligando desprotegido, **1**, (A), dibencilado, **2**, (B) y persustituido, **3**, (C) con la conformación plegada de la MK, mostrando las interacciones intermoleculares y sus tipos. Se muestra la conformación del ligando de mayor energía de interacción.

- Conformación abierta de la MK (MK0)

Los resultados del *docking* sobre la conformación MK0 muestran una clara predisposición por el dominio-N terminal (Figura 5.7). Si se comparan estos resultados con los descritos en el capítulo anterior para el complejo con PTN, se observan ciertas diferencias, ya que en PTN las diferentes poses se distribuían a lo largo de los bolsillos de unión con cierta predilección por el dominio C-terminal, mientras que, en este caso, todos los complejos se encuentran exclusivamente en el NTD (Figura 5.7). Aunque se conocen estudios que confirman la interacción de oligosacáridos de CS en este dominio,^{[17],[18]} para los cálculos sucesivos se usaron como estructuras de partida aquellas en las que el ligando estaba centrado en el dominio C-terminal. Éstas se obtuvieron repitiendo los cálculos de *docking* molecular usando un *grid* que abarcaba la zona central y dominio CTD, donde, por lo general, esta proteína presenta mayor afinidad.

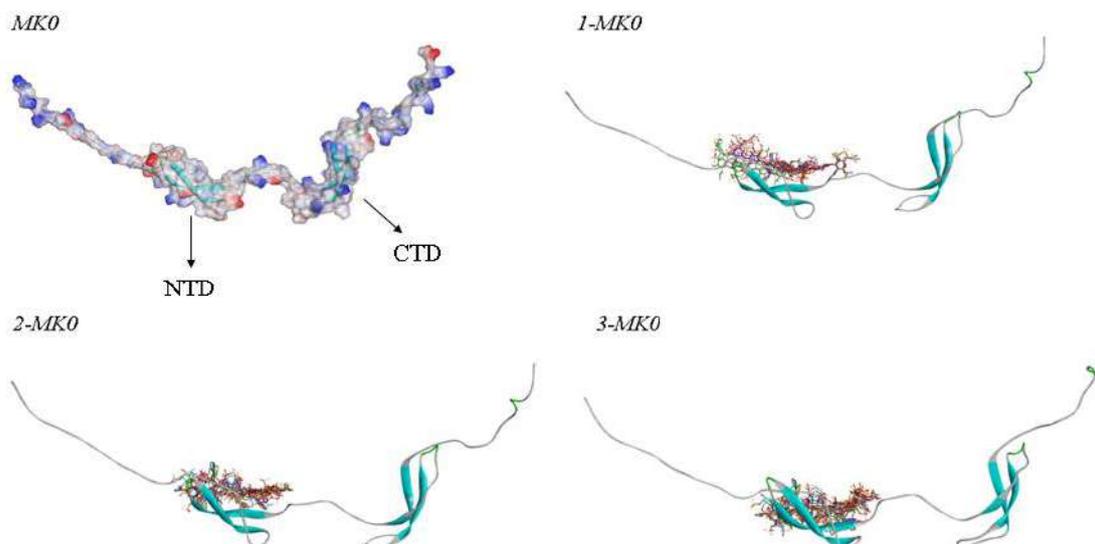
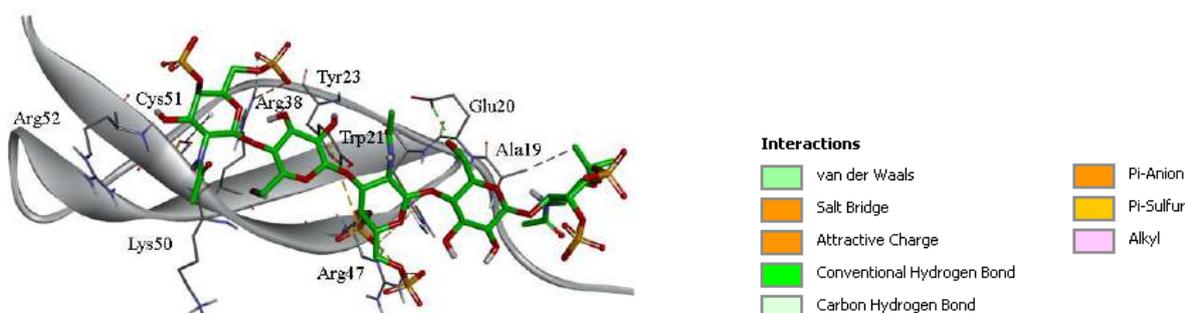


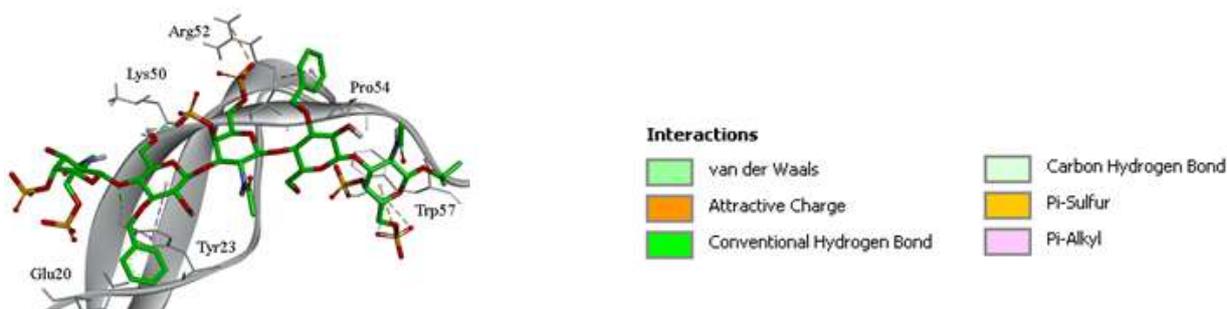
Figura 5.7. Representación de la estructura tridimensional de la MK en su forma lineal y sus complejos con 1, 2 y 3 obtenidos por cálculos de docking molecular (9 estructuras).

Al igual que para el complejo anterior, los aminoácidos básicos son los que presentan las principales interacciones ligando-proteína, con los grupos sulfato del carbohidrato. Además de las interacciones clásicas carga-carga con los residuos básicos de la MK (algunos ejemplos de éstos están descritos en el Anexo II) en este complejo aparecen nuevas interacciones entre sistemas aromáticos de los aminoácidos con los azufres de los grupos sulfatos, estableciendo uniones π -azufre como se puede ver reflejado en la Figura 5.8.

A)



B)



C)

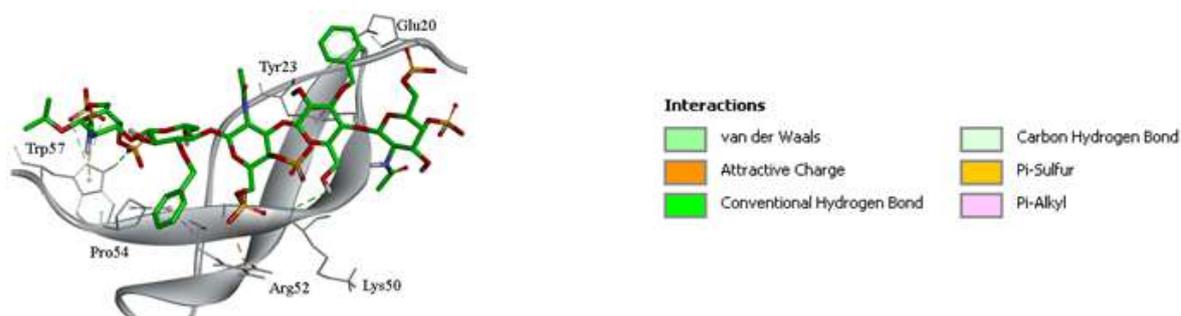


Figura 5.8. Imágenes de las mejores poses de docking obtenidas para los complejos con el ligando desprotegido, **1**, (A), dibencilado, **2**, (B) y persustituído **3**, (C) con la conformación lineal de la MK, mostrando las interacciones intermoleculares y sus tipos. Se muestra la conformación del ligando de mayor energía de interacción.

- Conformación bisagra de la MK (MK2)

La última conformación de la MK presenta diferentes comportamientos en función del mimético estudiado. En la Figura 5.9 se observa que para el complejo con el ligando completamente desprotegido **1** éste se concentra preferentemente sobre el dominio NTD con cierta contribución de la zona *hinge* y más escasa de la región CTD. Algo similar se observa para el compuesto dibencilado **2**, aunque en esta ocasión la distribución se encuentra más centrada entre ambos dominios. En el último caso, el ligando persustituído **3** interacciona mayoritariamente con el dominio NTD y la región de unión. También se aprecian, aunque en menor grado, interacciones con algunos residuos del dominio C-terminal.

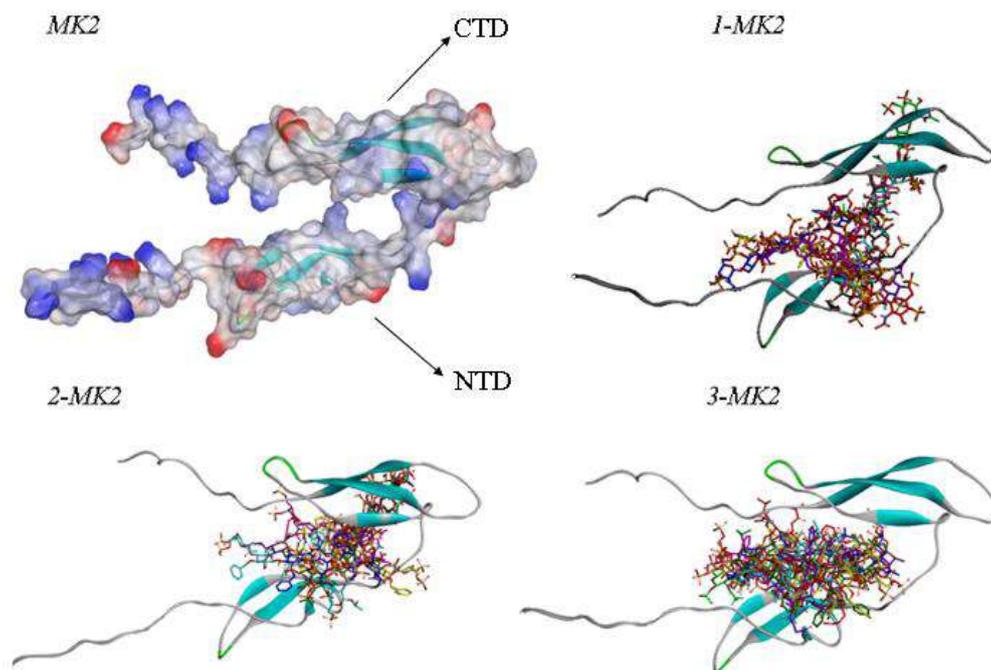


Figura 5.9. Representación de la estructura tridimensional de la MK en su forma “bisagra” y sus complejos con **1**, **2** y **3** obtenidos por cálculos de docking molecular (9 estructuras).

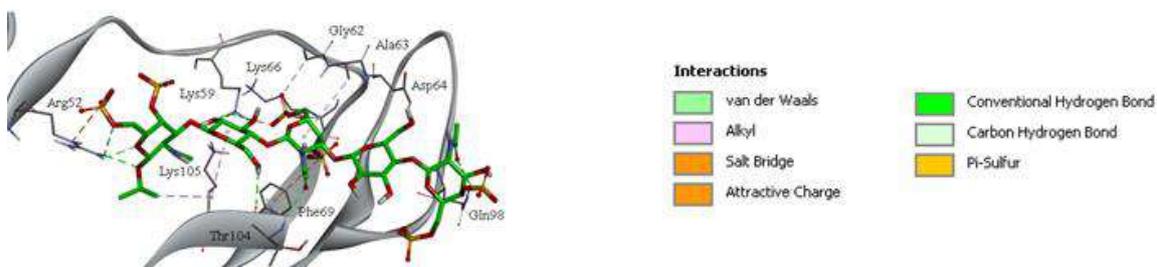
A diferencia de lo que ocurría con la anterior citoquina, PTN, la disposición de los dos dominios enfrentados entre sí para esta conformación bisagra en la MK va a permitir la interacción simultánea del ligando con los dos dominios NTD y CTD.

El ligando desprotegido **1** presenta uniones similares a las observadas para la conformación lineal de la MK. De nuevo se establecen interacciones tipo π -azufre entre los sulfatos del ligando y aminoácidos aromáticos, o entre los oxígenos de los sulfatos y aminoácidos básicos, con carga positiva.

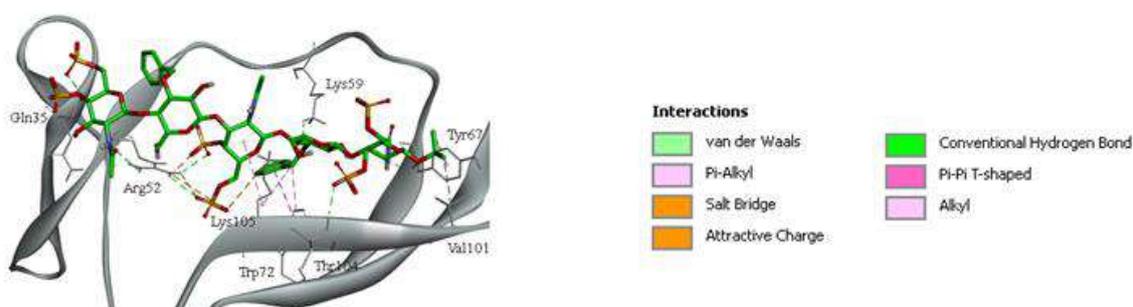
Para el compuesto dibencilado **2**, además de las interacciones alifáticas entre los residuos de azúcar y aminoácidos básicos, aparecen por primera vez para esta proteína las interacciones π - π T-shaped, entre sistemas aromáticos. La aromaticidad de ciertos residuos de la proteína promueve estas interacciones con el bencilo de la glucosa, junto con interacciones π -alquil del C-H bencílico. Aparte de estas uniones, vuelven a surgir los contactos entre residuos alquílicos o interacciones entre los oxígenos carbonílicos y los átomos de flúor del azúcar (Figura 5.10).

En el compuesto **3**, la introducción de un mayor número de residuos aromáticos se refleja en un incremento del número de interacciones entre los residuos aromáticos de la proteína con los introducidos en el carbohidrato. Este hecho explica la diferencia de afinidad medida por FP en el apartado anterior.

A)



B)



C)

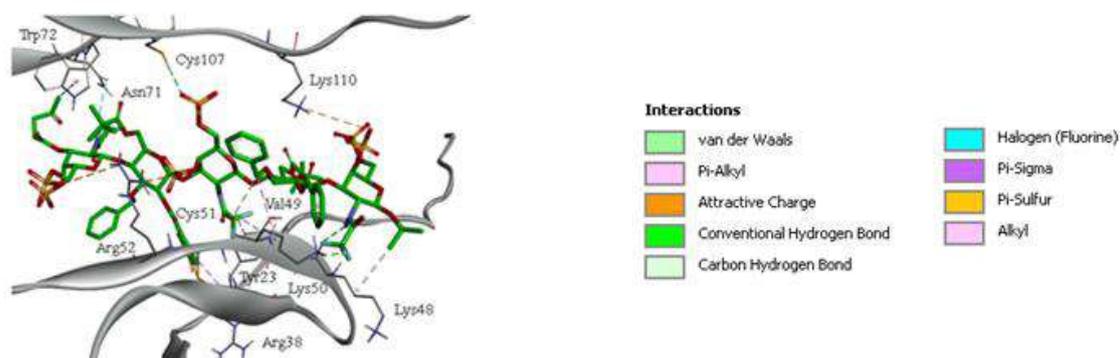


Figura 5.10. Imágenes de las mejores poses de docking obtenidas para los complejos con el ligando desprotegido **1**, (A), dibencilado **2**, (B) y persustituido **3**, (C) con la conformación bisagra de la MK, mostrando las interacciones intermoleculares y sus tipos. Se muestra la conformación del ligando de mayor energía de interacción.

5.4.2 Dinámica Molecular

Siguiendo la misma metodología descrita en el **Capítulo 4**, se realizó un análisis de los complejos formados por la MK y los ligandos de la serie pentasacarídica. Las estructuras de partida de los complejos de las tres conformaciones de MK con **1**, **2** y **3** se seleccionaron siguiendo el mismo procedimiento que el descrito en el capítulo anterior.

En este caso, se restringió la zona de búsqueda a la zona central y al dominio CTD para una mayor precisión. El complejo inicial empleado fue la pose de mayor energía de afinidad de acuerdo con los resultados de Autodock. Tras realizar las dinámicas moleculares en agua explícita y PME, se valoraron los datos generados y los cálculos se repitieron o no, en función de los resultados obtenidos. Para ello, se empleó como criterio de calidad el ajuste entre los resultados experimentales de NOESY transferido y las estructuras obtenidas a lo largo de la dinámica molecular. Posteriormente se realizaron los análisis de MM-PBSA/GBSA.

Los estudios previos demuestran que ambas citoquinas presentan un comportamiento muy similar respecto a su interacción con fragmentos de CS, DS y homólogos. Por ello, se limitó este estudio a la serie pentasacarídica **1**, **2** y **3** de la que se han obtenido más datos experimentales. Con este análisis, se pretende determinar si las diferencias estructurales entre PTN y MK se manifiestan en un diferente comportamiento de los ligandos en los complejos.

5.4.2.1 Complejo pentasacarídico 1-MK

Los resultados de la primera simulación de 200 ns, sin restricciones, del complejo formado por la conformación más plegada de la proteína (**1-MK4**) se analizaron usando *gnuplot* en bloques de 40 ns.^[25] Los resultados mostraron una disposición de tipo *anti-ψ* mayoritaria en el enlace GlcB-GalNAcC, con poblaciones minoritarias que presentan geometrías incompatibles con el efecto exoanomérico ($\phi = -60^\circ$). Por otro lado, las gráficas de Cremer-Pople para los anillos son compatibles con conformaciones de silla 4C_1 (Figura A.55 del Anexo).

En la Figura 5.11 se muestra la repetición de este cálculo incluyendo como restricciones experimentales promediadas en el tiempo las distancias procedentes de los experimentos de NOESY transferido del compuesto. Para este caso, se aumentó la constante de fuerza de 30 a 50 Kcal·mol⁻¹·Å⁻² en el enlace interglicosídico GlcB-GalNAc C, consiguiendo su corrección, mostrando en todos los enlaces una geometría *syn-ψ/exo-φ*, aunque en el caso del enlace GalNAc A-Glc B se encuentran también poblaciones que se extienden hacia una conformación que no cumpliría el efecto exoanomérico (Figura A.56 del Anexo).

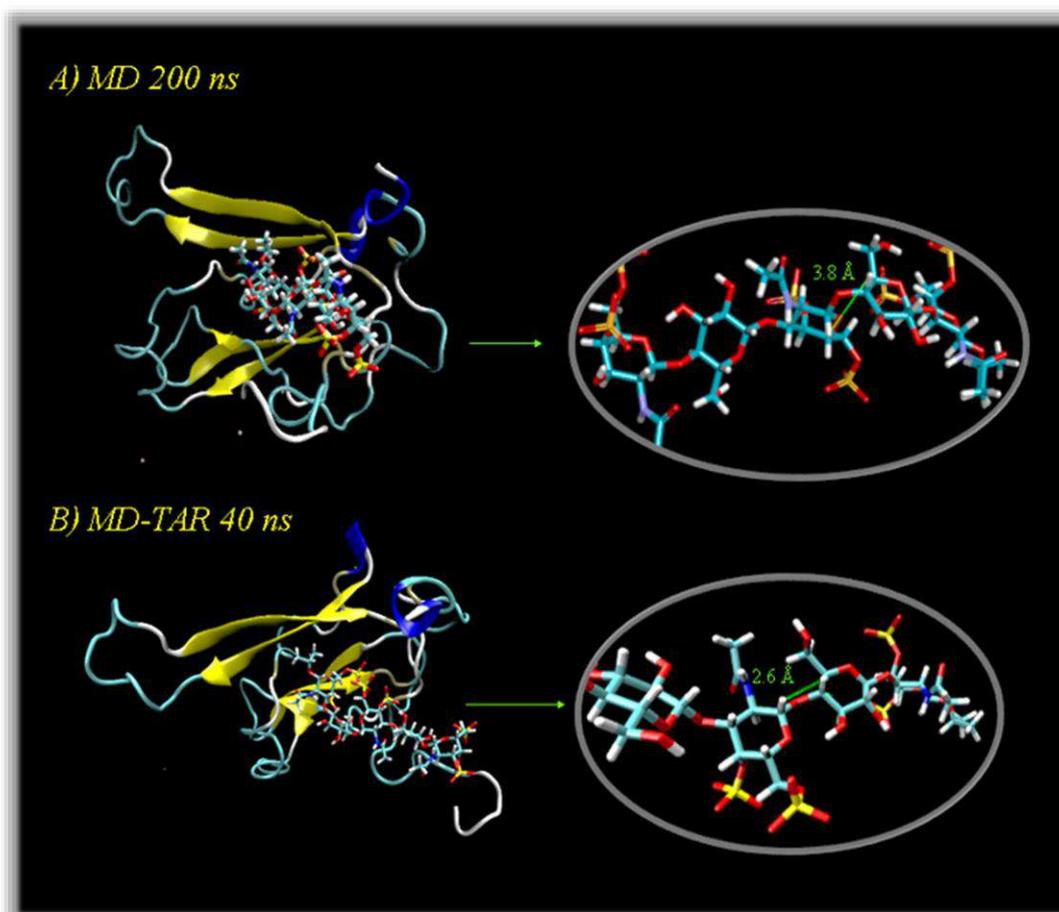


Figura 5.11. Representación tridimensional de las estructuras finales de dinámica molecular para el complejo pentasacarídico **1-MK4** y detalle del esqueleto pentasacarídico, para 200 ns de MD no restringida A) y 40 ns de MD-TAR B).

Para la segunda conformación, su forma lineal (MK0), cuando la dinámica molecular se hizo sin restricciones experimentales, la estructura de la MK experimentó un plegamiento sobre el ligando, flanqueándolo entre ambos dominios, NTD y CTD (Figura 5.12). Este comportamiento ya se vio en las simulaciones con la PTN, donde la proteína conservaba su conformación en las MD-TAR. En cuanto al ligando, todos los anillos presentaban un comportamiento monoconformacional de silla 4C_1 con los enlaces glicosídicos en una orientación *syn-ψ/exo-φ* mayoritaria (ver Figuras A.57 del Anexo).

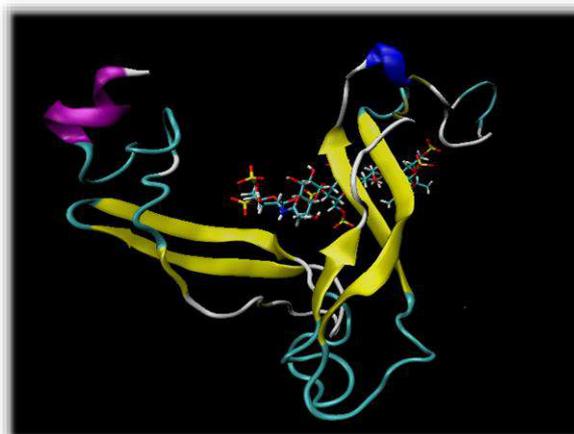


Figura 5.12. Representación tridimensional de la simulación de 200 ns para el complejo pentasacarídico 1-MK0.

Por último, se efectuó el análisis con la conformación bisagra de la MK, MK2. En este caso, fue necesario introducir restricciones experimentales porque a pesar de que ninguno de los anillos de azúcar mostraba distorsión, el enlace GalNAc A-Glc B se encontraba en una disposición *anti-ψ*. El empleo de MD-TAR tampoco corrige el comportamiento a pesar de aumentar hasta $50 \text{ Kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ la constante de fuerza (Figura 5.13). Este incremento no consigue corregir la disposición del enlace glicosídico sino deformar el plegamiento de los anillos (ver Figura A.58 del Anexo).

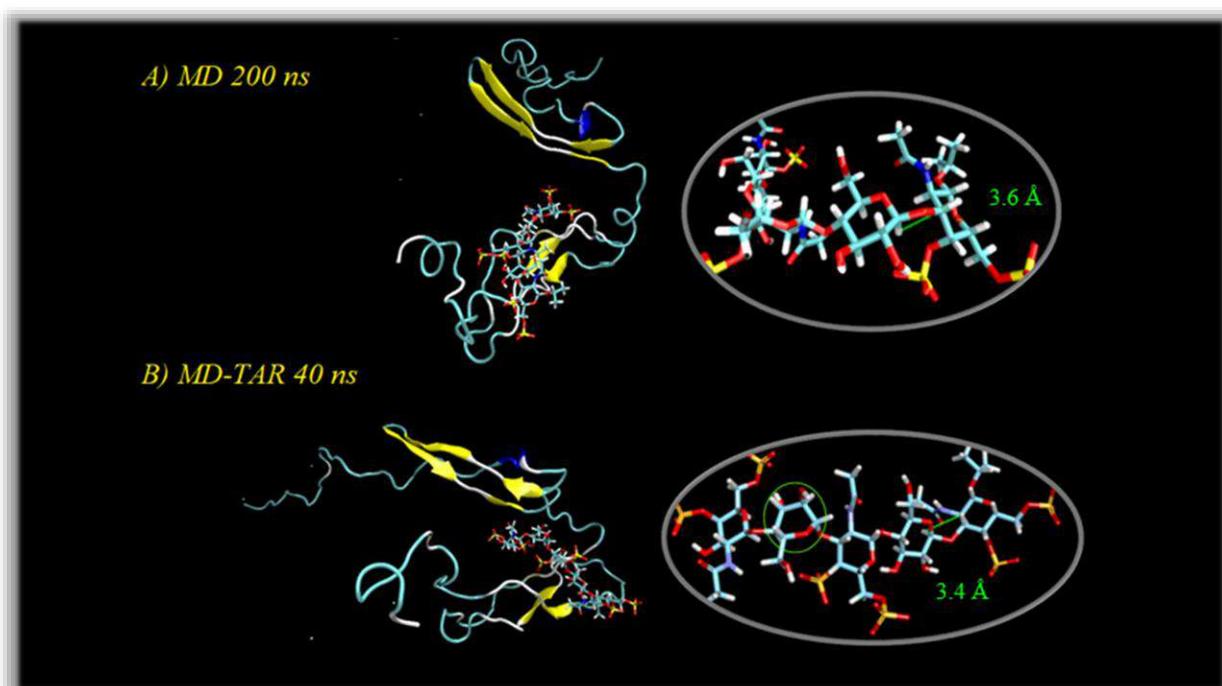


Figura 5.13. Representación tridimensional de las estructuras finales de dinámica molecular para el complejo pentasacarídico 1-MK2 y detalle del esqueleto pentasacarídico, para 200 ns de MD no restringida A) y 40 ns de MD-TAR B).

Basándonos en la experiencia previa, se repitió el cálculo empleando otro conformero localizado en la misma región de interacción, pero con diferente orientación, y una energía de afinidad similar al anterior (Figura 5.14). En este caso, los resultados de MD-TAR generan una estructura compatible con los datos experimentales: con geometrías preferentemente *syn-ψ/exo-φ* en todos los enlaces interglucosídicos y conformaciones 4C_1 en todos los anillos (ver Figura A.59 del Anexo).

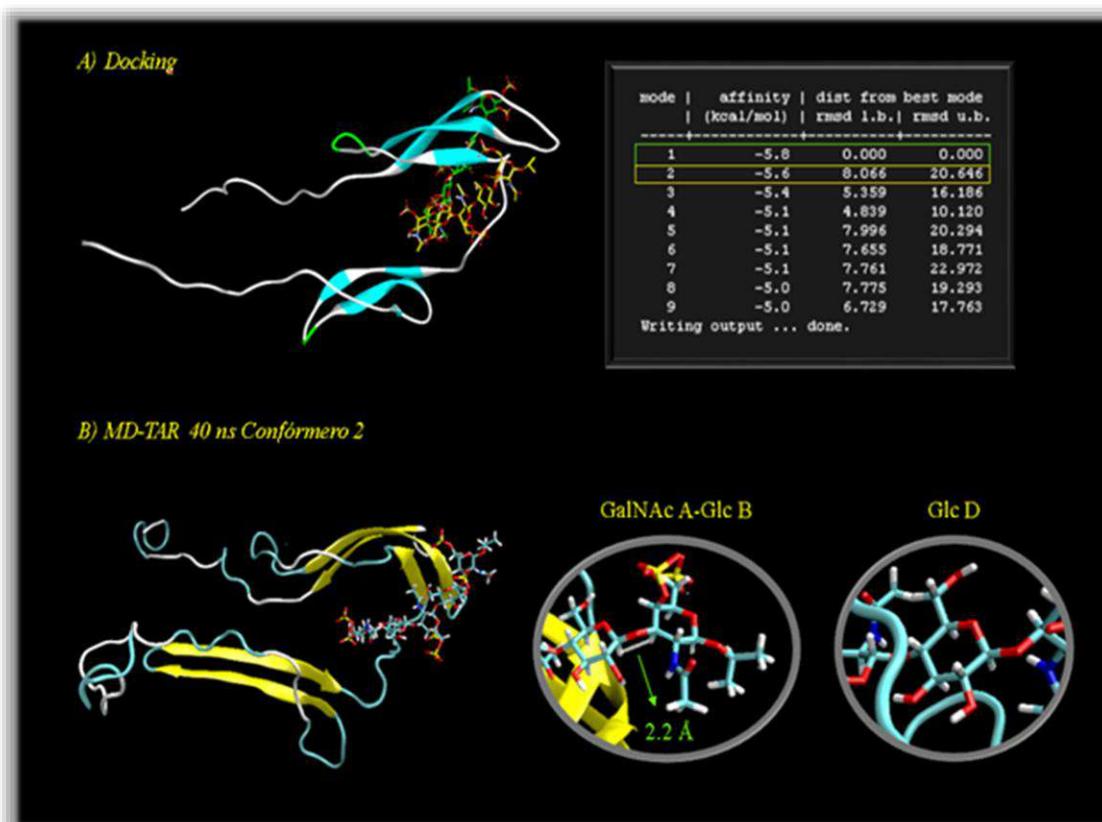


Figura 5.14. A) Representación del docking de los dos ligandos escogidos para las simulaciones llevadas a cabo para el complejo **I-MK2**. En color verde, el primer ligando más estable escogido para la simulación y en amarillo, el segundo más energéticamente favorable. B) Estructura tridimensional procedente de la MD-TAR de 40 ns con el conformero 2, observándose la conformación de silla 4C_1 de la Glc D y la disposición *syn-ψ/exo-φ* del enlace GalNAc A-Glc B.

Estos resultados fueron validados por la comparación de las distancias obtenidas del promedio de los frames individuales a lo largo de la dinámica con las experimentales obtenidas a partir de los NOESY transferidos (Tabla 5.2). Los dos tipos de distancias intra- e interresiduo indican tanto la conformación 4C_1 de los anillos como la disposición *syn-ψ/exo-φ* de los enlaces glicosídicos. Como se ha comentado en el **Capítulo 4**, las constantes de acoplamiento no permiten extraer los ángulos de torsión del estado enlazado al promediar en función de sus poblaciones. Sin embargo, en este caso, la similitud entre los valores del estado enlazado y del estado libre y la concordancia entre todos los valores de constantes del mismo anillo, sugieren que no se está produciendo ninguna variación conformacional relevante. Por ejemplo, la presencia de una conformación del plano ecuatorial de la esfera de Cremer-Pople implicaría una variación asimétrica de las constantes de acoplamiento, y esto no se observa (ver Tabla 5.3).

Distancias	1-MK4	1-MK4	1-MK0	1-MK2	1-MK2	1-MK2
	MD 200ns	MD-TAR 40ns	MD 200ns	MD 200 ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 2
A H1H3	2.3	2.5	2.6	2.5	2.6	2.6
A H1H5	2.5	2.6	2.6	2.7	2.6	2.6
A H3H5	2.4	2.3	2.4	2.4	2.3	2.3
Interg.BH1-AH3	2.1	2.3	2.2	3.6	3.6	2.2
B H1H3	2.6	2.7	2.7	2.6	2.7	2.6
B H1H5	2.5	2.5	2.6	2.4	2.5	2.6
B H3H5	2.8	2.7	2.6	2.6	2.6	2.6
Interg.CH1-BH4	3.5	2.2	2.3	2.2	2.2	2.3
CH1H3	2.5	2.5	2.6	2.5	2.6	2.5
CH1H5	2.6	2.6	2.7	2.5	2.6	2.6
CH3H5	2.5	2.4	2.4	2.4	2.4	2.3
Interg.DH1-CH3	2.5	2.4	2.4	2.4	2.2	2.6
D H1H3	2.6	2.6	2.7	2.7	3.0	2.5
D H1H5	2.5	2.5	2.5	2.5	3.2	2.5
D H3H5	2.6	2.6	2.6	2.6	2.7	2.7
Interg.EH1-DH4	2.4	2.2	2.3	2.2	2.2	2.4
E H1H3	2.7	2.6	2.6	2.6	2.6	2.7
E H1H5	2.6	2.6	2.5	2.5	2.5	2.6
E H3H5	2.4	2.5	2.5	2.5	2.6	2.6

Tabla 5.2. Distancias teóricas de cada pareja de protones intrarresiduales e interglucosídicos para el complejo **1-MK** promediadas de las simulaciones moleculares llevadas a cabo (con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

³ J _{HH} coupling	1-MK4	1-MK4	1-MK0	1-MK2	1-MK2	1-MK2
	MD 200ns	MD-TAR 40ns Conf.1	MD 200ns	MD 200 ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 2
A1A2	8.3	8.3	8.2	8.1	7.8	8.1
A2A3	10.8	10.7	10.6	10.5	10.5	10.5
A3A4	1.8	2.1	2.2	2.0	2.0	2.0
A4A5	1.2	1.1	1.2	0.9	1.0	0.9
B1B2	8.0	7.7	7.3	7.7	7.6	7.7
B2B3	9.4	8.1	8.9	9.5	8.0	9.5
B3B4	9.3	8.6	9	9.0	9.2	9.0
B4B5	9.7	9.5	9.6	9.5	9.7	9.5

Tabla 5.3. ³J_{HH} teóricas de cada pareja de protones calculadas para el complejo **1-MK** promediadas a lo largo de las simulaciones moleculares (con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

$^3J_{HH}$ coupling	1-MK4	1-MK4	1-MK0	1-MK2	1-MK2	1-MK2
	MD 200ns	MD-TAR 40ns Conf.1	MD 200ns	MD 200 ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 2
C1C2	8.4	8.2	8.2	8.2	8.3	8.2
C2C3	10.6	10.6	10.7	10.6	10.7	10.6
C3C4	2.5	2.3	2.3	2.0	2.3	2.0
C4C5	1.1	1.0	1.2	0.8	0.9	0.8
D1D2	7.7	7.5	7.5	8.0	5.5	8.0
D2D3	9.2	9.1	9.6	9.6	8.5	9.6
D3D4	8.8	9.1	9.2	8.5	9.1	8.5
D4D5	9.6	9.7	9.7	9.3	8.1	9.3
E1E2	8.4	8.3	8.3	8.3	8.4	8.3
E2E3	10.2	10.2	10.3	10.2	10.3	10.2
E3E4	3	2.9	2.9	3.0	3.1	3.1
E4E5	1.3	1.3	1.1	1.3	1.4	1.4

Tabla 5.3. Continuación.

5.4.2.2 Complejo pentasacárido 2-MK

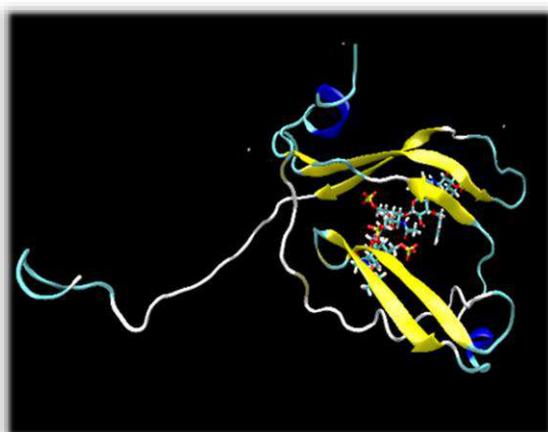


Figura 5.15. Representación tridimensional de la dinámica molecular 200 ns para el complejo 2-MK4.

A continuación, se expone el análisis del complejo del pentasacárido dibencilado **2** con las tres conformaciones de MK, plegada, lineal y bisagra. En la Figura 5.15 se observa la primera de las estructuras, donde la trayectoria del complejo **2-MK4** procedente de una dinámica libre de 200 ns es compatible con los resultados experimentales: se mantiene la estructura canónica 4C_1 de todos los residuos mayoritariamente. Además, los mapas de distribución de población muestran una mayor flexibilidad, especialmente para los enlaces interglucosídicos Glc B-GalNAc C y Glc D-GalNAc E, con algunos submínimos de población de ϕ que se alejan un poco del efecto exoanomérico, pero cumpliendo mayoritariamente una geometría *syn- ψ /exo- ϕ* (ver Figuras A.61 del Anexo).

Por el contrario, el análisis de la dinámica de este ligando con la conformación lineal de la MK sin restricciones, empleando el complejo de energía más favorable, no fue compatible con los resultados experimentales. Hubo que emplear MD-TAR para la corrección tanto de la orientación incompatible del enlace glicosídico como de la distorsión del anillo de la Glc D (ver Figura 5.16 y A.62 del Anexo).

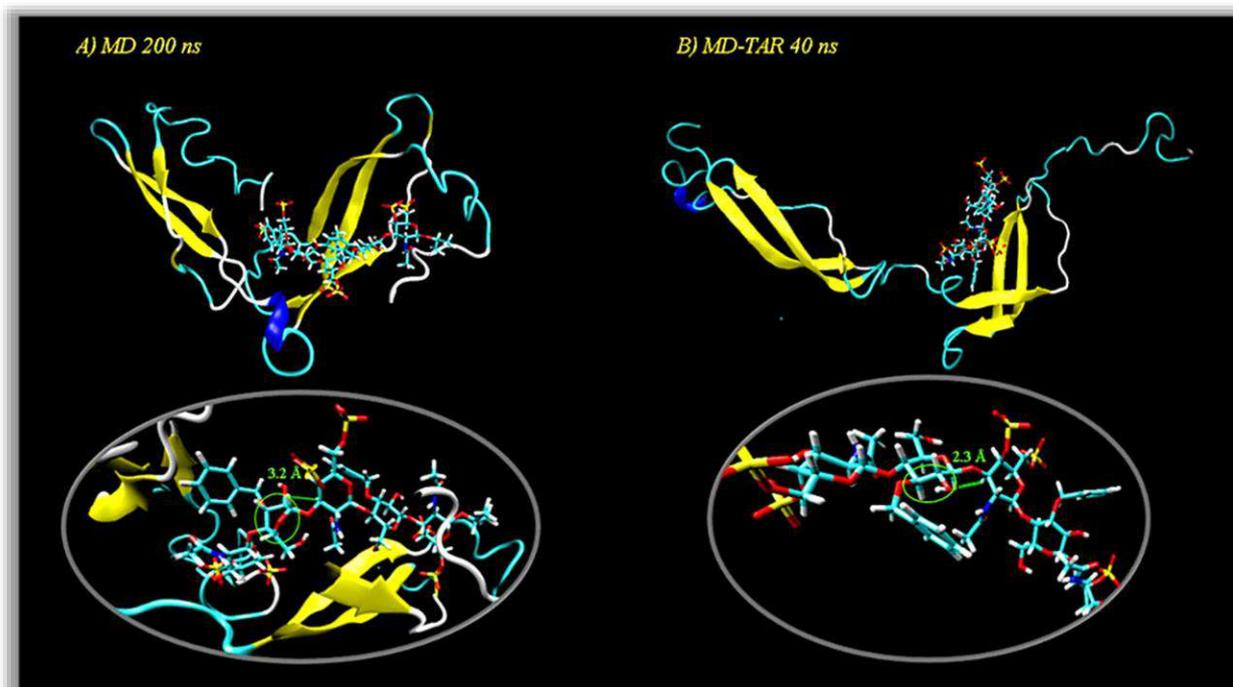
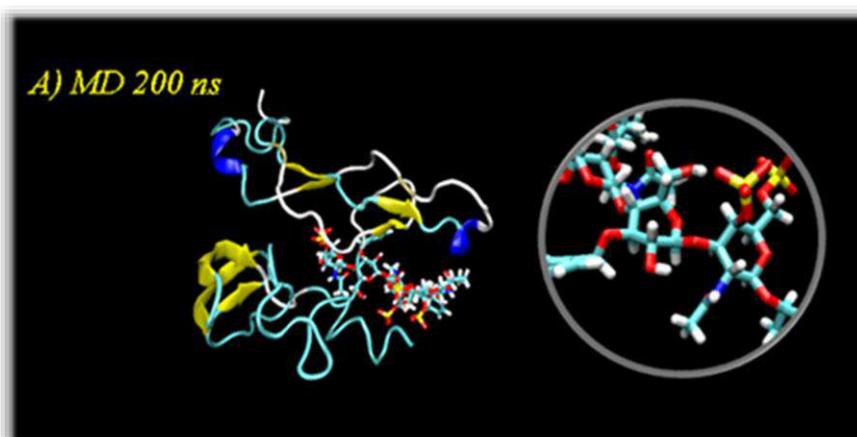


Figura 5.16. Representación tridimensional de las estructuras finales de dinámica molecular para el complejo pentasacarídico 2-MK0 y detalle del esqueleto pentasacarídico, para 200 ns de MD no restringida A) y 40 ns de MD-TAR B).

Para la conformación bisagra se necesitó de nuevo el empleo de otros conformémeros de menor probabilidad en los cálculos de *docking* para obtener resultados compatibles con los datos experimentales. La Figura 5.17 muestra los resultados de la MD y la MD-TAR del conformémero más estable, que fueron incompatibles con los resultados experimentales tanto a nivel de las conformaciones de los azúcares como de los enlaces glicosídicos (ver Figuras A.63 y A 64 del Anexo).



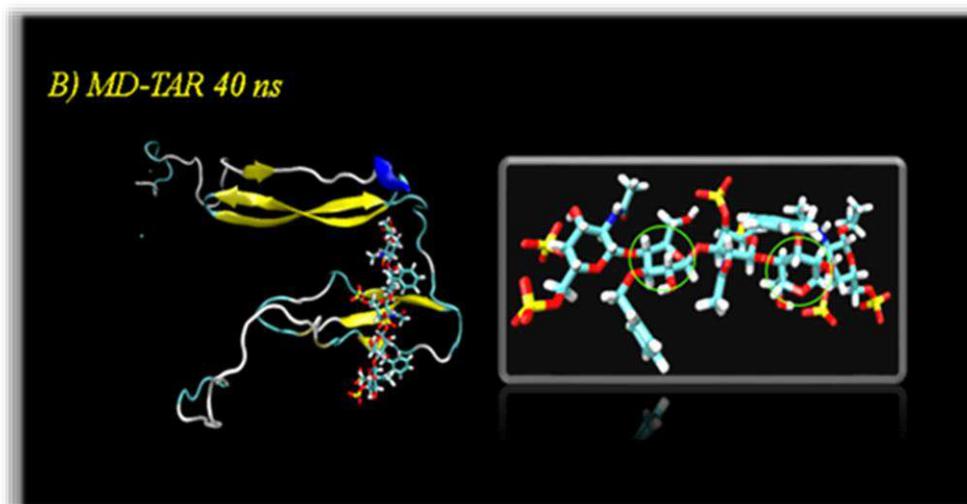


Figura 5.17. Representación tridimensional de las estructuras finales de dinámica molecular para el complejo pentasacarídico 2-MK2 y detalle del esqueleto pentasacarídico, para 200 ns de MD no restringida A) y 40 ns de MD-TAR B).

En este caso particular hubo que usar el tercer conformero según la probabilidad de *docking*, localizado en la zona central de la bisagra, aunque con una orientación distinta respecto al primero, para obtener resultados compatibles con los experimentales (Figura 5.18). En la Figura A.65 del Anexo se muestran las gráficas de *puckering* y los mapas (ϕ , ψ) que predicen una conformación 4C_1 para todos los anillos, con uniones interglucosídicas principalmente *syn- ψ /exo- ϕ* , aunque en el enlace Glc B-GalNAc C siguen apareciendo transiciones esporádicas que incumplen el efecto exoanomérico (valores de ϕ cercanos a -60°), como ocurría en el caso del conformero 1.

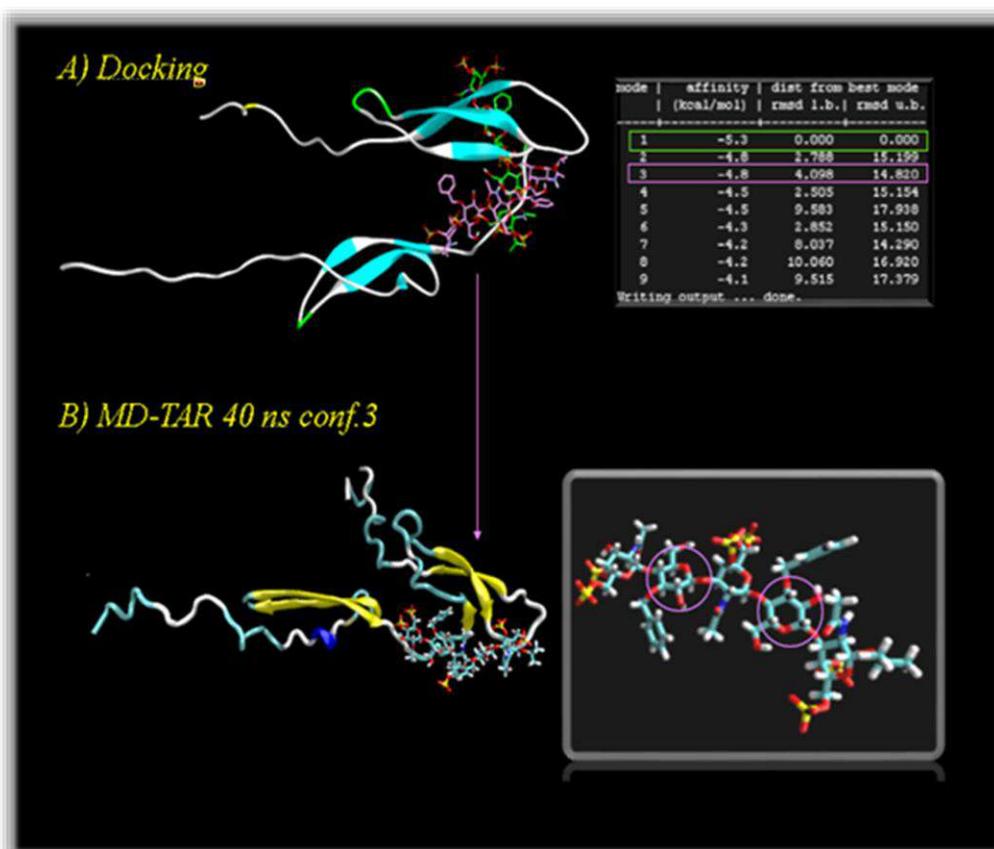


Figura 5.18. A) Representación de los dos conformémeros escogidos para las simulaciones llevadas a cabo en el complejo 2-MK2. En color verde, el más estable de todos y en morado, el tercero más energéticamente favorable. B) Resultado de la simulación MD-TAR de 40 ns para el conformémero 3 dando como resultado la corrección de ambos anillos.

Los resultados de todas las simulaciones para las tres conformaciones con el ligando dibencilado quedan recogidos en las Tablas 5.4 y 5.5.

Distancias	2-MK4	2-MK0	2-MK0	2-MK2	2-MK2	2-MK2
	MD 200ns	MD 200ns	MD-TAR 40ns	MD 200 ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 3
A H1H3	2.6	2.6	2.6	2.6	2.5	2.6
A H1H5	2.8	2.6	2.6	2.7	2.6	2.6
A H3H5	2.3	2.4	2.4	2.5	2.6	2.4
Interg.BH1-AH3	2.3	2.5	2.6	2.2	2.3	2.2

Tabla 5.4. Distancias teóricas de cada pareja de protones intrarresiduales e interglucosídicos para el complejo 2-MK promediadas de las simulaciones moleculares llevadas a cabo (con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

Distancias	2-MK4	2-MK0	2-MK0	2-MK2	2-MK2	2-MK2
	MD 200ns	MD 200ns	MD-TAR 40ns	MD 200 ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 3
B H1H3	2.4	2.7	2.5	3.6	3.2	2.7
B H1H5	2.6	2.4	2.5	3.8	3.3	2.6
B H3H5	2.5	2.8	2.7	3.1	2.7	2.7
Interg.CH1-BH4	2.4	3.9	2.8	2.7	2.7	3.5
C H1H3	2.6	2.7	2.9	2.5	2.6	2.6
C H1H5	2.6	2.6	2.6	2.5	2.5	2.7
C H3H5	2.4	2.5	2.4	2.3	2.4	2.3
Interg.DH1-CH3	2.5	2.9	2.5	2.2	2.3	2.1
D H1H3	2.6	2.9	2.5	2.5	2.9	2.7
D H1H5	2.6	4.1	2.5	2.8	3.8	2.5
D H3H5	2.7	3.4	2.9	2.7	3.3	2.7
Interg.EH1-DH4	2.4	2.6	2.4	2.4	2.3	2.2
E H1H3	2.5	2.4	2.6	2.6	2.7	2.6
E H1H5	2.5	2.5	2.6	2.6	2.6	2.6
E H3H5	2.4	2.6	2.5	2.4	2.5	2.6

Tabla 5.4. Continuación.

$^3J_{HH}$ coupling	2-MK4	2-MK0	2-MK0	2-MK2	2-MK2	2-MK2
	MD 200ns	MD 200ns	MD 40 ns Conf. 1	MD 200ns	MD-TAR 40ns Conf. 1	MD-TAR 40ns Conf. 3
A1A2	8.2	8.0	7.9	8.1	8.2	8.1
A2A3	10.6	10.4	10.3	10.5	10.6	10.6
A3A4	2.3	2.4	2.3	2.3	2.0	1.9
A4A5	1.2	1.0	0.9	1.1	1.0	0.7
B1B2	7.3	7.4	7.4	2.0	4.6	7.2
B2B3	8.1	7.5	9.0	5.7	8.9	8.9
B3B4	9.5	6.4	8.8	7.3	8.6	8.8
B4B5	9.6	7.5	9.5	7.1	5.3	9.4
C1C2	8.4	7.9	7.8	8.3	8.1	8.2
C2C3	10.7	10.2	10.1	10.7	10.2	10.6
C3C4	2.4	2.7	2.4	1.9	2.9	2.3
C4C5	1.1	0.7	0.7	0.9	1.1	0.8

Tabla 5.5. $^3J_{HH}$ teóricas de cada pareja de protones calculadas para el complejo 2-MK promediadas a lo largo de las simulaciones moleculares (con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

$^3J_{HH}$ coupling	2-MK4 MD 200ns	2-MK0 MD 200ns	2-MK0 MD 40 ns Conf. 1	2-MK2 MD 200ns	2-MK2 MD-TAR 40ns Conf. 1	2-MK2 MD-TAR 40ns Conf. 3
D1D2	7.5	7.8	7.5	7.4	4.7	8.0
D2D3	9.4	8.8	8.5	9.8	9.5	9.5
D3D4	8.6	6.2	8.1	8.2	7.1	8.4
D4D5	9.1	9.4	9.5	5.7	2.2	9.0
E1E2	8.4	8.2	8.3	8.1	8.1	8.4
E2E3	10.4	10.3	10.3	10.1	10.2	10.4
E3E4	3.0	2.9	2.8	2.8	2.7	2.8
E4E5	1.4	1.3	1.2	1.3	1.1	0.9

Tabla 5.5. Continuación.

5.4.2.3 Complejo pentasacárido 3-MK

El último estudio de los complejos con MK se realizó con el pentasacárido completamente protegido, **3**. Como se comentó en el punto **5.3.1**, el análisis experimental de este compuesto tuvo más complicaciones que el del resto de la serie. Para el análisis de los cálculos realizados con la estructura de MK se consideraron los resultados experimentales obtenidos para el complejo con PTN (ver **Capítulo 4**).

En primer lugar, se ejecutaron MD libres sobre el modelo plegado, MK-4, con el confórmero más estable obtenido por cálculos de *docking*. A pesar de que las orientaciones en los enlaces interglicosídicos cumplían la geometría esperada *syn-ψ/exo-φ*, el anillo de Glc D adoptaba una conformación de bote torcido tipo 2S_0 (Figura 5.19). Al realizar cálculos MD-TAR restringiendo únicamente la Glc D y aumentando la constante de fuerza K de 30 a 50 Kcal·mol⁻¹·Å⁻² se consigue corregir satisfactoriamente la conformación de dicho anillo (Figura A.67 del Anexo).

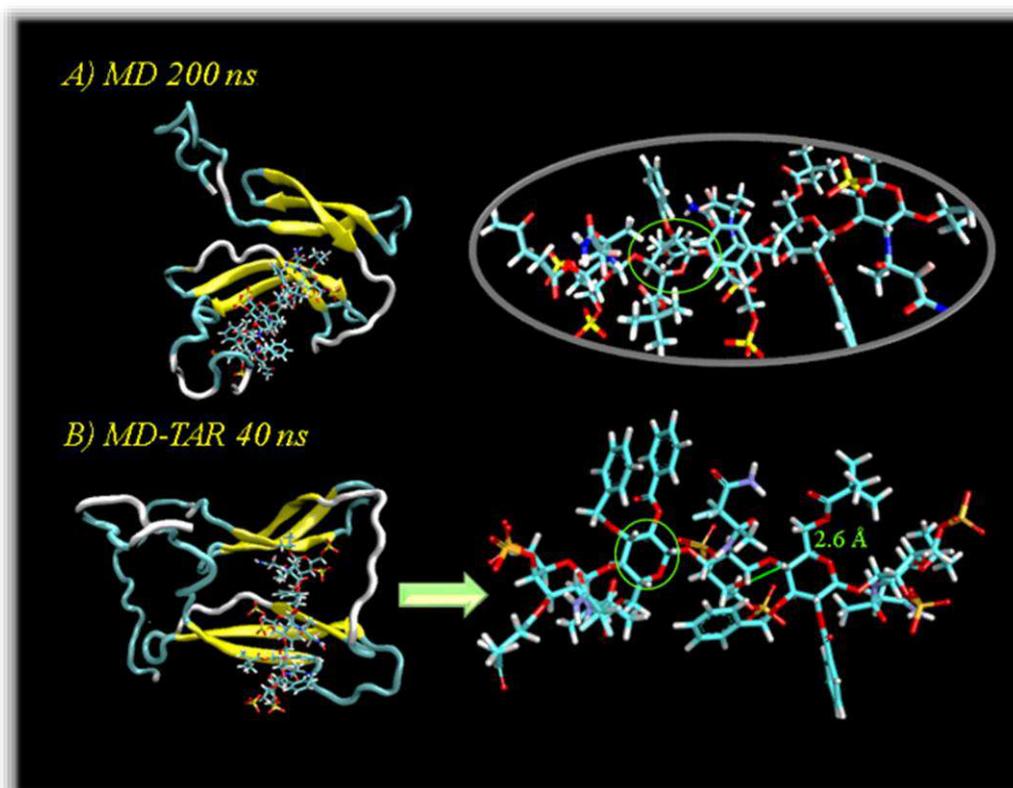


Figura 5.19. Representación tridimensional de las estructuras finales de dinámica molecular para el complejo pentasacáridico 3-MK4 y detalle del esqueleto pentasacáridico, para 200 ns de MD no restringida (A) y 40 ns de MD-TAR (B).

Para el siguiente complejo, con la estructura de la MK lineal, MK0, y con el conformero energéticamente más favorable no se obtuvo un acuerdo razonable entre los resultados teóricos y los experimentales, ni para la dinámica molecular libre ni para la restringida en el tiempo (Figura 5.20, y Figuras A.68 y A.69 del Anexo).

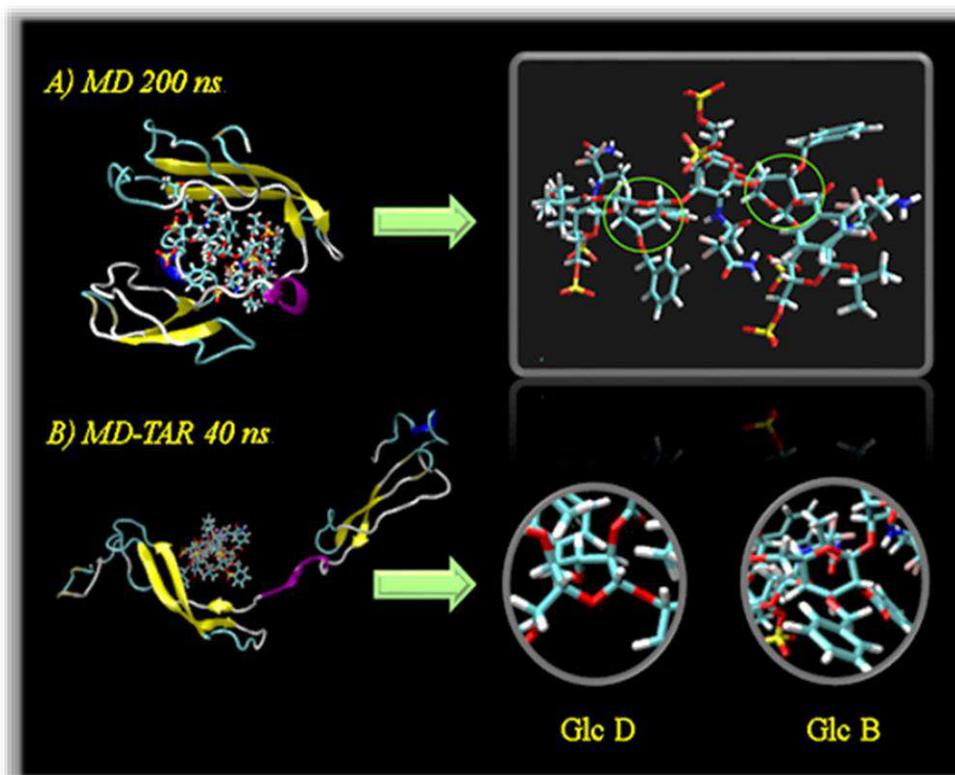


Figura 5.20. Representación tridimensional de las estructuras finales de dinámica molecular para el complejo pentasacarídico 3-MK0 y detalle del esqueleto pentasacarídico, para 200 ns de MD no restringida (A) y 40 ns de MD-TAR (B).

Tras emplear como estructura de partida dos conformémeros alternativos del ligando (conformaciones 2 y 5) de nuevo se obtienen estructuras en desacuerdo con los datos experimentales (Figura 5.21). En concreto, se observan, deformidades en las conformaciones de los anillos B y D respectivamente y, a pesar de aumentar las constantes de fuerza, no se consiguió reproducir una estructura razonable (ver Figuras A.70 y A.71 del Anexo). Con respecto a los enlaces interglucosídicos se observó en ambos casos (conformémeros 2 y 5), cierta flexibilidad en la unión GlcB-GalNAc C, con dos mínimos de población, con los ángulos ψ fluctuando entre -60° y 60° , y los ϕ entre -30° y 60° , pero en general se mantiene una geometría *syn- ψ /exo- ϕ* (Figura A.72 del Anexo). Por lo tanto, en este caso no se pudo reproducir una estructura en total acuerdo con los resultados experimentales.

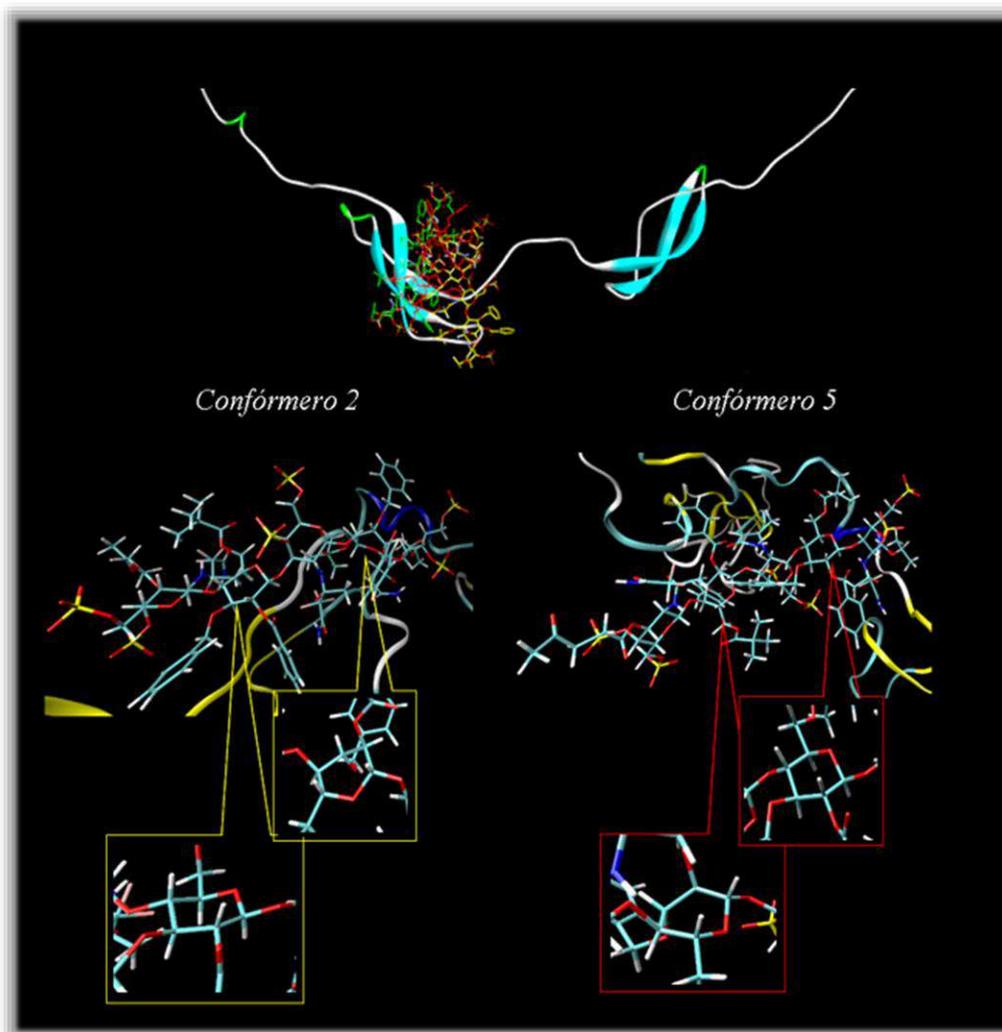


Figura 5.21. Representación tridimensional de los conformeros 1, 2 y 5 del docking (arriba) y las MD-TAR llevadas a cabo para el complejo pentasacáridico 3-MK0. A la izquierda utilizando el conformero 2 (B) observándose distorsión en la Glc B, y a la derecha con el conformero 5 mostrando distorsión en la Glc D.

Finalmente se realizaron los cálculos sobre la conformación de la MK tipo bisagra (MK2). Dados los antecedentes que sugieren una menor fiabilidad de la MD al aumentar el número de sustituyentes del ligando, únicamente se realizaron MD-TAR sobre los complejos de **3**. Tras múltiples e infructuosos intentos sobre las estructuras que los cálculos de *docking* predecían más estables (Figuras A.73 y A.74 del Anexo), finalmente se consiguió una trayectoria compatible con la información estructural disponible, tras varios cambios de conformero (Figura 5.22). Las gráficas de Cremer-Pople mostraron un comportamiento monoconformacional 4C_1 durante toda la trayectoria y una geometría *syn-ψ/exo-φ* en todas las uniones interglucosídicas (Figura A.75 del Anexo).

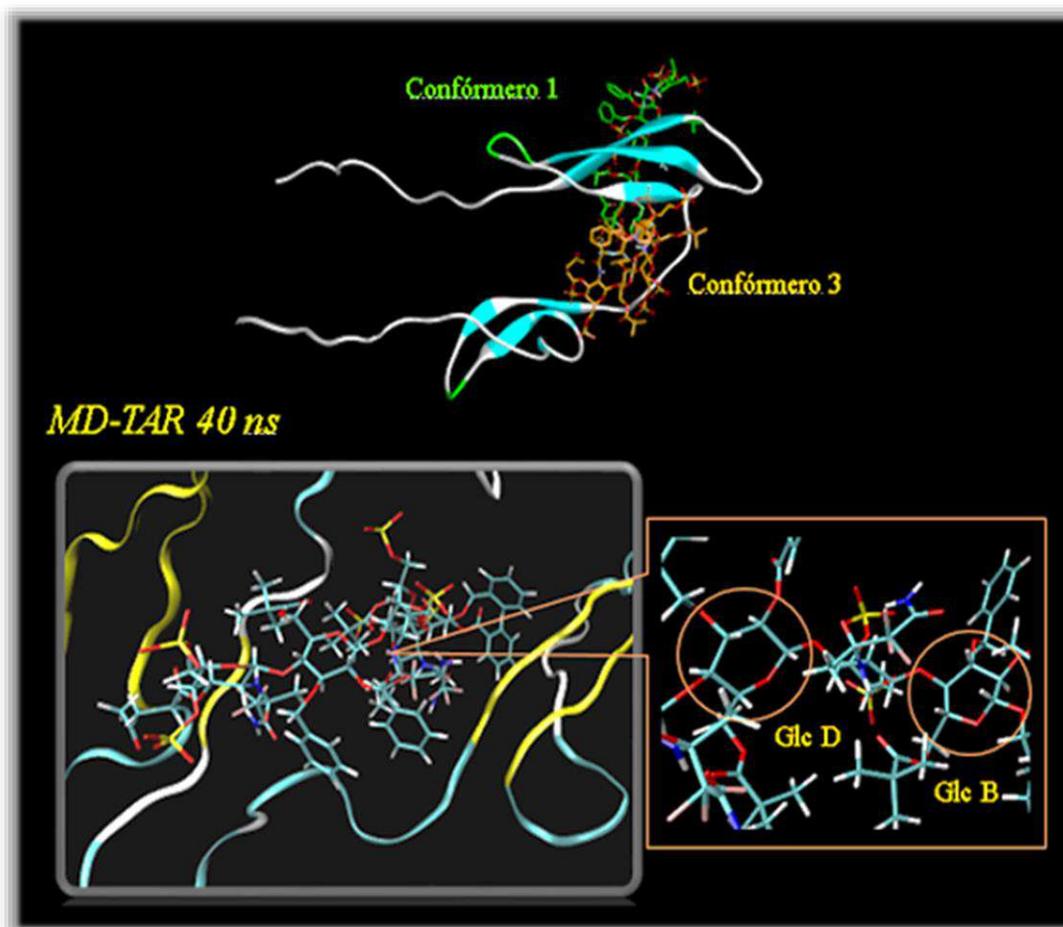


Figura 5.22. Arriba, representación de los dos ligandos escogidos para las simulaciones llevadas a cabo para el complejo **3-MK2**. En color verde, el primer ligando más estable escogido para la simulación y en naranja, el tercero más energéticamente favorable. Abajo, resultado de la simulación MD-TAR de 40 ns para el confórmero 3 dando como resultado la corrección de ambos anillos.

A continuación, se recogen en las Tablas 5.6 y 5.7 las distancias y $^3J_{HH}$ teóricas calculadas a lo largo de todas las trayectorias para el complejo **3-MK**, considerando los diferentes confórmers de ligando utilizados en cada caso con cada una de las conformaciones analizadas de la MK. Las distorsiones de los anillos quedan reflejadas en valores de distancias mayores a 3 Å y $^3J_{HH}$ pequeñas para unidades de Glc.

Distancias	3-MK4	3-MK4	3-MK0	3-MK0	3-MK0	3-MK2	3-MK2
	MD 200ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf.1	MD-TAR 40 ns Conf.2	MD-TAR 40 ns Conf.5	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 3
A H1H3	2.6	2.4	2.6	2.4	2.8	2.7	2.6
A H1H5	2.6	2.6	2.7	2.7	2.7	2.6	2.7
A H3H5	2.5	2.5	2.4	2.3	2.4	2.3	2.4
Interg.BH1-AH3	2.5	2.6	2.8	2.5	2.5	2.4	2.4

Tabla 5.6. Distancias teóricas de cada pareja de protones intrarresiduales e interglicosídicos para el complejo **3-MK** promediadas de las simulaciones moleculares llevadas a cabo (con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

Distancias	3-MK4	3-MK4	3-MK0	3-MK0	3-MK0	3-MK2	3-MK2
	MD 200ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf.1	MD-TAR 40 ns Conf.2	MD-TAR 40 ns Conf.5	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 3
B H1H3	3.0	2.9	3.2	3.9	2.9	2.5	2.4
B H1H5	2.4	2.5	2.5	2.7	2.4	3.3	2.5
B H3H5	2.7	2.7	3.6	4.0	2.8	3.8	2.8
Interg.CH1-BH4	2.1	2.1	2.1	2.2	2.1	2.1	2.1
C H1H3	2.5	2.7	2.4	2.6	2.4	2.3	2.5
C H1H5	2.7	2.6	2.7	2.6	2.7	2.5	2.6
C H3H5	2.4	2.5	2.3	2.4	2.5	2.6	2.4
Interg.DH1-CH3	2.2	2.5	2.1	2.5	2.5	2.6	2.6
D H1H3	4.0	2.5	3.6	2.7	3.9	3.3	2.7
D H1H5	2.4	2.5	2.4	2.7	2.4	2.4	2.4
D H3H5	3.8	2.8	4.0	2.4	3.9	3.1	2.8
Interg.EH1-DH4	2.8	2.5	2.6	2.5	2.6	2.5	2.8
E H1H3	2.6	2.4	2.5	2.6	2.5	2.6	2.4
E H1H5	2.6	2.6	2.6	2.5	2.6	2.6	2.6
E H3H5	2.4	2.5	2.4	2.5	2.4	2.5	2.4

Tabla 5.6. Continuación.

$^3J_{HH}$ coupling	3-MK4	3-MK4	3-MK0	3-MK0	3-MK0	3-MK2	3-MK2
	MD 200ns	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf.1	MD-TAR 40 ns Conf.2	MD-TAR 40 ns Conf.5	MD-TAR 40ns Conf.1	MD-TAR 40ns Conf. 3
A1A2	8.3	8.5	8.1	8.3	7.4	7.3	8.4
A2A3	10.7	10.8	10.7	10.7	10.2	10.5	10.4
A3A4	2.1	1.7	2.0	2.0	2.2	1.8	2.1
A4A5	1.0	0.9	0.9	1.0	0.9	1.0	1.1
B1B2	7.5	7.1	8.0	4.6	7.6	7.9	8.2
B2B3	8.4	7.5	5.8	1.3	8.7	9.0	9.8
B3B4	8.1	8.2	1.3	0.4	8.0	3.4	8.2
B4B5	4.1	6.5	2.8	4.8	6.1	1.5	7.7
C1C2	8.4	8.0	7.9	8.1	8.2	8.4	8.3
C2C3	10.7	10.5	10.6	10.5	10.7	10.7	10.7
C3C4	1.9	2.1	1.6	2.1	1.8	2.5	1.8
C4C5	1.0	1.1	2.3	2.1	2.5	2.6	1.1

Tabla 5.7. $^3J_{HH}$ teóricas de cada pareja de protones calculadas para el complejo 3-MK promediadas a lo largo de las simulaciones moleculares (con y sin restricciones) para las 3 conformaciones escogidas de la proteína.

$^3J_{HH}$ coupling	3-MK4 MD 200ns	3-MK4 MD-TAR 40ns Conf.1	3-MK0 MD-TAR 40ns Conf.1	3-MK0 MD-TAR 40 ns Conf.2	3-MK0 MD-TAR 40 ns Conf.5	3-MK2 MD-TAR 40ns Conf.1	3-MK2 MD-TAR 40ns Conf. 3
D1D2	1.7	8.1	7.8	8.0	4.2	6.3	8.1
D2D3	2.5	9.8	9.5	10.0	1.1	6.2	8.9
D3D4	1.0	8.0	7.9	8.2	1.6	6.1	7.1
D4D5	7.2	9.2	9.3	9.5	4.6	2.7	8.5
E1E2	8.4	8.4	8.5	8.4	8.4	8.4	8.5
E2E3	10.7	10.6	10.7	10.8	10.7	10.5	10.8
E3E4	2.0	2.2	2.4	2.2	2.2	2.5	2.1
E4E5	1.0	1.3	1.4	1.3	1.0	1.0	1.1

Tabla 5.7. Continuación.

Para resumir todos los resultados expuestos en este punto del capítulo, en la Tabla 5.8 se han esquematizado todos los cálculos de dinámica molecular llevados a cabo.

Métodos	MD 200 ns	MD-TAR 40 ns Confórmero 1	MD-TAR 40 ns (Cambio de confórmero)
Complejos Pentasacarídicos			
Conformación MK4			
1-MK4	x	✓	
2-MK4	✓		
3-MK4	x	✓	
Conformación MK0			
1-MK0	✓		
2-MK0	x	✓	
3-MK0	x	x	Confórmero 2 x Confórmero 5 x
Conformación MK2			
1-MK2	x	x	Confórmero 2 ✓
2-MK2	x	x	Confórmero 3 ✓
3-MK2	x	x	Confórmero 3 ✓

Tabla 5.8. Resumen de los resultados obtenidos de las dinámicas moleculares con y sin restricciones para los complejos con los ligandos pentasacarídicos y las tres conformaciones de la MK.

5.4.3 Métodos MM-PBSA/MM-GBSA

Finalmente, se realizaron los cálculos de energía libre del sistema mediante los métodos de MM-GBSA/PBSA sobre la serie pentasacárida **1**, **2** y **3** para conocer en más detalle la contribución energética de los residuos clave en el sitio de unión determinado al realizar las dinámicas moleculares. De nuevo, como en el **Capítulo 4**, se utilizaron las estructuras extraídas de las trayectorias de las simulaciones para obtener valores promedio de energía libre. Para ello se hizo uso del script de Python MMPBSA.py implementado en AMBER16, extrayéndose 1000 estructuras de los últimos 40 ns de las dinámicas libres, y 100 estructuras de los últimos 8 ns de las MD-TAR. Los detalles del cálculo siguen el mismo procedimiento que el que se ha descrito en el apartado **4.5** de metodología del **Capítulo 4**.

Primero se empleó el método de MM-GBSA para el cálculo de la energía libre ligando-MK de los complejos de la serie pentasacárida **1**, **2** y **3** con los tres modelos de MK (Tabla 5.9).

GB Energies (kcal/mol)	VDW	EEL	E _{GB}	E _{Surf}	ΔG_{gas}	ΔG_{solv}	ΔG_{T}
1-MK4	-18.61	-7224.44	7024.36	-12.14	-7243.05	7012.21	-230.83
1-MK4-TAR	-22.48	-3258.22	3249.74	-5.39	-3280.71	3244.34	-36.37
1-MK0	-46.29	-6088.00	5990.38	-13.57	-6134.30	5976.82	-157.48
1-MK2	-1.92	-5889.72	5773.32	-8.86	-5891.65	5764.46	-127.18
1-MK2-TAR	-22.41	-3348.46	3341.85	-5.72	-3370.87	3336.13	-34.73
2-MK4	-30.69	-2966.93	2949.44	-6.13	-2997.63	2896.62	-101.01
2-MK0	-55.73	-6752.61	6562.10	-16.64	-6808.33	6545.45	-262.88
2-MK0-TAR	-25.01	-2687.48	2698.46	-4.69	-2712.48	2693.77	-18.71
2-MK2	-48.97	-7009.41	6857.60	-15.52	-7058.39	6842.07	-216.32
2-MK2-TAR	-55.62	-3755.86	3746.62	-10.36	-3811.49	3736.26	-75.22
3-MK4	-50.25	-7015.37	6930.08	-14.95	-7065.62	6915.12	-150.49
3-MK4-TAR	-62.86	-3921.49	3908.60	-12.49	-3984.36	3896.11	-71.59
3-MK0	-94.17	-7861.87	7735.60	-22.25	-7956.05	7713.36	-242.69
3-MK0-TAR	-39.07	-3302.72	3321.39	-7.57	-3341.79	3313.82	-27.97
3-MK2	-56.42	-7101.54	7001.00	-17.43	-7157.95	6983.57	-174.38
3-MK2-TAR	-82.07	-5252.01	5261.85	-14.79	-5334.08	5247.06	-87.02

Tabla 5.9. Energía de unión (kcal/mol) para la serie **1**, **2** y **3** con el método MM-GBSA a partir de los resultados de dinámica molecular libre y MD-TAR.

Estos cálculos se han realizado sobre los archivos de coordenadas y topologías finales de las dinámicas moleculares libres y MD-TAR cuando se encontraron discrepancias entre la geometría del compuesto en la simulación y la obtenida experimentalmente (ver apartado anterior). A la vista de los resultados de la Tabla 5.9, destaca la discrepancia

lógica entre los valores obtenidos al emplear una u otra metodología, ya que en las MD con restricciones promediadas en el tiempo se genera una energía de penalización que disminuye el valor de ΔG .

En general, se puede establecer que al aumentar el número de grupos aromáticos aumenta la afinidad por la citoquina. Este efecto es más acusado cuando se comparan los valores de afinidad con la conformación lineal (**1-MK0**, **2-MK0** y **3-MK0**). En cambio, cuando la MK se encuentra en su conformación plegada, el ligando desprotegido (**1-MK4**) presenta una interacción considerablemente más fuerte con respecto al resto de la serie, en torno a 100 kcal/mol. Como se ha mencionado antes, tras establecer restricciones en la dinámica, los resultados varían completamente, observándose que la energía libre disminuye en algunos casos hasta 100-150 kcal/mol.

Se realizaron también cálculos de afinidad mediante el método MM-PBSA aplicado sobre las trayectorias obtenidas en el apartado **5.4.2** mostrando cuantitativamente las interacciones de cada residuo de la MK con los átomos presentes en el ligando mediante la descomposición de su energía de interacción. Si bien las estructuras de partida fueron construidas para estudiar las interacciones del dominio CTD y la zona *hinge* central, se han obtenido valores de energía de interacción apreciable para aminoácidos fuera de estas regiones, principalmente del extremo C terminal. De nuevo, para simplificar las tablas generadas, se han tenido en cuenta aquellas interacciones que presentan contribuciones energéticas por debajo de -1 kcal/mol.

En las Tablas A.T.7-T.9 se muestran los resultados de las energías de interacción MM-PBSA de las tres conformaciones de la MK, plegada, lineal y bisagra, con cada compuesto de la serie pentasacarídica, **1**, **2** y **3**. Comparando los resultados con los obtenidos para la PTN en el **Capítulo 4** se aprecia una mayor población de aminoácidos que muestran interacción en el caso de MK. Esto estaría de acuerdo con los valores de FP donde se aprecia una mayor capacidad de interacción de MK frente a PTN que se podría reflejar en el número de aminoácidos involucrados en la interacción. En general, se observa que existen fuertes interacciones con aminoácidos del primer dominio NTD para los compuestos **1**, **2** y **3**, en las tres conformaciones de la MK, MK4, MK0 y MK2, sin disminuir las interacciones con el resto de la proteína.

- Conformación plegada MK4

Los aminoácidos que residen en el dominio C-terminal presentan interacciones con los tres ligandos de la serie, siendo común en los tres casos la Lys66 y la Arg84. En los complejos con los compuestos **1** y **3** las interacciones más fuertes son principalmente de tipo carga-carga, entre aminoácidos básicos (Lys121, Lys123, Lys110), y los grupos sulfato del carbohidrato. Por otro lado, la presencia de los anillos aromáticos en **2** y **3** genera nuevos tipos de interacciones más débiles con valores de energía cercanos a -1 kcal/mol, pudiendo ser de tipo π - π como con Tyr67 y Phe69, o de van der Waals con aminoácidos polares como Thr41 y Gly40 entre otros. En la Figura 5.23 se han reproducido los datos de la Tabla A.T7 del Anexo para la serie pentasacarídica con la conformación de la MK más plegada, MK4, señalando los aminoácidos con mayor

afinidad por los residuos del azúcar.

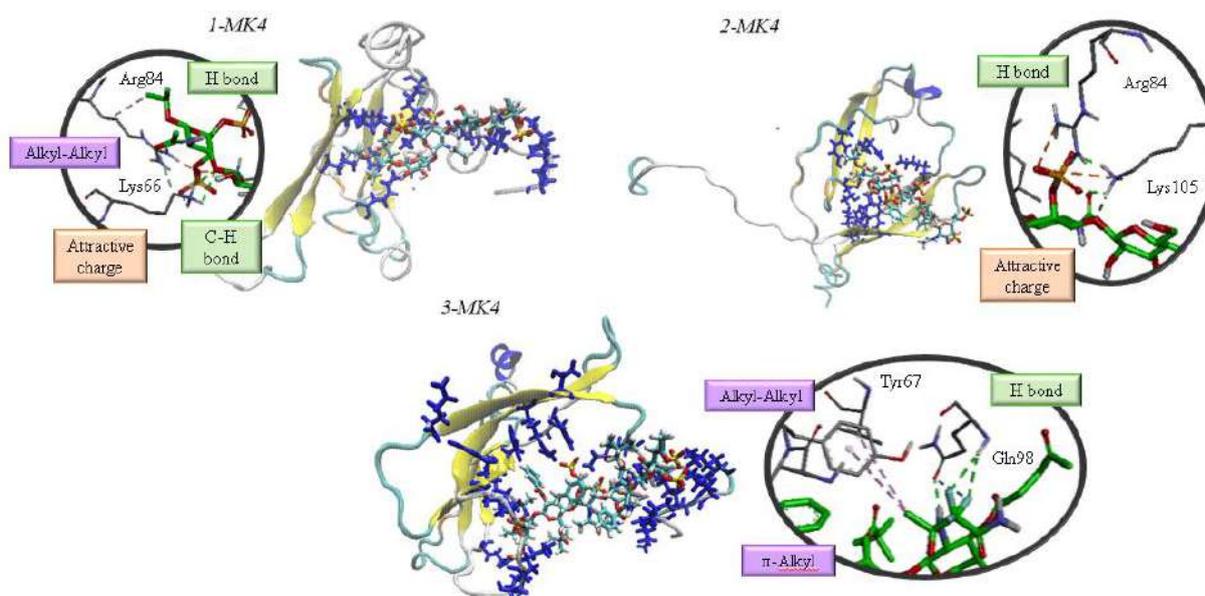


Figura 5.23. Representación tridimensional de las principales interacciones entre los ligandos pentasacarídicos y los aminoácidos que presentan mayor afinidad de la MK en su conformación plegada, MK4, mediante cálculos MM-PBSA. Ampliación de ciertas interacciones con algunos de los aminoácidos de mayor afinidad.

- Conformación lineal MK0

La MK en su forma lineal, MK0, resulta no ser estable y tiende a plegarse durante la dinámica adoptando estructuras similares a la anterior (MK4), aunque al analizar las trayectorias de los compuestos **1**, **2** y **3** para esta conformación se aprecian diferencias entre ellas (Figura 5.24). Se observó una considerable participación de los aminoácidos del dominio NTD, del dominio C-terminal, así como de los residuos situados en ambos extremos de la MK (especialmente en el caso del complejo **2-MK0**). El hecho de que tengan lugar interacciones con los extremos terminales de la MK se debe a que en las dinámicas sin restricciones observábamos cierto plegamiento de estas zonas que se aproximaban hasta establecer interacción con el azúcar. Del mismo modo que antes, aparecen interacciones de sistemas π - π débiles entre los grupos aromáticos del ligando y los residuos de Tyr67. También aparecen interacciones π -catión ligeramente más fuertes con aminoácidos como Lys 105 (complejo **3-MK0**). Al igual que antes, en la Figura 5.24 se han representado tridimensionalmente las trayectorias de los tres complejos para la conformación más abierta de la proteína, remarcando los aminoácidos que más interacción mostraban como se ha detallado en la Tabla A.T8 del Anexo.

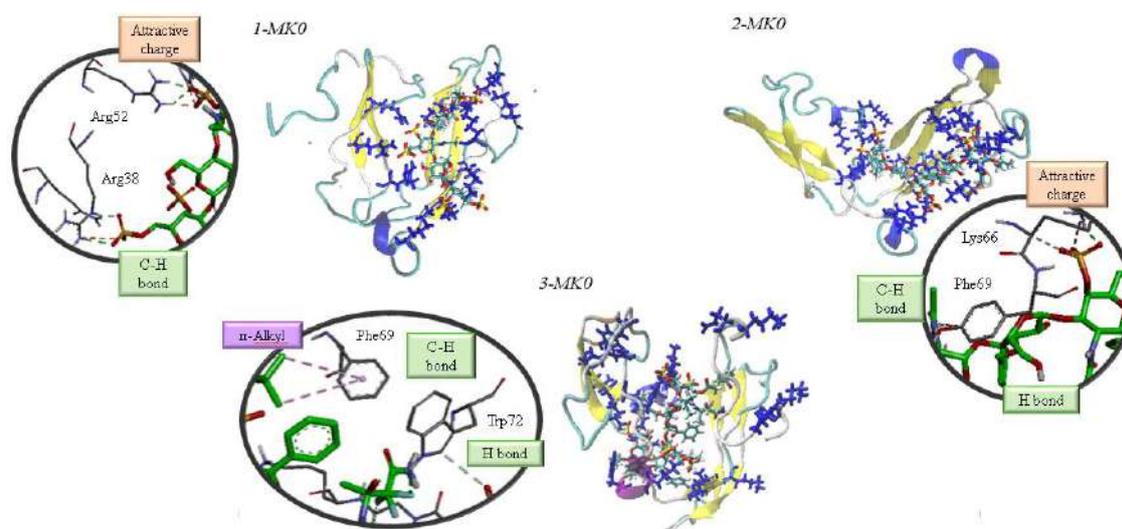
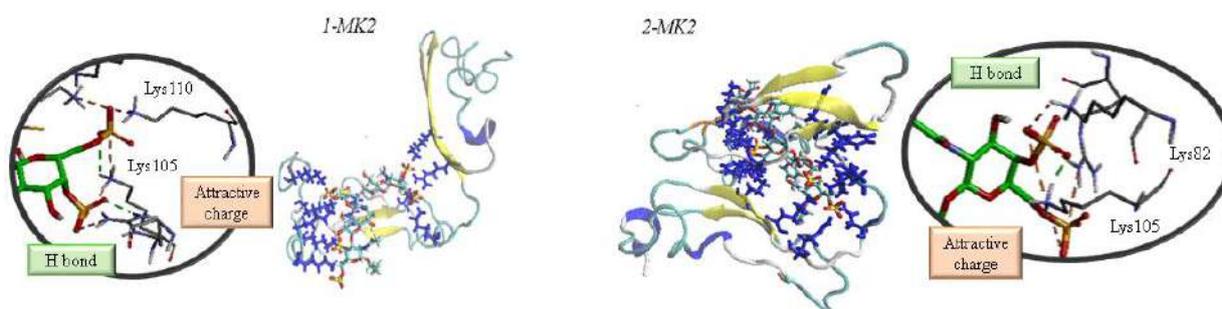


Figura 5.24. Representación tridimensional de las principales interacciones entre los ligandos pentasacáridos y los aminoácidos que presentan mayor afinidad de la MK en su conformación abierta, MK0, mediante cálculos MM-PBSA. Ampliación de ciertas interacciones con algunos de los aminoácidos de mayor afinidad.

- Conformación bisagra MK2.

En la figura 5.25 se muestran los resultados de los cálculos MM-PBSA sobre el tercer modelo de MK, MK2, indicando los resultados de la Tabla A.T9 del Anexo. Los aminoácidos básicos Arg52 y Lys50 del dominio NTD participan en interacciones carga-carga en todos los complejos con los ligandos **1**, **2** y **3**. Las interacciones en los tres sitios principales de unión son proporcionales y en este caso, al tener una alta población de lisinas, predominan las interacciones carga – carga con el ligando cargado negativamente, tanto en el área más central (Lys66 y Lys68), como en el extremo C-terminal (Lys105-Lys123), véase el complejo **3-MK2**. Nuevamente, la Tyr67 establece interacciones débiles con los grupos benzoato de los anillos de Glucosa.



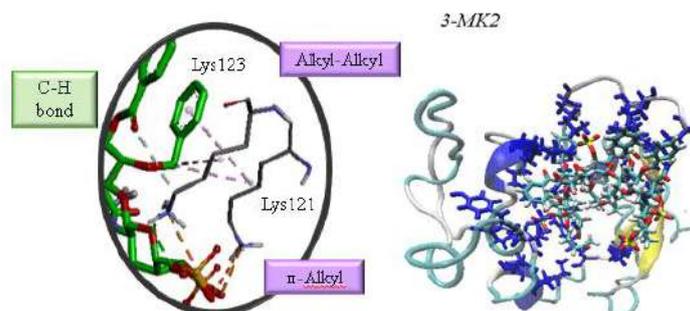


Figura 5.25. Representación tridimensional de las principales interacciones entre los ligandos pentasacarídicos y los aminoácidos que presentan mayor afinidad de la MK en su conformación bisagra, MK2, mediante cálculos MM-PBSA. Ampliación de ciertas interacciones con algunos de los aminoácidos de mayor afinidad

En general, se observan comportamientos comunes entre los tres complejos para las tres conformaciones proteicas. Como es de esperar, debido a la presencia de los grupos sulfatos, se establecen más interacciones electrostáticas entre aminoácidos básicos y los residuos de GalNAc, en comparación con los de Glc. En segundo lugar, de las tres áreas de unión existentes, las mayores interacciones tienen lugar en ambos dominios estructurados, NTD y CTD, aunque predomina este último. También se observan interacciones con los extremos C y N-terminal, especialmente con el primero de ellos. Con respecto a la región *hinge* o bisagra los residuos Asn56, Lys59, Lys58, Glu60 y Phe61 presentan interacción. Cabe añadir que, a pesar de haberse encontrado altas afinidades en diferentes estudios de GAGs con la MK para la secuencia Leu-Lys-Lys-Ala-Arg-Tyr (residuos 85-90), estos residuos no han presentado unos valores de energía remarcables en nuestro estudio. Por último, las interacciones carga-carga de los grupos sulfatos con residuos positivamente cargados son las que aportan mayores energías de interacción. En la Figura 5.26 se han resumido los datos de las Tablas A.T7, A.T8 y A.T9 del Anexo para los cálculos de MMPBSA mediante un diagrama de barras.

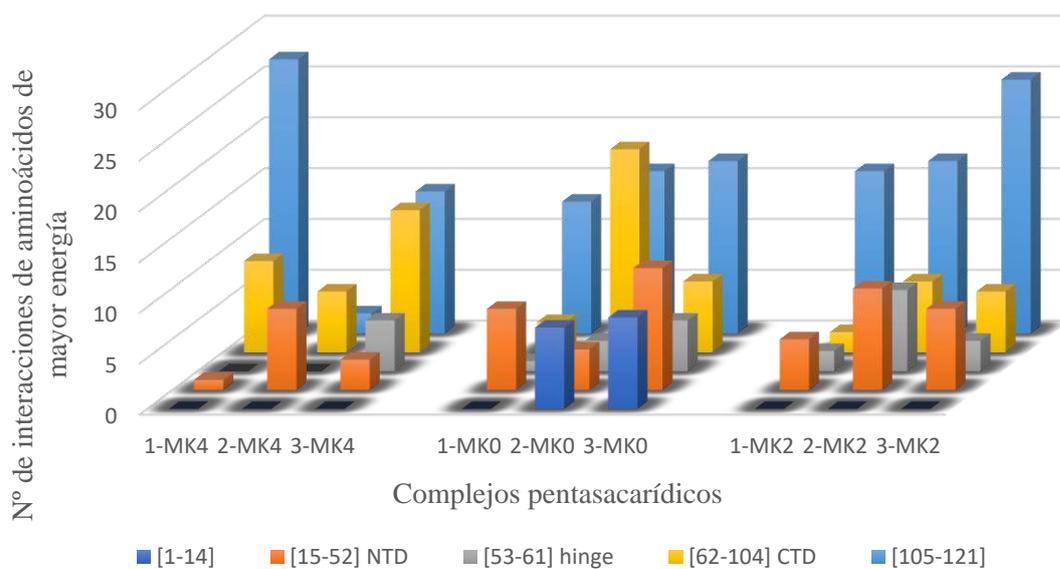


Figura 5.26. Representación gráfica de los resultados de las tablas de MM-PBSA comparando los complejos de la serie pentasacáridica para cada conformero de la MK, indicando las zonas de ésta que presentan mayor interacción con el carbohidrato.

5.5 Metodología

La metodología y los protocolos empleados en este capítulo, tanto para los cálculos teóricos como para las técnicas experimentales, han sido los mismos que los descritos en el punto 4.5 del **Capítulo 4**.

Bibliografía Capítulo 5

- [1] C. Winkler, S. Yao, *Br. J. Pharmacol.* **2014**, *171*, 905–912.
- [2] P. S. Filippou, G. S. Karagiannis, A. Constantinidou, *Oncogene*, **2020**, *39*, 2040–2054.
- [3] K. Maruyama, H. Muramatsu, N. Ishiguro, T. Muramatsu, *Arthritis Rheum.* **2004**, *50*, 1420–1429.
- [4] R. Choudhuri, H. T. Zhang, S. Donnini, M. Ziche, R. Bicknell, *Cancer Res.* **1997**, *57*, 1814–1819.
- [5] T. Muramatsu, *Br. J. Pharmacol.* **2014**, *171*, 814–826.
- [6] K. Sugahara, T. Mikami, *Curr. Opin. Struct. Biol.* **2007**, *17*, 536–545.
- [7] C. Ueoka, N. Kaneda, I. Okazaki, S. Nadanaka, T. Muramatsu, K. Sugahara, *J. Biol. Chem.* **2000**, *275*, 37407–37413.
- [8] P. Zou, K. Zou, H. Muramatsu, K. Ichihara-Tanaka, O. Habuchi, S. Ohtake, S. Ikematsu, S. Sakuma, T. Muramatsu, *Glycobiology* **2003**, *13*, 35–42.
- [9] A. M. J. Garcia-jimenez, S. Gil-caballero, F. Corzana, F. Juarez-vicente, J. R. Miles, K. Sakamoto, K. Kadomatsu, M. Dominguez-garcía, J. L. De Paz, P. M. Nieto, **2021**, *27*, 12395-12409.
- [10] E. Gramage, T. D’Cruz, S. Taylor, R. Thummel, P. F. Hitchcock, *PLoS One* **2015**, *10*, 1–17.
- [11] D. Liedtke, C. Winkler, *Dev. Dyn.* **2008**, *237*, 62–74.
- [12] A. A. Calinescu, T. S. Vihtelic, D. R. Hyde, P. F. Hitchcock, *J. Comp. Neurol.* **2009**, *514*, 1–10.
- [13] W. Iwasaki, *EMBO J.* **1997**, *16*, 6936–6946.
- [14] A. R. Griffith, C. J. Rogers, G. M. Miller, R. Abrol, L. C. Hsieh-Wilson, W. A. Goddard, *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 13697–13702.
- [15] J. D. Esko, J. H. Prestegard, R. J. Linhardt, in *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, **2017**.
- [16] C. Solera, G. Macchione, S. Maza, M. M. Kayser, F. Corzana, J. L. De Paz, P. M. Nieto, *Chem. A Eur. J.* **2016**, *22*, 2356–2369.
- [17] C. Bu, L. Jin, *Front. Mol. Biosci.* **2021**, *8*, 1-20.
- [18] J. Lim, S. Yao, M. Graf, C. Winkler, D. Yang, *Biochem. J.* **2013**, *451*, 407–415.
- [19] E. E. O. Caldwell, V. D. Nadkarni, J. R. Fromm, R. J. Linhardt, J. M. Weiler, *Int. J. Biochem. Cell Biol.* **1996**, *28*, 203–216.
- [20] M. Torres Rico, S. Maza, J. L. de Paz, P. M. Nieto, *Org. Biomol. Chem.* **2021**, *19*, 5312-5326.

- [21] S. Maza, M. Mar Kayser, G. MacChione, J. López-Prados, J. Angulo, J. L. De Paz, P. M. Nieto, *Org. Biomol. Chem.* **2013**, *11*, 3510–3525.
- [22] S. Maza, N. Gandia-Aguado, J. L. de Paz, P. M. Nieto, *Bioorganic Med. Chem.* **2018**, *26*, 1076–1085.
- [23] J. L. De Paz, P. M. Nieto, *Org. Biomol. Chem.* **2016**, *14*, 3506–3509.
- [24] J.-S. Gómez-Jeria, A. Robles-Navarro, G. Kpotin, N. Garrido-Sáez, G.-D. Nelson, *Chem. Res. J.* **2020**, *5*, 32–52.
- [25] J. Racine, *J. Appl. Econom.* **2006**, *21*, 133–141.

Capítulo 6

Conclusiones

6.1 Conclusiones

En esta Tesis Doctoral se ha abordado el estudio estructural y el análisis de las interacciones con MK y PTN de compuestos miméticos de CS de tipo E carentes de grupo carboxilato incluyendo o no grupos aromáticos en su estructura. La estructura de los ligandos se ha resuelto mediante RMN y su interacción con las citoquinas se ha cuantificado por Polarización de Fluorescencia. Estos datos se han completado con estudios teóricos de *docking*, dinámica molecular y cálculos MM-PBSA/GBSA. Este estudio podría dividirse en dos partes diferenciadas:

- *Estudio de los ligandos en su forma libre*

Se ha determinado la estructura tridimensional de los ligandos libres **1-6** mediante experimentos de RMN, completados con simulaciones de dinámica molecular. Los resultados experimentales indican que tanto la introducción de anillos aromáticos como la eliminación del grupo carboxilato no tienen efecto aparente alguno en la estructura tridimensional de los compuestos estudiados. Los análisis tanto de los NOE intraresiduales como de $^3J_{HH}$ confirman la disposición 4C_1 para ambos tipos de anillos, Glc y GalNAc. Del mismo modo, los NOE interglucosídicos reflejan una disposición *syn-ψ/exo-φ* sin evidencias de participación de hipotéticas disposiciones *anti-ψ*.

Las simulaciones de dinámica molecular de los miméticos sin grupos aromáticos adicionales corroboraron la irrelevancia de la eliminación del grupo carboxilato sobre la estructura tridimensional de éstos. No obstante, los ligandos que presentaban grupos voluminosos en su esqueleto tienden a deformarse a lo largo de la dinámica, causando distorsiones incompatibles con los resultados experimentales. Estas se pueden evitar usando restricciones experimentales promediadas en el tiempo en la simulación, MD-TAR, generando estructuras compatibles con los resultados de RMN.

- *Estudio de los complejos*

➤ *Complejos con Pleiotrofina*

En primer lugar, se obtuvieron datos experimentales de las estructuras de los ligandos complejados mediante experimentos de RMN de transferencia de magnetización y observación del ligando (NOESY transferido y *Saturation Transfer Difference*, STD). En los compuestos **1**, **2**, **4** y **5** se obtuvieron distancias experimentales que reflejaban una estructura tridimensional que no cambiaba respecto a su forma libre. Estas distancias no pudieron obtenerse para **3** y **6**, aunque en el caso del complejo **3-PTN**, el patrón de contactos dipolares era compatible con el de **1**, **2**, **4** y **5**. Es decir, ni la conformación de los anillos ni las orientaciones relativas de las uniones interglucosídicas cambian respecto a la de los miméticos libres al formarse los complejos.

Los experimentos de STD no son compatibles con una única estructura enlazada sino con múltiples conformeros uniéndose a la proteína que reciben la magnetización desde todas

las direcciones del espacio. Mediante Polarización de Fluorescencia se pudo determinar que la presencia de sistemas aromáticos incrementaba la afinidad por la citoquina.

También se emplearon métodos de observación del receptor registrando una valoración usando ^{15}N -HSQC. De la variación de desplazamientos químicos ponderada de los sistemas NH, se pudo deducir que el sitio activo de la PTN se encontraba preferentemente en el dominio CTD y parte de la zona *hinge* central, concordando con los resultados previamente descritos para compuestos con carboxilato.

Los estudios teóricos comenzaron con la preparación de los complejos con los seis miméticos empleando Autodock sobre tres estructuras representativas de la PTN obtenidas de la base de datos pdb. El análisis de nuestros resultados indica la versatilidad de los ligandos, que se asocian a la PTN en multitud de orientaciones, lo que concuerda con las observaciones experimentales de STD. En general, los diferentes confórmeros obtenidos se reparten entre los dos dominios y la zona central, sin una clara tendencia de afinidad. Tras la determinación de las zonas más favorables para la interacción, se generaron modelos más localizados para el resto de los cálculos de Dinámica Molecular en agua explícita. Los resultados de estas simulaciones se cotejaron con los datos de NOESY transferido que describen la conformación enlazada. Frecuentemente al aumentar el número de sustituyentes aromáticos no hubo acuerdo entre los datos experimentales y la simulación. En estos casos se emplearon métodos de MD-TAR para su corrección. En los casos en los que la discrepancia se mantuvo se procedió al cambio de la estructura de partida del complejo por otra solución de las simulaciones de *docking*, lo que en la mayoría de los casos condujo a modelos compatibles con los resultados experimentales.

Finalmente, la medida de Energía de interacción calculada empleando MMGBSA confirmó los valores obtenidos de los experimentos de Polarización de Fluorescencia, donde la presencia de grupos aromáticos aumentaba esta afinidad. Los cálculos de MMPBSA avalaron que los aminoácidos con mayor relevancia en esta interacción eran las lisinas y argininas situadas preferentemente en el dominio CTD y zona central de la PTN.

➤ *Complejos con Midkina*

Nuevamente, se emplearon métodos de RMN de observación del ligando para el estudio de los compuestos **1**, **2**, **4** y **5** con Midkina. Los NOESY transferidos y los STD son compatibles con la formación de complejos que mantienen la misma conformación que en el estado libre, pero con múltiples orientaciones relativas respecto a la MK. Las medidas de Polarización de Fluorescencia indican que la afinidad por esta proteína es mayor que por PTN, y de nuevo aumenta con el número de anillos aromáticos de la molécula.

Los estudios teóricos se restringieron a la serie pentasacáridica para las tres conformaciones distintas de MK. Al igual que para la PTN, para corregir las discrepancias con los datos de RMN, fue necesario recurrir frecuentemente a MD-TAR para reproducir

los resultados experimentales, e incluso cambiar la orientación relativa del ligando eligiendo otro conformero del proceso de *docking*.

Los estudios de *docking* molecular indican una cierta preferencia de los conformeros por el dominio NTD y la zona *hinge* central.

Los análisis de los cálculos MM-PBSA/GBSA se realizaron sobre la misma zona de interacción que en el caso de la PTN, sin apreciarse diferencias importantes respecto a los aminoácidos involucrados, siendo participes de nuevo lisinas y argininas que interaccionan con los grupos sulfato de los miméticos.
