



Universidad de Valencia y Centro de Investigación Príncipe Felipe

TESIS

Para obtener el grado de Doctor de la Universidad de Valencia

Programa de Doctorado: Biología Celular (030G) del Departamento de
Bioquímica y Biología Molecular de la Facultad de Medicina

Título del Proyecto de Tesis

CARACTERIZACIÓN ESTRUCTURAL DE HEPARANASA POR RMN.
NUEVAS ESTRATEGIAS EN LA BÚSQUEDA DE FÁRMACOS
ANTIMETASTÁSICOS

Doctorando: Silvia Mosulén Machuca

Director de Tesis: Antonio Pineda Lucena

Co-director de Tesis: Rodrigo J. Carbajo Martínez

Tutor de Tesis: Jose Enrique O'Connor Blasco.



Antonio Pineda Lucena, Doctor en Químicas contratado en el Centro de Investigación Príncipe Felipe y **Rodrigo José Carbajo Martínez**, Doctor en Químicas contratado en el Centro de Investigación Príncipe Felipe,

CERTIFICAN

Que la presente Memoria titulada “*Caracterización Estructural de heparanasa por RMN. Nuevas estrategias en la búsqueda de fármacos antimetastásicos*”, ha sido realizada bajo su dirección en el Laboratorio de Bioquímica Estructural del Centro de Investigación Príncipe Felipe de Valencia, por la Licenciada en Ciencias Químicas Silvia Mosulén Machuca, y autorizan su presentación para que sea calificada como Tesis Doctoral.

Fdo.: Antonio Pineda Lucena

Valencia, 8 de mayo de 2012

Fdo.:Rodrigo José Carbajo Martínez



PRINCIPE FELIPE

CENTRO DE INVESTIGACION



VNIVERSITAT
D VALÈNCIA

Jose Enrique O'Connor Blasco, Doctor en Farmacia, Catedrático de Bioquímica y Biología Molecular de la Universidad de Valencia,

CERTIFICA

Que la presente Memoria titulada “*Caracterización Estructural de heparanasa por RMN. Nuevas estrategias en la búsqueda de fármacos antimetastásicos*”, ha sido realizada bajo su tutela en el Laboratorio de Bioquímica Estructural del Centro de Investigación Príncipe Felipe de Valencia, por la Licenciada en Ciencias Químicas Silvia Mosulén Machuca, y autoriza su presentación para que sea calificada como Tesis Doctoral.

Valencia, 8 de mayo de 2012

Fdo.: Jose Enrique O'Connor Blasco

Agradecimientos

Agradecimientos

Agradecimientos

Este trabajo ha sido realizado en el Centro de Investigación Príncipe Felipe (CIPF) de Valencia, España, dentro del Programa de Doctorado Señalización Celular (030G), iniciado en el curso 2006/2007 en la Universidad de Valencia y bajo el soporte de una beca FPI concedida por el Ministerio de Educación y Ciencia.

Agradecimientos

Este trabajo ha conllevado un gran esfuerzo y por esta razón quería dar las gracias a todas aquellas personas que han contribuido de alguna manera, leyendo, opinando, corrigiendo, teniendo paciencia, dando ánimo y acompañando en los momentos difíciles y momentos de felicidad.

En primer lugar quería agradecer a mis directores de Tesis, Dr. Antonio Pineda-Lucena por haberme dado esta oportunidad, haberme aceptado en su grupo y por la dirección de este trabajo y al Dr. Rodrigo J. Carbajo por haberme ayudado con la RMN, por su apoyo y por haber creído en mí especialmente durante la última etapa de este trabajo, que ha sido la más dura.

Gracias al Prof. Santos Fustero y al Dr. Jose Luis Aceña por su contribución con la síntesis de uno de los compuestos utilizados en este trabajo.

Special thanks to the Dr. Hubber, director of Structural Biology Group Roche (Basel, Switzerland) and his group for their kind hospitality during the summer 2010 and for sharing their knowledge in Surface Plasmon Resonance with me.

I gratefully acknowledge the financial support provided by the Access to Research Infrastructures activity in the 6th FP of the EC (Contract #RII3-026145, EU-NMR), and the CERM for providing technical support.

Quería dar un agradecimiento especial a la Prof. Marta Bruix por haberme apoyado cuando más necesitaba que alguien confiara en mí. Gracias de corazón.

Gracias también al Prof. Jose M^a Vega, por confiar en mí cuando todavía era su alumna, por creer en mis posibilidades y darme la oportunidad de hacer el DEA en su laboratorio. Mil gracias.

Agradecimientos

Gracias a todos los compañeros de laboratorio y en prácticas que me han sacado alguna vez una sonrisa y me han ayudado a realizarme como persona y especialmente a mi compañero Rafael Gozalbes por los estudios in silico realizados en este trabajo.

Especial agradecimiento a mis compañeros Beatriz y David, por vuestra ayuda en mi trabajo, por vuestro apoyo y comprensión durante todos estos años, por no dejarme sola y sobretodo por vuestra sincera amistad. Sin vosotros, chicos, todo hubiera sido mucho más complicado.

Gracias también a Laura Cascales, que aunque coincidimos poco tiempo, la consideré siempre mi amiga. Gracias Laurita por la bici, por tu apoyo y amistad.

Y que podría decir de mi compañero y amigo de Tesis. Gracias Guillermo, por todo, por tu amistad, tu ayuda sincera, por haber estado siempre a mi lado, en los buenos y malos momentos, por entenderme, por sacarme la sonrisa cuando parecía que ya no podría reír, por apoyarme en todo y por estar siempre cuando lo he necesitado. En definitiva, gracias por tu cariño. Sabes que para mí tu amistad es lo más valioso que me llevo de estos años en Valencia. Mil gracias de verdad.

Un agradecimiento sincero a todos aquellos amigos que han compartido conmigo estos años en el plano personal especialmente a mis compañeros de piso Natxo, Arantxa y Silvia, por hacerme la estancia en Valencia más agradable y por compartir esos ratos de charla donde me podía sentir como en casa.

Por supuesto mi más profundo agradecimiento va para mi familia. Gracias a mis padres y hermanos por todo su apoyo incondicional, por entender mi momentos más difíciles, por creer en mí, apoyarme siempre y por su paciencia

Agradecimientos

infinita. Sin vuestro cariño y comprensión hubiera sido imposible llevar a cabo este trabajo. Os quiero.

Especial mención quería hacer a mi sobrinita, Mar, porque sin ella saberlo me ha llenado de alegría en momentos muy difíciles y porque simplemente mirando su carita y con sus besitos podía olvidar cualquier momento complicado. Gracias mi niña por nacer y por ser tan preciosa.

Finalmente, quería agradecer a mi marido, Franck, sin el que este sueño no hubiera podido hacerse realidad. Gracias por confiar siempre en mí, por levantarme siempre en mis momentos más duros, por creer siempre que soy la mejor, por no dudar ni un momento que esta tesis iba a salir adelante, por tu paciencia infinita, por comprenderme, por animarme, por aguantar la distancia, por estar siempre conmigo, presente o al otro lado del teléfono, por ilusionarme con pequeñas cosas y especialmente por tu amor. Te quiero cariño.

*A mis padres, mis hermanos y a mi sobrina Mar por todo
lo que habéis hecho por mí durante todos estos años*

A Franck, por estar siempre a mi lado

Índice

Índice	1
Abreviaturas	9
Introducción	15
1.- Cáncer	17
2.- Proteínas implicadas en metástasis y angiogénesis	31
2.1.- Proteínas implicadas en angiogénesis	31
2.2.- Quimioquinas	38
2.3.- Proteínas que intervienen en el proceso de invasión y en la transición epitelial-mesenquimal (EMT)	40
2.4.- Papel que juegan las proteínas CD44 en el proceso de metástasis	42
2.5.- Metaloproteasas de la matriz extracelular	42
3.-Heparanasa	44
3.1.- Proteoglicanos de heparán sulfato (HSPGs)	44
3.2.- Antecedentes de Heparanasa	45
3.3.- Propiedades moleculares y bioquímicas de la enzima	46
3.4.- Regulación	48
3.4.1.-Metilación del promotor	48
3.4.2.- Regulación por p53	48
3.4.3.- Regulación por mediadores de la inflamación	49
3.4.4.-Regulación por estrógeno y cáncer de mama	49
3.4.5.-Regulación por glucosa	50
3.5.- Localización celular de heparanasa	50
3.6.- Función fisiológica en tejidos sanos	51
3.7.- Heparanasa y progresión del cáncer	52
3.7.1.- Propiedades pro-angiogénicas	52
3.7.2.-Potencial metastásico	53

3.7.3.-Inflamación	54
3.7.4.-Actividades no enzimáticas de heparanasa	54
3.8.-Heparanasa en la terapia del cáncer	55
4.- Uso del cribado de fragmentos en el desarrollo de nuevos fármacos	58
4.1.-Definición y ventajas	58
4.2.-Técnicas utilizadas en el cribado de fragmentos	61
4.3.-Cribado de fragmentos por espectroscopia de Resonancia Magnética Nuclear (RMN)	65
4.3.1.-Experimentos basados en la observación de la Macromolécula. SAR por RMN	65
4.3.2.-Experimentos basados en la observación del ligando	66
Materiales y Métodos	69
1.-Biología Molecular	71
1.1.-Heparanasa	71
1.1.1.-Diseño y preparación de las construcciones de Heparanasa	71
1.1.2.-Expresión de las distintas construcciones	80
1.1.3.-Marcaje isotópico de las construcciones de heparanasa	82
1.1.4.-Marcaje isotópico selectivo de las construcciones de heparanasa	82
1.1.5.-Lisis celular de las construcciones de heparanasa	82
1.1.6.-Purificación de las construcciones clonadas en p11 y pET47b	83
1.1.7.-Purificación de HEP ₃₆₋₄₁₇ -pMAL	85
1.1.8.-Purificación de HEP ₃₆₋₄₁₇ -pET44A	86
1.1.9.- Identificación de la secuencia y peso molecular de HEP ₁₅₈₋₄₁₆ por espectrometría de masas MALDI-TOF	86
1.2.-Purificación de fondaparinux	88

2.-Resonancia Magnética Nuclear	88
2.1.-Preparación de la muestra	88
2.1.1.-Experimentos bi- y tri-dimensionales necesarios para la asignación de la proteína	88
2.1.2.-Ensayos de interacción	90
2.2.-Adquisición de los experimentos	91
2.2.1.-Experimentos bi- y tri-dimensionales necesarios para la asignación de la proteína	91
2.2.2.-Ensayos de interacción	96
2.3.-Asignación del esqueleto y cadenas laterales	97
2.4.-Cálculo de estructura secundaria	100
2.5.-Cálculo de estructura tridimensional	101
3.-Modelado molecular	102
3.1.-Desarrollo del modelo farmacofórico	102
3.2.-Cribado de fragmentos con el modelo farmacofórico	106
3.3.-Modelo de homología de heparanasa	112
3.4.-Docking	113
4.-Ensayos de interacción por Resonancia de Plasma Superficial (SPR)	114
4.1.- Inmovilización de heparanasa a la superficie del sensor	115
4.2.- Inyección del ligando	116
4.3.-Regeneración	117
5.-Estrategia de búsqueda y selección de nuevos inhibidores potenciales de heparanasa	117
5.1.- Selección de inhibidores a partir de la quimioteca de fragmentos	117
5.2.-Selección de inhibidores procedentes de la librería PRESTWICK	120

Discusión y Resultados	125
1.-Análisis del modelo estructural de heparanasa	127
2.-Clonación, expresión y purificación de distintas construcciones de heparanasa	130
3.-Optimización de las condiciones de solubilidad y estabilidad para los estudios de RMN	147
4.-Asignación de heparanasa por Resonancia Magnética Nuclear	147
5.-Cálculo de estructura secundaria de HEP₁₅₈₋₄₁₇ y HEP₃₇₋₁₀₉	159
5.1.-Estructura secundaria de HEP ₁₅₈₋₄₁₇	159
5.2.-Estructura secundaria de HEP ₃₇₋₁₀₉	163
5.3.-Asignación de NOEs y determinación de estructura tridimensional de HEP ₁₅₈₋₄₁₇	165
6.-Optimización de las condiciones de solubilidad y estabilidad para los estudios de SPR	167
7.-Validación de la construcción HEP₁₅₈₋₄₁₇ como diana de cribado por RMN Y SPR	171
8.-Cribado de fragmentos por combinación de química computacional y diferentes técnicas biofísicas	180
8.1.-Generación de un modelo farmacofórico para heparanasa	180
8.2.-Fiabilidad del modelo farmacofórico	184
8.3.-Análisis de la diversidad química de los compuestos que cumplían las hipótesis farmacofóricas	186
8.4.- Identificación de “hits” por Resonancia Magnética Nuclear y validación del modelo farmacofórico por RMN	188

8.5.-Identificación de “hits” por Resonancia de Plasma Superficial	192
9.-Cribado de compuestos con características de fármacos por combinación de química computacional y diferentes técnicas biofísicas	194
10.-Optimización de “hits”	198
10.1.-Optimización de “hits” que provienen del cribado de fragmentos	198
10.1.1.-Búsqueda de análogos comerciales	199
10.1.2.-Búsqueda de análogos mediante síntesis orgánica. Síntesis de una molécula simétrica	203
10.2.-Optimización de “hits” que provienen del cribado de compuestos con características de fármacos	205
Conclusiones	213
Bibliografía	217

Abreviaturas

- ADN: Ácido Desoxirribonucleico
- a-FGF: "Acid-Fibroblast Growth Factor" (Factor de Crecimiento para Fibroblastos ácido)
- ALT: Alargamiento Alternativo de los Telómeros
- Ang: Angiopoyetinas
- ARN: Ácido Ribonucleico
- b-FGF: "basic-Fibroblast Growth Factor" (Factor de Crecimiento para Fibroblastos básico)
- BM: "Subendothelial Basal Membrane" (Membrana Basal Subendotelial)
- BMRB: "Biological Magnetic Resonance Bank" (Banco de Resonancia Magnética Biológica)
- CARA "Computer-Aided Resonance Assignment" (Asignación de Resonancias Asistida por Ordenador)
- CM: "Carboxymethylated" (Carboximetilada)
- COMP: Compuesto
- CSF: "Colony Stimulating Factor" (Factor estimulante de colonias)
- DMSO: "Dimethyl sulfoxide" (Dimetil Sulfoxido)
- DTH: "Delayed-Test Hypersensitivity" (Ensayo de Hipersensibilidad Retardada)
- DTT: "Dithiothreitol" (Ditiotreitól)
- ECM: "Extra-Celular Matriz" (Matriz Extra-Celular)
- EDC: Hidrocloruro de Carbodiimida
- EDTA: "Ethylenediaminetetraacetic acid" (Ácido etilendiaminotetraacético)
- EGF: "Epidermal Growth Factor" (Factor de Crecimiento de la Epidermis)
- EGFR: "Epidermal Growth Factor Receptor" (Receptor del Factor de Crecimiento de la Epidermis)
- EMT: "Epithelial-Mesenchymal Transition" (Transición Epitelial-mesenquimal (EMT))
- FBDD: "Fragment Based Drug Design" (Descubrimiento de Fármacos Basado en el Cribado de Fragmentos)

- FGFR: “Fibroblast Growth Factor Receptor” (Receptor del Factor de Crecimiento para Fibroblastos)
- FRED: “Fast Rigid and Exhaustive Docking” (Acoplamiento Rápido, Rígido y Exhaustivo)
- GH-A: Glicosil Hidrosilasas A
- GndCl: Cloruro de Guanidinio
- GPCR: “G-protein coupled receptor” (Receptor Acoplado a proteínas G)
- HEP: Heparanasa
- HIV: “Human Immunodeficiency Virus” (Virus de Inmunodeficiencia Humano)
- HS: “Heparan Sulphate” (heparán sulfato)
- HSPG: “Heparan Sulphate Proteoglycans” (Proteoglicanos de Heparán Sulfato)
- HTS: “High Throughput Screening” (Cribado de Alto Rendimiento)
- IAM: “Iodoacetamide” (Yodoacetamida)
- IL-8: Interleuquina 8
- IPTG: Isopropil- β -Tiogalactopiranosido
- IMAC: “Immobilized Metal Chromatography” (Cromatografía de Metal Inmovilizado)
- INT- γ : Interferón-gamma
- IUPAC: “International Union of Pure and Applied Chemistry” (Química aplicada y pura de la Unión Internacional)
- K_d : Constante de disociación
- LE: “Ligand Efficiency” (Eficiencia de Unión)
- MMP: “Extracellular Matrix Metalloproteinase” (Metaloproteasas de la Matriz Extracelular)
- MOE: “Molecular Operating Environment” (Entorno Operativo Molecular)
- NHS: N-hidroxisuccinimida
- NOE: “Nuclear Overhauser Effect” (Efecto Nuclear Overhauser)

- NTA: “Immobilized Ni²⁺-Nitrilotriacetic Acid” (Ácido nitrilotriacético-Ni²⁺ inmovilizado)
- OMEGA: “Optimized Molecular Ensemble Generation Application” (Aplicación de la Generación del Ensamblaje Molecular Optimizado)
- OMS: Organización Mundial de la Salud
- PCR: “Polymerase Chain Reaction” (Reacción en Cadena de la Polimerasa)
- PDGF: “Platelets-Derived Growth Factor” (Factor de Crecimiento Derivado de Plaquetas)
- PI3K: “Phosphatidylinositol 3-Kinase” (Fosfatidilinositol 3-Quinasa)
- RMN: Resonancia Magnética Nuclear
- RTK: Receptor de Tirosina Kinasa
- RU: “Resonance Units” (Unidades de Resonancia)
- SAR “Structure Activity Relationships” (Relación Estructura-Actividad)
- SPR: “Surface Plasmon Resonance” (Resonancia de Plasma Superficial)
- STD: “Saturation transfer difference” (Diferencia de Transferencia de Saturación)
- TCA: “Trichloroacetic Acid” (Ácido Tricloroacético)
- TEV: *Tobacco etch virus*
- TGF: “Transforming Growth Factor” (Factor de Crecimiento Transformante)
- TIMP: “Tissue Inhibitor Metalloproteinase” (Inhibidores de Metaloproteasas de Tejidos)
- TFA: “Trifluoroacetic Acid” (Ácido Trifluoroacético)
- TSP-1: Trombospondina-1
- VEGFR: “Vascular Endothelial Growth Factor Receptor” (Receptor del Factor de Crecimiento Vascular Endotelial)
- VEGF: “Vascular Endothelial Growth Factor” (Factor de crecimiento Vascular Endotelial)

Introducción

1.- CÁNCER

Según datos de la OMS, el cáncer es el responsable de aproximadamente el 13% del total de defunciones anuales, lo que lo convierte en la segunda causa de mortalidad a nivel mundial después de las enfermedades cardiovasculares. Se prevé que la mortalidad por cáncer aumentará un 45% entre 2007 y 2030 (se pasará de 7,9 millones a 11,5 millones de defunciones) debido en parte al crecimiento demográfico y al envejecimiento de la población. Es conocido que la mayoría de las muertes por cáncer no se deben a neoplasias primarias, sino que estos tumores primarios invaden tejidos y/o vasos sanguíneos adyacentes transformándose en carcinomas infiltrantes o tumores secundarios. Esta capacidad de formar tumores secundarios en otros lugares del organismo recibe el nombre de “metástasis”, y es la causa del 90% de las muertes por cáncer (1).

Hasta la fecha se han identificado más de 200 tipos de cáncer, que se pueden clasificar según el tipo de tejido en el cual se originan en (2):

- Carcinomas: son tumores derivados de tejidos epiteliales. Son los más abundantes, aproximadamente un 80%. Se pueden clasificar en carcinomas escamosos, que derivan de células epiteliales que sellan la cavidad que recubren para proteger el interior de los órganos (cáncer de colon, ovarios, próstata, páncreas, estómago, esófago, pulmón) y los adenocarcinomas, que derivan de células epiteliales especializadas que secretan sustancias a los conductos/cavidades que recubren (cavidad nasal, laringe, esófago, cervicales, pulmón). Hay otro tipo de carcinomas que son mezclas de los dos anteriores como por ejemplo el carcinoma hepatocelular, el carcinoma de células de pulmón o los carcinomas de células renales.
- Sarcomas: derivan de células del tejido conectivo. No son muy abundantes (1%) y todos comparten un origen embrionario común, el mesodermo. Derivan de muchos tipos celulares mesenquimales (estroma) como fibroblastos, adipocitos, osteoblastos y miocitos.

- Leucemias/linfomas: incluyen tumores que derivan de los tipos celulares que forman los tejidos productores de sangre así como los del sistema inmune, eritrocitos, células plasmáticas y linfocitos B y T. Las leucemias son precursores de tejidos hematopoyéticos malignos que se mueven libremente y no forman masas tumorales sólidas mientras que los linfomas derivan de linajes linfoides, no son circulantes y se agregan en masas sólidas en los nódulos linfáticos.
- Tumores neuroectodérmicos: derivan de la capa externa del embrión. No son muy abundantes (aproximadamente un 1,3%), sin embargo muestran una mortalidad muy elevada (el 2,5% de todas las muertes).
- Melanomas: derivan de melanocitos, células pigmentadas de la piel y la retina.
- Carcinoma de pulmón microcítico: deriva de células con propiedades neurosecretoras y epiteliales. Su origen es probablemente endodérmico.
- Tumores desdiferenciados o anaplásicos: no preservan propiedades del tipo celular del que derivan.

Después de 60 años de grandes avances, la investigación sobre el cáncer ha revelado que ésta no es una enfermedad sino un grupo extenso de patologías que producen importantes cambios en el genoma (3). El cáncer es causado por alteraciones genéticas de una célula parental tumoral. Este daño genético no puede corregirse, de tal forma que es un rasgo hereditario de todas las células de generaciones subsecuentes. Sin embargo, un solo cambio genético muy rara vez es suficiente para el desarrollo de un tumor maligno y muchas evidencias demuestran que es un proceso de alteraciones genéticas sucesivas.

Estas alteraciones genéticas van asociadas a ciertos factores de riesgo comunes como el consumo de tabaco y alcohol, dieta inadecuada, falta de actividad física, exposición a carcinógenos en el ambiente laboral (por ejemplo, el amianto) o en el medio ambiente (por ejemplo, por contaminación del aire en

locales cerrados), radiación (ultravioleta o ionizante) y algunas infecciones (por ejemplo, hepatitis B o infección por virus del papiloma humano).

El daño genético encontrado en células cancerosas puede ser (4, 5):

- Dominante. Es el producido sobre los genes llamados proto-oncogenes y oncogenes. La distinción entre los términos proto-oncogén y oncogén se relaciona con la actividad de la proteína producto del gen. Un proto-oncogén es un gen cuya proteína promueve el crecimiento y división celular. Un oncogén es un gen que ha tenido algún daño genético y, por lo tanto, produce una proteína capaz de provocar transformación celular. El proceso de activación de un proto-oncogén a oncogén puede incluir transducción por retrovirus o integración retroviral, mutaciones puntuales, mutaciones de inserción, amplificación de genes, traslocación de cromosomas y/o interacciones proteína-proteína (6, 7). Los proto-oncogenes se pueden clasificar en muchos grupos diferentes basados en su función normal dentro de las células o en la homología de secuencia con otras proteínas conocidas. Según lo predicho, los proto-oncogenes se han identificado en todos los niveles de las varias vías de transducción de señal que controlan el crecimiento, la proliferación y la diferenciación celular.
- Recesivo. Es el producido sobre genes denominados supresores de tumores, supresores de crecimiento, oncogenes recesivos o anti-oncogenes.
- Genes de microARN. A diferencia de otros genes implicados en cáncer, no codifican para ninguna proteína. Los productos de estos genes son simples cadenas de ARN de unos 21-23 nucleótidos cuya función es regular la expresión de los genes. Una molécula de ARN puede unirse a una molécula de ARN mensajero que contenga la secuencia complementaria de un microARN. De esta forma el microARN bloquea la traducción de la proteína o lo que es lo mismo provoca la degradación del ARN mensajero.

El seguimiento de la expresión de estos genes ha permitido la clasificación de tumores, su diagnóstico y progresión así como el pronóstico y posible respuesta a tratamientos.

En el transcurso de la oncogénesis o carcinogénesis (paso de una célula sana a un tumor maligno), la célula puede sufrir hasta seis alteraciones en su fisiología (Fig.1), adquiriendo las siguientes capacidades:

- División celular en ausencia de señales externas que estimulan el crecimiento.

Las células de tejidos sanos sólo se dividen si reciben señales externas de otras células que estimulan el crecimiento a través de proteínas llamadas factores de crecimiento. Estos son detectados en la matriz extracelular (*Extra-Cellular Matrix*, ECM) a través de los llamados receptores de factores de crecimiento que pueden ser endoteliales vasculares (*Vascular Endothelial Growth Factor Receptor*, VEGFR) o de la epidermis (*Epidermal Growth Factor Receptor*, EGFR). Estos receptores son proteínas transmembranales que pertenecen a la familia de receptores de tirosina quinasas (RTKs). Cuando el EGFR se acopla a uno de sus ligandos (EGF, Her2, Neu, ErbB2) se activan ciertas rutas de transducción de señales que favorecen la proliferación o inhiben la apoptosis (8).

Sin embargo, las células tumorales no dependen de estos factores de crecimiento externos, sino que producen y segregan dentro de la matriz extracelular sus propios factores mitogénicos que finalmente estimularán la proliferación. Además, estas células tumorales son capaces de sobreexpresar receptores de tirosina quinasas o formar RTKs estructuralmente modificados que activan las rutas de transducción de señal incluso en ausencia de la estimulación de factores mitogénicos.

Por otro lado, la oncoproteína Ras que participa en rutas de transducción de señales, normalmente permanece en estado de reposo en ausencia de factores de crecimiento, sin embargo en el 20% de las células tumorales, esta oncoproteína

está mutada de forma que transmite constantemente señales de estimulación del crecimiento (9).

Todos estos mecanismos hacen que las células neoplásicas proliferen incluso en ausencia de factores de crecimiento exógenos.

- Crecimiento en presencia de señales inhibitorias de crecimiento.

Siempre que las células se dividen para construir órganos o tejidos, éstas pasan a través de otras células o tejidos. Con el fin de proteger la integridad y la función de los mismos, las células transmiten señales de inhibición del crecimiento. Este mecanismo llamado “inhibición de contacto” (10) es crucial para la viabilidad de todos los organismos superiores. Las células cancerígenas han perdido este mecanismo y se dividen incluso en presencia de señales inhibitorias del crecimiento (11), por lo tanto las neoplasias pueden invadir e incluso destruir tejidos durante su intensa proliferación.

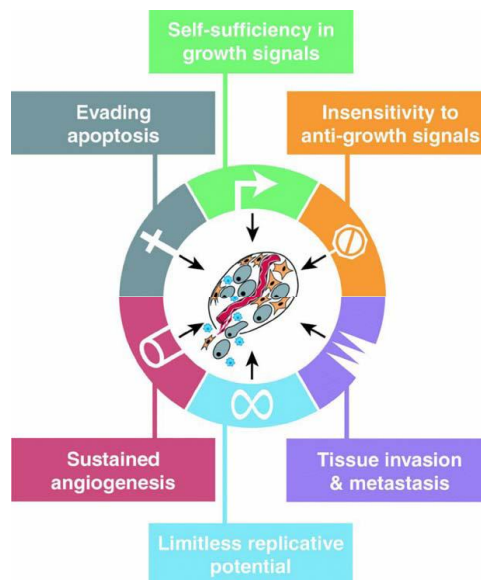


Figura 1. Capacidades adquiridas del cáncer (3).

- Evasión de apoptosis.

Existen al menos dos tipos de muerte celular, necrosis y apoptosis (muerte celular programada). La apoptosis es un proceso estrictamente regulado, que juega un papel muy importante en el desarrollo, morfogénesis, remodelación de tejidos y respuesta inmunitaria, mediando la eliminación ordenada de células viejas o dañadas. Al contrario que las células necróticas que pueden provocar una reacción inflamatoria, las células apoptóticas se eliminan de forma discreta principalmente por fagocitosis. Además la apoptosis se distingue por unos rasgos morfológicos y bioquímicos característicos que incluyen contracción celular, la fragmentación del ADN nuclear y la globulación de las membranas celulares (12).

En células y tejidos de organismos multicelulares, la desregulación de la apoptosis o una resistencia a la misma puede dar lugar a una proliferación celular excesiva y una alta acumulación de células que contribuye al desarrollo de tumores y enfermedades autoinmunes (13). Por lo tanto, la resistencia de las células tumorales a la apoptosis es esencial para el desarrollo futuro del cáncer. Esta hipótesis se confirma por el hecho de que la sola proliferación descontrolada no es suficiente para la formación del tumor, ya que por el contrario conduciría a la muerte celular (12). Por lo tanto, la progresión del tumor requiere además de la expresión de proteínas que promuevan la proliferación celular, la expresión de proteínas antiapoptóticas o la inactivación de proteínas proapoptóticas esenciales (14).

Bcl-2 fue la primera proteína implicada en la regulación de la apoptosis caracterizada molecularmente. Se tardó varios años en conocer cómo esta proteína ejercía su efecto anti-apoptótico. Hockenbery en 1990 demostró que esta proteína podía suprimir la producción de especies reactivas del oxígeno en la mitocondria sugiriendo que los efectos anti-apoptóticos se producían a nivel de orgánulos subcelulares (15). Posteriormente se demostró que Bcl-2 era capaz de inhibir la liberación de citocromo c uniéndose a una proteína proapoptótica BAX. El citocromo c tiene un papel crucial para la activación de las caspasas 9 y

3, que intervienen a su vez en la activación del proceso apoptótico (16). Durante la década de los 90 creció el número de miembros de la familia de Bcl-2 implicados en la regulación de apoptosis y se descubrió que proteínas que tenían sólo un dominio BH₃ tales como BIM (también conocida como Bcl2L11), podían interactuar con Bcl-2 para permitir la liberación de BAX y así permitir la liberación de citocromo c a la mitocondria (Fig. 2) (13).

Además de la familia de Bcl-2, se identificaron otros muchos oncogenes y supresores de tumores que estaban implicados en la regulación de la apoptosis. Entre ellos cabe destacar el oncogén MYC y el supresor de tumor p53. MYC es capaz de transformar y sensibilizar a las células en condiciones no favorables para ello, aumentando la proliferación celular lo que induce apoptosis. Quizás el ejemplo más famoso de gen supresor es TP53, que es el gen que aparece mutado con más frecuencia (50%) en cánceres humanos y que codifica la proteína p53 también llamada el “guardián del genoma” (17). En 1991, se demostró que la expresión de p53 en células leucémicas inducía apoptosis (18). Posteriormente también se comprobó que p53 regulaba la expresión de muchas proteínas clave en la regulación de apoptosis como por ejemplo BAX, BIM, NOXA o PUMA (19-22).

Por otro lado, también fue de particular interés la idea de tratar tumores con anticuerpos específicos. En el grupo de Krammer (German Cancer Research Center, Heidelberg), consiguieron producir un anticuerpo monoclonal que inducía apoptosis en células linfoides. Al antígeno reconocido por este anticuerpo se le llamó FAS/CD95 (23). Como se dijo anteriormente, las células cancerígenas desarrollan ciertas estrategias que inhiben la apoptosis incluso en condiciones en las que normalmente estaría activada, por ejemplo contienen mutaciones en el gen CD95 incrementando la síntesis de factores antiapoptóticos como Bcl-2, FLIPs, IAPs, survivin y también mutaciones que provocan la pérdida de funcionalidad en genes pro-apoptóticos tales como BAX (1). En la Fig. 2 se describe con detalle la ruta de señalización y regulación de apoptosis (13).

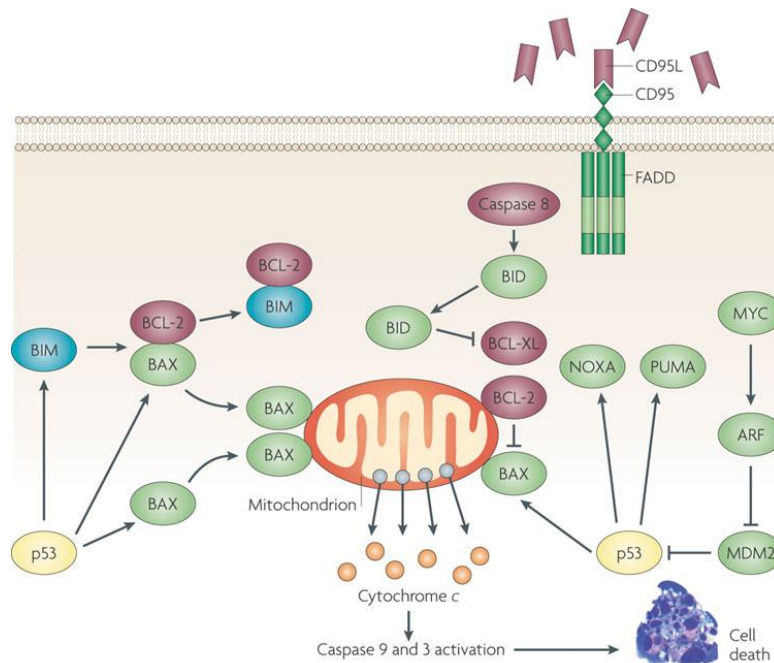


Figura 2. Ruta de señalización y regulación de apoptosis (13).

o Neovangiogénesis

La angiogénesis es la formación de nuevos vasos sanguíneos desde el endotelio de otros vasos ya existentes y es fundamental en el crecimiento de tumores, progresión y metástasis (24). En condiciones fisiológicas normales es un proceso esencial que incluye cicatrización de heridas, reparación de tejidos, reproducción, crecimiento y desarrollo. Los tejidos que son particularmente activos y proliferan muy rápidamente como es el caso de los tejidos tumorales, requieren una buena conexión con el sistema circulatorio ya que les garantiza un suministro adecuado de nutrientes y oxígeno y les asegura un drenaje eficiente de metabolitos y productos de desecho que provienen de las rutas metabólicas (25).

En 1945, Algire y Chalkley (26), fueron los primeros en concluir que el crecimiento de un tumor sólido estaba en estrecha relación con el desarrollo de vascularización propia. El proceso de vascularización de un tumor se definió con

el término de “neoangiogénesis”. Además de que este proceso es necesario para el crecimiento de un tumor primario, la metástasis depende de la neovascularización en dos de sus etapas fundamentales: la primera, cuando las células tumorales malignas abandonan un tumor primario y penetran en el interior del torrente sanguíneo y la segunda, después de la llegada a otro tejido u órgano, para poder proliferar y formar un tumor secundario.

En 1970, el cirujano Folkman (27) fue el primero en sugerir que la inhibición de la formación de nuevos vasos sanguíneos permitiría frenar el crecimiento de un tumor e incluso la disminución de su tamaño. La base fisiológica de esta hipótesis es que un tumor no puede crecer más de 1-2 mm³ (~10⁷ células) sin vascularización propia. Sin embargo, antes de producir neoangiogénesis, las células cancerígenas de un tumor primario pueden permanecer en un periodo de inactividad que puede durar años (28).

Actualmente se sabe que este proceso está muy bien regulado por un equilibrio entre factores pro- y anti-angiogénicos (Fig. 3) (29, 30).

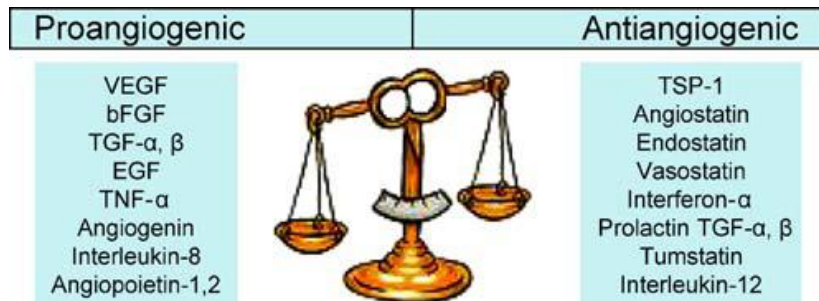


Figura 3. Balance entre factores pro- y anti-angiogénicos en la regulación de la angiogénesis (25).

Cuando los factores proangiogénicos superan el efecto de las moléculas antiangiogénicas, los tumores adquieren un fenotipo angiogénico que permite la formación de nuevos vasos sanguíneos (31). Este fenotipo angiogénico es el paso clave por el que una lesión microscópica puede transformarse en una gran masa con propiedades metastásicas. La expresión de proteínas derivadas de

oncogenes, así como un gran número de factores de estrés celular tales como hipoxia, bajo pH, privación de nutrientes o inductores de oxígeno reactivo, son muy importantes en la estimulación de señales angiogénicas (30). Los capilares sanguíneos están formados por las células endoteliales y la membrana basal; cuando los factores angiogénicos producidos por las células tumorales se unen a los receptores de las células endoteliales se inicia la secuencia de etapas angiogénicas o también llamada “cáscada angiogénica” que incluye una fase de activación y una de resolución. Cuando se estimula el crecimiento de las células endoteliales, éstas segregan proteasas y otras enzimas que digieren la membrana basal que rodea a los vasos sanguíneos. Esta rotura de la matriz extracelular permite la liberación de factores proangiogénicos y por tanto la formación de nuevos capilares. La etapa de resolución es aquella en la que se produce la maduración y estabilización de los nuevos capilares sanguíneos (25). La angiogénesis, por tanto, es el resultado de una serie de procesos celulares y moleculares que inducen la migración, proliferación y diferenciación de células endoteliales dentro de nuevos capilares que seguidamente pueden desarrollarse y formar vasos sanguíneos más grandes.

- Potencial infinito de replicación

Hayflick demostró que las células humanas en cultivos celulares paran de crecer y por lo tanto mueren después de ~50 mitosis (32). A este proceso se le llama “senescencia”. La causa principal de este proceso es el acortamiento progresivo de los telómeros que representan los extremos de los cromosomas. Se pierden entre 50-100 pares de bases de ADN telomérico en cada replicación, provocando inevitablemente la muerte celular. En células cancerosas se expresa una enzima llamada telomerasa que se caracteriza por solucionar el problema del acortamiento de los telómeros añadiendo secuencias repetitivas teloméricas manteniendo así la integridad de los mismos (33) o bien porque se activa un mecanismo llamado ALT (alargamiento alternativo de los telómeros) donde los telómeros mantienen su longitud por recombinación mediante cambios de información de secuencia entre cromosomas (34). La telomerasa, por tanto,

protege a las células cancerosas de la muerte apoptótica y estimula la continua proliferación y por tanto la inmortalización.

- Invasión y metástasis

La metástasis se puede definir como la capacidad que tienen algunas células cancerosas de abandonar un tumor primario y desplazarse a través de vasos sanguíneos o linfáticos a otros tejidos y formar un tumor secundario (35-37). La metástasis se puede producir en distintos órganos, a distinta velocidad y el tratamiento clínico es diferente dependiendo del tipo de cáncer. Así por ejemplo, el cáncer de mama tiende a producir colonias metastásicas en muchos tejidos del cuerpo incluyendo tejidos del cerebro, hígado, huesos y pulmones, los tumores de próstata a menudo tienden a producir metástasis en los huesos, mientras que los carcinomas de colon tienden a formar nuevas colonias preferentemente en el hígado (Tabla 1) (38, 39). Sin embargo, la capacidad de invasión y metástasis no es del todo entendida. Por ejemplo, hay algunos cánceres como los melanomas que tienen una alta probabilidad de metástasis mientras que otros como los carcinomas escamosos de la piel y tumores de células gliales del cerebro, que rara vez producen tumores secundarios (1, 39).

Por otro lado, la relación entre el tamaño del tumor primario y el riesgo de metástasis tampoco es muy clara. Normalmente se relaciona el tamaño de un tumor primario con la mayor probabilidad de desarrollar metástasis. Así, por ejemplo, se han realizado estudios que indican que el 22% de las mujeres con un cáncer primario de pulmón de menos de 1 cm, eventualmente desarrollaron metástasis. Sin embargo, el 77% de las mujeres con un tumor de más de 8 cm desarrollaron metástasis durante su enfermedad (40). Además hay otras evidencias que constatan este hecho, por ejemplo que solo un 4% de los tumores cerebrales de menos de 1 cm de diámetro tienen mutaciones en p53, mientras que un 42% de los tumores con un diámetro mayor de 3 cm tienen dichas mutaciones. Sin embargo, a veces, incluso tumores primarios en fase inicial pueden desarrollar micrometástasis que pueden difundirse por todo el organismo, aunque rara vez terminan siendo tumores secundarios sólidos y de

gran relevancia clínica. En base a estas afirmaciones se establecieron dos modelos de metástasis (41): modelo de progresión lineal y modelo de progresión paralelo. En el modelo lineal (42), la habilidad de un tumor primario por producir metástasis solo se adquiere después de un determinado número de mutaciones y alteraciones genéticas, cuando las células son capaces de proliferar relativamente de una forma autónoma y a una velocidad competitiva; por tanto esta habilidad está limitada a los tumores de mayor tamaño. En el modelo paralelo (43), las células individuales cancerosas tanto de tumores grandes como pequeños tienen la habilidad de producir metástasis y esta habilidad se desarrolla mucho antes de que los primeros síntomas del cáncer aparezcan o de que el tumor primario pueda ser diagnosticado. La progresión paralela no cuestiona los mecanismos de crecimiento del cáncer, como la selección de clones con aptitudes competitivas o la acumulación de alteraciones genéticas, sin embargo la difusión de células cancerosas por el organismo no se produce en las etapas finales de un tumor primario (44).

Tabla 1. Principales lugares de metástasis según el tipo de cáncer (38).

Tipo de tumor	Principales lugares de metástasis
Mama	Huesos, pulmones, hígado y cerebro
Adenocarcinoma	Cerebro, huesos, glándula suprarrenal e hígado
Melanoma de la piel	Pulmones, cerebro, piel e hígado
Colon	Hígado y pulmones
Páncreas	Hígado y pulmones
Próstata	Huesos
Sarcoma	Pulmones
Melanoma ocular	Hígado

La metástasis ocurre a través de una serie de pasos complejos definidos como la cascada metastásica (Fig 4) (45). En primer lugar, las células cancerosas se separan de un tumor primario e invaden tejidos sanos adyacentes. Para ello, segregan enzimas que rompen las paredes de la matriz extracelular, facilitándose así la migración. A la invasión le sigue la intravasación, fenómeno por el que las células tumorales se unen y segregan enzimas proteolíticas que les permiten su infiltración en un nuevo vaso sanguíneo o linfático.

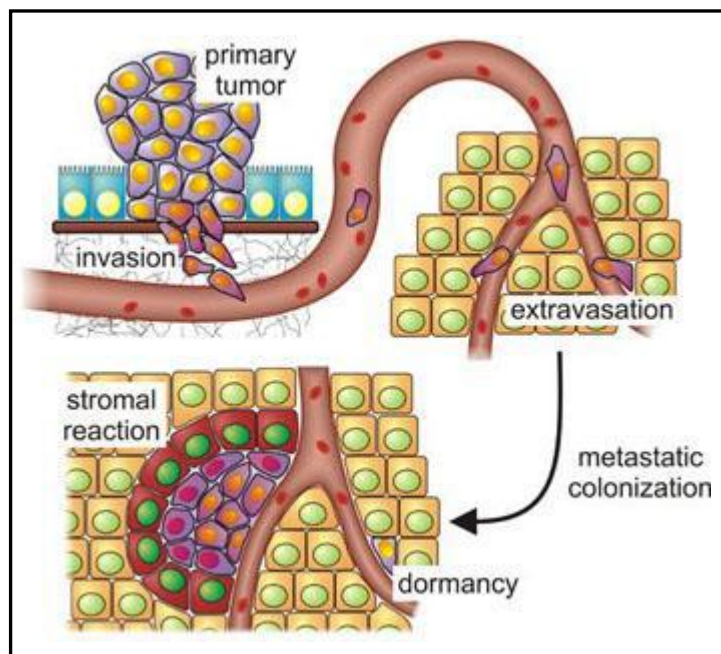


Figura 4. Principales pasos en la formación de la metástasis (45).

Una vez que estas células han llegado al lumen de un vaso sanguíneo o linfático, las células cancerígenas pueden viajar individualmente con la sangre o la linfa a otras áreas del cuerpo. Sin embargo, la circulación a través de estos vasos sanguíneos o linfáticos no es un proceso favorable; las células cancerosas al igual que las células sanas normales, pueden continuar su desplazamiento a través de estos vasos sanguíneos o linfáticos dependiendo de su anclaje a un

sustrato sólido como por ejemplo la matriz extracelular, por lo tanto el desprendimiento de la misma le puede ocasionar la muerte (llamada "muerte de anoikis"). Por otro lado, la sangre representa un medio hostil para las células cancerosas; la fuerza hidrodinámica puede ser muy grande, sobre todo en vasos sanguíneos pequeños, y puede desgarrar estas células. En el transcurso del siglo pasado se hicieron repetidos estudios que documentaron la presencia de células metastásicas dentro de pequeños coágulos llamados microtrombos, microembolia o tromboembolia; éstos se forman porque grupos de plaquetas y células de la sangre roja se adsorben sobre la superficie de células tumorales individuales o grupos de las mismas dentro de los vasos sanguíneos.

Algunos modelos experimentales de metástasis en ratones indican que la supervivencia de una célula metastásica es mucho mayor si se forman estos microtrombos ya que usan las plaquetas como vehículos protectores dentro del torrente sanguíneo (46-48). Si las células cancerosas sobreviven y consiguen acceder a vasos sanguíneos más grandes del sistema venoso, éstas viajarán dentro del torrente sanguíneo hacia el corazón y quedarán atrapadas con mucha probabilidad (debido a la diferencia de tamaño) en el primer grupo de capilares que encuentren después de su paso por el corazón; en la mayoría de los casos se quedan en los lechos capilares de los pulmones donde crean nuevas colonias. Sin embargo, se pueden encontrar metástasis en cualquier parte del cuerpo, lo que indica que con frecuencia, las células tumorales logran escapar de los pulmones y viajan hacia otras partes del cuerpo.

Una vez que las células cancerosas quedan atrapadas en los vasos sanguíneos o capilares de algún tejido, éstas deben salir del interior de estos vasos y penetrar en el tejido circundante. A este paso de la cascada metastásica se le conoce con el nombre de "extravasación". Este proceso depende de interacciones complejas entre las células cancerosas y las paredes de los vasos en los que estas células han quedado atrapadas. El microtrombo atrapado en el capilar desplaza a las células endoteliales de la pared del capilar, por lo que entra en contacto directo con la parte interior de la membrana basal del mismo; posteriormente las células de la sangre encargadas de disolver los coágulos

disuelven el microtrombo y las células cancerosas proliferan dentro del vaso sanguíneo. Después de un tiempo, algunas células cancerosas rompen la membrana basal del capilar e invaden el tejido adyacente (49).

Finalmente, la colonización en el nuevo tejido representa el paso más complejo y el mayor desafío para las células cancerosas, dentro de la cascada metastásica. Esto es debido a que el entorno del nuevo tejido no les proporciona los mismos factores de crecimiento y supervivencia que tenían sus células progenitoras en el tumor primario. Por tanto, es muy probable que estas células cancerígenas mueran o en el mejor de los casos para ellas, se quedan como células individuales o forman pequeños grupos de células, llamadas micrometástasis, que solamente pueden ser detectadas microscópicamente y rara vez aumentan su tamaño. Pueden pasar largos periodos de tiempo hasta que algún pequeño grupo de células cancerosas adquieren la capacidad de crecer y formar una masa detectable clínicamente (39). Las células que lo consiguen proliferan en su nuevo asentamiento, producen neoangiogénesis para asegurarse suficiente vascularización y forman un nuevo tumor, llamado tumor secundario (1).

Realmente, el proceso de invasión y metástasis es un proceso muy ineficiente y la probabilidad de que una célula individualmente complete todos los pasos con éxito es muy baja, por lo tanto en este proceso se seleccionan las células tumorales más agresivas (39).

2.- PROTEÍNAS IMPLICADAS EN METÁSTASIS Y ANGIOGÉNESIS

2.1.- Proteínas implicadas en angiogénesis.

La angiogénesis, como se ha descrito anteriormente, es un proceso complejo que está regulado por un delicado equilibrio entre diferentes moléculas proangiogénicas y antiangiogénicas (Fig. 3). Estas moléculas son liberadas al medio por distintas células tumorales tales como células endoteliales, macrófagos, plaquetas, mastocitos y componentes del estroma. Entre las moléculas que inducen angiogénesis o proangiogénicas se pueden destacar los

factores de crecimiento vasculares endoteliales (VEGF), factores de crecimiento para fibroblastos básico (b-FGF), factores de crecimiento de la epidermis (EGF), interleuquina 8 (IL-8) y factores de crecimiento transformantes β (TGF- β). Además, las plaquetas son capaces de liberar un gran número de moléculas proangiogénicas tales como VEGF-A, factores de crecimiento derivados de plaquetas (PDGF), bFGF, EGF, angiopoyetina-1 y heparanasa (25, 50, 51).

- Familia de VEGF

Los factores de crecimiento vasculares endoteliales (VEGF) son una de las citoquinas con mayor potencial angiogénico y fueron descubiertas y clonadas por primera vez por Napoleone Ferrara en 1989 (52). Los VEGF están implicados en la supervivencia de células endoteliales durante la formación de nuevos vasos sanguíneos. Además inducen la expresión de proteasas que producen la degradación de la matriz extracelular (ECM), proceso necesario para la migración de estas células endoteliales. Estas moléculas son también potentes estimuladores de la permeabilidad vascular permitiendo el posicionamiento de las fibrinas dentro de la ECM, que posteriormente servirán como peldaños para la formación de una nueva red de vasos sanguíneos (53).

La familia de VEGF está formada por seis factores de crecimiento, VEGF-A, VEGF-B, VEGF-C, VEGF-D, VEGF-E, el factor de crecimiento de la placenta (PlGF) y tres receptores, VEGFR-1 (Flt-1), VEGFR-2 (KDR/Flk-1) y VEGFR-3 (Flt-4) (54, 55). El mecanismo de acción de cada uno de los VEGF está descrito en la Tabla 2. El VEGF-A está implicado principalmente en angiogénesis mientras que VEGF-C y VEGF-D están implicados en linfangiogénesis. Se ha demostrado recientemente que VEGF-B está más involucrado en la supervivencia de los vasos sanguíneos que en el proceso de angiogénesis (56).

Los VEGFR son receptores de tirosina quinasa transmembranales que se caracterizan por tener siete dominios de inmunoglobulinas en su región

extracelular, un dominio transmembranal y un dominio intracelular de tirosina quinasa (57).

Tabla 2. Mecanismo de acción de los VEGF (58).

Tipo de VEGF	Mecanismo de acción
VEGF-A	Angiogénesis fisiológica y tumoral
VEGF-B	Vasculogénesis y activación de enzimas que invaden las células endoteliales
VEGF-C	Linfoangiogénesis y angiogénesis tumoral
VEGF-D	Angiogénesis
VEGF-E	Mitosis de las células endoteliales y angiogénesis

La unión de los VEGF a los receptores provoca la activación de los dominios tirosina quinasa iniciándose la cascada de señales que desencadena la angiogénesis. VEGFR-1 y VEGFR-2 se expresan en células endoteliales y son dos importantes mediadores de angiogénesis mientras que VEGFR-3 se expresa en células endoteliales linfáticas y está implicado en linfoangiogénesis (25, 51). VEGFR-1 es un receptor de VEGF-A y también tiene la capacidad de unirse a VEGF-B y PlGF (59-61). Este receptor es muy importante para el desarrollo de angiogénesis (62, 63), sin embargo su función específica no ha sido bien definida. VEGFR-2 es el regulador de angiogénesis más importante y media la mayoría de los efectos angiogénicos de VEGF-A, incluyendo la permeabilidad vascular, la proliferación, migración y supervivencia de las células endoteliales (64-66). El papel de VEGFR-2 en oncogénesis está bien definido y se ha demostrado que la inhibición de éste *in vivo* frena el crecimiento del tumor en ratones (67) (Fig. 5).

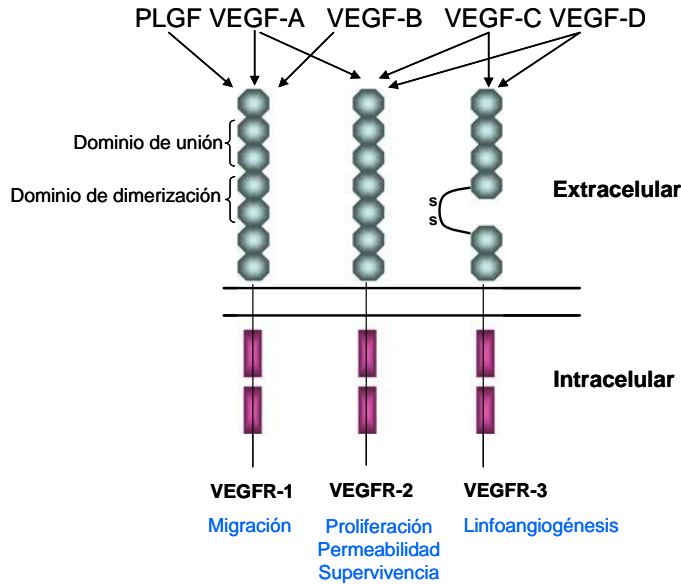


Figura 5. Isoformas de VEGF y posibles sitios de unión a VEGFR-1, VEGFR-2, y VEGFR-3 (25).

○ Angiopoyetinas

Las angiopoyetinas (Ang-1–Ang-4) están implicadas en la generación de nuevos vasos sanguíneos en muchos tipos de tumores y juegan un papel importante en el mantenimiento, crecimiento y estabilización de los vasos sanguíneos (68, 69). Las Ang-1 y Ang-2 son las angiopoyetinas mejor caracterizadas de las cuatro y ambas son citoquinas que ejercen su función biológica mediante la unión al receptor Tie-2 (70). Ang-1 promueve la supervivencia de las células endoteliales y estabilización de los vasos sanguíneos mediante la contracción de pericitos. En contraste, Ang-2, induce desestabilización vascular a altas concentraciones y se expresa en lugares de remodelación vascular lo que conlleva la inducción de angiogenesis e invasión de tumores. En presencia de VEGF, esta desestabilización induce una respuesta angiogénica. La expresión de Ang-1 está restringida normalmente a células tumorales mientras que Ang-2 se ha encontrado también en microvasos sanguíneos. En muchos casos, el nivel de Ang-1 que expresan las células

tumorales permanece inalterado mientras que Ang-2 incrementa su expresión particularmente en tumores con alto nivel de vascularización (25).

- Factores de crecimiento para fibroblastos (FGFs).

Los factores de crecimiento para fibroblastos constituyen una familia de más de veinte polipéptidos que están implicados en una gran variedad de procesos biológicos que incluyen la proliferación de células endoteliales, diferenciación celular, migración y angiogénesis (71-73). Los factores de crecimiento para fibroblastos ácido y básico (a-FGF o FGF-1 y b-FGF o FGF-2 respectivamente) son los miembros prototipo de esta familia (74).

En la actualidad se sabe que su mecanismo de acción involucra un sistema receptor dual que implica en primer lugar la interacción extracelular del FGF con un fragmento oligosacárido tipo heparina o heparán sulfato (o receptor de baja afinidad) que proporciona a la proteína estabilidad frente a agentes desnaturalizantes (75-78) y, mediante un mecanismo que aún no se conoce, con un receptor proteico transmembranal con actividad de tirosina quinasa denominado “receptor de alta afinidad” (FGFR). Una vez formado el complejo ternario, se transmite la señal a través de la membrana y se inicia la respuesta celular (79).

- Otros factores pro-angiogénicos

El mecanismo de acción de los factores de crecimientos derivados de plaquetas (PDGF) no es muy conocido hasta el momento. Sin embargo, se sabe que estimulan la angiogénesis *in vivo* y algunos experimentos con ratones sugieren que estos PDGF están implicados en el reclutamiento de pericitos necesarios para el desarrollo de capilares en tumores y además que intervienen en el proceso de envejecimiento vascular (80, 81). La señalización de paracrinas PDGF está implicada en cánceres epiteliales (80-83) mientras que la señalización de autocrinas PDGF es característica de tumores como leucemias (84), gliomas (85) y sarcomas (86).

Los factores de crecimiento transformantes β (TGF- β) están considerados como factores pro- y anti-angiogénicos. Durante la fase inicial de la tumorigénesis, TGF- β inhibe el crecimiento y desarrollo del tumor ya que frena la proliferación celular e induce apoptosis. En fases finales, el tumor llega a ser resistente a la actividad supresora de TGF- β , iniciando su papel pro-angiogénico (87).

Hay otros factores como el factor de necrosis tumoral α (TNF- α), el factor de crecimiento transformante α (TGF- α), el factor de crecimiento de la epidermis (EGF) y los factores estimulantes de colonias (CSFs) entre otros que también están involucrados en el proceso de angiogénesis (88, 89).

- Factores anti-angiogénicos

Sin embargo, la presencia de factores angiogénicos no es suficiente para iniciar el proceso de nueva vascularización, ya que también hay que tener en cuenta el efecto de un gran número de moléculas que inhiben la angiogénesis. Entre las moléculas antiangiogénicas se pueden destacar la trombospondina-1 (TSP-1) y la endostatina/angiostatina entre otras.

- Trombospondina-1

TSP-1 es una glicoproteína trimérica de 450 kDa que se descubrió en 1971. Esta proteína se encuentra almacenada en altas concentraciones en los gránulos de las plaquetas y constituye el 25% de las proteínas que segregan estas células (90), participando en la agregación de las mismas y en la formación de coágulos (91). Además, otros tipos de células producen TSP-1, muchas de ellas se encuentran en la matriz extracelular. Se han descrito gran variedad de funciones para esta proteína, entre ellas efectos de adhesión celular, migración, proliferación y angiogénesis. CD36 es un receptor celular muy importante para TSP-1 en el endotelio microvascular y necesario para su función antiangiogénica (92). La interacción entre TSP-1 y CD36 activa una secuencia de eventos intracelulares que desencadenan la apoptosis de las células

endoteliales. Además de los efectos angiogénicos mediados por la interacción con este receptor, TSP-1 puede inhibir la angiogénesis porque es responsable de modular los niveles de ciertas metaloproteasas MMP-2/9 responsables de la degradación de la matriz extracelular, proceso clave en angiogénesis. Sin embargo el papel de TSP-1 en angiogénesis genera ciertas controversias. Hay algunas enzimas como la trombina, plamina y tripsina que generan dos fragmentos de TSP-1 de 25 y 140 kDa. Aunque su actividad antiangiogénica mediada por la interacción con el receptor CD36 está localizada en el fragmento de 140 kDa, se ha encontrado un efecto proangiogénico en el fragmento que une a heparina de 25 kDa (25).

- Otros factores anti-angiogénicos

La angiostatina es un fragmento plasminógeno de 38 kDa que inhibe la producción de nuevos vasos sanguíneos en tumores e induce apoptosis en las células endoteliales. Además, inhibe la mejora de la matriz extracelular y la activación del plasminógeno t-Pa, lo que permite la migración y la invasión de las células endoteliales, proceso clave en angiogénesis (93). La endostatina es una citoquina que se genera por la rotura de un fragmento de 20 kDa de colágeno XVIII, un proteoglicano que se encuentra en las paredes de los vasos sanguíneos y las membranas basales. Actúa inhibiendo la migración de las células endoteliales e induce apoptosis y el fin del ciclo celular (58). La vasostatina es el extremo N-terminal de calreticulina y constituye un potente inhibidor de angiogénesis junto con PF-4 que se utilizan como un único péptido recombinante en el desarrollo de nuevos medicamentos (94). La osteopontina es una glicoproteína fosforilada implicada en angiogénesis, crecimiento de tumores, invasión, metástasis y cicatrización de heridas y su sobreexpresión se ha asociado a los cánceres de mama más agresivos (95). Los tejidos inhibidores de metaloproteasas (TIMP) comprenden una familia de 4 proteasas: TIMP-1, -2, -3, -4. Estos inhiben las metaloproteasas y juegan un papel importante en supresión de metástasis. Mutaciones en el gen TIMP-3 causan la distrofia de

fondo de ojo que se caracteriza por la neovascularización y pérdida de visión (96).

2.2.- Quimioquinas

Otra importante familia de proteínas que merecen ser mencionadas son las quimioquinas. Esta familia consta de aproximadamente 50 citoquinas quimioestáticas de bajo peso molecular (8-14 kDa) que están implicadas en diferentes procesos biológicos tales como la migración de leucocitos, embriogénesis, angiogénesis, hematopoyesis, aterosclerosis, crecimiento de tumores y metástasis e infección de HIV (97, 98). Todas ellas se caracterizan por tener una estructura similar basada en la conservación de los residuos de cisteína así como por su capacidad de unirse a receptores llamados GPCR (receptores acoplados a proteínas G) (99). Se pueden clasificar en 4 subgrupos dependiendo de la posición de los residuos de cisteína que se conservan en la parte amino terminal de estas proteínas: CXC, CC, CX3C y C, siendo los CC y CXC los más comunes (100). La posibilidad de que varias quimioquinas puedan unirse al mismo receptor y una quimioquina pueda unirse a distintos receptores, crea múltiples combinaciones dando lugar a funciones biológicas muy diversas (101).

- Quimioquinas que regulan angiogénesis:

Las quimioquinas pueden influir de una manera indirecta sobre el crecimiento de tumores y por lo tanto promover la angiogénesis (Fig. 6). Las moléculas responsables de esta función son proteínas con sitio de unión a heparina que pertenecen a la subfamilia de quimioquinas CXC. Estas proteínas reciben este nombre porque mantienen las cuatro cisteínas conservadas, estando las dos primeras separadas por otro aminoácido no conservado (97, 102). La actividad biológica de esta familia depende de un segundo dominio estructural y su unión a receptores específicos. Los extremos N-terminal de varias quimioquinas CXC contienen un motivo altamente conservado formado por Glu-Leu-Arg (motivo ELR o ELR+) que precede a la primera cisteína que se

conserva en este tipo de proteínas. Las quimioquinas CXC ELR+ son potentes promotores de angiogénesis, mientras que las quimioquinas CXC que no tienen este motivo ELR (quimioquinas CXC ELR-) son potentes inhibidores de angiogénesis (103).

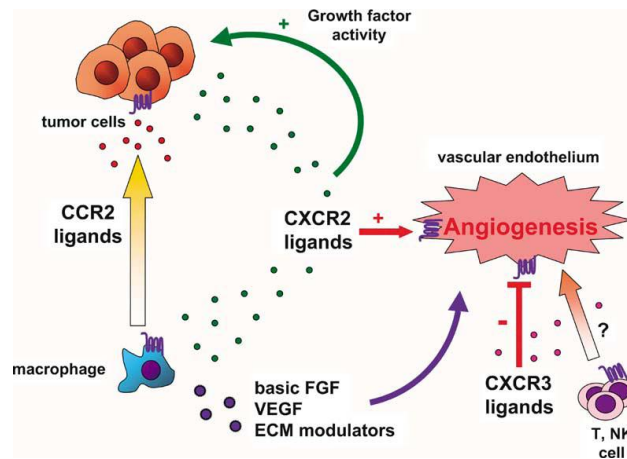


Figura 6. Regulación del crecimiento tumoral y angiogénesis por quimioquinas (101)

Las quimioquinas CXC ELR+ interactúan con receptores CXCR2 (por ejemplo CXCL1, CXCL2, CXCL3, CXCL8) que por un lado liberan al medio macrófagos que se infiltran en las células cancerígenas contribuyendo al crecimiento del tumor y por otro lado, producen factores de crecimiento como bFGF y VEGF que inducen angiogénesis. A la inversa, las quimioquinas CXC ELR- se unen a receptores CXCR3 (CXCL9, CXCL10, CXCL11 y CCL21 de ratón) que interactúan directamente con el endotelio reclutando células T o NK que destruyen la vascularización.

- Implicación de quimioquinas en metástasis

La invasión de las células tumorales a través de la matriz extracelular y la entrada dentro de la circulación sanguínea implica el desprendimiento celular, la movilidad hacia la matriz extracelular y la degradación de las membranas basales que son procesos que dependen de una gran variedad de enzimas que

incluyen metaloproteasas (MMPs) y serina y cisteína proteinasas. Las quimioquinas activan muchas células tumorales que producen estas proteinasas. La expresión de CXCL8 en células de melanomas humanos induce la activación transcripcional de la expresión del gen que codifica para MMP-2 (un tipo de colagenasa IV) aumentando la actividad colagenasa dentro del tumor y por tanto incrementando el poder de invasión de estas células (104). En cáncer de próstata, la sobreexpresión de CXCL8 induce la expresión de MMP-9 aumentando la capacidad de invasión y metástasis de estas células cancerosas en ratones (105). Las quimioquinas CC también juegan un papel importante en el paso de reclutamiento de fagocitos que a su vez liberan proteinasas que promueven la invasión de células cancerígenas.

Otro grupo importante de receptores de estas quimioquinas son los CXCR4 que se encuentran sobre la superficie de células de cáncer de mama. Los ligandos de este receptor se pueden encontrar en células de cáncer de pulmón, hígado y huesos, incrementándose el riesgo de metástasis en estos órganos. Además se ha encontrado que CXCR4 es vital para la metástasis de los nodos linfáticos de melanomas y carcinomas de mama (106, 107). Estos resultados entre otros, apoyan la hipótesis de que ciertas quimioquinas y sus receptores dirigen el destino de distintas células tumorales metastásicas.

2.3.- Proteínas que intervienen en el proceso de invasión y en la transición epitelial-mesenquimal (EMT).

La organización de los distintos niveles dentro de una célula epitelial es incompatible con la motilidad y la invasividad que caracterizan a las células cancerígenas. Para adquirir estas características, las células de un carcinoma sufren una serie de alteraciones, pudiendo cambiar su fenotipo epitelial a mesenquimal (proceso conocido como EMT). Por ejemplo, en muchos tumores primarios con propiedades invasivas, se reduce la adhesión intercelular, debido muy a menudo, a la pérdida de la expresión de E-cadherina, una proteína que media la interacción y la adhesión célula-célula. En una célula sana, la cola citoplasmática de la E-cadherina se mantiene unida a la “actina de citoesqueleto”

(una de las actinas que se encargan de mantener las uniones entre células). Esta unión se mantiene a través de α -catenina y β -catenina (1, 108). Cuando la expresión de E-cadherina disminuye, gran cantidad de β -catenina queda libre en el citosol. Ésta puede posteriormente migrar al núcleo celular dónde puede unirse a factores de transcripción como Tcf/LEF que promueven la EMT (1).

Además, durante EMT, se produce la expresión de vimentina, un componente filamentoso de la célula mesenquimal citoesqueleto de actina. Por otro lado, las células epiteliales bajo este proceso EMT, segregan fibronectina, una proteína de la matriz extracelular que normalmente es fabricada por células mesenquimales como por ejemplo los fibroblastos. Al mismo tiempo, se segrega un marcador de fibroblastos, la N-cadherina en lugar de E-cadherina. Todas las alteraciones que se producen durante EMT vienen descritas en la Tabla 3.

Tabla 3. Alteraciones producidas en las células epiteliales de un carcinoma durante EMT.

Pérdida de:
Expresión de citoqueratina
Expresión de E-cadherina
Polaridad de la célula epitelial
Adquisición de:
Forma de fibroblastos
Motilidad
Invasividad
Programa de expresión de genes mesenquimales
Expresión de N-cadherina
Secreción de proteasas como MMP-2 y MMP-9
Expresión de vimentina
Secreción de fibronectina
Expresión de receptores de PDGF
Expresión de integrina α v β 6

2.4.- Papel que juegan las proteínas CD44 en el proceso de metástasis

Las proteínas CD44 son una clase de glicoproteínas de membrana que se encuentran sobre la superficie de varias células y juegan un papel vital en la adhesión celular, migración y en la regulación del tráfico de linfocitos (109). Las CD44 actúan como receptores de ácido hialurónico, un glicosaminoglicano que forma parte de la matriz extracelular. Además, une también a otras moléculas como colágeno, osteopontina, laminina, MMPs (metaloproteasas de la matriz extracelular), fibronectina, condroitina sulfato y heparán sulfato. Estas proteínas están formadas por 20 exones de los cuales 10 son variables dando lugar a distintas isoformas de la proteína (109, 110). Las proteínas CD44 interactúan con el citoesqueleto de actina y el reordenamiento que se produce en esta actina como consecuencia de la interacción promueve la invasión y la migración. Se ha encontrado que varias isoformas de la proteína, especialmente aquellas que contienen el exón V6, están presentes en neoplasias. Además se ha descubierto que hay una conexión entre la presencia de CD44v6+ y la habilidad de las células cancerosas para producir metástasis (109). Por lo tanto la presencia y la sobreexpresión de algunas isoformas de estas proteínas se han usado como indicadores del grado de desarrollo y agresividad de un cáncer (109-111).

2.5.- Metaloproteasas de la matriz extracelular.

Las células de un tumor maligno tienen que adquirir la capacidad de atravesar la matriz extracelular (ECM) y varias capas de tejidos para poder invadir y empezar el proceso de metástasis. Para ello, estas células pueden activar enzimas proteolíticas que degradan la matriz extracelular. Un grupo muy importante dentro de este tipo de proteínas son las metaloproteasas de la matriz extracelular (MMPs) que degradan no solo proteínas como fibronectinas y vitronectinas, sino también otros componentes de la ECM como glicosaminoglicanos, colágenos y lamininas. Además, también están implicadas en el corte de receptores de membrana incluyendo transglutaminasas tisulares,

CD44, pro- α_v integrina, β -glicanos y lipoproteínas de baja densidad relacionadas con receptores proteicos.

Las metaloproteasas de la matriz extracelular constituyen una familia de proteínas perteneciente al grupo de las endopeptidasas dependientes de Zn^{2+} y la mayor parte de ellas son expresadas por células del estroma como macrófagos, mastocitos y fibroblastos. La capacidad de las MMPs para degradar distintos componentes de la matriz extracelular (ECM) es ampliamente conocida, de ahí su implicación en procesos que requieren remodelado tisular, como por ejemplo, el desarrollo embrionario, la cicatrización de heridas o la transición de cartílago a hueso que tiene lugar durante la osificación. Sin embargo, estas proteínas están sobreexpresadas en condiciones patológicas como la periodontitis, la artritis reumatoide y durante la invasión celular y metástasis de células tumorales (112).

La estructura básica de las MMPs presenta una serie de dominios característicos: un pre-dominio necesario para su desplazamiento intracelular hasta la superficie de la célula; un pro-péptido que mantiene al enzima en su forma latente; un dominio catalítico donde tiene lugar el corte proteolítico de los sustratos; una región rica en prolinas que aporta flexibilidad a la estructura denominada "*linker*" y un dominio homólogo de la vitronectina o de la hemopexina (PEX) que media en la especificidad del sustrato y las interacciones con inhibidores endógenos (113). Las MMPs solubles son siempre secretadas como una pro-enzima inactiva dentro de la ECM y posteriormente son activadas por otras serina proteasas como plasmina y furina, mientras que algunas de las MMPs son activadas por miembros de su propia familia (39).

Además de la implicación de estas proteínas en invasión y metástasis por la degradación de la ECM, las MMPs afectan a los procesos de proliferación y angiogénesis ya que están implicadas en la activación/desactivación de varios factores de crecimiento como b-FGF, VEGF y TGF. El llamado tejido inhibidor de metaloproteinasas TIMPs juega un papel importante en la regulación de las MMPs y el balance entre TIMPs y MMPs es de crucial relevancia en los procesos de invasión y metástasis. Hoy en día se conocen más de 20 MMPs diferentes que se pueden clasificar en: colagenasas (MMP-1/8/13/18),

gelatinasas (MMP-2/9), estromalinas (MMP-3/10/11), matrilinas (MMP-7/26), MMP que unen a la membrana (MMP-14/15/16/17/24/25) y otras MMPs que no están incluidas en ninguno de estos grupos (114). La función de las colagenasas es la degradación de la triple hélice de colágeno; las gelatinasas digieren el colágeno desnaturalizado y las gelatinas; las estromalinas activan pro-MMPs para el corte proteolítico y digieren otros componentes de la ECM; las matrilinas procesan moléculas de la superficie celular como el ligando de CD95 (CD178) y E-cadherina; las MMPs que unen a la membrana activan la enzima pro-MMP2 y otras moléculas de la ECM. Además la desregulación de las MMPs puede promover la conversión de una célula sana en una célula de un tumor maligno (1).

3.-HEPARANASA

3.1.- Proteoglicanos de heparán sulfato (HSPGs).

Los proteoglicanos de heparán sulfato (HSPGs) son componentes muy abundantes de los vasos sanguíneos, concentrándose mayoritariamente en el interior de los vasos sanguíneos largos, mientras que en capilares se encuentran principalmente en la membrana basal subendotelial (BM). Estos HSPGs se pueden definir como macromoléculas asociadas con la superficie celular y la matriz extracelular de un gran número de células de tejidos tanto de vertebrados como invertebrados (115, 116) y su estructura básica es un núcleo proteico al que se unen varias cadenas de heparán sulfato (HS). Esta unión tiene lugar a través de un tetrasacárido con una secuencia xilosa-galactosa-galactosa-ácido urónico que se une covalentemente a la proteína central mediante un enlace O-glicosídico (116-120). El heparán sulfato está formado por la repetición de unidades disacáridicas compuestas de ácido urónico y D-glucosamina modificadas en distintas posiciones con distinto patrón de sulfatación, epimerización y N-acetilación (116-119). Estos HSPGs interactúan con fibronectina, laminina, colágeno, citoquinas y factores de crecimiento tales como bFGF, VEGF, KGF, HGF y TGFβ ayudando a mantener el ensamblado y la

integridad de la ECM. Debido a la multitud de funciones de los HSPGs en la fisiología celular y tisular, su degradación altera la integridad de los tejidos y proporciona un mecanismo por el que las células pueden separarse unas de otras y provocar cambios en el microambiente de un tejido. Por tanto la degradación enzimática del HS y como consecuencia, de la matriz extracelular, tiene un papel clave en distintos procesos fisiológicos normales como morfogénesis, embarazo, cicatrización de heridas, crecimiento, migración y diferenciación celular, así como en algunos procesos patológicos como por ejemplo, angiogénesis, inflamación, invasión de las células tumorales y metástasis (121-123). Durante mucho tiempo se ha investigado intensivamente sobre las enzimas capaces de degradar componentes proteicos de la matriz extracelular (124-127), sin embargo, se ha prestado menos atención a enzimas que son capaces de degradar las cadenas de glicosaminoglicanos como la heparanasa, debido seguramente a la falta de sondas moleculares apropiadas para estudiar su papel en la progresión de tumores.

3.2.- Antecedentes de Heparanasa

En 1975, Höök *et al.* observaron por primera vez en tejidos de hígado de rata la actividad de una endoglicosidasa que degradaba el HS (128). Años más tarde se relacionó la degradación del HS con las propiedades metastásicas e invasivas de las células de melanoma de ratón (129). Con el objetivo de caracterizar esta endoglicosidasa, en 1984 Nakajima *et al.* analizaron los productos de degradación del HS concluyendo que la endoglicosidasa que degradaba el HS era una endoglucuronidasa, que fue llamada “heparanasa” (130).

La heparanasa humana (heparanasa-1) fue clonada por primera vez en 1999 de manera independiente por los grupos de Vlodaysky (121) y Parish (131) a partir de líneas celulares de la placenta y hepatoma humanos. Posteriormente, otros grupos realizaron la clonación del mismo gen a partir de células humanas (132, 133) y otras especies (pollo, rata, ratón, bovino y topo) (134-137), siendo la homología con la heparanasa humana mayor en el caso de la enzima de

mamíferos y el topo (138). La heparanasa-1 es una endo- β -D-glucosidasa capaz de cortar específicamente las cadenas de HS en un número limitado de sitios dando lugar a fragmentos de HS de apreciable tamaño (~5-7 kDa). Además es posible diferenciarla de otra enzima que degrada HS en bacterias, la heparitinasa, que despolimeriza el HS por cortes inespecíficos. McKenzie *et al.* publicaron en el año 2000 la clonación de cADN humano con un 40% de homología con heparanasa-1. La expresión de este clon (llamado heparanasa-2), resultó en una proteína sin actividad catalítica y cuya función es aún desconocida. A pesar de la existencia de distintas heparanasas de mamíferos (topo y humanos) encontradas recientemente (136-139), heparanasa-1 es la única enzima capaz de degradar el HS.

3.3.- *Propiedades moleculares y bioquímicas de la enzima.*

La heparanasa humana se expresa inicialmente como un polipéptido de 543 aminoácidos con un peso molecular de 61,2 kDa (122, 140, 141). La heparanasa activa purificada de placenta, plaquetas y varias líneas celulares no contiene el extremo N-terminal de 156 aminoácidos lo que sugiere una proteólisis post-transcripcional de este polipéptido (142). Se ha demostrado que la forma activa de heparanasa es un heterodímero formado por una subunidad de 50 kDa (Lys¹⁵⁸-Ile⁵⁴³) asociada no covalentemente con otra subunidad de 8 kDa (Gln³⁶-Gln¹⁰⁹). El péptido de 6 kDa que contiene la secuencia S¹¹⁰-Q¹⁵⁷ se elimina por proteólisis (143-145). Esta proteína también contiene una secuencia N-terminal de 35 aminoácidos (Met¹-Ala³⁵) y un dominio C-terminal hidrofóbico transmembrana muy próximo al extremo carboxilo terminal (Pro⁵¹⁵-Ile⁵³⁴) (146). La naturaleza precisa de las proteasas responsables de la generación de las dos subunidades es un área que está bajo intensa investigación. Estudios de mutagénesis indican que el voluminoso e hidrofóbico aminoácido Y¹⁵⁶, es esencial para el correcto procesamiento de heparanasa y la activación de la proteasa catepsina L que corta entre los aminoácidos Q¹⁵⁷ y K¹⁵⁸; sin embargo mutaciones realizadas en el sitio E¹⁰⁹-S¹¹⁰ no han mostrado efectos relevantes, lo

que sugiere que esta región contiene un sitio de reconocimiento para otra proteasa (147). Por otro lado y basándose en la predicción de secuencias, es conocido que la unidad de 50 kDa contiene 6 sitios putativos de N-glicosidación. El análisis de heparanasa purificada ha confirmado la existencia de estos sitios de N-glicosidación y aunque estos no son requeridos para la actividad enzimática, la secreción de heparanasa está regulada por glicosidación (148). También se ha demostrado que existen dos puentes disulfuro Cys¹²⁷-Cys¹⁷⁹ y Cys⁴³⁷-Cys⁵⁴² y S-cisteinilación de la Cys²¹¹. Mientras que el puente Cys¹²⁷-Cys¹⁷⁹ y el residuo S-cisteinilado Cys²¹¹ ejercen un efecto pequeño en la función de heparanasa, el puente Cys⁴³⁷-Cys⁵⁴² es necesario para la secreción y activación de esta enzima (149).

La heparanasa se ha relacionado con miembros de una familia de glicosil hidrosilasas A (GH-A). El alineamiento de secuencias junto con predicciones de estructura secundaria indican que la heparanasa es homóloga a las familias 10, 39 y 51 de las GH-A, especialmente en la región del sitio catalítico (150). Las enzimas pertenecientes a esta familia usan un mecanismo general de catálisis para hidrolizar enlaces glucosídicos. El mecanismo requiere dos residuos críticos, un donador de protones y un nucleófilo, los cuales parecen estar conservados en la heparanasa, siendo estos el Glu²²⁵ y el Glu³⁴³, respectivamente. La mutagénesis dirigida de esos dos ácidos glutámicos inhibe completamente la actividad de la enzima indicando que la heparanasa usa un mecanismo de catálisis característico de GH-A (150). Como el sitio catalítico de la enzima se encuentra dentro de la subunidad de 50 kDa, y se ha demostrado que la formación del heterodímero es necesaria para la actividad de la enzima (144) se puede suponer que la subunidad de 8 kDa provoca un cambio conformacional dentro del sitio catalítico que facilita la catálisis. Además se han identificado dentro de la subunidad de 50 kDa, dos regiones de este dominio que unen al sustrato HS, necesarias para la actividad catalítica de la enzima. Estas dos regiones son Lys¹⁵⁸-Asn¹⁶² y Gln²⁷⁰-Lys²⁸⁰ (151).

3.4.- Regulación.

3.4.1.-Metilación del promotor.

Los procesos epigenéticos marcados por hipermetilación de la región del promotor, junto con las alteraciones genéticas que gobiernan la transformación celular, contribuyen significativamente en la transcripción génica y en la progresión del cáncer. Varios estudios han demostrado convincentemente que la metilación del promotor juega un papel importante en la transcripción de heparanasa. El examen de varias líneas celulares derivadas de tumores ha revelado que las células que muestran expresión o actividad de heparanasa albergan al menos un alelo no metilado, mientras que las células que no muestran actividad de la enzima tienen todos los alelos metilados. Cuando se tratan estas células con agentes desmetilantes como la 5-azacitidina recuperan la actividad de heparanasa lo que está acompañado de un aumento de la capacidad metastásica *in vivo* (152). Además se ha demostrado que este aumento es suprimido cuando se tratan estas células con laminarén sulfato, compuesto que inhibe la actividad de heparanasa y la diseminación metastásica (153), apoyando la implicación de heparanasa en metástasis. Estudios posteriores han revelado una correlación similar usando líneas celulares derivadas de cáncer de próstata y vejiga (154, 155). Todos estos datos indican que la metilación del promotor de heparanasa juega un papel importante en la regulación de su expresión. Este modo de regulación puede estar perturbado en tejidos tumorales malignos, donde el promotor de heparanasa está hipometilado, incrementándose por ello la expresión del gen de heparanasa y favoreciéndose la metástasis del tumor y la angiogénesis (138).

3.4.2.- Regulación por p53.

p53, como se explicó en apartados anteriores, es un factor de transcripción que regula una gran variedad de promotores celulares y además sufre alteraciones génicas en el 40-50% de los cánceres en humanos. En condiciones normales p53 suprime el gen de heparanasa por unión directa con su promotor,

inhibiéndose así la actividad de la enzima. Sin embargo, la inactivación de p53 en células cancerosas provoca la inducción de la expresión de heparanasa. En conclusión, la regulación de la expresión de heparanasa por p53 puede darnos una explicación de los altos niveles de heparanasa observados en una gran variedad de tejidos cancerosos, en comparación con los tejidos sanos (156).

3.4.3.- Regulación por mediadores de la inflamación.

Compuestos polianiónicos que inhiben la actividad enzimática de la heparanasa (por ejemplo la heparina), también inhiben respuestas inflamatorias. (157-159). Este efecto puede ser atribuido a la inhibición de heparanasa por la acción de los linfocitos T activados, considerados como la principal fuente celular de la heparanasa en el proceso de inflamación (160). Se ha examinado la implicación de heparanasa en inflamación utilizando un ensayo de hipersensibilidad retardada (DTH) en un modelo de ratón. Se ha comprobado que el factor-alfa (TNF- α) y el interferón-gamma (INT- γ) de la necrosis tumoral inducen la expresión de heparanasa en cultivos de células endoteliales (la expresión aumenta 2 o 3 veces) y por tanto aumentan su actividad enzimática. La administración de inhibidores de la actividad de heparanasa detiene de forma efectiva la respuesta inflamatoria de la DTH (161).

3.4.4.-Regulación por estrógeno y cáncer de mama.

Un ejemplo de regulador sistemático es la hormona estrógeno, una de las principales fuerzas conductoras en el cáncer de mama. Esta hormona induce transcripción del ARNm de heparanasa dentro del receptor positivo del estrógeno de células de cáncer de mama, pero no la induce en el receptor negativo. Se han identificado cuatro elementos putativos de respuesta a estrógeno en la región del promotor de heparanasa y se ha demostrado que la transcripción de un gen de luciferasa aumenta en presencia de un receptor de estrógeno positivo MCF-7 de células de carcinoma después de un tratamiento con estrógeno (162). El efecto que ejerce el estrógeno sobre los niveles de expresión del ARNm de heparanasa se elimina en presencia de un antiestrógeno

puro (ICI 182780), indicando que la ruta del receptor del estrógeno clásica está implicada en la activación transcripcional de heparanasa. Además, la expresión de heparanasa y por tanto su actividad en células endoteliales aumenta cuando se exponen a 17β -estradiol (163). El tamoxifen estimula la transcripción de heparanasa, induciendo la proliferación de la células del cáncer de mama en la matriz extracelular (164).

3.4.5.-Regulación por glucosa.

La expresión de heparanasa en células endoteliales está sobrerregulada cuando las células están expuestas a alta concentración de glucosa, mientras que en presencia simultánea de alta concentración de glucosa e insulina la expresión no se ve afectada. Se ha demostrado que la expresión de heparanasa glomerular aumenta en ratas y ratones que sufren diabetes inducida por estreptozotocina (165).

3.5.- Localización celular de heparanasa.

La heparanasa ha sido encontrada en los compartimentos endosomal y lisosomal de tejidos sanos y tumorales, así como en los gránulos terciarios de neutrófilos humanos. El proceso proteolítico de heparanasa tiene lugar intracelularmente dentro de los endosomas y los lisosomas tardíos (147, 166, 167). Se ha demostrado que el pH óptimo para que se produzca este proceso se encuentra en un rango entre 4 y 5 (168). El hecho de que sea necesario un microambiente con pH bajo para que este proceso sea eficiente, sugiere que esta reacción tiene lugar dentro los orgánulos ácidos intracelulares. Se ha comprobado que agentes químicos que aumentan el pH como la cloroquina y la bafilomicina A1 inhiben este proceso proteolítico de la enzima (169). Por tanto, está claro que la heparanasa tiene una importante función dentro de estos lisosomas y endosomas tardíos donde tiene lugar el procesamiento y reciclado del HS desde los HSPGs internos. Sin embargo la localización de heparanasa no está restringida a estos orgánulos, ya que también se ha encontrado en el núcleo celular, las regiones perinucleares y la membrana del plasma (135, 166, 170,

171). La heparanasa puede también desplazarse dentro de las células tumorales y reorganizarse en la superficie celular o secretarse donde pueda degradar el HS (Fig. 7). Algunos trabajos parecen indicar que la heparanasa es inicialmente excretada y posteriormente incorporada y procesada en los lisosomas, posiblemente por la acción de catepsinas (151, 169).

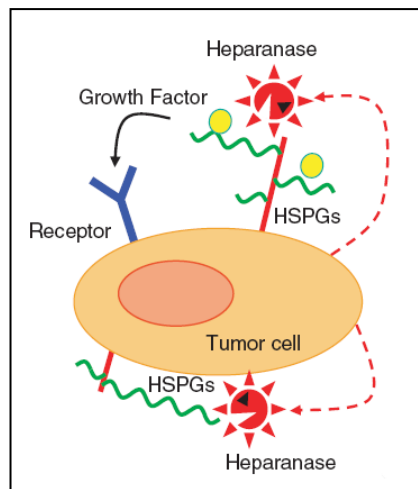


Figura 7. La heparanasa es producida por las células tumorales y posteriormente excretada. Una vez en el exterior, la proteína puede hidrolizar las cadenas de heparán sulfato (141).

3.6.- Función fisiológica en tejidos sanos.

La función de heparanasa en tejidos normales no es muy conocida. La expresión del ARNm de heparanasa en tejidos sanos está restringida en primer lugar a la placenta, órganos linfoides y queratinocitos (121, 123, 171). También se han detectado altos niveles de la enzima en las plaquetas de la sangre. Mientras que la expresión de heparanasa en tejidos humanos es limitada, esta enzima se expresa en gran cantidad en diversos tejidos de topo. En este animal, la heparanasa ha sido detectada en tejidos de hígado, corazón, cerebro y ojo y su papel en este mamífero subterráneo puede estar relacionado con su adaptación a vivir bajo la tierra, contribuyendo al incremento de la densidad de los vasos

sanguíneos que se ha encontrado en algunos de sus tejidos comparado con la densidad de estos vasos en mamíferos que viven en la superficie (136). Además la enzima parece estar implicada en la implantación y desarrollo embrionario (135, 172), reparación de heridas, movimiento del HS, remodelización de los tejidos, vigilancia inmunitaria y crecimiento del cabello (173, 174).

3.7.- Heparanasa y progresión del cáncer.

3.7.1.- Propiedades pro-angiogénicas.

La angiogénesis representa un proceso multicelular coordinado en el que están implicados una gran variedad de moléculas, incluyendo factores de crecimiento, componentes de la matriz extracelular, receptores de adhesión y enzimas que degradan la matriz extracelular como MMPs, serina y cisteína proteasas y endoglucosidasas. Los HSPGs interaccionan con estas biomoléculas y aseguran que permanezcan unidas a la superficie celular y a la matriz extracelular. Por lo tanto, la enzima heparanasa al degradar el HS libera todas estas biomoléculas que están involucradas en procesos de invasión, migración, adhesión, diferenciación y proliferación (118, 119, 175), es decir, en todos los procesos que se asocian con la angiogénesis (Fig. 8). La heparina y el heparán sulfato secuestran, estabilizan y protegen a los FGFs y VEGFs. Además estas moléculas pueden funcionar como co-receptores que estimulan la formación del complejo HS-FGF-FGFR, facilitando así la dimerización del receptor y la señalización (115, 176). La heparanasa nativa y recombinante expresada en células tumorales, con inflamación y plaquetas, liberan un complejo activo de bFGF y un fragmento de HS desde la matriz extracelular y las membranas basales (176). Se ha demostrado que la longitud de los fragmentos de HS requeridos para la activación del complejo FGF-receptor es similar a los fragmentos que libera la heparanasa (176, 177). Por otro lado, los niveles de expresión de esta enzima se han correlacionado directamente con la densidad de microvasos en una gran variedad de tumores malignos (138).

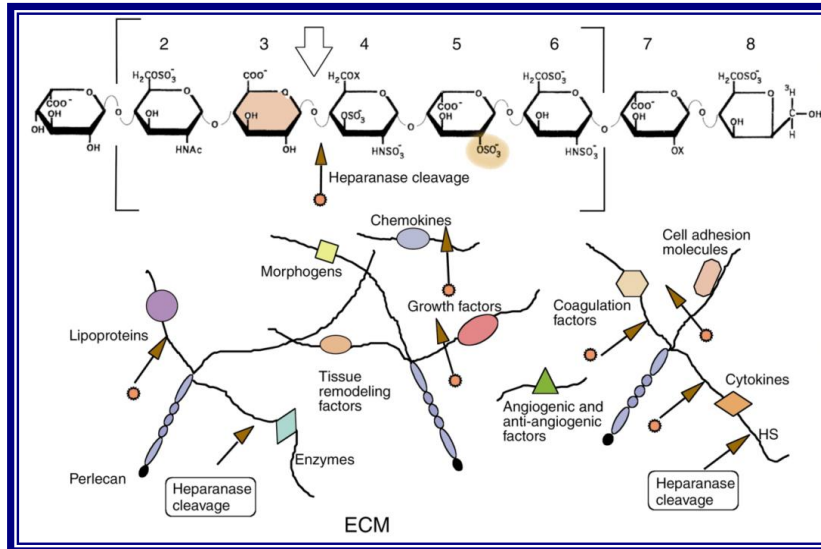


Figura 8. La heparanasa corta las cadenas de HS y libera gran variedad de moléculas bioactivas (122)

3.7.2.-Potencial metastásico.

Los procesos de invasión y metástasis en cáncer conllevan la degradación de los constituyentes de la matriz extracelular, por lo tanto todas las enzimas implicadas en este proceso promueven la extravasación, invasión, migración y todos aquellos procesos que conforman la cascada metastásica. En la Fig. 9 se puede observar cómo la heparanasa promueve los procesos de intra y extravasación celular. Los niveles de expresión de la heparanasa se incrementan de manera significativa en numerosos procesos cancerosos, tal como ha sido descrito en melanomas, tumores de mama, colon, pulmón, próstata, hígado, vejiga, intestino, ovario o páncreas, entre otros. Además existe una correlación entre los niveles de expresión y el grado de malignidad tumoral, y en casos como el del cáncer pancreático, la supervivencia del paciente se correlaciona significativamente con los niveles de expresión de la enzima (122, 178). La invasión tumoral y consiguiente metástasis requieren la degradación de componentes de la matriz extracelular y el HS es uno de los más prominentes.

De ahí que la expresión de heparanasa se correlacione con el potencial metastásico de las células tumorales. Además se ha observado que el suero de pacientes y animales portadores de tumores metastásicos muestra altos niveles de la enzima, apareciendo incluso en orina en casos de enfermedades metastásicas agresivas (122, 146).

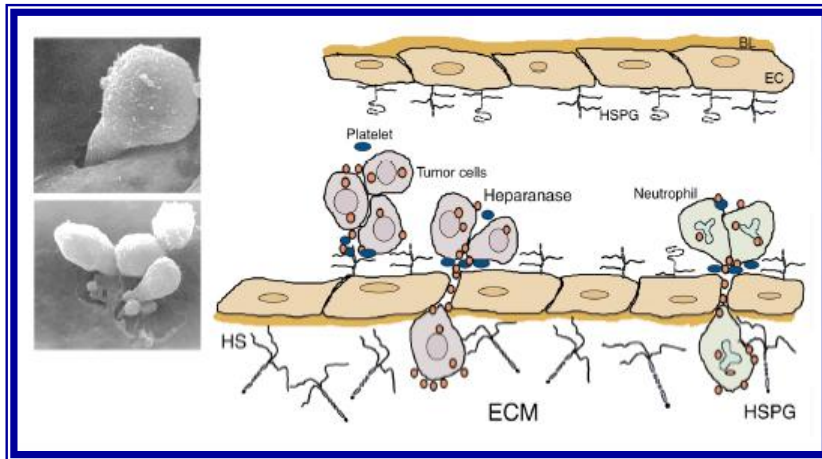


Figura 9. Heparanasa expresada por neutrófilos y células tumorales media la intra y extravasación dentro de la matriz extracelular por desestabilización de las paredes celulares de los vasos sanguíneos (122).

3.7.3.-Inflamación.

La heparanasa está implicada en inflamación, debido principalmente a su capacidad de degradar la ECM, ya que permite que las células con características inflamatorias alcancen su tejido objetivo. Así por ejemplo, las moléculas que inhiben heparanasa inhiben también la hipersensibilidad de las células T tipo retardada (DTH), la encefalomiелitis autoinmune experimental y la artritis (138). Además la heparanasa es producida en el endotelio donde se encuentra la inflamación asociada con las células DTH en respuesta a necrosis tumoral, factor-alfa e interferón-gamma (161).

3.7.4.-Actividades no enzimáticas de heparanasa

Heparanasa ejerce también actividades no enzimáticas que son independientes de su capacidad para degradar y alterar la matriz extracelular.

Por ejemplo la expresión de heparanasa inactiva en la superficie celular provoca la adhesión celular lo que refleja su implicación en dicho fenómeno. Por otro lado, la heparanasa inactiva mejora la señalización de Akt y media la fosforilación de ésta, estimula la invasión/migración de las células endoteliales dependientes de fosfatidilinositol 3-quinasa PI3K y además la heparanasa es promotor de la expresión de VEGF mediante la activación de Scr (138, 179). Recientemente se ha demostrado que la región de la proteína que contribuye con estas funciones no enzimáticas es el dominio de la proteína localizado en el extremo C-terminal de la misma. En la actualidad, se está llevando a cabo una intensa investigación para conocer si esta función de la proteína está mediada por algún receptor proteico que podría estar localizado en la superficie celular. Recientemente, experimentos de entrecruzamiento han revelado alguna información sobre la naturaleza y el peso molecular de este receptor, habiéndose encontrado dos complejos formados por proteínas de 130 y 170 kDa asociadas a heparanasa (180).

3.8.-Heparanasa en la terapia del cáncer

La implicación de heparanasa en la progresión del cáncer (181) y el hecho de que parece ser una enzima única en claro contraste con otras proteasas relacionadas con el mismo fenómeno, la convierte en una diana clave en tratamientos antitumorales. Aunque hace ya dos décadas que se identificó a esta enzima como una potencial diana terapéutica (129, 182), la investigación se ha visto ralentizada por tres factores: dificultades en obtener una proteína activa y pura, la falta de un método robusto para medir la actividad enzimática y la falta de la estructura tridimensional de la proteína y por tanto de un conocimiento estructural detallado del sitio catalítico. Aunque las dos primeras parecen haberse solventado, la falta de estructura 3D sigue siendo un obstáculo a la hora de buscar inhibidores de la enzima. Este problema se ha intentado solucionar con modelos estructurales basados en homología de secuencia con otras glicosidasas hidrolasas como el modelo basado en la estructura de la endo-1,4- β -xilanasas que pertenece a la familia 10 de glicosidasas hidroxilasas (150) y el modelo basado

en la estructura de 1,4- β -xilosidasa perteneciente a la familia 39 de glicosidasas hidroxilasas (183). Sin embargo, aunque el mayor porcentaje de similitud se encuentra en la región del sitio catalítico de la enzima, la homología entre la heparanasa y estas familias de hidrolasas glicosidasas es muy baja, lo que ha impedido el diseño racional de inhibidores.

A pesar de la falta de información estructural se han identificado una gran variedad de inhibidores gracias a la disponibilidad de heparanasa y al desarrollo de métodos como el cribado de alto rendimiento. Entre ellos, se puede destacar:

- Heparina y derivados de heparina de bajo peso molecular. Estos oligosacáridos son potentes agentes anti-metástasicos pero también tienen gran poder anticoagulante, por tanto, no pueden utilizarse en altas concentraciones. Esto ha llevado al uso de derivados de heparina no anticoagulantes como por ejemplo dalteparina, el cual ha sido demostrado que inhibe heparanasa (184).

- Heparinoides o polímeros que imitan la estructura de heparina. Los heparinoides pueden ser de dos tipos: polisacáridos sulfatados natural y químicamente modificados entre los que se encuentran λ -carrageenan, fucodán, pentosán polisulfato, dextrán sulfato, 6-*O*-carboximetil quitina III, laminarín sulfato y spirulán cálcico y el segundo tipo, los heparinoides que son polímeros polianiónicos lineales sintéticos. La mayoría de ellos tienen IC₅₀ en el rango de 1-10 μ M (185, 186).

- Oligosacáridos sulfatados y derivados. Han sido definidos como potentes inhibidores de heparanasa con valores de IC₅₀ de entre 1-30 μ M. Entre ellos destaca JG3 que es un oligosacárido sulfato semisintético derivado de los bloques marinos de oligomannurrate que inhibe la actividad de heparanasa por unión con las regiones que unen con su sustrato (187) y por otro lado hay que mencionar PI-88, que es el único inhibidor de heparanasa que ha llegado a fase clínica (Fig. 10). Actualmente se encuentra en fase clínica II y se ha demostrado que inhibe el crecimiento tumoral, la metástasis y la angiogénesis (188, 189).

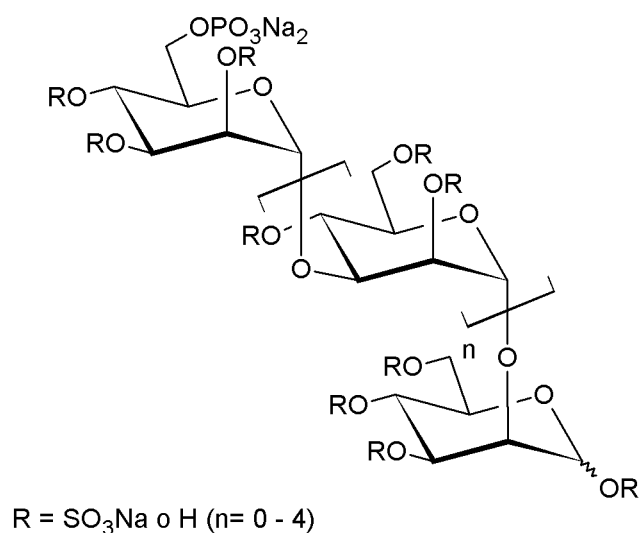


Figura 10. Estructura del inhibidor de heparanasa PI-88

- Inhibidores descubiertos vía cribado de distintas librerías de compuestos. Son compuestos aromáticos más pequeños que tienen IC₅₀ de 1-50 μM y que se pueden clasificar en cuatro grupos: derivados de difenil éter, de carbazoles, de fluorenos y de benzo-1,3-azoles (190).
- Productos naturales y sus derivados. Algunos ejemplos son el (+)-ácido trachyspic, RK-682 y derivados CRM646-A y -B con IC₅₀ entre 10 y 40 μM (190).
- Péptidos que inhiben la actividad de heparanasa. Levy-Adam *et al.* caracterizaron un dominio de unión a heparina en el extremo N-terminal de la subunidad de 50 kDa de heparanasa (Lys¹⁵⁸-Asp¹⁷¹). Un péptido compuesto por estos aminoácidos denominado KKDC inhibe la actividad enzimática de heparanasa *in vitro* e interfiere en la unión de heparanasa con su sustrato (151).
- Anticuerpos anti-heparanasa. Un ejemplo es el anticuerpo #733 que reconoce preferiblemente la subunidad de 50 kDa frente al precursor latente de

65 kDa y neutraliza parcialmente la actividad de heparanasa *in vitro*. En contraste, el anticuerpo 6F8 mejora la actividad enzimática de heparanasa y facilita *in vitro* la invasión de las células tumorales y la curación de heridas en modelos de ratones (138).

- Gen silenciador de heparanasa. Se ha demostrado que un método para suprimir la metástasis y la angiogénesis relacionada con heparanasa consiste en silenciar el gen de heparanasa. Algunos ejemplos son el silenciamiento genético mediado por la ribozima o el siRMN (138).

4.- USO DEL CRIBADO DE FRAGMENTOS EN EL DESARROLLO DE NUEVOS FÁRMACOS

4.1.-Definición y ventajas.

La gran dificultad que conlleva la identificación de un nuevo fármaco se ha visto reflejada en la disminución del número de compuestos que han llegado al mercado o que han sido propuestos para su aceptación durante los últimos años. Los cambios de normativas impuestos por las autoridades y la rigidez de los protocolos empleados han influido notablemente en las estrategias desarrolladas y forzado a las compañías a encontrar alternativas a los métodos ya existentes. Por ejemplo, Lipinski *et al.* examinaron distintas moléculas que habían llegado a fase clínica con el objetivo de encontrar características comunes entre ellas, llegando a la conclusión de que un fármaco debía cumplir lo que se conoce como la “regla de 5”: el peso molecular no debe ser superior a 500, $cLogP < 5$, número de donadores de puentes de hidrógeno < 5 , número de aceptores de puentes de hidrógeno < 10 y el área de superficie polar en la molécula debe ser menor de 100 \AA^2 (191).

Es indudable que como punto de partida, es importante disponer de compuestos que tengan una buena afinidad y selectividad por la diana terapéutica de interés. En este sentido el descubrimiento de fármacos basado en el cribado de fragmentos (FBDD) se ha desarrollado significativamente en los últimos 10-15 años y ahora se ha reconocido como una alternativa tangible a los

métodos más tradicionales de identificación de fármacos utilizados por las compañías farmacéuticas como el cribado de alto rendimiento (HTS) (192).

En el diseño de fármacos basado en el cribado de fragmentos se considera que el mejor camino para llegar a tener compuestos que cumplan la regla de Lipinski es empezar con colecciones de fragmentos altamente solubles con peso molecular menor de 300 y que cumplen la regla de 3 (peso molecular < 300 Da, cLoP < 3, núcleos donadores de puentes de H < 3, número de enlaces con capacidad de rotación < 3 y superficie total polar < 60Å²) (193). Así se identifican estructuras químicas pequeñas llamadas “hits” que muestran baja afinidad por una diana terapéutica ($K_d \sim 100\text{-}1000 \mu\text{M}$). Posteriormente se intenta mejorar el potencial de unión de estos compuestos y esto se lleva a cabo optimizando primero fragmentos individuales para posteriormente unirlos con “linkers” sintéticos, fusionarlos omitiendo las partes que tienen en común u optimizarlos hacia las regiones del sitio de unión de la diana terapéutica que queden por cubrir. Con esta técnica se intenta construir un ligando pieza por pieza con ayuda de la información estructural de la diana con la que se una (194, 195) (Fig. 11).

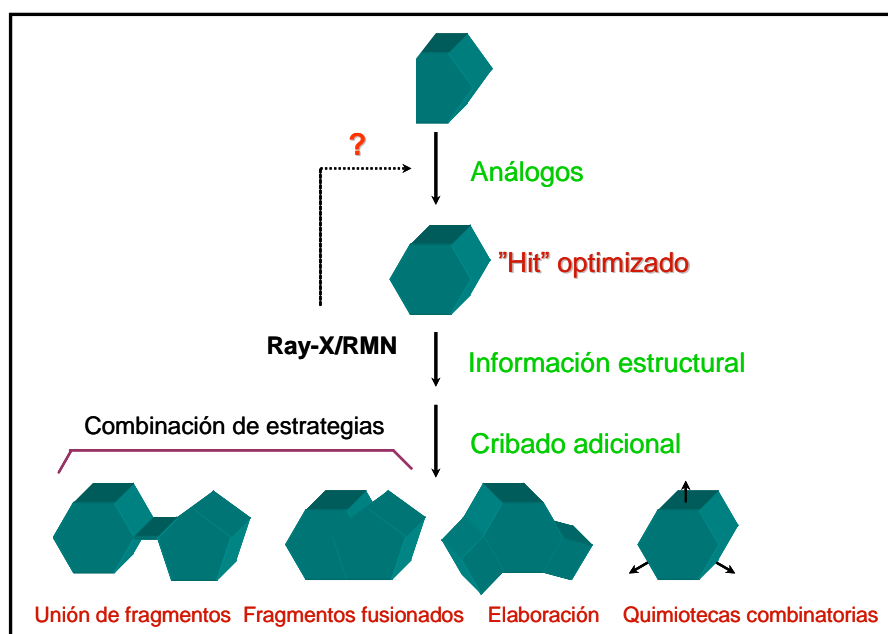


Figura 11. Estrategias en el diseño de fármacos basado en el cribado de fragmentos.

El FBDD tiene principalmente tres ventajas con respecto a las técnicas tradicionales de descubrimiento de fármacos como el HTS (Fig. 12):

- Igual diversidad química con menor número de compuestos.

El número de compuestos que se prueban frente a una diana terapéutica en HTS es del orden de 10^5 - 10^6 . Si tenemos en cuenta que el número de compuestos posibles que cumplan las propiedades de fármacos está estimado en 10^{60} y el número total de moléculas disponibles tanto en instituciones comerciales como académicas de todo el mundo es 10^8 , aunque todo el planeta hiciera un esfuerzo por hacer el cribado de estos compuestos solamente se rozaría una pequeña porción del total de la diversidad del espacio químico posible.

En FBDD, se suelen escoger quimiotecas de menor número de compuestos que en HTS, del orden de 10^2 - 10^4 compuestos con un peso molecular < 250 Da y que presenten una diversidad química muy alta. Estos compuestos que suelen presentar baja afinidad pueden ser modificados añadiéndoles diferentes grupos o pueden combinarse entre ellos para llegar a tener un compuesto de mayor afinidad frente una diana terapéutica. De esta forma se reduce enormemente el número de compuestos manteniendo la diversidad química (196).

- Eficiencia de unión más alta

Algunos de los parámetros que se han establecido en los últimos años como una buena medida para evaluar el potencial de un fragmento “hit” de partida es la eficiencia de unión (LE) (197). La eficiencia de unión se define como el cociente entre la energía libre de unión de un compuesto por una diana concreta y el número de átomos pesados que tenga ese compuesto (HAC).

$$LE = \Delta G/HAC = -RT\ln(IC_{50})/HAC$$

Se estima que un compuesto tiene buen potencial cuando el valor de la eficiencia de unión es igual o superior a 0,3. En el caso de los “hits” encontrados utilizando el cribado de fragmentos, aunque inicialmente son poco potentes, son más pequeños y por tanto una gran parte de la molécula participa en la interacción, de forma que normalmente presentan eficiencias de unión elevadas.

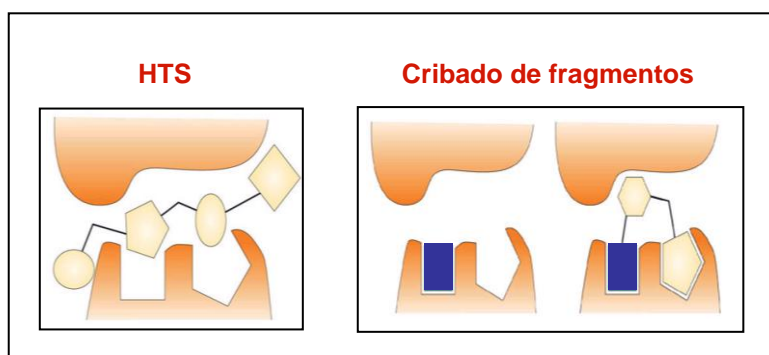


Figura 12. Diferencias en la eficiencia de unión en los compuestos identificados con HTS y FBDD (198).

- Mayor facilidad y velocidad de optimización de “hits”

Finalmente, además de la mayor posibilidad de encontrar un espacio químico libre de patentes cuando se trabaja con compuestos más pequeños, otra ventaja es que el número de posiciones dentro de la molécula susceptibles de ser modificadas decrece considerablemente si lo comparamos con el tipo de compuestos que se usan en HTS. Además son moléculas con una complejidad química menor y por lo tanto las rutas sintéticas utilizadas para modificarlas son más cortas y más simples (199).

4.2.-Técnicas utilizadas en el cribado de fragmentos.

Existe una gran variedad de métodos biofísicos tales como la resonancia magnética nuclear, la cristalografía de rayos X, la espectrometría de masas o la combinación de algunos de ellos, que se usan para investigar e identificar interacciones débiles entre los fragmentos y la diana de interés. Antes de empezar el cribado, muchas compañías generan sus propias quimiotecas de fragmentos para su posterior aplicación con todo tipo de dianas terapéuticas.

Esta selección está basada en maximizar la diversidad química. En una primera etapa generan una base de datos con millones de compuestos disponibles en las distintas casas comerciales. Posteriormente, se consideran las propiedades moleculares de los compuestos y se eliminan todos aquellos que estén fuera del espacio químico de los fragmentos y que no cumplan la regla del 3 (193). Los compuestos que cumplan este primer filtro se evalúan de nuevo para eliminar todos aquellos que contengan grupos funcionales inestables o que pudiesen reaccionar en las condiciones de los ensayos posteriores. Finalmente se agrupan los compuestos por familias y se eligen representantes de cada una de ellas, de forma que se obtenga la mayor diversidad química posible. De esta forma estas quimiotecas exploran mejor el espacio químico disponible que 1 millón de compuestos usados en HTS clásico.

A continuación se describen algunas de las técnicas más usadas en el cribado de fragmentos (200, 201):

- Análisis *in silico* y elección de fragmentos

Se trata de un cribado virtual en el que mediante programas informáticos se evalúa la afinidad proteína-ligando. Esta técnica permite realizar cribado de millones de compuestos en muy poco tiempo, a bajo coste. El análisis de fragmentos *in silico* contra una diana terapéutica concreta representa el punto de partida en muchas compañías farmacéuticas. Este análisis inicial permite reducir enormemente el número de fragmentos necesarios para hacer el cribado.

- Cribado a altas concentraciones

Es el más común de los métodos utilizados para hacer el cribado de compuestos. Típicamente la concentración de ligando que se utiliza está entre 250 y 1000 μM frente a 10-30 μM utilizado en HTS. Es un método rápido que emplea gran variedad de técnicas de detección por lo que no necesita instrumentación muy cara o muy compleja. Sin embargo tiene varios inconvenientes debido a la alta concentración del ligando como por ejemplo el

número elevado de falsos positivos detectados, la desnaturalización de la proteína y la toxicidad en ensayos celulares.

- Resonancia magnética nuclear

Es una técnica utilizada por muchas compañías farmacéuticas en la que se ha puesto mucho esfuerzo para mejorar su sensibilidad y su rendimiento mediante la aplicación del registro digital, criosondas, automuestreadores y campos magnéticos mayores. Tiene la ventaja de que el número de falsos positivos es muy pequeño y se pueden identificar fácilmente. Además es una técnica con alta sensibilidad, se pueden detectar compuestos con constantes de disociación del orden de mM y con algunos experimentos se puede caracterizar la región de unión a la proteína.

- Cristalografía de rayos X

La cristalografía de rayos X es una técnica bien establecida con la que se estudian las estructuras de macromoléculas de interés biológico. Esta técnica permite obtener información estructural de la interacción y en ella se detalla perfectamente el modo de unión del ligando a la proteína.

- Resonancia de plasma superficial

Es una técnica que detecta interacción entre moléculas midiendo las variaciones en el índice de refracción entre dos fases separadas por un sensor (normalmente una película de oro). Normalmente, el cambio en el índice de refracción está relacionado con el cambio de masa en la superficie del sensor cuando las dos moléculas se unen. Tiene la ventaja de que no sólo se obtienen datos de afinidad sino que también da información de la cinética de la interacción. El límite de detección de la técnica está en constantes de disociación del orden de 1 mM.

- “Tethering”

Los fragmentos deben estar marcados como compuestos disulfuro y las dianas deben poseer una cisteína dentro o cerca del sitio de interés o de unión al

ligando. A veces, si la proteína carece de cisteínas en dicho sitio, puede ser introducida por mutagénesis controlando que la actividad de la proteína no haya cambiado. Durante la incubación del ligando con la proteína en condiciones reductoras se rompe el puente disulfuro de los fragmentos y estos pueden interaccionar con la cisteína de la proteína formando nuevos puentes disulfuro. Aquellos fragmentos que interaccionen no covalentemente con los residuos cerca de la cisteína de interés y formen un complejo suficientemente estable son detectados por espectrometría de masas. Puede detectar complejos con constantes de disociación del orden de 1-5 mM.

Recientemente, se han propuesto cuatro nuevas técnicas biofísicas para medir interacciones, sin embargo su uso se ha restringido principalmente para el cribado de quimiotecas que contienen compuestos más elaborados y con una afinidad por la diana de nM- μ M. Estas técnicas son las siguientes:

- Cromatografía frontal de afinidad-espectrometría de masas

Esta técnica se utiliza en muchas fases del proceso de descubrimiento de fármacos y en los últimos años también se ha utilizado para la identificación de ligandos de bajo peso molecular. Con este método, la diana es inmovilizada en un capilar y se pasan mezclas de diferentes compuestos. La interacción se detecta por un incremento del tiempo de retención del ligando usando espectrometría de masas.

- Termofluór

Esta técnica detecta cambios en la temperatura de fusión de una diana cuando un ligando se une y la estabiliza. Aunque claramente se puede utilizar para medir interacciones, la gran utilidad de esta técnica reside en encontrar moléculas que puedan estabilizar proteínas.

- Electroforesis capilar de afinidad

La electroforesis capilar es una técnica de separación en la que las sustancias a analizar se separan en función de su diferente movilidad, en sentido

y velocidad, bajo la acción de un campo eléctrico. Además la utilización de un capilar minimiza los efectos de convección de las técnicas electroforéticas convencionales. Se pueden medir constantes de disociación detectando diferencias en la movilidad de la proteína sola o el complejo.

- Calorimetría de titulación isotérmica

Es una técnica biofísica con la que se puede calcular la constante de disociación entre dos compuestos al medir el calor absorbido o liberado cuando existe una interacción. En un mismo experimento se puede determinar la constante de afinidad (K_a), la estequiometría (n), y los parámetros termodinámicos que definen dicha interacción como la energía libre de Gibbs (ΔG), la entalpía (ΔH) y la entropía (ΔS).

4.3.-Cribado de fragmentos por espectroscopia de Resonancia Magnética Nuclear (RMN)

Utilizando el cribado de fragmentos por RMN se han podido identificar moléculas pequeñas que interaccionan con una diana de interés. Esto es posible mediante la observación de los cambios que se producen durante la interacción en ciertos parámetros de RMN. Los métodos de cribado por RMN se pueden clasificar en aquellos que detectan la interacción por observación de la macromolécula o diana de interés y los que la detectan por la observación del ligando. En el primer caso los cambios que pueden ser monitorizados son los desplazamientos químicos y en el segundo caso los cambios se detectan en otro tipo de parámetros como por ejemplo, en la relajación longitudinal, transversal y de doble cuanto, en el coeficiente de difusión y en la transferencia de magnetización inter e intra molecular.

4.3.1.-Experimentos basados en la observación de la macromolécula. SAR por RMN

SAR (“Structure Activity Relationships”) por RMN (202) consiste en el diseño de fármacos a partir de varios fragmentos que interaccionan en sitios distintos de la proteína. Para ello se utiliza la información que proporcionan los cambios producidos en el espectro de RMN durante la interacción junto con la información estructural y conformacional tanto de la proteína como del ligando.

En primer lugar, la interacción se detecta al comparar el espectro 2D-[^1H - ^{15}N]-HSQC de la proteína enriquecida con ^{15}N en ausencia y presencia del ligando, al observar los cambios producidos en los desplazamientos químicos de los residuos de la diana implicados en la interacción. Se puede calcular la constante de disociación del complejo monitorizando los cambios de desplazamientos químicos en función de la concentración del ligando. Una vez identificado un ligando se intenta modificar o buscar análogos para mejorar su afinidad por la diana en cuestión.

Si se comprueba que otro fragmento interacciona en otro sitio diferente de la proteína, la localización y orientación de ambos en el complejo puede ser determinada experimentalmente mediante experimentos “transfer NOE” o “isotope-edited”. En base a esta información, los dos compuestos pueden ser unidos sintéticamente de forma que se mantenga la orientación espacial de los dos ligandos entre ellos mismos y con respecto a la proteína en cuestión. De esta forma se pueden obtener compuestos de alta afinidad.

Por consiguiente el marcaje isotópico de la proteína con ^{15}N , la asignación de al menos el esqueleto de la proteína que está limitado a proteínas con peso molecular menor de 30 kDa y el conocimiento de la estructura de la proteína son los requisitos mínimos para poder utilizar esta técnica y realizar SAR por RMN.

4.3.2.-Experimentos basados en la observación del ligando

En un esfuerzo por ampliar el número de dianas para hacer cribado, se diseñaron experimentos por RMN basados en la observación del ligando. En este caso, no hay que marcar isotópicamente la proteína y el peso molecular de la proteína debe ser mayor de 20 kDa. Otra ventaja es que la sensibilidad de la

técnica es mayor, ya que se pueden detectar interacciones más débiles. Una desventaja es que en algunos casos puede aumentar el riesgo de falsos positivos. Entre todos los experimentos posibles, los basados en “diferencia de transferencia de saturación” (STD) (203) y los experimentos “WATERLOGSY” (204) son los más utilizados. Ambos experimentos son fáciles de ejecutar y se basan en la transferencia de magnetización selectiva desde la proteína al ligando que está unido. Con este tipo de experimento también se puede hacer competición frente a un ligando conocido pudiéndose de esta forma determinar el sitio de unión con la diana. Si el compuesto que estamos probando compite con el ligando conocido por el mismo sitio de unión con la proteína, se detectará una disminución de intensidad de las señales de uno de ellos o de ambos. De esta forma se minimiza también el número de falsos positivos. Existe también la posibilidad de que se obtengan falsos negativos en el caso de compuestos que tengan afinidad muy alta por la proteína. Estos compuestos no resultarían positivos en el STD y WATERLOGSY. Sin embargo esta posibilidad es muy baja para compuestos de pequeño tamaño como son los fragmentos.

Materiales y Métodos

1.-BIOLOGÍA MOLECULAR

1.1.-Heparanasa

1.1.1- Diseño y preparación de las construcciones de Heparanasa

La clonación del ADN recombinante de heparanasa completa (*Homo sapiens*, referencia uniprot Q9Y251) (HC) fue realizada a partir del plásmido HC-PGEX cedido por el Dr. Luis Manuel Quirós (Instituto Universitario de Oncología del Principado de Asturias). La clonación fue realizada en el plásmido p11 (205), una versión modificada del plásmido pET15b (Fig. 1). Este plásmido contiene una cola de seis histidinas que fusiona con la proteína por su extremo N-terminal y una secuencia de aminoácidos susceptible a un corte proteolítico (indicado como SCTEV), para posteriormente poder aislar la proteína utilizando la proteasa de “Tobacco etch virus” (TEV) (206).

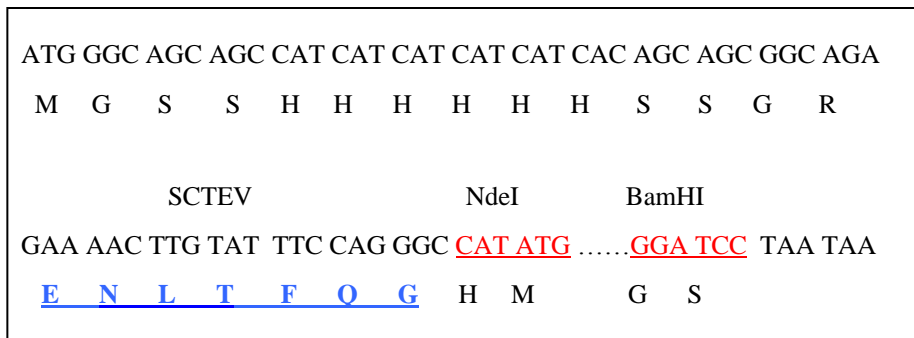


Figura 1. Región modificada del vector pET15b que incluye una cola de 6 histidinas que fusiona por el extremo N-terminal. En azul está señalada la secuencia de nucleótidos que reconoce la proteasa TEV (SCTEV) y en rojo la secuencia de nucleótidos que codifica para los sitios de restricción de las enzimas NdeI y BamHI (205).

Los pasos que se siguieron para realizar la clonación de HC en p11 (HC-p11) fueron:

- Amplificación del inserto deseado a partir del plásmido de partida:

Para ello se empleó el método de la doble reacción en cadena de la polimerasa (doble PCR) (207). La enzima utilizada fue la “Pwo DNA polymerase” (Roche) y los oligonucleótidos diseñados para ello fueron los siguientes:

N_{terminal} largo: 5'-T ATG CAG GAC GTC GTG GAC CTG GAC TTC-3'

N_{terminal} corto: 5'-TG CAG GAC GTC GTG GAC CTG GAC TTC-3'

C_{terminal} largo: 5'-GAT CCT TAT CAG ATG CAA GCA GCA ACT TTG GCA TTT-3'

C_{terminal} corto: 5'-C TTA CTA GAT GCA AGC AGC AAC TTT GGC ATT TCT-3'

Para llevar a cabo la doble PCR, ambas reacciones contenían 5 µL de 10x Pwo Buffer con 1 M MgSO₄ (Roche), 1 µL de 10 mM dNTP, 1 µL de “primer” N_{terminal} (100 pmol/µL), 1 µL de “primer” C_{terminal} (100 pmol/µL) (SIGMA), 1 µL de HC-PGEX (5 ng en la reacción) y 0,5 µL de enzima Pwo DNA Polymerase (Roche) en un volumen final de 50 µL de agua. El programa de temperaturas que se siguió fue el siguiente: una desnaturalización a 94 °C durante 5 minutos seguida de un ciclo que consistía en una desnaturalización a 94 °C durante 30 segundos, una fase de hibridación a 63 °C durante 30 segundos y una extensión a 72 °C durante 30 segundos. Este ciclo fue repetido 30 veces. Finalmente se realizó una extensión a 72 °C durante 10 minutos.

A continuación se detalla la secuencia del inserto de heparanasa completa en nucleótidos y en aminoácidos:

```
caggacgtcgtggacctggacttctcaccaggagccgctgcacctggtgagcccctcgttctgtccgtcacca
ttgacgccaacctggccacggacccgcggttctcctcctctgggttctcaaagcttcgtacctggccagagg
cttctctctcgttacctgaggttgggtggcaccacagacttctaattttcgatccaagaaggaaggtagcg
gatccggctctaaaaagtcaagaacagcacctactcaagaagctctgtagatgtgctatacactttgcaaactgct
caggactggacttgatctttggcctaatgcgttattaagaacagcagattgagtggaacagttctaagtctcagtt
gctcctggactactgctctccaaggggtataacatttctgggaactaggcaatgaacctaacagtttcttaagaag
gctgatatttcatcaatgggtcgcagttaggagaagattttattcaattgcataaacttcaagaagtcaccttcaa
```

aaatgcaaaactctatggtcctgatgttgctcagcctcgaagaaagacggctaagatgctgaagagcttctgaag
gctggtggagaagtgattgattcagttacatggcatcactactatttgaatggacggactgctaccaggaagatttt
ctaaacctgatgtattggacattttttcatctgtgcaaaaagtccaggtggtgagagcaccaggcctggcaa
gaaggctggttaggagaacaagctctgcatatggaggcggagcgccttgctatccgacacctttgcagctgg
ctttatgtgctggataaattggcctgtcagcccgaatgggaatagaagtggatgaggaagattctttggag
caggaaactaccatttagtggatgaaaactcgcatttacctgattattggctatcttctgttcaagaaattggg
gcaccaaggtgtaatggcaagcgtgcaaggtcaagagaaggaagcttcgagtatacttcattgcacaacac
tgacaatccaaggtataaagaaggagatttaactctgtatgccataaacctccataatgccaagtagtctgcggtta
ccctatccttttctaacaagcaagtgataaataccttctaagaccttgggacctcatggattacttccaaatctgc
caactcaatggtctaacttaagatgggtgatgataaaccttggcactttaatggaaaaccttccggccagg
aagttcactgggcttccagcttctcatatagtttttggataagaatgccaagttgctgcttgcattctga

QDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSP
AYLRFGGTKTDFLI FDPKKEGSGSGSKKFKNSTYSRSSVDVLYTFANC SGLDL
IFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFI
NGSQLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVI
DSVTWHHYLNGRTATREDFLNPVDLDIFISSVQKVFQVVESTRPGKKVWLGE
TSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDE
NFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPYKEGD
LTLYAINLHNVTKYLRLPYPF SNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKM
VDDQTLPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI

- Tratamiento del plásmido receptor con las enzimas correspondientes:

En este caso las enzimas utilizadas fueron NdeI y BamHI (NEW ENGLAND BioLabs). En primer lugar se llevó a cabo el corte con NdeI, el volumen final de la reacción fue de 100 μ L con una concentración de ADN plasmídico de 0,12 μ g/ μ L, al que se le añadió 1 μ L de enzima y 5 μ L de buffer 10x compatible con la enzima (NEW ENGLAND BioLabs). A continuación se dejó a 37 °C durante el tiempo necesario hasta que el plásmido se encontrara en su forma lineal; en este caso 12 horas. Una vez comprobado que el plásmido estaba mayoritariamente en su forma lineal se corrió un gel de agarosa con todo el ADN, se cortó la banda correspondiente y posteriormente se purificó

utilizando para ello el kit de purificación Wizard^R SV Gel and PCR Clean-Up SYSTEM (Promega) para eliminar impurezas. Una vez que el DNA estaba puro se cortó con la enzima de restricción BamHI de la misma forma que con la enzima anterior pero en este caso el tiempo de incubación fue 3 horas a 37 °C. El plásmido se volvió a purificar utilizando el kit anteriormente descrito. Posteriormente el plásmido se trató con la enzima fosfatasa alcalina (Roche) para desfosforilar los nucleótidos de los extremos del plásmido. Para ello se volvió a utilizar un volumen final de reacción de 100 µL que contenía 1 µL de enzima, 5 µL de buffer apropiado para esta enzima y el resto de agua Milli-Q. Finalmente el plásmido lineal y desfosforilado fue purificado usando para ello el kit anteriormente descrito dejándolo ya preparado para introducir el inserto correspondiente.

- Ligación:

Se puede definir como la incorporación del inserto entre los sitios NdeI–BamHI del plásmido receptor. Para llevar a cabo la ligación se utilizó una concentración de ADN total de entre 1-10 ng en un volumen total de reacción de 20 µL manteniendo una proporción de ADN inserto:plásmido de 5:1. La mezcla de reacción se puso en presencia de T4 DNA Ligase (Amersham Biosciences) y se incubó a 16 °C durante 30 min.

- Transformación del HC-p11 en células supercompetentes comerciales XL10-GOLD (Stratagene Products):

El protocolo que se siguió para la transformación es el siguiente:

1. Se añadieron 3 µL de la mezcla de reacción de la ligación a 100 µL de células previamente tratadas con β-mercaptoetanol y se dejaron en hielo durante 30 min.
2. Choque térmico: se incubó a 42 °C durante 30 s y luego se dejó en hielo 2 min.

3. Se añadieron 900 μL de medio SOC (Super Optimal broth with Catabolite repression) a las células y se incubó a 37 °C durante una hora.

4. El cultivo anterior se centrifugó a 3500 rpm durante 3 min, se retiraron 900 μL y el sólido resultante se resuspendió en los 100 μL restantes. A continuación estos 100 μL de cultivo se utilizaron para sembrar una placa de medio sólido Luria-Bertani (LB) conteniendo ampicilina como antibiótico a una concentración 100 $\mu\text{g}/\text{mL}$ y se incubó toda la noche a 37 °C.

Todas las construcciones utilizadas en este trabajo fueron clonadas a partir del plásmido anteriormente descrito HC-p11 y usando en todas ellas la doble PCR y el mismo método que para HC-p11 excepto para la subunidad de 45 kDa en la que se utilizó PCR normal utilizando las mismas condiciones de reacción que para la doble PCR. Los plásmidos que se utilizaron fueron p11 (descrito anteriormente), pET47b que contiene una cola de seis histidinas que fusiona con la proteína por su extremo N-terminal (en este plásmido no hay sitio de corte de TEV), pMal que contiene MBP como proteína de fusión y un sitio de corte con TEV y pET44a que contiene Nus A como proteína de fusión y un sitio de corte con TEV.

A continuación se detallan las secuencias en nucleótidos y aminoácidos de todas las construcciones:

○ HEP₃₆₋₁₀₉ (8 kDa)

```
caggacgtcgtggacctgactcttcacccaggagccgctgcacctggtgagccctcgttctgtccgtcac  
cattgacgccaacctggccacggaccgcgggtcctcctcctcctgggttctccaaagctcgtaccttggccag  
aggcttctcctcgtctacctgaggttgggtggcaccaagacagacttcttaatttcgatccaagaaggaa
```

```
QDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRTRLARGL  
SPAYLRFGGTKTDFLIFDPKKE
```

○ HEP₁₅₈₋₄₁₇ (29 kDa)

aaaaagtcaagaacagcacctactcaagaagctctgtagatgtgctatacacttttgcaaaactgctcaggactgg
acttgatctttggcctaataatgcgftattaagaacagcagatttgcaagtgaacagttctaagctcagttgctcctgg
actactgctctccaaggggtataacatttctgggaactaggcaatgaacctaacagtttcctaagaaggctgat
atfttcatcaatgggtcgcagttaggagaagattttattcaattgcataaacttctaagaaagtccacctcaaaaatg
caaaactctatggtcctgatgttggtcagcctcgaagaagacggctaagatgctgaagagcttctgaaggctg
gtggagaagtgattgattcagttacatggcatcactactatttgaatggacggactgctaccaggaagattttcta
aacctgatgtattggacattttatttcatctgtgcaaaaagtttccaggtggttgagagcaccaggcctggcaa
gaaggctctggttaggagaaacaagctctgcatatggaggcggagcgccttgctatccgacaccttgcagctg
gctttatgtgctggataaattggcctgtcagcccgaatgggaatagaagtggatgaggcaagtattcttgg
agcaggaaactaccatttagtgatgaaaacttcgatcctttacctgattattggctatctcttctgttcaaaaattg
gtgggcaccaag

KKFKNSTYSRSSVDVLYTFANC SGLDLIFGLNALLRTADLQWNSSNAQLLLD
YCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDFIQLHKLLRKSTFKNA
KLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYLNGRTATREDFLNP
DVLDI FISSVQKVFQVVESTRP GK V LWGETSSAYGGGAPLLSDTFAAGFMW
LDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTK

○ HEP₂₃₃₋₄₁₇ (21 kDa)

gctgatatttcatcaatgggtcgcagttaggagaagattttattcaattgcataaacttctaagaaagtccacctca
aaaatgcaaaactctatggtcctgatgttggtcagcctcgaagaagacggctaagatgctgaagagcttctga
aggctggtggagaagtgattgattcagttacatggcatcactactatttgaatggacggactgctaccaggaag
atfttctaaacctgatgtattggacattttatttcatctgtgcaaaaagtttccaggtggttgagagcaccaggcct
ggcaagaaggctggttaggagaaacaagctctgcatatggaggcggagcgccttgctatccgacaccttgc
agctggctttatgtgctggataaattggcctgtcagcccgaatgggaatagaagtggatgaggcaagtatt
cttggagcaggaaactaccatttagtgatgaaaacttcgatcctttacctgattattggctatctcttctgttcaag
aaattggtgggcaccaag

ADIFINGSQLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLK
AGGEVIDSVTWHHYLNGRTATREDFLNP DVLDI FISSVQKVFQVVESTRP G

KKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGA
GNYHLVDENFDPLPDYWLSLLFKKLVGTK

○ HEP₁₅₈₋₅₄₃ (45 kDa)

aaaaagtcaagaacagcacactactcaagaagctctgtagatgtgctatacacttttgc aaactgctcaggactgg
acttgatctttggcctaataatgcgttattaagaacagcagatttgagtggaacagttctaagctcagttgctcctgg
actactgctctccaaggggtataacatttctgggaactaggcaatgaacctaacagttcctaagaaggctgat
atfttcatcaatgggtcgcagttaggagaagatfttattcaattgcataaacttctaagaaagtcacactcaaaaaatg
caaaactctatggtcctgatgttggtcagcctcgaagaagacggctaagatgctgaagagcttctgaaggctg
gtggagaagtgattgattcagttacatggcatcactactatttgaatggacggactgctaccaggaagattttcta
aacctgatgtattggacatttttattcatctgtgcaaaaagtttccaggtggtgagagcaccaggcctggcaa
gaaggtctggttaggagaacaagctctgcatatggaggcggagcgccttgctatccgacaccttgacagctg
gctttatgtggtgataaattggcctgtcagcccgaatgggaatagaagtggatgaggcaagtattctttgg
agcaggaaactaccatttagtgatgaaaacttcgatcctttacctgattattggctatcttctgttcaagaaattg
gtgggcaccaaggtgtaattggaagcgtgcaaggttcaagagaaggaagcttcgagataccttcattgcac
aaactgacaatccaaggtataaagaaggagatttaactctgatgcataaacctccataatgtccaagtagc
ttgcggttaccctatccttttctaacaagcaagtgataaataccttctaagaccttgggacctcatggattactttc
caaatctgtccaactcaatggtctaactctaagatggtggatgatcaaaccttgccaccttaatggaaaaacctc
tccggccaggaagttcactgggctgccagcttctcatatagtttttgtgataagaaatgccaagttgctgcttg
catctga

KKFKNSTYSRSSVDVLYTFANC SGLDLIFGLNALLRTADLQWNSSNAQLLLD
YCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDFIQLHKLLRKSTFKNA
KLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYLNGRTATREDFLNP
DVLDFISSVQKVFQVVESTRP GK K V L G E T S S A Y G G G A P L L S D T F A A G F M W
L D K L G L S A R M G I E V V M R Q V F F G A G N Y H L V D E N F D P L P D Y W L S L L F K K L V G T K
V L M A S V Q G S K R R K L R V Y L H C T N T D N P R Y K E G D L T L Y A I N L H N V T K Y L R L P Y P
F S N K Q V D K Y L L R P L G P H G L L S K S V Q L N G L T L K M V D D Q T L P P L M E K P L R P G S S
L G L P A F S Y S F F V I R N A K V A A C I

- HEP₃₆₋₄₁₇ (37 kDa)

caggacgtcgtggacctggacttctcaccaggagccgctgcacctggtgagccccctgftcctgtccgtcac
 cattgacgccaacctggccacggaccgcggttctcctcctctgggttctccaaagcttcgtacctggccag
 aggttctcctcctcgtacctgaggttggggcaccacagactcctaatttcgatcccaagaaggaaggt
 agcggatccggtcctaaaaagtcaagaacagcacctactcaagaagctctgtagatgtctatacactttgcaa
 actgctcaggactggactgatctttggcctaataatcggttattaagaacagcagattgcagtggacacgttcta
 gctcagttgctcctggactactgctctccaaggggtataacattcttgggaactaggcaatgaacctaacgttt
 ccttaagaaggctgatatttcatcaatgggtcgcagttaggagaagattttattcaattgcataaacttcaagaaa
 gtccacctcaaaaatgcaaaactctatggctctgatgttggtcagcctcgaagaagacggctaagatgctgaa
 gagcttctgaaggctggtggagaagtattgattcagttacatggcatcactactatttgaatggacggactgct
 accaggaagatttttaaacctgatgtattggacatttttattctctgtcaaaaagtttccaggtggttga
 gcaccaggcctggcaagaaggtctggttaggagaaacaagctctgcatatggaggcggagcgccttctatc
 cgacaccttgcagctggctttatgtggtggataaattggcctgtcagcccgaatgggaatagaagtggat
 gaggcaagtattcttggagcaggaaactaccatttagtggtgataaaaactcagatcctttacctgattattggctatc
 tcttctgtcaagaattggtgggcaccaagg

QDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRRTLARGL
 SPAYLRFGGTKTDFLIFDPKKEGSGSGSKKFKNSTYSRSSVDVLYTFANCSG
 LDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKA
 DIFINGSQLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKA
 GGEVIDSVTWHHYLNGRTATREDFLNPVLDIFISSVQKVFQVVESTRPGK
 KVWLGETSSAYGGAPLLSDTFAGFMWLDKLGLSARMGIEVVMRQVFFGAG
 NYHLVDENFDPLPDYWLSLLFKKLVGK

Los oligonuclótidos utilizados para las clonaciones fueron diferentes según la construcción y el plásmido utilizado:

1. Construcciones clonadas en p11

- HEP₃₆₋₁₀₉

N_{terminal} largo: 5'-T ATG CAG GAC GTC GTG GAC CTG GAC TTC-3'

N_{terminal} corto: 5'-TG CAG GAC GTC GTG GAC CTG GAC TTC-3'

C_{terminal} largo: 5'-GA TCC TTA TTC CTT CTT GGG ATC GAA AAT TAG-3'

C_{terminal} corto: 5'-CTTATTCCTTCTTGGGATCGAAAATTAG-3'

○ HEP₁₅₈₋₄₁₇

N_{terminal} largo: 5'-T ATG AAA AAG TTC AAG AAC AGC ACC TAC TCA AGA-3'

N_{terminal} corto: 5'-TG AAA AAG TTC AAG AAC AGC ACC TAC TCA AGA-3'

C_{terminal} largo: 5'-GAT CCT TAC TTG GTG CCC ACC AAT TTC TTG AAC AGA-3'

C_{terminal} corto: 5'-C TTA CTT GGT GCC CAC CAA TTT CTT GAA CAG AAG-3'

○ HEP₂₃₃₋₄₁₇

N_{terminal} largo: 5'-TA TGG CTG ATA TTT TCA TCA ATG GGT CGC-3'

N_{terminal} corto: 5-TGG CTG ATA TTT TCA TCA ATG GGT CGC-3'

C_{terminal} largo: 5'-GAT CCT TAC TTG GTG CCC ACC AAT TTC TTG AAC AGA-3'

C_{terminal} corto: 5'-C TTA CTT GGT GCC CAC CAA TTT CTT GAA CAG AAG-3'

○ HEP₁₅₈₋₅₄₃

N_{terminal} largo: 5'-T ATG AAA AAG TTC AAG AAC AGC ACC TAC TCA AGA- 3'

N_{terminal} corto: 5'-TG AAA AAG TTC AAG AAC AGC ACC TAC TCA AGA-3'

C_{terminal} largo: 5'-GAT CCT TAT CAG ATG CAA GCA GCA ACT TTG GCA TTT-3'

C_{terminal} corto: 5'-C TTA CTA GAT GCA AGC AGC AAC TTT GGC ATT TCT-3'

○ HEP₃₆₋₄₁₇

N_{terminal} largo: 5'-T ATG CAG GAC GTC GTG GAC CTG GAC TTC-3'

N_{terminal} corto: 5'-TG CAG GAC GTC GTG GAC CTG GAC TTC-3'

C_{terminal} largo: 5'-GAT CCT TAC TTG GTG CCC ACC AAT TTC TTG AAC AGA-3'

C_{terminal} corto: 5'-C TTA CTT GGT GCC CAC CAA TTT CTT GAA CAG AAG-3'

2. HEP₃₆₋₄₁₇ clonada en pMal

N_{terminal} largo: 5'-AA TTC GAA AAC TTG TAT TTC CAG GGC CAT ATG CAG
GAC 3'

N_{terminal} corto: 5'-CGA AAA CTT GTA TTT CCA GGG CCA TAT GCA GGA 3'

C_{terminal} largo: 5'-GAT CCT TAC TTG GTG CCC ACC AAT TTC TTG AAC AGA-3'

C_{terminal} corto: 5'-C TT ACT TGG TGCC CAC CAATTT CTT GAA CAG AAG-3'

3. HEP₃₆₋₄₁₇ clonada en pET44a

N_{terminal} largo: 5'-GCG GAA AAC TTG TAT TTC CAG GGC CAT ATG 3'

N_{terminal} corto: 5'-G GAA AAC TTG TAT TTC CAG GGC CAT ATG 3'

C_{terminal} largo: 5'-GAT CCT TAC TTG GTG CCC ACC AAT TTC TTG AAC AGA-3'

C_{terminal} corto: 5'-C TTA CTT GGT GCC CAC CAA TTT CTT GAA CAG AAG-3'

4. HEP₁₅₈₋₅₄₃ (45 kDa) clonada en pET47b

N_{terminal}: 5'-GTA GAG GGA TCC CTA CTA GAT GCA AGC AGG-3'

C_{terminal} : 5'-CGC TCT CCG CGG AAA AAG TTC AAG AAC ACG-3'

1.1.2.-Expresión de las distintas construcciones

Una vez que las distintas construcciones de heparanasa se clonaron en el vector de expresión adecuado, se llevó a cabo la transformación para cada una de ellas en células de expresión de *Escherichia coli* (*E. coli*) BL21 (DE3) (Stratagene) excepto para HEP₁₅₈₋₄₁₇ en la que se usó BL21-CodonPlus(DE3)-RIL (Stratagene). La transformación fue realizada utilizando 1 µL de plásmido en 50 µL de las células anteriormente descritas siguiendo el mismo protocolo descrito en el apartado 1.1.1.

El cultivo de las células se realizó en condiciones aeróbicas en medio líquido Luria-Bertani (LB). Todos los cultivos fueron preparados utilizando ampicilina como antibiótico a una concentración final de 100 µg/mL, excepto para el caso de HEP₁₅₈₋₄₁₇ en el que se añadió además cloranfenicol también a una concentración final de 100 µg/mL. La resistencia específica a la ampicilina se la confiere el vector de expresión p11 y al cloranfenicol las células de expresión BL21-CodonPlus(DE3)-RIL

Para la correcta oxigenación del cultivo se utilizaron matraces de 2 L con 500 mL de medio cada uno y se mantuvieron en agitación continua durante todo el

tiempo de incubación. Tras la inoculación del medio con 5 mL de un precultivo crecido a 37 °C en LB durante la noche anterior, se mantuvieron a 37 °C hasta alcanzar una densidad óptica a 600 nm de 0,6-0,8 excepto para la construcción HEP_{158-417-p11} en la que se utilizó 0,3. Los plásmidos utilizados para cada una de las construcciones contenían en su secuencia el gen regulador “LacI”. El producto del gen regulador es un represor que impide la expresión de los genes en ausencia del inductor. Para inducir la expresión de la proteína se utilizó isopropil-β-tiogalactopiranosido (IPTG) (SIGMA-ALDRICH) capaz de inactivar el represor del operón *Lac* e inducir la expresión de los genes de interés. Para ello se utilizó una concentración de 0,8 mM de IPTG y una temperatura de 25 °C. A continuación se dejó que la expresión tuviese lugar durante un tiempo aproximado de 18 horas. Transcurrido este tiempo, se recogieron las células por centrifugación de los cultivos a 3500 rpm, durante 15 min a 4 °C, y se almacenaron los pellets a -20 °C hasta su uso. Todo este proceso viene representado en la Fig. 2.

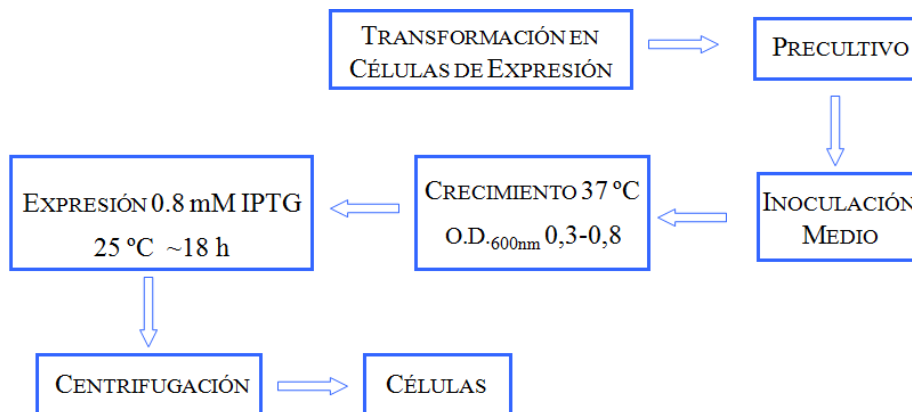


Figura 2. Esquema para la expresión de la proteína.

1.1.3.-Marcaje isotópico de las construcciones de heparanasa

Para la preparación de muestras marcadas uniformemente en ^2H , ^{15}N y/o ^{13}C , se utilizó un medio mínimo M9 utilizando D_2O 70%, $^{15}\text{NH}_4\text{Cl}$ y ^{13}C -glucosa (Eurisotop) como únicas fuentes de deuterio, nitrógeno y carbono respectivamente (208).

Las condiciones para expresar la proteína marcada en ^{15}N y/o ^{13}C fueron las mismas que las descritas anteriormente. Sin embargo en el caso de muestras marcadas en ^2H , hubo que hacer una adaptación previa de las células. En primer lugar se puso un precultivo en medio LB hasta que las células alcanzaron una densidad óptica a 600 nm de 0,6-0,8. Este precultivo fue utilizado para inocular un cultivo de 100 mL de medio mínimo M9 conteniendo 0% D_2O ; cuando las células alcanzaron la densidad óptica deseada, se volvió a utilizar este cultivo para inocular otro de 30% D_2O y se volvió a repetir la operación posteriormente inoculando cultivos de 50% y finalmente del 70% de D_2O . De esta manera nos asegurábamos de que un 70% de la proteína estaba marcada con ^2H (208). La expresión de la proteína se indujo de la misma forma que se ha indicado en el apartado 1.1.2.

1.1.4.-Marcaje isotópico selectivo de las construcciones de heparanasa

La proteína fue también expresada marcada en ^{15}N de forma selectiva en algunos aminoácidos. Para ello, se añadió 1 g de aminoácido marcado en ^{15}N a 500 mL de medio de cultivo LB (209). Los aminoácidos utilizados fueron ^{15}N -Gly, ^{15}N -Leu, ^{15}N -Lys, ^{15}N -Glu y ^{15}N -Ser (Tracer tec). Las células se crecieron a 37 °C hasta una O.D. de 0,6-0,8 a 600 nm. La expresión se indujo con 0,8 mM de IPTG y se llevó a cabo a 37 °C durante 4 horas.

1.1.5.-Lisis celular de las construcciones de heparanasa

La ruptura de las células se realizó en dos etapas, una primera etapa basada en una ruptura enzimática (utilizando lisozima), y una segunda etapa basada en una

ruptura física (sometiendo a las células a diferentes ciclos de sonicación). Para la ruptura enzimática, las células que procedían de un litro de cultivo se resuspendieron en 50 mL de un tampón de lisis cuya composición variaba dependiendo del plásmido utilizado. Estos tampones fueron:

- Tampón A: 25 mM Tris-HCl pH 8, 300 mM de cloruro sódico, 10% (v/v) glicerol (MERCK), 1% (v/v) Triton (FLUKA) utilizados para todas las construcciones clonadas en p11 y pET47b.
- Tampón B: 25 mM Tris-HCl pH 7,4, 200 mM de cloruro sódico, 1 mM EDTA (Roche) para HEP₃₆₋₄₁₇-pMal.
- Tampón C: 25 mM Tris-HCl pH 8, 300 mM de cloruro sódico, 10% (v/v) glicerol (MERCK) para HEP₃₆₋₄₁₇-pET44a.

Todos estos tampones fueron suplementados con 0,5 mg/mL de lisozima (Fluka), 10 µg/mL de DNasa (Roche) y media tableta de cóctel inhibidor de proteasas libre de EDTA (Roche).

Posteriormente las células se dejaron en agitación moderada durante media hora a 4 °C. Una vez transcurridos los 30 min, fueron sometidas a 10 ciclos de sonicación (Microson™ Ultrasonic cell disruptor XL, Misonix) que consistían cada uno de ellos en sonicar durante 30 s a máxima potencia seguido de un periodo de descanso de 10 s. Una vez rotas las células, se separaron el sobrenadante y el pellet mediante centrifugación a 9000 rpm a 4 °C.

1.1.6.-Purificación de las construcciones clonadas en p11 Y pET47b.

Después de la lisis celular, todas las construcciones de heparanasa fueron encontradas en el pellet. Posteriormente, estas construcciones fueron sometidas a distintos ciclos de sonicación con distintos tampones de lisis con el objeto de purificarlas y eliminar del pellet otras proteínas más solubles que heparanasa. Finalmente las construcciones de heparanasa ya prácticamente puras fueron solubilizadas añadiendo al tampón 0,125% (p/v) de N-lauroylsarcosine. Todas

las sonicaciones se llevaron a cabo con el mismo sonicador y usando los mismos ciclos que para la ruptura celular. Los tampones usados para cada sonicación fueron los siguientes:

- 2ª sonicación: 25 mM Tris-HCl pH 7,4, 300 mM de cloruro sódico, 1% (p/v) Triton, 10% (v/v) glicerol.
- 3ª sonicación: 25 mM Tris-HCl pH 7,4, 300 mM de cloruro sódico, 0,06% (p/v) N-lauroylsarcosine (SIGMA-ALDRICH).
- 4ª sonicación: 25 mM Tris-HCl pH 7,4, 300 mM de cloruro sódico, 0,125% (p/v) de N-lauroylsarcosine.

Después de cada sonicación se separó sobrenadante y pellet mediante centrifugación a 9000 rpm durante 30 min. El pellet fue resuspendido en el nuevo tampón dejándolo en agitación durante 30 min a 4 °C. Todo este proceso aparece esquematizado en la Fig. 3

Posteriormente se refinó la purificación de la proteína mediante cromatografía de afinidad de metal inmovilizado (IMAC). Para ello se aplicó el sobrenadante de la 4ª sonicación a una columna que previamente había sido cargada con 2 mL de resina TALON (Clontech) y equilibrada con el mismo tampón en el que se encontraba la proteína (25 mM Tris-HCl pH 7,4, 300 mM de cloruro sódico, 0,125% de N-lauroylsarcosine). La proteína se dejó en contacto con la resina durante dos horas con agitación usando una noria con una rotación de 18 rpm. Una vez transcurrido este tiempo se dejó pasar el tampón a través de la columna y la proteína que quedó unida a la resina fue lavada primero con 25 mL del mismo tampón de equilibración y en segundo lugar con el mismo tampón suplementado con 5 mM imidazol (SIGMA) para eliminar posibles uniones inespecíficas de otras proteínas a la resina. Seguidamente la proteína fue eluida con 10 mL de este mismo tampón suplementado con 500 mM imidazol. La pureza de la proteína se valoró sobre SDS-PAGE (Bis/Tris NuPAGE polyacrylamide gels) al 12% (p/v).

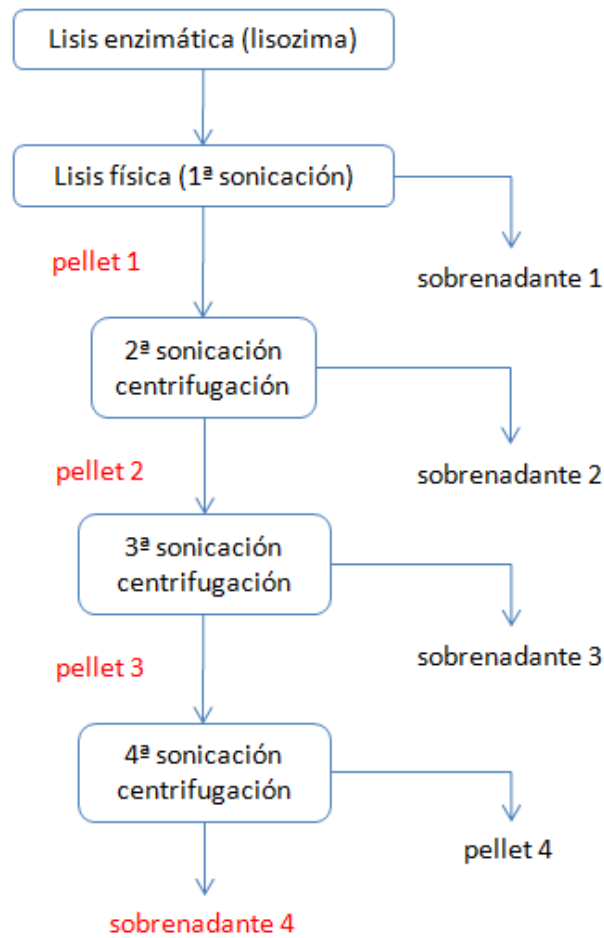


Figura 3. Esquema de tratamiento de lisis y purificación de las construcciones de heparanasa clonadas en el plásmido p11 y pET47b. En rojo está señalado el recorrido de la proteína desde la célula hasta tenerla en condiciones solubles.

1.1.7.-Purificación de HEP₃₆₋₄₁₇-pMAL.

Después de la lisis celular la proteína fue encontrada en el sobrenadante. Teniendo en cuenta que esta construcción contiene fusionada la proteína MBP (proteína de unión a maltosa), la purificación de HEP₃₆₋₄₁₇ se llevó a cabo mediante cromatografía de afinidad aprovechando la interacción entre la MBP y la resina de amilosa. De esta forma, el sobrenadante fue aplicado sobre una

columna que contenía 5 mL de resina de amilosa (New England BioLabs) equilibrado con tampón B. La columna que contenía la proteína unida a la resina fue lavada con tampón B y posteriormente eluida con el mismo tampón suplementado con 10 mM maltosa. Se comprobó la pureza de la proteína utilizando SDS-PAGE al 12% (p/v). Toda la purificación se llevó a cabo a 4 °C.

1.1.8.-Purificación de HEP₃₆₋₄₁₇-pET44a

Se utilizó cromatografía de metal inmovilizado (IMAC) para purificar esta construcción. Para ello, el sobrenadante que se obtuvo después de la lisis celular se aplicó a una columna cargada con 2 mL de resina TALON (Clontech) equilibrada con buffer C. La proteína que quedó unida a la resina fue lavada en primer lugar con 25 mL del mismo tampón de equilibración y en segundo lugar con el mismo tampón suplementado con 5 mM imidazol para eliminar uniones inespecíficas. Seguidamente la proteína fue eluida con 10 mL de este tampón suplementado con 500 mM imidazol. La pureza de la proteína fue valorada por SDS-PAGE al 12% (p/v).

1.1.9.- Identificación de la secuencia y peso molecular de HEP₁₅₈₋₄₁₆ por espectrometría de masas MALDI-TOF

El peso molecular y la secuencia de aminoácidos fueron confirmadas mediante espectrometría de masas MALDI-TOF. Para ello se utilizó una muestra disuelta en 25 mM fosfato pH 6, 50 mM NaCl y 0,125% de N-lauroyl-sarcosine (p/v) con una concentración de 0,3 µg/µL en un volumen final de 200 µL.

- Identificación de la secuencia de aminoácidos de la proteína.

La muestra fue precipitada con una mezcla de ácido tricloroacético (TCA)/acetona. Una vez que la proteína fue precipitada, se procedió a la digestión de la muestra con tripsina. Para ello, la proteína precipitada fue tratada en primer lugar con 5 mM de DTT disuelto en una disolución 50 mM de bicarbonato amónico (NH₄HCO₃) durante 30 s a 60 °C, a continuación con 35 mM de IAM (yodoacetamida) disuelto en una disolución 50 mM de bicarbonato

amónico (NH_4HCO_3) durante 30 s en oscuridad y posteriormente se incubó con 100 ng de tripsina a 37 °C para llevar a cabo la digestión. La reacción de digestión fue parada con ácido trifluoroacético (TFA) al 10% (pH 1). Finalmente se llevó a cabo el espoteo directo y el Zip Tip de 10 μL de muestra. Se incluyó en el procesado de las muestras un blanco de 75 ng de BSA como control de la digestión y un blanco con los reactivos para determinar el origen de posibles contaminaciones por queratinas. Antes del análisis de la muestra, se calibraron la placa y los métodos de adquisición. Los espectros de MALDI TOF se adquirieron en modo reflectón positivo utilizando un 4700 Proteómica Analyzer, ABI. Se seleccionaron automáticamente 5 iones entre aquellos de mayor intensidad, excluyendo los que correspondían a contaminaciones conocidas. Posteriormente se realizó una calibración interna con los picos de autodigestión de la tripsina y se obtuvo el espectro de fragmentación de cada uno de estos iones utilizando la disociación inducida por colisión (“CID on”) con 1 kV de energía de colisión.

Las muestras fueron analizadas con un Applied Biosystems en un servidor de MASCOT residente (http://www.matrixscience.com/search_intro.html). Las bases de datos utilizadas fueron SwissProt y HomaMade.

○ Identificación del peso molecular de la proteína

Se hicieron 6 diluciones de la muestra (1:1, 1:2, 1:5, 1:10, 1:20) en matriz de ácido α -ciano-4-hidroxicinámico y se depositaron en una placa de MALDI. El peso molecular se determinó en un espectrómetro de masas con ionización MALDI y detectores de vuelo de tándem (MALDI TOFTOF, 4700 PA, ABI). Los espectros se adquirieron en modo lineal positivo. Los métodos de adquisición fueron calibrados previamente al análisis de las muestras y se acumularon entre 2000 a 4000 disparos a una intensidad de láser de 6000 (YAG LASER, 200 Hz). La adquisición de los datos se realizó de forma aleatoria por toda la muestra.

Estos experimentos fueron realizados por el Servicio de Proteómica (Proteored) del Centro de Investigación Príncipe Felipe (Valencia).

1.2.-Purificación de fondaparinux

Se partió de una solución salina de 3,2 mL que contenía Fondaparinux (Arixtra) a una concentración de 12,5 mg/mL. La muestra fue congelada en nitrógeno líquido y liofilizada durante toda la noche para eliminar el agua. Posteriormente fue disuelta en 400 μ L de agua MilliQ. Una vez que la muestra fue concentrada, se eliminó el cloruro sódico mediante cromatografía de exclusión molecular. Para ello se utilizó una columna PD MidiTrap G-10 previamente lavada y equilibrada con 16 mL de agua MilliQ. Posteriormente se aplicaron los 400 μ L de muestra que contenían el Fondaparinux sobre la columna y se dejó pasar el volumen de tampón que ocupaba la resina. Se eluyó con 1,3 mL de agua MilliQ y se recogieron 13 fracciones de 100 μ L. Para saber si las distintas fracciones contenían cloruro sódico, se utilizó una solución de nitrato de plata que añadida a las fracciones que contenían sal formaban un precipitado blanco de cloruro de plata. De esta forma solo se escogieron las fracciones libres de cloruro sódico y que contenían el Fondaparinux puro. Finalmente se unieron todas estas fracciones y se liofilizaron para ser almacenadas hasta su utilización en experimentos posteriores.

2.-RESONANCIA MAGNÉTICA NUCLEAR

2.1.-Preparación de la muestra

2.1.1.-Experimentos bi- y tri-dimensionales necesarios para la asignación de la proteína

Las muestras utilizadas para adquirir los experimentos estaban marcadas isotópicamente en ^{15}N , ^{13}C , ^2H o combinaciones de estos isótopos dependiendo del tipo de experimentos (*Tabla 1*) aunque también se utilizaron muestras marcadas selectivamente con ^{15}N en algunos aminoácidos. Para adquirir los experimentos bi y tridimensionales se usaron muestras con una concentración de entre 150 μM y 1 mM en un volumen de entre 450-500 μL , por lo que en el caso de las construcciones de heparanasa fue necesaria la concentración de la proteína

tras su purificación. Para ello, una vez terminada la purificación, se eligieron las fracciones más limpias y fueron dializadas frente a 25 mM de fosfato sódico, 50 mM de cloruro sódico y 0,125% de N-lauroyl-sarcosine (tampón utilizado para la adquisición de los experimentos). Finalmente, la proteína fue concentrada centrifugando a 7.000 rpm usando concentradores (Vivaspin) con un corte de peso molecular entre 2.000 y 10.000 Da dependiendo del tamaño de la construcción con la que se estaba trabajando (*Tabla 2*). Todas las muestras contenían una cantidad de D₂O entre el 5-10% del volumen de la muestra. Algunos experimentos como el ¹³C-HSQC, el [¹H-¹³C]-HSQC-NOESY y el HCCH-TOCSY se adquirieron utilizando D₂O como disolvente para reducir la señal del H₂O. El disolvente utilizado para adquirir cada uno de los experimentos se detalla en la *Tabla 1*.

Tabla 1. Marcaje isotópico y disolvente de las muestras según el tipo de experimentos

	MARCAJE ISOTÓPICO	DISOLVENTE
[¹⁵ N, ¹ H]-HSQC	¹⁵ N	H ₂ O + 5-10% D ₂ O
[¹⁵ N, ¹ H]-HSQC-TOCSY	¹⁵ N	H ₂ O+ 5-10% D ₂ O
[¹⁵ N, ¹ H]-HSQC-NOESY	¹⁵ N	H ₂ O+ 5-10% D ₂ O
HNCO*	¹⁵ N + ¹³ C	H ₂ O+ 5-10% D ₂ O
HNCA*	¹⁵ N + ¹³ C	H ₂ O+ 5-10% D ₂ O
CBCA(CO)NH*	¹⁵ N + ¹³ C	H ₂ O+ 5-10% D ₂ O
HNCACB*	¹⁵ N + ¹³ C	H ₂ O+ 5-10% D ₂ O
HNCACO*	¹⁵ N + ¹³ C	H ₂ O+ 5-10% D ₂ O
HNHA	¹⁵ N	H ₂ O+ 5-10% D ₂ O
HNHB	¹⁵ N	H ₂ O+ 5-10% D ₂ O
HBHA(CO)NH	¹⁵ N	H ₂ O+ 5-10% D ₂ O
[¹³ C, ¹ H]-HSQC	¹⁵ N + ¹³ C	D ₂ O
[¹³ C, ¹ H]-HSQC-NOESY	¹⁵ N + ¹³ C	D ₂ O

(H)CCH-TOCSY	$^{15}\text{N} + ^{13}\text{C}$	D ₂ O
H(C)CH-TOCSY	$^{15}\text{N} + ^{13}\text{C}$	D ₂ O

* Las muestras de HEP₁₅₈₋₄₁₇ estaban marcadas adicionalmente en ^2H para las versiones TROSY de estos experimentos

Tabla 2. Corte de peso molecular de la membrana utilizada para cada construcción

	TAMAÑO PROTEÍNA (Da)	TAMAÑO DE CORTE DE MEMBRANA (Da)
HEP ₁₅₈₋₄₁₇	29000	10000
HEP ₂₃₃₋₄₁₇	21000	10000
HEP ₃₆₋₁₀₉	8000	2000
HEP ₁₅₈₋₅₄₃	45000	10000

2.1.2.-Ensayos de interacción

Los tipos de experimentos que se usaron para medir la interacción entre proteína y ligando fueron waterLOGSY (210), STD (203) y mapeo de desplazamientos químicos utilizando [^1H - ^{15}N]-HSQC (211).

En todos los casos se partió de un stock de proteína entre 0,5 y 1 mM (preparado como se explicó en el punto 2.1.1) disuelto en tampón 25 mM fosfato sódico pH 6, 50 mM de cloruro sódico y 0,125% (p/v) de N-lauroylsarcosine. El stock utilizado para el ligando tenía una concentración de 50 mM y se encontraba disuelto en el mismo tampón que la proteína o en caso de que no fuera soluble en agua, en DMSO deuterado. El volumen total de todas las muestras era de 450-500 μl y todas ellas contenían entre un 5 y 10% de D₂O. La preparación de la muestra fue diferente según el tipo de experimento. A continuación se detalla para cada uno de ellos.

- WaterLOGSY

El ratio proteína/ligando para este tipo de experimento fue 1:60 con una concentración de proteína de 5 μM y una concentración de ligando de 300 μM .

- STD

En este caso el ratio proteína/ligando fue 1:100 con una concentración de proteína de 5 μM y una concentración de ligando de 500 μM .

- [^1H - ^{15}N]-HSQC

Para este tipo de experimento se usó proteína marcada en ^{15}N a una concentración de 50 μM como muestra control. Los ligandos utilizados en este tipo de experimento estaban todos disueltos en el mismo tampón que la proteína. Para ver la interacción fue necesario hacer una titulación con ratios proteína: ligando de 1:1, 1:2, 1:5, 1:10, 1:25, 1:50.

2.2.-Adquisición de los experimentos

2.2.1.-Experimentos bi- y tri-dimensionales necesarios para la asignación de la proteína

Los experimentos tridimensionales de RMN realizados con la muestra de HEP₁₅₈₋₄₁₇ se llevaron a cabo a frecuencias de 899,58, 91,16 y 226,23 MHz para ^1H , ^{15}N y ^{13}C respectivamente en un espectrómetro Bruker Avance 900 que operaba a un campo electromagnético de 21,06 Teslas y equipado con una criosonda TCI de 5 mm.

Los experimentos bi y tridimensionales correspondientes al resto de las construcciones se realizaron a frecuencias de 600,13, 60,81 y 150,90 MHz para ^1H , ^{15}N y ^{13}C respectivamente en un espectrómetro Bruker Avance 600 que operaba a un campo electromagnético de 14,1 Teslas equipado con una criosonda TCI de 5 mm.

Todos los experimentos bi y tridimensionales de las muestras de heparanasa fueron realizados a una temperatura de 45 °C y adquiridos con el programa

TOPSPIN (Bruker GmbH). Las anchuras espectrales (SW) de ^1H fueron en todos los casos 14 ppm centrados a una frecuencia de 4,7 ppm (o1); en el caso del ^{15}N , se utilizó un SW de 30 ppm y un o2 de 118 ppm para las muestras de HEP₁₅₈₋₄₁₇ y HEP₂₃₃₋₄₁₇ y de 36 y 117 ppm respectivamente para la muestra de HEP₃₆₋₁₀₉. La anchura espectral de ^{13}C dependía de cada experimento y vienen indicados a continuación además de otros parámetros espectrales.

Los experimentos bidimensionales adquiridos para la asignación de las proteínas fueron los siguientes:

- [^1H - ^{15}N]-HSQC (212)
 - TD (^1H): 2048
 - TD (^{15}N): 756
 - d1: 1,2 s

- [^1H - ^{13}C]-HSQC (212)
 - TD (^1H): 2048
 - TD (^{13}C): 1024
 - d1: 1,2 s
 - SW (^{13}C): 90 ppm
 - o2p: 45 ppm

Para la asignación del esqueleto peptídico, se adquirieron los siguientes experimentos de triple resonancia:

- HNCO (213)
 - TD (^1H): 2048
 - TD (^{15}N): 48
 - TD (^{13}C): 80
 - d1: 1,1 s
 - SW (^{13}C): 18 ppm

- o2p: 176 ppm

- HNCA (213)
 - TD (^1H): 2048
 - TD (^{15}N): 60
 - TD (^{13}C): 120
 - d1: 1,1 s
 - SW (^{13}C): 26,7 ppm
 - o2p: 54 ppm

- CBCA(CO)NH (214)
 - TD (^1H): 2048
 - TD (^{15}N): 60
 - TD (^{13}C): 180
 - d1: 1,1 s
 - SW (^{13}C): 62,5 ppm
 - o2p: 39 ppm

- HNCACB (215)
 - TD (^1H): 2048
 - TD (^{15}N): 48
 - TD (^{13}C): 154
 - d1: 1,1 s
 - SW (^{13}C): 53.4 ppm
 - o2p: 39 ppm

- HN(CA)CO (216)
 - TD (^1H): 2048
 - TD (^{15}N): 48
 - TD (^{13}C): 110
 - d1: 1,1 s

- SW (^{13}C): 18 ppm
- o2p: 176 ppm
- HNHA (217)
 - TD (^1H): 2048
 - TD (^{15}N): 58
 - TD indirecta (^1H): 128
 - d1: 1,1 s

- HNHB (218)
 - TD (^1H): 2048
 - TD (^{15}N): 58
 - TD indirecta (^1H): 128
 - d1: 1,1 s

- HBHA(CO)NH (219)
 - TD (^1H): 2048
 - TD (^{15}N): 58
 - TD indirecta (^1H): 128
 - d1: 1,1 s

- [^1H - ^{15}N]-HSQC-TOCSY (220)
 - TD (^1H): 2048
 - TD (^{15}N): 72
 - TD indirecta (^1H): 182
 - d1: 1,1 s
 - d9: 50 ms

Los experimentos adquiridos para asignar las cadenas laterales fueron los siguientes:

- (H)CCH-TOCSY (221)

- TD (^1H): 2048
- TD (^{13}C): 120
- TD indirecta (^1H): 120
- d1: 1,2 s
- SW (^{13}C): 125 ppm
- o2p: 45 ppm

- H(C)CH-TOCSY (221)

- TD (^1H): 2048
- TD (^{13}C): 190
- TD indirecta (^1H): 110
- d1: 1,2 s
- SW (^{13}C): 125 ppm
- o2p: 45 ppm

A continuación se muestran los experimentos tridimensionales que daban información estructural de la proteína:

- [^1H - ^{15}N]-HSQC-NOESY (222)

- TD directa (^1H): 2048
- TD (^{15}N): 80
- TD indirecta (^1H): 344
- d1: 1,2 s
- d8: para la muestra HEP₁₅₈₋₄₁₇: 110 ms
- d8 para la muestra de HEP₂₃₃₋₄₁₇: 150 ms
- d8 para la muestra de HEP₃₆₋₁₀₉: 150 ms

- [^1H - ^{13}C]-HSQC-NOESY (223)

- TD directa (^1H): 2048

- TD (^{13}C): 80
- TD indirecta (^1H): 222
- d1: 1,2 s
- D8 para todas las muestras: 80 ms
- SW (^{13}C): 90 ppm
- o2p: 45 ppm

Para las muestras de HEP₁₅₈₋₄₁₇, se usaron las versiones TROSY (“Transverse Relaxation-Optimized SpectroscopY”) de experimentos convencionales (224). Por tanto los experimentos tridimensionales utilizados para esta construcción fueron [^1H - ^{15}N]-TROSY-HNCO, [^1H - ^{15}N]-TROSY-HNCA (225, 226), [^1H - ^{15}N]-TROSY-CBCA(CO)NH, [^1H - ^{15}N]-TROSY-HNCACB y [^1H - ^{15}N]-TROSY-HN(CA)CO (227). Todos ellos tenían como base el experimento [^1H - ^{15}N]-TROSY-HSQC (224). Los parámetros para este tipo de experimentos fueron los mismos que anteriormente se han descrito.

Una vez adquiridos todos los experimentos, estos se procesaron con el programa TOPSPIN (Bruker GmbH) utilizando función senoidal con un desplazamiento de campana de seno de 6 para los experimentos bidimensionales y 4 para los experimentos tridimensionales.

2.2.2.-Ensayos de interacción

Los experimentos se adquirieron con los espectrómetros Bruker Avance 600 equipado con una criosonda TCI de 5 mm y Bruker Avance 500 con una sonda TBI de 5 mm y gradientes en el eje z. La adquisición y procesado de estos experimentos se realizó con el programa TOSPIN (Bruker GmbH). La temperatura de medida fue de 27 °C en el caso de water-LOGSY y STD y 45 °C para los [^1H - ^{15}N]-HSQC.

A continuación se detallan los parámetros espectrales para cada tipo de experimento:

- WaterLOGSY

Para la adquisición de este experimento se utilizaron 8 K puntos para un ancho de barrido de 8,6 kHz. El número total de experimentos acumulados fue 1 K para cada waterLOGSY cuando fueron adquiridos con un espectrómetro Bruker Avance 600 y 2 K cuando fueron adquiridos con Bruker Avance 500. Se aplicó un tiempo de relajación de 2,5 s y un tiempo de mezcla de 1,5 s. El pulso selectivo para el agua fue 5 ms en la parte de preparación mientras que la secuencia llamada “excitation sculpting” (228) se realizó con un pulso selectivo de 2 ms. Para procesar el experimento se usó una función exponencial con un ensanchamiento de línea de 3 Hz.

- STD

Un tren de pulsos con forma gaussiana entre 2 y 0,8 ms fue aplicado en este tipo de experimentos con una saturación final de 2,5 s y frecuencias en resonancia y fuera de resonancia de 0,3 y 33 ppm respectivamente. El número de experimentos acumulados fue 1 K cuando fueron adquiridos con un espectrómetro Bruker Avance 600 y 2 K cuando fueron adquiridos con Bruker Avance 500.

- Mapeo de desplazamientos químicos utilizando el [^1H - ^{15}N]-HSQC

Estos experimentos se realizaron todos en un espectrómetro Bruker Avance 600 con una acumulación de experimentos de 16. El resto de los parámetros espectrales ya fueron definidos en la sección 2.2.

2.3.-Asignación del esqueleto y cadenas laterales

El proceso de asignación de una proteína consiste en establecer la relación existente entre cada una de las señales identificadas en los espectros de RMN y los átomos que forman parte de la proteína.

La metodología que se siguió para asignar cada una de las construcciones tenía varias etapas (229).

1. En primer lugar fue necesario conocer la información que daban cada uno de los experimentos. Esta información se resume en la *Tabla 3*.
2. Se identificaron cada una de las señales visibles en el espectro de [^1H - ^{15}N]-HSQC que correspondían con un $-\text{NH}$ de un aminoácido de la proteína. Para ello se marcaron cada una de las señales que aparecían en el espectro que posteriormente se usó como base para los espectros tridimensionales.
3. Identificación de las señales de la tercera dimensión en los espectros tridimensionales (proceso conocido como “peak picking”). Cada valor de $-\text{NH}$ anteriormente identificado en el espectro bidimensional está relacionado con un conjunto de señales que aparecen en la tercera dimensión de los espectros tridimensionales, esta información aparece en la *Tabla 3*.
4. Posteriormente se intentó enlazar aminoácidos que estaban seguidos en la secuencia de la proteína. Para ello se debía encontrar por ejemplo en el experimento HNCA un aminoácido (i) con una señal de $\text{C}\alpha_i$ que coincidiese en desplazamiento químico con el $\text{C}\alpha_{i-1}$ de otro aminoácido (i-1). En este caso y si no hubiera dudas entre varias señales se enlazaba el aminoácido (i-1) con su predecesor (i). Esta unión se podía comprobar con el experimento HNCACB que nos permitía ver si a su vez el desplazamiento químico de $\text{C}\beta_i$ coincidía con el de $\text{C}\beta_{i-1}$. También se pudieron hacer correlaciones secuenciales con los desplazamientos químicos de protones usando los experimentos HNHA, HNHB Y HBHA(CO)NH. De esta forma se iban obteniendo fragmentos de aminoácidos dentro de la secuencia de la proteína.

Tabla 3. Información aportada por cada uno de los experimentos tridimensionales utilizados en la asignación de la proteína.

	DIMENSIONES (x-y-z)	INFORMACIÓN (3ª DIMENSION)
HNCO	$^1\text{H} - ^{15}\text{N} - ^{13}\text{C}$	CO_{i-1}
HNCA	$^1\text{H} - ^{15}\text{N} - ^{13}\text{C}$	$\text{C}\alpha_i, \text{C}\alpha_{i-1}$
CBCA(CO)NH	$^1\text{H} - ^{15}\text{N} - ^{13}\text{C}$	$\text{C}\alpha_{i-1}, \text{C}\beta_{i-1}$
HNCACB	$^1\text{H} - ^{15}\text{N} - ^{13}\text{C}$	$\text{C}\alpha_i, \text{C}\beta_i, \text{C}\alpha_{i-1}, \text{C}\beta_{i-1}$
HNCACO	$^1\text{H} - ^{15}\text{N} - ^{13}\text{C}$	$\text{CO}_i, \text{CO}_{i-1}$
HNHA	$^1\text{H} - ^{15}\text{N} - ^1\text{H}$	$\text{H}\alpha_i$
HNHB	$^1\text{H} - ^{15}\text{N} - ^1\text{H}$	$\text{H}\beta_i$
HBHA(CO)NH	$^1\text{H} - ^{15}\text{N} - ^1\text{H}$	$\text{H}\alpha_{i-1}, \text{H}\beta_{i-1}$
$[^{15}\text{N}, ^1\text{H}]$ -HSQC-NOESY	$^1\text{H} - ^{15}\text{N} - ^1\text{H}$	H (1,6-6 Å)
$[^{13}\text{C}, ^1\text{H}]$ -HSQC-NOESY	$^1\text{H} - ^{15}\text{N} - ^1\text{H}$	H (1,6-6 Å)
(H)CCH-TOCSY	$^1\text{H} - ^{13}\text{C} - ^{13}\text{C}$	C (cadena lateral)
H(C)CH-TOCSY	$^1\text{H} - ^{13}\text{C} - ^1\text{H}$	H (cadena lateral)

- Identificación y asignación de aminoácidos. A partir de los desplazamientos químicos típicos para las señales de $\text{C}\alpha$, $\text{C}\beta$, $\text{H}\alpha$ y $\text{H}\beta$, se identificaba la naturaleza de algunos residuos. Si la naturaleza de cada uno de los aminoácidos para un fragmento dado coincidía con algún trozo de la secuencia de la proteína se asignaba. Si había más de una opción posible, se debía buscar más información antes de asignar el fragmento.
- Asignación de las cadenas laterales. Para ello se usaban los experimentos (H)CCH-TOCSY que permiten asignar la cadena lateral de C ($\text{C}\gamma$, $\text{C}\delta$, etc)

y el H(C)CH-TOCSY que permite asignar la cadena lateral de H (H_γ , H_δ , etc).

Para el proceso de asignación se utilizó el programa CARA (Computer-Aided Resonance Assignment) (www.nmr.ch) como soporte informático.

La *Figura 4* resume todo el proceso utilizado para la asignación de todas las construcciones de Heparanasa.

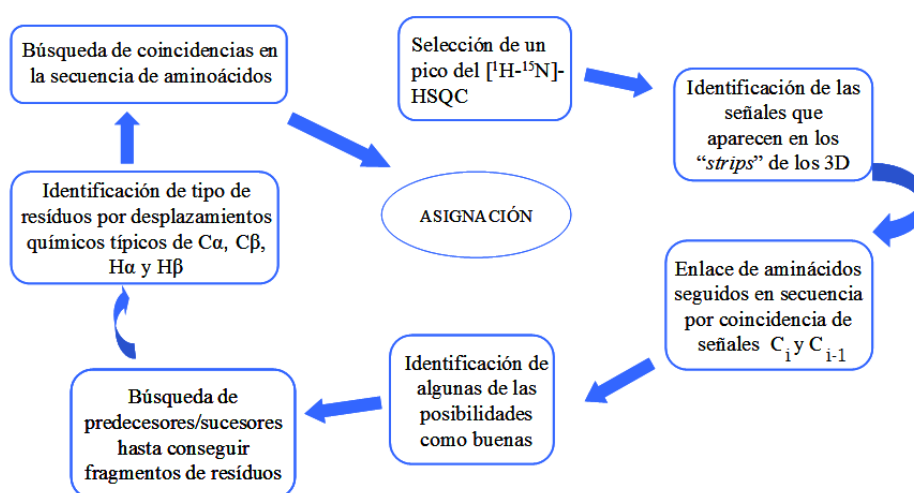


Figura 4. Esquema de los pasos utilizados en la asignación de Heparanasa

2.4.-Cálculo de estructura secundaria

El cálculo de estructura secundaria fue realizado con el programa TALOS (<http://www.nmr.chem.uu.nl/~abonvin/tutorials/Talos-Data/talos.html>) (230). Este programa calcula restricciones de ángulos de torsión de proteínas (ángulos Phi y Psi) a partir de los desplazamientos químicos de los átomos que forman el esqueleto de la proteína (H_α , C_α , C_β , CO y N) junto con información disponible en bases de datos como BMRB y PDB. Este programa se basa en que los desplazamientos químicos del esqueleto de las proteínas son extraordinariamente sensibles a la conformación local y por eso muchas proteínas homólogas muestran patrones similares de estructura secundaria.

Por tanto, este programa precisa de información sobre los desplazamientos químicos de todos los aminoácidos y de la secuencia de la proteína. Para ello, se utilizó un fichero .prot (obtenido a partir de CARA) con los desplazamientos químicos de los átomos necesarios anteriormente mencionados y que fue transformado en un fichero .tab (formato requerido para ejecutar el programa TALOS). Después del cálculo se obtuvo un fichero .aco con la información de los ángulos Phi y Psi de la proteína.

2.5.-Cálculo de estructura tridimensional.

Para el cálculo automático de NOEs y para el cálculo de la estructura tridimensional de la construcción de heparanasa HEP₁₅₈₋₄₁₇ se utilizó el programa Cyana 2.1 (<http://www.enmr.eu/webportal/cyana.html>). Antes de llevar a cabo el cálculo de la estructura se corrigieron los desplazamientos químicos de carbono de las señales del espectro 3D [¹³C, ¹H]-HSQC-NOESY según el calibrado de temperatura para el espectrómetro Bruker Avance 900 MHz del centro de Resonancia Magnética Nuclear CERM de Florencia donde se realizaron los experimentos. Por esta razón se sumó 2,75 ppm a todos los desplazamientos de carbono de este experimento. Esto se realizó utilizando como soporte informático el programa de asignación de espectros de RMN, CARA (Computer-Aided Resonance Assignment) (www.nmr.ch) ejecutando la macro *ShiftSpinAndAliases*. Posteriormente se marcaron cada una de las señales en los espectros y a continuación se generaron los ficheros que contenían el listado de desplazamientos químicos para cada una de estas señales marcadas en los espectros 3D [¹³C, ¹H]-HSQC-NOESY y [¹⁵N, ¹H]-HSQC-NOESY.

Los ficheros de entrada que se utilizaron para el cálculo de estructura fueron los siguientes:

- Heparanasa.seq: contiene la secuencia de la proteína. Este fichero se obtuvo con el programa CARA al ejecutar la macro *AssignmentReport*.

- ^{13}C .peaks: contiene los desplazamientos químicos de las señales del espectro 3D [^{13}C , ^1H]-HSQC-NOESY. Este fichero se obtuvo con el programa CARA
- ^{15}N .peaks: contiene los desplazamientos químicos de las señales del espectro 3D [^{15}N , ^1H]-HSQC-NOESY. Este fichero se obtuvo utilizando el programa CARA.
- Angulos.aco: contiene los ángulos Phi y Psi de la proteína. Se obtiene con el programa TALOS como se ha explicado en el apartado 5 de materiales y métodos.
- ^{13}C .prot y ^{15}N .prot que contienen la asignación de la proteína. Estos ficheros se obtuvieron mediante el programa CARA al ejecutar la macro *AssignmentReport*.

Las macros que se ejecutaron para poder empezar el cálculo de estructura fueron: init.cya y CALC.cya.

Las estructuras resultantes se visualizaron con el programa Molmol 2K.2.

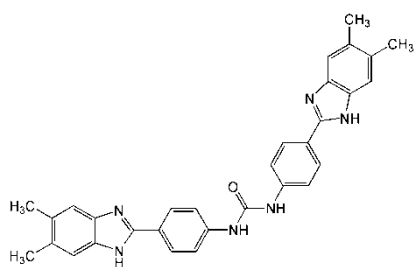
3.-MODELADO MOLECULAR*

3.1.-Desarrollo del modelo farmacofórico

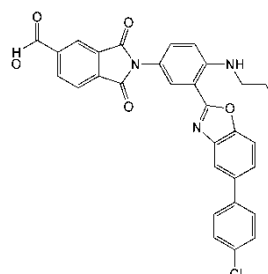
Según la definición de la IUPAC, un modelo farmacofórico es un conjunto de rasgos estéricos y electrónicos necesarios para asegurar las óptimas interacciones supramoleculares con una diana biológica específica y para bloquear su respuesta biológica. La generación del modelo farmacofórico se hizo a partir de los 27 inhibidores conocidos de heparanasa más activos y más pequeños escogidos de Hammond *et al.* (190) (Tabla 4). Los compuestos se dibujaron con el programa MDL Isis Draw 2.5 (www.mdl.com) y se exportaron al software “Molecular Operating Environment” (MOE) (www.chemcomp.com). El desarrollo del modelo farmacofórico en MOE permite la caracterización de átomos por sus características químicas (unidos a hidrógenos aceptores o

donadores de electrones, carácter aromático o naturaleza hidrofóbica), la generación posterior de hipótesis farmacofóricas y la aplicación de estas hipótesis a bases de datos conformacionales. Esto permite encontrar compuestos con una conformación determinada que cumplan este modelo. En este caso, antes de desarrollar el modelo farmacofórico, se generaron múltiples conformaciones de los 27 inhibidores. Para ello se utilizó un campo de fuerza MMFF94 y una energía de corte de 4 kcal/mol con un máximo de 250 conformaciones por compuesto. Las hipótesis farmacofóricas se generaron a partir del alineamiento de varios subconjuntos de estos 27 inhibidores de heparanasa. La validez de los modelos que se generaron se verificó aplicando el modelo al conjunto de compuestos formado por los 27 inhibidores conocidos y la quimioteca Prestwick en la cual un 90% de los compuestos son fármacos que han llegado a mercado (www.prestwickchemical.com).

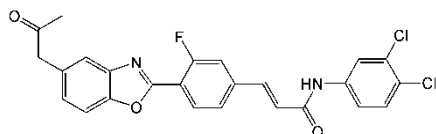
Tabla 4. 27 inhibidores conocidos de heparanasa escogidos de Hammond et al. (190).



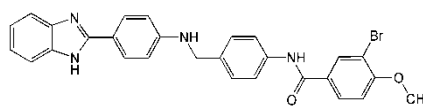
INH-1 IC₅₀ (0.075 μM)



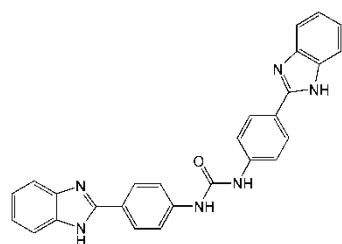
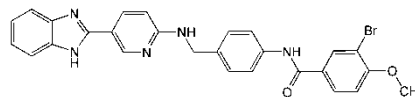
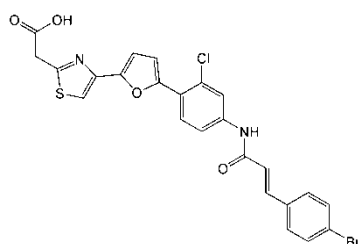
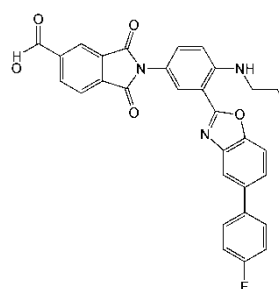
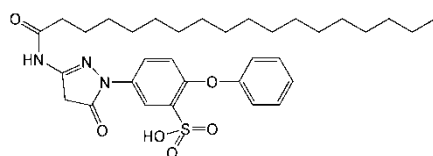
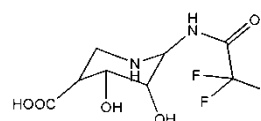
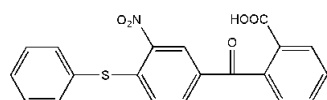
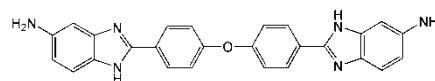
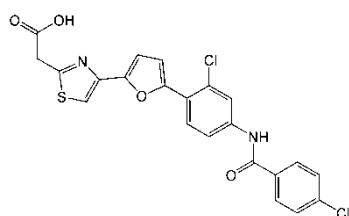
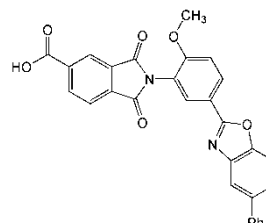
INH-2 IC₅₀ (0.20 μM)

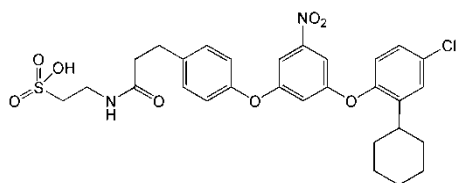
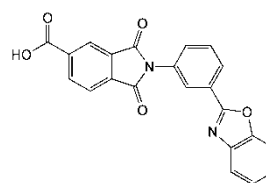
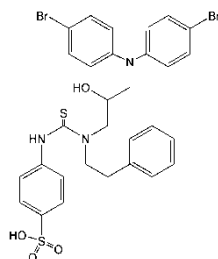
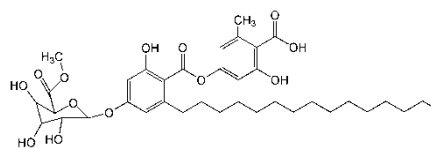
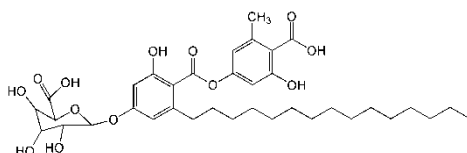
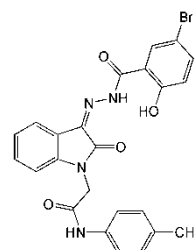
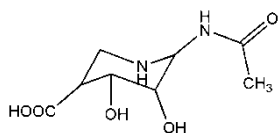
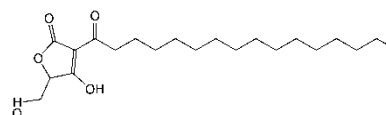
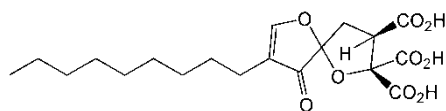
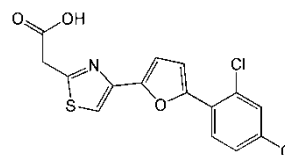
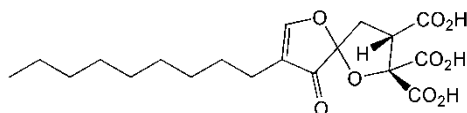
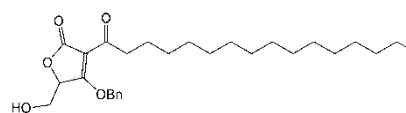


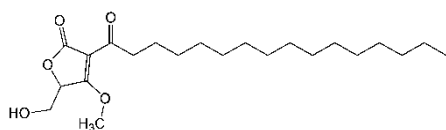
INH-3 IC₅₀ (0.20 μM)



INH-4 IC₅₀ (0.23 μM)

**INH-5** IC₅₀ (0.27 μM)**INH-6** IC₅₀ (0.29 μM)**INH-7** IC₅₀ (0.40 μM)**INH-8** IC₅₀ (0.50 μM)**INH-9** IC₅₀ (0.13-19 μM)**INH-10** IC₅₀ (1 μM)**INH-11** IC₅₀ (1-10 μM)**INH-12** IC₅₀ (1.50-36 μM)**INH-13** IC₅₀ (2.50 μM)**INH-14** IC₅₀ (3 μM)

**INH-15** IC₅₀ (5-45 μM)**INH-16** IC₅₀ (8 μM)**INH-17** IC₅₀ (8-25 μM)**INH-18** IC₅₀ (10 μM)**INH-19** IC₅₀ (10 μM)**INH-20** IC₅₀ (10-26 μM)**INH-21** IC₅₀ (12 μM)**INH-22** IC₅₀ (17 μM)**INH-23** IC₅₀ (24-32 μM)**INH-24** IC₅₀ (25 μM)**INH-25** IC₅₀ (26-36 μM)**INH-26** IC₅₀ (>100 μM)

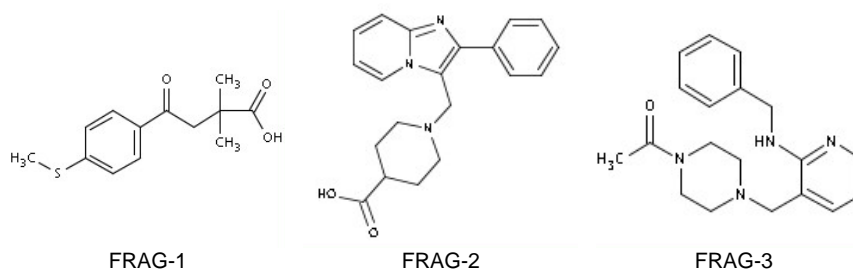


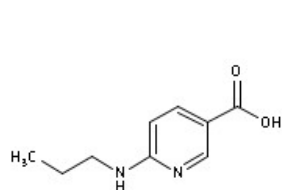
INH-27 IC₅₀ (>100 μM)

3.2.-Cribado de fragmentos con el modelo farmacofórico

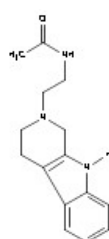
El modelo farmacofórico previamente determinado se aplicó a una quimioteca comercial de 686 fragmentos (CEREP, www.cerep.fr) usando el mismo protocolo anteriormente descrito. Esta quimioteca cumplía la llamada “regla del tres” (193) y contenía una alta diversidad química determinada por el algoritmo de Tanimoto de 0,3. Para ello se usaron descriptores llamados “MACCS keys” incluidos en el programa MOE. De esta forma, esta quimioteca de fragmentos fue clasificada en función del número de puntos del modelo farmacofórico que cumplía. Finalmente se eligieron 100 fragmentos que cumplieran total o parcialmente el modelo y con alta diversidad química (índice de Tanimoto de 0,27) para realizar el cribado por Resonancia Magnética Nuclear. Las estructuras de estos 100 fragmentos se muestran en la Tabla 5.

Tabla 5. Conjunto de 100 fragmentos seleccionados para el cribado por Resonancia Magnética Nuclear

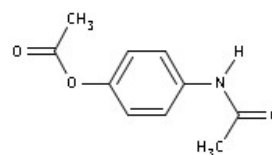




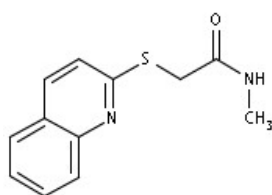
FRAG-4



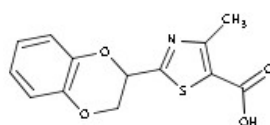
FRAG-5



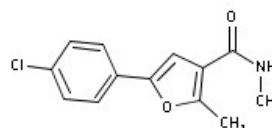
FRAG-6



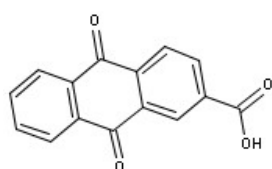
FRAG-7



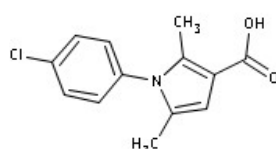
FRAG-8



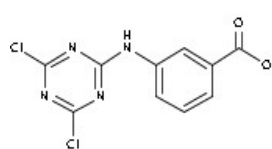
FRAG-9



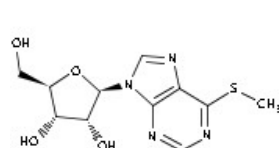
FRAG-10



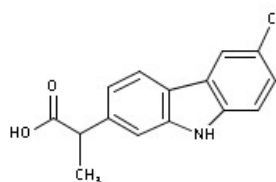
FRAG-11



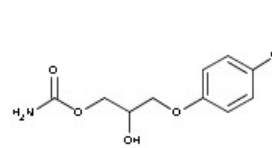
FRAG-12



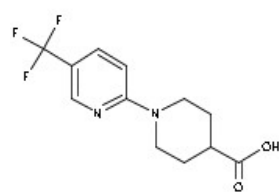
FRAG-13



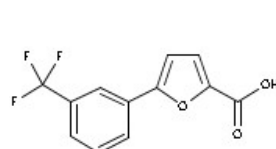
FRAG-14



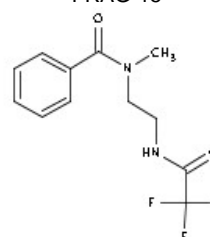
FRAG-15



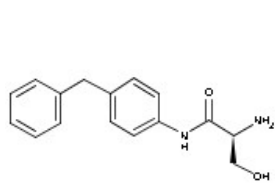
FRAG-16



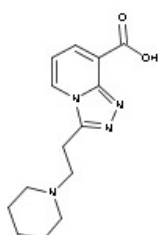
FRAG-17



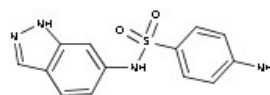
FRAG-18



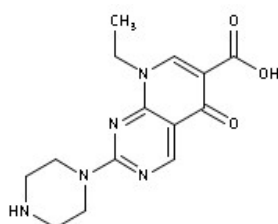
FRAG-19



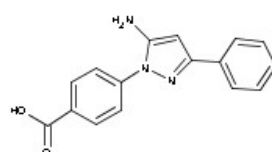
FRAG-20



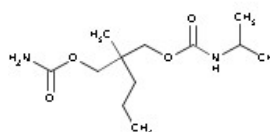
FRAG-21



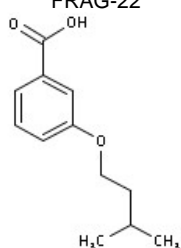
FRAG-22



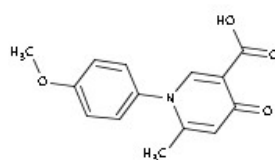
FRAG-23



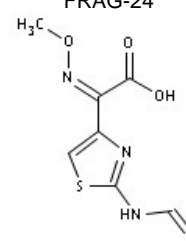
FRAG-24



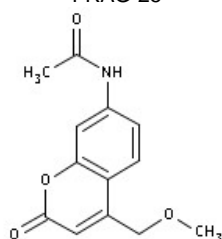
FRAG-25



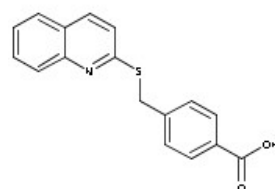
FRAG-26



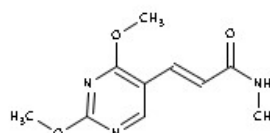
FRAG-27



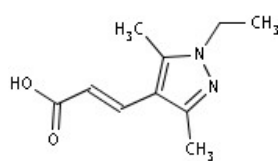
FRAG-28



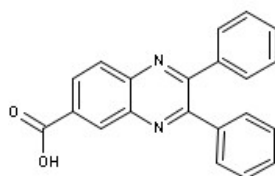
FRAG-29



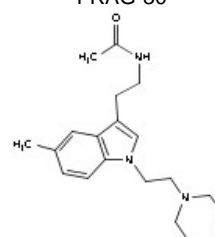
FRAG-30



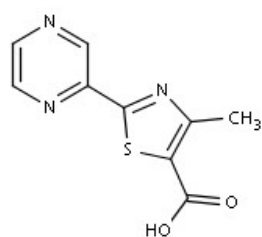
FRAG-31



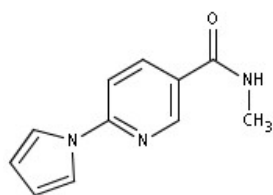
FRAG-32



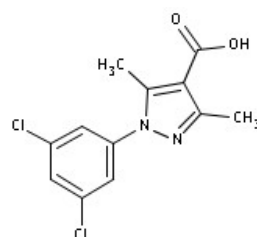
FRAG-33



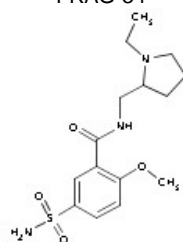
FRAG-34



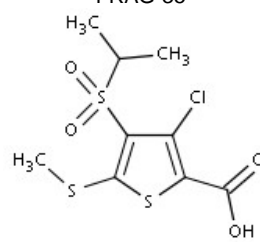
FRAG-35



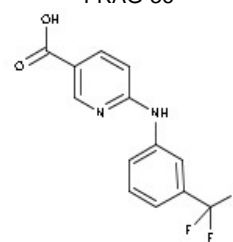
FRAG-36



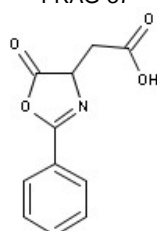
FRAG-37



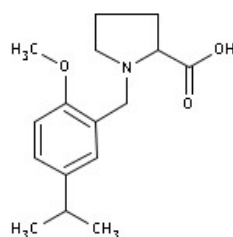
FRAG-38



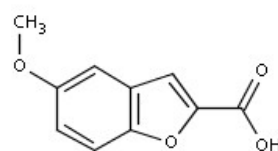
FRAG-39



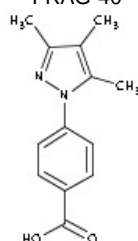
FRAG-40



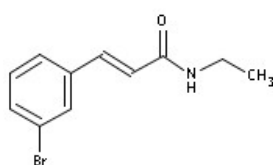
FRAG-41



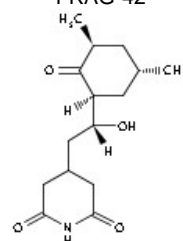
FRAG-42



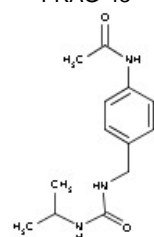
FRAG-43



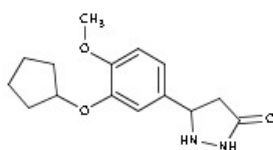
FRAG-44



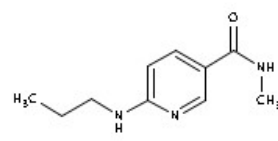
FRAG-45



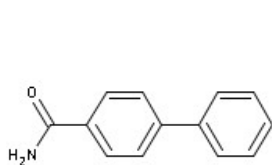
FRAG-46



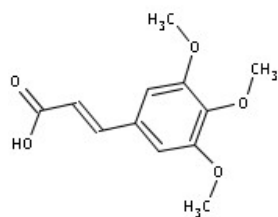
FRAG-47



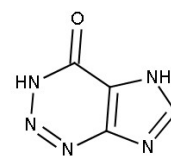
FRAG-48



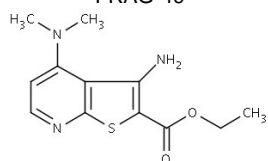
FRAG-49



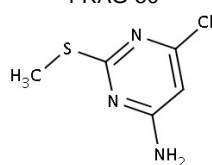
FRAG-50



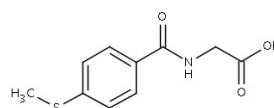
FRAG-51



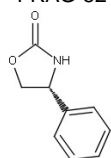
FRAG-52



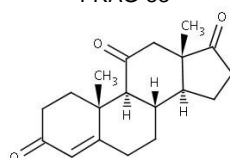
FRAG-53



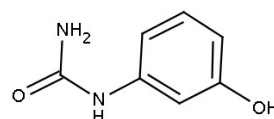
FRAG-54



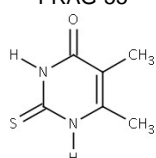
FRAG-55



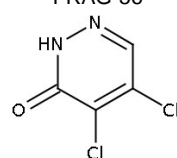
FRAG-56



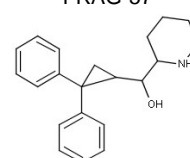
FRAG-57



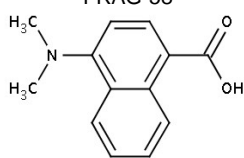
FRAG-58



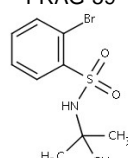
FRAG-59



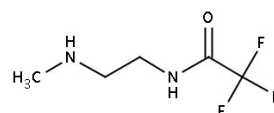
FRAG-60



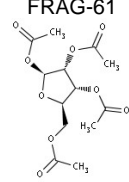
FRAG-61



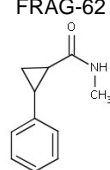
FRAG-62



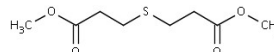
FRAG-63



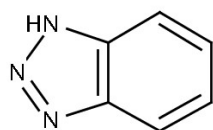
FRAG-64



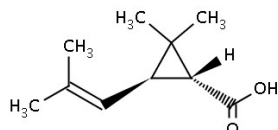
FRAG-65



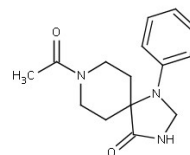
FRAG-66



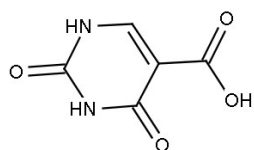
FRAG-67



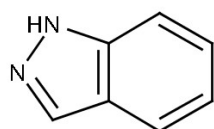
FRAG-68



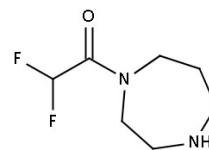
FRAG-69



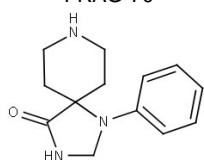
FRAG-70



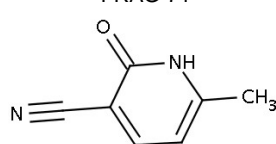
FRAG-71



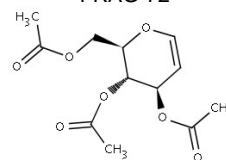
FRAG-72



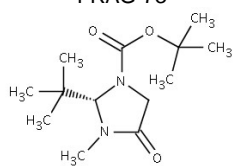
FRAG-73



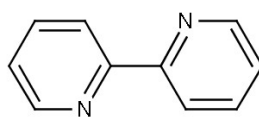
FRAG-74



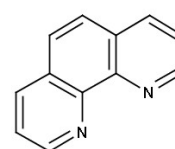
FRAG-75



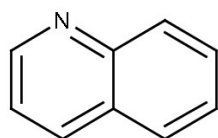
FRAG-76



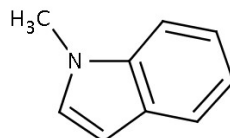
FRAG-77



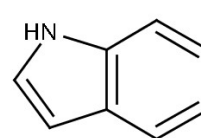
FRAG-78



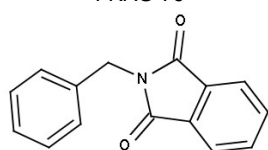
FRAG-79



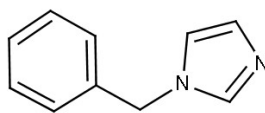
FRAG-80



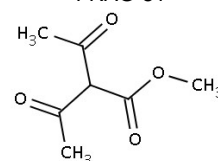
FRAG-81



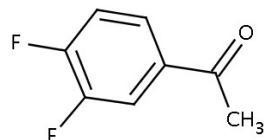
FRAG-82



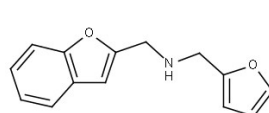
FRAG-83



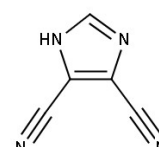
FRAG-84



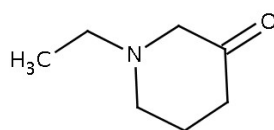
FRAG-85



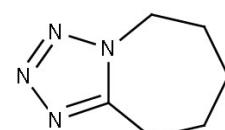
FRAG-86



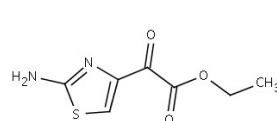
FRAG-87



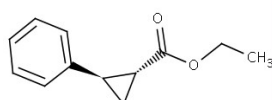
FRAG-88



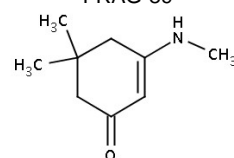
FRAG-89



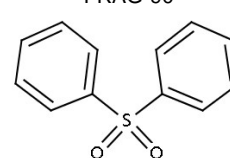
FRAG-90



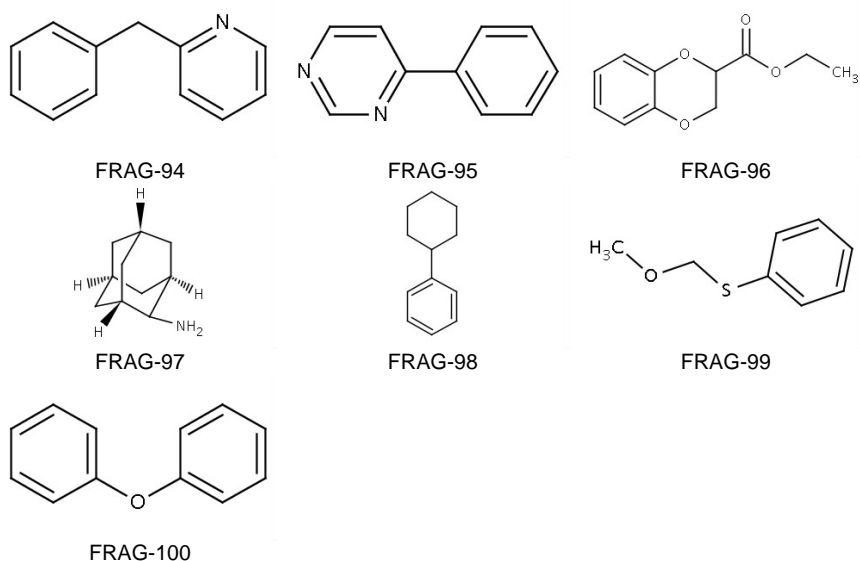
FRAG-91



FRAG-92



FRAG-93



3.3.-Modelo de homología de heparanasa

El modelo de homología de heparanasa utilizado para posteriores estudios se obtuvo mediante el programa MODELLER (<http://www.salilab.org/modeller>) (231-234). Para ello, se utilizó como punto de partida una proteína cuya estructura era conocida y cuya secuencia mostraba cierta similitud con la heparanasa en el sitio catalítico de la enzima. En este caso la proteína de referencia fue α -L-arabinofuranosidase de *Geobacillus stearothermophilus* T-6, que pertenece a la familia 51 de las glicosidasas. MODELLER automáticamente calcula un modelo basado en las restricciones espaciales entre los heteroátomos de la proteína empleada como punto de partida. Para la utilización de este programa se usó la interfase gráfica ModWeb (<https://modbase.compbio.ucsf.edu/scgi/modweb.cgi>).

3.4.-Docking

Los ensayos de “docking” fueron llevados a cabo con el conjunto de herramientas de modelización de OpenEye (versión 2.1) (<http://www.eyesopen.com>). El módulo “OpenEye docking” permite la predicción de la orientación y conformación más favorable de un ligando en el sitio de unión de un receptor para formar un complejo estable. Para ello se utilizaron la estructura tridimensional de heparanasa que previamente se había generado por homología de secuencia y dos quimiotecas virtuales de compuestos: una formada por los 27 inhibidores conocidos de heparanasa (190) y la “Prestwick Chemical Library”, una base de datos comercial que integra 1120 estructuras de fármacos (<http://www.prestwickchemical.com/>).

Este ensayo se llevó a cabo en varias etapas:

- Generación de distintas conformaciones de cada uno de los compuestos de las quimiotecas. Para ello se utilizó el módulo OMEGA (Optimized Molecular Ensemble Generation Application) de OpenEye (235).
- Definición del centro activo (236). Posteriormente se definió la región teórica de la proteína con la que los compuestos pudieran interactuar. Para ello se utilizó el módulo “FRED” de OpenEye (Fast Rigid and Exhaustive Docking), un programa específicamente concebido para el estudio de las interacciones potenciales de ligandos con receptores proteicos, concretamente la utilidad gráfica “FRED receptor”. En este caso se utilizó el sitio catalítico de la heparanasa como centro de esta región que se extendió en un cubo de 1500 Å³. A continuación se definió el volumen comprendido entre los residuos que forman este sitio catalítico, en este caso los dos ácidos glutámicos activos de la enzima Glu²²⁵ y el Glu³⁴³ y los dos sitios de unión a heparán sulfato Lys¹⁵⁸-Asn¹⁶² y Gln²⁷⁰-Lys²⁸⁰ para excluir posteriormente aquellas poses de los compuestos (conformación de los compuestos después de hacer el docking) que se salían de este volumen.

*El trabajo experimental descrito en este apartado ha sido realizado en colaboración con el Dr. Rafael Gozalbes.

4-ENSAYOS DE INTERACCIÓN POR RESONANCIA DE PLASMA SUPERFICIAL (SPR)

Todos los experimentos se realizaron con un instrumento Biacore T100. Estos ensayos se basan en el fenómeno de resonancia de plasma superficial y permiten medir de forma cuantitativa la interacción entre heparanasa y diferentes ligandos a tiempo real.

La resonancia de plasma superficial es un fenómeno que ocurre dentro de una fina capa de un material conductor que se encuentra situado en una interfase entre dos medios de diferente índice refractivo. En sistemas Biacore los dos medios son el vidrio del *sensor chip* y la solución de la muestra. La película conductora es una fina capa de oro que se encuentra en la superficie del *sensor chip*. Si la luz incide sobre esta interfase con un cierto ángulo y una cierta energía (longitud de onda), se provoca la excitación de los plasmones (cuanto de oscilación del plasma) que se encuentran en la película de oro provocando un descenso en la intensidad de la luz reflejada. Este cambio es proporcional a la masa del complejo formado entre la proteína y el ligando unido, por lo tanto la respuesta será distinta para cada uno de los compuestos que se unan a la proteína (Fig. 5). Pueden medirse diferencias de masa por debajo de pocos picogramos por milímetro cuadrado de superficie, lo que corresponde a concentraciones del rango de picomolar a nanomolar.

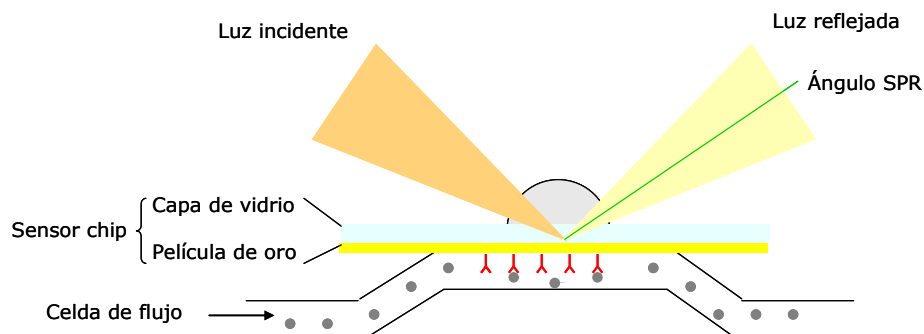


Figura 5. Esquema del fenómeno de resonancia de plasma superficial (SPR).

Estos experimentos de interacción se realizaron en diferentes etapas:

- Inmovilización de la proteína a la superficie del sensor
- Inyección del ligando
- Regeneración

4.1.- Inmovilización de heparanasa a la superficie del sensor

La inmovilización se realizó a 25 °C sobre un *sensor chip* del tipo CM5. Este sensor pertenece a la familia de sensores CM que contienen una matriz de dextrano carboximetilada unida covalentemente a la superficie de la película de oro; es el más utilizado y se diferencia de otros chips CM por el nivel de carboximetilación y la longitud de sus cadenas de dextranos.

Antes de inmovilizar la proteína hubo que activar la superficie del chip y esto se realizó inyectando una mezcla de 115 mg/mL de N-hidroxisuccinimida (NHS) y 750 mg/mL hidrocloreuro de carbodiimida (EDC) durante un tiempo de 7 min con un flujo de 10 μ L/min. Los grupos carboxilos que se encontraban en la matriz de dextrano reaccionaban con la mezcla de EDC y NHS para dar grupos ésteres succinimidas que a su vez reaccionaban espontáneamente con grupos aminos y otros grupos nucleófilos permitiendo la inmovilización directa de biomoléculas que contenían esos grupos (Fig.6).

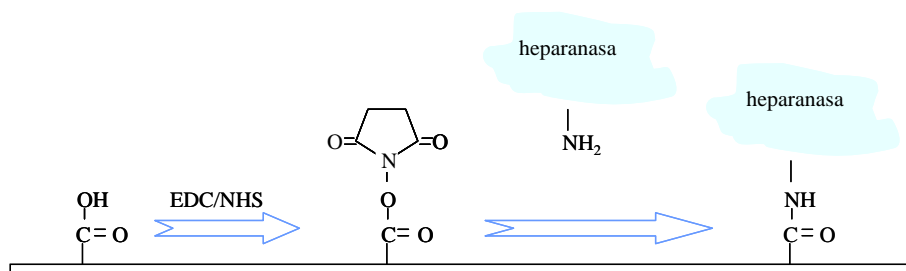


Figura 6. Reacción de activación de la superficie del chip

Posteriormente la proteína fue diluida con tampón 25 mM fosfato sódico pH 6, 50 mM NaCl hasta una concentración final de 50 µg/mL (para ello se partió de un stock de concentración entre 0,5 y 1 mM) y se inyectó hasta conseguir la inmovilización deseada. La cantidad de heparanasa inmovilizada en el chip variaba entre 2500 y 8000 RU dependiendo del tipo de experimento y del peso molecular de la proteína y del ligando. La respuesta del instrumento venía dada en RU (unidades de resonancia) que era directamente proporcional a la concentración de proteína inmovilizada. Una respuesta de 1 RU era equivalente al cambio de concentración en la superficie del sensor de 1 pg/mm².

Una vez que la proteína estaba inmovilizada en el chip, se procedió a bloquear los grupos ésteres activos de succinimidias donde la proteína no se había unido. Para ello se utilizó una solución 1 M de etanolamina de pH 8,5. Se inyectó con un flujo de 10 µL/min durante 700 s.

4.2.- Inyección del ligando

El Biacore requiere un flujo continuo de un tampón durante todo el experimento. El tampón utilizado fue 10 mM HEPES (pH 7,4), 150 mM NaCl, 50 µM EDTA y 0,005% (v/v) de detergente P20.

Los compuestos que se midieron iban disueltos en este mismo tampón para evitar problemas de diferencias en la luz refractada. Cuando los compuestos eran insolubles en agua, las muestras contenían un 2% de DMSO que también se añadió al tampón utilizado para correr los experimentos.

Para el caso de la suramina y la heparina que son compuestos más grandes con una constante de disociación baja se utilizaron concentraciones de 0,01, 0,025, 0,05, 0,1, 0,3, 0,5, 1, 2 y 5 µM de ligando. Sin embargo para compuestos pequeños como en el caso de los fragmentos, las concentraciones utilizadas fueron 0,001, 0,002, 0,005, 0,01, 0,025, 0,05, 0,1, 0,3, 0,5 y 1 mM. En todos los experimentos una de las muestras se hacía por duplicado para comprobar la reproducibilidad del experimento, en el primer caso fue 0,1 µM y en el segundo caso fue 0,1 mM. Además cada tres inyecciones se volvía a inyectar solo el

tampón como control. El flujo que se empleó para todos los compuestos fue 30 $\mu\text{L}/\text{min}$ y el tiempo de contacto variaba entre 120 y 400 s según el compuesto utilizado. La temperatura de medida fue de 25 °C en todos los casos.

Para medir la constante de disociación de un gran número de compuestos se utilizaron placas de 96 pocillos. Después de la inyección de las diferentes concentraciones de ligando y con el objetivo de comprobar la estabilidad de la proteína unida covalentemente al chip, se inyectaba un compuesto control que en este caso fue el FRAG-86 cuya K_D se había medido anteriormente.

4.3.-Regeneración.

El objetivo de la regeneración era eliminar todo el ligando unido a la proteína después de cada medida. Para ello se usó una disolución de 10 mM de glicina pH 2,5 durante 30 s a un flujo de 30 $\mu\text{L}/\text{min}$.

En el caso de ligandos pequeños para los que el tiempo de disociación era más pequeño y por tanto las constantes cinéticas de disociación k_{off} (s^{-1}) eran más altas, no era necesario utilizar la regeneración para eliminar el ligando unido a la proteína. En estos casos, se dejaron tiempos de disociación proteína-ligando de 900 segundos con lo que se conseguía que la respuesta en el sensograma fuera 0 RU. Este método era menos agresivo para la proteína y permitía que la proteína unida covalentemente al chip permaneciera más tiempo estable.

El análisis de todos los resultados se llevó a cabo con el software suministrado por Biacore T100 como soporte informático.

5.-ESTRATEGIA DE BÚSQUEDA Y SELECCIÓN DE NUEVOS INHIBIDORES POTENCIALES DE HEPARANASA

5.1.- Selección de inhibidores a partir de la quimioteca de fragmentos

La quimioteca de fragmentos formada por 686 compuestos fue clasificada en función del número de puntos del modelo farmacofórico que cumplía. Finalmente se eligieron 100 fragmentos que cumplían total o parcialmente el modelo para realizar el cribado por Resonancia Magnética Nuclear. Las

estructuras de estos 100 fragmentos se muestran en la Tabla 5 de materiales y métodos.

Después de llevar a cabo el cribado de estos 100 fragmentos por RMN (utilizando experimentos water-LOGSY, STD y experimentos de STD de competición con suramina) y medir la constante de disociación por SPR, se eligió el FRAG-86 como punto de partida para buscar análogos que tuvieran una interacción más fuerte con la construcción HEP₁₅₈₋₄₁₆ de heparanasa. Todo este proceso está representado en la Fig. 7.

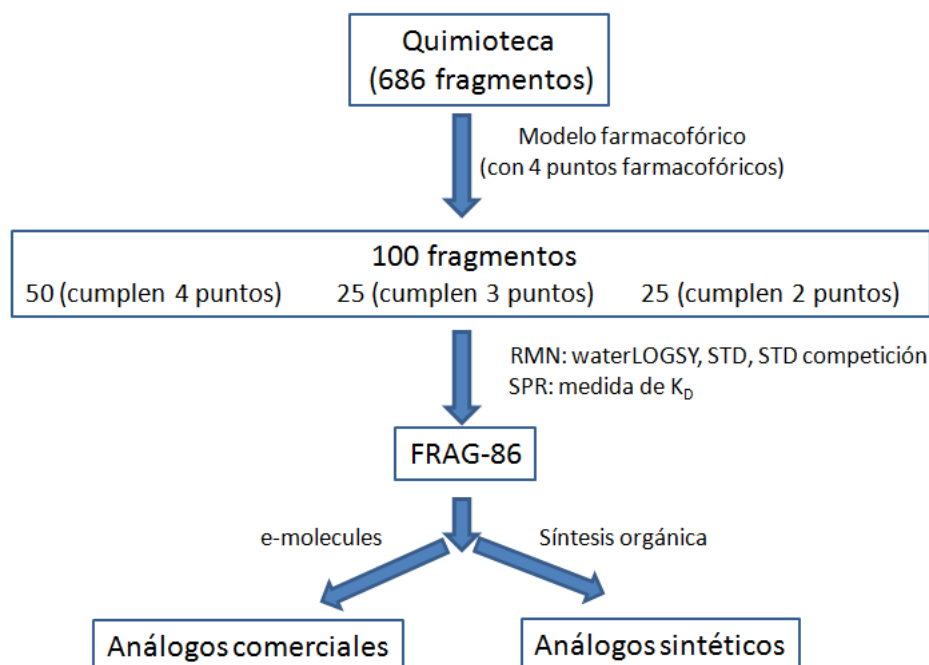


Figura 7. Estrategia de selección de inhibidores a partir de la quimioteca de fragmentos CEREP.

Para la búsqueda de análogos del FRAG-86 se desarrollaron dos estrategias. Por un lado se llevó a cabo la búsqueda de análogos comerciales (casa comercial ChemBridge) y se realizó a través de la página web <http://www.emolecules.com>, usando la estructura de FRAG-86 como referencia. Estos análogos están representados en la Tabla 6.

Por otro lado, se propuso la síntesis de compuestos que contenían este fragmento por duplicado con las siguientes combinaciones: cabeza-cabeza (COMP-1), cabeza-cola (COMP-2) y cola-cola (COMP-3) unidos mediante “linkers” de distintos tamaños (Fig. 8). Estos compuestos están siendo sintetizados en el laboratorio de Síntesis Orgánica dirigido por el Dr. Santos Fustero del Centro de Investigación Príncipe Felipe. Hasta la fecha se ha sintetizado solo COMP-1.

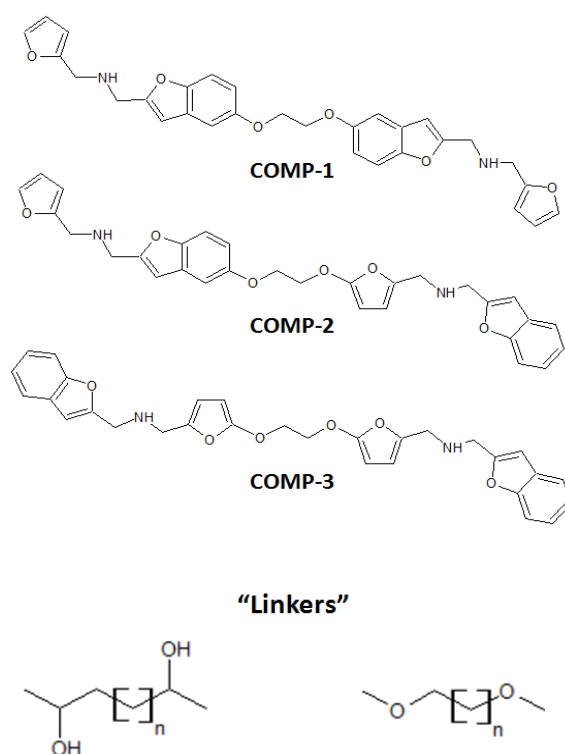
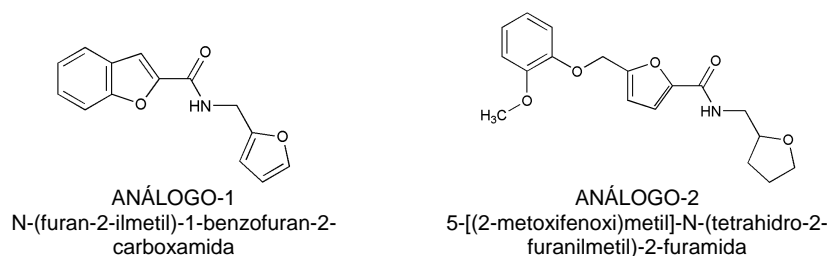
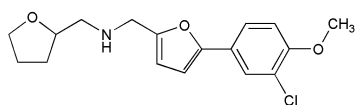


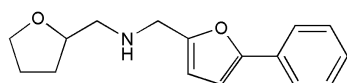
Figura 8. Compuestos sintéticos análogos del FRAG-86.

Tabla 6. Análogos comerciales del FRAG-86

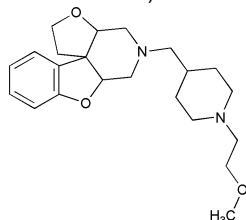




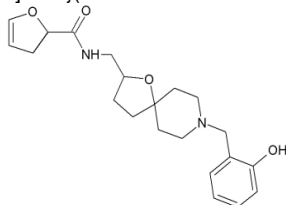
ANALOGO-3
1-[5-(3-cloro-4-metoxifenil)furan-2-il]-N-(tetrahydrofuran-2-ilmetil)metanamina



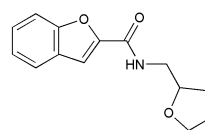
ANALOGO-5
N-[(5-fenil-2-furil)metil]-N-(tetrahydro-2-furanilmetil)amina



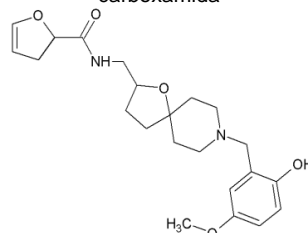
ANALOGO-7
(1-benzofuran-2-ilmetil){[1-(2-metoxietil)-4-piperidinil]metil}(tetrahydro-2-furanilmetil)amina



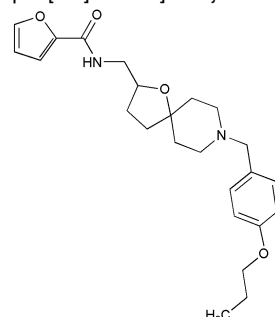
ANÁLOGO-9
N-[[8-(2-hidroxibenzil)-1-oxa-8-azaspiro[4.5]dec-2-il]metil]-2-furamida



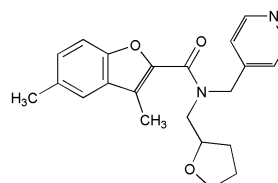
ANALOGO-4
N-(tetrahydro-2-furanilmetil)-1-benzofuran-2-carboxamida



ANALOGO-6
N-[[8-(2-hidroxi-5-metoxibenzil)-1-oxa-8-azaspiro[4.5]dec-2-il]metil]-2-furamida



ANALOGO-8
N-[[8-(2-hidroxi-5-metoxibenzil)-1-oxa-8-azaspiro[4.5]dec-2-il]metil]-2-furamida



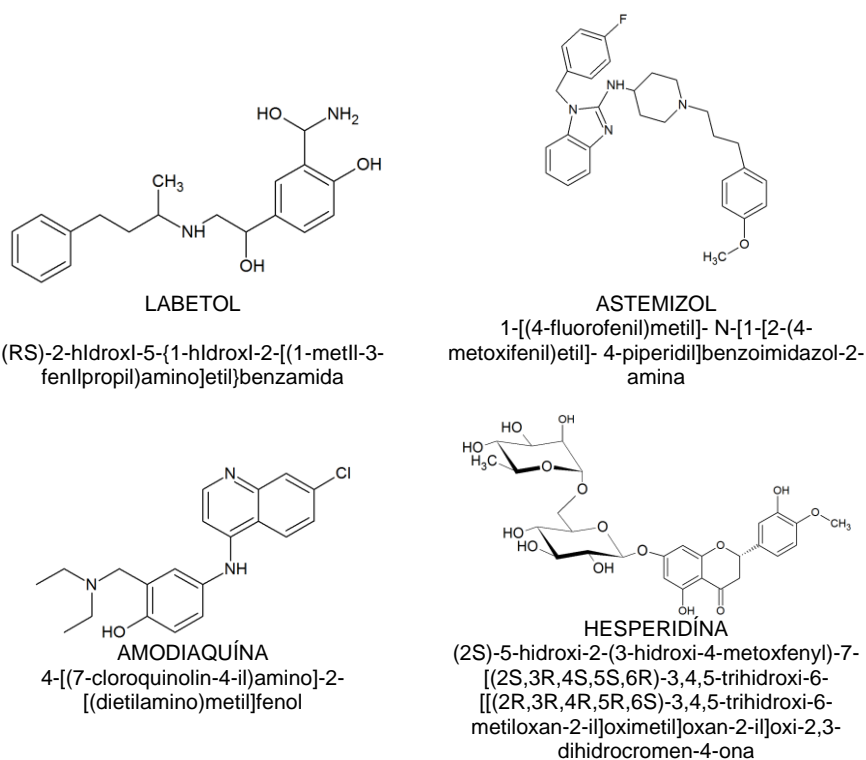
ANALOGO-10
3,5-dimetil-N-(4-piridinilmetil)-N-(tetrahydro-2-furanilmetil)-1-benzofuran-2-carboxamida

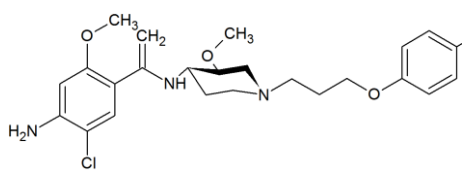
5.2.-Selección de inhibidores procedentes de la librería PRESTWICK

La quimioteca de compuestos PRESTWICK formada por 1120 moléculas fue clasificada en función del número de puntos del modelo farmacofórico que cumplía y según la clasificación que se obtuvo después de hacer docking con estos compuestos y el modelo por homología de secuencia de heparanasa.

Finalmente se eligieron los 7 mejores compuestos para llevar a cabo los experimentos de interacción por RMN y SPR. Estos 7 compuestos fueron: astemizol, amodiaquina, hesperidina, cisaprida, naringina, metergolina y labetol. Todos ellos se obtuvieron de SIGMA-ALDRICH y sus estructuras están representadas en la Tabla 7. Finalmente, y basándonos en los resultados obtenidos se eligió la amodiaquina como punto de partida para la búsqueda de análogos (Tabla 8) con mejor afinidad química por la proteína. Todo este procedimiento está esquematizado en la Fig. 9. Estos análogos sintéticos fueron sintetizados en el laboratorio de la Dra. Patricia Melnyk del Institute Pasteur (CNRS) de Lille (Francia).

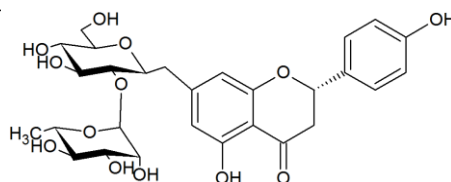
Tabla 7. Compuestos seleccionados para RMN y SPR





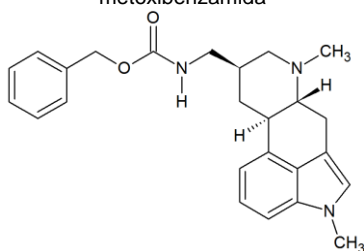
CISAPRIDE

4-amino-5-cloro-N-((3S,4S)-1-[3-(4-fluorofenoxi)propil]-3-metoxipiperidin-4-il)-2-metoxibenzamida



NARINGINA

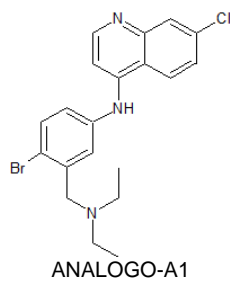
7-[[2-O-(6-Deoxi- α -L-manopiranosil)- β -D-glucopiranosil]oxi]-2,3-dihidro-5-hidroxi-2-(4-hidroxfenil)-4H-1-benzopiran-4-ona



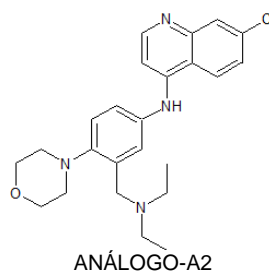
METERGOLINA

[(8 β)-1,6-dimetilergolin-8-il]metil]carbamato

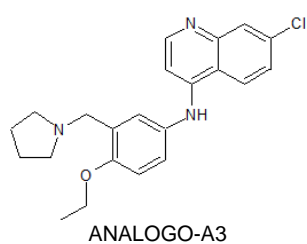
Tabla 8. Análogos sintéticos de la amodiaquina



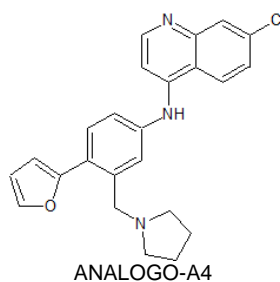
ANÁLOGO-A1



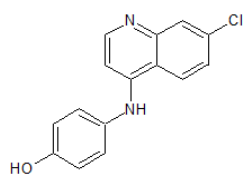
ANÁLOGO-A2



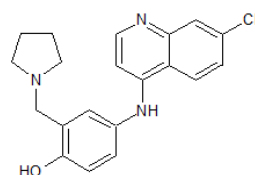
ANÁLOGO-A3



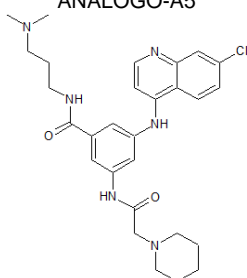
ANÁLOGO-A4



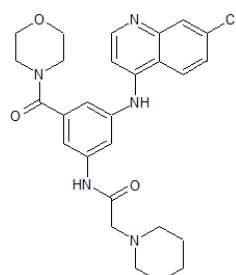
ANALOGO-A5



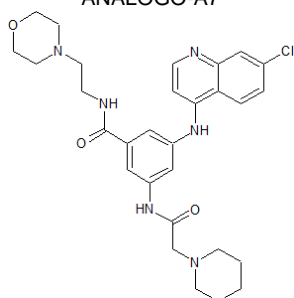
ANALOGO-A6



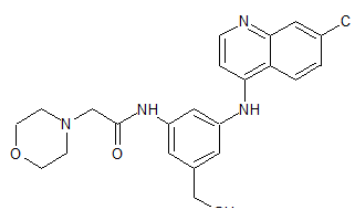
ANALOGO-A7



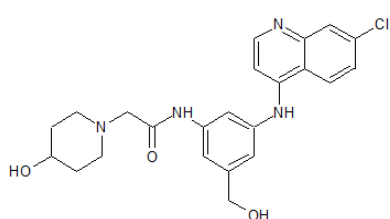
ANALOGO-A8



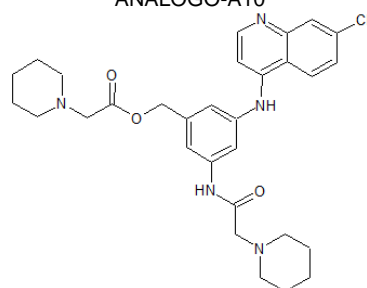
ANALOGO-A9



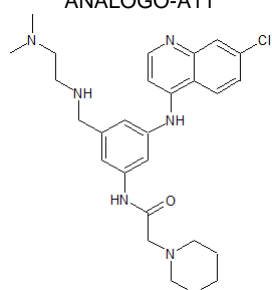
ANALOGO-A10



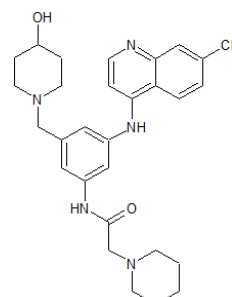
ANALOGO-A11



ANALOGO-A12



ANALOGO-A13



ANALOGO-A14

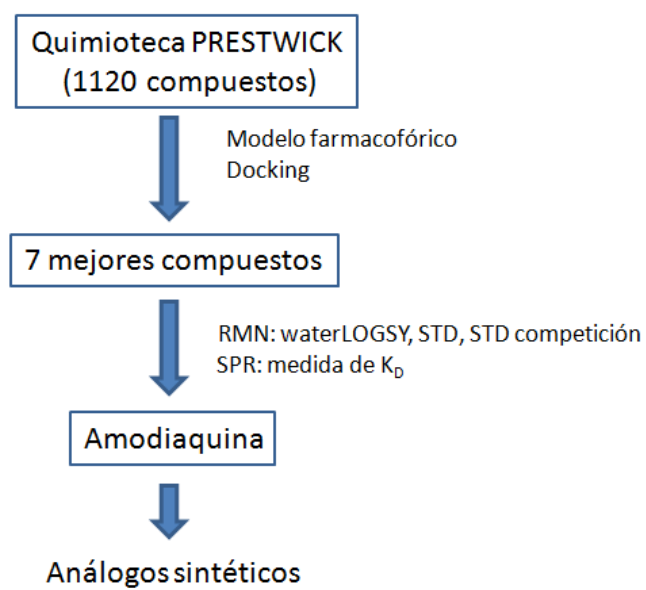


Figura 9. Estrategia de selección de inhibidores de heparanasa a partir de la quimioteca Prestwick.

Discusión y Resultados

1.-ANÁLISIS DEL MODELO ESTRUCTURAL DE HEPARANASA

El hecho de que hasta el momento no se haya descrito la estructura tridimensional de heparanasa nos llevó a pensar en la necesidad de obtener un modelo estructural de la proteína. Este modelo fue utilizado posteriormente, en primer lugar, para el diseño de las construcciones de heparanasa que se utilizarían para el cálculo la estructura tridimensional de la proteína mediante Resonancia Magnética Nuclear y en segundo lugar para llevar a cabo paralelamente el diseño de nuevos inhibidores de la proteína mediante técnicas *in silico*. La heparanasa se ha relacionado con miembros de una familia de glicosil hidrosilasas A (GH-A) y se ha demostrado que mediante el alineamiento de secuencias junto con predicciones de estructura secundaria, la heparanasa es homóloga a las familias 10, 39 y 51 de las GH-A, especialmente en la región del sitio catalítico (150).

El modelo estructural de heparanasa humana se obtuvo por homología de secuencia utilizando el programa MODELLER (<http://www.salilab.org/modeller/>) (231). En primer lugar el programa comparó la secuencia de la heparanasa con proteínas de estructura tridimensional conocida. La proteína con mayor similitud a la heparanasa y que se utilizó como punto de partida para el cálculo del modelo estructural de heparanasa fue la α -L-arabinofuranosidasa de *Geobacillus stearothermophilus* T-6 (237) que pertenece a la familia 51 de las glicosidasas. El resultado se muestra en la Fig.1.

Otros modelos estructurales de heparanasa realizados por homología de secuencia han sido publicados (150, 183, 238). Entre ellos, Huttler *et al.* utilizaron como punto de partida la secuencia entera de la proteína antes de la proteólisis (150) y Abboud-Jarrous *et al.* la secuencia del fragmento “linker” más la secuencia de la subunidad de 45 kDa (183); ambas construcciones representan formas inactivas de la enzima. La secuencia de la proteína está descrita en el apartado 3.3 de la Introducción. En nuestro trabajo se ha utilizado, al igual que Fux *et al.* (238), una secuencia de heparanasa en la que se ha introducido un puente entre las dos subunidades de heparanasa, HEP₃₆₋₁₀₉ (8

kDa) y HEP₁₅₈₋₅₄₃ (45 kDa), que consiste en tres repeticiones Gly-Ser, (GS)₃. Previamente se había demostrado que esta construcción era catalíticamente activa (239) y por ello se supuso que era la construcción más válida para hacer el modelo.

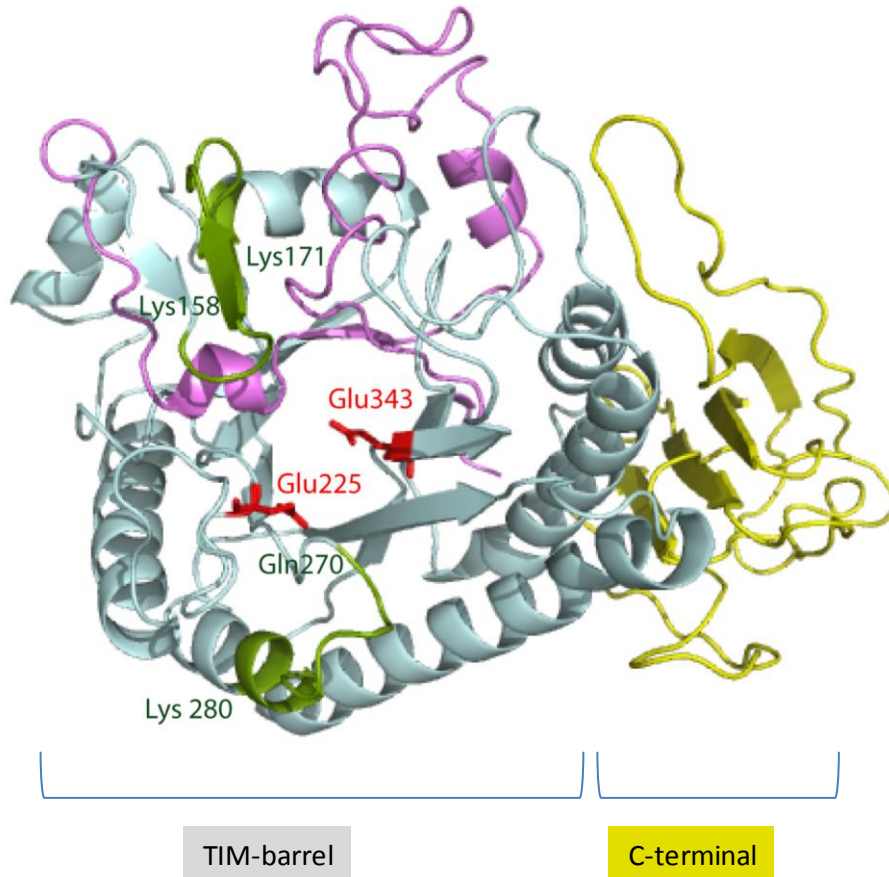


Figura 1. Modelo tridimensional de la enzima heparanasa. El modelo incluye la subunidad de 8 kDa en violeta, y la subunidad de 45 kDa que corresponde con el resto de la proteína. El extremo C-terminal está marcado en amarillo cuya estructura de 4 hojas β se diferencia con el plegamiento TIM-barrel que tiene el resto de la proteína. También se muestran los dos ácidos glutámicos activos Glu²²⁵ y Glu³⁴³ en rojo y los dos sitios de unión a heparán sulfato Lys¹⁵⁸-Asp¹⁷¹ y Gln²⁷⁰-Lys²⁸⁰ en verde.

La estructura que se obtuvo con el programa MODELLER muestra claramente un plegamiento de nueve hélices alfa alternadas con nueve hojas-β denominado “TIM-barrel” (Fig. 2) que también ha sido predicho para la enzima en otros modelos publicados (150, 183, 238) y que es común para toda la familia de las glicosil hidrolasas del tipo GH-A (150). Además, se puede observar que los dos sitios de unión a heparán sulfato Lys¹⁵⁸-Asp¹⁷¹ y Gln²⁷⁰-Lys²⁸⁰ (151) (Fig. 1, representados en verde) están próximos a los dos ácidos glutámicos responsables de la actividad catalítica de la enzima Glu²²⁵ y Glu³⁴³ (150) (Fig. 1, representados en rojo). La estructura del extremo C-terminal de la enzima se diferencia del resto por contener 4 hojas-β y no pertenecer al plegamiento TIM-barrel que contiene el resto de la proteína.

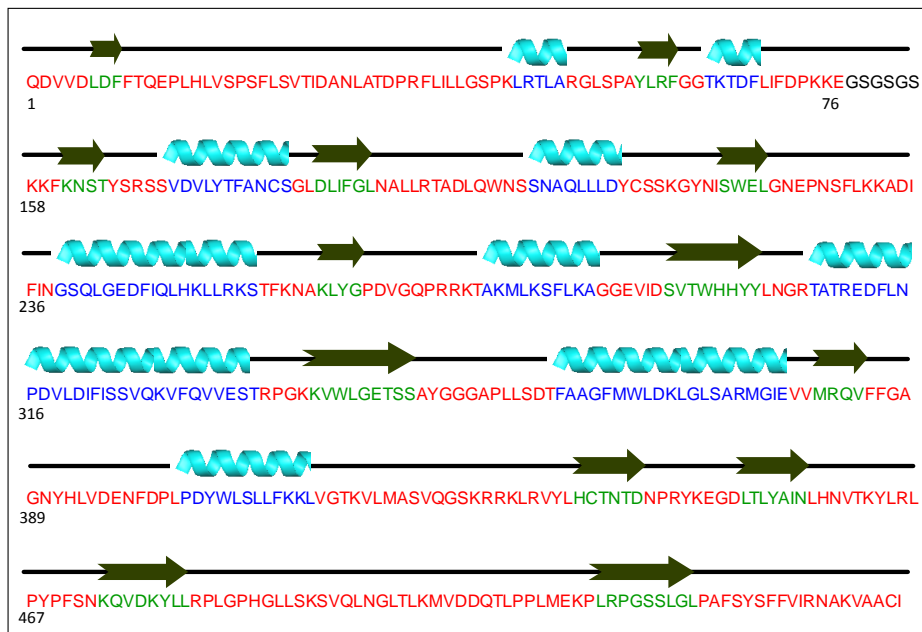


Figura 2. Esquema de la estructura secundaria de heparanasa predicha por homología de secuencia con el programa MODELLER. En azul están representadas las hélices-α y en verde las hojas-β. El resto de la proteína no tiene estructura secundaria.

2.-CLONACIÓN, EXPRESIÓN Y PURIFICACIÓN DE DISTINTAS CONSTRUCCIONES DE HEPARANASA

Se diseñaron distintas construcciones de heparanasa según el interés que podían tener tanto desde el punto de vista estructural como para la búsqueda de nuevos inhibidores de la enzima. El hecho de dividir la proteína en diferentes construcciones fue debido a que la resolución de estructuras por Resonancia Magnética Nuclear para proteínas de tamaños superiores a 30 kDa puede generar complicaciones de asignación debido al solapamiento y/o desaparición de señales en los espectros a causa de la relajación de las proteínas de mayor tamaño. Otra razón por la que se diseñaron diferentes construcciones es porque la heparanasa es una proteína muy hidrofóbica y a veces se puede eliminar parte de esta hidrofobicidad diseñando construcciones más pequeñas. Estas construcciones se diseñaron teniendo en cuenta el modelo estructural de la proteína descrito anteriormente y fueron la heparanasa completa (HEP₃₆₋₅₄₃), la subunidad de 8 kDa (HEP₃₆₋₁₀₉) y la subunidad de 45 kDa (HEP₁₅₈₋₅₄₃). Según el modelo, la heparanasa estaba dividida claramente en dos subdominios desde el punto de vista estructural, uno que comprendía la secuencia Glu³⁶-Lys⁴¹⁷ que muestra el plegamiento de “TIM-barrel” y otro con el resto de la secuencia de heparanasa Val⁴¹⁸-Ile⁵⁴³ que no formaba parte de este plegamiento de “TIM-barrel”. Para evaluar esta hipótesis se creyó interesante el diseño de dos construcciones más de heparanasa, HEP₃₆₋₄₁₇ de 37 kDa y HEP₁₅₈₋₄₁₇ de 29 kDa las cuales contenían los dos ácidos glutámicos activos de la enzima y los dos sitios de unión a heparán sulfato. En ambas construcciones se había eliminado el extremo C-terminal que, según el modelo por homología, parecía diferenciarse estructuralmente del resto de la proteína. Además, en la construcción de 29 kDa se había eliminado también la subunidad de 8 kDa. Finalmente también se consideró interesante el diseño de una construcción que solo contuviera uno de los dos ácidos glutámicos activos y uno de los dos sitios de unión a heparán sulfato, con el fin de evaluar y comprobar la importancia del sitio catalítico de la enzima en los estudios de interacción para la búsqueda de nuevos fármacos. Esta

construcción, HEP₂₃₃₋₄₁₇, de 21 kDa, contenía el residuo Glu³⁴³ y una de las secuencias de unión a heparán sulfato (Gln²⁷⁰-Lys²⁸⁰). Esta construcción es la homóloga humana a una proteína que se ha identificado en un topo, *Spalax judaei* (240), como una variante negativa de la enzima silvestre que inhibe la degradación de heparán sulfato y el crecimiento y colonización de tumores en modelo de ratón (136, 241). Esta construcción se diseñó para compararla con aquellas construcciones que contenían los dos ácidos glutámicos activos y las dos secuencias de unión a heparán sulfato y evaluar las diferencias en cuanto a su interacción con inhibidores conocidos de la enzima utilizando la Resonancia Magnética Nuclear y la Resonancia de Plasma Superficial.

HEP₃₆₋₄₁₇ fue elegida como referencia entre las diferentes construcciones de heparanasa para comparar los niveles de expresión y solubilidad de la proteína. Se utilizó esta construcción debido a que tenía un tamaño inferior a la proteína completa y solo se había eliminado la parte del extremo C-terminal que era estructuralmente diferente según el modelo por homología. Para ello, HEP₃₆₋₄₁₇ se subclonó a partir de HC-p11, que a su vez se había clonado a partir de HC-PGEX como se indica en el apartado 1.1 de Materiales y Métodos, en tres plásmidos con diferentes proteínas de fusión: p11, pMAL y pET44a. Las proteínas de fusión que contiene cada uno de los plásmidos aparecen en la Tabla 1. MBP y NusA se eligieron porque se ha demostrado anteriormente que ayudan en muchos casos a solubilizar proteínas expresadas en *E. coli* favoreciendo que se plieguen correctamente (242-244) y la cola de histidinas porque añadía menos complejidad a la purificación y obtención de la proteína. Se evaluaron distintas cepas de expresión de *E. coli* (BL21 (DE3), BL21-CodonPlus (DE3)-RIL, CV47, Rosetta (DE3) y Rosetta Blue K12 (DE3)), distintas temperaturas de inducción de la expresión (20, 25 y 37 °C) a distintos tiempos y distintas concentraciones de IPTG (0,4, 0,6 y 0,8 mM). Los niveles más altos de expresión se obtuvieron con las células BL21 (DE3) y 0,8 mM IPTG. La expresión de la proteína fue muy baja cuando las condiciones eran 20 °C durante toda la noche. Sin embargo, se obtenían altos niveles de expresión a 25 °C durante toda la noche y 37 °C durante 4 horas. No se encontraron diferencias

significativas entre ambas condiciones por lo que se eligió 25 °C para que la expresión fuera más lenta y promover así un correcto plegamiento de la proteína.

Tabla 1. Plásmidos utilizados en la clonación de HEP₃₆₋₄₁₇ y resultados de solubilidad/purificación obtenidos

Vector	Cola N-terminal	P. M. cola N-terminal	Solubilidad	Purificación	Corte TEV
p11	6xHis-TEV	1,8 kDa	Insoluble	–	–
pMAL	6xHis-NusA-TEV	58 kDa	Soluble	Muy pobre	no
pET44a	MBP-TEV	26 kDa	Soluble	Degradación	no

Después de la lisis celular realizada mediante sonicación, la proteína que contenía fusionada por el extremo N-terminal una cola de seis histidinas se encontraba en el pellet; por lo tanto era insoluble y formaba cuerpos de inclusión incluso cuando la temperatura de inducción se bajaba a 20 °C (Fig. 3). Por el contrario, las construcciones que contenían HEP₃₆₋₄₁₇ fusionada a MBP o NusA eran altamente solubles. Sin embargo, estas construcciones dieron problemas durante la purificación. En el primer caso, se usó cromatografía de afinidad aprovechando la interacción entre la MBP y la resina de amilosa pero se conseguía purificar muy poca cantidad de proteína y en el segundo, en el que se usó cromatografía de afinidad con resina TALON, la proteína se iba degradando durante la purificación. Además todos los intentos de cortar la cola MBP/NusA a distintas temperaturas y a distintos tiempos con la proteasa *Tobacco etch virus* (TEV) no tuvieron éxito.

La imposibilidad de eliminar las proteínas de fusión MBP (26 kDa) o NusA (58 kDa), hacía inviable el uso de estas construcciones para su caracterización biofísica por RMN ya que el tamaño total de la proteína hubiera sido demasiado

grande y hubiéramos tenido problemas de relajación y solapamiento de señales en el espectro. Por tanto, y a pesar de formar cuerpos de inclusión, todas las estrategias para mejorar la expresión, solubilidad y purificación de la proteína se enfocaron en la utilización del plásmido p11, en el que la cola de histidinas solo aportaba 1,8 kDa al total del peso molecular de la construcción y probablemente no interferiría en los experimentos de RMN.

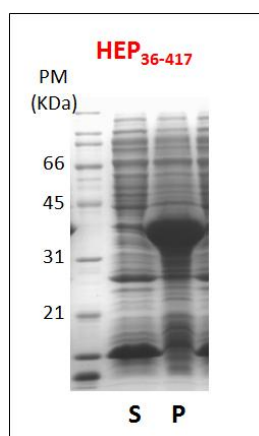


Figura 3. SDS-PAGE de la lisis de HEP₃₆₋₄₁₇ clonada en p11. S representa el sobrenadante y P el “pellet” después de la centrifugación que sigue a la sonicación.

Todas las construcciones anteriormente descritas fueron clonadas en el plásmido p11 excepto la subunidad de 45 kDa que fue clonada en pET47b ya que no se consiguió la clonación de esta construcción en p11. Este plásmido contiene al igual que p11, una cola de seis histidinas que fusiona con la proteína clonada por el extremo N-terminal. Sin embargo, a este plásmido le falta el sitio de corte con TEV. De todas las cepas celulares de *E. coli* que se probaron, BL21(DE3) fue la más eficiente para todas las construcciones excepto para HEP₁₅₈₋₄₁₇ donde BL21-CodonPlus (DE3)-RIL daba los mejores resultados. Aunque los niveles de expresión eran muy altos (Fig. 4), después de la lisis celular todas las construcciones mostraban muy baja solubilidad, quedándose casi la totalidad de la proteína en el pellet, incluso cuando la temperatura de expresión se bajaba a 20 °C.

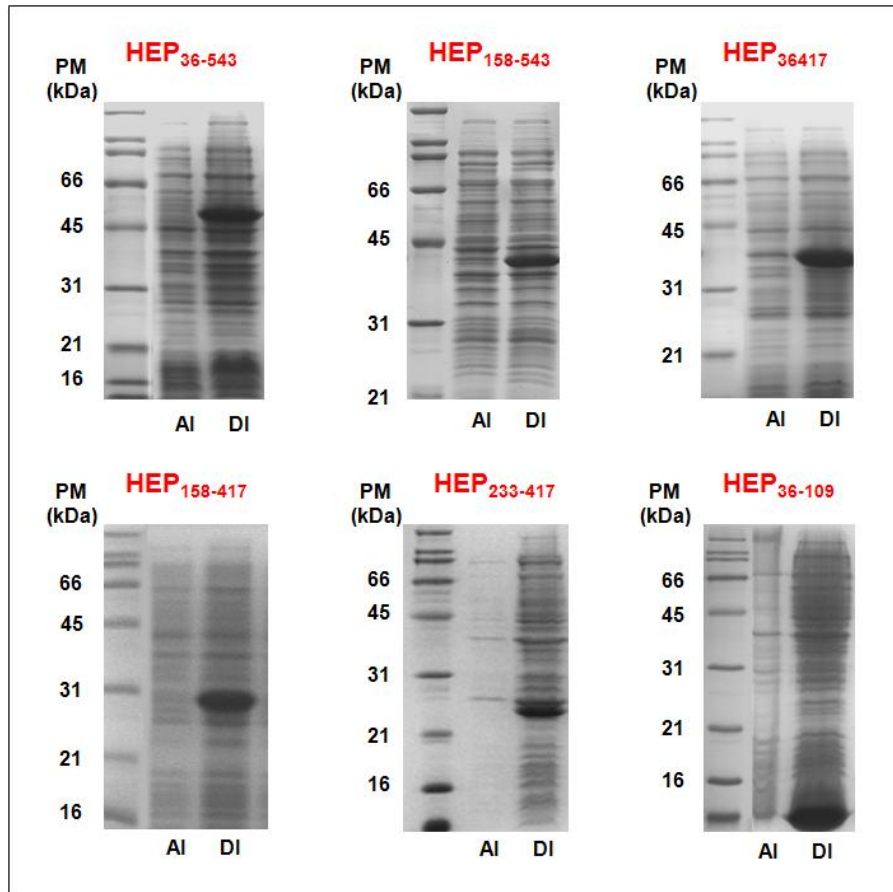


Figura 4. SDS-PAGE de la expresión de las distintas construcciones de heparanasa ordenadas de mayor a menor tamaño. El primer carril en cada uno de los geles indica el marcador de peso molecular, el segundo representa la expresión de la proteína antes de inducir (AI) y el tercer carril la expresión después de inducir con 0,8 mM IPTG (DI).

Se probaron distintas condiciones para poder obtener la proteína soluble, pura y en concentración suficiente para su caracterización biofísica. Estas condiciones del tampón de lisis (sonicación) y purificación vienen resumidas en la Tabla 2. El mejor resultado se obtuvo cuando se usaba una concentración de 0,125% (p/v) de N-Lauroyl-sarcosine en el tampón de lisis resultando que un 45-50% de la proteína era soluble. A continuación se probaron tres purificaciones

diferentes, las dos primeras con cromatografía de afinidad de metal inmovilizado, utilizando en el primer caso resina TALON y en el segundo resina NTA aprovechando la afinidad entre la cola de histidinas y los iones Co^{2+} y Ni^{2+} , respectivamente. El tercer tipo de purificación fue una cromatografía de afinidad utilizando una resina de heparina teniendo en cuenta la afinidad que tiene la heparanasa por heparina. En ninguno de los casos se obtuvieron resultados positivos. Posiblemente la presencia de detergente en el caso de la resina de Co^{2+} y Ni^{2+} , aunque no afectaba la interacción con la resina, impedía separar la heparanasa de otras proteínas y en el caso de la resina de heparina disminuía la interacción de la proteína con la resina impidiendo que la proteína quedara retenida en la columna.

Tabla 2. Resumen de las condiciones de lisis para aumentar la solubilidad de la proteína.

Detergente	NaCl	Otros aditivos	Condiciones desnaturalizantes	Observaciones
1% Triton	50 mM	10% glicerol	-	insoluble
1% Triton	100 mM	10% glicerol	-	insoluble
1% Triton	150 mM	10% glicerol	-	insoluble
1% Triton	300 mM	10% glicerol	-	insoluble
1% Triton	500 mM	10% glicerol	-	insoluble
2% Triton	300 mM	10% glicerol	-	insoluble
0,25% N-octilglucósido	300 mM	10% glicerol	-	insoluble
0,5% N-octilglucósido	300 mM	10% glicerol	-	insoluble
1% Nonidet	300 mM	10% glicerol	-	insoluble
3 mL BugBuster comercial	300 mM	10% glicerol	-	insoluble
1% CHAPS	300 mM	10% glicerol	-	insoluble
1% SDS	300 mM	-	-	insoluble

0,06 % N-Lauroylsarcosine	300 mM	-	-	10% soluble
0,125 % N-Lauroylsarcosine	300 mM	-	-	45-50% soluble, problemas en la purificación
0,25 % N-Lauroylsarcosine	300 mM	-	-	80-100% soluble, problemas en la purificación
-	300 mM	-	2 M urea	insoluble
-	300 mM	-	4 M urea	25% soluble, problemas en el replegamiento
-	300 mM	-	6 M urea	50% soluble, problemas en el replegamiento
-	300 mM	-	8 M urea	100% soluble, problemas en el replegamiento
-	300 mM	-	2M GndCl	25% soluble, problemas en el replegamiento
-	300 mM	-	4 M GndCl	50% soluble, problemas en el replegamiento
-	300 mM	-	6 M GndCl	100% soluble, problemas en el replegamiento
-	100 mM	25 mM heparina	-	insoluble
-	100 mM	25 mM suramina	-	insoluble
-	300 mM	1 mM PMFS	-	insoluble
-	-	pN-20	-	insoluble

El análisis de la secuencia de heparanasa confirmaba que el grado de solubilidad de la proteína es posiblemente debido a su propia naturaleza ya que se trata de una proteína que presenta un número muy elevado de aminoácidos de carácter hidrofóbico: leucinas (14,6%), isoleucinas (3,2%), fenilalaninas (6%), valinas (6,2%) y triptófanos (3,9%); en total un 33,9 % de la secuencia de

heparanasa completa. Este análisis junto con los resultados de solubilidad y purificación anteriormente mencionados hizo necesario buscar métodos alternativos de purificación de la proteína. En este sentido y aprovechando la alta hidrofobicidad de la enzima se intentó purificar la proteína eliminando del “pellet” otras proteínas más solubles que heparanasa.

Con este objetivo, el “pellet” obtenido después de la lisis celular en la que se usó como tampón 25 mM Tris-HCl, pH 8, 300 mM NaCl, 10% glicerol y 1% Triton se sometió a 3 sonicaciones más, intercalando entre cada sonicación la resuspensión del “pellet” con 25 mM Tris-HCl, pH 8, 300 mM NaCl además de: 10% glicerol y 1% Triton para la segunda sonicación, 0,06% (w/v) N-Lauroyl sarcosine y 0,125% (w/v) N-Lauroyl sarcosine para la tercera y cuarta sonicación, respectivamente. Cada sonicación consistió en 10 ciclos de 30 s de sonicación y 30 s de descanso. De esta forma, se pudieron eliminar muchas de las proteínas presentes en la muestra durante la lisis celular y las dos siguientes sonicaciones. Después de la última sonicación se obtuvo la proteína soluble y con una mayor pureza (Fig. 5A). Este proceso se aplicó a todas las construcciones de heparanasa con buenos resultados.

Finalmente, para refinar la purificación de la proteína, se utilizó cromatografía de afinidad de metal inmovilizado (IMAC) usando una resina TALON, que como se ha dicho anteriormente, tiene afinidad por la cola de histidinas fusionada a la proteína. Para minimizar la cantidad de detergente en la muestra, la proteína se eluyó en 10 fracciones de 1 mL de tampón con una concentración de 500 mM imidazol. Con este procedimiento se pudo obtener cada una de las construcciones con una pureza >90% como se puede observar en SDS-PAGE de la Fig. 5B. La cantidad de proteína pura obtenida a partir de 1 L de cultivo fue de 15-20 mg para cada una de las construcciones.

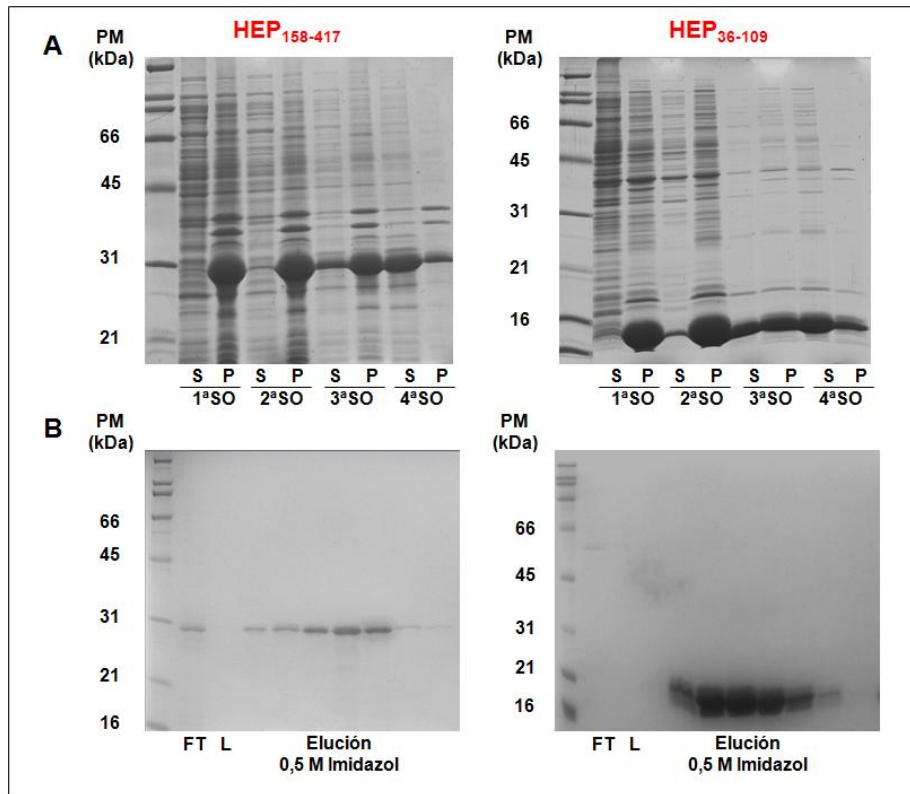


Figura 5. SDS-PAGE de la lisis y purificación de dos de las construcciones clonadas en p11, HEP₁₅₈₋₄₁₇ a la izquierda y HEP₃₇₋₁₀₉ a la derecha. (A) Muestra el resultado de la lisis celular (1^a SO) y todas las sonicaciones posteriores (2^a SO, 3^a SO y 4^a SO), S representa el sobrenadante y P el “pellet” después de la centrifugación que sigue a cada sonicación. (B) Muestra el resultado de la purificación de las dos construcciones utilizando cromatografía de afinidad de metal inmovilizado (resina TALON), FT “flow through”, L es el lavado de la resina con el tampón y los posteriores carriles representan la elución con 500 mM imidazol. En todos los casos el primer carril corresponde al marcador de peso molecular (PM).

El peso molecular y la secuencia de cada una de las construcciones fueron confirmadas mediante espectrometría de masas MALDI TOF. En la figura 6 se puede observar un ejemplo de espectro de masas después de realizar la digestión de la proteína con tripsina que corta por el extremo C-terminal de lisina-arginina

excepto cuando el aminoácido que le precede es una prolina. A continuación se escogieron los picos de mayor intensidad como precursores para hacer el MS/MS. Los péptidos que se encontraron en el espectro cubren un 69 % de la secuencia de la proteína y están marcados en rojo en la secuencia de la proteína que se muestra a continuación.

```

1  KKFKNSTYSR  SSVDVLYTFANCSDLIFGLNALLRTADLQWNSSNAQLL
51  LDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDFIQLHKLLRKSTF
102 KNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYLNGRTATRED
153 FLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA
204 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDLSELLFKKL
256 VGTK
    
```

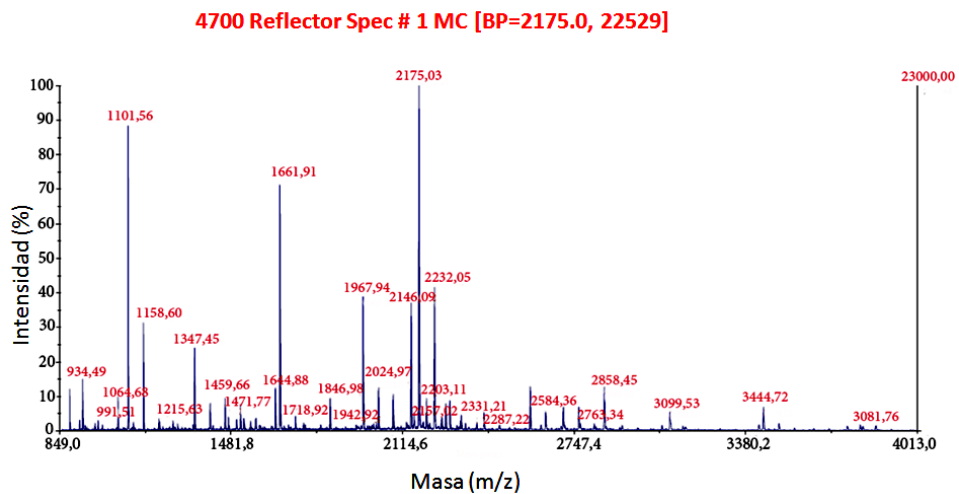


Figura 6. Espectro de masas MALDI-TOF de heparanasa después de la digestión de la proteína con tripsina.

A continuación se muestra el resultado del análisis realizado con el programa MASCOT. Se pueden observar la masa y secuencia de los picos obtenidos por MS y aquellos que se escogieron como precursores para MS/MS.

Nº aa	Masa observada	Masa esperada	Masa calculada	Delta	Secuencia	
11 - 36	2858.45	2857.44	2857.46	-0.0227	R.SSVDVLYTFANCSGLDLIFGLNALLR.T	MS
37 - 57	2414.13	2413.12	2413.12	0.0034	R.TADLQWNSSNAQLLLDYCSSK.G	MS
58 - 74	1967.94	1966.93	1966.94	-0.0054	K.GYNISWELGNEPNSFLK.K	MS/MS
58 - 74	1967.94	1966.93	1966.94	-0.0054	K.GYNISWELGNEPNSFLK.K	MS
75 - 94	2273.18	2272.18	2272.18	-0.0047	K.KADIFINGSQLGEDFIQLHK.L	MS
76 - 94	2145.09	2144.08	2144.09	-0.0049	K.ADIFINGSQLGEDFIQLHK.L	MS
76 - 94	2146.08	2145.07	2145.07	0.0058	K.ADIFINGSQLGEDFIQLHK.L	MS
106 - 115	1101.57	1100.57	1100.56	0.0092	K.LYGPVVGQPR.R	MS/MS
106 - 115	1101.57	1100.57	1100.56	0.0092	K.LYGPVVGQPR.R	MS
128 - 146	2174.02	2173.02	2173.03	-0.0127	K.AGGEVIDSVTWHHYLNGR.T	MS/MS
128 - 146	2174.02	2173.02	2173.03	-0.0127	K.AGGEVIDSVTWHHYLNGR.T	MS
128 - 146	2175.02	2174.01	2174.01	0.0003	K.AGGEVIDSVTWHHYLNGR.T	MS
151 - 168	2079.05	2078.05	2078.05	-0.0051	R.EDFLNPVDLDFISSVQK.V	MS
169 - 180	1346.74	1345.73	1345.73	0.0043	K.VFQVVESTRPQK.K	MS
169 - 180	1347.75	1346.74	1346.71	0.0274	K.VFQVVESTRPQK.K	MS
218 - 225	934.49	933.48	933.47	0.0090	R.MGIEVVMR.Q	MS
226 - 254	3444.71	3443.71	3443.68	0.0259	R.QVFFGAGNYHLVDENFDPLPDYWLSLFLK.K	MS

3.-OPTIMIZACIÓN DE LAS CONDICIONES DE SOLUBILIDAD Y ESTABILIDAD PARA LOS ESTUDIOS DE RMN

Antes de abordar los estudios de asignación e interacción por RMN de heparanasa fue necesario optimizar algunos parámetros como la solubilidad, concentración y estabilidad de la muestra. Con este fin, se probaron diferentes tampones con un rango de pH de 5 a 8, concentraciones de NaCl (50-200 mM) y diferentes temperaturas (25-45 °C).

En primer lugar, todas las construcciones fueron dializadas frente a distintos tampones: 25 mM acetato sódico pH 5,5, 25 mM fosfato sódico pH 6 y 25 mM Tris-HCl pH 7,4 y 8, todos ellos incluyendo 50 mM NaCl y 0,125% N-Lauroyl sarcosine. A continuación y con el objetivo de poder evaluar si la proteína estaba bien plegada se realizaron experimentos 1D ¹H para cada una de las condiciones anteriores. Debido a la presencia de detergente N-lauroyl-sarcosine en la muestra este espectro no dio mucha información ya que las señales correspondientes al detergente resuenan en la misma zona del espectro que la parte alifática de la proteína. En la Fig. 7 se muestra un ejemplo de este experimento para la construcción Hep₂₃₃₋₄₁₇. La dispersión de las señales en la región aromática y de

los -NH va de 6.8 a 8.9 ppm, lo que nos hace pensar en una estructura mayoritariamente en hélice α .

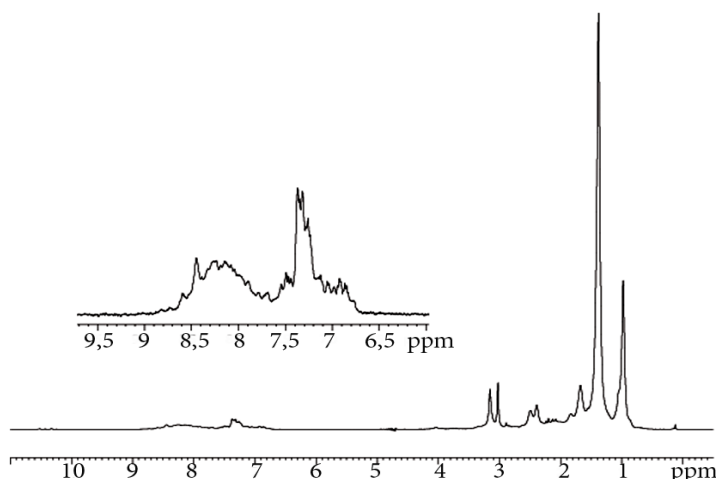


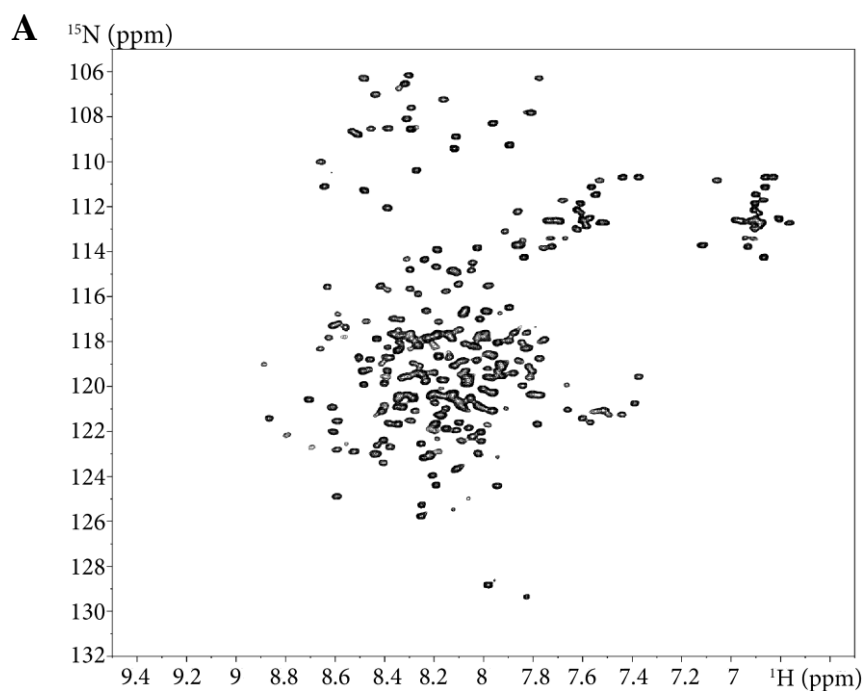
Figura 7. Espectro 1D ^1H de $\text{Hep}_{233-417}$ adquirido en un espectrómetro Bruker 600 MHz con criosonda TCI. La temperatura de la muestra era de 45 °C y el pH 6.

Finalmente cada una de las condiciones experimentales mencionadas anteriormente fueron evaluadas mediante el experimento de RMN 2D [^1H , ^{15}N]-HSQC que proporciona importante información cualitativa sobre el plegamiento de una proteína. Globalmente, los peores resultados en términos de estabilidad se observaron en 25 mM acetato sódico, pH 5,5 debido en primer lugar a que la muestra se degradaba a las 12-24 horas, lo que hacía imposible utilizar este tampón para experimentos tridimensionales largos de RMN que duraban entre 2-4 días y en segundo lugar la calidad del espectro no es buena ya que muestra aproximadamente la mitad de las señales esperadas, probablemente como resultado de la agregación de la proteína. El tampón que dio mejores resultados para todas las construcciones de heparanasa fue 25 mM fosfato sódico, pH 6 donde el número de picos coincide con el número de residuos que tiene la proteína en su secuencia. La diferencia en la calidad del espectro a pH 5,5 y pH 6 se puede observar en el espectro [^1H , ^{15}N]-HSQC de $\text{HEP}_{158-417}$ (Fig. 8).

El espectro correspondiente a la muestra preparada con tampón a pH 8 también mostraba menos señales de las que correspondían al número de aminoácidos de la proteína. Esto puede explicarse porque el intercambio NH-H₂O aumentaba por encima de pH 7,5 lo que da lugar a la reducción o desaparición de señales en el espectro [¹H, ¹⁵N]-HSQC (245) (Fig. 8).

En todos los experimentos adquiridos, la concentración de NaCl no parecía afectar significativamente a la estabilidad de la proteína ni a la calidad de los espectros, por lo tanto se usó la menor concentración de NaCl posible (50 mM). Esta concentración era suficiente para que la proteína fuera estable y para evitar una reducción de la relación señal/ruido en el espectro debido a la mayor conductividad de la muestra provocada por una alta concentración de NaCl en la misma (246).

Por otro lado, los experimentos adquiridos a 45 °C dieron los mejores resultados en todos los casos. En la figura 9 se muestra un ejemplo con la construcción HEP₁₅₈₋₄₁₇ a 37 °C y 45 °C. Se observó que los experimentos adquiridos a menor temperatura mostraban señales muy anchas en el espectro de RMN, quizás debido a la viscosidad de la muestra por la presencia de detergente.



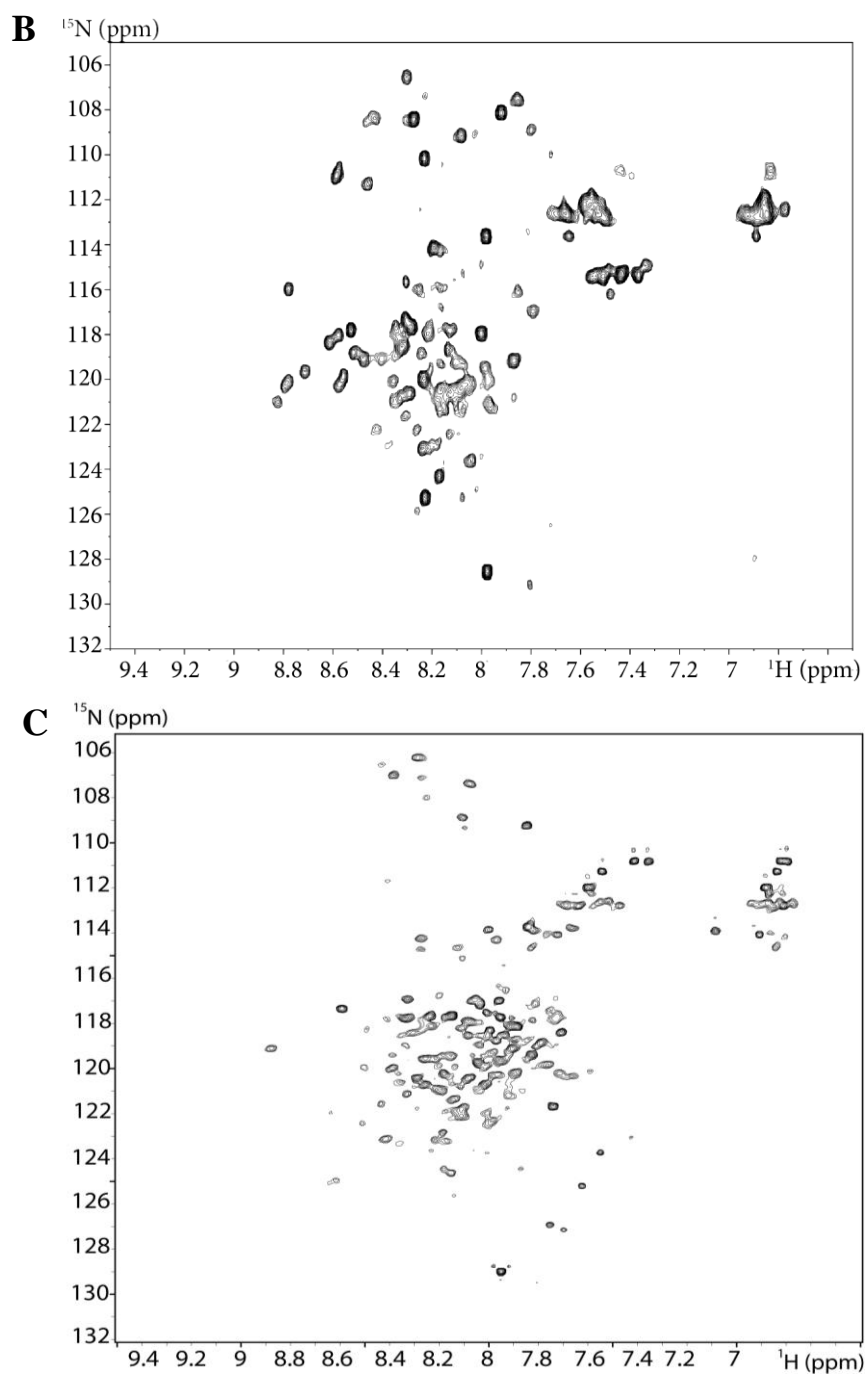


Figura 8. Espectro [^1H , ^{15}N]-HSQC adquirido a 600 MHz con criosonda TCI. 300 μM HEP₁₅₈₋₄₁₇, 45 °C. Los tampones utilizados, fueron: (A) fosfato sódico pH 6 (B) acetato

sódico pH 5,5, (C) Tris-HCl pH 8. En todos los casos con 50 mM NaCl y 0,125% N-Lauroyl sarcosine.

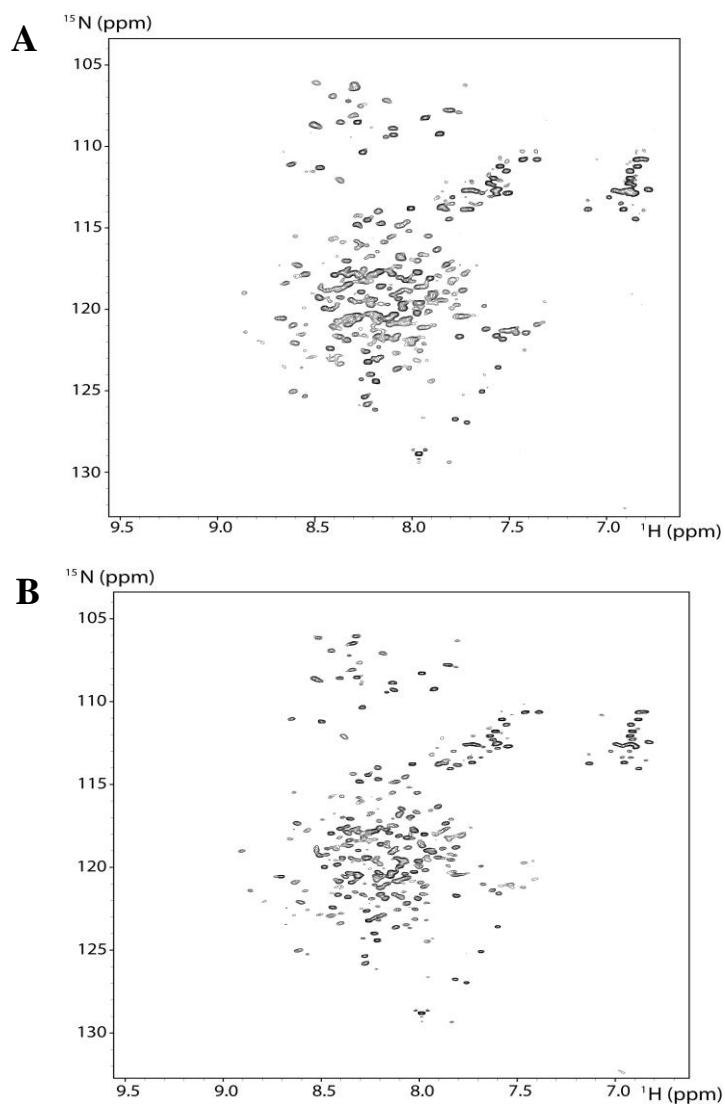


Figura 9. Espectro [^1H , ^{15}N]-HSQC de $\text{HEP}_{158-417}$ 300 μM , en fosfato sódico pH 6, 50 mM NaCl y 0,125% N-Lauroyl sarcosine adquiridos en un espectrómetro Bruker 600 MHz con criosonda TCI dos temperaturas (A) 37 °C (B) 45 °C.

Al optimizar las condiciones se llevaron a cabo experimentos con todas las construcciones de heparanasa y los resultados revelaban que había cambios significativos entre las construcciones de mayor y menor tamaño. Los mejores espectros de RMN (número de picos, calidad y forma de las señales) se observaron para las construcciones HEP₁₅₈₋₄₁₇, HEP₂₃₃₋₄₁₇ y HEP₃₆₋₁₀₉ de 29, 21 y 8 kDa respectivamente, en comparación con los espectros realizados para las construcciones con peso molecular mayor de 30 kDa, HEP₃₆₋₄₁₇, HEP₁₅₈₋₅₄₃ y HEP₃₆₋₅₄₃ de 37, 45 y 52 kDa respectivamente. Los espectros de las proteínas más grandes muestran señales más anchas e incluso muchas llegan a desaparecer, producido por un decrecimiento de la relación señal/ruido del espectro. Esto se debe a que las proteínas más grandes tienen una velocidad de relajación transversal muy rápida especialmente debido a los heteronúcleos $^{13}\text{C}^\alpha$ que relajan eficientemente por interacciones dipolo-dipolo con $^1\text{H}^\alpha$ y a la relajación transversal de ^{15}N durante los pasos de transferencia de la coherencia (247). A ello se une el problema de la falta de resolución espectral debido al gran número de señales de RMN que las proteínas grandes presentan.

La Fig. 10 muestra los espectros [$^1\text{H}, ^{15}\text{N}$]-HSQC para cuatro de estas construcciones. La Fig. 10A corresponde al espectro de HEP₃₆₋₅₄₃ de 52 kDa donde se puede apreciar el ensanchamiento de las señales al que se ha hecho mención anteriormente. Aunque el ensanchamiento de los picos no es tan pronunciado en el espectro de HEP₃₆₋₄₁₇ (37 kDa) (Fig. 10 B), el solapamiento de señales es muy grande, sobre todo en la zona entre 8 y 8,5 ppm. Es muy posible que el tamaño aparente de estas proteínas sea mayor de lo que realmente indica su peso molecular debido a la presencia del detergente en la muestra. Se puede observar cómo la calidad de los espectros es mayor en las construcciones de menor peso molecular, HEP₁₅₈₋₄₁₇ y HEP₃₆₋₁₀₉ de 29 y 8 kDa respectivamente (Fig. 10C y 10D).

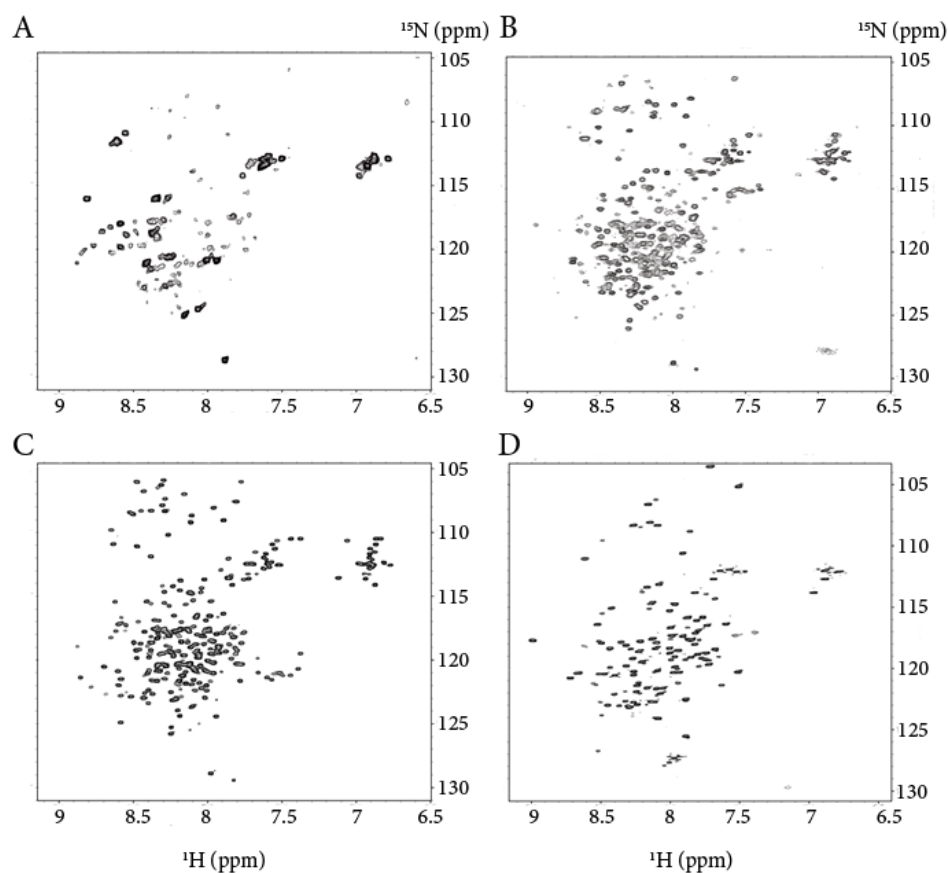


Figura 10. Espectros $[^1\text{H}, ^{15}\text{N}]$ -HSQC adquiridos en un espectrómetro Bruker 600 MHz con criosonda TCI a 45 °C de (A) HEP₃₇₋₅₄₃, (B) HEP₃₇₋₄₁₇, (C) HEP₁₅₈₋₄₁₇ y (D) HEP₃₇₋₁₀₉.

La dispersión de las señales de ^1H en estos espectros se ve reducida a un rango de desplazamientos químicos entre 6,5 y 8,8 ppm, que es característico de las proteínas con estructura mayoritaria de hélice- α . En los espectros $[^1\text{H}, ^{15}\text{N}]$ -HSQC de las construcciones de 29, 21 y 8 kDa se contabilizaron casi tantos picos como aminoácidos tenía la proteína lo que sugería que los espectros de RMN eran de buena calidad para empezar la asignación. Para comprobar si la proteína estaba plegada se adquirió el espectro $[^1\text{H}, ^{15}\text{N}]$ -HSQC de cada una de las construcciones disueltas en 25mM fosfato sódico, 50 mM NaCl y 6 M de GndCl. El resultado se muestra en la Fig. 11 para la construcción HEP₁₅₈₋₄₁₇ en donde se puede observar que el rango de desplazamientos químicos es más

reducido, característica típica de proteínas desplegadas. Al poner la proteína en diálisis frente al mismo tampón sin GndCl la proteína vuelve a recuperar su estructura siempre y cuando se añada 0,125 mM N-Lauroyl sarcosine al tampón de diálisis. Por lo tanto se demostró con este experimento que la proteína estaba plegada aunque el rango de desplazamientos químicos no fuera demasiado extenso.

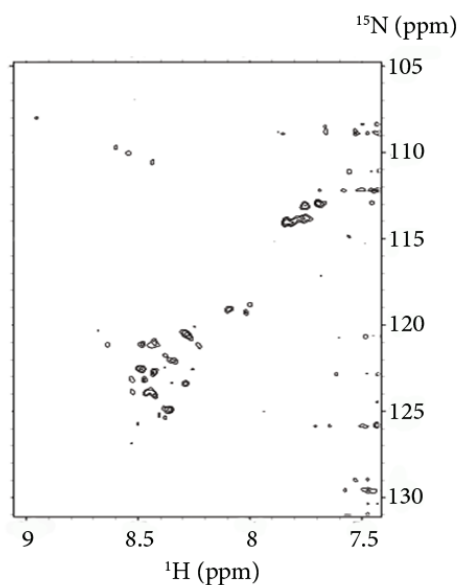


Figura 11. Espectro [^1H , ^{15}N]-HSQC de una muestra de HEP₁₅₈₋₄₁₇ preparada en un tampón 25 mM fosfato sódico, 50 mM NaCl y 6 M GndCl. Este experimento fue adquirido en un espectrómetro Bruker 600 MHz con criosonda TCI a 45 °C.

4.-ASIGNACIÓN DE HEPARANASA POR RESONANCIA MAGNÉTICA NUCLEAR

Se eligieron las tres construcciones de heparanasa más pequeñas, HEP₁₅₈₋₄₁₇ (29 kDa), HEP₂₃₃₋₄₁₇ (21 kDa) y HEP₃₇₋₁₀₉ (8 kDa) debido a que son las que presentan condiciones más favorables en RMN como se ha demostrado con los espectros de [^1H , ^{15}N]-HSQC.

Para el inicio de la asignación, los primeros experimentos que se adquirieron fueron los espectros bidimensionales [^1H , ^{15}N]-HSQC y [^1H , ^{13}C]-HSQC, en $\text{H}_2\text{O}:\text{D}_2\text{O}$ (90:10) y 100% D_2O respectivamente. En la Fig. 12, se pueden observar ambos espectros para $\text{HEP}_{233-417}$. Las señales que aparecen dentro de un círculo en el [^1H , ^{15}N]-HSQC (Fig. 12A) a aproximadamente 10,5 ppm corresponden al NH del anillo indólico del triptófano ($\text{H}_\epsilon\text{-N}_\epsilon$). En este caso se pueden observar 4 picos que corresponden al total de los Trp de la proteína. Lo mismo ocurre con $\text{HEP}_{158-417}$ en donde se pudieron observar las 6 señales de Trp. En el caso de HEP_{37-109} no se observó ninguna señal en esta zona ya que esta construcción no contenía Trp en su secuencia. El número de resonancias en la región de este mismo experimento indicada con un rectángulo es coherente con el número de glicinas presentes en la secuencia de cada una de las construcciones 22, 21 y 4 para $\text{HEP}_{158-417}$, $\text{HEP}_{233-417}$ y HEP_{37-109} respectivamente. Por otro lado, el número de señales en la región del espectro señalado con un recuadro de líneas punteadas concuerda con el número de cadenas laterales de Asn y Gln en estas construcciones. Por último, los NH de la cadena lateral de arginina y lisina aparecen normalmente a desplazamientos químicos de ^{15}N alrededor de 73 ppm y de ^1H alrededor de 6,8 ppm en el caso de arginina y de 54 ppm y 7,5 ppm respectivamente en el caso de lisina. Estos desplazamientos químicos están fuera del rango de desplazamientos que se utiliza normalmente para el [^1H , ^{15}N]-HSQC. Algunos de ellos se pueden observar en el espectro a 123/7.5 ppm debido al reflejo de la señal sobre el rango de frecuencias aplicado en el experimento.

En la Fig. 12B se muestra el espectro [^1H , ^{13}C]-HSQC de $\text{HEP}_{233-417}$ donde se observan las correlaciones $^1J_{\text{CH}}$ para todos los carbonos no cuaternarios de una proteína. En esta figura están indicadas las regiones del espectro correspondientes a las correlaciones entre $^1\text{H}^\alpha\text{-}^{13}\text{C}^\alpha$, $^1\text{H}^\beta\text{-}^{13}\text{C}^\beta$ y entre los ^1H y los ^{13}C de las cadenas laterales (aquí están incluidas las correlaciones $^1\text{H}^\gamma\text{-}^{13}\text{C}^\gamma$, $^1\text{H}_i^\delta\text{-}^{13}\text{C}^\delta$ y $^1\text{H}^\epsilon\text{-}^{13}\text{C}^\epsilon$). También se ha indicado la región correspondiente a los C^δ de las prolinas.

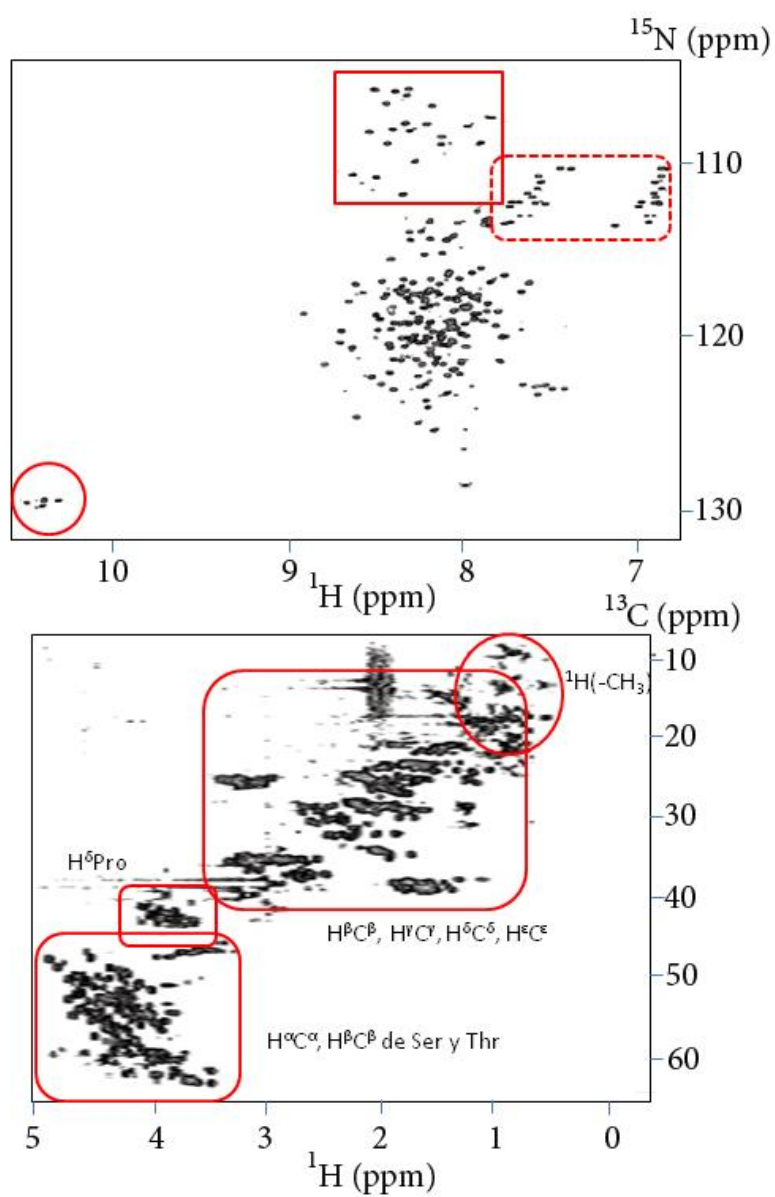


Figura 12. Experimentos (A) $[\text{}^1\text{H}, \text{}^{15}\text{N}]$ -HSQC y (B) $[\text{}^1\text{H}, \text{}^{13}\text{C}]$ -HSQC de HEP₂₃₃₋₄₁₇ en $\text{H}_2\text{O}/\text{D}_2\text{O}$ y D_2O respectivamente adquiridos en un espectrómetro AV600 Bruker 600 MHz con criosonda TCI a 45 °C. La concentración de la muestra fue de 1 mM.

Como se puede observar en la Fig. 10 y 12, en espectros de proteínas como estas construcciones de heparanasa, donde las predicciones dan un resultado mayoritario de estructura secundaria de α -hélice y además tienen un peso molecular elevado (principalmente las construcciones HEP₁₅₈₋₄₁₇ y HEP₂₃₃₋₄₁₇ de 29 kDa y 21 kDa respectivamente), el solapamiento de las señales es muy elevado y resulta esencial el uso de experimentos tridimensionales para realizar la asignación. Aunque el solapamiento es menor para la construcción HEP₃₇₋₁₀₉ de 8 kDa, también fue necesaria la adquisición de experimentos tridimensionales. Esto requirió como se ha descrito en apartados anteriores el marcaje isotópico de la proteína en ^{15}N y ^{13}C (248). No se contempló la posibilidad de utilizar experimentos bidimensionales como la secuencia 2D NOESY/TOCSY con los que además no se necesita marcaje isotópico de la proteína, debido a que las señales correspondientes al detergente solaparían con las señales de la parte alifática de la proteína.

Posteriormente, se adquirieron los experimentos de triple resonancia: HNCO, HNCA, HN(CA)CO, HNCACB, CBCA(CO)NH, HNHA, HNHB, HBHA(CO)NH, [^1H - ^{15}N]-HSQC-TOCSY, H(C)CH-TOCSY, (H)CCH-TOCSY, [^1H - ^{15}N]-HSQC-NOESY, [^1H - ^{13}C]-HSQC-NOESY. A modo de ejemplo en la Fig. 13 se puede observar el espectro 3D HNCO de HEP₁₅₈₋₄₁₇, donde puede apreciarse como al incluir la tercera dimensión de ^{13}C se produce una dispersión de las señales con respecto al [^1H , ^{15}N]-HSQC en el que se basa el experimento 3D. La tercera dimensión del espectro HNCO muestra los desplazamientos químicos de $^{13}\text{C}_i$ de carbonilo (CO) para un determinado NH_{i+1} . Este experimento nos permitió determinar con precisión el número de resonancias presentes en el espectro que coincidía con el número de aminoácidos en la secuencia de la proteína.

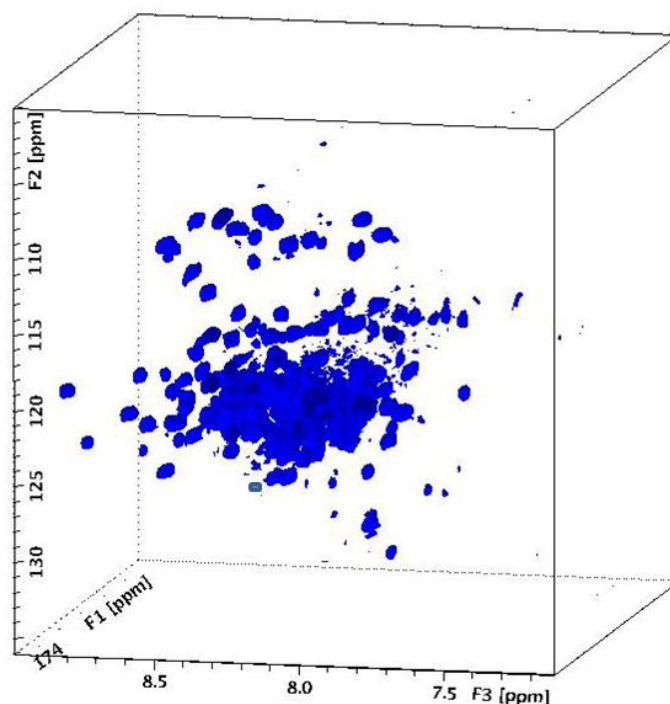


Figura 13. Espectro HNCOCY de la construcción $HEP_{158-417}$ de heparanasa. El experimento fue adquirido a 45 °C en un espectrómetro Bruker 900 MHz. La concentración de la proteína fue de 800 μ M.

Posteriormente se realizó la identificación de las señales a través de tiras de espectros como se especifica en el apartado 2.3 de materiales y métodos. En la Fig. 14 se muestra como se llevó a cabo la identificación de las señales que aparecen en el HNCACB para el aminoácido Val₄₁₄. Una vez identificada la señal en el espectro [¹H, ¹⁵N]-HSQC (Fig. 10C) y fijadas las dimensiones de ¹H y ¹⁵N, se estudió todo el ancho de ventana para la tercera dimensión de ¹³C. En la Fig. 14A se muestra la tira del espectro del HNCACB para el desplazamiento químico de Val₄₁₄ en el [¹H, ¹⁵N]-HSQC. Pueden distinguirse 4 picos en esta tira del HNCACB. Como consecuencia del diseño de la secuencia, las señales de los C ^{α} y los C ^{β} aparecen de distinto signo para poder ser identificadas más fácilmente. Puesto que las señales propias del sistema (i) aparecen normalmente

a una intensidad mayor que las pertenecientes al sistema (i-1), las señales de C^α y C^β que aparecen a desplazamientos de 59 y 28,5 ppm respectivamente, corresponden al sistema (i, Val₄₁₄). Por el contrario, las que aparecen a 53,6 y 39,5 ppm de menor intensidad corresponden al sistema (i-1, Leu₄₁₃). La naturaleza de la Val no resultó difícil de determinar ya que estos aminoácidos tienen una combinación de desplazamientos químicos de C^α y C^β diferente al resto de aminoácidos. Esta combinación de desplazamientos es $62,46 \pm 2,95$ ppm para C^α y $32,65 \pm 1,82$ ppm para C^β (249). Como existía más de una combinación Leu-Val en la secuencia de la proteína, fue necesario utilizar otros experimentos tridimensionales como el [1H , ^{15}N]-HSQC-NOESY descrito en la Fig. 16.

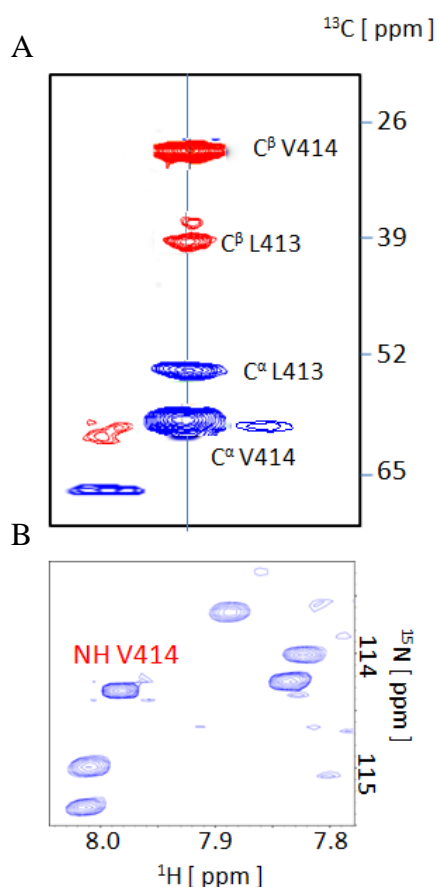


Figura 14. "Peak-picking" del espectro del HNCACB de HEP₁₅₈₋₄₁₇ para la Val₄₁₄. (A) Dimensión de ^{13}C del HNCACB. Las señales de los C^α aparecen como positivas en azul y las señales de los C^β como negativas y se muestran en rojo. (B) Dimensiones de ^{15}N y 1H del HNCACB. Este experimento fue adquirido a 45 °C en un espectrómetro Bruker, 900 MHz. La concentración de la proteína fue de 800 μM .

En la Fig. 15 se muestra un ejemplo de cómo las tiras del HNCA fueron esenciales para poder unir de forma secuencial los residuos Val₄₁₄-Gly₄₁₅-Thr₄₁₆-Lys₄₁₇ en la construcción HEP₁₅₈₋₄₁₇. En esta figura, las tiras del HNCA que corresponden al plano ¹H-¹³C muestran las señales del C^α_i y C^α_{i-1} para sus NH_i. Como en el caso anterior, las señales que muestran mayor intensidad son las correspondientes al sistema (i) mientras que las que tienen una intensidad menor pertenecen al sistema (i-1), ya que la constante de acoplamiento escalar entre el H_i-N_i y el C^α_i es mayor que entre H_i-N_i y el C^α_{i-1}. Esto permitió identificar fácilmente las señales de C^α_i y C^α_{i-1} para los residuos Val₄₁₄, Gly₄₁₅, Thr₄₁₆, Lys₄₁₇. Para poder unir estos sistemas de forma secuencial, fue necesario buscar coincidencias entre la señal C^α del residuo (i) y la señal de C^α_{i-1} del residuo (i+1). Por lo tanto, la asignación secuencial de estos residuos se realizó identificando aquellas señales que presentan las coincidencias esperadas en los desplazamientos químicos de ¹⁵N y ¹³C. En el caso del fragmento Val₄₁₄-Gly₄₁₅-Thr₄₁₆-Lys₄₁₇, su asignación se vio facilitada por los particulares desplazamientos químicos de C^α que presentan los aminoácidos Val, Gly y Thr (62,46 ± 2,95, 45,33 ± 1,33 y 62,17 ± 2,71 ppm respectivamente (249)), así como la ausencia de un fragmento similar en la secuencia de la proteína.

Para resolver las ambigüedades producidas por segmentos secuenciales idénticos dentro de la proteína, se utilizó la información proporcionada por el experimento [¹H, ¹⁵N]-HSQC-NOESY donde además de otras muchas correlaciones se pueden observar correlaciones N_{H_i}-N_{H_(i+1)}. En este espectro, también se pueden observar las correlaciones, H^α_i-NH_{i+1} y H^α_i-NH_{i+3} para las zonas de la proteína con estructura secundaria en hélice-alfa, lo que ayuda a unir secuencialmente los aminoácidos. En la Fig.16 están representadas las tiras de espectros de [¹H, ¹⁵N]-HSQC-NOESY para la construcción HEP₃₇₋₁₀₉. Con este experimento fue posible unir el fragmento Asp₈₅-Ile₈₆-Phe₈₇-Ile₈₈-Ser₈₉ basándose en las correlaciones NH_i-NH_{i+1}. En este experimento hay dos dimensiones de ¹H y una de ¹⁵N. Para poder ver la tira correspondiente a la tercera dimensión se fijan las dimensiones de ¹⁵N y la primera de ¹H. Como hay dos dimensiones de ¹H, el espectro presenta una diagonal que permite ver el pico

de cruce del $^1\text{H}_\text{N}$ consigo mismo y por tanto se puede ver en la tira, la señal del $^1\text{H}_i\text{-}^{15}\text{N}_i$. Como las correlaciones existentes en el NOESY están basadas en distancias pueden observarse correlaciones no solo secuenciales sino también con otros aminoácidos lejanos en secuencia. En la Fig. 16 las coincidencias de desplazamientos químicos están representadas con líneas de color rojo.

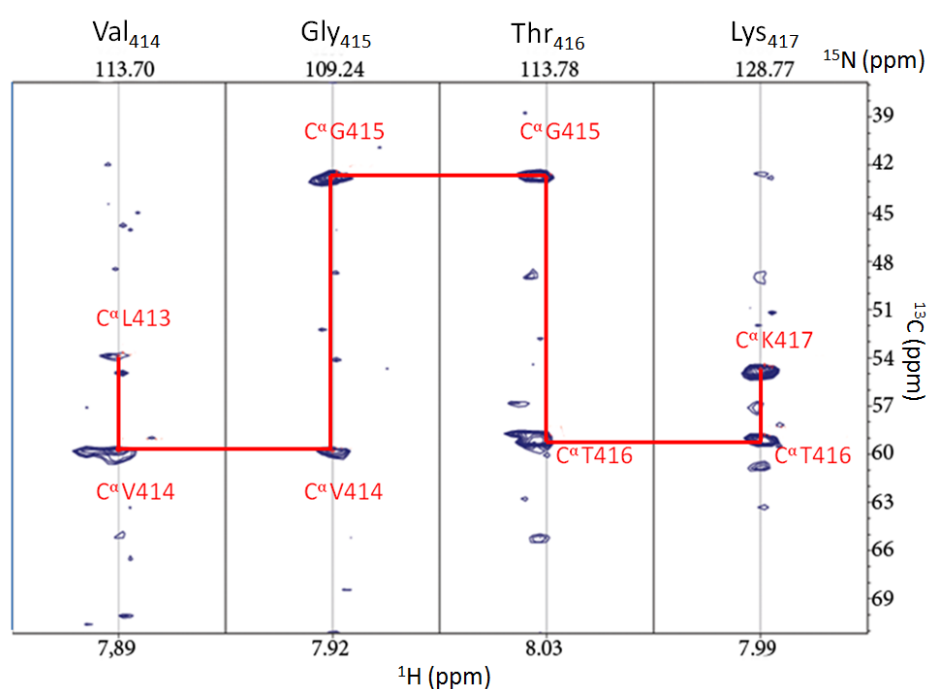


Figura 15. Unión secuencial de los residuos Val₄₁₄-Gly₄₁₅-Thr₄₁₆-Lys₄₁₇ de la construcción HEP₁₅₈₋₄₁₇ utilizando las tiras del espectro HNCA. Las coincidencias en los desplazamientos químicos de las señales C $^\alpha_i$ y C $^\alpha_{i-1}$ se han representado en la figura con una línea roja.

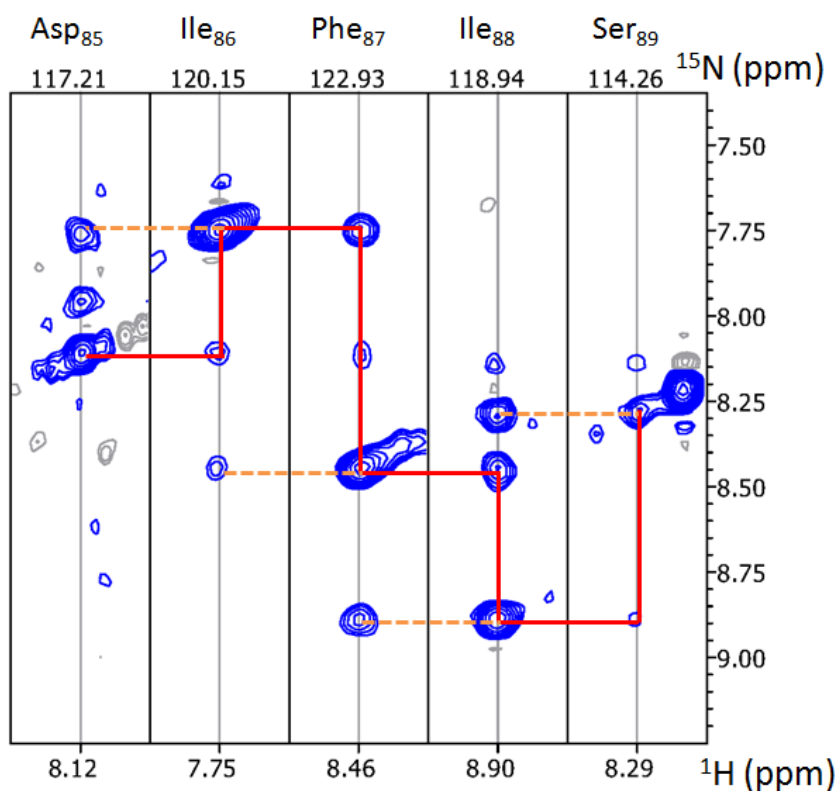


Figura 16. Unión secuencial de los residuos Asp₈₅-Ile₈₆-Phe₈₇-Ile₈₈-Ser₈₉ en la construcción HEP₃₇₋₁₀₉ utilizando los desplazamientos químicos de $^1\text{H}_\text{N}$ - $^{15}\text{N}_\text{H}$ que aparecen en el espectro de $[^1\text{H}, ^{15}\text{N}]$ -HSQC-NOESY. Las correlaciones entre las señales de $(\text{H-N})_i$ - $(\text{H-N})_{i+1}$ se han marcado con una línea roja y las correlaciones $(\text{H-N})_i$ - $(\text{H-N})_{i-1}$ se han marcado con una línea naranja punteada.

En los casos de las construcciones HEP₂₃₃₋₄₁₇ (21 kDa) y HEP₃₇₋₁₀₉ (8 kDa) se asignó el 100 % de la proteína siguiendo este procedimiento. Sin embargo, en el caso de una construcción más grande como HEP₁₅₈₋₄₁₇ de 29 kDa se encontraron problemas de ensanchamiento de señales debido a la relajación transversal de los heteroátomos ^{13}C y ^{15}N y además problemas debidos al solapamiento de señales que era bastante más acusado que en los dos casos anteriores. En este caso se

hicieron varias modificaciones en la preparación de la muestra, espectrómetro usado para la adquisición de los experimentos y por último en la secuencia de pulsos de los espectros tridimensionales para evitar en la medida de lo posible estos problemas. La primera de estas modificaciones fue marcar la proteína con ^2H , para ello se creció la proteína en medio de cultivo preparado en D_2O , consiguiendo así marcar la proteína en aproximadamente un 70%. De esta forma se mejoró el problema de la rápida relajación transversal del heteronúcleo de ^{13}C , principalmente $^{13}\text{C}^\alpha$ que es eficientemente relajado por las interacciones dipolo-dipolo con $^1\text{H}^\alpha$ de la proteína (250, 251). Por otro lado, los experimentos de $\text{HEP}_{158-417}$ se adquirieron en un espectrómetro que operaba a 900 MHz en lugar del 600 MHz utilizado para las otras dos construcciones, lo que mejoró la resolución y sensibilidad de los espectros, y por último se utilizaron las versiones TROSY de todos los experimentos tridimensionales, lo que producía un decrecimiento de la velocidad de la relajación transversal de todos los sistemas ^{15}N - ^1H mediante la interferencia que provoca la secuencia de este experimento en las interacciones dipolo-dipolo y de anisotropía de desplazamientos químicos (224, 227, 247, 252, 253). Al utilizar las versiones TROSY se mejora más de tres veces la sensibilidad de los espectros. En la Fig. 17 se muestra la diferencia entre dos planos ^1H - ^{13}C del experimento tridimensional HNCACB para la construcción $\text{HEP}_{158-417}$ cuando se aplican todos los parámetros anteriormente citados. Se puede observar un aumento del número de señales provocada en parte por la mayor resolución y sensibilidad que conlleva el uso de un espectrómetro de mayor campo y por otro lado debido a la disminución del ensanchamiento de las señales.

Incluso adoptando esta serie de medidas para mejorar la calidad de los espectros solo se consiguió asignar el 80% de los desplazamientos químicos de ^1H y el 70% de ^{13}C debido principalmente al solapamiento de señales.

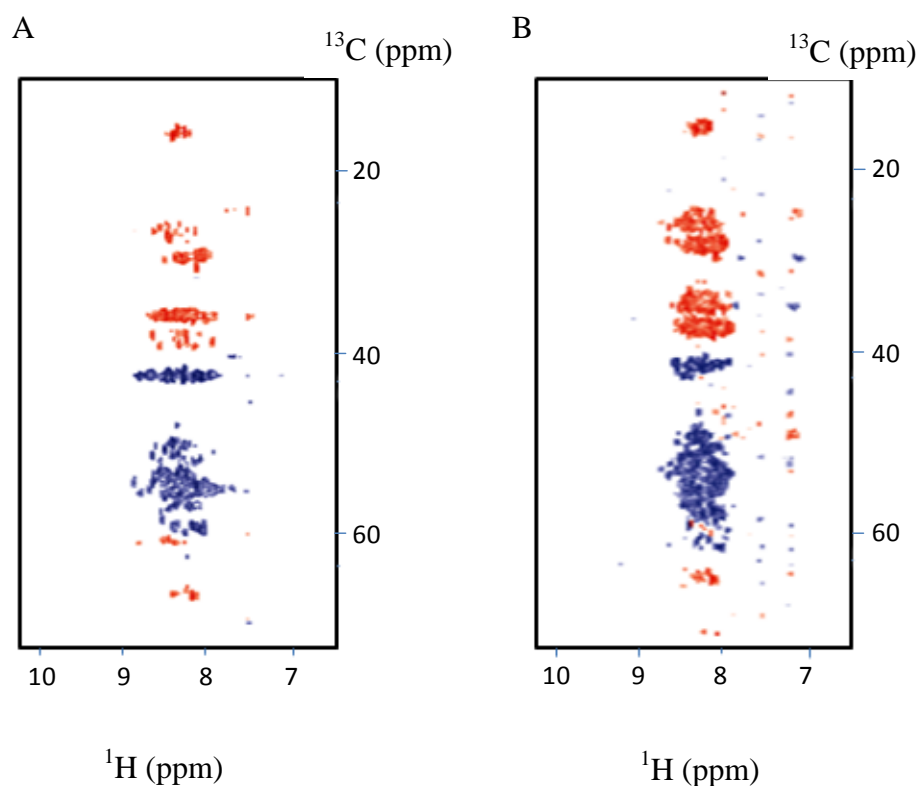


Figura 17. (A) Experimento HNCACB de una muestra marcada en ^{13}C y ^{15}N de la construcción $\text{HEP}_{158-417}$ adquirido en un espectrómetro Bruker 600 MHz con criosonda TCI a 45 °C. (B) Experimento TROSY-HNCACB de una muestra marcada en ^{13}C , ^{15}N y ^2H adquirido en un espectrómetro Bruker 900 MHz con criosonda TCI a 45 °C. La concentración de la proteína en ambos casos fue de 800 μM .

Por esta razón y con el objetivo de completar la asignación de $\text{HEP}_{158-417}$ se marcó la proteína selectivamente con distintos aminoácidos enriquecidos isotópicamente en ^{15}N . Los aminoácidos que se utilizaron fueron aquellos que interesaban para completar la asignación: Lys, Gly, Glu, Phe, Val, Met, Ile, Arg. En muchos de los espectros con la proteína marcada solo con este tipo de aminoácidos se observó que se había producido el fenómeno que se denomina “scrambling” que se produce cuando algunos aminoácidos quedan marcados inespecíficamente, debido al metabolismo de las bacterias donde se expresa la proteína. Sin embargo en todos los casos, las señales correspondientes a estos

aminoácidos marcados inespecíficamente eran menos intensas y podían distinguirse fácilmente. En la Fig. 18 se muestran los marcajes selectivos de Leu, Lys y Val superpuestos en distintas regiones del $[^1\text{H}, ^{15}\text{N}]$ -HSQC. Las señales de color rojo son aquellas que aparecen en el $[^1\text{H}, ^{15}\text{N}]$ -HSQC de la muestra marcada selectivamente, mientras que las que aparecen en el espectro uniformemente marcado aparecen en azul. Al superponer los espectros HSQC de las muestras marcadas uniformemente en ^{15}N y las muestras marcadas selectivamente, se puede identificar fácilmente la naturaleza de los residuos, principalmente cuando estaban superpuestos con otros que ya habían sido identificados. De esta forma se consiguió asignar el 98% del esqueleto y cadena lateral de protón de la proteína. Sin embargo, el porcentaje de cadena lateral de carbono no se pudo mejorar quedando finalmente en un 70%. La asignación de esta construcción se encuentra en la base de datos BMRB <http://www.bmrwisc.edu/> y cuya referencia es 17126

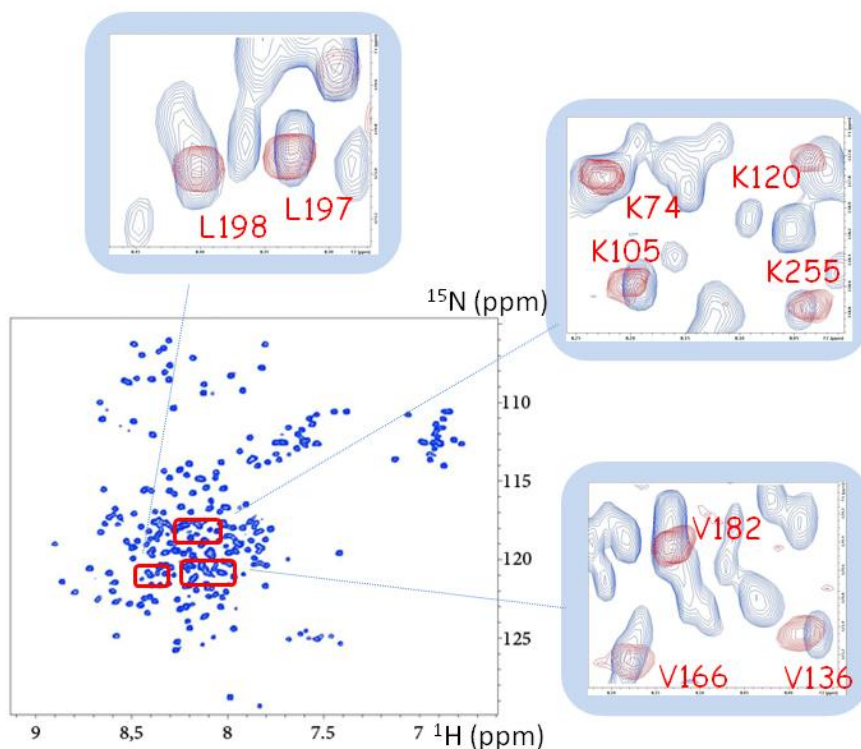


Figura 18. Superposición del marcaje selectivo de Leu, Lys y Val sobre distintas regiones del espectro $[^1\text{H}, ^{15}\text{N}]$ -HSQC para la construcción HEP₁₅₈₋₄₁₇. Las señales de

los aminoácidos marcados en ^{15}N se muestran en rojo mientras que las de la proteína marcada uniformemente en ^{15}N se muestran en azul.

5.- CALCULO DE ESTRUCTURA SECUNDARIA DE HEP₁₅₈₋₄₁₇ Y HEP₃₇₋₁₀₉

5.1.-Estructura secundaria de HEP₁₅₈₋₄₁₇

En primer lugar se determinaron las zonas de la secuencia donde la proteína tenía restricciones de NOE características de hélice-alfa (254). Estos datos se obtuvieron del espectro 3D de RMN ^{15}N -HSQC-NOESY. Estas restricciones eran las siguientes:

- NOE entre H_{Ni} y $\text{H}\alpha_{i-1}$ (d: 3,5 Å)
- NOE entre H_{Ni} y $\text{H}_{\text{Ni}+1}$ (d: 2,8 Å)
- NOE entre H_{Ni} y $\text{H}\alpha_{i-3}$ (d: 3,4 Å)

Los resultados están reflejados en la Fig. 19 donde se puede observar que las zonas con restricciones de NOE para hélice-alfa, representadas en forma de espiral azul, no son muchas y además no implican un número elevado de aminoácidos. Adicionalmente, en la Fig. 19A se pueden comparar estos resultados con los que se obtienen del modelo estructural por homología de heparanasa. Los aminoácidos de la secuencia representados en azul tienen estructura hélice-alfa en el modelo; los aminoácidos representados en verde tienen estructura de hoja plegada β y los representados en rojo no tienen estructura. La mayoría de las zonas de la proteína que tienen restricciones de hélice-alfa también aparecen como hélice-alfa en el modelo. Estas zonas son:

- Val¹⁷⁰-Ser¹⁸⁰
- Ala²⁷⁶-Ala²⁸⁵
- Thr³⁰⁴-Val³²⁹
- Phe³⁵⁹-Met³⁷⁵
- Tyr⁴⁰⁴-Leu⁴¹³

Sin embargo, también se pueden apreciar diferencias entre el resultado experimental y la predicción del modelo, ya que por ejemplo hay zonas que el modelo predice como estructura de hoja plegada β y experimentalmente aparece como estructura hélice-alfa como son las zonas Lys²⁶²-Tyr²⁶⁴, Met²⁸¹-Val²⁸⁴ y His²⁹⁷-Tyr²⁹⁹. Además también hay zonas predichas como “random coil” que presentan estructura secundaria de hélice alfa en RMN principalmente Lys¹⁹²-Ser²⁰¹ y Gly²²³-Phe²²⁹.

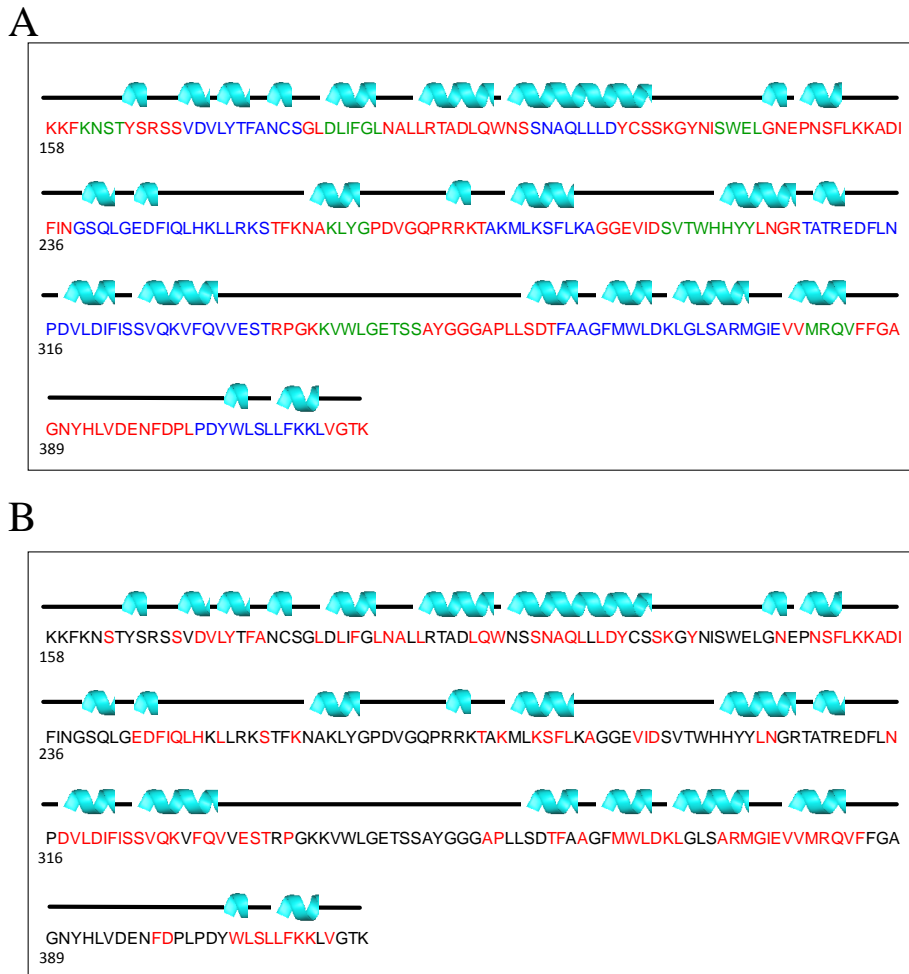


Figura 19. (A) Diferencia entre las zonas con estructura en hélice-alfa determinadas a partir de las restricciones de NOE experimentales del 3D ¹⁵N-HSQC-NOESY de HEP₁₅₈₋₄₁₇ y la estructura secundaria predicha por el modelo por homología de heparanasa. En forma de espirales están representadas las zonas con restricciones de NOE en hélice-

alfa. Las zonas del modelo con estructura secundaria en hélice-alfa están representadas en azul, las zonas con estructura en hoja plegada β están representadas en verde y las zonas sin estructura están representadas en rojo. (B) Diferencia entre las zonas con estructura en hélice-alfa determinadas a partir de las restricciones de NOE experimentales del 3D ^{15}N -HSQC-NOESY (espirales en azul) de HEP₁₅₈₋₄₁₇ y las zonas que podrían tener estructura secundaria estructura secundaria calculada a partir de los ángulos Phi y Psi utilizando el programa TALOS (en rojo).

En la Fig. 19B están representadas las zonas con estructura secundaria calculada a partir de los ángulos Phi y Psi utilizando el programa TALOS. Este programa calcula restricciones de ángulos de torsión de proteínas (ángulos Phi y Psi) a partir de los desplazamientos químicos de los átomos que forman el esqueleto de la proteína (H α , C α , C β , CO y N) junto con información disponible en bases de datos como BMRB y PDB. En esta figura se puede observar que las zonas con estructura secundaria calculadas con TALOS se corresponden en su mayoría con las que se obtienen a partir del experimento 3D ^{15}N -HSQC-NOESY.

Por otro lado, se calculó la diferencia entre los desplazamientos químicos de los C α y H α de HEP₁₅₈₋₄₁₇ y estos desplazamientos para una proteína desplegada o en “random coil” (255, 256). Los resultados se muestran en la Fig. 20 donde se puede observar que las zonas de mayores cambios en los desplazamientos son comunes para C α y H α .

Adicionalmente, en esta figura se puede apreciar que las zonas con mayor diferencia de desplazamientos químicos de C α y H α coinciden en gran medida con las zonas anteriormente definidas con estructura de hélice-alfa. En azul se muestran las zonas con estructura de hélice-alfa calculada a partir de las restricciones de NOEs del experimento de RMN 3D ^{15}N -HSQC-NOESY y a la vez que coinciden con las predichas por el modelo. Estas zonas son de izquierda a derecha Val¹⁷⁰-Ser¹⁸⁰, Ala²⁷⁶-Ala²⁸⁵, Thr³⁰⁴-Val³²⁹, Phe³⁵⁹-Met³⁷⁵, Tyr⁴⁰⁴-Leu⁴¹³. En naranja está representada una zona de la proteína (Gly²³⁹-Phe²⁵⁸) con estructura de hélice-alfa predicha por el modelo y el fragmento de secuencia

representado en verde (Arg¹⁹³-Leu²⁰⁸) corresponde con una secuencia que según las restricciones de NOE tiene estructura secundaria de hélice-alfa.

Estos resultados indican que la mayor parte de las zonas de la proteína que muestran mayor diferencia de desplazamientos químicos C α y H α con respecto a una proteína desplegada tendrían estructura secundaria en hélice-alfa.

Por otro lado, los resultados obtenidos con los tres métodos muestran que al menos un 46% de la proteína presenta una estructura secundaria de hélice-alfa.

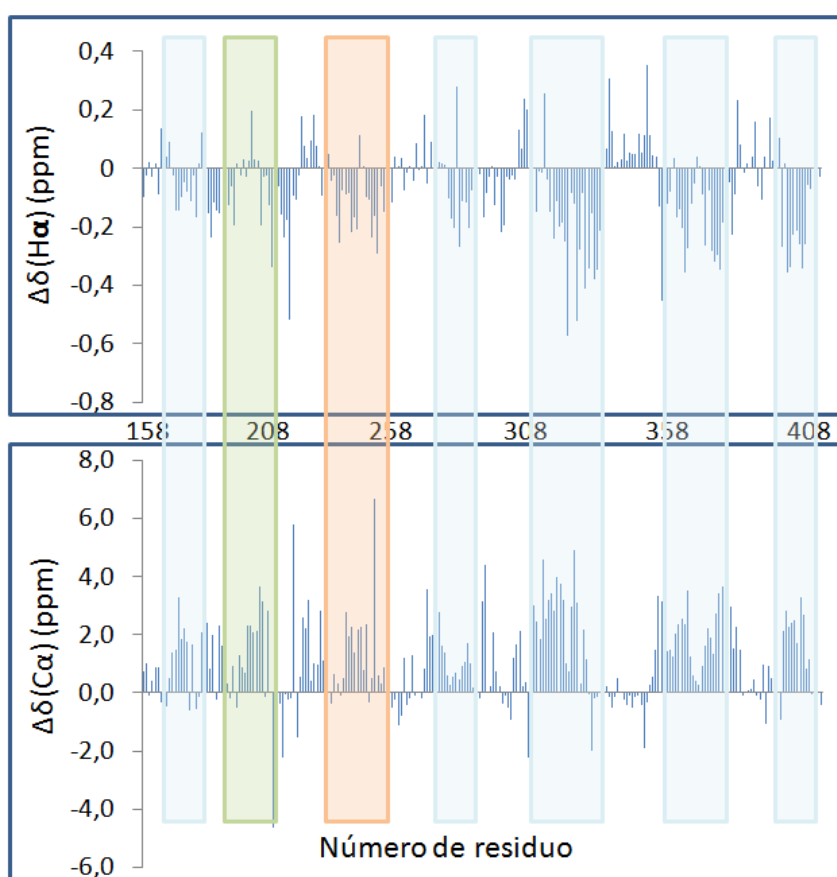


Figura 20. Desviaciones de desplazamientos químicos con respecto a “random coil” para (arriba) los H α y (abajo) los C α de Hep₁₅₈₋₄₁₇. Las desviaciones se muestran para cada residuo de la secuencia. En azul se muestran las zonas con estructura de hélice-alfa calculada a partir de las restricciones de NOEs del experimento de RMN 3D ¹⁵N-HSQC-NOESY y a la vez que coinciden con las predichas por el modelo. Estas zonas

son de izquierda a derecha Val¹⁷⁰-Ser¹⁸⁰, Ala²⁷⁶-Ala²⁸⁵, Thr³⁰⁴-Val³²⁹, Phe³⁵⁹-Met³⁷⁵, Tyr⁴⁰⁴-Leu⁴¹³. En naranja se muestra una zona de la proteína (Gly²³⁹-Phe²⁵⁸) con estructura de hélice-alfa predicha por el modelo y el trozo de secuencia representado en verde (Arg¹⁹³-Leu²⁰⁸) corresponde con una secuencia de la proteína que según las restricciones de NOE tiene estructura secundaria de hélice-alfa.

5.2.-Estructura secundaria de HEP₃₇₋₁₀₉

La estructura secundaria de HEP₃₇₋₁₀₉ se determinó utilizando el mismo método que para HEP₁₅₈₋₄₁₇. Se compararon los resultados obtenidos de TALOS, las diferencias de desplazamientos químicos para C α y H α y aquellos obtenidos a partir de las restricciones de NOEs para hélice-alfa del experimento 3D ¹⁵N-HSQC-NOESY. Los resultados se muestran en las Fig. 21 y 22.

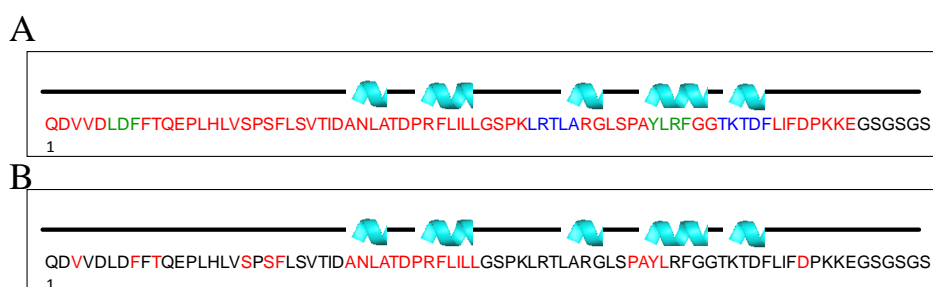


Figura 21. (A) Comparación entre las zonas con estructura en hélice-alfa determinadas a partir de las restricciones de NOE experimentales del 3D ¹⁵N-HSQC-NOESY de HEP₃₇₋₁₀₉ y la estructura secundaria predicha por el modelo por homología de heparanasa. En forma de espirales están representadas las zonas con restricciones de NOE en hélice-alfa. Las zonas del modelo con estructura secundaria en hélice-alfa están representadas en azul, las zonas con estructura en hoja plegada β están representadas en verde y las zonas sin estructura están representadas en rojo. (B) Comparación entre las zonas con estructura en hélice-alfa determinadas a partir de las restricciones de NOE experimentales del 3D ¹⁵N-HSQC-NOESY de HEP₃₇₋₁₀₉ y las zonas con estructura secundaria calculada con el programa TALOS. En forma de espirales están representadas las zonas con restricciones de NOE en hélice-alfa. Las zonas de HEP₃₇₋₁₀₉ con estructura secundaria calculada a partir de los ángulos Phi y Psi utilizando el programa TALOS en rojo.

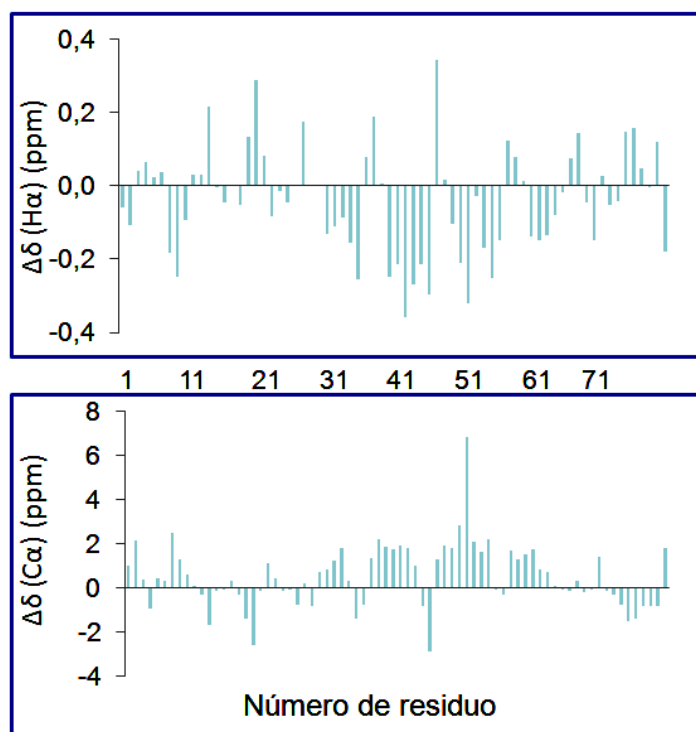


Figura 22. Desviaciones de desplazamientos químicos con respecto a “random coil” para (arriba) los $\text{H}\alpha$ y (abajo) los $\text{C}\alpha$ de HEP_{36-109} . Las desviaciones se muestran para cada residuo de la secuencia.

Experimentalmente, hay tres zonas diferenciadas con estructura secundaria, $\text{Asn}^{29}\text{-Leu}^{40}$, $\text{Ala}^{40}\text{-Gly}^{51}$ y $\text{Tyr}^{56}\text{-Gly}^{60}$ (Fig. 21A) que se obtienen a partir de las restricciones de NOE para estructura hélice-alfa. Estas tres zonas coinciden con las mayores desviaciones de desplazamientos químicos $\text{C}\alpha$ y $\text{H}\alpha$ de la proteína frente a los valores de desplazamientos químicos “random coil” según se puede observar en la Fig. 22 y las zonas con estructura secundaria obtenidas con el programa TALOS (Fig. 21B). Sin embargo, no hay ninguna similitud con la estructura secundaria predicha por el modelo para esta zona de la proteína. Por lo tanto se puede concluir que el modelo estructural para esta zona de la heparanasa (HEP_{37-109}) no predice los resultados experimentales de estructura secundaria. Esto podría deberse a la falta de homología entre esta zona de la

heparanasa con la proteína usada para el cálculo del modelo. El uso de la estructura de una proteína de mayor homología de secuencia con la heparanasa permitiría la obtención de un modelo con el que se obtendrían resultados más próximos a los resultados experimentales. Sin embargo, no existe hasta el momento, una estructura de proteína que presente un mayor grado de homología a la heparanasa.

5.3.-Asignación de NOEs y determinación de estructura tridimensional de HEP₁₅₈₋₄₁₇

Los espectros de RMN que se utilizaron para poder calcular la estructura de la proteína, en este caso de la construcción de heparanasa HEP₁₅₈₋₄₁₇, fueron el experimento 3D ¹⁵N-HSQC-NOESY que da información principalmente sobre la estructura secundaria de la proteína aunque puede haber algunos NOEs de larga distancia y el experimento 3D ¹³C-HSQC-NOESY que da información tanto de la estructura secundaria como tridimensional de la proteína. Para poder calcular la estructura lo primero que se hizo fue intentar asignar los NOEs de cada uno de estos experimentos. En primer lugar se llevó a cabo la asignación de los NOEs de forma automática utilizando el programa Cyana 2.1 (<http://www.cyana.org>). Para poder llevar a cabo esta asignación automática, se marcaron cada una de las señales de ambos espectros utilizando como soporte informático el programa de asignación de espectros de RMN, CARA (Computer-Aided Resonance Assignment) (www.nmr.ch) y con este mismo programa se generaron los ficheros que contenían el listado de desplazamientos químicos para cada una de estas señales marcadas en ambos espectros. También se generaron los ficheros que contenían información de la asignación tanto del esqueleto como de las cadenas laterales de la proteína. Aunque esta información debe ser la misma para ambos espectros, se generaron dos ficheros diferentes por las pequeñas diferencias de desplazamientos químicos que se pueden encontrar en cada uno de los espectros debido principalmente a la utilización de muestras diferentes o pequeñas variaciones de temperatura. Con toda esta información junto con el fichero generado por el programa TALOS con los ángulos entre los diferentes

aminoácidos de la proteína, se intentó hacer la asignación automática de los NOEs y a su vez el cálculo de la estructura. Los problemas que surgieron fueron los siguientes:

- La calidad del espectro 3D ^{13}C -HSQC-NOESY era insuficiente para la asignación de los NOEs debido a que la presencia de detergente en la muestra (N-lauroyl-sarcosine) daba lugar a la presencia de señales en el espectro que solapaban con muchos de los NOEs entre aminoácidos de la proteína y además los -CH de las moléculas del detergente daban a su vez NOEs con los aminoácidos de la proteína que estaban interaccionando con dicho detergente.
- El espectro 3D ^{15}N -HSQC-NOESY contenía mayoritariamente NOEs entre aminoácidos próximos en la secuencia y muy pocos NOEs de larga distancia por lo que sólo con este espectro no era posible calcular la estructura tridimensional.
- Al tratarse de una proteína de 29 kDa, había un gran solapamiento de señales en el espectro 3D ^{13}C -HSQC-NOESY, por lo que al problema anterior se sumaba el hecho de que muchos de las señales correspondientes a la cadena lateral de aminoácidos del mismo tipo (no solapados por las señales del detergente) tenían el mismo desplazamiento químico y por lo tanto no había una asignación única y segura para la mayoría de los NOEs de este espectro.

Estos problemas se traducían a que el cálculo de estructura daba un RMSD muy alto, del orden de 10 Å.

Se intentó hacer la asignación manual pero se tuvo el mismo problema que con la asignación automática. La mayoría de los NOEs asignados pertenecían al espectro 3D ^{15}N -HSQC-NOESY que como ya hemos dicho anteriormente daba mayoritariamente información de la estructura secundaria.

Este solapamiento se debía no solo al tamaño de la proteína sino también a que como hemos visto en apartados anteriores de este trabajo un 60% de la proteína

era “random coil” y un 40% aproximadamente tenía estructura en hélice-alfa, por lo que el rango de desplazamientos químicos de los aminoácidos de la proteína era mucho más reducido que para proteínas que tienen un tanto por ciento mayor de estructura secundaria, principalmente aquellas proteínas que tienen mayoritariamente estructura de hoja β .

6.-OPTIMIZACIÓN DE LAS CONDICIONES DE SOLUBILIDAD Y ESTABILIDAD PARA LOS ESTUDIOS DE SPR

Al igual que para los estudios de RMN, hubo que optimizar las condiciones de la muestra (concentración de proteína, tampón, temperatura, presencia de detergente) para abordar los experimentos mediante SPR. Para ello, la construcción que se utilizó como referencia fue HEP₁₅₈₋₄₁₇ que es la misma que se eligió para llevar a cabo los estudios de interacción por RMN. Se usó esta construcción porque como ya se ha comprobado en apartados anteriores es la construcción de mayor peso molecular que contiene el sitio catalítico de la enzima y los dos sitios de unión a heparán sulfato con la que se obtienen espectros de RMN de calidad suficiente para poder llevar a cabo este tipo de experimentos.

Inicialmente se utilizó un chip NTA que es un chip con una matriz de dextrano carboximetilado pre-inmovilizado con ácido nitrilotriacético. Al utilizar este chip, la proteína era capturada sobre una superficie activada con NiCl₂ debido a la afinidad existente entre la cola de histidinas que tenía fusionada la proteína y los cationes Ni²⁺ (257, 258). Sin embargo, la afinidad de una única cola de histidinas por el Ni²⁺ inmovilizado en este chip es relativamente débil ($K_D \sim 0,5-3 \mu\text{M}$) y después de la inyección la línea base no se mantenía estable probablemente debido a la disociación de la cola de histidinas de la superficie del chip (257). Además, a este problema se une el hecho de que la presencia de detergente en la muestra tampoco favorecía esta interacción. Por tanto, en este caso, el chip NTA no era óptimo para medir y cuantificar interacciones con heparanasa.

Posteriormente se intentó inmovilizar la proteína covalentemente en un chip CM5, que contiene una matriz de dextrano carboximetilada; el 5 hace referencia al nivel de carboximetilación y la longitud de sus cadenas de dextransos (259). Lo primero que se optimizó fue el tampón de la proteína para que ésta pudiera ser correctamente inmovilizada. Se tenía que usar un tampón en el que la proteína quedara cargada positivamente, por lo tanto se debía mantener el pH por debajo del punto isoeléctrico de la proteína. En este caso el pI de la proteína era 9, por lo que se podía usar cualquier tampón con pH por debajo de este valor. En primer lugar se utilizó 25 mM acetato sódico pH 5,5, 50 mM NaCl, pero al igual que se ha comentado en el apartado 3 de este capítulo, la proteína en estas condiciones no era estable durante más de 12 horas por lo que se probó el mismo tampón utilizado para RMN, 25 mM fosfato sódico pH 6 50 mM NaCl, con buenos resultados. Para inmovilizar la proteína lo primero que se usó fue una mezcla de NHS/EDC (N-hidroxisuccinimida/hidrocloruro de carbodiimida) para la activación de la superficie del chip. Posteriormente se llevó a cabo la inyección de la muestra y los resultados fueron que para una concentración de proteína de 50 µg/mL y un flujo de 420 µL/min a 25 °C se obtenían niveles de inmovilización de entre 2000-3000 RU (Fig. 23). Si se quería obtener un nivel de inmovilización más alto, la proteína podía ser inyectada varias veces. Finalmente la superficie del chip fue bloqueada utilizando etanolamina pH 8,5. Esta proteína era estable a 25 °C durante varias semanas por lo que no fue necesario bajar la temperatura para inmovilizarla a la superficie del chip. Normalmente, cuando las proteínas no son estables a 25 °C, la inmovilización se hace a baja temperatura que tiene el inconveniente de que la inmovilización de la proteína es más lenta, requiere mayor tiempo de contacto con la superficie de chip y por lo tanto mayor cantidad de muestra (Biacore Sensor Surface Handbook). Por lo tanto en nuestro caso fue una ventaja poder trabajar a 25 °C.

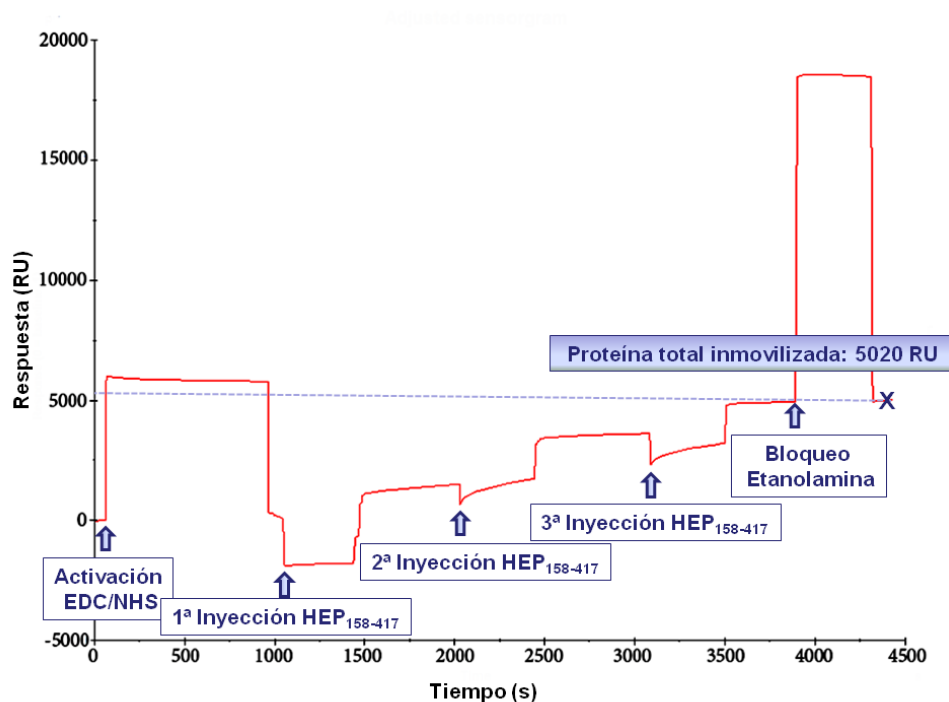


Figura 23. Inmovilización de *Hep*₁₅₈₋₄₁₇ en un chip CM5 utilizando un instrumento Biacore T100 a 25 °C. Cada etapa está señalada en recuadros: activación del chip con EDC/NHS durante 700 s seguido de tres inyecciones de *Hep*₁₅₈₋₄₁₇ a una concentración de 50 µg/mL en 25 mM fosfato sódico, 50 mM NaCl y 0,025% (p/v) de N-lauroyl-sarcosine durante 420 s cada una y finalmente bloqueo de la superficie del chip con etanolamina pH 8,5 durante 420 s.

A continuación, se optimizaron las condiciones del experimento de interacción con inhibidores conocidos de heparanasa como la suramina y el fondaparinux (151, 260). Lo primero que se comprobó fue si la proteína unida covalentemente al chip necesitaba la presencia de detergente para ser estable. Para ello se realizaron tres experimentos a la misma temperatura y concentración de ligando pero utilizando 0,25%, 0,05% y 0% de N-lauroyl sarcosine, en el tampón utilizado durante el experimento. Los resultados fueron que en el primer caso no había interacción y los dos siguientes, la constante de disociación era la misma. Por lo tanto se eliminó el detergente del tampón en experimentos posteriores.

Además se tuvo también que optimizar el rango de concentración del ligando para llevar a cabo este experimento de interacción en el Biacore. Normalmente, este rango oscila alrededor del valor de la constante de disociación del complejo.

El experimento de interacción también se llevó a cabo a dos temperaturas, 25 y 37 °C ya que esta última se asemejaba más a la temperatura empleada para los experimentos de RMN. Como se puede observar en la Fig. 24, la cinética es más lenta a la temperatura más baja y la actividad de interacción de la proteína es superior en el experimento realizado a 25 °C. La actividad de la proteína se pudo calcular por la siguiente fórmula:

$$R_{\max} = r \times (PM_{\text{ligando}}/PM_{\text{proteína}}) \times R_{\text{real}}$$

donde R_{\max} es la respuesta máxima que el ligando puede dar y corresponde con el 100% de actividad de interacción de la proteína en la superficie del chip, r es el ratio estequiométrico de la interacción, PM_{ligando} es el peso molecular del ligando, $PM_{\text{proteína}}$ es el peso molecular de la proteína y R_{real} es el nivel de inmovilización de la proteína en la superficie del chip.

En el caso de la interacción con suramina $r=1$, $PM_{\text{suramina}}=1429$ Da, $PM_{\text{proteína}}=31000$ Da ($PM_{\text{proteína}}=29000$ Da + cola de histidinas) y $R_{\text{real}}=5020$ RU. Por lo tanto, el 100% de actividad se lograría cuando el ligando diera una respuesta de 231,4 RU para la concentración de ligando a la que se llegue a la saturación.

En este experimento (Fig. 24) se consiguió una respuesta de 215 RU a 25 °C a una concentración de 1 μ M de suramina (como se puede observar en la figura a 2 μ M de suramina se ha llegado a la saturación), mientras que a 37 °C se obtenía una respuesta de 108 RU para la misma concentración de suramina. Esto significaba que la proteína inmovilizada a la superficie del chip tenía una actividad del 92,9% a 25 °C y del 46,7% a 37 °C, lo que demostraba que la proteína a 25 °C y con las condiciones descritas para este experimento estaba bien plegada y con sus dos sitios de unión accesibles al ligando. El modelo utilizado para el ajuste de las curvas se hizo teniendo en cuenta una interacción

1:1. En la Fig. 24 se puede observar como la cinética se ajustaba perfectamente a este modelo.

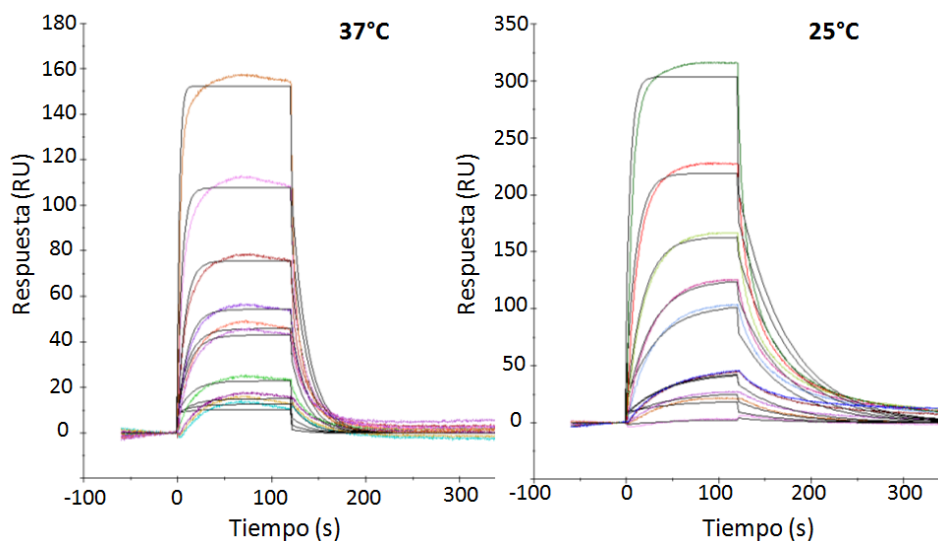


Figura 24. *Sensogramas que muestran la cinética de la interacción entre HEP₁₅₈₋₄₁₇ y la suramina a 25 °C (derecha) y 37 °C (izquierda). En este caso se utilizó un chip CM5 donde se había inmovilizado covalentemente HEP₁₅₈₋₄₁₇. Cada curva representa la interacción con suramina a una concentración diferente. Las concentraciones que se usaron, indicadas desde abajo hacia arriba, fueron 0,01, 0,025, 0,05, 0,1, 0,25, 0,5, 1 y 2 μM. La concentración de 0,1 μM se hizo por duplicado. Estos experimentos fueron realizados con un Biacore T100.*

7.-VALIDACIÓN DE LA CONSTRUCCIÓN HEP₁₅₈₋₄₁₇ COMO DIANA DE CRIBADO POR RMN Y SPR

Con el objetivo de saber si HEP₁₅₈₋₄₁₇ podía ser utilizada para identificar nuevos inhibidores de heparanasa sustituyendo a la enzima completa HEP₃₇₋₅₄₃, hubo que validar esta construcción para estos fines. En este sentido se utilizaron dos técnicas biofísicas, la Resonancia Magnética Nuclear y la Resonancia de Plasma Superficial para comprobar en primer lugar si existía interacción de esta

construcción de heparanasa con inhibidores conocidos de la enzima y en segundo lugar para saber si esta interacción era específica y con una afinidad similar a la encontrada en la bibliografía para la enzima completa.

Los experimentos de Resonancia Magnética Nuclear fueron adquiridos en presencia de suramina y heparina. La suramina es un polisulfonato naftilurea conocido como agente antiproliferativo que inhibe la actividad de heparanasa *in vitro* ($IC_{50} = 46 \mu\text{M}$) (260). La heparina es un glicosaminoglicano altamente sulfatado utilizado comúnmente como un agente anticoagulante y que además se ha demostrado que inhibe la actividad de heparanasa (260). La validación por RMN se basó en experimentos de tipo STD (203). Estos experimentos permiten medir la interacción proteína-ligando debido a la transferencia de magnetización desde la proteína al ligando unido a través de un pulso de radiofrecuencias. Se detectan las resonancias correspondientes al ligando libre que está en equilibrio con el ligando unido a la proteína. La figura 25 muestra el experimento STD adquirido para una mezcla HEP₁₅₈₋₄₁₇-suramina en relación 1:100. El resultado positivo de este experimento demostró la interacción entre suramina y esta construcción. Resultados similares se obtuvieron para la interacción entre esta construcción y la heparina. Aunque los experimentos de STD presentan algunas ventajas sobre otros experimentos de interacción tales como la baja concentración de proteína requerida (1-5 μM) y el uso de proteína no marcada, no informan acerca del sitio de unión a la proteína donde se produce la interacción con la consecuencia de que pueden aparecer uniones inespecíficas como falsos positivos.

Con el objetivo de determinar si esta interacción se estaba produciendo en el sitio catalítico de la enzima y por lo tanto demostrar que esta interacción era específica y selectiva se utilizó el espectro 2D [¹H-¹⁵N]-HSQC para el experimento de mapeo de desplazamientos químicos. Para este tipo de experimentos se necesitaba que la proteína estuviera previamente asignada (ver apartado 4). Se realizó una titulación mediante RMN incrementando la concentración de ligando (suramina o heparina) y dejando la proteína con la misma concentración inicial. La relación proteína/ligando fue 1:1, 1:2, 1:5, 1:10,

1:25 y 1:50. Esta metodología permitió la monitorización de los cambios en algunas de las señales del esqueleto de la proteína debido a la interacción con la suramina o heparina. En este caso fue necesaria una relación proteína:ligando de 1:50 probablemente debido al efecto del N-Lauroyl-sarcosine utilizado para solubilizar la proteína que posiblemente bloqueaba el acceso del ligando a la superficie de la misma.

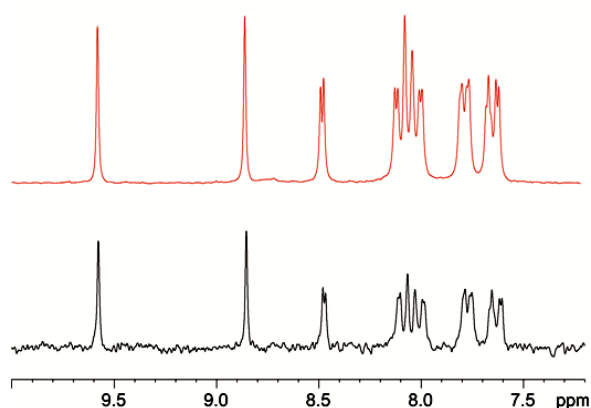
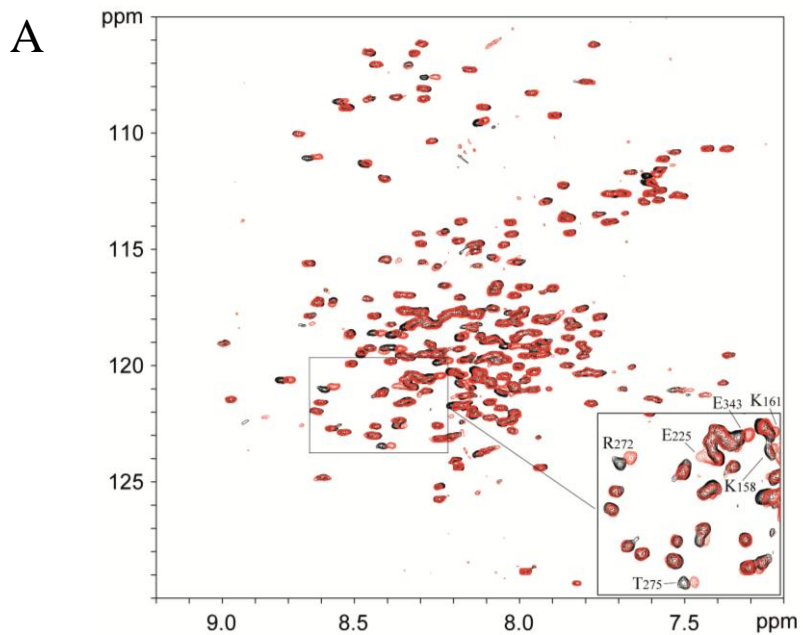


Figura 25. Validación de $HEP_{158-417}$ por RMN. En rojo se muestra el espectro 1D de 1H donde se pueden observar las señales aromáticas de suramina; en negro se muestra la misma región para el experimento de STD de la mezcla $HEP_{158-417}$ -suramina en relación 1:100. Este experimento fue adquirido en un espectrómetro Bruker 600 MHz con criosonda TCI con un total de 1K acumulaciones.

Tanto la suramina como la heparina provocaron cambios de desplazamientos químicos en los ácidos glutámicos que forman el sitio catalítico de la enzima (Glu^{225} y el Glu^{343}), así como los aminoácidos implicados en la unión con el sustrato natural de la enzima, el heparán sulfato (Lys^{158} - Asp^{171} y Gln^{270} - Lys^{280}). En la Fig. 26A se pueden observar los cambios producidos por la suramina. En esta figura se muestra ampliado el cambio de desplazamiento químico de los dos ácidos glutámicos activos así como de algunos de los aminoácidos implicados en la unión a heparán sulfato, Lys^{158} and Lys^{161} localizados en el primer sitio de unión y Arg^{272} and Thr^{275} en el segundo sitio de unión.



B

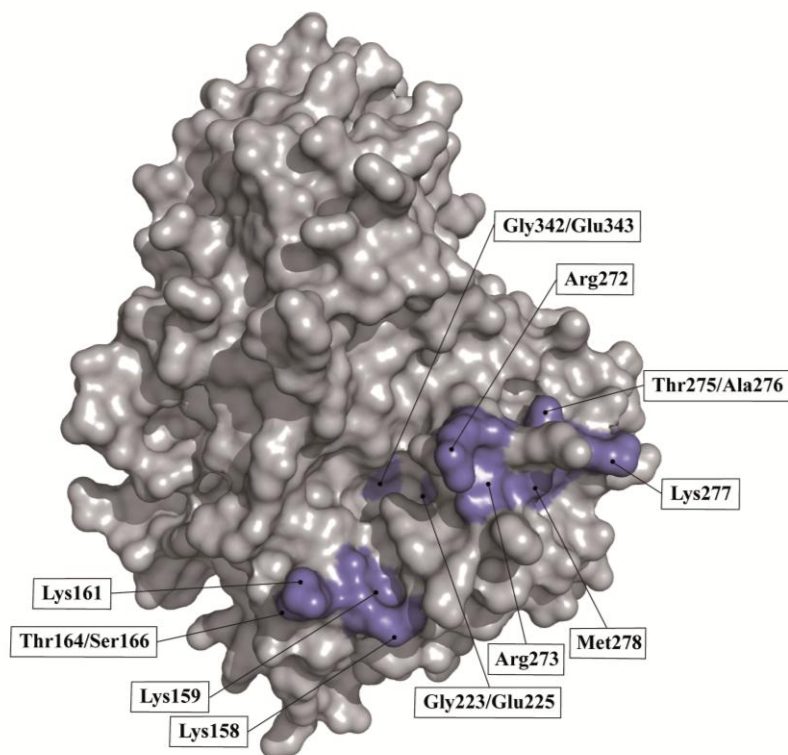


Figura 26. Mapeo de desplazamientos químicos. (A) Espectro 2D [^1H - ^{15}N]-HSQC (600 MHz, criosonda TCI). El espectro de HEP₁₅₈₋₄₁₇ en ausencia de suramina se muestra en negro y la mezcla HEP₁₅₈₋₄₁₇-suramina (1:50) en rojo. La región del espectro ampliada muestra los cambios de desplazamientos químicos observados para el sitio catalítico Glu²²⁵, Glu³⁴³ y algunos de los residuos implicados en el sitio de unión a heparán sulfato: Lys¹⁵⁸, Lys¹⁶¹, Arg²⁷² y Thr²⁷⁵. (B) Mapeo de las principales perturbaciones de desplazamientos químicos causadas por la suramina sobre el modelo estructural por homología de heparanasa.

En la Fig. 26B se pueden observar dónde están localizados estos aminoácidos en el modelo estructural de heparanasa. Tanto estos resultados como los del STD demuestran la especificidad y selectividad de la interacción entre suramina con la construcción de HEP₁₅₈₋₄₁₇ y por tanto confirman la validez de esta construcción para la búsqueda de nuevos inhibidores de la enzima.

Por otro lado, se aplicó una metodología complementaria para medir cuantitativamente la afinidad de la interacción. Para ello, se utilizó la Resonancia de Plasma Superficial (SPR) en la que la proteína fue inmovilizada a la superficie de un chip mientras que el ligando en solución pasaba a través del chip donde la proteína había sido inmovilizada. De esta forma se pudo medir la constante de disociación entre HEP₁₅₈₋₄₁₇ y la suramina.

Como se mencionó en el apartado 6 de este capítulo se utilizó el mismo tampón utilizado para RMN (25 mM fosfato sódico pH 6, 50 mM NaCl) para inmovilizar la proteína, sin embargo para medir la interacción se utilizó el tampón recomendado por Biacore para llevar a cabo estos experimentos: 10 mM HEPES pH 7,4, 150 mM NaCl, 50 μM EDTA y 0,005% P20. Contrariamente a lo que ocurre en RMN, fue posible llevar a cabo los experimentos de SPR en ausencia de N-Lauroyl-sarcosine ya que la construcción HEP₁₅₈₋₄₁₇ fue covalentemente unida a la superficie del chip y era estable en esas condiciones.

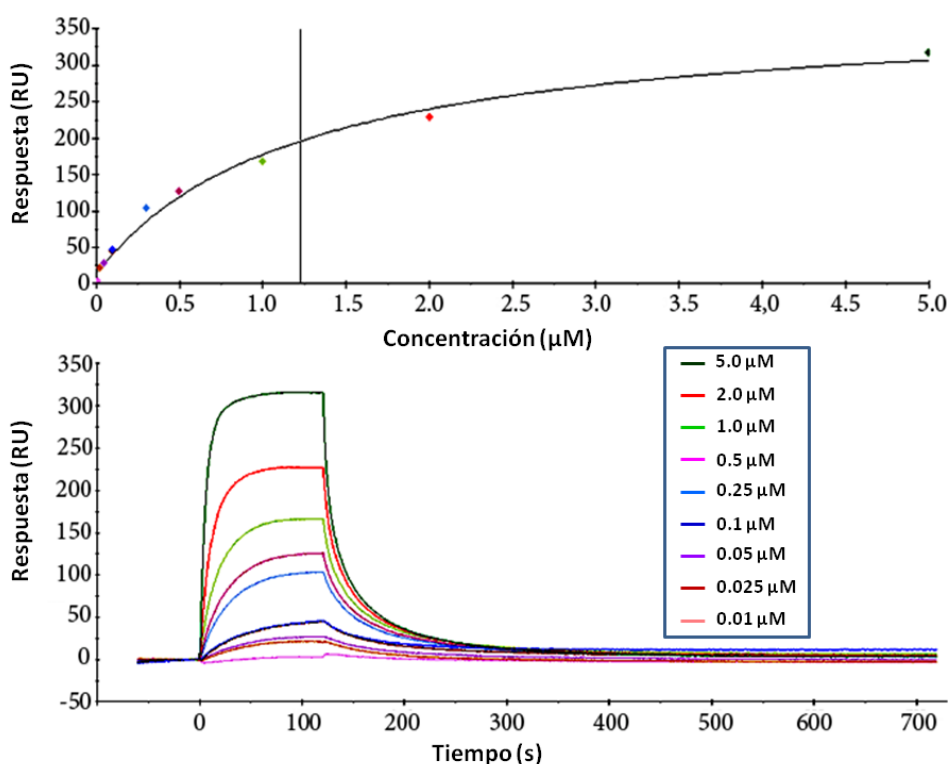


Figura 27. Medida de afinidad de la interacción $HEP_{158-417}$:suramina mediante SPR. Se muestra en la parte de abajo el sensograma utilizado para las constantes cinéticas y en la parte superior el ajuste para el cálculo de la constante de disociación. Las concentraciones de suramina que se usaron indicadas desde abajo hacia arriba fueron 0,01, 0,025, 0,05, 0,1, 0,25, 0,5, 1, y 5 μM . La concentración de 0,1 μM se hizo por duplicado. Estos experimentos fueron realizados con un instrumento Biacore T100.

La proteína fue inmovilizada en un chip sensor carboximetilado CM5 con un nivel de 5020 RU. Para llevar a cabo el experimento de interacción con suramina se usó un rango de concentraciones de 0,01-5 μM . La constante de disociación que se midió para la interacción $HEP_{158-417}$ -suramina fue 0,5 μM que es comparable a la IC_{50} publicada en la bibliografía para la heparanasa humana completa recombinante (260, 261) (Fig. 27). Estos resultados confirman que $HEP_{158-417}$ es una construcción válida como diana para encontrar nuevos inhibidores de la enzima.

Con el objetivo de comprobar si la presencia de los dos sitios de unión a heparán sulfato es necesaria para buscar nuevos inhibidores, se midió la interacción entre la suramina y la construcción HEP₂₃₃₋₄₁₇ que como se ha dicho en apartados anteriores solo contenía un sitio de unión a heparán sulfato (Gln²⁷⁰-Lys²⁸⁰) y un solo ácido glutámico activo de la enzima Glu³⁴³. Esta interacción se midió tanto por RMN como por SPR en las mismas condiciones de concentración y temperatura que para HEP₁₅₈₋₄₁₇.

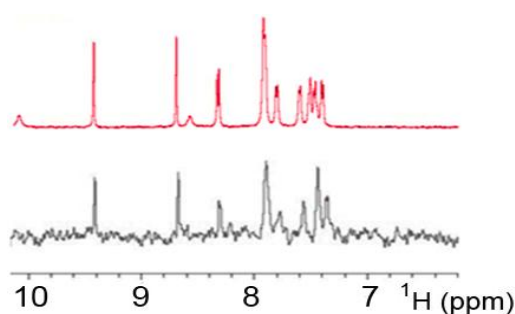


Figura 28. Validación de HEP₂₃₃₋₄₁₇ por RMN. En rojo se muestra el espectro 1D de ¹H donde se pueden observar las señales aromáticas de suramina; en negro se muestra la misma región para el experimento de STD de la mezcla HEP₂₃₃₋₄₁₇-suramina.

En la Fig. 28 se muestra el experimento STD y en la Fig. 29 se puede observar el experimento 2D [¹H,¹⁵N]-HSQC que dan información de la interacción entre esta construcción y la suramina. La interacción como se puede ver en el STD es positiva y además es selectiva como puede observarse en el [¹H,¹⁵N]-HSQC ya que todos los residuos implicados en el segundo sitio de interacción de heparanasa, Gln²⁷⁰-Lys²⁸⁰, cambian sus desplazamientos químicos en presencia de suramina. En la Fig. 29 están representados los cambios en los residuos Arg²⁷², Arg²⁷³, Lys²⁷⁴, Thr²⁷⁵ y Ala²⁷⁶. Estos cambios en desplazamientos químicos son de la misma proporción o incluso mayor que en el caso de la construcción que contenía los dos sitios catalíticos. Para completar estos resultados se midió la constante de disociación de HEP₂₃₃₋₄₁₇ con suramina por SPR.

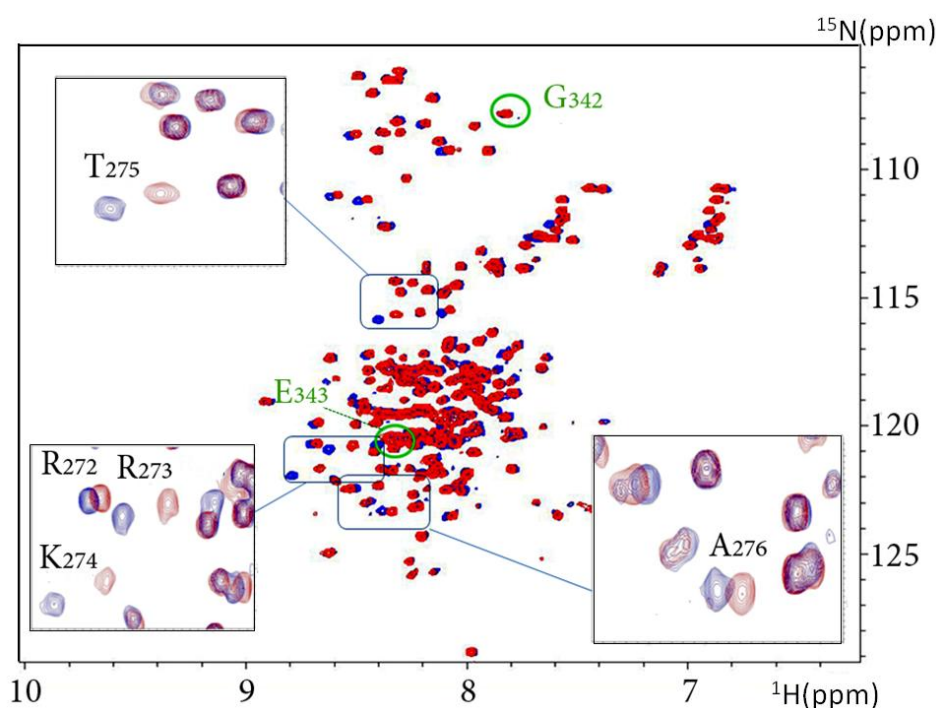


Figura 29. Mapeo de desplazamientos químicos para la construcción HEP₂₃₃₋₄₁₇. Espectro 2D [¹H,¹⁵N]-HSQC (600 MHz, criosonda TCI). El espectro de HEP₂₃₃₋₄₁₇ en ausencia de suramina se muestra en azul y la mezcla HEP₂₃₃₋₄₁₇-suramina (1:50) en rojo. La región del espectro ampliada muestra los cambios de desplazamientos químicos observados para algunos de los residuos implicados en el sitio de unión a heparán sulfato: Arg²⁷², Arg²⁷³, Lys²⁷⁴, Thr²⁷⁵ y Ala²⁷⁶. El Glu³⁴³ del sitio catalítico de la enzima y el residuo que le precede en la secuencia Gly³⁴² aparecen representados dentro de un círculo verde.

La constante de disociación de la interacción HEP₂₃₃₋₄₁₇-suramina medida por SPR dio un valor de K_D : 0,5 μ M (Fig. 30) que es el mismo que el obtenido para la interacción HEP₁₅₈₋₄₁₇-suramina. Para poder entender estos resultados se midieron las constantes cinéticas de las dos construcciones (HEP₁₅₈₋₄₁₇ y HEP₂₃₃₋₄₁₇) con suramina que se obtenía del experimento realizado por SPR. Las constantes cinéticas resultaron ser diferentes a pesar de que la constante de disociación era la misma en los dos casos (Tabla 3), siendo la k_{on} y la k_{off} más pequeñas en el caso de la construcción con solo un sitio de unión a heparán

sulfato. El hecho de que la interacción de HEP₂₃₃₋₄₁₇ con suramina tenga una k_{off} más pequeña implica que el tiempo en el que la suramina está unida a la proteína es más largo en este caso que en el caso de la interacción entre HEP₁₅₈₋₄₁₇ y suramina, lo que podría explicar que los cambios de desplazamientos químicos del HSQC sean mayores en el caso de la interacción entre HEP₂₃₃₋₄₁₇ y suramina. Esto también podría explicar el hecho de que las señales del STD para la interacción HEP₂₃₃₋₄₁₇-suramina sea más débil que para la interacción HEP₁₅₈₋₄₁₇-suramina ya que en el STD lo que se detecta es la magnetización del ligando libre. A esto se une el hecho de que HEP₂₃₃₋₄₁₇ es una proteína de 21 kDa en comparación con los 29 kDa de HEP₁₅₈₋₄₁₇, que da problemas en la saturación de la proteína por difusión de espín (262).

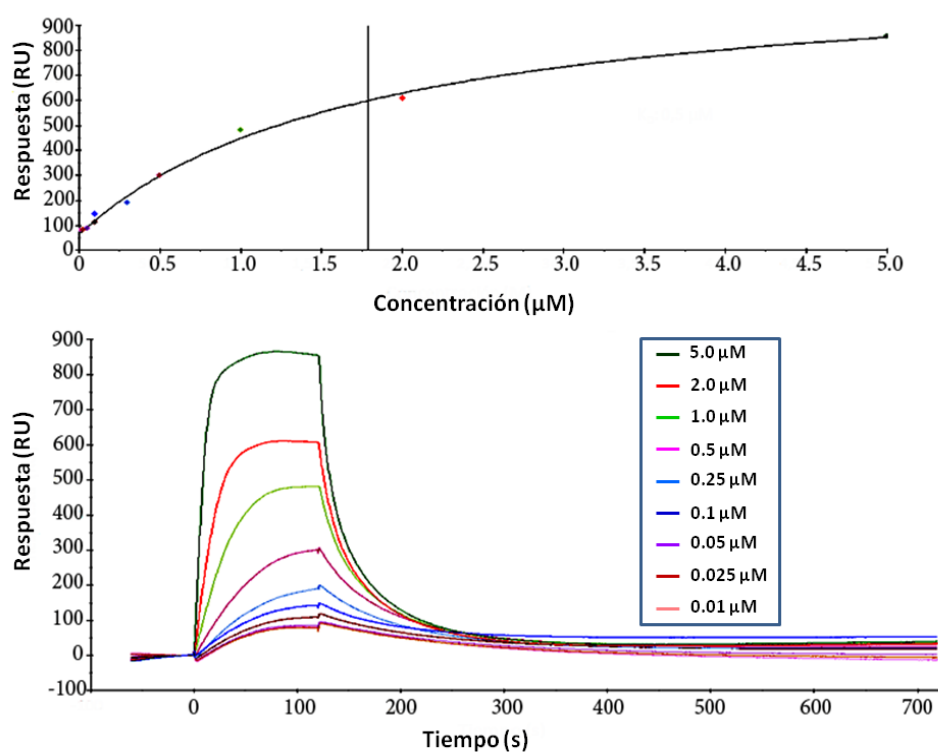


Figura 30. Medida de afinidad de la interacción HEP₂₃₃₋₄₁₇:suramina mediante SPR. Se muestra en la parte de abajo el sensograma utilizado para las constantes cinéticas y en la parte superior el ajuste para el cálculo de la constante de disociación. Las

concentraciones que se usaron indicadas desde abajo hacia arriba fueron 0,01, 0,025, 0,05, 0,1, 0,25, 0,5, 1, 2 y 5 μM . La concentración de 0,1 μM se hizo por duplicado. Estos experimentos fueron realizados con un instrumento Biacore T100.

Todos los resultados apuntaban a que no era necesario tener los dos sitios de unión a heparán sulfato; sin embargo el hecho de que el ácido glutámico activo Glu³⁴³ y el residuo que le precede Gly³⁴² de HEP₂₃₃₋₄₁₇ no cambien su desplazamiento químico en presencia de suramina (Fig. 29B indicados con un círculo verde) podría llevarnos a la hipótesis de que esta construcción mantenía la afinidad por el ligando, pero podría no conservar la buena orientación del mismo. El hecho de que existiera esta posibilidad, justificaba el hecho de que se utilizara para los estudios de interacción la construcción con los dos sitios de unión a heparán sulfato y los dos ácidos glutámicos del sitio catalítico.

Tabla 3. Resumen de las constantes cinéticas para las construcciones HEP₁₅₈₋₄₁₇ y HEP₂₃₃₋₄₁₇

Construcción de heparanasa	K_{ON} (1/M·s)	K_{OFF} (1/s)	K_{D} (μM)
HEP ₁₅₈₋₄₁₇	36339,37	0,01893	0,52
HEP ₂₃₃₋₄₁₇	27585,15	0,015095	0,54

8.-CRIBADO DE FRAGMENTOS POR COMBINACIÓN DE QUÍMICA COMPUTACIONAL Y DIFERENTES TÉCNICAS BIOFÍSICAS

8.1.-Generación de un modelo farmacofórico para heparanasa

En esta parte del trabajo, el objetivo era desarrollar un modelo farmacofórico para heparanasa que se pudiera aplicar a una quimioteca de fragmentos. Al tratarse de fragmentos se escogió un modelo en el que las distancias entre los diferentes puntos farmacofóricos fueran relativamente cortas. Con este propósito, se desarrollaron diferentes hipótesis farmacofóricas utilizando una colección de los 27 inhibidores de heparanasa más pequeños que se conocen (190) que están representados en la Tabla 4 de Materiales y Métodos. Finalmente, se seleccionó el modelo que está representado en la Fig. 31 que resultó del alineamiento de los tres inhibidores más activos de los 27 mencionados anteriormente (INH-1, INH-2 y INH-3). Estos tres inhibidores están representados en la Fig. 32. Este modelo se consideró el más robusto y predictivo al dejar mejor clasificados a los 27 inhibidores conocidos de heparanasa cuando se aplicó a un conjunto de compuestos formados por estos 27 inhibidores y la quimioteca Prestwick (formada por 1120 compuestos). El modelo seleccionado estaba formado por dos posiciones hidrofóbicas/aromáticas, un donador de hidrógenos y un aceptor de hidrógenos como se puede observar en la Fig. 31. La función que se usó para evaluar la habilidad que tenían los diferentes modelos en proponer como mejores candidatos a los inhibidores conocidos de heparanasa fue el “root of the mean square deviation” o “la raíz cuadrada de la desviación del promedio al cuadrado” (rmsd) entre las distancias de los puntos farmacofóricos en el modelo y las distancias calculadas en cada ligando tal como se define en las herramientas MOE de búsqueda de farmacóforos.

Finalmente, el modelo farmacofórico seleccionado fue aplicado a una quimioteca comercial (CEREP, www.cerep.fr) compuesta por 686 fragmentos, encontrando que 114 compuestos (16% de la quimioteca) satisfacían los cuatro puntos farmacofóricos. La llamada “regla del 3” (peso molecular < 300 Da, cLoP < 3, núcleos donadores de puentes de H < 3, número de enlaces que pueden ser rotados < 3 y superficie total polar < 60Å²) (193) que tienen que cumplir los compuestos para ser fragmentos fue evaluada para toda la quimioteca, y se pudo comprobar que los compuestos que pasaban los cuatro

puntos farmacofóricos eran a su vez los más complejos, con un mayor peso molecular y mayor número de donadores y aceptores de hidrógeno como se puede comprobar en la Fig. 33. Esto era un resultado esperado, especialmente considerando que la distancia más larga entre dos puntos farmacofóricos del modelo estaba cerca de los 10 Å (Fig. 31), por lo que resultaba imposible que los fragmentos más pequeños cumplieran este requisito.

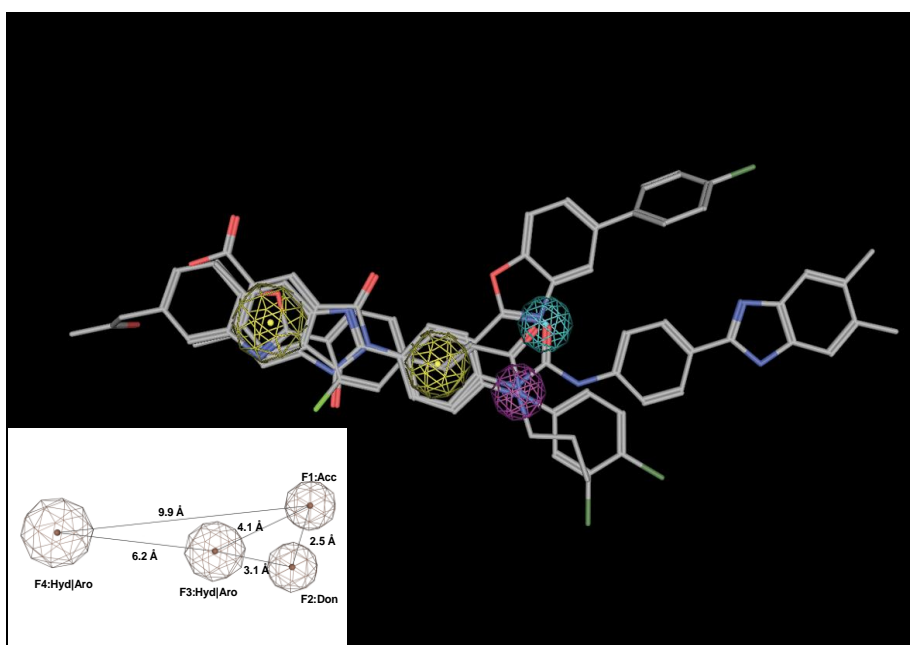
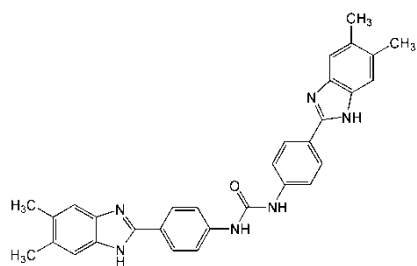
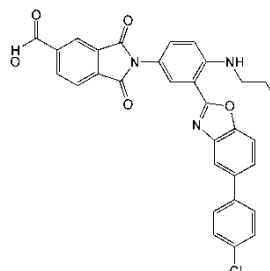


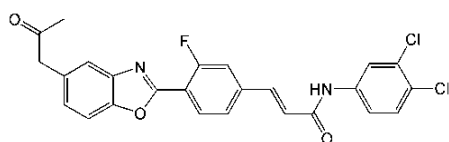
Figura 31. Modelo farmacofórico obtenido del alineamiento de los tres inhibidores conocidos de heparanasa más activos (INH-1, INH-2, INH-3). Este modelo incluye un donador de hidrógeno (esfera morada), un aceptor de hidrógenos (esfera azul) y dos regiones hidrofóbicas/aromáticas (esferas amarillas).



INH-1 IC₅₀ (0.075 μM)



INH-2 IC₅₀ (0.20 μM)



INH-3 IC₅₀ (0.20 μM)

Figura 32. Tres inhibidores más activos de heparanasa de los 27 inhibidores más pequeños conocidos de la proteína.

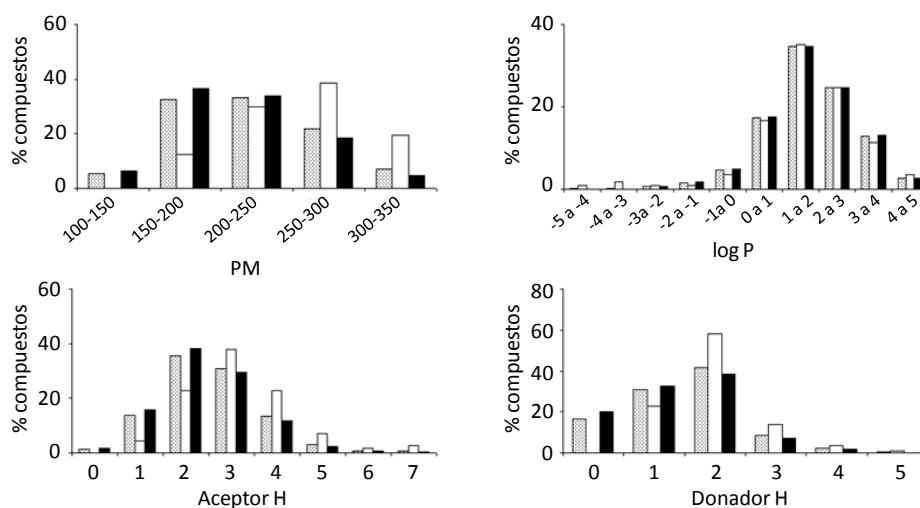


Figura 33. Porcentaje de compuestos en función del peso molecular (PM), rango de log P, número de aceptores de hidrógeno y número de donadores de hidrógeno. Gris, blanco y negro representan la quimioteca entera de fragmentos, compuestos que

satisfacen los cuatro puntos farmacofóricos y aquellos que no satisfacen el modelo farmacofórico respectivamente.

8.2.-Fiabilidad del modelo farmacofórico

Como se muestra en la Tabla 4, 23 de los 27 inhibidores escogidos de heparanasa satisfacían las hipótesis del modelo farmacofórico, estando entre los cuatro compuestos que no la cumplían los dos que mostraban menor actividad inhibitoria de heparanasa (190). Además se encontró una clara tendencia de actividad con los compuestos mejor clasificados por el modelo, siendo los mejores compuestos predichos aquellos inhibidores más activos.

Por otro lado se analizaron los resultados obtenidos cuando el modelo se aplicaba a estos 27 inhibidores de heparanasa junto con los 1120 compuestos de la quimioteca Prestwick para evaluar la fiabilidad del modelo. Se encontró que 13 de los 23 inhibidores de heparanasa que cumplían el modelo estaban entre el 10% de los compuestos mejor clasificados y que entre los ocho inhibidores que tenían actividad inhibitoria en el rango de submicromolar, cuatro de ellos estaban entre el 2% mejor clasificado (incluyendo tres compuestos que no se usaron para la generación del modelo farmacofórico) (Tabla 4). Por lo tanto, parecía que escogiendo los compuestos mejor clasificados en este cribado *in silico*, se encontraría un alto porcentaje de “hits”. Por otro lado, un resultado interesante fue el obtenido de la aplicación del modelo farmacofórico al sustrato natural de heparanasa, el heparán sulfato (Fig. 34). Para ello y debido a que el heparán sulfato es un glicosaminoglicano de cadena muy larga, se utilizó el mínimo fragmento del heparán sulfato capaz de ser reconocido por heparanasa, resultando que esta secuencia de heparán sulfato satisfacía todas las hipótesis farmacofóricas del modelo.

Tabla 4. Clasificación de los inhibidores conocidos de heparanasa

	Inhibidor de heparanasa	IC ₅₀ (μM)	Clasificación ^a
	INH-1	0,075	2**
	INH-2	0,20	6*
	INH-3	0,20	17
	INH-4	0,23	4**
	INH-5	0,27	1**
	INH-6	0,29	3**
	INH-7	0,40	11*
	INH-8	0,50	5*
	INH-9	0,13-19,00	22
	INH-10	1,00	18
	INH-11	1,00-10,00	7*
	INH-12	1,50-36,00	23
	INH-13	2,50	12*
	INH-14	3,00	8*
	INH-15	5,00-45,00	21
	INH-16	8,00	20
	INH-17	8,00-25,00	16
Una	INH-18	10,00	15
	INH-19	10,00	14
	INH-20	10,00-26,00	N.S
	INH-21	12,00	19
	INH-22	17,00	N.S
	INH-23	24,00-32,00	10*
	INH-24	25,00	13*
	INH-25	26,00-36,00	9*
	INH-26	>100,00	N.S
	INH-27	>100,00	N.S

descripción detallada de estos inhibidores se recoge en la revisión bibliográfica Hammond et al. (190). "N.S" no satisfacen el modelo. ^a Un conjunto de 27 inhibidores conocidos de heparanasa más 1120 estructuras de la quimioteca Prestwick se usaron

para validar el modelo. El número en la columna indica el orden en el que los inhibidores fueron clasificados con respecto a otros inhibidores cuando se usaron los 1147 compuestos. * Compuestos que se encuentran entre el 10% mejor clasificado por el modelo farmacofórico; ** aquellos que se encuentran entre el 2% mejor clasificado.

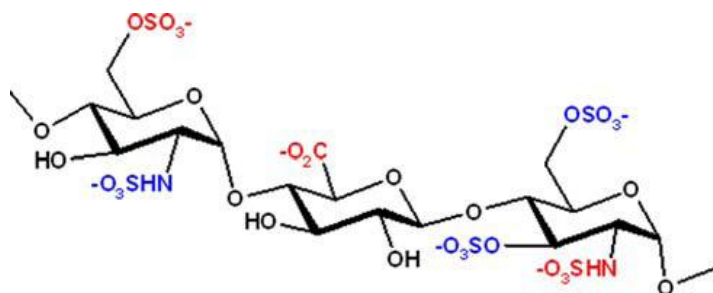


Figura 34. Mínimo fragmento de heparán sulfato que es reconocido por la enzima heparanasa. Los grupos cargados esenciales para que la enzima reconozca al sustrato están representados en rojo. En azul están representados otros grupos cargados que mejoran la afinidad con heparanasa.

8.3.-Análisis de la diversidad química de los compuestos que cumplen las hipótesis farmacofóricas

Un aspecto interesante de esta técnica es la posibilidad de identificar estructuras que, aun cumpliendo las hipótesis farmacofóricas, son diferentes químicamente (263) lo que ocurrió también en este caso. Como se muestra en la Fig. 35, existían compuestos (INH-1, INH-9 y INH-10) de distinta naturaleza química que satisfacían este modelo a pesar de que fue generado a partir de 3 compuestos con características químicas muy similares entre sí. La Fig. 36 muestra las estructuras de los tres fragmentos mejor clasificados FRAG-8, FRAG-28 y FRAG-3 (representados en Tabla 5 de Materiales y Métodos) así como su superposición con el modelo farmacofórico demostrando que la misma observación (diferentes estructuras químicas son compatibles con el mismo modelo farmacofórico) puede ser extrapolada a una quimioteca de fragmentos.

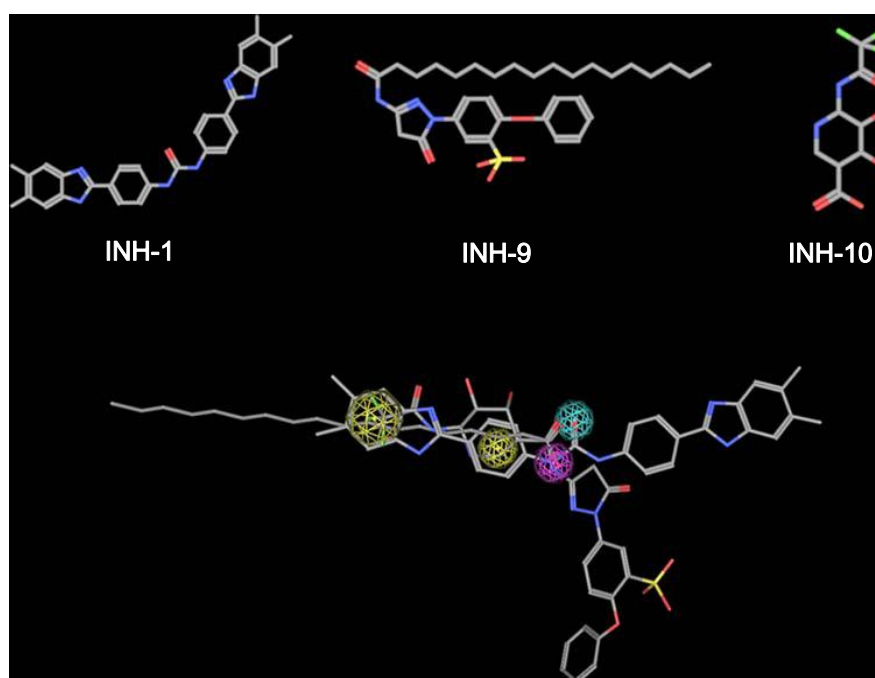


Figura 35. En la parte de arriba se muestran las estructuras de tres inhibidores de heparanasa, de diferente familia química, de izquierda a derecha INH-1, INH-9, INH-10, y en la parte de abajo se muestra la superposición de estos compuestos con el modelo farmacofórico.

Por otro lado, se llevó a cabo un análisis adicional de los inhibidores conocidos de heparanasa basado en su naturaleza química. Entre estos compuestos se incluyen algunos que a primera vista podrían tener un comportamiento similar frente a heparanasa teniendo en cuenta que todos contienen una larga cadena alifática. El modelo farmacofórico fue capaz de predecir que INH-9, INH-18 y INH-19 (Tabla 4 de Materiales y Métodos) eran efectivos inhibidores de heparanasa ya que estos compuestos satisfacían los cuatro puntos farmacofóricos y que INH-26 y INH-27 eran de hecho compuestos menos activos con una IC_{50} más alta de 100 μ M. Sin embargo, este análisis también mostró las limitaciones de esta técnica ya que el compuesto INH-22 no cumplía el modelo y sin embargo tenía una estructura química similar a INH-26

y INH-27, a pesar de tener una IC_{50} de un orden de magnitud menor que estos dos inhibidores. Parece evidente que el modelo no era capaz de apreciar leves diferencias químicas entre los compuestos.

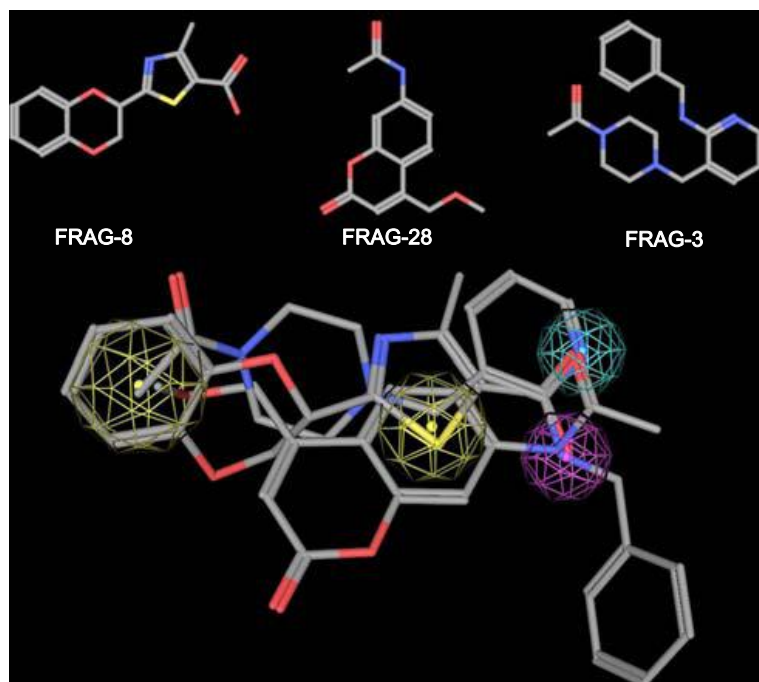


Figura 36. Estructuras de los fragmentos mejor clasificados por el modelo, en la parte superior de la figura de izquierda a derecha FRAG-8, FRAG-28 y FRAG-3. En la parte inferior está representada la superposición de estos tres fragmentos con el modelo farmacofórico

8.4.- Identificación de “hits” por Resonancia Magnética Nuclear y validación del modelo farmacofórico por RMN

Con el objetivo de validar el modelo, se llevó a cabo un cribado de fragmentos mediante Resonancia Magnética Nuclear. Para ello, se eligió una colección de 100 fragmentos (Tabla 5 de Materiales y Métodos) que contenía 50 compuestos de los 114 fragmentos que cumplían los cuatro puntos del modelo farmacofórico (44% de este grupo), 25 compuestos que cumplían tres puntos y 25 que cumplían dos. Los compuestos pertenecientes a los dos últimos grupos fueron

seleccionados por diversidad química y el grupo de los 50 fragmentos contenía los 18 compuestos mejor clasificados por el modelo y los 32 restantes fueron escogidos igualmente por diversidad química. Esto se hizo utilizando la herramienta MACCS de MOE y el algoritmo de Tanimoto (264) como se describe en Materiales y Métodos.

Inicialmente se realizó un cribado de estos 100 fragmentos utilizando el experimento WaterLOGSY por RMN (204, 210, 265) que identifica las interacciones proteína-ligando detectando la magnetización que se transfiere a través de las moléculas de agua desde la proteína al ligando. Este experimento permitía la identificación de interacciones en el rango de mM- μ M, con la ventaja de que no requería gran concentración de la muestra y no era necesaria la utilización de proteína marcada. El resultado de estos estudios mostró que el mayor porcentaje de “hits” se identificaron entre los 50 fragmentos que cumplían los cuatro puntos farmacofóricos. De esta forma, se comprobó que el 66% de los compuestos que cumplían todas las hipótesis farmacofóricas daban resultado positivo en el WaterLOGSY, representando un 6 y un 30% de enriquecimiento frente a los compuestos que cumplían 3 o 2 puntos del modelo farmacofórico respectivamente.

Estos resultados mostraban el potencial de la combinación de estas dos técnicas para la generación de “hits” en cribado de fragmentos ya que se realiza un filtro previo con herramientas de química computacional reduciendo el número inicial de compuestos para hacer el cribado por técnicas experimentales con la RMN. Esto es una ventaja adicional para aquellos investigadores que trabajan en academia donde la posibilidad de manejar grandes colecciones de fragmentos está más limitada.

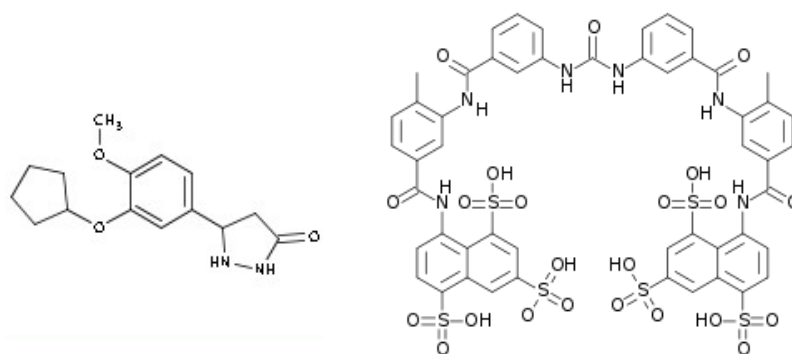
Aunque el experimento WaterLOGSY es extremadamente útil para hacer cribado de fragmentos y para encontrar compuestos que interaccionan con una proteína diana, tiene el inconveniente de que no proporciona información directa sobre el sitio de la proteína que une al ligando. Además la intensidad de la señal en el experimento WaterLOGSY depende enormemente de la orientación de la

moléculas de agua (266), por lo que es muy difícil obtener información sobre la afinidad de la interacción en este tipo de experimento.

Con el objetivo de encontrar nuevos inhibidores de heparanasa se realizó el cribado para este mismo conjunto de 100 fragmentos utilizando experimentos de interacción STD y experimentos de competición también utilizando el STD por Resonancia Magnética Nuclear. El propósito de este tipo de experimentos fue en primer lugar, reducir el número de falsos positivos que puede producir el WaterLOGSY y el STD, debido a los cambios de pH, agregación o a interacciones no específicas (267), al medir la interacción por dos métodos diferentes (268). En segundo lugar, este tipo de experimentos permitían obtener información sobre si los compuestos se unían al sitio catalítico de heparanasa utilizando experimentos de competición (200) con ligandos conocidos de la enzima. El ligando que se utilizó para hacer los experimentos de competición fue la suramina, K_D medida mediante SPR: 0,5 μM (269). Como era de esperar, cuando se hizo el cribado de fragmentos utilizando dos tipos de experimentos diferentes (WaterLOGSY y STD), hubo un descenso en la cantidad de “hits” encontrados. Sin embargo, cuando se comparaban fragmentos de tamaño similar (peso molecular y distancia topológica (270)), se encontró que todavía había una fuerte correlación entre el cumplimiento del modelo farmacofórico y el número de “hits” encontrados, de forma que el 30% de los fragmentos que satisfacían las 4 hipótesis farmacofóricas del modelo daban resultado positivo en los experimentos STD de competición por RMN. Estos resultados representaban un 9% y 20% de enriquecimiento cuando se compara con los compuestos que eran positivos experimentalmente y que solo cumplían tres y dos puntos del modelo, respectivamente.

De los 100 fragmentos que se probaron por RMN, se encontraron 37 compuestos que mostraban interacción con heparanasa y además competían con la suramina por el sitio catalítico de la enzima. Estos fragmentos son FRAG-1, 3, 5, 7, 8, 10, 11, 13, 14, 16, 19, 22, 24, 32, 33, 35, 37, 41, 44, 47, 52, 55, 57, 58, 61, 66, 67, 70, 73, 77, 79, 82, 83, 86, 92, 95 y 97 cuyas estructuras están representadas en la Tabla 5 de Materiales y Métodos. En la Fig. 37 se muestran

los experimentos monodimensionales de ^1H (negro), WaterLOGSY (rojo) y STD normal y de competición (azul) para uno de estos fragmentos, FRAG-47.



FRAG-47

SURAMINA

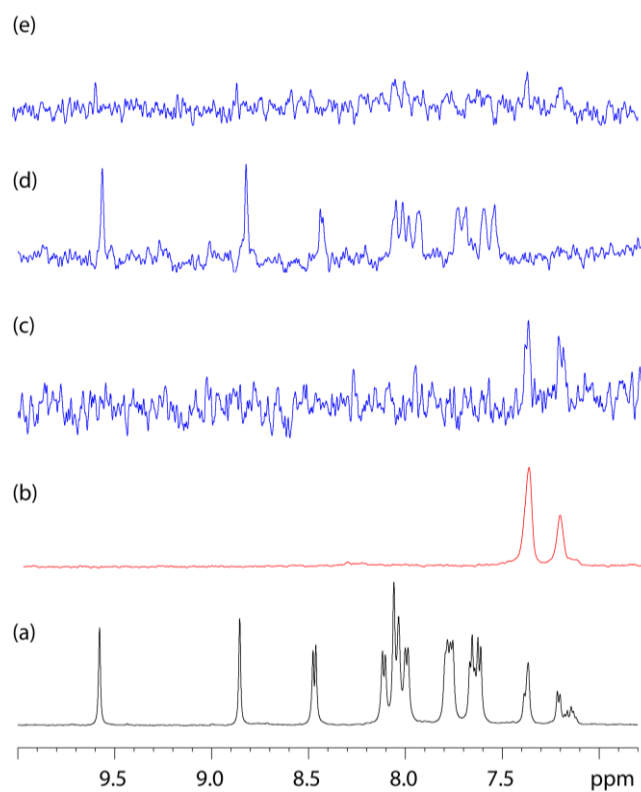


Figura 37. Ejemplo de experimentos de interacción proteína-ligando por RMN llevados a cabo para cada fragmento, mostrándose con detalle la región aromática del espectro. (a) Espectro de ^1H de una mezcla de suramina y FRAG-47 (cada uno a una concentración de $500\ \mu\text{M}$); (b) Espectro water-LOGSY del FRAG-47 ($5\ \mu\text{M}$ de HEP₁₅₈).

417, 300 μM del fragmento) mostrando interacción positiva; (c) experimento STD del FRAG-47 (5 μM de HEP₁₅₈₋₄₁₇, 500 μM del fragmento), mostrando su interacción con HEP₁₅₈₋₄₁₇, (d) experimento STD de la suramina (5 μM de HEP₁₅₈₋₄₁₇, 500 μM de suramina), mostrando su interacción con HEP₁₅₈₋₄₁₇ (e) experimento de competición por STD (5 μM de HEP₁₅₈₋₄₁₇, 500 μM del fragmento y 500 μM de suramina) mostrando la competición del FRAG-47 y la suramina por el sitio catalítico de la enzima. En la parte superior de la figura están representadas la estructura del FRAG-47 y la suramina.

Se puede comprobar que el experimento WaterLOGSY (Fig. 37b) muestra interacción positiva. Un compuesto que no interaccionara con la proteína daría como resultado las mismas señales pero invertidas, mientras que compuestos que interaccionan con la proteína dan señales positivas o resonancias menos negativas que el WaterLOGSY realizado en ausencia de la proteína. La Fig. 37c muestra el experimento STD para el FRAG-47 mostrando su interacción con HEP₁₅₈₋₄₁₇. El experimento de competición por STD se muestra en la Fig. 37e; el decrecimiento de la intensidad de las señales de resonancia del fragmento y la suramina comparada con la intensidad de las señales del STD del fragmento (Fig. 37c) y de la suramina (Fig. 37d) indica que FRAG-47 y suramina están compitiendo por el sitio catalítico de heparanasa.

8.5.-Identificación de “hits” por Resonancia de Plasma Superficial

Finalmente se midió la constante de disociación para este grupo de 37 compuestos que resultaron ser positivos por RMN, pudiéndose comprobar que el método de SPR anteriormente optimizado para ligandos conocidos de heparanasa también era válido para medir interacciones con moléculas de bajo peso molecular obteniéndose respuestas de más de 100 RU. En la Fig. 38 se puede apreciar el sensograma para el compuesto FRAG-86 cuya estructura viene representada en esta figura. Para este fragmento se obtiene una respuesta de aproximadamente 120 RU para una concentración de 500 μM de compuesto y 30 RU para una concentración de 100 μM que es la concentración que se suele

usar para hacer cribado de fragmentos por SPR. Esta respuesta es considerada como suficiente para poder detectar compuestos de bajo peso molecular (< 350 Da) como es este caso donde el compuesto tiene un peso molecular de 263,71 Da. Además también se demostró que la enzima inmovilizada en el chip era estable durante 48 horas que es otro de los requisitos indispensables para poder hacer el cribado de una librería de compuestos. Esto condujo a la conclusión de que se ha optimizado el método para poder hacer cribado de fragmentos por SPR con la enzima heparanasa.

Los resultados que se obtuvieron para este grupo de 37 compuestos detectados como positivos por RMN, mostraban K_D en el rango de 0,2-5 mM y eficiencia de unión equivalentes a las encontradas en la bibliografía para fragmentos del mismo tamaño (271).

Por lo tanto usando este conjunto de técnicas *in silico* y biofísicas se han identificado varios fragmentos que podrían ser potencialmente un buen punto de partida para la optimización de estos “hits” hasta ser “leads” de esta enzima. Los mejores resultados se obtuvieron para el FRAG-86 cuya K_D fue 236 μM .

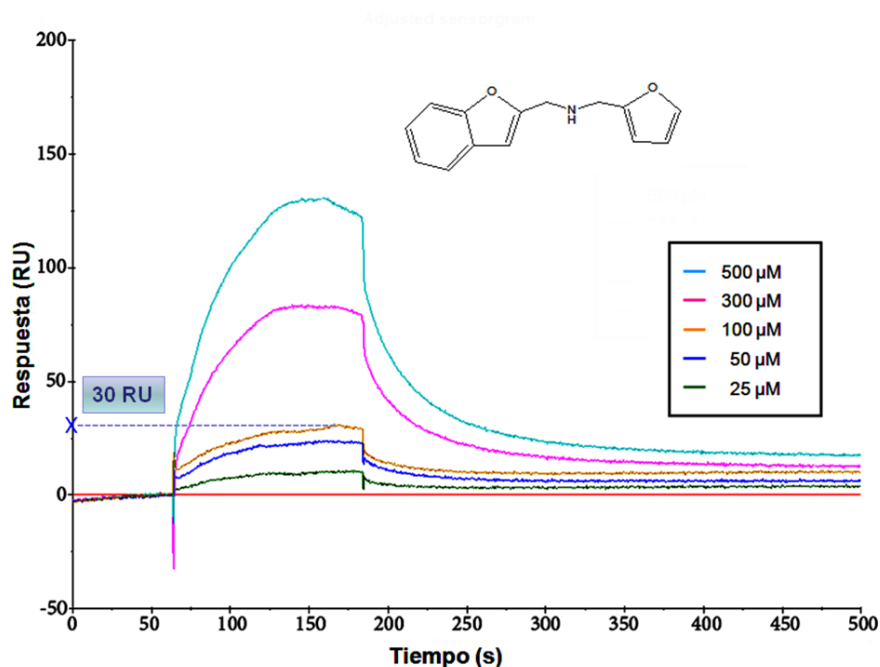


Figura 38. Sensograma para el fragmento FRAG-86. Las concentraciones de este compuesto que se usaron indicadas desde abajo hacia arriba fueron 25, 50, 100, 300, 500 μ M. Este experimento fue realizado a 25 °C en un instrumento Biacore T100.

9.-CRIBADO DE COMPUESTOS CON CARACTERÍSTICAS DE FÁRMACOS POR COMBINACIÓN DE QUÍMICA COMPUTACIONAL Y DIFERENTES TÉCNICAS BIOFÍSICAS

En esta parte del trabajo se utilizaron compuestos de mayor peso molecular (Peso molecular > 350 Da) y en su mayoría con características de fármacos con el objetivo de encontrar compuestos de mayor afinidad por la proteína. Estos compuestos pertenecían a la quimioteca Prestwick (1120 compuestos).

Se hizo un primer cribado utilizando técnicas *in silico*. Para ello se aplicó el modelo farmacofórico ya descrito en el apartado 8 de Discusión y Resultados a esta quimioteca Prestwick y posteriormente los compuestos se clasificaron en función del número de puntos farmacofóricos que cumplían.

Posteriormente se quería hacer una nueva clasificación de estos compuestos utilizando el “docking” a partir del modelo estructural de heparanasa (descrito en el apartado 1 de Discusión y Resultados). Para ello, esta técnica se validó utilizando inhibidores conocidos de heparanasa para comprobar que el modelo estructural de heparanasa se podía utilizar para este propósito. En la Fig. 39 se puede observar el docking realizado con el inhibidor de heparanasa utilizado para llevar a cabo el modelo farmacofórico INH-1 (Tabla 4 de Materiales y Métodos) y el modelo de homología por heparanasa. En azul están representados los dos sitios de unión a heparán sulfato (Lys¹⁵⁸-Asp¹⁷¹ y Gln²⁷⁰-Lys²⁸⁰) y en rojo los dos ácidos glutámicos del sitio catalítico de la enzima (Glu²²⁵, Glu³⁴³). Se puede observar cómo el inhibidor de heparanasa bloquea el sitio catalítico de la enzima e interacciona con parte de los aminoácidos que pertenecen al sitio de unión a heparán sulfato. Por lo tanto, se concluyó que el modelo estructural de heparanasa se puede utilizar para llevar a cabo el docking de la enzima con nuevos ligandos.

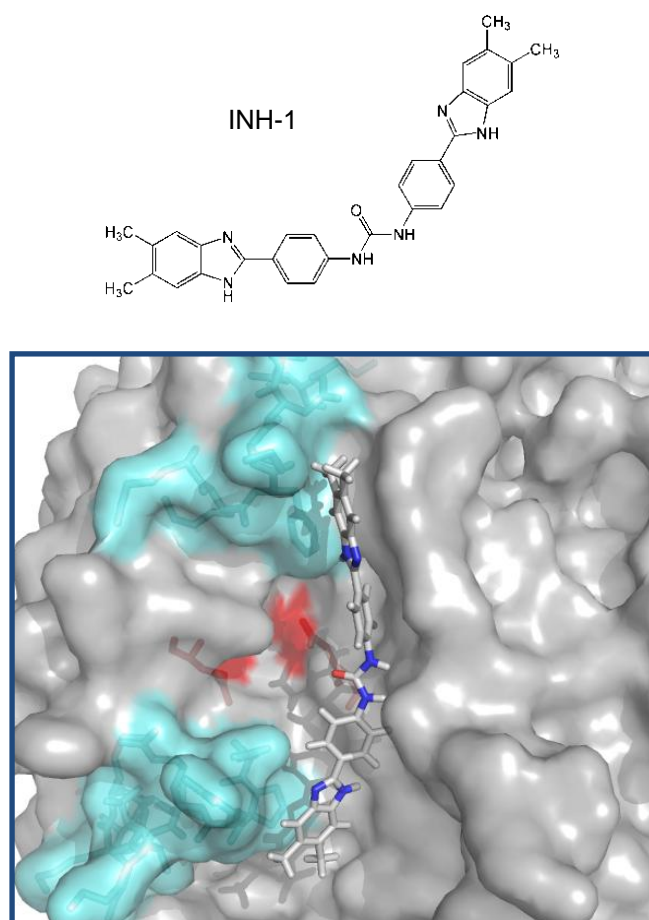


Figura 39. Docking entre el modelo estructural de heparanasa por homología y un inhibidor conocido de heparanasa INH-1. En azul están representados los dos sitios de unión a heparán sulfato (Lys¹⁵⁸-Asp¹⁷¹ y Gln²⁷⁰-Lys²⁸⁰) y en rojo los dos ácidos glutámicos del sitio catalítico de la enzima (Glu²²⁵, Glu³⁴³). El docking se llevó a cabo con el conjunto de herramientas de modelización de OpenEye (v. 2.1) (<http://www.eyesopen.com>).

A continuación, se realizó el docking con el modelo estructural de heparanasa y los 1120 compuestos de la quimioteca Prestwick y estos compuestos fueron clasificados en función de la energía de interacción calculada.

La combinación de estos dos métodos de química computacional, uno basado en la estructura del receptor (“docking”) y el otro basado en la estructura del ligando (modelo farmacofórico) nos permitió realizar una nueva clasificación de los compuestos de la quimioteca Prestwick teniendo en cuenta los mejores resultados de las dos técnicas (Tabla 5).

Tabla 5. Compuestos Prestwick mejor clasificados in silico. Resultados experimentales por RMN y SPR.

Compuesto	Clasificación <i>in silico</i>	Clasificación farmacofórico	Clasificación docking	WaterLOGSY	STD	K _D (SPR)
Astemizol	1	6	4	+	+	> 1mM
Amodiaquina	2	1	33	+	+	52,6 μM
Hesperidina	3	36	11	+	+	> 1mM
Cisapride	4	25	30	+	+	> 1mM
Naringina	5	34	29	+	+	> 1mM
Metergoline	6	72	6	+	+	> 1mM
Labetol	7	4	76	+	+	1 mM

Posteriormente se escogieron los 7 compuestos mejor clasificados para evaluarlos experimentalmente por RMN y SPR. En la Tabla 5 aparecen los resultados de estos experimentos y como se puede observar todos eran positivos en WaterLOGSY y STD pero la mayoría tenían constantes de disociación muy débiles (>1 mM) excepto la amodiaquina y labetol con K_D de 52,6 μM y 1 mM, respectivamente. A continuación se muestran los experimentos de RMN (Fig. 40) y SPR (Fig. 41) para la amodiaquina.

Se intentó conseguir el epítipo de la amodiaquina que interaccionaba con la proteína utilizando el experimento STD. Para ello, en primer lugar se asignaron todas las señales del experimento 1D de ¹H de la amodiaquina. A continuación se integraron todas las señales en el experimento ¹H y en el de interacción STD.

Se midió la relación integral de las señales STD/integrales de las señales ^1H para identificar el epítipo de la amodiaquina que interaccionaba con $\text{HEP}_{158-417}$. Sin embargo no se obtuvieron resultados coherentes debido al tamaño de la molécula de amodiaquina o a la posibilidad de que tenga distintas orientaciones al unirse a la proteína y por tanto tenga distintos epítipos de afinidades diferentes.

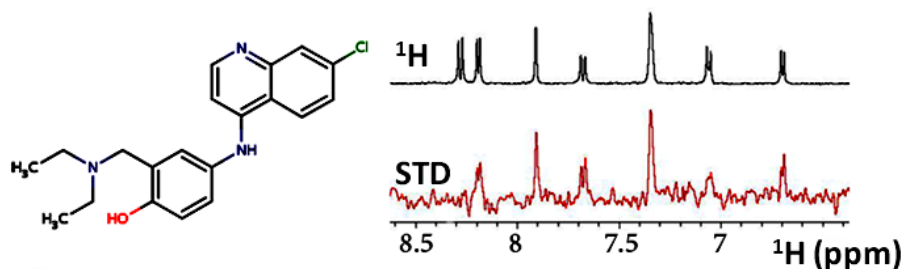


Figura 40. Interacción de la amodiaquina con $\text{HEP}_{158-417}$. Estructura de amodiaquina y experimento STD; en negro espectro de ^1H de la amodiaquina a una concentración de $500 \mu\text{M}$ y en rojo experimento STD de la amodiaquina ($5 \mu\text{M}$ de $\text{HEP}_{158-417}$, $500 \mu\text{M}$ de la amodiaquina), mostrando su interacción con $\text{HEP}_{158-417}$. Estos experimentos fueron realizados con un espectrómetro Bruker 600 MHz con criosonda TCI.

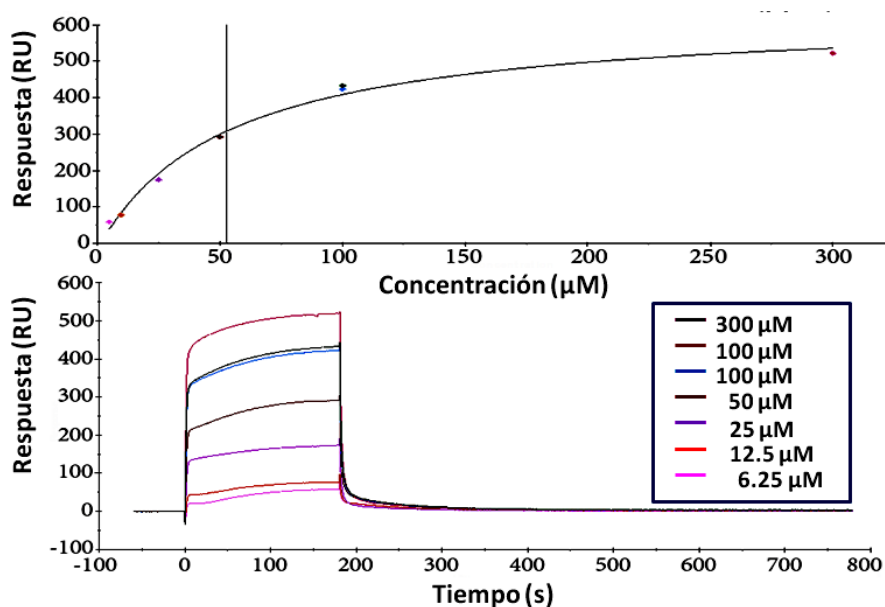


Figura 41. Interacción de amodiaquina con $\text{HEP}_{158-417}$. Medida de afinidad de la interacción $\text{HEP}_{158-417}$:amodiaquina mediante SPR. Se muestra en la parte de abajo el

sensograma utilizado para las constantes cinéticas y en la parte superior el ajuste para el cálculo de la constante de disociación, la barra vertical indica el valor de la K_D . Las concentraciones que se usaron indicadas desde abajo hacia arriba fueron 6,25, 12,5, 25, 50, 100 y 300 μM . La concentración de 100 μM se hizo por duplicado. Estos experimentos fueron realizados con un instrumento Biacore T100.

10.-OPTIMIZACIÓN DE “HITS”

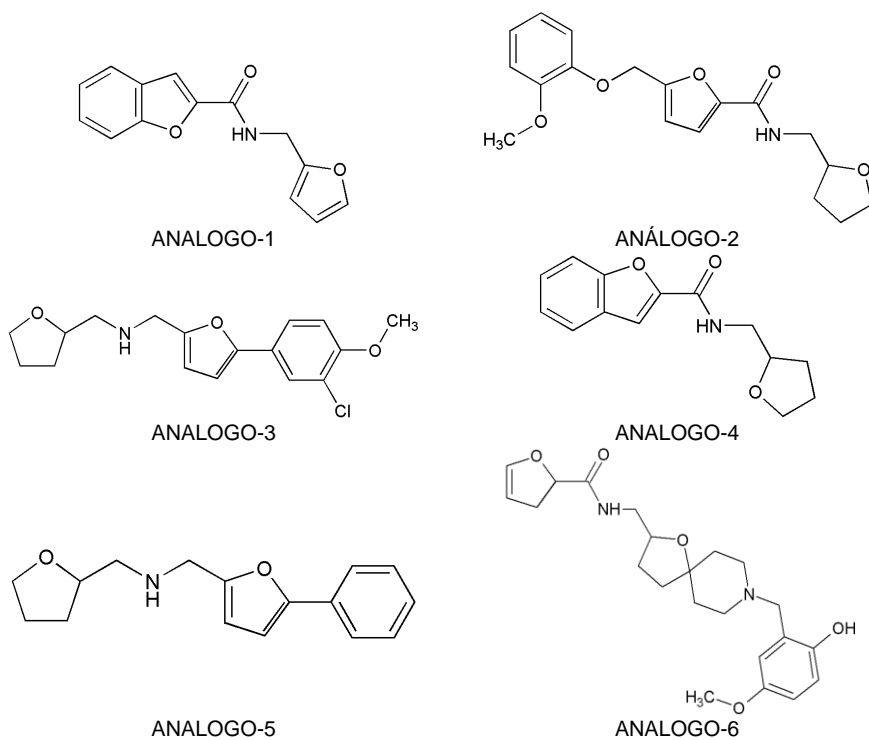
10.1.-Optimización de “hits” que provienen del cribado de fragmentos

El objetivo de esta parte del trabajo fue encontrar compuestos que tuvieran mayor afinidad por heparanasa a partir de uno de los hits que se obtuvieron del cribado de fragmentos por RMN. El mejor de los fragmentos que se obtuvo a partir de este cribado (apartado 8 de Discusión y Resultados) fue el FRAG-86, Este compuesto fue positivo en el experimento WaterLOGSY y el STD y utilizando el STD de competición con suramina se demostró que este fragmento competía con la suramina por el sitio catalítico de la enzima. Además la K_D que se midió por SPR entre este fragmento y la construcción HEP₁₅₈₋₄₁₇ fue 236 μM como se mencionó en el apartado 8.5 (Fig. 38) demostrándose así que la interacción entre esta construcción de heparanasa y el FRAG-86 era mayor que cualquier otro fragmento de los 100 que se eligieron inicialmente. Por otro lado, era un compuesto fácil de modificar desde el punto de vista de la síntesis química y por lo tanto fácil de optimizar.

Con el objetivo de mejorar la constante de disociación de este fragmento con heparanasa se llevaron a cabo dos estrategias. La primera de ellas fue la búsqueda de análogos comerciales en los que se añadían o eliminaban distintos grupos en diferentes partes de la molécula y la segunda fue unir dos fragmentos mediante síntesis química que tuvieran afinidad con la proteína o lo que se conoce con el nombre de “linking”. En este caso se sintetizó un compuesto simétrico formado por dos moléculas del FRAG-86 ya que como se demostrará posteriormente podría tener mayor afinidad por heparanasa.

10.1.1.-Búsqueda de análogos comerciales

En la Fig. 42 están representadas las estructuras de estos análogos comerciales que tienen como base la estructura del FRAG-86 al que se le han añadido, sustituido o eliminado diferentes grupos químicos de mayor o menor tamaño y complejidad con el objetivo al menos de determinar qué parte del fragmento es importante en la interacción. La búsqueda de estos análogos comerciales (casa comercial ChemBridge) se realizó a través de la página web <http://www.emolecules.com>, usando la estructura de FRAG-86 como referencia.



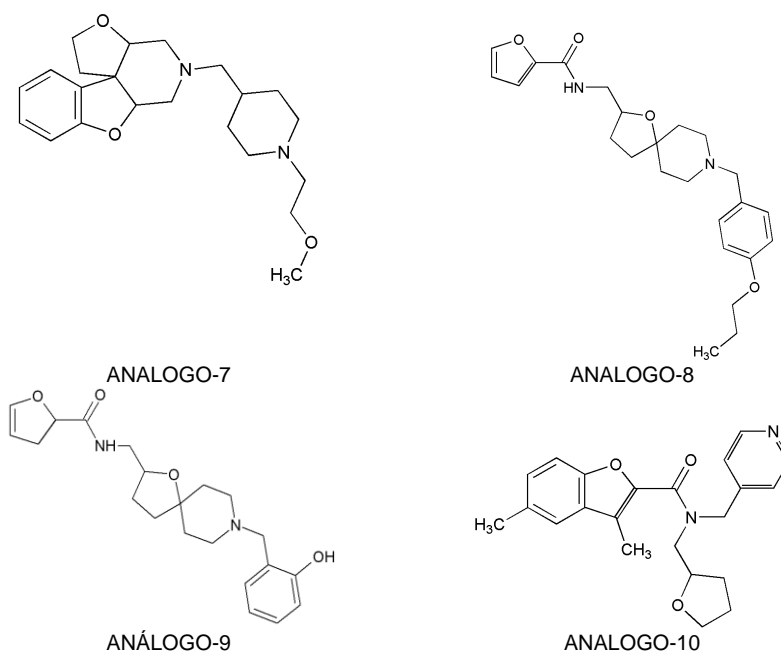


Figura. 42. Estructura de los análogos comerciales del FRAG-86.

Estos análogos fueron evaluados experimentalmente mediante RMN y SPR y los resultados se recogen en la Tabla 6.

Tabla 6. Resultados de interacción de los análogos comerciales del FRAG-86.

Compuesto	STD	STD competición	K _D (SPR)
ANÁLOGO-1	+	compite	> 1mM
ANÁLOGO-2	+	no compite	> 1mM
ANÁLOGO-3	+	compite	470 μM
ANÁLOGO-4	+	compite	176 μM
ANÁLOGO-5	+	compite	79 μM
ANÁLOGO-6	+	compite	> 1mM
ANÁLOGO-7	+	compite	> 1mM
ANÁLOGO-8	+	no compite	> 1mM
ANÁLOGO-9	+	compite	> 1mM
ANÁLOGO-10	+	compite	> 1mM

En la Fig. 43 se comparan algunas de estas estructuras para determinar qué cambios son aquellos que afectan a la interacción del compuesto con heparanasa. Si comparamos el FRAG-86 con el análogo 1 podemos observar que la sustitución de un grupo amino por un grupo amida disminuye la potencia de la interacción al menos 4 veces. Al comparar el análogo 4 con el análogo 1 donde además de la introducción del grupo amida se ha sustituido un anillo furano por un oxolano, la potencia aumenta al menos cinco veces, por lo que se observa que esta sustitución del anillo de furano por un oxolano mejora la interacción a pesar de tener la presencia negativa del grupo amida. Por lo tanto una de las propuestas a sintetizar sería una molécula similar a este mismo FRAG-86 donde se sustituya el anillo de furano por un oxolano. Si se compara el análogo 4 y 5 se podría esperar que mejoraría como mínimo cuatro veces la potencia de interacción del análogo 5 ya que le falta el grupo amida, sin embargo como la K_D del análogo 5 con respecto al análogo 4 es dos veces menor se podría deducir que la adición del enlace entre los anillos de benceno y furano no favorece la interacción o la disminuye dos veces.

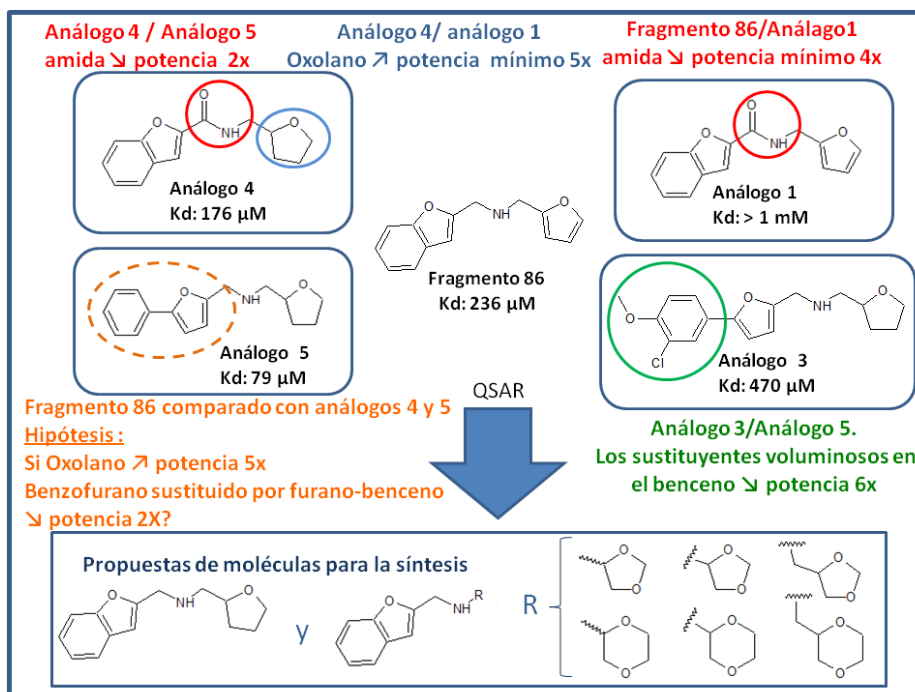


Figura 43. Comparación de estructuras de los análogos del FRAG-86.

Si seguimos comparando el análogo 3 y 5 podemos observar que el anillo de benceno sustituido con grupos voluminosos como el Cl y un grupo metoxi, disminuyen la interacción del orden de seis veces. Por lo tanto podemos deducir que la adición de grupos voluminosos en el benzofurano del fragmento original disminuiría la interacción. En consecuencia se propone en primer lugar la sustitución del anillo furano del FRAG-86 por un anillo oxolano y en segundo lugar el aumento del número de oxígenos en anillo de cinco y seis enlaces para favorecer los puentes de hidrógeno con la proteína manteniendo el benzofurano en la estructura del compuesto.

Además se calculó la eficiencia de unión (272) para el compuesto original y los análogos con mejores constantes de disociación. La fórmula utilizada fue

$$LE = \Delta G/HAC$$

donde LE es la eficiencia de unión del ligando a la proteína en $\text{Kcal}\cdot\text{mol}^{-1}$ de cada átomo pesado, HAC es el número de átomos pesados y $\Delta G = -RT \ln K_D$, T es la temperatura en K, R es la constante de gases perfectos cuyo valor es $2,38 \times 10^{-4}$ ($\text{kcal}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$) y K_D es la constante de disociación en M.

La eficiencia de unión para los mejores análogos del FRAG-86 se puede observar en la Tabla 7.

Tabla 7. Eficiencia de unión para FRAG-86 y sus mejores análogos.

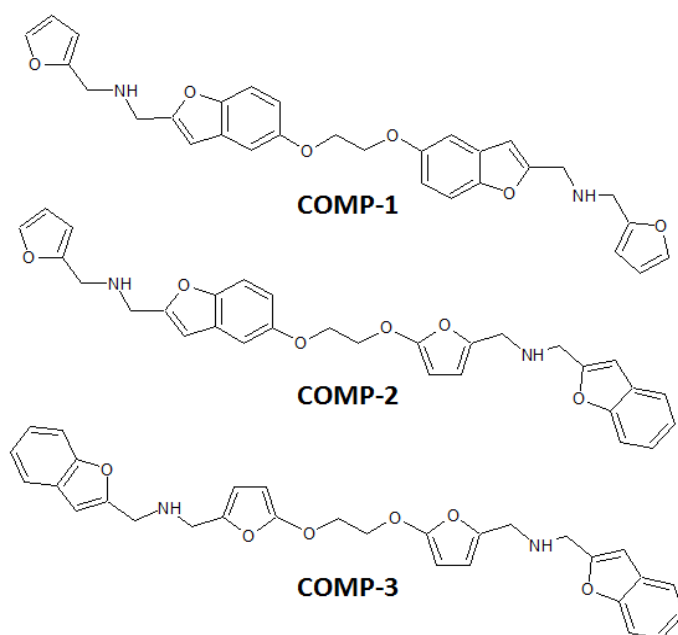
Compuesto	K_D	LE
FRAG-86	236 μM	0,26
ANÁLOGO-3	470 μM	0,20
ANÁLOGO-4	176 μM	0,28
ANÁLOGO-5	79 μM	0,29

En el proceso de optimización de “hits” a “leads”, se recomienda que por cada átomo que se añada al compuesto, la eficiencia de unión no debe bajar de 0,3 aproximadamente (272). Por tanto, el análogo 5 del FRAG-86 sería un buen punto de partida para obtener un compuesto más potente por dos razones; en primer lugar porque es un compuesto sencillo, fácil de modificar sintéticamente

y en segundo lugar porque ya parte con una eficiencia de unión de 0,29, y por tanto es más fácil mantenerla que si se parte de una eficiencia más baja.

10.1.2.-Búsqueda de análogos mediante síntesis orgánica. Síntesis de una molécula simétrica.

Paralelamente al análisis anterior y puesto que se sabe que una sola molécula de FRAG-86 es demasiado pequeña para unirse a los dos sitios de unión a heparán sulfato que distan entre ellos según el modelo por homología un mínimo de 16 Å, se propuso la síntesis de compuestos que contenían este fragmento por duplicado con las siguientes combinaciones: cabeza-cabeza (COMP-1), cabeza-cola (COMP-2) y cola-cola (COMP-3) unidos mediante “linkers” de distintos tamaños. En la Fig. 44 están representadas las tres combinaciones posibles con uno de los “linkers” propuestos. Se ha considerado que la cabeza del compuesto es el anillo benzofurano y la cola del compuesto, el anillo de furano



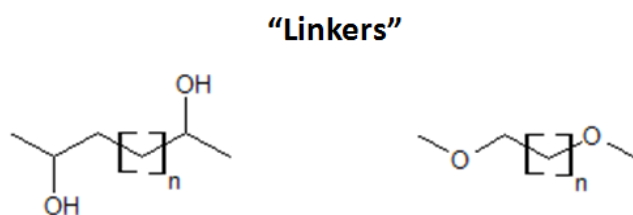


Figura 44. Estructura de los compuestos sintéticos propuestos, análogos del FRAG-86.

Estos compuestos están siendo sintetizados en el laboratorio de Síntesis Orgánica dirigido por el Dr. Santos Fustero del Centro de Investigación Príncipe Felipe.

Hasta la fecha se ha sintetizado solo uno de ellos (COMP-1) y los resultados experimentales mostraban que este compuesto daba positivo por STD y competía con la suramina por la interacción con el sitio catalítico de heparanasa en el STD de competición (Fig. 45). La K_D medida mediante SPR fue de 92,7 μM (Fig. 46).

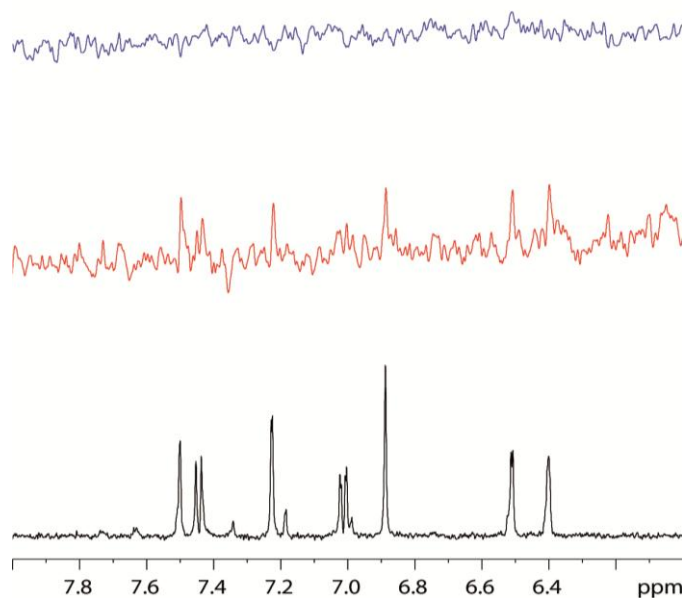


Figura 45. Interacción de COMP-1 con $\text{HEP}_{158-417}$. En negro está representado el espectro de ^1H del compuesto COMP-1 (a una concentración de 500 μM). En rojo está

representado el experimento STD del COMP-1 (5 μM de HEP₁₅₈₋₄₁₇, 500 μM del fragmento), mostrando su interacción con HEP₁₅₈₋₄₁₇ y en azul se puede observar el experimento de competición por STD (5 μM de HEP₁₅₈₋₄₁₇, 500 μM de COMP-1 y 500 μM de suramina) mostrándose la competición de COMP-1 y la suramina por el sitio catalítico de la enzima. Estos experimentos fueron realizados con un espectrómetro Bruker 600 MHz con criosonda TCI.

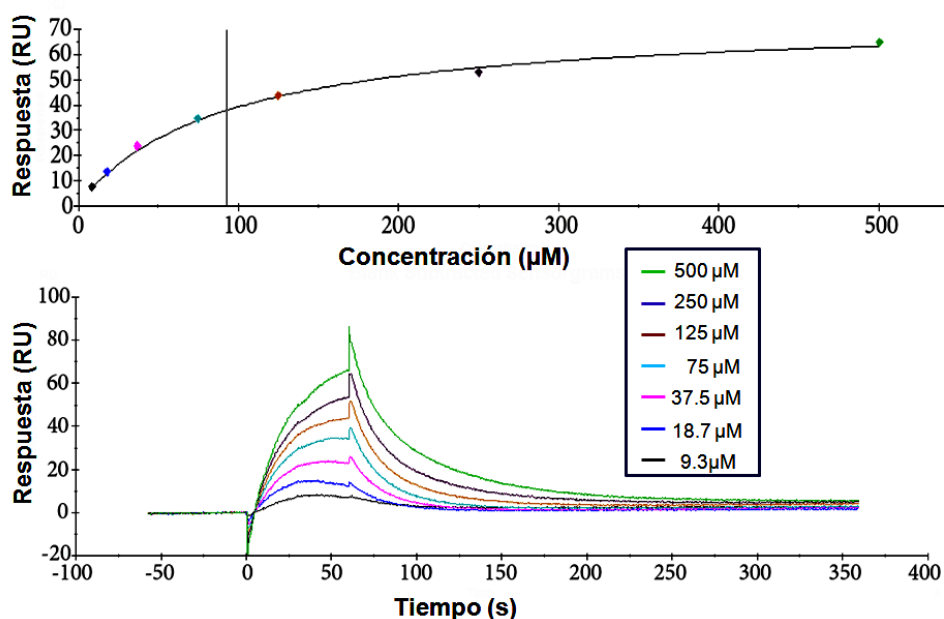
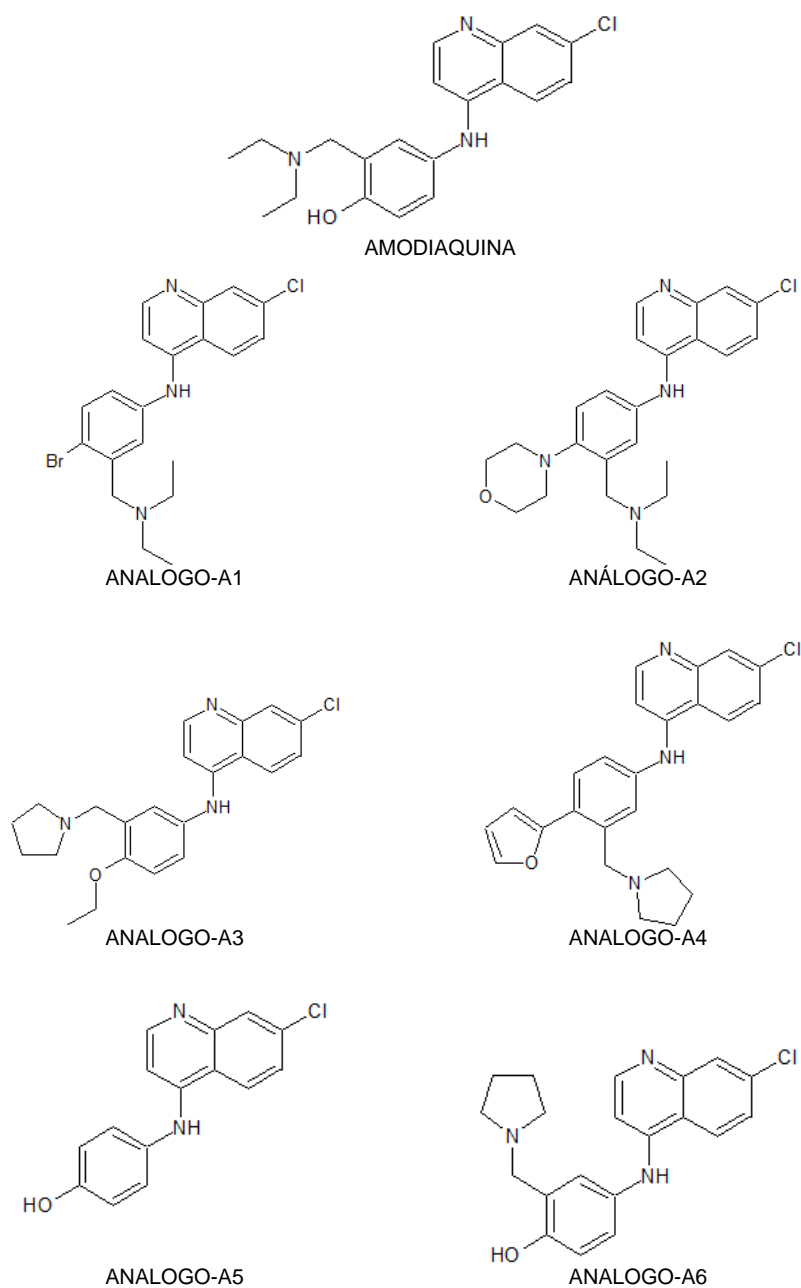


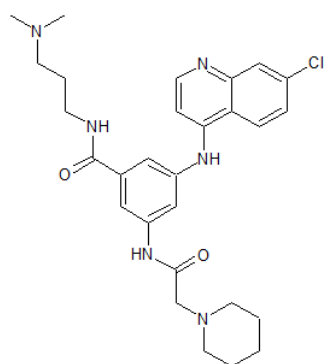
Figura 46. Interacción de COMP-1 con HEP₁₅₈₋₄₁₇ por SPR. Se muestra en la parte de abajo el sensograma utilizado para las constantes cinéticas y en la parte superior el ajuste para el cálculo de la constante de disociación, la barra vertical indica el valor de la K_D . Las concentraciones que se usaron, indicadas desde abajo hacia arriba, fueron 9,3, 18,7, 37,5, 75, 125, 250 y 500 μM . Estos experimentos fueron realizados con un instrumento Biacore T100.

10.2.-Optimización de “hits” que provienen del cribado de compuestos con características de fármacos.

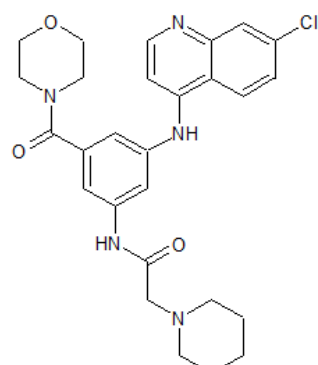
Se utilizó la amodiaquina como compuesto de partida para buscar análogos con mayor afinidad por la proteína y para saber qué parte de la molécula estaba

implicada en esta interacción con heparanasa. Para ello, se utilizaron análogos sintetizados en el laboratorio de la Dra. Patricia Melnyk del Institute Pasteur (CNRS) de Lille (Francia). Estos análogos están representados en la Fig.47.

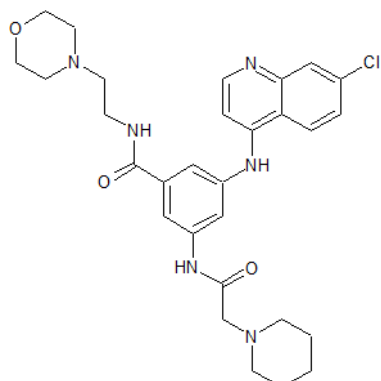




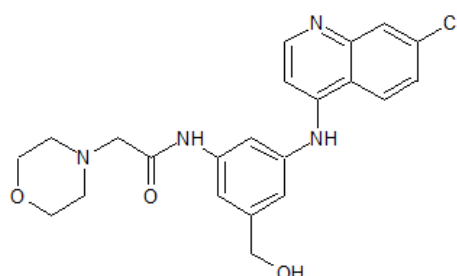
ANALOGO-A7



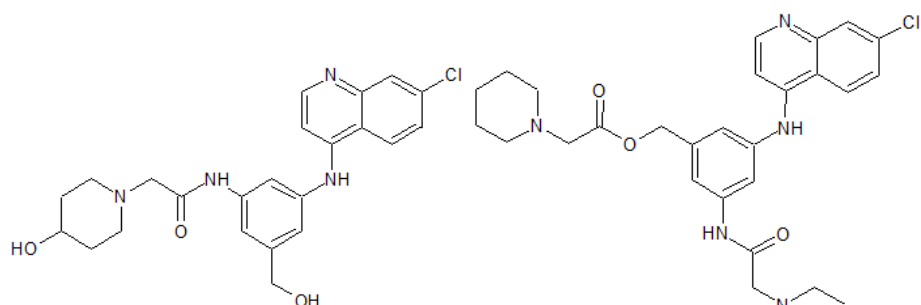
ANALOGO-A8



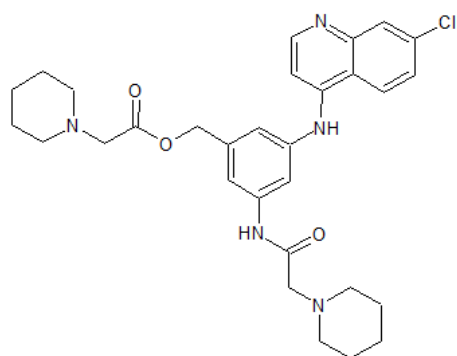
ANALOGO-A9



ANALOGO-A10



ANALOGO-11



ANALOGO-12

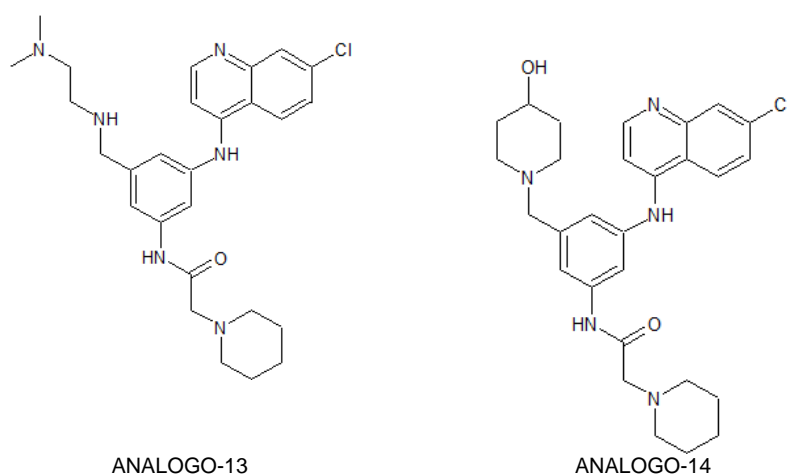


Figura 47. Compuestos análogos de la amodiaquina.

Se realizaron experimentos de interacción por RMN (STD) y se midieron las constantes de disociación por SPR. Los resultados están recogidos en la Tabla 8.

Tabla 8. Resultados de interacción de los análogos sintéticos de la amodiaquina.

Compuesto	STD	K _D (SPR)
ANÁLOGO-1	+	> 1 mM
ANÁLOGO-2	+	239 μM
ANÁLOGO-3	+	169 μM
ANÁLOGO-4	+	> 1 mM
ANÁLOGO-5	-	-
ANÁLOGO-6	+	precipita
ANÁLOGO-7	+	> 1 mM
ANÁLOGO-8	+	> 1 mM
ANÁLOGO-9	-	-
ANÁLOGO-10	+	75 μM
ANÁLOGO-11	-	-
ANÁLOGO-12	+	79 μM
ANÁLOGO-13	+	639 μM
ANÁLOGO-14	+	1 mM

Los resultados recogidos en la Tabla 8 muestran que ninguno de estos compuestos mejora los obtenidos para la amodiaquina cuya K_D era de $52,6 \mu\text{M}$ (apartado 9 de Discusión y Resultados). Sin embargo se puede llegar a algunas conclusiones sobre qué parte de la molécula es importante para la interacción. En este análisis no se ha tenido en cuenta los resultados obtenidos para el análogo A6 ya que al precipitar a las concentraciones necesarias para RMN y SPR no se puede tener seguridad de que los resultados sean correctos. Al comparar los 5 primeros análogos (Fig. 48) podemos deducir que es importante que la molécula tenga un heteroátomo en la posición *para* de la benzamina ya que al eliminarlo como ocurre en el análogo-A4 se pierde la interacción. Ocurre lo mismo cuando se sustituye este heteroátomo por un bromo (análogo-A1). Además se puede observar que es necesaria la presencia del grupo que hay en la posición *meta* del anillo posiblemente porque ayude a orientar la molécula en la interacción. Cuando se elimina este grupo se puede también observar que se pierde la interacción.

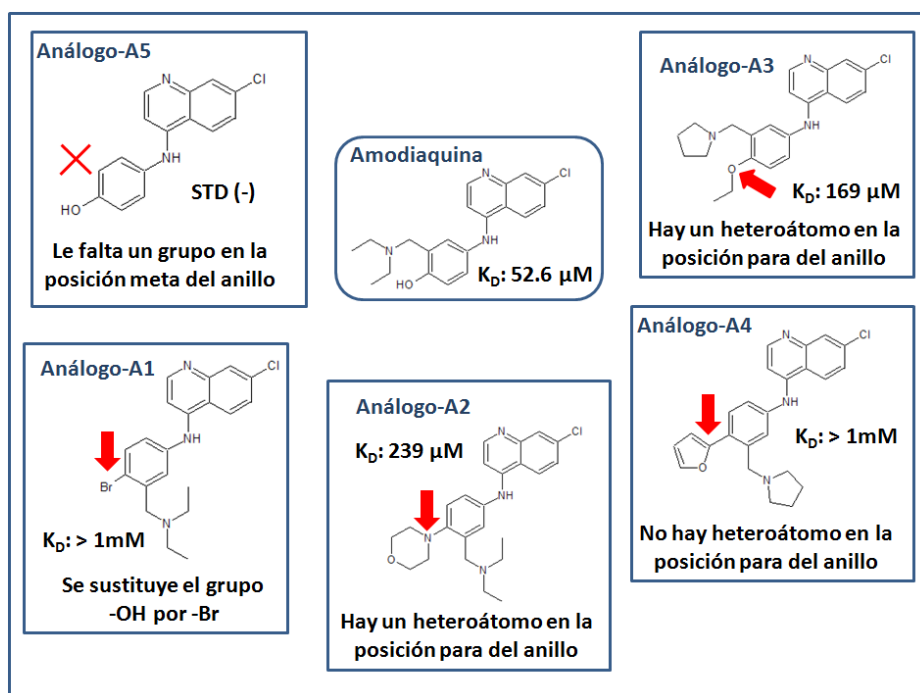


Figura 48. Comparación de estructuras de análogos sintéticos de lamodiaquina (A1-A5).

En los análogos 10 y 11 (Fig. 49), el heteroátomo que se encuentra en la posición para de la benzamina podría verse compensado con el oxígeno del grupo amida de la posición meta o/y el CH₂OH de la segunda posición meta. Sin embargo, la sustitución del oxígeno del morfolino por un 4-hidroxipiperidina (grupos señalados con un anillo de color rojo en la figura 49) hace que la molécula pierda su actividad.

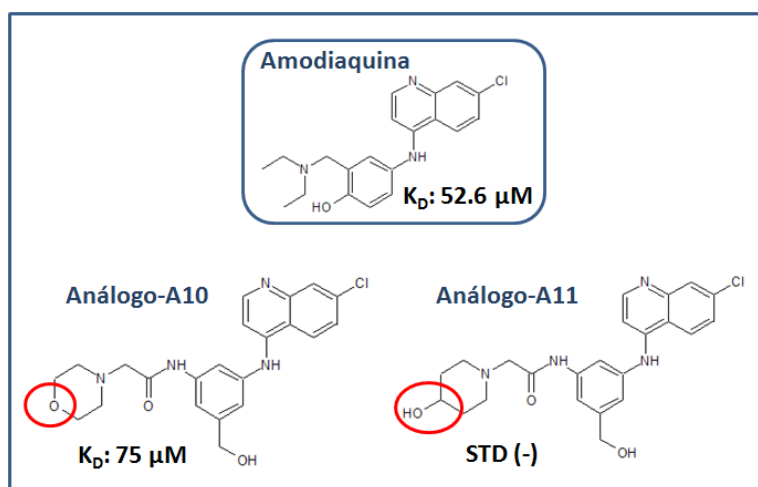


Figura 49. Comparación de estructuras de análogos sintéticos de amodiaquin A10 Y A11

Por último como se puede ver en la Fig. 50, se puede concluir que la adición de grupos demasiado voluminosos que dan rigidez a la molécula hacen que disminuya o se pierda la interacción con la proteína. Sin embargo si observamos el análogo-A12 que contiene de nuevo un grupo amida en la posición meta cuya orientación es igual que el análogo-A10 y un CH₂-O- (grupo éster) en la segunda posición meta se recupera la actividad. Esta configuración sólo existe en los compuestos 10, 11 y 12, lo que nos lleva de nuevo a concluir que la ausencia del heteroátomo en la posición para de la benzamina se compensa con estos dos grupos en las dos posiciones meta de la benzamina. Sin embargo, sería necesario realizar estos experimentos con un mayor número de compuestos con

diferente diversidad química para completar este análisis de la relación entre la estructura y la afinidad de estos compuestos por la proteína.

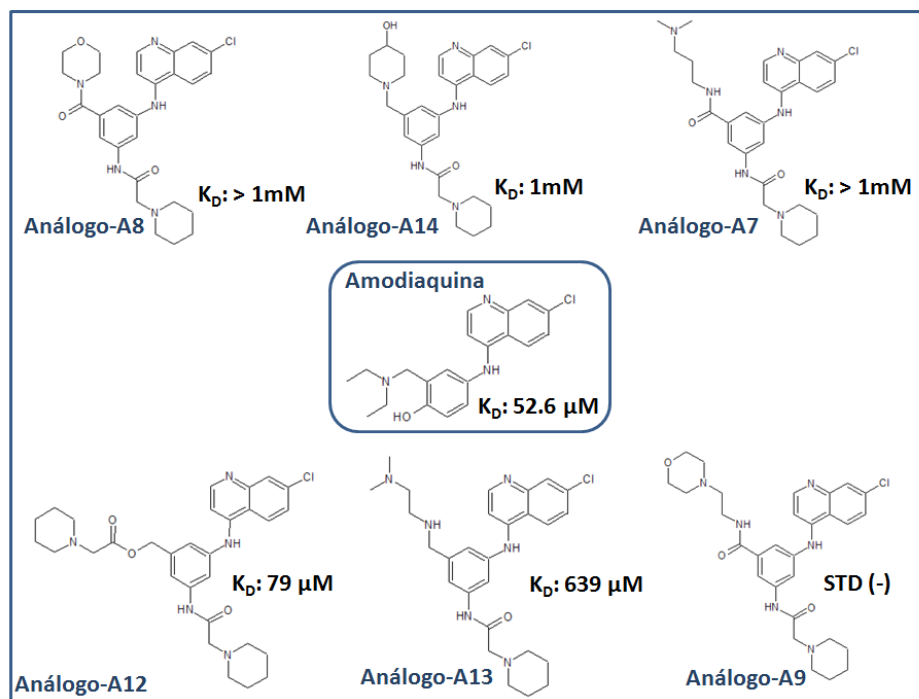


Figura 50. Comparación de estructuras de análogos sintéticos de amodiaquin A7, A8, A9, A12, A13, A14

Conclusiones

- Se han diseñado y clonado diferentes construcciones de heparanasa: HEP₃₆₋₅₄₃ (52 kDa), HEP₁₅₈₋₅₄₃ (45 kDa), HEP₁₅₈₋₄₁₇ (37 kDa), HEP₁₅₈₋₄₁₇ (29 kDa), HEP₂₃₃₋₄₁₇ (21 kDa), HEP₃₆₋₁₀₉ (8 kDa), y se ha optimizado el método de expresión y purificación de estas construcciones. Con este procedimiento se ha obtenido la proteína soluble, estable y en concentración adecuada para poder llevar a cabo experimentos mediante diferentes técnicas biofísicas como RMN, SPR, X-ray, calorimetría, etc.
- Se ha realizado un estudio estructural por RMN de varias de estas construcciones: HEP₁₅₈₋₄₁₇ (37 kDa), HEP₁₅₈₋₄₁₇ (29 kDa), HEP₂₃₃₋₄₁₇ (21 kDa) y se ha conseguido asignar el esqueleto y las cadenas laterales de las mismas, lo que ha permitido obtener información de la estructura secundaria de la heparanasa.
- Se ha validado la construcción HEP₁₅₈₋₄₁₇ para utilizarla en campañas de cribado y buscar nuevos inhibidores de la enzima. Además se han optimizado las condiciones para realizar este cribado de compuestos mediante RMN y SPR.
- Se han identificado varios “hits” con afinidad por debajo de 1 mM, entre ellos un fragmento con una K_D de 236 μM y LE de 0,26 $\text{Kcal}\cdot\text{mol}^{-1}$ y un análogo del anterior con una K_D de 79 μM y LE de 0,29 $\text{Kcal}\cdot\text{mol}^{-1}$. Estas dos moléculas representan un buen punto de partida para la síntesis de nuevos inhibidores de la enzima.
- La amodiaquina, un fármaco conocido para el tratamiento de la malaria, mostró una K_D de 52,6 μM frente a la heparanasa lo que podría significar la búsqueda futura de una nueva indicación para la amodiaquina y compuestos de su familia. Un análisis preliminar de diferentes análogos de amodiaquina, nos ha permitido establecer relaciones entre la estructura de estos compuestos y afinidad por la heparanasa, clave para el desarrollo de inhibidores más potentes.

Bibliografía

- (1) Leber, M. F., and Efferth, T. (2009) Molecular principles of cancer invasion and metastasis (review). *Int J Oncol* 34, 881-95.
- (2) Gabriel, J. A. (2007) *The Biology of Cancer*.
- (3) Hanahan, D., and Weinberg, R. A. (2000) The hallmarks of cancer. *Cell* 100, 57-70.
- (4) Bishop, D., and Weinberg, R. A. (1996) *Molecular Oncology*.
- (5) Croce, C. M. (2008) Oncogenes and cancer. *N Engl J Med* 358, 502-11.
- (6) Konopka, J. B., Watanabe, S. M., Singer, J. W., Collins, S. J., and Witte, O. N. (1985) Cell lines and clinical isolates derived from Ph1-positive chronic myelogenous leukemia patients express c-abl proteins with a common structural alteration. *Proc Natl Acad Sci U S A* 82, 1810-4.
- (7) Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E., and Croce, C. M. (1985) The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science* 229, 1390-3.
- (8) Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J., and Haber, D. A. (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350, 2129-39.
- (9) Rajalingam, K., Schreck, R., Rapp, U. R., and Albert, S. (2007) Ras oncogenes and their downstream targets. *Biochim Biophys Acta* 1773, 1177-95.
- (10) Wieser, R. J., Renauer, D., Schafer, A., Heck, R., Engel, R., Schutz, S., and Oesch, F. (1990) Growth control in mammalian cells by cell-cell contacts. *Environ Health Perspect* 88, 251-3.
- (11) Hahn, W. C., and Weinberg, R. A. (2002) Rules for making human tumor cells. *N Engl J Med* 347, 1593-603.
- (12) Hengartner, M. O. (2000) The biochemistry of apoptosis. *Nature* 407, 770-6.
- (13) Cotter, T. G. (2009) Apoptosis and cancer: the genesis of a research field. *Nat Rev Cancer* 9, 501-7.
- (14) Igney, F. H., and Krammer, P. H. (2002) Death and anti-death: tumour resistance to apoptosis. *Nat Rev Cancer* 2, 277-88.
- (15) Hockenbery, D., Nunez, G., Milliman, C., Schreiber, R. D., and Korsmeyer, S. J. (1990) Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature* 348, 334-6.

- (16) Yang, J., Liu, X., Bhalla, K., Kim, C. N., Ibrado, A. M., Cai, J., Peng, T. I., Jones, D. P., and Wang, X. (1997) Prevention of apoptosis by Bcl-2: release of cytochrome c from mitochondria blocked. *Science* 275, 1129-32.
- (17) Lane, D. P. (1992) Cancer. p53, guardian of the genome. *Nature* 358, 15-6.
- (18) Yonish-Rouach, E., Resnitzky, D., Lotem, J., Sachs, L., Kimchi, A., and Oren, M. (1991) Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6. *Nature* 352, 345-7.
- (19) Miyashita, T., Krajewski, S., Krajewska, M., Wang, H. G., Lin, H. K., Liebermann, D. A., Hoffman, B., and Reed, J. C. (1994) Tumor suppressor p53 is a regulator of bcl-2 and bax gene expression in vitro and in vivo. *Oncogene* 9, 1799-805.
- (20) Miyashita, T., and Reed, J. C. (1995) Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell* 80, 293-9.
- (21) Zhan, Q., Fan, S., Bae, I., Guillouf, C., Liebermann, D. A., O'Connor, P. M., and Fornace, A. J., Jr. (1994) Induction of bax by genotoxic stress in human cells correlates with normal p53 status and apoptosis. *Oncogene* 9, 3743-51.
- (22) Lapenko, O., and Prives, C. (2006) Transcriptional regulation by p53: one protein, many possibilities. *Cell Death Differ.* 13, 951-61.
- (23) Trauth, B. C., Klas, C., Peters, A. M., Matzku, S., Moller, P., Falk, W., Debatin, K. M., and Krammer, P. H. (1989) Monoclonal antibody-mediated tumor regression by induction of apoptosis. *Science* 245, 301-5.
- (24) Folkman, J. (1990) What is the evidence that tumors are angiogenesis dependent? *J Natl Cancer Inst* 82, 4-6.
- (25) Eichhorn, M. E., Kleespies, A., Angele, M. K., Jauch, K. W., and Bruns, C. J. (2007) Angiogenesis in cancer: molecular mechanisms, clinical impact. *Langenbecks Arch Surg* 392, 371-9.
- (26) Algire, G. H., and HW, C. (1945) Vascular reactions of normal and malignant tissue in vivo. *N Engl J Med* 285, 1182-6
- (27) Folkman, J. (1971) Tumor angiogenesis: therapeutic implications. *N Engl J Med* 285, 1182-6.
- (28) Folkman, J., and Kalluri, R. (2004) Cancer without disease. *Nature* 427, 787.
- (29) Muehlbauer, P. M. (2003) Anti-angiogenesis in cancer therapy. *Semin Oncol Nurs* 19, 180-92.

- (30) Bergers, G., and Benjamin, L. E. (2003) Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* 3, 401-10.
- (31) Auguste, P., Lemiere, S., Larrieu-Lahargue, F., and Bikfalvi, A. (2005) Molecular mechanisms of tumor vascularization. *Crit Rev Oncol Hematol* 54, 53-61.
- (32) Hayflick, L. (1997) Mortality and immortality at the cellular level. A review. *Biochemistry (Mosc)* 62, 1180-90.
- (33) Bryan, T. M., and Cech, T. R. (1999) Telomerase and the maintenance of chromosome ends. *Curr Opin Cell Biol* 11, 318-24.
- (34) Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., and Reddel, R. R. (1995) Telomere elongation in immortal human cells without detectable telomerase activity. *Embo J* 14, 4240-8.
- (35) Christofori, G. (2006) New signals from the invasive front. *Nature* 441, 444-50.
- (36) Gupta, G. P., and Massague, J. (2006) Cancer metastasis: building a framework. *Cell* 127, 679-95.
- (37) Steeg, P. S. (2006) Tumor metastasis: mechanistic insights and clinical challenges. *Nat Med* 12, 895-904.
- (38) Nguyen, D. X., Bos, P. D., and Massague, J. (2009) Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer* 9, 274-84.
- (39) Weinberg, R. A. (2006) *The Biology of Cancer*.
- (40) Heimann, R., and Hellman, S. (1998) Aging, progression, and phenotype in breast cancer. *J Clin Oncol* 16, 2686-92.
- (41) Klein, C. A. (2009) Parallel progression of primary tumours and metastases. *Nat Rev Cancer* 9, 302-12.
- (42) Klein, G. (1998) Foulds' dangerous idea revisited: the multistep development of tumors 40 years later. *Adv Cancer Res* 72, 1-23.
- (43) Collins, V. P., Loeffler, R. K., and Tivey, H. (1956) Observations on growth rates of human tumors. *Am J Roentgenol Radium Ther Nucl Med* 76, 988-1000.
- (44) Friberg, S., and Mattson, S. (1997) On the growth rates of human malignant tumors: implications for medical decision making. *J Surg Oncol* 65, 284-97.
- (45) Fidler, I. J. (2003) The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer* 3, 453-8.
- (46) Camerer, E., Qazi, A. A., Duong, D. N., Cornelissen, I., Advincula, R., and Coughlin, S. R. (2004) Platelets, protease-activated receptors, and fibrinogen in hematogenous metastasis. *Blood* 104, 397-401.

- (47) Borsig, L., Wong, R., Feramisco, J., Nadeau, D. R., Varki, N. M., and Varki, A. (2001) Heparin and cancer revisited: mechanistic connections involving platelets, P-selectin, carcinoma mucins, and tumor metastasis. *Proc Natl Acad Sci U S A* 98, 3352-7.
- (48) Muller, Y. A., Kelley, R. F., and de Vos, A. M. (1998) Hinge bending within the cytokine receptor superfamily revealed by the 2.4 Å crystal structure of the extracellular domain of rabbit tissue factor. *Protein Sci* 7, 1106-15.
- (49) Crissman, J. D., Hatfield, J. S., Menter, D. G., Sloane, B., and Honn, K. V. (1988) Morphological study of the interaction of intravascular tumor cells with endothelial cells and subendothelial matrix. *Cancer Res* 48, 4065-72.
- (50) Ferrara, N., and Kerbel, R. S. (2005) Angiogenesis as a therapeutic target. *Nature* 438, 967-74.
- (51) Sun, S., and Schiller, J. H. (2007) Angiogenesis inhibitors in the treatment of lung cancer. *Crit Rev Oncol Hematol* 62, 93-104.
- (52) Leung, D. W., Cachianes, G., Kuang, W. J., Goeddel, D. V., and Ferrara, N. (1989) Vascular endothelial growth factor is a secreted angiogenic mitogen. *Science* 246, 1306-9.
- (53) Dvorak, H. F., Brown, L. F., Detmar, M., and Dvorak, A. M. (1995) Vascular permeability factor/vascular endothelial growth factor, microvascular hyperpermeability, and angiogenesis. *Am J Pathol* 146, 1029-39.
- (54) Ferrara, N., Gerber, H. P., and LeCouter, J. (2003) The biology of VEGF and its receptors. *Nat Med* 9, 669-76.
- (55) Rosen, L. S. (2005) VEGF-targeted therapy: therapeutic potential and recent advances. *Oncologist* 10, 382-91.
- (56) Zhang, F., Tang, Z., Hou, X., Lennartsson, J., Li, Y., Koch, A. W., Scotney, P., Lee, C., Arjunan, P., Dong, L., Kumar, A., Rissanen, T. T., Wang, B., Nagai, N., Fons, P., Fariss, R., Zhang, Y., Wawrousek, E., Tansey, G., Raber, J., Fong, G. H., Ding, H., Greenberg, D. A., Becker, K. G., Herbert, J. M., Nash, A., Yla-Herttuala, S., Cao, Y., Watts, R. J., and Li, X. (2009) VEGF-B is dispensable for blood vessel growth but critical for their survival, and VEGF-B targeting inhibits pathological angiogenesis. *Proc Natl Acad Sci U S A* 106, 6152-7.
- (57) Veikkola, T., and Alitalo, K. (1999) VEGFs, receptors and angiogenesis. *Semin Cancer Biol* 9, 211-20.
- (58) Makrilia, N., Lappa, T., Xyla, V., Nikolaidis, I., and Syrigos, K. (2009) The role of angiogenesis in solid tumours: an overview. *Eur J Intern Med* 20, 663-71.

- (59) de Vries, C., Escobedo, J. A., Ueno, H., Houck, K., Ferrara, N., and Williams, L. T. (1992) The fms-like tyrosine kinase, a receptor for vascular endothelial growth factor. *Science* 255, 989-91.
- (60) Park, J. E., Chen, H. H., Winer, J., Houck, K. A., and Ferrara, N. (1994) Placenta growth factor. Potentiation of vascular endothelial growth factor bioactivity, in vitro and in vivo, and high affinity binding to Flt-1 but not to Flk-1/KDR. *J Biol Chem* 269, 25646-54.
- (61) Olofsson, B., Korpelainen, E., Pepper, M. S., Mandriota, S. J., Aase, K., Kumar, V., Gunji, Y., Jeltsch, M. M., Shibuya, M., Alitalo, K., and Eriksson, U. (1998) Vascular endothelial growth factor B (VEGF-B) binds to VEGF receptor-1 and regulates plasminogen activator activity in endothelial cells. *Proc Natl Acad Sci U S A* 95, 11709-14.
- (62) Fong, G. H., Rossant, J., Gertsenstein, M., and Breitman, M. L. (1995) Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium. *Nature* 376, 66-70.
- (63) Hiratsuka, S., Maru, Y., Okada, A., Seiki, M., Noda, T., and Shibuya, M. (2001) Involvement of Flt-1 tyrosine kinase (vascular endothelial growth factor receptor-1) in pathological angiogenesis. *Cancer Res* 61, 1207-13.
- (64) Gille, H., Kowalski, J., Li, B., LeCouter, J., Moffat, B., Zioncheck, T. F., Pelletier, N., and Ferrara, N. (2001) Analysis of biological effects and signaling properties of Flt-1 (VEGFR-1) and KDR (VEGFR-2). A reassessment using novel receptor-specific vascular endothelial growth factor mutants. *J Biol Chem* 276, 3222-30.
- (65) Dvorak, H. F. (2002) Vascular permeability factor/vascular endothelial growth factor: a critical cytokine in tumor angiogenesis and a potential target for diagnosis and therapy. *J Clin Oncol* 20, 4368-80.
- (66) Zeng, H., Dvorak, H. F., and Mukhopadhyay, D. (2001) Vascular permeability factor (VPF)/vascular endothelial growth factor (VEGF) peceptor-1 down-modulates VPF/VEGF receptor-2-mediated endothelial cell proliferation, but not migration, through phosphatidylinositol 3-kinase-dependent pathways. *J Biol Chem* 276, 26969-79.
- (67) Millauer, B., Longhi, M. P., Plate, K. H., Shawver, L. K., Risau, W., Ullrich, A., and Strawn, L. M. (1996) Dominant-negative inhibition of Flk-1 suppresses the growth of many tumor types in vivo. *Cancer Res* 56, 1615-20.

-
- (68) Yancopoulos, G. D., Klagsbrun, M., and Folkman, J. (1998) Vasculogenesis, angiogenesis, and growth factors: ephrins enter the fray at the border. *Cell* 93, 661-4.
- (69) Plank, M. J., Sleeman, B. D., and Jones, P. F. (2004) The role of the angiopoietins in tumour angiogenesis. *Growth Factors* 22, 1-11.
- (70) Maisonpierre, P. C., Suri, C., Jones, P. F., Bartunkova, S., Wiegand, S. J., Radziejewski, C., Compton, D., McClain, J., Aldrich, T. H., Papadopoulos, N., Daly, T. J., Davis, S., Sato, T. N., and Yancopoulos, G. D. (1997) Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis. *Science* 277, 55-60.
- (71) Gospodarowicz, D., Abraham, J. A., and Schilling, J. (1989) Isolation and characterization of a vascular endothelial cell mitogen produced by pituitary-derived folliculo stellate cells. *Proc Natl Acad Sci U S A* 86, 7311-5.
- (72) Nishimura, T., Nakatake, Y., Konishi, M., and Itoh, N. (2000) Identification of a novel FGF, FGF-21, preferentially expressed in the liver. *Biochim Biophys Acta* 1492, 203-6.
- (73) Terranova, V. P., DiFlorio, R., Lyall, R. M., Hic, S., Friesel, R., and Maciag, T. (1985) Human endothelial cells are chemotactic to endothelial cell growth factor and heparin. *J Cell Biol* 101, 2330-4.
- (74) Thomas, K. A. (1987) Fibroblast growth factors. *Faseb J* 1, 434-40.
- (75) Ornitz, D. M., Yayon, A., Flanagan, J. G., Svahn, C. M., Levi, E., and Leder, P. (1992) Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Mol Cell Biol* 12, 240-7.
- (76) Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., and Ornitz, D. M. (1991) Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell* 64, 841-8.
- (77) Rapraeger, A. C., Krufka, A., and Olwin, B. B. (1991) Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science* 252, 1705-8.
- (78) Pineda-Lucena, A., Jimenez, M. A., Lozano, R. M., Nieto, J. L., Santoro, J., Rico, M., and Gimenez-Gallego, G. (1996) Three-dimensional structure of acidic fibroblast growth factor in solution: effects of binding to a heparin functional analog. *J Mol Biol* 264, 162-78.

- (79) Schlessinger, J., Plotnikov, A. N., Ibrahimi, O. A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., Linhardt, R. J., and Mohammadi, M. (2000) Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol Cell* 6, 743-50.
- (80) Andrae, J., Gallini, R., and Betsholtz, C. (2008) Role of platelet-derived growth factors in physiology and medicine. *Genes Dev* 22, 1276-312.
- (81) Yang, X. P., Pei, Z. H., and Ren, J. (2009) Making up or breaking up: the tortuous role of platelet-derived growth factor in vascular ageing. *Clin Exp Pharmacol Physiol* 36, 739-47.
- (82) Crawford, Y., Kasman, I., Yu, L., Zhong, C., Wu, X., Modrusan, Z., Kaminker, J., and Ferrara, N. (2009) PDGF-C mediates the angiogenic and tumorigenic properties of fibroblasts associated with tumors refractory to anti-VEGF treatment. *Cancer Cell* 15, 21-34.
- (83) Kong, D., Banerjee, S., Huang, W., Li, Y., Wang, Z., Kim, H. R., and Sarkar, F. H. (2008) Mammalian target of rapamycin repression by 3,3'-diindolylmethane inhibits invasion and angiogenesis in platelet-derived growth factor-D-overexpressing PC3 cells. *Cancer Res* 68, 1927-34.
- (84) Zhang, R., Shah, M. V., Yang, J., Nyland, S. B., Liu, X., Yun, J. K., Albert, R., and Loughran, T. P., Jr. (2008) Network model of survival signaling in large granular lymphocyte leukemia. *Proc Natl Acad Sci U S A* 105, 16308-13.
- (85) di Tomaso, E., London, N., Fuja, D., Logie, J., Tyrrell, J. A., Kamoun, W., Munn, L. L., and Jain, R. K. (2009) PDGF-C induces maturation of blood vessels in a model of glioblastoma and attenuates the response to anti-VEGF treatment. *PLoS One* 4, e5123.
- (86) Liegl, B., Reich, O., Nogales, F. F., and Regauer, S. (2006) PDGF-alpha and PDGF-beta are expressed in endometrial stromal sarcoma: a potential therapeutic target for tyrosine kinase inhibitors? *Histopathology* 49, 545-6.
- (87) Pardali, E., and ten Dijke, P. (2009) Transforming growth factor-beta signaling and tumor angiogenesis. *Front Biosci* 14, 4848-61.
- (88) Michalaki, V., Syrigos, K., Charles, P., and Waxman, J. (2004) Serum levels of IL-6 and TNF-alpha correlate with clinicopathological features and patient survival in patients with prostate cancer. *Br J Cancer* 90, 2312-6.
- (89) Karayiannakis, A. J., Syrigos, K. N., Polychronidis, A., Pitiakoudis, M., Bounovas, A., and Simopoulos, K. (2001) Serum

- levels of tumor necrosis factor-alpha and nutritional status in pancreatic cancer patients. *Anticancer Res* 21, 1355-8.
- (90) Tuszynski, G. P., Srivastava, S., Switalska, H. I., Holt, J. C., Cierniewski, C. S., and Niewiarowski, S. (1985) The interaction of human platelet thrombospondin with fibrinogen. Thrombospondin purification and specificity of interaction. *J Biol Chem* 260, 12240-5.
- (91) Lawler, J. (1986) The structural and functional properties of thrombospondin. *Blood* 67, 1197-209.
- (92) Sargiannidou, I., Zhou, J., and Tuszynski, G. P. (2001) The role of thrombospondin-1 in tumor progression. *Exp Biol Med (Maywood)* 226, 726-33.
- (93) Kirsch, M., Strasser, J., Allende, R., Bello, L., Zhang, J., and Black, P. M. (1998) Angiostatin suppresses malignant glioma growth in vivo. *Cancer Res* 58, 4654-9.
- (94) Li, X., Jiang, L., Wang, Y., Xiao, Y., Huang, Y., Yao, Q., Yang, Y., and Wu, X. (2007) Inhibition of angiogenesis by a novel small peptide consisting of the active fragments of platelet factor-4 and vasostatin. *Cancer Lett* 256, 29-32.
- (95) Ribeiro-Silva, A., and Oliveira da Costa, J. P. (2008) Osteopontin expression according to molecular profile of invasive breast cancer: a clinicopathological and immunohistochemical study. *Int J Biol Markers* 23, 154-60.
- (96) Qi, J. H., Dai, G., Luthert, P., Chaurasia, S., Hollyfield, J., Weber, B. H., Stohr, H., and Anand-Apte, B. (2009) S156C mutation in tissue inhibitor of metalloproteinases-3 induces increased angiogenesis. *J Biol Chem* 284, 19927-36.
- (97) Luster, A. D. (1998) Chemokines--chemotactic cytokines that mediate inflammation. *N Engl J Med* 338, 436-45.
- (98) Struyf, S., Proost, P., and Van Damme, J. (2003) Regulation of the immune response by the interaction of chemokines and proteases. *Adv Immunol* 81, 1-44.
- (99) Rossi, D., and Zlotnik, A. (2000) The biology of chemokines and their receptors. *Annu Rev Immunol* 18, 217-42.
- (100) Zlotnik, A., and Yoshie, O. (2000) Chemokines: a new classification system and their role in immunity. *Immunity* 12, 121-7.
- (101) Vicari, A. P., and Caux, C. (2002) Chemokines in cancer. *Cytokine Growth Factor Rev* 13, 143-54.
- (102) Balkwill, F. (1998) The molecular and cellular biology of the chemokines. *J Viral Hepat* 5, 1-14.

- (103) Strieter, R. M., Polverini, P. J., Kunkel, S. L., Arenberg, D. A., Burdick, M. D., Kasper, J., Dzuiba, J., Van Damme, J., Walz, A., Marriott, D., and et al. (1995) The functional role of the ELR motif in CXC chemokine-mediated angiogenesis. *J Biol Chem* 270, 27348-57.
- (104) Luca, M., Huang, S., Gershenwald, J. E., Singh, R. K., Reich, R., and Bar-Eli, M. (1997) Expression of interleukin-8 by human melanoma cells up-regulates MMP-2 activity and increases tumor growth and metastasis. *Am J Pathol* 151, 1105-13.
- (105) Inoue, K., Slaton, J. W., Eve, B. Y., Kim, S. J., Perrotte, P., Balbay, M. D., Yano, S., Bar-Eli, M., Radinsky, R., Pettaway, C. A., and Dinney, C. P. (2000) Interleukin 8 expression regulates tumorigenicity and metastases in androgen-independent prostate cancer. *Clin Cancer Res* 6, 2104-19.
- (106) Wong, S. Y., and Hynes, R. O. (2006) Lymphatic or hematogenous dissemination: how does a metastatic tumor cell decide? *Cell Cycle* 5, 812-7.
- (107) Muller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M. E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S. N., Barrera, J. L., Mohar, A., Verastegui, E., and Zlotnik, A. (2001) Involvement of chemokine receptors in breast cancer metastasis. *Nature* 410, 50-6.
- (108) Chiang, A. C., and Massague, J. (2008) Molecular basis of metastasis. *N Engl J Med* 359, 2814-23.
- (109) Rudzki, Z., and Jothy, S. (1997) CD44 and the adhesion of neoplastic cells. *Mol Pathol* 50, 57-71.
- (110) Isacke, C. M., and Yarwood, H. (2002) The hyaluronan receptor, CD44. *Int J Biochem Cell Biol* 34, 718-21.
- (111) Stauder, R., Eisterer, W., Thaler, J., and Gunthert, U. (1995) CD44 variant isoforms in non-Hodgkin's lymphoma: a new independent prognostic factor. *Blood* 85, 2885-99.
- (112) Brinckerhoff, C. E., and Matrisian, L. M. (2002) Matrix metalloproteinases: a tail of a frog that became a prince. *Nat Rev Mol Cell Biol* 3, 207-14.
- (113) Page-McCaw, A., Ewald, A. J., and Werb, Z. (2007) Matrix metalloproteinases and the regulation of tissue remodelling. *Nat Rev Mol Cell Biol* 8, 221-33.
- (114) Rundhaug, J. E. (2003) Matrix metalloproteinases, angiogenesis, and cancer: commentary re: A. C. Lockhart et al., Reduction of wound angiogenesis in patients treated with BMS-275291, a broad spectrum matrix metalloproteinase inhibitor. *Clin. Cancer Res.*, 9: 00-00, 2003. *Clin Cancer Res* 9, 551-4.

- (115) Iozzo, R. V., and San Antonio, J. D. (2001) Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena. *J Clin Invest* 108, 349-55.
- (116) Kjellen, L., and Lindahl, U. (1991) Proteoglycans: structures and interactions. *Annu Rev Biochem* 60, 443-75.
- (117) Hacker, U., Nybakken, K., and Perrimon, N. (2005) Heparan sulphate proteoglycans: the sweet side of development. *Nat Rev Mol Cell Biol* 6, 530-41.
- (118) Bernfield, M., Gotte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J., and Zako, M. (1999) Functions of cell surface heparan sulfate proteoglycans. *Annu Rev Biochem* 68, 729-77.
- (119) Iozzo, R. V. (1998) Matrix proteoglycans: from molecular design to cellular function. *Annu Rev Biochem* 67, 609-52.
- (120) Korner, G., Bjornsson, T. D., and Vlodavsky, I. (1993) Extracellular matrix produced by cultured corneal and aortic endothelial cells contains active tissue-type and urokinase-type plasminogen activators. *J Cell Physiol* 154, 456-65.
- (121) Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R., Bitan, M., Pappo, O., Peretz, T., Michal, I., Spector, L., and Pecker, I. (1999) Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nat Med* 5, 793-802.
- (122) Vlodavsky, I., and Friedmann, Y. (2001) Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J Clin Invest* 108, 341-7.
- (123) Parish, C. R., Freeman, C., and Hulett, M. D. (2001) Heparanase: a key enzyme involved in cell invasion. *Biochim Biophys Acta* 1471, M99-108.
- (124) Stetler-Stevenson, W. G., Aznavoorian, S., and Liotta, L. A. (1993) Tumor cell interactions with the extracellular matrix during invasion and metastasis. *Annu Rev Cell Biol* 9, 541-73.
- (125) Werb, Z. (1997) ECM and cell surface proteolysis: regulating cellular ecology. *Cell* 91, 439-42.
- (126) Liotta, L. A., and Kohn, E. C. (2001) The microenvironment of the tumour-host interface. *Nature* 411, 375-9.
- (127) Stetler-Stevenson, W. G., and Yu, A. E. (2001) Proteases in invasion: matrix metalloproteinases. *Semin Cancer Biol* 11, 143-52.
- (128) Hook, M., Wasteson, A., and Oldberg, A. (1975) A heparan sulfate-degrading endoglycosidase from rat liver tissue. *Biochem Biophys Res Commun* 67, 1422-8.

- (129) Nakajima, M., Irimura, T., Di Ferrante, D., Di Ferrante, N., and Nicolson, G. L. (1983) Heparan sulfate degradation: relation to tumor invasive and metastatic properties of mouse B16 melanoma sublines. *Science* 220, 611-3.
- (130) Nakajima, M., Irimura, T., Di Ferrante, N., and Nicolson, G. L. (1984) Metastatic melanoma cell heparanase. Characterization of heparan sulfate degradation fragments produced by B16 melanoma endoglucuronidase. *J Biol Chem* 259, 2283-90.
- (131) Hulett, M. D., Freeman, C., Hamdorf, B. J., Baker, R. T., Harris, M. J., and Parish, C. R. (1999) Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis. *Nat Med* 5, 803-9.
- (132) Toyoshima, M., and Nakajima, M. (1999) Human heparanase. Purification, characterization, cloning, and expression. *J Biol Chem* 274, 24153-60.
- (133) Kussie, P. H., Hulmes, J. D., Ludwig, D. L., Patel, S., Navarro, E. C., Seddon, A. P., Giorgio, N. A., and Bohlen, P. (1999) Cloning and functional expression of a human heparanase gene. *Biochem Biophys Res Commun* 261, 183-7.
- (134) Miao, H. Q., Navarro, E., Patel, S., Sargent, D., Koo, H., Wan, H., Plata, A., Zhou, Q., Ludwig, D., Bohlen, P., and Kussie, P. (2002) Cloning, expression, and purification of mouse heparanase. *Protein Expr Purif* 26, 425-31.
- (135) Goldshmidt, O., Zcharia, E., Aingorn, H., Guatta-Rangini, Z., Atzmon, R., Michal, I., Pecker, I., Mitrani, E., and Vlodaysky, I. (2001) Expression pattern and secretion of human and chicken heparanase are determined by their signal peptide sequence. *J Biol Chem* 276, 29178-87.
- (136) Nasser, N. J., Nevo, E., Shafat, I., Ilan, N., Vlodaysky, I., and Avivi, A. (2005) Adaptive evolution of heparanase in hypoxia-tolerant *Spalax*: gene cloning and identification of a unique splice variant. *Proc Natl Acad Sci U S A* 102, 15161-6.
- (137) Kizaki, K., Nakano, H., Nakano, H., Takahashi, T., Imai, K., and Hashizume, K. (2001) Expression of heparanase mRNA in bovine placenta during gestation. *Reproduction* 121, 573-80.
- (138) Nasser, N. J. (2008) Heparanase involvement in physiology and disease. *Cell Mol Life Sci* 65, 1706-15.
- (139) Nasser, N. J., Avivi, A., Shushy, M., Vlodaysky, I., and Nevo, E. (2007) Cloning, expression, and characterization of an alternatively spliced variant of human heparanase. *Biochem Biophys Res Commun* 354, 33-8.

- (140) Bame, K. J. (2001) Heparanases: endoglycosidases that degrade heparan sulfate proteoglycans. *Glycobiology* 11, 91R-98R.
- (141) Simizu, S., Ishida, K., and Osada, H. (2004) Heparanase as a molecular target of cancer chemotherapy. *Cancer Sci* 95, 553-8.
- (142) Parish, C. R., Freeman, C., Brown, K. J., Francis, D. J., and Cowden, W. B. (1999) Identification of sulfated oligosaccharide-based inhibitors of tumor growth and metastasis using novel in vitro assays for angiogenesis and heparanase activity. *Cancer Res* 59, 3433-41.
- (143) Fairbanks, M. B., Mildner, A. M., Leone, J. W., Cavey, G. S., Mathews, W. R., Drong, R. F., Slightom, J. L., Bienkowski, M. J., Smith, C. W., Bannow, C. A., and Heinrikson, R. L. (1999) Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer. *J Biol Chem* 274, 29587-90.
- (144) Levy-Adam, F., Miao, H. Q., Heinrikson, R. L., Vlodaysky, I., and Ilan, N. (2003) Heterodimer formation is essential for heparanase enzymatic activity. *Biochem Biophys Res Commun* 308, 885-91.
- (145) McKenzie, E., Young, K., Hircock, M., Bennett, J., Bhaman, M., Felix, R., Turner, P., Stamps, A., McMillan, D., Saville, G., Ng, S., Mason, S., Snell, D., Schofield, D., Gong, H., Townsend, R., Gallagher, J., Page, M., Parekh, R., and Stubberfield, C. (2003) Biochemical characterization of the active heterodimer form of human heparanase (Hpa1) protein expressed in insect cells. *Biochem J* 373, 423-35.
- (146) Vlodaysky, I., Goldshmidt, O., Zcharia, E., Atzmon, R., Rangini-Guatta, Z., Elkin, M., Peretz, T., and Friedmann, Y. (2002) Mammalian heparanase: involvement in cancer metastasis, angiogenesis and normal development. *Semin Cancer Biol* 12, 121-9.
- (147) Nadav, L., Eldor, A., Yacoby-Zeevi, O., Zamir, E., Pecker, I., Ilan, N., Geiger, B., Vlodaysky, I., and Katz, B. Z. (2002) Activation, processing and trafficking of extracellular heparanase by primary human fibroblasts. *J Cell Sci* 115, 2179-87.
- (148) Simizu, S., Ishida, K., Wierzba, M. K., and Osada, H. (2004) Secretion of heparanase protein is regulated by glycosylation in human tumor cell lines. *J Biol Chem* 279, 2697-703.
- (149) Simizu, S., Suzuki, T., Muroi, M., Lai, N. S., Takagi, S., Dohmae, N., and Osada, H. (2007) Involvement of disulfide bond formation in the activation of heparanase. *Cancer Res* 67, 7841-9.
- (150) Hulett, M. D., Hornby, J. R., Ohms, S. J., Zuegg, J., Freeman, C., Gready, J. E., and Parish, C. R. (2000) Identification of active-site

- residues of the pro-metastatic endoglycosidase heparanase. *Biochemistry* 39, 15659-67.
- (151) Levy-Adam, F., Abboud-Jarrous, G., Guerrini, M., Beccati, D., Vlodaysky, I., and Ilan, N. (2005) Identification and characterization of heparin/heparan sulfate binding domains of the endoglycosidase heparanase. *J Biol Chem* 280, 20457-66.
- (152) Shteper, P. J., Zcharia, E., Ashhab, Y., Peretz, T., Vlodaysky, I., and Ben-Yehuda, D. (2003) Role of promoter methylation in regulation of the mammalian heparanase gene. *Oncogene* 22, 7737-49.
- (153) Miao, H. Q., Elkin, M., Aingorn, E., Ishai-Michaeli, R., Stein, C. A., and Vlodaysky, I. (1999) Inhibition of heparanase activity and tumor metastasis by laminarin sulfate and synthetic phosphorothioate oligodeoxynucleotides. *Int J Cancer* 83, 424-31.
- (154) Ogishima, T., Shiina, H., Breault, J. E., Tabatabai, L., Bassett, W. W., Enokida, H., Li, L. C., Kawakami, T., Urakami, S., Ribeiro-Filho, L. A., Terashima, M., Fujime, M., Igawa, M., and Dahiya, R. (2005) Increased heparanase expression is caused by promoter hypomethylation and up-regulation of transcriptional factor early growth response-1 in human prostate cancer. *Clin Cancer Res* 11, 1028-36.
- (155) Ogishima, T., Shiina, H., Breault, J. E., Terashima, M., Honda, S., Enokida, H., Urakami, S., Tokizane, T., Kawakami, T., Ribeiro-Filho, L. A., Fujime, M., Kane, C. J., Carroll, P. R., Igawa, M., and Dahiya, R. (2005) Promoter CpG hypomethylation and transcription factor EGR1 hyperactivate heparanase expression in bladder cancer. *Oncogene* 24, 6765-72.
- (156) Baraz, L., Haupt, Y., Elkin, M., Peretz, T., and Vlodaysky, I. (2006) Tumor suppressor p53 regulates heparanase gene expression. *Oncogene* 25, 3939-47.
- (157) Bartlett, M. R., Underwood, P. A., and Parish, C. R. (1995) Comparative analysis of the ability of leucocytes, endothelial cells and platelets to degrade the subendothelial basement membrane: evidence for cytokine dependence and detection of a novel sulfatase. *Immunol Cell Biol* 73, 113-24.
- (158) Lider, O., Baharav, E., Mekori, Y. A., Miller, T., Naparstek, Y., Vlodaysky, I., and Cohen, I. R. (1989) Suppression of experimental autoimmune diseases and prolongation of allograft survival by treatment of animals with low doses of heparins. *J Clin Invest* 83, 752-6.
- (159) Parish, C. R., Hindmarsh, E. J., Bartlett, M. R., Staykova, M. A., Cowden, W. B., and Willenborg, D. O. (1998) Treatment of

- central nervous system inflammation with inhibitors of basement membrane degradation. *Immunol Cell Biol* 76, 104-13.
- (160) Naparstek, Y., Cohen, I. R., Fuks, Z., and Vlodaysky, I. (1984) Activated T lymphocytes produce a matrix-degrading heparan sulphate endoglycosidase. *Nature* 310, 241-4.
- (161) Edovitsky, E., Lerner, I., Zcharia, E., Peretz, T., Vlodaysky, I., and Elkin, M. (2006) Role of endothelial heparanase in delayed-type hypersensitivity. *Blood* 107, 3609-16.
- (162) Elkin, M., Cohen, I., Zcharia, E., Orgel, A., Guatta-Rangini, Z., Peretz, T., Vlodaysky, I., and Kleinman, H. K. (2003) Regulation of heparanase gene expression by estrogen in breast cancer. *Cancer Res* 63, 8821-6.
- (163) Xu, X., Ding, J., Rao, G., Shen, J., Prinz, R. A., Rana, N., and Dmowski, W. P. (2007) Estradiol induces heparanase-1 expression and heparan sulphate proteoglycan degradation in human endometrium. *Hum Reprod* 22, 927-37.
- (164) Cohen, I., Maly, B., Simon, I., Meirovitz, A., Pikarsky, E., Zcharia, E., Peretz, T., Vlodaysky, I., and Elkin, M. (2007) Tamoxifen induces heparanase expression in estrogen receptor-positive breast cancer. *Clin Cancer Res* 13, 4069-77.
- (165) Han, J., Woytowich, A. E., Mandal, A. K., and Hiebert, L. M. (2007) Heparanase upregulation in high glucose-treated endothelial cells is prevented by insulin and heparin. *Exp Biol Med (Maywood)* 232, 927-34.
- (166) Goldshmidt, O., Nadav, L., Aingorn, H., Irit, C., Feinstein, N., Ilan, N., Zamir, E., Geiger, B., Vlodaysky, I., and Katz, B. Z. (2002) Human heparanase is localized within lysosomes in a stable form. *Exp Cell Res* 281, 50-62.
- (167) Vreys, V., Delande, N., Zhang, Z., Coomans, C., Roebroek, A., Durr, J., and David, G. (2005) Cellular uptake of mammalian heparanase precursor involves low density lipoprotein receptor-related proteins, mannose 6-phosphate receptors, and heparan sulfate proteoglycans. *J Biol Chem* 280, 33141-8.
- (168) Cohen, E., Atzmon, R., Vlodaysky, I., and Ilan, N. (2005) Heparanase processing by lysosomal/endosomal protein preparation. *FEBS Lett* 579, 2334-8.
- (169) Zetser, A., Levy-Adam, F., Kaplan, V., Gingis-Velitski, S., Bashenko, Y., Schubert, S., Flugelman, M. Y., Vlodaysky, I., and Ilan, N. (2004) Processing and activation of latent heparanase occurs in lysosomes. *J Cell Sci* 117, 2249-58.

- (170) Marchetti, D., Li, J., and Shen, R. (2000) Astrocytes contribute to the brain-metastatic specificity of melanoma cells by producing heparanase. *Cancer Res* 60, 4767-70.
- (171) Goldshmidt, O., Zcharia, E., Abramovitch, R., Metzger, S., Aingorn, H., Friedmann, Y., Schirmacher, V., Mitrani, E., and Vlodaysky, I. (2002) Cell surface expression and secretion of heparanase markedly promote tumor angiogenesis and metastasis. *Proc Natl Acad Sci U S A* 99, 10031-6.
- (172) Dempsey, L. A., Brunn, G. J., and Platt, J. L. (2000) Heparanase, a potential regulator of cell-matrix interactions. *Trends Biochem Sci* 25, 349-51.
- (173) Zcharia, E., Zilka, R., Yaar, A., Yacoby-Zeevi, O., Zetser, A., Metzger, S., Sarid, R., Naggi, A., Casu, B., Ilan, N., Vlodaysky, I., and Abramovitch, R. (2005) Heparanase accelerates wound angiogenesis and wound healing in mouse and rat models. *Faseb J* 19, 211-21.
- (174) Zcharia, E., Philp, D., Edovitsky, E., Aingorn, H., Metzger, S., Kleinman, H. K., Vlodaysky, I., and Elkin, M. (2005) Heparanase regulates murine hair growth. *Am J Pathol* 166, 999-1008.
- (175) Folkman, J., and Shing, Y. (1992) Control of angiogenesis by heparin and other sulfated polysaccharides. *Adv Exp Med Biol* 313, 355-64.
- (176) Vlodaysky, I., Miao, H. Q., Medalion, B., Danagher, P., and Ron, D. (1996) Involvement of heparan sulfate and related molecules in sequestration and growth promoting activity of fibroblast growth factor. *Cancer Metastasis Rev* 15, 177-86.
- (177) Pikas, D. S., Li, J. P., Vlodaysky, I., and Lindahl, U. (1998) Substrate specificity of heparanases from human hepatoma and platelets. *J Biol Chem* 273, 18770-7.
- (178) Rohloff, J., Zinke, J., Schoppmeyer, K., Tannapfel, A., Witzigmann, H., Mossner, J., Wittekind, C., and Caca, K. (2002) Heparanase expression is a prognostic indicator for postoperative survival in pancreatic adenocarcinoma. *Br J Cancer* 86, 1270-5.
- (179) Ilan, N., Elkin, M., and Vlodaysky, I. (2006) Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis. *Int J Biochem Cell Biol* 38, 2018-39.
- (180) Fux, L., Ilan, N., Sanderson, R. D., and Vlodaysky, I. (2009) Heparanase: busy at the cell surface. *Trends Biochem Sci* 34, 511-9.
- (181) Zhang, Y. F., Tang, X. D., Gao, J. H., Fang, D. C., and Yang, S. M. (2011) Heparanase: a universal immunotherapeutic target in human cancers. *Drug Discov Today* 16, 412-7.

- (182) Vlodaysky, I., Fuks, Z., Bar-Ner, M., Ariav, Y., and Schirmmacher, V. (1983) Lymphoma cell-mediated degradation of sulfated proteoglycans in the subendothelial extracellular matrix: relationship to tumor cell metastasis. *Cancer Res* 43, 2704-11.
- (183) Abboud-Jarrous, G., Rangini-Guetta, Z., Aingorn, H., Atzmon, R., Elgavish, S., Peretz, T., and Vlodaysky, I. (2005) Site-directed mutagenesis, proteolytic cleavage, and activation of human proheparanase. *J Biol Chem* 280, 13568-75.
- (184) Takahashi, H., Ebihara, S., Okazaki, T., Asada, M., Sasaki, H., and Yamaya, M. (2005) A comparison of the effects of unfractionated heparin, dalteparin and danaparoid on vascular endothelial growth factor-induced tumour angiogenesis and heparanase activity. *Br J Pharmacol* 146, 333-43.
- (185) Ferro, V., Hammond, E., and Fairweather, J. K. (2004) The development of inhibitors of heparanase, a key enzyme involved in tumour metastasis, angiogenesis and inflammation. *Mini Rev Med Chem* 4, 693-702.
- (186) Kumar, N., Bentolila, A., and Domb, A. J. (2005) Structure and biological activity of heparinoid. *Mini Rev Med Chem* 5, 441-7.
- (187) Zhao, H., Liu, H., Chen, Y., Xin, X., Li, J., Hou, Y., Zhang, Z., Zhang, X., Xie, C., Geng, M., and Ding, J. (2006) Oligomannuric acid sulfate, a novel heparanase inhibitor simultaneously targeting basic fibroblast growth factor, combats tumor angiogenesis and metastasis. *Cancer Res* 66, 8779-87.
- (188) Fairweather, J. K., Hammond, E., Johnstone, K. D., and Ferro, V. (2008) Synthesis and heparanase inhibitory activity of sulfated manooligosaccharides related to the antiangiogenic agent PI-88. *Bioorg Med Chem* 16, 699-709.
- (189) Lewis, K. D., Robinson, W. A., Millward, M. J., Powell, A., Price, T. J., Thomson, D. B., Walpole, E. T., Haydon, A. M., Creese, B. R., Roberts, K. L., Zalberg, J. R., and Gonzalez, R. (2008) A phase II study of the heparanase inhibitor PI-88 in patients with advanced melanoma. *Invest New Drugs* 26, 89-94.
- (190) Hammond, E., Bytheway, I., and Ferro, V. (2006) Heparanase as a target for anticancer therapeutics: New developments and future prospects. *New Developments in Therapeutic Glycomics* 81-7736-266-6, 251-282.
- (191) Lipinski, M. (2001) Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol.* 44, 235-249.

- (192) Congreve, M., Chessari, G., Tisi, D., and Woodhead, A. J. (2008) Recent developments in fragment-based drug discovery. *J Med Chem* 51, 3661-80.
- (193) Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003) A 'rule of three' for fragment-based lead discovery? *Drug Discov Today* 8, 876-7.
- (194) Carr, R., and Jhoti, H. (2002) Structure-based screening of low-affinity compounds. *Drug Discov Today* 7, 522-7.
- (195) Hartshorn, M. J., Murray, C. W., Cleasby, A., Frederickson, M., Tickle, I. J., and Jhoti, H. (2005) Fragment-based lead discovery using X-ray crystallography. *J Med Chem* 48, 403-13.
- (196) Jahnke, W., and Erlanson, D. A. (2006) *Fragment-based Approaches in Drug Discovery*.
- (197) Hopkins, A. L., Groom, C. R., and Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* 9, 430-1.
- (198) Rees, D. C., Congreve, M., Murray, C. W., and Carr, R. (2004) Fragment-based lead discovery. *Nat Rev Drug Discov* 3, 660-72.
- (199) Carr, R. A., Congreve, M., Murray, C. W., and Rees, D. C. (2005) Fragment-based lead discovery: leads by design. *Drug Discov Today* 10, 987-92.
- (200) Siegal, G., Ab, E., and Schultz, J. (2007) Integration of fragment screening and library design. *Drug Discov Today* 12, 1032-9.
- (201) Bartoli, S., Fincham, C. I., and Fattori, D. (2006) The fragment-approach: An update. *Drug Discov Today: Technologies* 3, 425-31.
- (202) Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531-4.
- (203) Mayer, M., and Meyer, B. (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed. Print* 38, 1784-88.
- (204) Dalvit, C., Pevarello, P., Tato, M., Veronesi, M., Vulpetti, A., and Sundstrom, M. (2000) Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J Biomol NMR* 18, 65-8.
- (205) Savchenko, A., Yee, A., Khachatryan, A., Skarina, T., Evdokimova, E., Pavlova, M., Semesi, A., Northey, J., Beasley, S., Lan, N., Das, R., Gerstein, M., Arrowmith, C. H., and Edwards, A. M. (2003) Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous

- proteins and of using both NMR and X-ray crystallography approaches. *Proteins* 50, 392-9.
- (206) Deborah, A., Polayes, A. G., Gary Ward, A., and Hohn Hughes, J. (1994) TEV Protease, recombinant: A site-specific protease for efficient cleavage of affinity tags from expressed proteins. *Focus* 16, 2-5.
- (207) Zeng, G. (1998) Sticky-end PCR: new method for subcloning. *Biotechniques* 25, 206-8.
- (208) Venters, R. A., Huang, C. C., Farmer, B. T., 2nd, Trolard, R., Spicer, L. D., and Fierke, C. A. (1995) High-level ²H/¹³C/¹⁵N labeling of proteins for NMR studies. *J Biomol NMR* 5, 339-44.
- (209) Englander, J., Cohen, L., Arshava, B., Estephan, R., Becker, J. M., and Naider, F. (2006) Selective labeling of a membrane peptide with ¹⁵N-amino acids using cells grown in rich medium. *Biopolymers* 84, 508-18.
- (210) Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M., and Stockman, B. (2001) WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J Biomol NMR* 21, 349-59.
- (211) Aszalos, A., Bax, A., Burlinson, N., Roller, P., and McNeal, C. (1985) Physico-chemical and microbiological comparison of nystatin, amphotericin A and amphotericin B, and structure of amphotericin A. *J Antibiot (Tokyo)* 38, 1699-713.
- (212) Bodenhausen, G., and Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced hetero-nuclear spectroscopy. *Chem. Phys. Lett.* 69, 185-189.
- (213) Kay, L. E., Ikura, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* 89, 496-514.
- (214) Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* 114, 6291-93.
- (215) Wittekind, M., and Mueller, L. (1993) HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *J. Magn. Reson., Series B.* B101, 201-5.
- (216) Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three dimensional triple-resonance pulse scheme to correlate intraresidue ¹H, ¹⁵N labeled proteins. *J Magn Reson* 100, 209-13.
- (217) Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond

- J(HNH.alpha.) coupling constants in ^{15}N -enriched proteins. *J. Am. Chem. Soc.* *115*, 7772-7777.
- (218) Düx, P., Whitehead, B., Boelens, R., Kaptein, R., and Vuister, G. W. (1997) Measurement of ^{15}N - ^1H coupling constants in uniformly ^{15}N -labeled proteins: Application to the photoactive yellow protein. *J Biomol NMR* *10*, 301-306.
- (219) Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J. Biomol. NMR* 185-204.
- (220) Sattler, M., Schwalbe, H., and Griesinger, C. (1992) Stereospecific assignment of leucine methyl groups with carbon-13 in natural abundance or with random ^{13}C labeling. *Journal of the American Chemical Society* *114*, 1126-7.
- (221) Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) ^1H - ^1H correlation via isotropic mixing of ^{13}C magnetization. A new three-dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C -enriched proteins. *J Magn Reson* *88*, 425-431.
- (222) Jahnke, W., Baur, G., Gemmecker, G., and Kessler, H. (1995) Improved Accuracy of NMR Structures by a Modified NOESY-HSQC Experiment. *J Magn Reson* *106 B*, 86-88.
- (223) Fesik, S. W., and Zuiderweg, E. R. P. (1988) Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. *J. Magn. Reson.* *78*, 588-593.
- (224) Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* *94*, 12366-71.
- (225) Salzmänn, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci U S A* *95*, 13585-90.
- (226) Salzmänn, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1999) ^{13}C -constant-time ^{15}N , ^1H -TROSY-HNCA for sequential assignments of large proteins. *J Biomol NMR* *14*, 85-8.
- (227) Salzmänn, M., Wider, G., Pervushin, K., Senn, H., and Wuthrich, K. (1999) TROSY-type Triple-Resonance Experiments for Sequential NMR Assignments of Large Proteins. *J Am Chem Soc* *121*, 844-48.
- (228) Hwang, T. L., and Shaka, A. J. (1998) Multiple-pulse mixing sequences that selectively enhance chemical exchange or cross-

- relaxation peaks in high-resolution NMR spectra. *J Magn Reson* 135, 280-7.
- (229) Ikura, M., Kay, L. E., and Bax, A. (1990) A novel approach for sequential assignment of ¹H, ¹³C, and ¹⁵N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* 29, 4659-67.
- (230) Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13, 289-302.
- (231) Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics Chapter 5*, Unit 5 6.
- (232) Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* 29, 291-325.
- (233) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234, 779-815.
- (234) Fiser, A., Do, R. K., and Sali, A. (2000) Modeling of loops in protein structures. *Protein Sci* 9, 1753-73.
- (235) Bostrom, J. (2001) Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. *J Comput Aided Mol Des* 15, 1137-52.
- (236) Schulz-Gasch, T., and Stahl, M. (2003) Binding site characteristics in structure-based virtual screening: evaluation of current docking tools. *J Mol Model* 9, 47-57.
- (237) Hovel, K., Shallom, D., Niefind, K., Belakhov, V., Shoham, G., Baasov, T., Shoham, Y., and Schomburg, D. (2003) Crystal structure and snapshots along the reaction pathway of a family 51 alpha-L-arabinofuranosidase. *Embo J* 22, 4922-32.
- (238) Fux, L., Feibish, N., Cohen-Kaplan, V., Gingis-Velitski, S., Feld, S., Geffen, C., Vlodaysky, I., and Ilan, N. (2009) Structure-function approach identifies a COOH-terminal domain that mediates heparanase signaling. *Cancer Res* 69, 1758-67.
- (239) Nardella, C., Lahm, A., Pallaoro, M., Brunetti, M., Vannini, A., and Steinkuhler, C. (2004) Mechanism of activation of human heparanase investigated by protein engineering. *Biochemistry* 43, 1862-73.

- (240) Nevo, E., Ivanitskaya, I., and Beiles, A. (2001) Adaptive Radiation of Blind Subterranean Mole Rats. *Backhuys, Leiden, The Netherlands*.
- (241) Nasser, N. J., Avivi, A., Shafat, I., Edovitsky, E., Zcharia, E., Ilan, N., Vlodayky, I., and Nevo, E. (2009) Alternatively spliced Spalax heparanase inhibits extracellular matrix degradation, tumor growth, and metastasis. *Proc Natl Acad Sci U S A* 106, 2253-8.
- (242) Kapust, R. B., and Waugh, D. S. (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* 8, 1668-74.
- (243) Nallamsetty, S., and Waugh, D. S. (2006) Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr Purif* 45, 175-82.
- (244) Waugh, D. S. (2005) Making the most of affinity tags. *Trends Biotechnol* 23, 316-20.
- (245) Gupta, R. K., and Koenig, S. H. (1971) Some aspects of pH and temperature dependence of the NMR spectra of cytochrome C. *Biochem Biophys Res Commun* 45, 1134-43.
- (246) Kelly, A. E., Ou, H. D., Withers, R., and Dotsch, V. (2002) Low-conductivity buffers for high-sensitivity NMR measurements. *J Am Chem Soc* 124, 12013-9.
- (247) Foster, M. P., McElroy, C. A., and Amero, C. D. (2007) Solution NMR of large molecules and assemblies. *Biochemistry* 46, 331-40.
- (248) Clore, G. M., and Gronenborn, A. M. (1994) Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol* 239, 349-63.
- (249) Biological Magnetic Resonance Data Bank. http://www.bmrb.wisc.edu/ref_info/statsel.htm.
- (250) LeMaster, D. M. (1994) Isotope labelling in solution protein assignment and structural analysis. *Prog. NMR Spectrosc.* 26, 371--419.
- (251) Gardner, K. H., and Kay, L. E. (1998) The use of ^2H , ^{13}C , ^{15}N multidimensional NMR GTO study the structure and dynamics of proteins. *Annual Review of Biophysics and Biomolecular Structure* 27, 357-406.
- (252) Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1998) Transverse relaxation-optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in C-13-labeled proteins. *J Am Chem Soc* 120, 6394-6400.

- (253) Pervushin, K., Wider, G., and Wuthrich, K. (1998) Single transition-to-single transition polarization transfer (ST2-PT) in (N-15,H-1)-TROSY. *J Biolmol. NMR* 12, 345-348.
- (254) Wuthrich, K., Billeter, M., and Braun, W. (1984) Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J Mol Biol* 180, 715-40.
- (255) Carbajo, R. J., Silvester, J. A., Runswick, M. J., Walker, J. E., and Neuhaus, D. (2004) Solution structure of subunit F(6) from the peripheral stalk region of ATP synthase from bovine heart mitochondria. *J Mol Biol* 342, 593-603.
- (256) Schwarzingler, S., Kroon, G. J., Foss, T. R., Wright, P. E., and Dyson, H. J. (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J Biomol NMR* 18, 43-8.
- (257) Nieba, L., Nieba-Axmann, S. E., Persson, A., Hamalainen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A. F., and Pluckthun, A. (1997) BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Anal Biochem* 252, 217-28.
- (258) Yoshitani, N., Saito, K., Saikawa, W., Asanuma, M., Yokoyama, S., and Hirota, H. (2007) NTA-mediated protein capturing strategy in screening experiments for small organic molecules by surface plasmon resonance. *Proteomics* 7, 494-9.
- (259) Johnsson, B., Lofas, S., and Lindquist, G. (1991) Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal Biochem* 198, 268-77.
- (260) Nakajima, M., DeChavigny, A., Johnson, C. E., Hamada, J., Stein, C. A., and Nicolson, G. L. (1991) Suramin. A potent inhibitor of melanoma heparanase and invasion. *J Biol Chem* 266, 9661-6.
- (261) Marchetti, D., Reiland, J., Erwin, B., and Roy, M. (2003) Inhibition of heparanase activity and heparanase-induced angiogenesis by suramin analogues. *Int J Cancer* 104, 167-74.
- (262) Meyer, B., and Peters, T. (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl* 42, 864-90.
- (263) Wolber, G., Seidel, T., Bendix, F., and Langer, T. (2008) Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discov Today* 13, 23-9.
- (264) Willett, P., Barnard, J. M., and Downs, G. M. (1998) Chemical similarity searching. *J. Chem. Inf. Comput. Sci.* 38, 983-96.

- (265) Dalvit, C., Fasolini, M., Flocco, M., Knapp, S., Pevarello, P., and Veronesi, M. (2002) NMR-Based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands. *J Med Chem* 45, 2610-4.
- (266) Ludwig, C., Michiels, P. J., Wu, X., Kavanagh, K. L., Pilka, E., Jansson, A., Oppermann, U., and Gunther, U. L. (2008) SALMON: solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy. *J Med Chem* 51, 1-3.
- (267) Stockman, B. J., Kothe, M., Kohls, D., Weibley, L., Connolly, B. J., Sheils, A. L., Cao, Q., Cheng, A. C., Yang, L., Kamath, A. V., Ding, Y. H., and Charlton, M. E. (2009) Identification of allosteric PIF-pocket ligands for PDK1 using NMR-based fragment screening and 1H-15N TROSY experiments. *Chem Biol Drug Des* 73, 179-88.
- (268) Klages, J., Coles, M., and Kessler, H. (2007) NMR-based screening: a powerful tool in fragment-based drug discovery. *Analyst* 132, 693-705.
- (269) Mosulen, S., Orti, L., Bas, E., Carbajo, R. J., and Pineda-Lucena, A. Production of heparanase constructs suitable for nuclear magnetic resonance and drug discovery studies. *Biopolymers* 95, 151-60.
- (270) Petitjean, M. (1992) Applications of the radius-diameter diagram to the classification of topological and geometrical shapes of chemical compounds. *J Chem Inf Comput Sci* 32, 331-337.
- (271) Dalvit, C. (2007) Ligand- and substrate-based ¹⁹F NMR screening: principles and applications to drug discovery. *Prog Nucl Magn Reson Spectrosc* 51, 243-271.
- (272) Abad-Zapatero, C. (2007) Ligand efficiency indices for effective drug discovery. *Expert Opin Drug Discov* 2, 469-488.