

11281  
38



**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO**

**INSTITUTO DE FISILOGIA CELULAR  
DOCTORADO EN CIENCIAS BIOMEDICAS**

**CLONACION Y CARACTERIZACION DEL  
PROMOTOR Y DEL GENE DEL BETAGLICANO,  
EL RECEPTOR III DEL TGF-beta**

**TESIS**

QUE PARA OBTENER EL GRADO DE

**DOCTOR EN CIENCIAS BIOMEDICAS**

PRESENTA LA MEDICA CIRUJANA

MARTHA VERONICA PONCE CASTAÑEDA

CIUDAD UNIVERSITARIA, 2002

TESIS CON  
FALLA DE ORIGEN



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**PAGINACION**

**DISCONTINUA**

El desarrollo de esta tesis se realizó bajo la dirección del Dr. Fernando López Casillas, en el Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México y con el financiamiento del Consejo Nacional de Ciencia y Tecnología (CONACYT), la Dirección General de Asuntos del Personal Académico de la UNAM (DGAPA) y del Howard Hughes Medical Institute, USA.

## **Miembros del Jurado**

**Dr. Jesús Chimal Monroy**

**Dr. Roberto Coria Ortega**

**Dra. Alicia González Manjarrez**

**Dr. Fernando López Casillas**

**Dr. Félix Recillas Targa**

**Dr. Ruy Pérez Montfort**

**Dr. Angel Zarain Herzberg**

**A todos aquellos y aquellas que me han hecho sentir fuerte en algún momento**

**A Fernando y a Francisco Fernando**

## INDICE DE CONTENIDOS

Indice de Figuras	i
Abreviaturas	ii
Resumen	iii
Abstract	iv
<b>I. INTRODUCCION</b>	<b>1</b>
Características del Betaglicano	
Función del Betaglicano	
Betaglicano Soluble	
TGF- $\beta$	
Funciones del TGF- $\beta$	
Estudio del gene de Betaglicano	
Clonación del cDNA Murino del Betaglicano	
Caracterización del extremo 5' del mRNA de Betaglicano	
Identificación del Promotor	
Ensayos Funcionales	
Construcciones Alb-Betaglicano	
<b>II. PLANTEAMIENTO DEL PROBLEMA E HIPOTESIS</b>	<b>14</b>
<b>III. OBJETIVOS Y ABORDAJE EXPERIMENTAL</b>	<b>15</b>
<b>IV. MATERIAL Y METODOS</b>	<b>16</b>
<b>V. CLONACION DEL cDNA MURINO</b>	<b>19</b>
<b>VI. CARACTERIZACION DEL EXTREMO 5' DEL mRNA</b>	<b>26</b>
<b>VII. ESTRUCTURA DEL GENE HUMANO</b>	<b>38</b>
<b>VIII. CLONACION DEL PROMOTOR Y ENSAYOS FUNCIONALES</b>	<b>45</b>
<b>IX. CONSTRUCCION DE TRANSGENES Alb-BETAGLICANO</b>	<b>55</b>
<b>X. PERSPECTIVAS Y CONSIDERACIONES</b>	<b>63</b>
<b>XI. BIBLIOGRAFIA</b>	<b>65</b>
<b>XII. PUBLICACION</b>	<b>70</b>

## INDICE DE FIGURAS

Figura		Página
Figura 1	Estructura del Betaglicano	2
Figura 2	Similitud entre Betaglicano y Endoglina	3
Figura 3	Marcado por afinidad en células mink	4
Figura 4	Dendrograma I	5
Figura 5	Vía de señalización del TGF- $\beta$	7
Figura 6	Marcado por afinidad a las 3 isoformas de TGF- $\beta$	8
Figura 7	Modelo en listones del TGF- $\beta$ 1	10
Figura 8	Marcado por afinidad del cDNA murino	20
Figura 9	Oligonucleótidos usados para mutagénesis	21
Figura 10	Estrategia general de mutagénesis	22
Figura 11	Marcado por afinidad de mutantes	23
Figura 12	Dendrograma II	24
Figura 13	Southernblot de colonias recombinantes	28
Figura 14	Secuencias RACE 8 y 9	29
Figura 15	Ideograma y mapas genéticos I	30
Figura 16	Comparación de secuencias obtenidas con BLAST	31
Figura 17	Contenido exónico de RACE 8 y 9	32
Figura 18	Secuencias IRES y LOX-DICE	33
Figura 19	IRES viral y estructuras secundarias RACE	35
Figura 20	Alineamiento múltiple de 5'UTRs	36
Figura 21	Ideograma y mapas genéticos II	38
Figura 22	Contig NT_004686.5	39
Figura 23	Mapa exónico de Betaglicano	41
Figura 24	Secuencia nucleotídica y de a.a.	43
Figura 25	Secuencia PCRtemp	47
Figura 26	Mapa y estrategia de clonación	49
Figura 27	Productos de amplificación, gel de agarosa	51
Figura 27'	pGL3B - pGL2.1hBG	52
Figura 28	Inducción transcripcional, actividad de luciferasa	53
Figura 29	Factores sobre la diferenciación	54
Figura 30	Posibles sitios de reconocimiento de FT miogénicos	56
Figura 31	pAlb e/p y pAlb TGFbeta1	57
Figura 32	Estrategia de construcción de transgenes	59
Figura 33	Marcado por afinidad de transgenes	60
Figura 34	Marcado por afinidad del medio condicionado	61

## ABREVIATURAS

$\alpha$ -dCTP	alfa dioxicitidina trifosfato (alpha deoxy citidine triphosphate)
AMV RT	Transcriptasa Reversa del virus de aves Maloney (Avian Maloney Virus Reverse Transcriptase)
BAC	Cromosoma Bacteriano Artificial (Bacterial Artificial Chromosome)
BG	Betaglicano
BLASTn	Algoritmo para comparar secuencias nucleotídicas
cDNA	Acido Desoxi-Ribonucleico complementario (complementary Desoxy   Ribonucleic Acid)
CMV	Citomegalovirus
COS1	Línea celular de riñón de mono
DMEM	Medio Dulbeco Modificado de Eagle (Dulbecco Modified Eagle's Medium)
ddNTP	Dideoxinucleótidos fosfatos (dideoxynucleotides phosphate)
DNA	Acido Desoxi-Ribonucleico (Desoxy Ribonucleic Acid)
DSS	Disuccimidil-suberato
FGF	Factor de Crecimiento Fibroblástico (Fibroblast Growth Factor)
FISH	Hibridación <i>In Situ</i> Fluorescente (Fluorescent In Situ Hybridization)]
GSP 3'	Oligonucleótido Específico del Gene (Gene Specific Primer)
HepG2	Línea celular de hepatocarcinoma humano
hGH	Hormona de Crecimiento Humana (Human Growth Hormone)
MEM	Medio Mínimo Esencial de Eagle (Minimum Essential Medium Eagle)
mRNA	Acido Ribonucleico mensajero (messenger Ribonucleic Acid)
NCBI	Centro Nacional para Información Biotecnológica (National Center for Biotechnology Information)
nt	nucleótido
ORI'	Apertura del Marco de Lectura (Open Reading Frame)
PAGE / SDS	Electroforesis en Gel de poliacrilamida / Sodio Dodecyl Sulfato (Polyacrilamide Gel Electrophoresis Sodium Dodecyl Sulphate)
pAlb/Enh	Vector con secuencias del elemento distal y promotor del gene de albúmina
pBSKS	Vector de clonación Bluescript KS
pCDNA3	Vector de expresión usado en células eucariontes con el gene que confiere resistencia a neomicina
pCMV5	Vector de expresión usado para transfecciones transitorias en células eucariontes.
PCR	Reacción de Polimerasa en Cadena (Polymerase Chain Reaction)
pGL3-B	Vector de expresión con el gene de luciferasa de luciérnaga
RACE	Amplificación Rápida de Extremos de cDNA)Rapid Amplification of cDNA Ends
RNA	Acido Ribonucleico (Ribonucleic Acid)
TGF- $\beta$	Factor de Crecimiento Transformante (Transforming Growth Factor Beta)
5' UTR	Región 5' no traducida (Untranslated Region)

## Resumen

El betaglicano o receptor III del TGF beta, es una proteína con un solo pase transmembranal componente del sistema de receptores del TGF- $\beta$  en la superficie celular. Existe una forma membranal y una forma soluble del betaglicano, ambas formas unen al TGF- $\beta$  con alta afinidad y se ha propuesto que constituyen un interruptor molecular. La forma membranal facilita la unión del TGF- $\beta$  a los receptores señaladores funcionando como agonista, mientras que la forma soluble secuestra al TGF- $\beta$  funcionando como antagonista. El TGF- $\beta$  es el prototipo de una familia de citocinas involucradas en funciones celulares críticas del desarrollo embrionario y participante central de la homeostasis de tejidos adultos. Se han identificado mutaciones en su vía de señalización en diversos cánceres y ciertos trastornos del desarrollo, por otro lado un exceso de TGF- $\beta$  ha sido identificado como mediador de fenómenos fibróticos en un número importante de enfermedades humanas. Debido a esto el papel del betaglicano como regulador de esta citocina es de gran relevancia biológica y médica. Para estudiar los mecanismos de regulación del betaglicano, en esta tesis se reporta la clonación del cDNA de ratón debido a nuestro interés de estudiar estos mecanismos en el modelo murino. La caracterización del extremo 5' UTR del mRNA humano, la identificación del promotor y la distribución exónica del gene humano. Y la construcción de dos transgenes de betaglicano bajo el control del promotor murino de albúmina, y los ensayos funcionales en células en cultivo.

## Abstract

Betaglycan or TGF- $\beta$  type III receptor, is a transmembrane protein component of the TGF- $\beta$  receptor system on the cell surface. There are two forms of betaglycan, a transmembrane form and a soluble form, both maintain high affinity for TGF- $\beta$  and they are proposed to function as a molecular switch. The transmembrane form facilitates TGF- $\beta$  binding to signaling receptors, functioning as an agonist, while the soluble form keeps away TGF- $\beta$  from signaling receptors, functioning as an antagonist. TGF- $\beta$  is the prototype of a cytokine family involved in critical cellular functions during embryonic development and a central player in adult tissues homeostasis. Mutations in its signaling pathway have been identified in different cancer types and development defects, and TGF- $\beta$  excess has been identified as a fibrotic mediator in a considerable number of important human diseases. Thus betaglycan's regulatory role on TGF- $\beta$  has great medical and biological relevance. To approach betaglycan's own regulatory mechanisms from a genetic point of view, in this thesis is reported: mouse betaglycan cDNA cloning due to our interest in studying regulatory mechanism in the mouse model. Human 5' UTR mRNA characterization, human promoter identification, exon-intron structure of the gene and the construction of two betaglycan transgenes under control of murine albumin promoter, and its functional assays in tissue culture.

## I. INTRODUCCION

El interés en la exploración y estudio del Betaglicano se deriva de su papel regulador sobre las funciones del TGF- $\beta$ , las cuales son críticas durante el desarrollo embrionario y durante el mantenimiento de la homeostasis en tejidos adultos.

### Características del Betaglicano

El Betaglicano o receptor III del TGF- $\beta$  es una proteína glicosilada de membrana presente en casi todas las células humanas en donde se le ha buscado. Además de los carbohidratos complejos unidos a residuos de Asparagina, posee también glicosaminoglicanos de ahí que se le clasifique como un proteoglicano [1]. Tiene un gran dominio extracelular en el cual se encuentra el extremo amino, un solo pase transmembranal y un pequeño dominio citoplasmático (Figura 1). La secuenciación del primer cDNA de BG clonado, el de rata, reveló que tiene dos regiones en la porción extracelular, denominados U y E por su similitud con Uromodulina y Endoglina respectivamente [2] (Figura 2). El pequeño dominio citoplasmático no tiene motivos funcionales reconocibles aunque tiene 12 Serinas que potencialmente se pueden fosforilar. De hecho, un estudio reciente ha demostrado la fosforilación *in vitro* de esta cola citoplasmática por la cinasa del receptor tipo II, aunque la relevancia fisiológica de este fenómeno aún está por demostrarse [3]. En total, el BG tiene 17 Cisteínas las cuales tienen el potencial de formar puentes disulfuro intra o intermoleculares. El tipo de glicosaminoglicanos (GAGs) que posee son cadenas de heparán y condroitín sulfato; las primeras tienen además la capacidad de unir el TGF ("Fibroblast Growth Factor"). Estas cadenas se enlazan covalentemente a las Serinas de los dipéptidos Ser-Gly flanqueados por residuos ácidos: en la rata Ser<sup>535</sup>-Gly y Ser<sup>546</sup>-Gly [2]

La cadena polipeptídica de BG, desprovista totalmente de carbohidratos, tiene una Mr de aproximadamente 110 kD [4]. Los carbohidratos unidos a Asparagina contribuyen aproximadamente con 20 kD a su masa, mientras que los glicosaminoglicanos hacen que su masa sea heterogénea y muy por arriba de los 300 kD. En ensayos de marcado por afinidad al TGF- $\beta$ , el BG migra como proteína medular de aproximadamente 130 kD (llamado "core protein") que tiene carbohidratos unidos a Asparagina, y como proteoglicano, el cual se exhibe como un barrido desde 300 kD en geles de poliacrilamida debido a la heterogénea glicosilación de la cual es objeto esta proteína, como se aprecia en la figura 3.

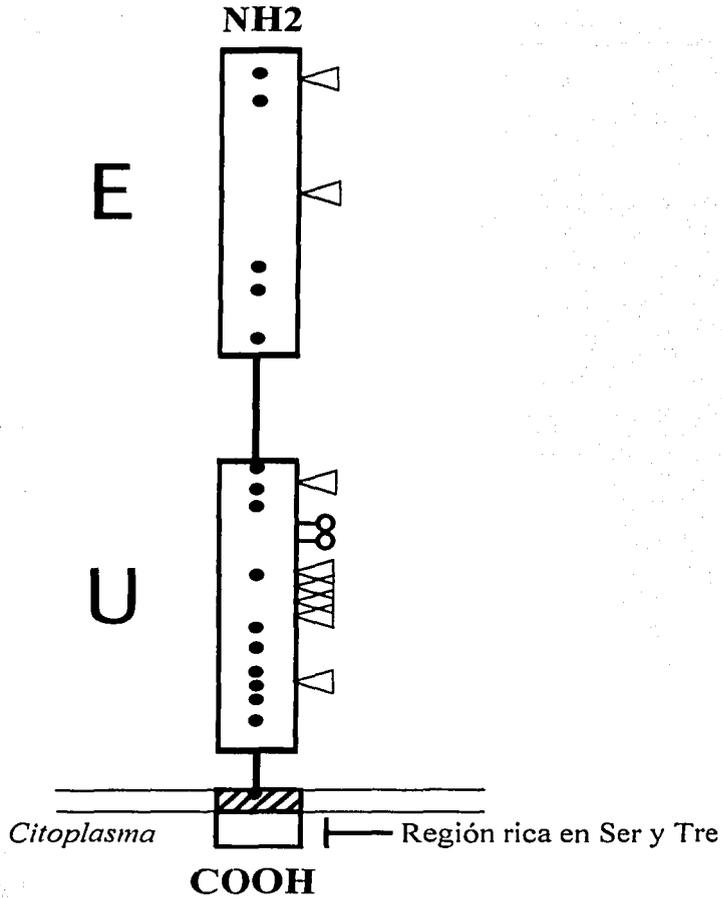


Figura 1. Esquema de la estructura de Betaglicano. Los círculos cerrados indican (●) Cisteínas, los círculos abiertos (○) los lugares donde se unen los GAGs y los triángulos (◁) las Asparaginas que unen azúcares.

TESIS CON  
FALLA DE ORIGEN

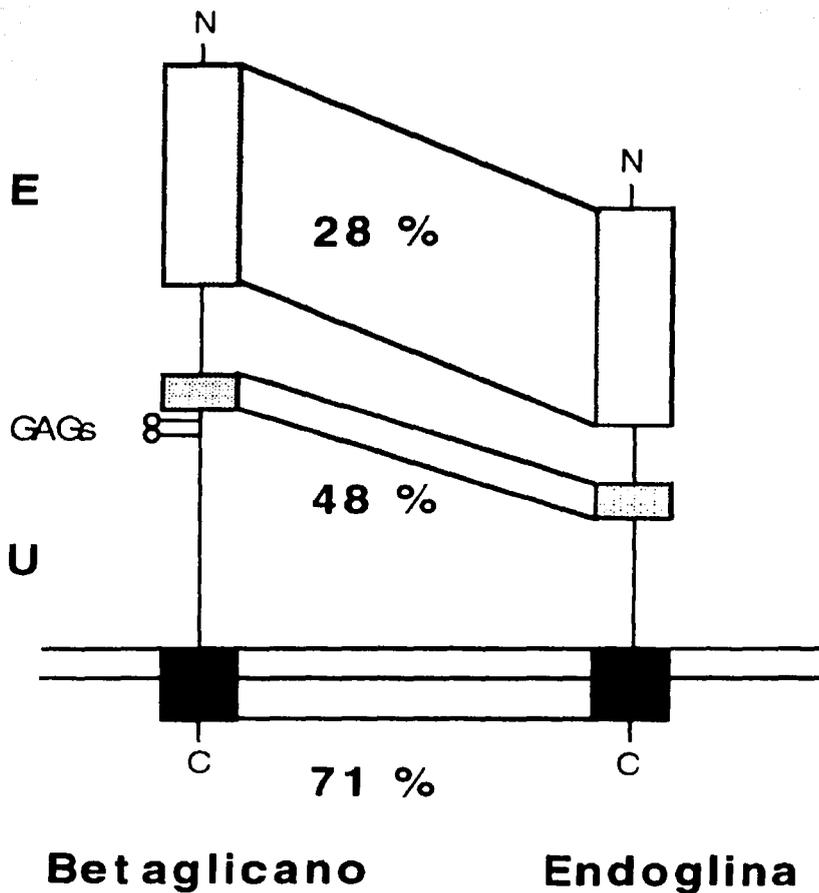


Figura 2. Regiones estructurales de similitud entre Betaglicano y Endoglina a nivel de secuencia nucleotídica.

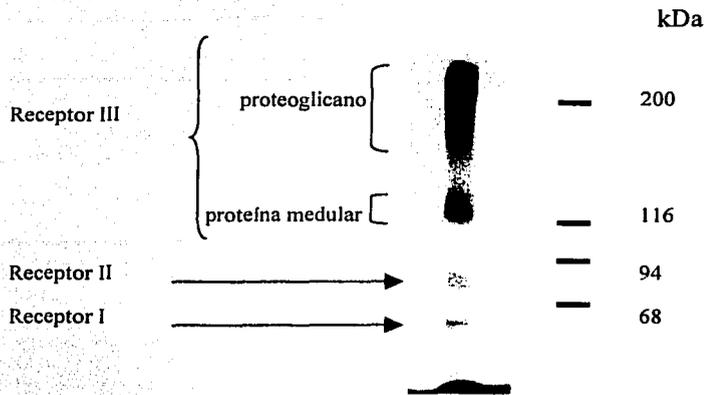


Figura 3. Células mink (epiteliales de pulmón) cultivadas en monocapa se marcaron por afinidad al  $^{125}\text{I}$ -TGF- $\beta$ . Las células se solubilizaron y el extracto total se corrió en un PAGE/SDS. La señal fue detectada por autoradiografía.

La clonación del cDNA de Betaglicano reveló que está constituido en la rata por 853 aminoácidos [2], en el cerdo por 848, en humanos por 849 [5] y en el pollo por 841 [6]. El grado de similitud que existe entre las secuencias del BG de estas especies se ilustra en el dendrograma de la Figura 4. La búsqueda por similitud de este gene en otras especies, reportó que no está presente en *Saccharomyces cerevisiae*, ni en *Caenorabditis elegans*, ni en *Drosophila melanogaster*, lo que sugiere que se trata de un gene que probablemente haya aparecido en vertebrados. El interés de usar tecnología transgénica para investigar otras funciones tanto de la proteína completa como de las cadenas GAG en el modelo murino, definió la necesidad de clonar el cDNA de ratón.

### Funciones del Betaglicano

El receptor III o Betaglicano de aquí en adelante, es un proteoglicano que forma parte del sistema de receptores del TGF- $\beta$  ("Transforming Growth Factor  $\beta$ "). Este sistema incluye a los receptores I, II, Betaglicano y Endogлина [7] (Figura 3).

TESIS CON  
FALLA DE ORIGEN

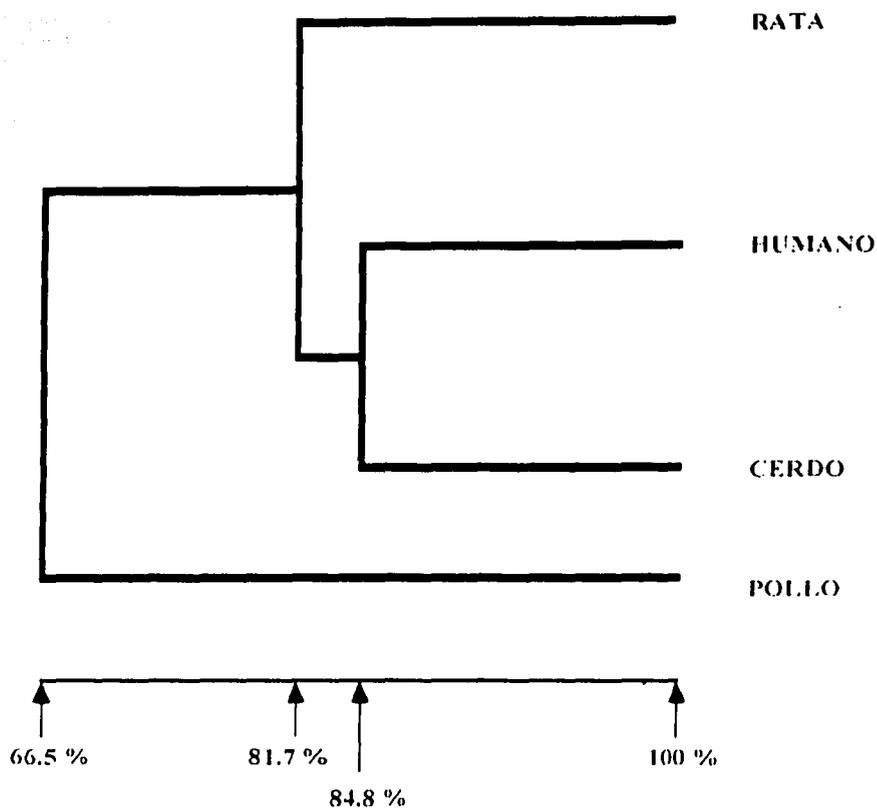


Figura 4. Dendrograma de las secuencias de Betaglicano. La barra inferior indica el porcentaje de similitud entre las secuencias proteicas incluidas en cada "rama". Elaborado con el programa GCG Wisconsin Package, versión 9.1. Genetic Computer Group, Madison Wisconsin.

Los receptores I y II son también glicoproteínas integrantes de membrana con un solo pase transmembranal, y con una movilidad relativa de 53 y 83 kDa respectivamente. En contraste con el Betaglicano, los receptores I y II tienen pequeñas regiones extracelulares y en la región intracelular cada uno tiene un dominio catalítico. Estos dominios son cinasas de Serinas y Treoninas [8][9]. El receptor II por su parte, tiene una afinidad por TGF- $\beta$  menor que el Betaglicano, mientras que el receptor I por sí mismo no une TGF- $\beta$ . Sin embargo los receptores I y II juntos, forman un sitio de alta afinidad por el TGF- $\beta$ .

La señalización del TGF- $\beta$  al interior celular se inicia cuando el TGF- $\beta$  encuentra y se asocia primero al Betaglicano en la superficie de las células. La unión del TGF- $\beta$  ocurre con la proteína medular y no con los GAGs del Betaglicano. Esta unión crea un efecto concentrador focalizado que facilita el acceso o "presentación" del TGF- $\beta$  al receptor II [10][11]. El establecimiento del complejo trimérico TGF- $\beta$ -Betaglicano-Receptor II facilita la interacción del TGF- $\beta$  con el receptor tipo II, y promueve el establecimiento de un complejo con el receptor tipo I. Entonces, la cinasa del receptor II que es constitutivamente activa, fosforila residuos en la región "GS", una región reguladora que precede a la cinasa del receptor tipo I ocasionando con ello la activación de su cinasa [12][13].

La señal o el mensaje del TGF- $\beta$  es propagado a través del citoplasma por mediadores intracelulares que forman una novedosa familia de proteínas conocidas como Smad, ver Figura 5. Un tipo de miembros de esta familia, los R-Smad, son los sustratos de fosforilación del receptor I. Cuando un R-Smad es fosforilado por el receptor I, se favorece la heterodimerización de R-smad con otro tipo de miembro de la familia Smad: los Co-Smad. Así heterodimerizados, R-Smad y Co-Smad se translocan al núcleo en donde funcionan junto con los factores basales de la maquinaria de transcripción, como activadores o represores de la transcripción regulando los genes blanco que responden al TGF- $\beta$ .

La modificación de la expresión transcripcional de los genes blanco del TGF- $\beta$ , constituye el inicio de la respuesta celular que induce esta citocina [14][15][16] ver Figura 5. Para la vía de transducción del TGF- $\beta$  los Smad2 y Smad3 funcionan como R-Smad, mientras que Smad 4 (también llamado DCP) funciona como Co-Smad [17]. El número de genes que responden a esta citocina depende de un número de factores como lo son el estado de la cromatina, el contexto transcripcional en que se recibe la señal [18] esto es, el conjunto de represores y activadores con que cuenta la célula en cuestión, los promotores accesibles a la maquinaria transcripcional, y la integración de señales de otras citocinas que el núcleo debe considerar [19]; de ahí que el tipo de respuesta al TGF- $\beta$  sea tan compleja y tan dependiente del tejido que está recibiendo la señal.

Además de crear el efecto de "presentación" del TGF- $\beta$ , explicado anteriormente, el betaglicano tiene distintas afinidades por cada una de las 3 isoformas del TGF- $\beta$  presentes en mamíferos. El Betaglicano tiene mayor afinidad por la isoforma del TGF- $\beta$  2 como se muestra en la Figura 6. Por lo que se ha considerado como el receptor específico del TGF- $\beta$  2 [11].

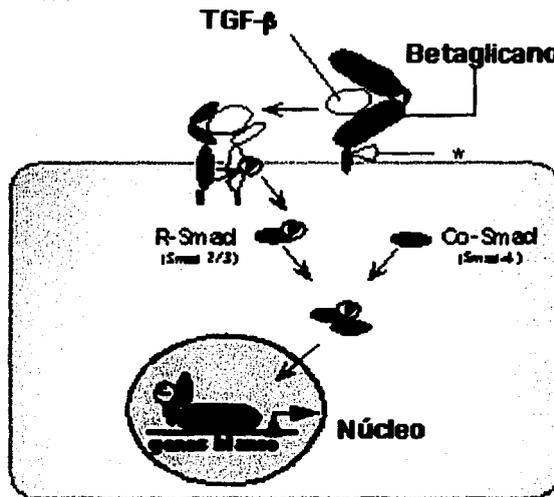


Figura 5. Cuando el TGF-beta une al Betaglicano, se crea un efecto concentrador que facilita el encuentro y la unión del TGF-beta con los receptores I y II, quienes transducen la señal al interior celular fosforilando a R-Smad. Esta fosforilación favorece la unión R-Smad /Co-Smad, juntos se translocan al núcleo donde regulan la transcripción. \* Indica el corte proteolítico del cual se deriva la forma soluble del Betaglicano

TESIS CON  
FALLA DE ORIGEN

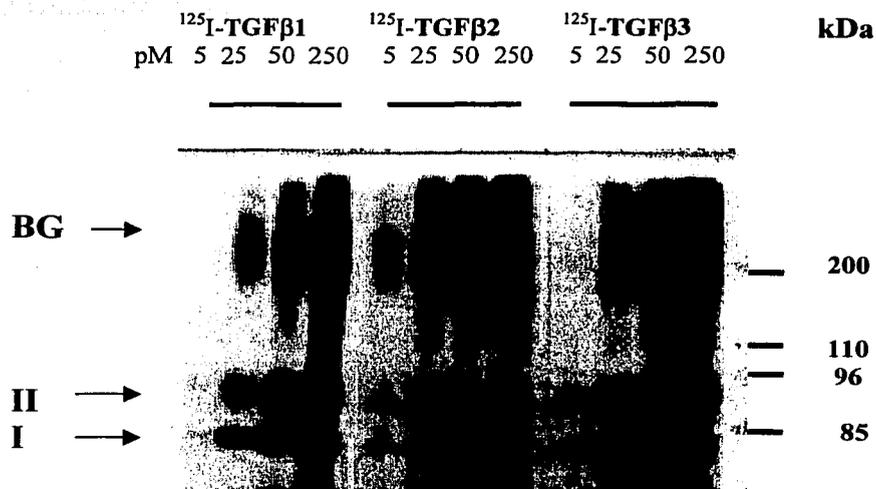


Figura 6. Marcado por afinidad de células en monocapa con 3 distintas isoformas del TGF $\beta$ . Con TGF $\beta_2$ , la saturación del Betaglicano se logra con 25 pM, mientras que la saturación con TGF- $\beta$  1 y 3 se logra con casi 250 pM.

### **Betaglicano Soluble**

Existe además de la forma transmembranal, una forma soluble del Betaglicano identificada por inmunoprecipitación en medios de cultivo de células en monocapa [1][10]. Esta forma soluble mantiene su alta afinidad por el TGF- $\beta$  y se deriva de la forma transmembranal a través de un corte proteolítico en la región yuxtamembranal (Figura 5). Estas observaciones han llevado a proponer que las dos formas del Betaglicano, la membranal y la soluble constituyen un interruptor molecular del TGF- $\beta$ . Este modelo establece que la forma transmembranal funciona como un facilitador o agonista del TGF- $\beta$ , mientras que la forma soluble lo secuestra alejándolo de los receptores I y II funcionando así como su antagonista [10].

TESIS CON  
FALLA DE ORIGEN

## Otras funciones

El Betaglicano es considerado un co-receptor en el sistema de receptores de TGF- $\beta$ , ya que su presencia no es indispensable para transducir la señal al interior celular. Se ha demostrado por otra parte, que el Betaglicano funciona también como un co-receptor de inhibina, otro miembro de la superfamilia del TGF- $\beta$ . En esos hallazgos la presencia del Betaglicano confiere sensibilidad a inhibina a las líneas celulares que responden pobremente a esta hormona [20]. Otro hallazgo que demuestra la relevancia biológica de BG, es el fenotipo letal en etapas tempranas del desarrollo del ratón Knock Out de Betaglicano [21], lo cual hace pensar en otras funciones muy importantes para este receptor aún no identificadas. Sin embargo la relevancia de investigar al Betaglicano actualmente, se deriva por un lado del potencial papel regulador que el Betaglicano tiene sobre el TGF- $\beta$  y por otro lado del papel central que el exceso de TGF- $\beta$  tiene en un número importante de patologías humanas. Para establecer claramente la importancia del Betaglicano en el contexto del TGF- $\beta$  a continuación se mencionan brevemente los aspectos más relevantes de esta citocina.

## TGF- $\beta$

El TGF- $\beta$  es el prototipo de una super-familia de proteínas de secreción que regulan el crecimiento, el desarrollo y la diferenciación en diversos organismos incluyendo nemátodos, insectos, roedores y humanos [22][23]. Originalmente el TGF- $\beta$  fue caracterizado en sobrenadantes de células cancerosas en cultivo, como la fracción responsable de producir fenómenos de transformación maligna en otras células en cultivo, de allí su nombre: Factor de Crecimiento Transformante ("Transforming Growth Factor") [24][25]. En mamíferos se han clonado las isoformas TGF- $\beta$ 1, 2 y 3 [26]. Las secuencias peptídicas de estas isoformas presentan una homología entre 70 y 80 % [27][29]. Existen 9 Cisteínas idénticas, 6 de las cuales forman una estructura central rígida conocida como nudo de Cistinas [30]. Las formas biológicamente activas del TGF- $\beta$  son homodímeros enlazados por puentes disulfuro y las unidades monoméricas están constituidas por 112 aminoácidos [31] ver Figura 7. El TGF- $\beta$ 1 está presente en casi todos los tejidos, encontrándose altas concentraciones en plaquetas y hueso.

El TGF- $\beta$ 1 es secretado por una gran variedad de tipos celulares como un precursor de alto peso molecular conocido como LAP ("Latent Associated Peptide"). Las isoformas activas del TGF- $\beta$  se derivan a través de un corte proteolítico en el extremo carboxilo de estos precursores más grandes [32], aunque el mecanismo fisiológico de activación del TGF- $\beta$  no se conoce detalladamente. Estos precursores varían de tamaño según la isoforma que se trate. TGF- $\beta$ 1: 390 aminoácidos, TGF- $\beta$ 2: 412 aminoácidos, TGF- $\beta$  3: 412, TGF- $\beta$ 4 : 304 aminoácidos, TGF- $\beta$ 5 382 aminoácidos [33].

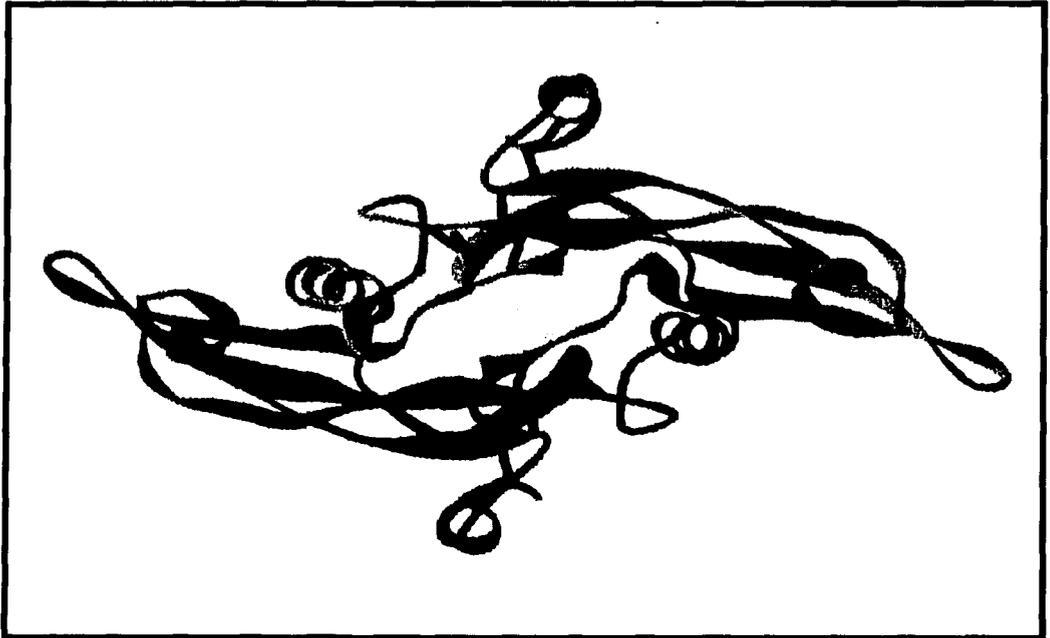


Figura 7. Estructura del TGF- $\beta$  en modelo de listones. Se enfatizan con anaranjado y azul los homodimeros que lo forman y los puentes disulfuro que estabilizan su unión.

## **Funciones del TGF- $\beta$**

Los estudios de los muchos y variados efectos del TGF- $\beta$ , tanto como los estudios de expresión de esta proteína, han demostrado que se trata de una citocina de secreción parácrina y autócrina. Está involucrada en la regulación de la diferenciación celular, en la morfogénesis durante el desarrollo embrionario, en la proliferación, en la reparación de heridas y en la respuesta inmune [34][39]. Debido al impacto médico que tienen todas estas funciones, particularmente en el campo oncológico, se está llevando a cabo un gran esfuerzo para elucidar sus mecanismos de acción y regulación. Su papel en la reparación de heridas se deriva de su capacidad de inducir deposición de matriz extracelular. Y resulta importante hacer notar que las acciones sobre la proliferación celular de esta citocina incluyen tanto la estimulación como la inhibición y esto depende del contexto transcripcional de la célula blanco sobre la que está actuando, así como de la presencia de otras moléculas con otros mensajes que esté recibiendo e integrando la célula [40][41]. Los fenotipos de los ratones en que han sido interrumpidos los genes correspondientes del TGF- $\beta$ , indican que las funciones de estas isoformas no son redundantes y son críticas para el desarrollo y viabilidad de los animales [42]-[44]. Estos hallazgos han sido consistentes con las acentuadas diferencias encontradas en las regiones promotoras de estos genes, lo que también ha demostrado que existe una regulación transcripcional espacio-temporal diferenciada entre las isoformas del TGF- $\beta$  a pesar de su similitud estructural [45]-[48].

Además de las funciones durante el desarrollo ya mencionadas, en tejidos adultos el exceso de TGF- $\beta$  ha sido identificado como un mediador central en diversos padecimientos fibróticos humanos como en la cicatrización queloide, la cirrosis hepática, la fibrosis pulmonar, la glomeruloesclerosis secundaria a diabetes y en otras nefropatías [49]-[51]. Consistentemente con estas observaciones, el fenotipo de los ratones transgénicos Alb- $\beta$ 1 que sobreexpresan la forma activa del TGF- $\beta$ 1 bajo el control del promotor de albúmina, presentan diversas lesiones fibróticas en hígado, en corazón y en páncreas, y lesiones particularmente severas en riñón [52]. Es importante destacar aquí que los ratones transgénicos Alb- $\beta$ 1, constituyen un modelo de experimentación muy adecuado en el cual se puede explorar *in vivo* la función antagonista del Betaglicano sobre el TGF- $\beta$ , como se discutirá más ampliamente en el último capítulo de esta tesis.

## **Estudio del gene de Betaglicano**

En resumen, existen varias razones por las cuales es importante estudiar el gene de Betaglicano. Como ya se mencionó en la sección Funciones de Betaglicano, a nivel de proteína se ha establecido que la versión soluble del Betaglicano mantiene su alta afinidad por el TGF- $\beta$ , creando un efecto antagonista en relación a la versión membranal y esto constituye un nivel de regulación del TGF- $\beta$  por el Betaglicano. Debido al potencial papel regulador que el Betaglicano tiene sobre esta citocina y al importante número de entidades patológicas humanas en que un exceso de TGF- $\beta$  está involucrado, resulta relevante identificar los mecanismos de regulación que a su vez tiene el Betaglicano.

Por otro lado existen razones que hacen pensar en otras funciones del Betaglicano aún no descubiertas, y un modo de explorar esas posibles funciones es estudiando el promotor del gene. Cuando, dónde y durante qué eventos se expresa este gene-proteína en organismos multicelulares, son preguntas que pueden ser contestadas experimentalmente si primero se tiene caracterizado bioquímica y molecularmente su promotor. La disponibilidad de estas estructuras y su empleo con metodología transgénica, pueden

aportar pistas muy informativas sobre otras funciones del Betaglicano. Existen ya evidencias de regulación transcripcional del Betaglicano [53][54], como son los estudios de hibridación *in situ* que demuestran la inducción del mRNA del Betaglicano en transiciones epitelio-mesenquimatosas, durante el desarrollo del paladar en ratón y durante la formación de las válvulas del corazón en el embrión de pollo [55]. También se ha establecido en nuestro laboratorio (datos por publicarse) que el mensajero del Betaglicano se induce durante el proceso de diferenciación de mioblastos C2C12 a miotubos, lo que constituye un sistema indispensable y al alcance en el cual se pueden estudiar los mecanismos de regulación transcripcional del Betaglicano.

Cabe señalar además que es muy probable que otras funciones del Betaglicano, probablemente estén relacionadas con los diferentes módulos o regiones estructurales de la proteína. Para probar estas ideas es posible usando tecnología transgénica en el modelo murino, generar deleciones modulares o mutaciones en regiones específicas del gene de Betaglicano para correlacionar el fenotipo obtenido con la estructura y la función correspondiente. Sin embargo para llevar a cabo esta estrategia es necesario primero conocer la estructura del gene normal y el tipo de relaciones que puedan existir entre las secuencias exónicas y la secuencia peptídica.

### **Clonación del cDNA murino de Betaglicano**

Dentro de este contexto, el primer objetivo en esta tesis doctoral fue clonar el cDNA murino de Betaglicano para poder continuar con la clonación de las secuencias genómicas correspondientes. Una vez clonado el mensajero murino y habiendo probado la expresión y los sitios de unión de los glicosaminoglicanos con mutagénesis dirigida, se procedió al tamizaje comercial de una genoteca murina de cósmidos de la compañía Research Genetics usando como sonda el cDNA obtenido de ratón. En el primer tamizaje se obtuvo un cósmido con aproximadamente 150 kpb en el cual no se encontraron los primeros exones. Considerando por un lado, que los intrones de este gene son muy largos lo que técnicamente complica el mapeo y localización de las regiones reguladoras y por otro lado, que el proyecto del genoma humano logró definir la región en la que se encuentra el Betaglicano en el cromosoma 1, decidimos aprovechar el esfuerzo de mapeo logrado en el cromosoma 1 e identificar las regiones del promotor humano, para lo cual fue necesario caracterizar el extremo 5' no traducido del mRNA humano de Betaglicano.

### **Caracterización del extremo 5' del mRNA de Betaglicano**

A continuación en esta tesis doctoral se reporta la caracterización del extremo 5' no traducido del mensajero humano, usando el método RACE (Rapid Amplification of cDNA Ends). La secuencias obtenidas y la comparación de estas secuencias contra el genoma humano nos permitió ubicar el primer exón y la región del promotor humano del Betaglicano.

### **Identificación del Promotor**

Habiendo identificado las secuencias del primer exón y la distribución exónica completa en el genoma humano, procedimos a diseñar una estrategia de amplificación por PCR para obtener un fragmento de 2.1 kpb corriente arriba de la unidad transcripcional del Betaglicano humano. Una vez amplificado el fragmento a partir de ADN genómico humano, se procedió a la clonación, secuenciación y análisis de las secuencias obtenidas.

### **Ensayos funcionales**

Con el fin de probar funcionalmente que el fragmento amplificado de 2.1 kpb contiene el promotor del Betaglicano, el fragmento fué subclonado en el vector de expresión de la serie pGL3-Basic que tiene como gene reportero la enzima luciferasa de la luciérnaga. Esta construcción se transfectó en células C2C12 que son mioblastos de ratón y en los cuales durante condiciones de diferenciación a miotubos se induce la acumulación del mRNA del Betaglicano. Con estos ensayos detectamos una inducción en la actividad de luciferasa 10 veces mayor, comparada con las células transfectadas con el vector control. Estos ensayos demostraron la actividad promotora contenida en el fragmento de 2.1 kpb, así como su regulación durante la miogénesis.

### **Construcciones Alb-Betaglicano**

La caracterización del cDNA murino nos ha permitido iniciar la manipulación genética del Betaglicano en el ratón. Considerando el fenotipo de los animales transgénicos Alb- $\beta$ 1 mencionado anteriormente, y su utilidad como modelo en el cual estudiar los efectos *in vivo* del Betaglicano, en esta tesis doctoral también se reporta la construcción de dos transgenes de Betaglicano bajo el control del promotor de albúmina así como las pruebas de su funcionalidad *in vitro* mediante ensayos de transfección transitoria. Los objetivos inmediatos con la generación de estos animales, son por una parte revertir el fenotipo de los animales Alb- $\beta$ 1 al cruzarlos con el ratón transgénico que sobreexpresa el Betaglicano soluble. Y por otra parte, probar el efecto antagónico del Betaglicano soluble sobre el TGF- $\beta$  en un modelo de inflamación y fibrosis crónica mediado por TGF- $\beta$ , como lo es la fibrosis hepática producida por tetracloruro de carbono (CCl<sub>4</sub>).

## II. PLANTEAMIENTO DEL PROBLEMA E HIPÓTESIS

### Planteamiento del problema

En contraste con los receptores I y II que poseen dominios catalíticos lo que los involucra funcionalmente en la vía de señalización del TGF- $\beta$ , los motivos estructurales del Betaglicano no revelaron una función bioquímica obvia. El hecho de que Betaglicano puede unir tres distintos factores de crecimientos, TGF- $\beta$ , Inhibina y FGF habla más sobre nuestro desconocimiento que su real utilidad para el organismo. La función propuesta para el Betaglicano es la de una molécula antagónica del TGF- $\beta$  en su forma soluble, y la de una molécula agonista en su forma transmembranal. Por otro lado el fenotipo letal de los ratones Knock Out del Betaglicano sugiere fuertemente que el Betaglicano tiene funciones aún no descubiertas muy importantes durante el desarrollo.

El abordaje genético, es decir el conocimiento de la estructuración del gene, de sus regiones reguladoras y de los patrones de expresión, aportarán información y herramientas valiosas para explorar estas funciones desconocidas.

Existen evidencias de inducción del mRNA del Betaglicano en un ensayo de diferenciación de mioblastos a miotubos, por lo que estratégicamente la identificación del promotor de Betaglicano constituye el paso inicial crítico, para abordar el estudio genético del Betaglicano. La clonación del promotor abrirá dos grandes avenidas de investigación: por un lado el estudio de la expresión diferencial de este gene en condiciones normales durante el desarrollo, y por otro lado la definición y el estudio de la regulación transcripcional del gene. Estas dos avenidas de investigación aportarán valiosas pistas sobre los procesos biológicos en que el Betaglicano participa y las funciones bioquímicas que esta proteína tiene.

En este contexto, el problema central propuesto en esta tesis, consiste en la identificación y comprobación funcional del promotor del gene de Betaglicano.

### Hipótesis

Si las secuencias genómicas identificadas con el potencial promotor lo son, entonces inducirán significativamente la expresión de un gene reportero en el ensayo de diferenciación a miotubos de células C2C12, en el cual sabemos que existe un aumento del mRNA de Betaglicano endógeno.

### III. OBJETIVOS Y ABORDAJE EXPERIMENTAL

El modelo mamífero sobre el cual se pueden hacer manipulaciones genéticas es el murino por lo que el primer objetivo de esta tesis lo constituyó la clonación del cDNA de Betaglicano de ratón. El siguiente objetivo como un paso intermedio para poder identificar el promotor, consistió en caracterizar el extremo 5' no traducido del mensajero de Betaglicano. Una vez identificado el primer exón del gene, procedimos a la amplificación y clonación de secuencias genómicas que potencialmente contienen la región promotora. A continuación demostramos la funcionalidad del promotor, induciendo la expresión de un gene reportero en ensayos de transfección transitoria. Por último objetivo en esta tesis, se planteó la construcción y el establecimiento de pruebas funcionales *in vitro* de dos transgenes del Betaglicano: una versión soluble y otra versión membranaral bajo el control del promotor murino de albúmina. Estas construcciones serán usadas en la generación de animales transgénicos, los cuales permitirán el estudio de las funciones *in vivo* del Betaglicano.

#### Objetivos Específicos

- III.1 Clonación y caracterización del cDNA del Betaglicano murino.
- III.2 Caracterización del extremo 5' UTR del mRNA humano.
- III.3 Localización del primer exón y las secuencias reguladoras en el cromosoma 1.
- III.4 Establecimiento de la distribución exónica del gene de Betaglicano humano.
- III.5 Identificación y clonación por PCR de secuencias genómicas corriente arriba de la unidad transcripcional.
- III.6 Transfección y cuantificación de la actividad transcripcional del fragmento genómico obtenido.
- III.7 Construcción de Betaglicano soluble y membranaral bajo el control del promotor de albúmina murino.

#### IV. MATERIAL Y MÉTODOS

##### **Clonación del cDNA murino**

Se llevaron a cabo 3 rondas de tamizaje primario de una biblioteca de cDNA en bacteriófago  $\lambda$ gt11 de corazón murino (Clontech LM5002b), se usó como sonda la región codificadora del cDNA de Betaglicano de rata [2]. La sonda se marcó con sistema de oligómeros aleatorios (Random Primed DNA Labeling Kit, Boehringer Mannheim, Cat. No. 1004 760), el isótopo radioactivo que se usó para marcar la sonda fue  $[\alpha\text{-}^{32}\text{P}]$  dCTP (370MBq/ml 10mCi/ml  $\sim$ 3000Ci/mmol Cat. No. AA0005 Amersham Pharmacia Biotec). Las sondas se marcaron con una actividad específica promedio de  $1 \times 10^8$  cpm/ $\mu$ g. Se usaron diversas enzimas de restricción y enzimas de modificación (Roche Molecular Biochemicals) para purificar insertos y subclonarlos. Las sales y otros reactivos para propagar los bacteriófagos son de Sigma Aldrich / Fisher. El inserto se subclonó inicialmente en pBluescript. Se utilizó el sistema para generar fragmentaciones progresivas con exonucleasa III (Erase-A-Base System, Promega, Cat. No. E5750), para facilitar la secuenciación de los insertos >3 Kb. La secuenciación de los insertos se llevó a cabo manualmente con Thermosequenase y ddNT- $^{33}\text{P}$  (Radiolabeled Terminator Cycle Sequencing Kit, Cat. No. P/N 79750,  $5'[\alpha\text{-}^{33}\text{P}]\text{ddNTP}$ , Cat. No. AH9539 Amersham). El inserto se subclonó en el vector pcDNA3 que tiene el promotor del citomegalovirus.

##### **Transfección Transitoria**

Las células COS1 en monocapa se transfectaron por el método dietilaminoetil-dextrán (DEAE-dextrán) [47] y fueron ensayadas con marcado por afinidad 48 hrs después de la transfección.

##### **Mutagénesis de Serina 533 y 544**

Con la finalidad de comprobar que las Serinas 533 y 544 corresponden a los sitios en donde se unen los GAGs, decidimos mutar los codones correspondientes para lo cual: Se diseñaron 4 oligonucleótidos (1027, 1028, 1029 y 1030) con las secuencias necesarias para cambiar codones AGT por GCT, y GTC por GGC. Se utilizaron además los oligonucleótidos HP4 y MN1 que flanquean la región de interés. El fragmento comprendido entre oligonucleótidos MN1 y HP4 es de 939 nt. Los productos de amplificación HP4-1028 (539 nt) y MN1-1027 (419 nt) así como HP4-1030 (571 nt) y MN1-1029 (386 nt) sirvieron como moldes con las mutaciones incorporadas para la amplificación del fragmento HP4 y MN1. Dentro del fragmento HP4 y MN1 se encuentran los sitios de restricción EcoRV y Pst I que se usaron para extraer el segmento mutado y reemplazarlo en el cDNA original. Para las reacciones de amplificación se utilizó el termociclador GeneAmp PCR System, Perking Elmers. La mutación de cada Serina se hizo por separado y para obtener la doble mutante, se generó una segunda mutagénesis en la mutante Serina 533, usando la misma estrategia. Las mutantes se secuenciaron y se transfectaron por el método de fosfato de calcio en células COS1 que fueron cultivadas en DMEM con 10 % Suero Fetal Bovino. 48 hr después las células transfectadas se ensayaron con marcado por afinidad al  $^{125}\text{I}$ -TGF- $\beta$ . El TGF- $\beta$  (TGF- $\beta$ 1 se obtuvo de RyD Systems, TGF- $\beta$ 2 se obtuvo de Ciba Geigy AG.) se marcó con el método de cloramina T y Disuccinimidyl-suberato (Pierce Chemicals).

### **Caracterización del extremo 5' del mRNA**

Se extrajo RNA total de células HepG2 con Ultraspec RNA system, Cat. No. BL-10050, BIOTECX Laboratories, Inc. El 5'UTR se obtuvo con el Sistema de Amplificación Rápida de Extremos de cDNA (GeneRacer Kit Version B, Cat. No. L1500-01 Invitrogen) y los fragmentos de amplificación se clonaron con el Sistema de Clonación TOPO (TOPO TA Cloning for Sequencing Version E2, Cat. No. K4575-01, Invitrogen). Los productos amplificados se secuenciaron con Thermosequenase y ddNT-<sup>33</sup>P (Radiolabeled Terminator Cycle Sequencing Kit, Cat. No. P/N 79750, 5' [ $\alpha$ -<sup>33</sup>P]ddNTP], Cat.No. AH9539 Amersham).

### **Amplificación de los fragmentos genómicos A, B, C, D**

Se usaron como molde 100 ng de DNA extraído de sangre humana total periférica. Con base en las secuencias del proyecto del genoma humano se diseñaron y usaron los siguientes oligonucleótidos 2417 = 5' atggtaaccgagaaacagcacaggaaaaa 3', 2418 = 5' atggtacctgctgacctagctccattattg 3', 2419 = 5' catcgcttcctcactcgctgggaagag 3', y 2420 = cataagcttgcaagtgaggaaagc 3', a una concentración de 10  $\mu$ M, para amplificar la región corriente arriba del primer exón del gene de Betaglicano humano. Los oligonucleótidos diseñados, comprenden una región máxima de 2.1 kpb incluyendo 184 pb corriente abajo del supuesto sitio de inicio de la transcripción, en esta región esperamos encontrar los elementos centrales del promotor. Se establecieron las siguientes condiciones de amplificación: amortiguador 1x, dNTPs 0.2 mM cada uno, DMSO 5 %, BSA 0.5  $\mu$ g totales, MgCl<sub>2</sub> 0.5, 1.5 y 3mM, Taq de High Expand<sup>TM</sup> Fidelity (Cat. No. 1 732 650 Boehringer Mannheim). Utilizando el siguiente protocolo: se desnaturalizó el DNA en presencia de MgCl<sub>2</sub>, DMSO y BSA por 10 min a 94° C, posteriormente se añadieron los demás componentes de la reacción y se llevaron a cabo 30 ciclos con los siguientes parámetros: 94°C 1 min, 40, 50 ó 60° C 45 seg., 72°C 3 min, seguidos de una extensión final a 72°C por 10 min. Se utilizó el termociclador de gradiente Master Cycler Gradient, Eppendorff. Los fragmentos amplificados se clonaron en el vector de clonación pBSKS de Stratagene. La identidad de los fragmentos amplificados se confirmó por secuenciación.

### **Ensayos funcionales del fragmento de 2.1 Kb**

Una vez confirmada la secuencia de los fragmentos amplificados y clonados en pBluescript, el fragmento de 2186 pb se subclonó en el vector pGL3-Basic (pGeneLight<sup>TM</sup> Promega) que tiene el gene reportador de la luciferasa de luciérnaga. La construcción se transformó en células C2C12 que para proliferar se mantuvieron en DMEM F12 (y suplementos: alto en glucosa, Suero Fetal Bovino 10 %, Extracto de Embrión de Pollo 0.5 %), y para diferenciarlas en miotubos se cambiaron los suplementos por Suero de Caballo 10 %. La actividad transcripcional del fragmento genómico clonado se detectó usando los sistemas de detección de luciferasa (Luciferase Assay System Cat. No. E1501 y Dual-Luciferase Reporter Assay System<sup>TM</sup> Cat.No, 040 Promega), y un luminómetro Lumat LB9501.

### **Construcción y ensayos de expresión de transgenes Alb-Betaglicano**

Se utilizó como base de las construcciones el vector pAlb c/p donado amablemente por el Dr. Richard D. Palmiter el cual contiene varios sitios de restricción únicos justo corriente abajo del promotor de albúmina lo que convenientemente facilitó la inserción de los diferentes componentes. También se utilizó el vector Alb- $\beta$ 1 donado amablemente por la Dra. Anita Roberts [63]. Del plásmido Alb- $\beta$ 1, obtuvimos el extremo 3'UTR de la hormona humana de crecimiento (hGH) con las enzimas Sal I – Eco RI que fué subclonado en pBluescript. Este extremo 3'UTR contiene la señal de poliadenilación del gene de la hormona de crecimiento y su empleo se recomienda para estabilizar y aumentar la vida media de los mRNAs. Sobre el plásmido Alb c/p se insertó en un sitio EcoRV único, el fragmento hGH. En el sitio único Bam HI ubicado entre el promotor y el segmento hGH se introdujeron la versión soluble o transmembranal de Betaglicano, con la finalidad de inducir niveles altos de expresión (comparables a los de albúmina) de una u otra forma de Betaglicano. Una vez confirmadas las construcciones por secuenciación, se transfectaron en células HepG2 que se cultivaron en MEM con 10 % Suero Fetal Bovino. 24 hr después de la transfección las células se marcaron por afinidad al  $^{125}$ I-TGF- $\beta$ 1. Se realizó inmunoprecipitación de los sobrenadantes de las células transfectadas con la versión soluble del Betaglicano con anticuerpo anti-myc (monoclonal de Sta Cruz, Ca), y de los extractos totales con el anticuerpo #822 (policlonal de conejo) contra el ectodominio del Betaglicano. Los extractos totales y los inmunoprecipitados se sometieron a PAGE/SDS. Se utilizó el sistema de revelado por quimioluminiscencia (ECL Western blotting detection systems, RPN 2106, Amersham Pharmacia Biotech.). Para la detección de señales radiactivas en matrices sólidas se usó el sistema Storm , Molecular Dynamics.

### **Programas utilizados para el análisis de secuencias**

Para la determinación de la distribución exónica y la localización genómica del primer exón de Betaglicano humano se usó BLASTnr, del NCBI <http://www.ncbi.blast.human/genome/> para análisis comparativos se utilizó MacVector Version 7 y GCG Wisconsin Package, versión 9.1 Genetic Computer Group, Madison Wisconsin, USA. Para el análisis de elementos funcionales en el extremo 5' no traducido: se usó el programa UTRScan en: <http://bigarea.area.ba.cnr.it:8000/EmblIT/UTRHome> Para el análisis de las secuencias del promotor se utilizó: MatInspector, AliBaba2.1 y la base de datos de TransFac, Software GmbH en: <http://genomatix.gsf.de>. Para la búsqueda de promotores humanos caracterizados experimentalmente por la determinación del inicio de la transcripción se utilizó la Eukaryotic Promotor Database en: <http://www.epd.isb-sib.ch/>

Para la síntesis de oligonucleótidos, secuenciación automática, sistemas de registro de imágenes, y equipo de amplificación se contó con el apoyo de la Unidad de Biología Molecular del Instituto de Fisiología Celular UNAM.

En los capítulos V - IX se presentan resultados específicos y metodologías pertinentes de cada objetivo.

## V. CLONACION DEL cDNA MURINO DE BETAGLICANO

### Abordaje Experimental

El cDNA de rata de Betaglicano [2] se utilizó como sonda y se tamizó una biblioteca de cDNA de corazón murino. Se obtuvieron varias clonas con insertos de diversos tamaños. El inserto más grande fue el de la clona  $\lambda$ CII-2 con 3977 pb y fue subclonado en pBluescript (Stratagene) para ser secuenciado. La secuenciación reveló que el inserto comprende toda la región codificadora del Betaglicano en 2550 pb, incluyendo además 230 pb del extremo 5' no traducido (5'UTR) así como 1192 pb del extremo 3' no traducido (3'UTR). Para comprobar que el cDNA clonado contiene un Betaglicano con actividad biológica, el inserto de 3977 pb fue subclonado en el vector de expresión pcDNA3. Esta construcción tanto como el vector vacío se transfectaron transitoriamente en células COS1. Las células transfectadas fueron ensayadas con marcado por afinidad al  $^{125}\text{I}$ -TGF- $\beta$ 1 48 hr después de la transfección. Posteriormente las células se lisaron y usando el anticuerpo policlonal #822 de conejo generado contra el ectodominio de Betaglicano de rata, estos lisados se inmunoprecipitaron y se separaron en geles de policarilamida / SDS [11]. Las señales se revelaron por autoradiografía.

### Secuenciación

La estrategia de secuenciación consistió en generar fragmentaciones progresivas sobre el inserto de la clona  $\lambda$ CII-2 con exonucleasa III (Erase-A-Base System, Promega, Cat. No. E5750) o con diversas enzimas de restricción y religando los fragmentos producidos. La secuenciación se llevó a cabo usando los oligonucleótidos T3 y T7 que flanquean a los insertos o con oligonucleótidos sintetizados para "caminar" sobre los fragmentos generados por las digestiones.

### Marcado por afinidad e inmunoprecipitación.

El TGF- $\beta$ 1 fue marcado radiactivamente con  $^{125}\text{I}$  usando cloramina T [48]. Brevemente: las células en monocapa (aproximadamente 80% de confluencia) se incubaron por 3 hr a 4°C con  $^{125}\text{I}$ -TGF- $\beta$  a una concentración de 50 pM, se entrecruzaron los receptores que unieron al ligando con DSS (disuccinimidil-suberato). Las células se solubilizaron en presencia de Triton X-100. Los extractos celulares resultantes fueron sujetos a electroforesis en geles de poliacrilamida y SDS (PAGE). La detección y análisis de la radioactividad se realizó con el sistema de detección STORM. Para la inmunoprecipitación los extractos celulares se diluyeron en amortiguador (25 mM Tris-HCl [pH 7.5] 150 mM NaCl, 2.5 mM EDTA, 1 % Nonident P-40). Se pre-absorbieron con 50  $\mu$ l de Sepharosa CL-4B por 10 minutos a 4° C, se centrifugaron brevemente para separar detritus. Los sobrenadantes se incubaron con el anticuerpo # 822 generado en conejo contra el ectodominio del Betaglicano de rata por un periodo de 2-4 hr a 4° C. Los inmunoprecipitados se recuperaron con una incubación de 60 min a 4° C con Proteína A-Sepharosa, se lavaron 3 veces con amortiguador PBS-Triton X-100 0.1 %- SDS 0.02 %, y se resuspendieron directamente en amortiguador de carga para someterse a electroforesis en gel de poliacrilamida al 7 % (Figura 8).

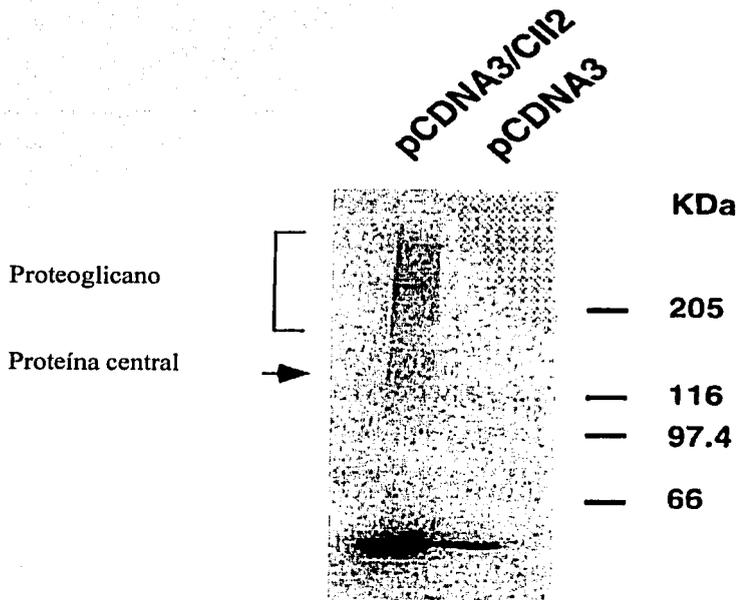


Figura 8. En esta autoradiografía se demuestra que el vector con el inserto (pcDNA3/CII2) condujo a la expresión de un barrido característico del Betaglicano por arriba de 200 kDa, junto con una proporción de la proteína central.

## Mutagénesis de Serinas 533 y 544

Para comprobar que las Ser 533 y 544 del cDNA murino están modificadas con GAGs, y por lo tanto son homólogos funcionales de las Ser 533 y 546 de rata, se realizó mutagénesis dirigida basada en PCR. Con el objetivo de cambiar los codones que codifican para Serina por codones que codifican para Alanina es decir AGT por GCT y GTC por GGC, se usaron oligonucleótidos sintéticos para introducir las mutaciones deseadas como se indica en la secuencia de la Figura 9:

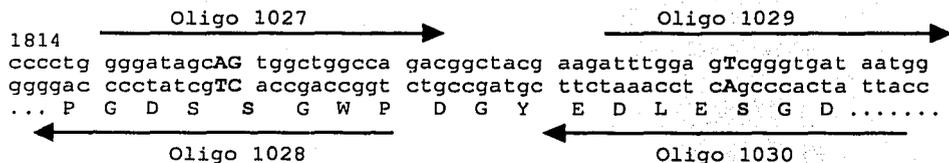


Figura 9. Con mayúsculas y sobresaltadas en negro se indican los nt que se modificaron durante la síntesis de los oligonucleótidos para introducir las mutaciones.

Las secuencias correspondientes son:

Oligo 1027	5' gggatagcGCtggetggccaga 3'	$t_m = 73^\circ \text{C}$
Oligo 1028	5' ggccagccaGCgctatccccagg 3'	$t_m = 75^\circ \text{C}$
Oligo 1029	5' gatttggagGcgggtgataatg 3'	$t_m = 61^\circ \text{C}$
Oligo 1030	5' tatcaccocgCctccaaatcttc 3'	$t_m = 61^\circ \text{C}$

Además se utilizaron los oligonucleótidos HP4 y MN1 que flanquean la zona de mutagénesis para amplificar los fragmentos de 939 pb y los sitios EcoRV y PstI, que se usaron para substituir los fragmentos amplificados por PCR en el cDNA original. La estrategia general se ilustra en el siguiente esquema:

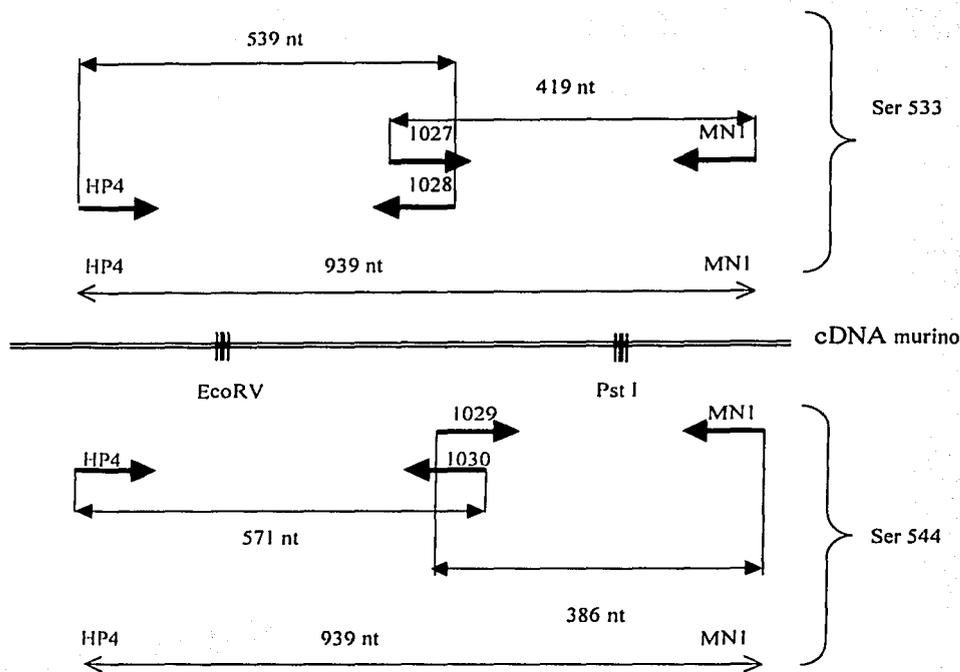


Figura 10. En la parte superior se ilustra la estrategia de mutagénesis dirigida para la Ser 533, y en la parte inferior la estrategia de mutagénesis para la Ser 544. Por ejemplo, los productos de 539 y 419 pb obtenidos por las PCRs: HP4 $\leftrightarrow$ 1028 y 1027  $\leftrightarrow$  MN1 respectivamente, sirvieron de molde para la amplificación final con HP4 $\leftrightarrow$ MN1, este fragmento amplificado se cortó con EcoRV y Pst I y el fragmento comprendido entre esos sitios de restricción se substituyó en el cDNA original, produciéndose así la versión mutante en Ser 533.

Las versiones mutantes se subclonaron en el vector pCMV5. Cada mutante se confirmó a través de secuenciación, y para obtener la doble mutante, sobre la mutante Ser<sup>533</sup>Ala se practicó la misma estrategia de mutagénesis empleada para la Ser<sup>544</sup>Ala.

#### Expresión del cDNA y de las versiones mutantes.

Para probar las tres versiones mutantes (Ser<sup>533</sup>Ala, Ser<sup>544</sup>Ala, Ser<sup>533,544</sup>Ala), los vectores construidos se transfectoron por el método DEAE-dextran en células COS1 y se marcaron por afinidad con <sup>125</sup>I-TGF2. Como se aprecia en la Figura 11, la versión silvestre

en pCMV5 muestra predominantemente el Betaglicano en la forma de proteoglicano. Las versiones mutantes muestran una disminución del proteoglicano y la doble mutante muestra una proteína central preponderante como un doblete entre 125-130 kDa versión proteica desprovista de proteoglicanos, aunque muy probablemente unida a otros azúcares.

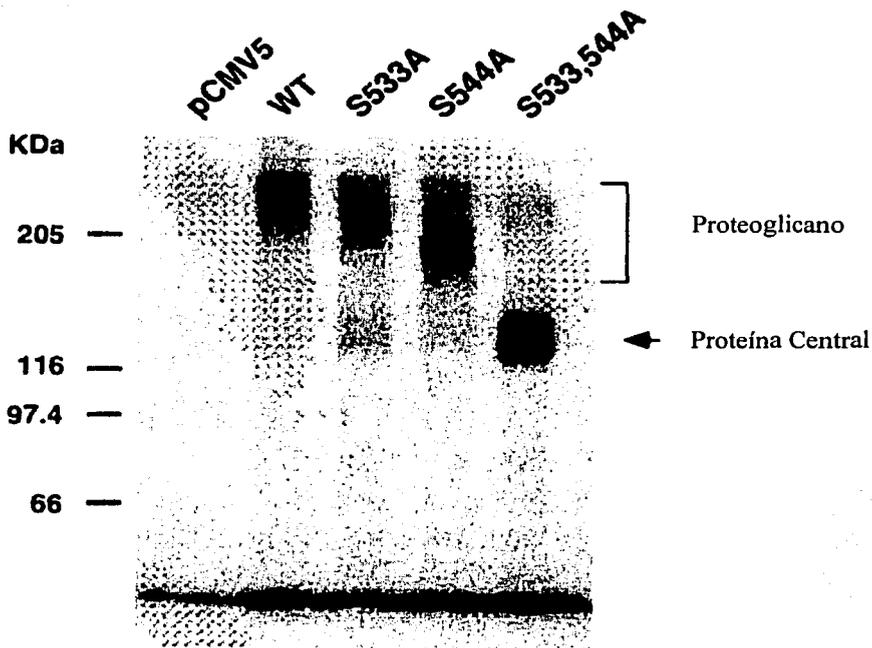


Figura 11. Células COS1 transfectadas con las versiones mutadas de Betaglicano, marcadas por afinidad con  $^{125}\text{I}$ -TGF- $\beta$ , solubilizadas y sometidas a PAGE-SDS. Tanto en el carril de la mutante S 533 A, como el de la mutante S 544 A se observa una disminución en el barrido característico del Betaglicano respecto a la forma silvestre. En el carril de la doble mutante S 533,544 A esta disminución se acentúa, con lo cual se confirma que estas Serinas son funcionalmente homólogas a las descritas en otras especies.

Con la secuencia del cDNA del Betaglicano de ratón nos fué posible crear un alineamiento múltiple nuevo, incluyendo las otras especies en las que se conoce la secuencia del Betaglicano y actualizar en un dendrograma el tipo de relaciones existentes entre las secuencias de las especies caracterizadas, Figura 12.

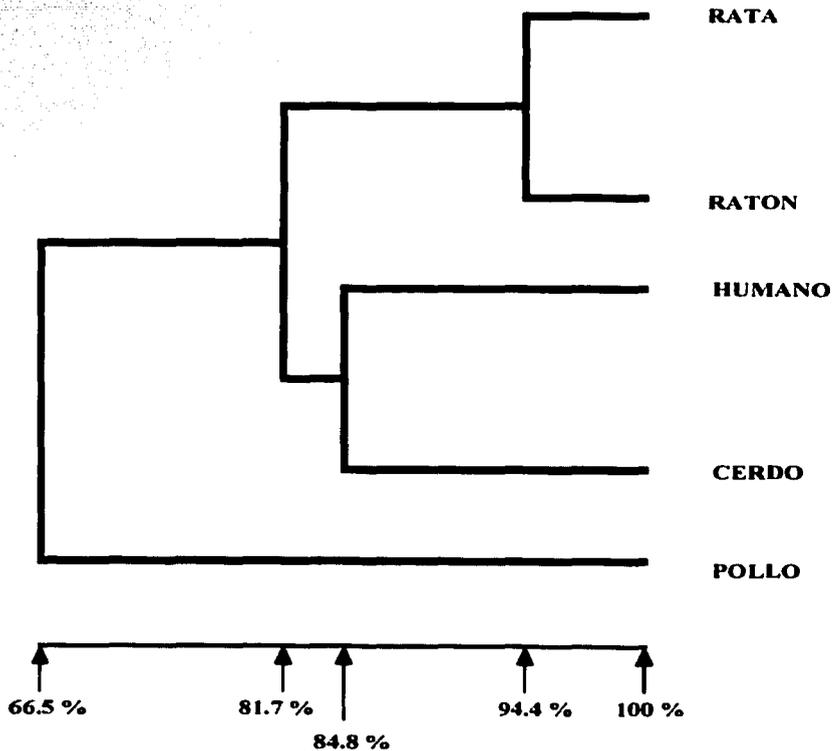


Figura 12. La longitud de las líneas horizontales permite saber en términos porcentuales el grado de similitud entre las secuencias o grupo de secuencias como lo muestra la barra inferior del dendrograma. Así, entre la secuencia de ratón y de rata existe un 94.4% de similitud. Resulta interesante comentar que desde el punto de vista comparativo, la secuencia humana de Betaglicano es más parecida a la secuencia del cerdo (84.8%) que a la del ratón (81.7%).

## **Comentario y Discusión**

Utilizando como sonda el cDNA de Betaglicano de rata, clonamos a partir de una biblioteca de cDNA de corazón murino, el cDNA de Betaglicano de ratón. La mutagénesis dirigida de las Serinas 533 y 544 nos permitió comprobar que a estos residuos se unen las cadenas de glicosaminoglicanos (GAGs). Además de ampliar nuestro conocimiento en términos filogenéticos sobre el Betaglicano, debido a la relevancia médica que esta proteína puede tener, la referencia en el ratón nos aporta criterios firmes sobre la validez o limitaciones que tendrá el abordaje de nuevos estudios del Betaglicano *in vivo* usando el modelo murino.

Los resultados presentados en este capítulo, han sido publicados en **BBA Biochimica et Biophysica Acta 1384**, *Murine betaglycan primary structure, expression and glycosaminoglycan* M. Verónica Ponce-Castañeda, José Esparza-López, M. Magdalena Vilchis-Landeros, Valentín Mendoza R., Fernando López-Casillas. La copia del artículo se encuentra al final de esta tesis.

## VI. CARACTERIZACION DEL EXTREMO 5' DEL mRNA HUMANO

### Abordaje Experimental

Con la clonación del cDNA murino nos dimos a la búsqueda de las secuencias genómicas reguladoras del Betaglicano, para lo cual recurrimos a los servicios de tamizaje de bibliotecas genómicas en BACs (Bacterial Artificial Chromosomes) de la compañía Molecular Genetics. Usando como sonda el inserto completo de la clona  $\lambda$ CII -2, el cual incluye además de la región codificadora 230 nt de la región 5' no traducida (5'UTR), se obtuvo de Molecular Genetics un BAC que a pesar de tener aproximadamente 150 Kb de inserto no contenía las secuencias del 5'UTR. Por otro lado la búsqueda de las secuencias genómicas humanas en una biblioteca genómica en  $\lambda$ gt10, nos había demostrado que los exones cercanos al inicio de la traducción están espaciados por intrones muy grandes (el exón que contiene el ATG que codifica la primera Medicinan, se encontró en un inserto de aproximadamente 10 kpb), lo cual también dificultaba la identificación de los primeros exones con esa estrategia. Mientras la compañía Molecular Genetics realizaba un segundo tamizaje en la biblioteca murina de BACs usando como sonda solamente los ~230 nt del 5'UTR del mensajero murino, la disponibilidad y el ordenamiento preliminar a través del proyecto del genoma humano de las secuencias en el cromosoma 1, permitió resolver el obstáculo de los grandes intrones con otro abordaje: procedimos a la identificación de los primeros exones del gene a través de la caracterización del extremo 5' UTR del mensajero humano. Para clonar las secuencias de los primeros exones del gene humano de Betaglicano, caracterizamos y clonamos el extremo 5' UTR del mensajero usando la metodología RLM-RACE ("T4 RNA Ligase Mediated Rapid Amplification of cDNA Ends"). Una vez obtenidos los fragmentos de amplificación, se subclonaron y se secuenciaron.

### Protocolo GeneRacer

La técnica se basa en la ligación de un oligonucleótido (RACE 5') que selectivamente se une a los extremos 5' de mRNA que han sido descasquetados, y en el empleo de un oligonucleótido específico del gene de interés (GSP3'), para transcribir con Transcriptasa Reversa el mensajero a cDNA, y posteriormente amplificar por PCR el segmento comprendido entre ambos oligonucleótidos.

### Purificación de RNA humano total

Usando células en cultivo HepG2 derivadas de un hepatoma humano y en donde se expresa el Betaglicano, se procedió a la extracción y purificación del RNA total usando el sistema de purificación Ultraspec RNA system, Cat. No. BL-10050, BIOTECHX Laboratories, Inc.

### Diseño del oligonucleótido gene específico

Los criterios para el diseño del Oligonucleótido Gene Específico del Betaglicano (GSP 3') son:

- Contenido GC de 50-70 % para obtener una temperatura de alineamiento alta,  $>70^{\circ}$  C.
- 23-28 nt de longitud para incrementar la especificidad de alineamiento.

- Secuencias sin complementaridad dentro del oligonucleótido especialmente en el extremo 3'.

El oligonucleótido GSP 3' que cubrió los criterios mencionados tiene la siguiente secuencia:

5' **gctcatcagggcaaagatggcaatcac** 3'

y está ubicado en las bases 42 –16 de la región codificadora del Betaglicano humano como se ilustra a continuación :

5' acgagacgcactgttgagaaataaaa**ATG**acttccattatgtgattgccatctttgcccctgatgagcttct 3'



Brevemente, el protocolo RACE consiste en lo siguiente:

Usando CIP (“Calf Intestine Phosphatase”) se desfosforilaron 4µg de RNA total con lo cual se quitaron los grupos fosfato en las posiciones 5' de los RNA no mensajeros y de los mRNA truncados, CIP no tiene efecto sobre los mensajeros encasquetados.

A continuación el RNA tratado con CIP se trató con TAP (“Tobacco Acid Pyrophosphatase”) para liberar el casquete de los extremos 5' de los mRNA, este tratamiento deja libre los fosfatos 5' que se necesitan para la ligación del oligonucleótido RACE 5'. Posteriormente se ligó el oligo RACE 5' a estos mRNA usando T4 RNA ligasa. Usando el GSP 3' como iniciador se realizó transcripción reversa con AMV RT, el cDNA resultante constituye el molde con extremos conocidos sobre el cual se practicó PCR. Los productos de amplificación, en este caso una banda de ~400 pb, fue purificada del gel de agarosa.

### Protocolo TOPO-Cloning

Usando el vector pCR4-TOPO / EcoRI activado con topoisomerasa, se ligaron los fragmentos purificados del gel, se transformó la reacción de ligación en células *E.coli* DH5-α competentes y se analizaron las colonias recombinantes produciendo minipreparaciones de plásmidos y digiriendo con la enzima EcoRI que libera los insertos. El análisis preliminar incluyó también un estampado de colonias sobre membranas de nitrocelulosa. Para identificar el contenido de las secuencias se llevó a cabo la hibridación tipo southern usando como sondas, oligonucleótidos sintéticos basados en la secuencia publicada del Betaglicano humano, correspondientes a las secuencias conocidas del 5' UTR y en la región inmediata del inicio de la traducción. Figura 13.

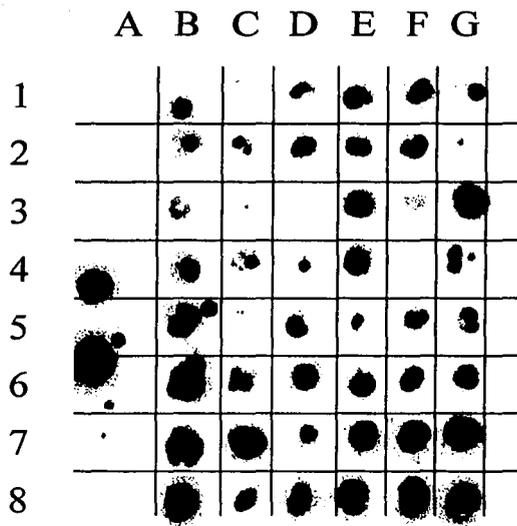


Figura 13. Hibridación tipo Southern de colonias recombinantes con los productos de amplificación del sistema RACE. El oligonucleótido empleado como sonda para la hibridación se ubica a -78-101 nt corriente arriba del ORF de las secuencias humanas del extremo 5'UTR publicadas [5]. Las colonias correspondientes a las posiciones A1, A2, A3, A7, A8, C1, C3, y D3 no hibridaron con con la sonda.

### Secuenciación y análisis de clonas recombinantes

Las secuencias de las clonas recombinantes nos permitió identificar dos poblaciones de extremos 5' UTR. Identificamos las dos poblaciones con los nombres RACE 8 y RACE 9. La mayoría de las clonas analizadas por secuenciación (14/18) contienen la secuencia RACE 8 y sólo 4 contienen la secuencia RACE 9. (Figura 14).

Las secuencias respectivas son:

#### RACE 8 (359pb)

CTTTCCTGTT CCCAGCGAGT GAAGGAGGGC AGTCGGCGGC TCTCGCGCCC CGGCCACTTT CCCTGCGCGA  
TTCCCGGAGC TCCCTGCAGG AGGTGAGAGT CCCCAGCGGG TCCGGATGGC GTAGTTTTGC CGCGGCGCAG  
CAGCTGCCGA GCTCGCCGCC GCCGAGCGCT GGGCGGGGAA ACTTGCCGCC GCTTTCCTCC AACTTGCTGC  
GGGTGGATCT CCGTGGACA CACCGCTCC GAGGAGGCAG TTTGAAAAT GCAAGGAGGG ACTTTAAGAC  
TACTTCTGAT TTGCAAAGAT GGTCTGTGCT CCGAGCAGC TAAAGTGACT GGACGAGAC CACTGTTGGA  
GAAATAAAA TG

#### RACE 9 (322pb)

GCGTGATTG CTGTGGCGTC CTGCCCGTCC CCGCCCGCGT GTGTGCGAGG GAGGGCGAGT GCGCCGGGTC  
GGCTTGATGG GGCTAAATCGA GGGTTTCGGG GATGCCGAGC GGCACTTTCC TCTTCCGAGC GCGTGAAGGA  
GGGCAGTCGG CGGCTCTCGC GCCCGGCCA CTTTCCCTGC GCGATTCCC GAGCTCCCTG CAGGAGGAGG  
CAGTTTGAAA ATTGCAAGGA GGGACTTTAA GACTACTTCT GATTTGCAAA GATGGTCTGT GCTCCGAGCA  
GGCTAAAGTG ACTGGACGAG ACGCACTGTT GGAGAAATAA AATG

Figura 14. El ATG que está resaltado al final en cada secuencia corresponde al codón de iniciación del ORF. Las secuencias subrayadas corresponden a secuencias comunes en ambas poblaciones de RACE identificadas a simple vista.

El análisis comparativo entre las secuencias RACE 8 y 9 usando el programa BLAST, nos permitió ir ubicando y refinando la localización de estas secuencias en el genoma humano. Primeramente este análisis estableció que las secuencias RACE 8 y 9 se encuentran exclusivamente en el cromosoma 1 en el locus del Betaglicano como se ilustra en la Figura 15, también estableció que se encuentran corriente arriba (5') de los exones codificadores del Betaglicano lo cual es consistente con la posición que se espera de las posiciones +1 respecto al resto de la unidad transcripcional, datos que confirman la identidad de estos extremos 5'.

Los resultados del programa BLAST indican también, que cada secuencia RACE está formada por dos segmentos que encuentran su contraparte en la secuencias genómicas (exones) separados por aproximadamente 80 Kpb. Esto demuestra que cada 5' UTR del Betaglicano humano está constituido por dos exones.

Tanto las secuencias de RACE 8 y 9 como el resto del gene de Betaglicano se encuentran en el Contig NT\_004686.5 panel B de la Figura 15. Un Contig (del inglés

Contiguous) es un segmento grande entre 1y 3 Mpb de DNA secuenciado, ensamblado a partir de segmentos más pequeños y orientado respecto al centrómero y constituyen los mapas físicos de DNA de más alta resolución que publica NCBI a través del proyecto del genoma humano.

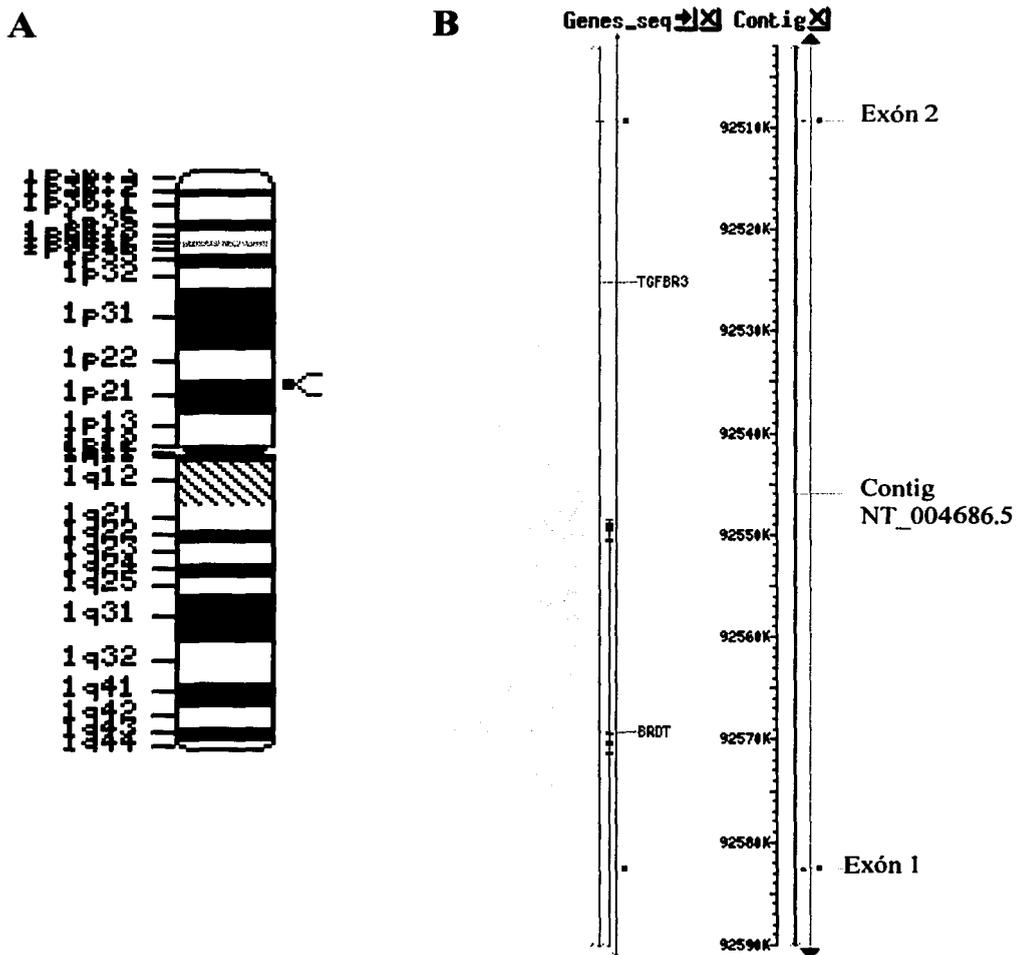


Figura 15. Resultados gráficos de BLAST con la secuencia RACE 8, se obtienen imágenes similares con RACE 9. A Ideograma del cromosoma 1, la marca roja indica el sitio en que se encontraron las secuencias genómicas correspondientes a RACE 8. B A la izquierda, está el mapa de los genes identificados en la zona. BRDT es un gene vecino y está sobrepuesto en el Betaglicano (TGFBR3) debido a que el mapa es de baja resolución. A la derecha está

el mapa del Contig NT\_4686.5, las dos marcas rojas indican la ubicación de los dos segmentos o exones que forman la secuencia de RACE 8.

Las dos posiciones +1 (una de RACE8 y otra de RACE9) se encuentran muy cercanas una de la otra: a 115 pb de distancia. Sin embargo un análisis cuidadoso de la localización de las secuencias RACE 8 y RACE 9 indica que existe una sobreposición en el contenido de las secuencias genómicas en el primer exón.

Lo anterior se aprecia mejor al visualizar la comparación de las secuencias RACE y las secuencias genómicas, base por base en un mapa de alta resolución sobre la secuencia del Contig NT\_004686.5 Figura 16.

### RACE 8

INT\_004686.5|Hs1\_4843 Homo sapiens chromosome 1 working draft sequence segment  
Length = 2128609  
Identities = 243/245 (99%), Gaps = 2/245 (0%)  
Strand = Plus / Minus

```
Query: 4      ctttcctcttccacgcagtgagtgaggaggcagtcggcggtctctcgccccggccacttt- 62
                |||
Sbjct: 440782 ctttcctcttccacgcagtgagtgaggaggcagtcggcggtctctcgccccggccacttt 440723
```

```
Query: 63      ccctgcgcgattccccggagctccctgcaggaggtgagagtcac-agcgggtccggatggc 121
                |||
Sbjct: 440722 ccctgcgcgattccccggagctccctgcaggaggtgagagtcacagcgggtccggatggc 440663
```

```
Query: 122     gtagttttgcccggcgagcagctgccggagctcgccgcccggagcgctggggcgggga 181
                |||
Sbjct: 440662 gtagttttgcccggcgagcagctgccggagctcgccgcccggagcgctggggcgggga 440603
```

```
Query: 182     aacttgcccgcctttctccaacttgctgcgggtggatctccgctggacacaccgcctc 241
                |||
Sbjct: 440602 aacttgcccgcctttctccaacttgctgcgggtggatctccgctggacacaccgcctc 440543
```

### RACE 9

INT\_004686.5|Hs1\_4843 Homo sapiens chromosome 1 working draft sequence segment  
Length = 2128609  
Identities = 207/208 (99%)  
Strand = Plus / Minus

```
Query: 3      cgcgctgattgctgtggcgtcctgccctccccgcccgctgtgtgctgaggaggggcgag 62
                |||
Sbjct: 440897 cgcgctgattgctgtggcgtcctgccctccccgcccgctgtgtgctgaggaggggcgag 440838
```

```
Query: 63      tgcgcccgggtcggcctgatggggtaatcgagggtttcggggatgccgagcggcactttc 122
                |||
Sbjct: 440837 tgcgcccgggtcggcctgatggggtaatcgagggtttcggggatgccgagcggcactttc 440778
```

```
Query: 123     ctcttcccagcagtgaggaggcagtcggcggtctctcgccccggccactttccctg 182
                |||
Sbjct: 440777 ctcttcccagcagtgaggaggcagtcggcggtctctcgccccggccactttccctg 440718
```

```
Query: 183     cgcgattcccggagctccctgcaggagg 210
                |||
Sbjct: 440717 cgcgattcccggagctccctgcaggagg 440690
```

\*pie de figura en la siguiente página

Figura 16. Las secuencias en el renglón "query" corresponden a RACE, las secuencias "subject" corresponden al Contig NT\_004686.5. La numeración en la secuencia del Contig revela un sobreposición entre ambas secuencias.

Para determinar el lugar en el que este sobreposición ocurre y definir con precisión los límites y los contenidos exónicos que tienen estos mensajeros, se realizó un análisis comparativo entre las dos secuencias RACE. El análisis estableció que estas secuencias se sobreponen en una región de 92 pb como se ilustra a continuación:

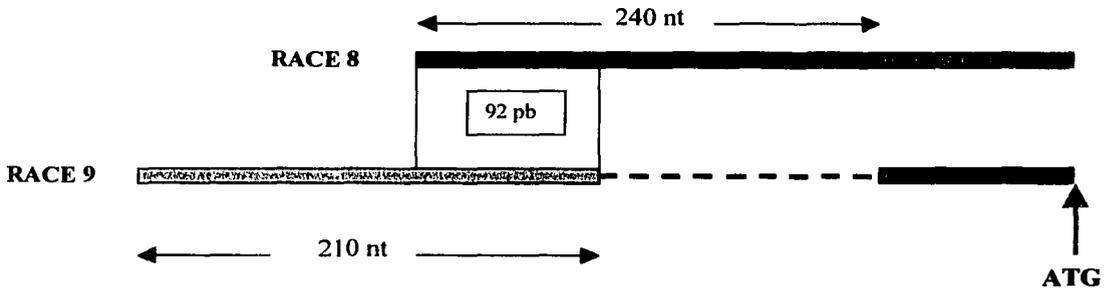
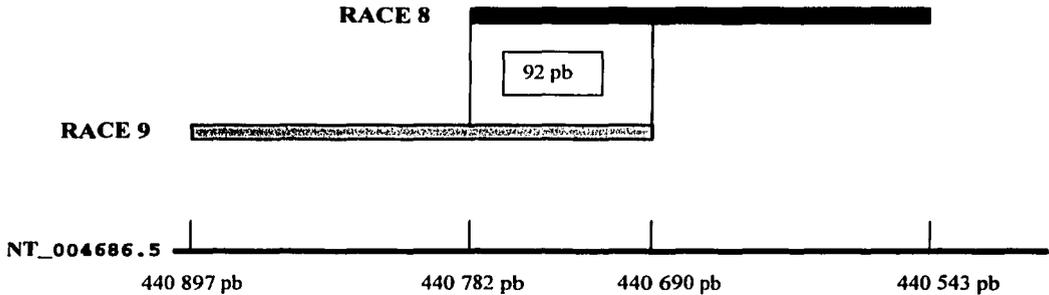


Figura 17 La relación que hay entre las secuencias RACE8, RACE9 y el Contig NT\_004686.5, sugiere que pueden existir dos probables sitios de inicio de la transcripción. Además de la sobreposición de estas 92 pb, a partir de la base 241 en RACE8 y de la base 211 de RACE9 las dos poblaciones tienen las mismas secuencias.

Estos hallazgos nos permiten proponer que estas dos poblaciones de RACE comparten el segundo exón que incluye el ORF, y nos permiten sugerir que en el primer

exón de este gene existen dos lugares donde puede iniciar la transcripción, separados por 115 pb.

No obstante que la orientación general de estas secuencias respecto al resto del gene de Betaglicano es consistente, la similitud encontrada entre las secuencias RACE8 y 9 en el Contig NT\_004686.5, es con la hebra de DNA "minus". Esto quiere decir que están en el sentido opuesto respecto a la orientación de la hebra codificadora del gene de Betaglicano en ese Contig. Esto es probablemente debido a que la secuencia de este Contig es una versión preliminar, y todavía está sujeta a revisiones continuas, como advierte la nota en el encabezado de la secuencia del Contig NT\_004686.5.

### Hallazgos en el 5'UTR de los mRNA amplificados por RACE

La búsqueda de motivos funcionales en bases de datos especializadas en regiones no traducidas de mRNA con el programa UTRScan [49], reveló la existencia de una secuencia IRES del inglés "Internal Ribosome Entry Site", próximo al ATG tanto en las secuencias de RACE 8 como RACE 9, y una secuencia 15-LOX.DICE del inglés "Differentiation Control Element" en RACE 8 aunque incompleta, como se indica en la Figura 18.

#### RACE 8 (359pb)

1	30	50	70
CTTTTGGTGTTC CCCACGGAGT GAAGGAGGGC ACTTCGGGGC TCTCCGGGCC CGGCACTTTT GCCTCGGGGA			
80	100	120	140
TTCCCGGAGC TCCCTGCAGG AGGTGAGAGT CCCCAGCGGG TCCGGATGGC GTAGTTTTGC CGCGGGCGAG			
150	170	190	210
CAGT'GCCGA GCTCGCCGCC GCCGAGCGCT GGGCGGGGAA ACTTGCCGCC GC'TTTCTCC AACTTGCTGC			
220	240	260	280
GGTGGATCT CCGCTGGACA <b>CACCGCTCC</b> <u>GAGGAGGCAG</u> TTTGAAAATT GCAAGGAGGG <u>ACTTTAAGAC</u>			
290	310	330	350
<u>TACTTCTGAT</u> <u>TTGCAAAGAT</u> <u>GGTCTGTGCT</u> <u>CCGAGCAGGC</u> <u>TAAAGTGACT</u> <u>GGACGAGACG</u> <u>CACTGTGGGA</u>			
360			
<u>GAAATAAAAA</u> TG			

#### RACE 9 (322pb)

1	30	50	70
CGCTGATTG CTGTGGCGTC CTGCCCGTCC CCGCCCGCGT GTGTGCGAGG GAGGGCGAGT GCGCCGGGTC			
80	100	120	140
GGCTGATGG GGGTAATCGA GGGTTTCGGG GATGCCGAGC GGCAC'TTCC TCTTCCAGC GCGTGAAGGA			
150	170	190	210
GGCAGTCGG CGGCTCTCGC GCCCCGGCCA CTTTCCTGCG GCGATTCCCG GAGCTCCCTG <u>CAGGAGGAGG</u>			
220	240	260	280
CAGTTGAAA ATTGCAAGGA <u>GGGACTTTAA</u> <u>GACTACTTCT</u> <u>GATTTGCAA</u> <u>GATGGTCTGT</u> <u>GCTCCGAGCA</u>			
290	310	330	
<u>GGCTAAAGTG</u> <u>ACTGGACGAG</u> <u>ACGCACTGTT</u> <u>GGAGAAATAA</u> <u>AAATG</u>			

Figura 18. En letras cursivas se señala la secuencia que forma el IRES, y en la secuencia RACE 8 se señala en negritas la estructura 15-LOX-DICE. El exón 2 está subrayado y sólo incluye hasta el ATG de la Metionina inicial..

### Definición de IRES.

Las secuencias IRES en los mRNA forman una estructura secundaria que controla la traducción del mRNA. La mayoría de los mRNA eucariones son traducidos a través de un mecanismo más común que es el dependiente del casquete de metilguanosa del extremo 5'. En este casquete se ensamblan los factores de iniciación con la subunidad pequeña del ribosoma, el que inicia un rastreo sobre el mensajero hasta encontrar en un contexto propicio el codón AUG que codifica la Metionina inicial. En ese momento empieza la fase de elongación del péptido [50]. Existen mecanismos alternativos de iniciación independientes del casquete del extremo 5'. La unión del ribosoma a un sitio interno del mRNA (IRES) es un mecanismo de traducción alternativo al mecanismo de traducción dependiente del reconocimiento del casquete de metilguanosa del extremo 5', y fué descrito inicialmente en picornavirus [51]. El inicio de la traducción interno le permite a los picornavirus, cuyos mensajeros no tienen casquete en el extremo 5', escapar del apagado de la maquinaria traduccional del huésped que ocurre durante la infección por estos virus. Los IRES pueden discriminarse funcionalmente de otras estructuras secundarias de los 5'UTRs, por su habilidad de mediar la traducción de segundos ORFs en mRNA bicistrónicos reporteros, independientemente del status traduccional del primer ORF [52][53]. Los IRES en mRNA eucariontes son similares pero menos complejos que los virales, pueden usar componentes en *trans* distintos a los que utilizan los IRES virales y pueden ser tan pequeños como 35 nucleótidos de largo [54]. El IRES encontrado en los 5'UTR de Betaglicano es de 82 nt de largo por lo que es muy probablemente funcional. La traducción de mRNA celulares mediada por IRES provee un mecanismo de traducción de un mRNA específico en circunstancias por ejemplo, en las que el mecanismo dependiente del casquete de guanosa esté inhabilitado. Además constituye un artefacto estructural con un potencial regulatorio que puede funcionar en varios contextos biológicos [55].

La búsqueda de motivos estructurales también identificó un elemento 15-LOX-DICE ("Differentiation Control Element"), aunque incompleto con sólo 14 nt de largo, lo que probablemente lo haga no funcional. El LOX-DICE completo, es decir de 38 nt es un elemento silenciador de la traducción, descrito inicialmente en el extremo 3'UTR del gene 15-lipoxigenasa. La proteína LOX, se expresa en células eritroides justo antes de madurar en eritrocitos, y está involucrada en la desintegración de la mitocondria. Esta restricción en la expresión se logra silenciando la traducción del mRNA en los precursores eritroides, a través de un complejo específico LOX-DICE-proteínas [56]. El hecho de que el elemento encontrado en el Betaglicano esté incompleto, limita las posibilidades de especular sobre el papel que en la regulación de la traducción pueda tener.

La predicción de estructuras secundarias estables de RNA, basada en las secuencias RACE 8 y RACE 9 (GCG Wisconsin Package -RNA Secondary Structure Function, FoldRNA y Squiggles-) produjo varios modelos de estructuras ver Figura 19 en la siguiente página. En la misma figura se muestra el modelo más estudiado y detallado de un IRES, el del virus de la hepatitis C.

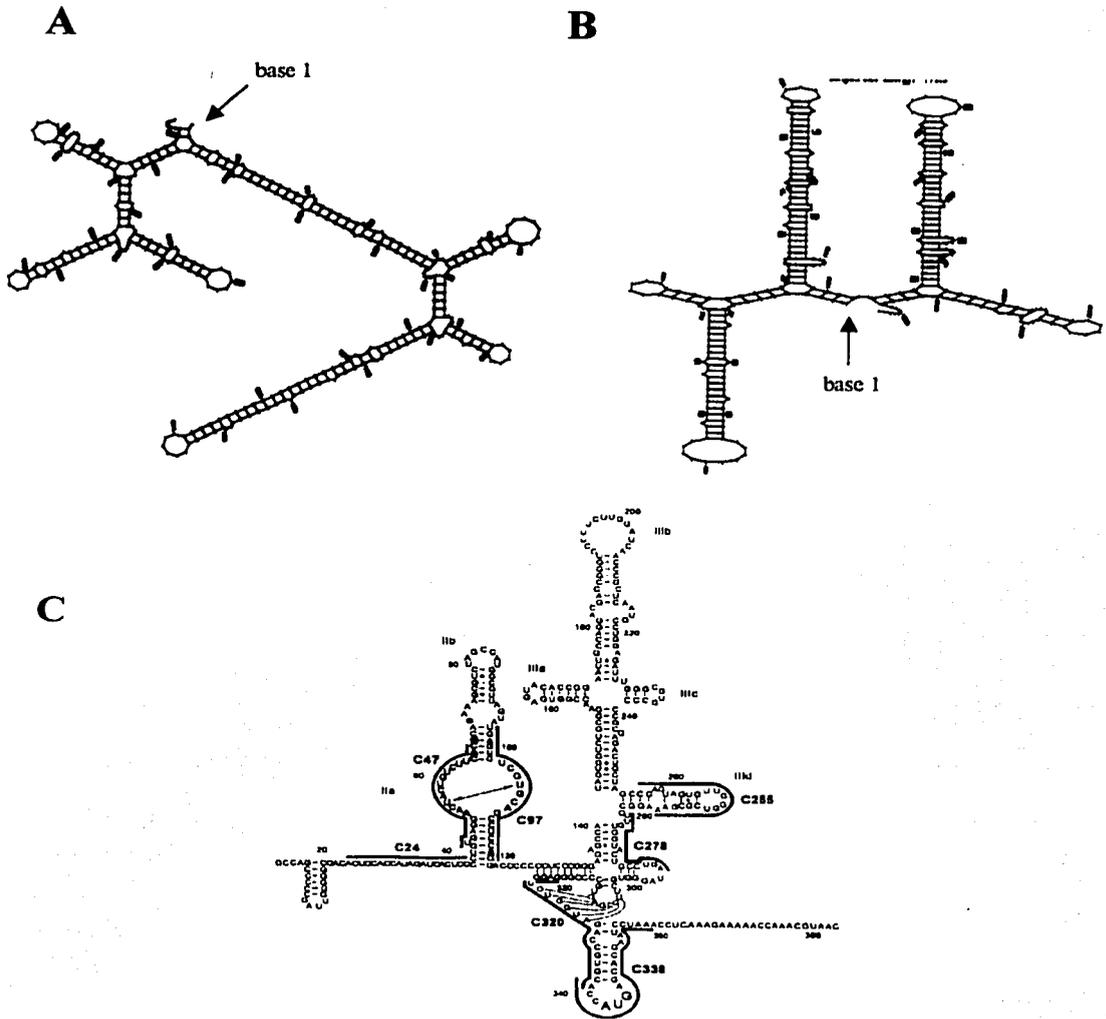


Figura 19. A Estructura secundaria RACE8, B Estructura secundaria RACE9, C Modelo del IRES viral.

ANÁLISIS CON  
FALLA DE ORIGEN

## Alineamiento múltiple de los 5'UTR de Betaglicano

Los resultados del alineamiento progresivo base por base de las secuencias 5' UTR de Betaglicano disponibles, se muestran a continuación en la Figura 20. El algoritmo con el que hicieron estos alineamientos introduce espacios para encontrar el alineamiento óptimo y esto permite visualizar los residuos estrictamente conservados en todas las secuencias de las especies incluidas, esos residuos están marcados en rojo. Además es posible ver a través de estos alineamientos que hay regiones conservadas, aunque no estrictamente en cada posición, lo que sugiere que estas secuencias pueden también estar jugando un papel en la regulación de la expresión de este gene al funcionar como secuencias blanco de proteínas que regulen la traducción.

	1				50
5'utr-ratón	-----	-----	-----	-----	-----
5'utr-rata	cctggcagtg	agagacagtg	gggagaggag	aaagcgggct	gctgtccttc
RACE-9	-----	-----	-----	~gcgctga	ttgctgtggc
RACE-8	--ctttcctc	ttccagcgag	tgaaggaggg	cagtcggcgg	ctctcgcgaa
5'utr-pollo	-----	-----	-----	-----	-----
	51				100
5'utr-ratón	-----	-----	-----	-----	~gaattc
5'utr-rata	caggccaactt	tcgctgcgcg	attccgcgag	ctccccgcag	gaggtgaaaag
RACE-9	gtcctgcccg	tccccgcccg	cgtgtgtgcg	agggagggcg	agtgcgccgg
RACE-8	aaggccaactt	tccttgcgcg	attccccgag	ctccctgcag	gaggtgagag
5'utr-pollo	-----	-----	-----	-----	-----
	101				150
5'utr-ratón	gcgcccgctg	cgaccggatg	gcgcagttga	actgc.gctg	ctgagctcgc
5'utr-rata	tccccggcgg	ggtccggatg	gcgcagttgc	actgcggctg	ctgagctcgc
RACE-9	gtcggcctga	tgggggtaat	cgagggtttc	ggggatgccg	agcggcactt
RACE-8	tcccagcggg	tcgggatggc	gtagttttgc	cgccggcgag	cagtgcccg
5'utr-pollo	-----	-----	-----	-----	-----
	151				200
5'utr-ratón	ggccgcctgc	gcgcgctgg.	...ggggact	cgtctgggct	...actccc
5'utr-rata	ggccgcctgc	gcgcgctggg	cgccggggact	cgcttcggct	agtaactcct
RACE-9	tcctcttccc	agcgagtgaa	ggagggcagt	cgccggctct	cgcccccgg
RACE-8	agctcgcgcg	cgccgagcgc	tgggcgggga	aacttgcgcg	cgctttcctc
5'utr-pollo	-----	-----	-----	-----	-----
	201				250
5'utr-ratón	ccacctcgcg	ttggacgacc	ggtcctggac	acaccgcttg	cgaggcaagt
5'utr-rata	ccacctcgcg	geggacgacc	ggtcctggac	acgtgctctg	cgaggcaagt
RACE-9	ccactttccc	tgcgcgattc	ccggagctcc	ctgcag..ga	ggaggcagtt
RACE-8	caacttgctg	cgggtggatc	tccgctggac	acaccgcctc	gaggcagtt.
5'utr-pollo	-----	~cggc	gcgcgggacg	aggcgtggac	ggagcccctg
	251				300
5'utr-ratón	tgaacattgc	agagaaggat	cctaaggcta	cacccgactt	gccacacttg
5'utr-rata	tgaacattgc	agagaaggat	cctaaggcta	cacccgactt	gccacgattg
RACE-9	tgaaaattgc	aagga.ggac	<u>tttaagacta</u>	<u>cttctgattt</u>	<u>gc.aaagatg</u>
RACE-8	tgaaaattgc	aaggaggac	<u>tttaagacta</u>	<u>cttctgattt</u>	<u>gcaaaagatg</u>
5'utr-pollo	agaagcctgc	aagatctgat	ttcaaagcac	accacaattt	tcaagggaag
	301				350
5'utr-ratón	ccatccagct	gaagaaccaa	aggctgttgg	agagatggca	<b>atg</b> -----
5'utr-rata	ccttcaatct	gaagaaccaa	aggctgttgg	agagatg	-----
RACE-9	<u>gtctgtgctc</u>	<u>cgagcagact</u>	<u>aaagtgactg</u>	<u>ccacqagacg</u>	<u>cactgtttga</u>
RACE-8	<u>gtctgtgctc</u>	<u>cgagcagact</u>	<u>aaagtgactg</u>	<u>gacqagacg</u>	<u>cactgtttga</u>
5'utr-pollo	taaatctaca	tgaagagcta	cagaagaac	tgatttaaga	agacggtgct

TESIS CON  
FALLA DE ORIGEN

	351	362
5'utr-raton	~~~~~	~~
5'utr-rata	~~~~~	~~
RACE-9	<u>gaaataaaaa</u>	tg
RACE-8	<u>gaaataaaaa</u>	tg
5'utr-pollo	tgcaaaaatg	~~

Figura 20. Con cursivas y subrayado se señala la secuencia que da lugar al IRES en RACE 8 y RACE 9, como se observa. este elemento estructural no está conservado en el resto de las secuencias, tampoco fue detectado con el programa UTRScan.

### Comentario y Discusión

La mayoría de los mRNA eucariontes tienen extremos 5'UTR entre 20 y 100 nucleótidos de largo. Bajo ese criterio el Betaglicano humano con 359 pb en una versión y 322 pb en otra, posee extremos 5'UTR largos. Experimentalmente, los recortamientos de los 5'UTR a menos de 12 nucleótidos imposibilitan la traducción, en contraste el aumento de la longitud del 5'UTR puede incrementar la eficiencia de traducción, debido a que se pueden cargar más complejos de preiniciación sobre la hebra 5'UTR. Sin embargo muchos mRNA celulares con extremos 5'UTR excepcionalmente largos, se traducen pobremente debido a la presencia de estructuras secundarias. Esto parece ser especialmente común en mRNA que codifican proto-oncogenes, factores de transcripción, factores de crecimiento [57] y sus receptores como es el caso de Betaglicano. Lo que sugiere que su traducción es rigurosamente regulada. Bajo qué circunstancias y en qué lugares del organismo entran en juego estos mecanismos regulatorios, son preguntas por abordarse.

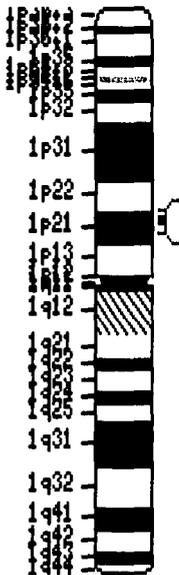
La confirmación de la existencia de estas dos especies de 5'UTR (RACE8 y 9) y su significado funcional son temas por explorar. Los experimentos de "primer extension" no han sido concluyentes, en parte debido a que los oligonucleótidos empleados corresponden a secuencias del segundo exón y es probable que las estructuras secundarias que muy probablemente existen obstaculicen la transcripción con AMV RT. Generar nuevos oligonucleótidos y condicionarlos para que funcionen en los ensayos de "primer extension" aportará pruebas valiosas sobre la existencia de RACE 8 y 9 *in vivo*. También serán de utilidad ensayos de digestión a la nucleasa S1 para determinar el o los dos sitios de inicio de la transcripción. En este momento no es posible determinar si las dos variaciones del exón 1 se derivan del proceso de empalme o si efectivamente se trata de dos inicios de la transcripción independientes.

Queda por comentar que el segundo tamizaje con la compañía Molecular Genetics aportó un nuevo BAC con otras 150 kpb corriente arriba de la región codificadora. El arribo de este BAC, se sobrepuso con los momentos en que la amplificación a partir de DNA genómico humano se estaba llevando a cabo, por lo que la caracterización del promotor murino, quedó a cargo de otro miembro del laboratorio.

## VII. ESTRUCTURA DEL GENE HUMANO DE BETAGLICANO

El gene humano de Betaglicano se encuentra en el cromosoma humano número 1. Usando técnicas de citogenética y FISH (Fluorescent in situ Hibridization) se ha reportado su localización en la banda 1 p32-33 [58]. Sin embargo por mapeo físico y secuenciación, el proyecto del genoma humano reporta la ubicación del Betaglicano en el contig NT\_004686.5 cuya localización genética corresponde a la banda 1 p21-22, como se ilustra en la Figura 21.

**A**



**B**

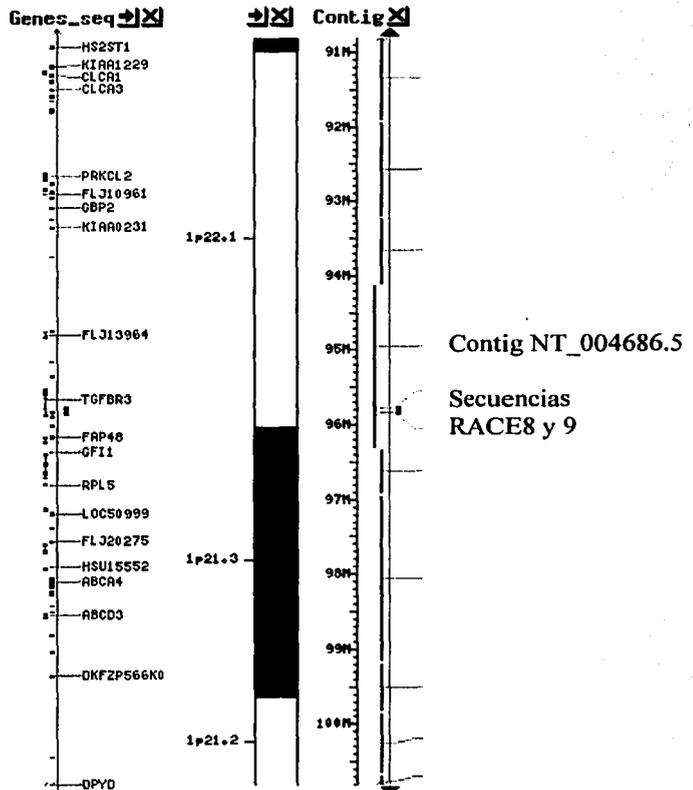
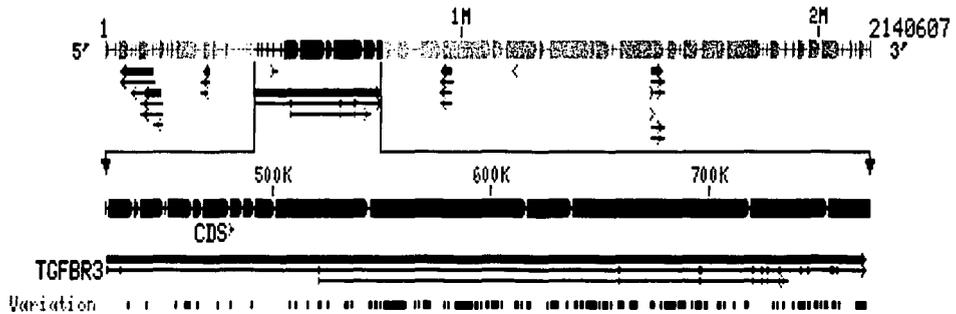


Figura 21 A Cromosoma 1 teñido con Giemsa. B Tres mapas equivalentes de la región sombreada en A. Izquierda, mapa con algunos genes identificados en la región; centro, mapa citogenético banda 1p21-22; derecha, mapa de los Contigs disponibles en la zona. Las marcas rojas indican la posición de las secuencias RACE.

Dentro del Contig NT\_004686.5, el gene de Betaglicano está esparcido en un segmento de aproximadamente 300 000 pb, como se ilustra en la figura 22.



**Legend:**

--- - segment boundaries    ——— - CDS    ——— - RNA    ——— - gene

Figura 22 En azul oscuro se representan los segmentos que comprenden el gene de Betaglicano ya ensamblados. "Variation" indica a las variantes polimórficas de nucleótido sencillo (SNP) encontradas en la secuencia del gene de Betaglicano.

Las secuencias RACE 8 y RACE 9 que obtuvimos al caracterizar el extremo 5' no traducido del mRNA y discutidas en el capítulo VI, apoyan el modelo de dos variaciones del exón 1. Considerando estas evidencias experimentales y la información publicada a través del proyecto del genoma humano, el gene de Betaglicano estaría constituido por 17 exones y en la tabla siguiente se resumen las características generales.

TESIS CON  
FALLA DE ORIGEN

Ubicación en NT_004686.5	Exón No.	Tamaño en pb	Intrón tamaño en pb
440897-440690	1	210	
440782-440538	1 <sup>a</sup>	240	
			80 547
521085-521258	2	173	
			137 382
658640 - 658820	3	180	
			37 294
696114-696251	4	137	
			23 665
719916-720099	5	183	
			4 803
724902-725070	6	168	
			1 999
727069-727216	7	147	
			5 515
732731-732920	8	189	
			1 725
734645-734982	9	337	
			430
735412-735563	10	151	
			2 604
738167-738307	11	140	
			174
738481-738639	12	158	
			3 693
742332-742631	13	299	
			3 460
746091-746211	14	120	
			10 534
756745-756786	15	41	
			2 310
759096-759203	16	107	
			11 825
771028-772459	17	1431	

Tabla 1. La representación de estos datos se encuentran en la Figura 23, en la siguiente página.

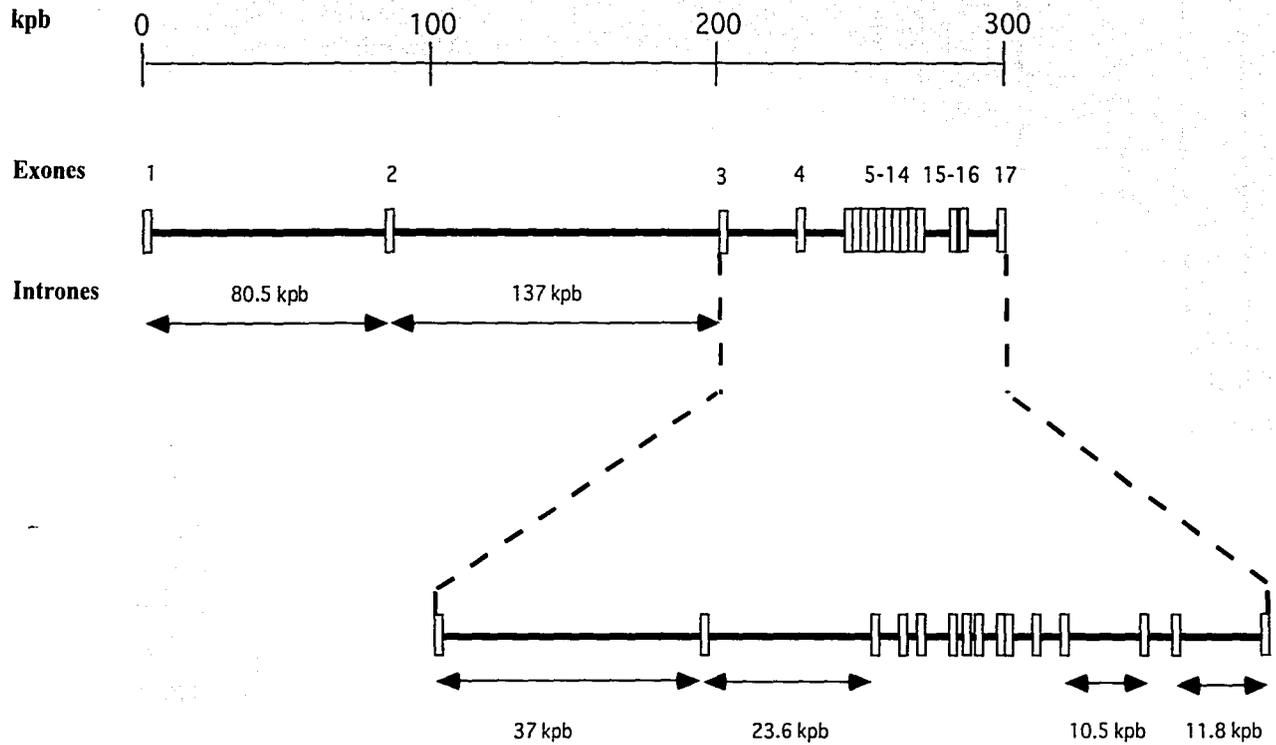


Figura 23. Representación de la estructura del gene de Betaglicano, en el texto se indican los tamaños de los intrones más pequeños.

## Relaciones estructurales entre las secuencias exónicas y peptídicas del Betaglicano

En las secuencias de la Figura 24 se ilustran alternando entre los colores amarillo y verde los límites exónicos en el cDNA, y en la región codificadora sobre la secuencia peptídica se han sobresalido en azul las estructuras  $\alpha$ -hélice y en rojo las estructuras  $\beta$ -{plegadas predichas con el programa ExPasy. Esta ilustración permite visualizar la relación que hay entre la estructura exónica del gene y las estructuras secundarias que por lo menos a nivel teórico se pueden deducir a partir de la secuencia primaria, lo que se puede llamar la modularidad del diseño de Betaglicano a nivel exónico. Como puntos sobresalientes hay que destacar que la secuencia del péptido señal constituido por los primeros 20 residuos, está incluido en un sólo exón (el exón 2), así como toda la región transmembrana (exón 16), y prácticamente toda la región citoplasmática (exón 17). 30 de 37  $\alpha$ -hélices o  $\beta$ -plegadas se encuentran codificadas dentro de regiones exónicas y solo 7 subestructuras se encuentran codificadas entre los límites de dos exones.

### Comentario y Discusión

La discrepancia que existe entre la ubicación del Betaglicano por métodos citogenéticos y el mapeo físico, puede deberse a limitaciones en la técnica citogenética empleada y a la complejidad del cromosoma 1 por ser el más grande de todos los cromosomas humanos. Los cromosomas mitóticos que están en metafase temprana muestran mas bandas de condensación que los cromosomas que están en metafase más tardía. Por lo que la ubicación citogenética pudo haber variado dependiendo si la tinción se realizó en metafase temprana o tardía.

El tamaño de los intrones en organismos eucariontes, tiene un rango amplísimo que va entre 80 nt a > 10 000 nt [59]. Los primeros tres intrones del gene de Betaglicano caen dentro del rango de intrones muy grandes. Sin embargo hay que tener en cuenta que las secuencias del Contig NT\_004686.5, no son versiones definitivas en cuanto a su orientación. La reubicación exónica dentro de algún segmento variará las distancias intrónicas establecidas actualmente si esas orientaciones cambian.

El gene de Betaglicano está dentro de las categorías mediano-grande si lo comparamos con algunos otros genes humanos:

Gene	Tamaño del gene miles de nucleótidos	tamaño del mRNA miles de nucleótidos	# de intrones (*) #
$\beta$ -globina	1.5	0.6	2
insulina	1.7	0.4	2
albúmina	25	2.1	14
factor VIII	186	9	25
tiroglobulina	300	8.7	36
<b>Betaglicano</b>	<b>300</b>	<b>7.5</b>	<b>16</b>
Distrofina	>2000	17	> 50

\* Tomado de [59] p. 340.

En cuanto a la modularidad exónica, se puede decir que con las herramientas de predicción estructural disponibles hoy en día, las posibilidades son altas de que

10 20 30 40 50 60 70 80 90 100  
 actttcctcttcccagcgagtgaggagggcagtcggcggtctcgcgccccggccaatttcctgcgcatcccggagctccctgcaggaggtgagag  
 110 120 130 140 150 160 170 180 190 200  
 tccccagcgggtccggatggcgtagttttgcccggcgagcagcagctgcggagctgcgcgcccggagcgtggcgggaaacttgcgccgcttccct  
 210 220 230 240 250 260 270 280 290 300  
 ccaacttgctgcggtggatctccgttgacacaccgctccgaggcagtttgaaaaatgcaaggaggactttaagactacttctgatttgcaagatg  
 310 320 330 340 350 360 370 380 390 400  
 gtctgtgctccgagcaggctaaagtgactggacgagacgcactgttgaggaaaaataaaatgacttcccattatgtgattggccatctttgcctgatgagc  
 M T S R Y V I A P A L M S  
 410 420 430 440 450 460 470 480 490 500  
 ttctgtttagccaactgca  
 F C L A T A G P E P G A L C E L S P V S A S H P V Q A L M E S F T  
 510 520 530 540 550 560 570 580 590 600  
 V L S G C A S R G T T G L P Q E V H V L N L A L R Q G P G Q L Q R E  
 610 620 630 640 650 660 670 680 690 700  
 ggtcacacttcacctgaaatcccctcctcagtcacatccaccacaagtctgtgtgctcctgctcaactccccaccctccctgggtggcatctgag  
 V T L H L N P I S S V H I H H K S V V F L L N S P H P L V W H L K  
 710 720 730 740 750 760 770 780 790 800  
 acagagacttgccaactgggtctccagactgtttt  
 T E R L A T G V S R L F L V S E G S V V Q F S S A N F S L T A E T  
 810 820 830 840 850 860 870 880 890 900  
 E E R N F P H G N E H L L N W A R K E Y G A V T S F T E L K I A R N  
 910 920 930 940 950 960 970 980 990 1000  
 agatcaagtgttccctccaaagtccaacatagggaaagaattttctcactcaattaccctgctgagtaccttcaacc  
 I Y I K V G E D Q V F P P K C N I G K N F L S L N Y L A E Y L Q P  
 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100  
 aaagcagcagaagggtgtgtgatgtccagcccagaaatgaggaagtacacatcatcgagctaatcaccccaactctaaccctac  
 K A A E G C V M S S Q P Q N E E V H I I E L I T P N S N P Y S A F  
 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200  
 Q V D I T I D I R P S Q E D L E V V K N L I L I L K C K K S V N W V  
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300  
 ttgctcctaacagttatggctttggaaaagagagtgaaagatctatgacaatgaccaaatca  
 I K S F D V K G S L K I I A P N S I G F G K E S E R S M T M T K S  
 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400  
 ataagagatgacattccttcaaccgaagggaatctgtggaagtggcctttggacaatggctatagccaataacttcatacacaatggctcctgtggcaa  
 I R D D I P S T Q G N L V K W A L D N G Y S P I T S Y T M A P V A  
 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500  
 tagtattcatcttcggcttgaataaa  
 I V F H L R L E N N E E M G D E E V H T I P P E L R I L L D P G A L  
 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600  
 P A L Q N P P I R G G E G Q N G G L P F P F P D I S R R V W N E E  
 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700  
 G E D G L P R P K D P V I P S I Q L F P G L R E P E E V Q G S V D  
 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800  
 ccagtggtactcgggatggcgtcacctgtt  
 I A L S V K C D N E K M I V A V E K D S F Q A S G Y S G M D V T L L  
 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900

TESIS CON  
 FALLA DE ORIGEN

ggatcctacctgcaaggccaagatgaatggcacacacttgttttggagtctcctctgaaatggctgaggactcggcccccgggtgagcccttgatggg  
 D P T C K A K M N G T H F V L E S P L N G C G T R P R W S A L D G

1910 1920 1930 1940 1950 1960 1970 1980 1990 2000  
 V V Y Y N S I V I Q V P A L G D S S G W P D G Y E D L E S G D N G

2010 2020 2030 2040 2050 2060 2070 2080 2090 2100  
 ttccgggagatgagatggaaggagatgcttccctgttccaccgacctgaaatcgtggtgtttaaattgcagcctcagcaggtaggaaccccagcagctt  
 F P G D M D E G D A S L F T R P E I V V F N C S L Q Q V R N P S S F

2110 2120 2130 2140 2150 2160 2170 2180 2190 2200  
 ccaggaacgcccacggaacatcaccttcaacatggagctatacaaacactgacctcttttggctccctcccaggcgtcttctctgtgccagagaat  
 Q E Q P H G N I T F N M E L Y N T D L F L V P S Q G V F S V P E N

2210 2220 2230 2240 2250 2260 2270 2280 2290 2300  
 ggacacgtttatgttga  
 G H V Y V E V S V T K A E Q E L G F A I Q T C F I S P Y S N P D R

2310 2320 2330 2340 2350 2360 2370 2380 2390 2400  
 M S H Y T I I E N I C P K D E S V K F Y S P K R V H F P I P Q A D M

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500  
 D K K R F S F V F K P V F N T S L L F L Q C E L T L C T K M E K H

2510 2520 2530 2540 2550 2560 2570 2580 2590 2600  
 cccccagatgttcctaagtgtgctcctcctgacgagcctgcacctcgctggagcctcgataaacttgggccatgacagaataagaagacgttcacca  
 P Q K L P K C V P P D E A C T S L D A S I I W A M M Q N K K T F T

2610 2620 2630 2640 2650 2660 2670 2680 2690 2700  
 agccccctgtgtgatccaccatgaagcagaatctaa  
 K P L A V I H H E A E S K E K G P S M K E P N P I S P P I F H G L D

2710 2720 2730 2740 2750 2760 2770 2780 2790 2800  
 caccctaaccgtgatgggcattgcttggcagccttgtgatcgggagcactcctgacggggccttgggtacatctattctcacacaggggagacagca  
 T L T V M G I A F A A F V I G A L L T G A L W Y I Y S H T G E T A

2810 2820 2830 2840 2850 2860 2870 2880 2890 2900  
 gcaggaaggcagcaagtccccacctccccgcagcctcggaacacagcagtgctgcccacagcatcgccagcagcagcagcagcagccttctccagcagca  
 G R Q Q V P T S P P A S E N S S A A H S I G S T Q S T P C S S S S

2910 2920 2930 2940 2950 2960 2970 2980 2990 3000  
 gcacggcctagcccaacccagcccaaccccgcccaacccagccagccagcctcagctcagctactccaaggcaggaaccaatggctgagcctcgtgct  
 T A \*

3010 3020 3030 3040 3050 3060 3070 3080 3090 3100  
 cagactcagagggctggatttgggttcccttgaagaca gagtgaatttcagataaagatcaccggttattcacccccacccagggctagtataa

3110 3120 3130 3140 3150 3160 3170 3180 3190 3200  
 acatgacctgggtctctgtaccacactagaattcattgtgagaagcctaaatgggtgcttctccaccagccctcagagcctgggggtttctctatgt

3210 3220 3230 3240 3250 3260 3270 3280 3290 3300  
 gaaacacatgccagttttaaataatgctgcttgtccaggtgagaacatccataattggggccctgagttttaccagactcaaggagttgtgtaaagggt

3310 3320 3330 3340 3350 3360 3370 3380 3390 3400  
 taatagccagatagtagaacacagtgaggagatgcccgaagattctttatctatctgaaaccaagatgtaaaacaagaaatgctttgagggctttctaaagca

3410 3420 3430 3440 3450 3460 3470 3480 3490 3500  
 tcctcctgtctaatttgcaccttgtctggatgacctcttctgaccttctgcccacaaccttaaggcaacatctcttctgtgacctggggcccaaaacca

3510 3520 3530 3540 3550 3560 3570 3580 3590 3600  
 atgctgatgacctatcagctcctgttcttcccaactgcatacaccactgcaaaatgctttaaagcaaaattgtattcttaccagggcctacagaaa

3610 3620 3630 3640 3650 3660 3670 3680 3690 3700  
 ttgaaaatgacaaaatcaggaaccagagatttggcccttccctaatttggctcgtcaaaatattgt ataattgaggtgaaatcagttataaagt

3710 3720 3730 3740 3750 3760 3770 3780 3790 3800  
 caaggacgaatttgcacagtgatatttc tatgtgatgcaagtacaagtataaattttagagttgcaacttaagtatttggtaggtgagtg

3810 3820 3830 3840  
 ttccactcaaaatgtgca acttaaaaaaaa  
 3850 3860 3870 3880

Figura 24. Los residuos peptídicos codificados de la base 350 a la 417 y resaltados en gris corresponden al péptido señal. Los residuos codificados de la base 2701 a 2775 corresponde a la región transmembranal. El extremo 5' no traducido del primer exón, corresponde a las secuencias RACE 8.

\* Tomado de [59] p. 340.

En cuanto a la modularidad exónica, se puede decir que con las herramientas de predicción estructural disponibles hoy en día, las posibilidades son altas de que efectivamente las secuencias exónicas correspondan a motivos estructurales específicos dentro de la proteína como queda claramente definido para el péptido señal, la región transmembranal, la cola citoplasmática y el número importante (30 de 37  $\alpha$ -hélices o  $\beta$ -plegadas) de estructuras secundarias codificadas dentro de regiones exónicas.

En la medida que los bancos de datos experimentales sobre estructuras proteicas aumenten de tamaño, la búsqueda de la relación que puede existir entre exones y dominios específicos de las proteínas aportará evidencias más sólidas que puedan confirmar las correlaciones incipientes encontradas en el análisis reportado aquí.

## VIII. CLONACIÓN DEL PROMOTOR Y ENSAYOS FUNCIONALES

### Abordaje experimental

La comparación de las secuencias RACE 8 y RACE 9 usando el programa BLAST contra el genoma humano, demostró la ubicación exclusiva de las secuencias correspondientes en el locus en donde está localizado el Betaglicano. Las secuencias RACE se encuentran en el Contig denominado NT\_004686.5 constituido por 2 128 609 pb. En ese Contig como ya se comentó en el capítulo VI, también se encuentran las secuencias genómicas del gene de Betaglicano. Las secuencias RACE constituyen muy probablemente dos inicios de la transcripción del Betaglicano y por lo tanto las secuencias inmediatas "corriente arriba" contienen el promotor y otras secuencias reguladoras. Para probar esto y utilizando como base la secuencia publicada de la región, se diseñaron una serie de oligonucleótidos con el objetivo de amplificar el área de interés usando la reacción de polimerasa en cadena (PCR) a partir de DNA genómico humano. Una vez amplificados varios fragmentos se clonaron en pTOPO, se identificaron por secuenciación y posteriormente los fragmentos genómicos se subclonaron en el vector pGL3-Basic que contiene el gene reportero de luciferasa. Siendo la luciferasa una enzima monomérica y activa sin necesidad de modificaciones postraduccionales, su actividad catalítica y cinética enzimática es considerada un reflejo casi directo de la actividad transcripcional del gene y por ende de las regiones reguladoras colocadas corriente arriba de la región codificadora.

Debido a que un ensayo funcional provee los medios para identificar y valorar la relevancia de los elementos regulatorios de una secuencia de DNA, es determinante en las fases iniciales de estudios transcripcionales, contar con un sistema de inducción del gene en cuestión en este caso Betaglicano, para poder delinear los elementos de control críticos de su promotor o de regiones de control distales.

El proceso de diferenciación de las células C2C12 constituye un ensayo funcional adecuado ya que durante este fenómeno se observa un aumento en el mRNA de Betaglicano. Las células C2C12, son mioblastos de ratón en los que se puede inducir un proceso de diferenciación a miotubos por deprivación de factores de crecimiento. Este sistema constituye un ensayo funcional ideal en el cual probar inicialmente la supuesta actividad promotora del fragmento genómico amplificado por PCR y subclonado en pGL3-Basic. Mediante experimentos de transfección transitoria en las células C2C12, se comprobó la inducción estadísticamente significativa del gene de luciferasa. Estos resultados sugieren fuertemente que las secuencias amplificadas y clonadas contienen el promotor del Betaglicano humano.

### Ubicación del primer exón y de las secuencias corriente arriba

Asumiendo que el extremo 5' del transcrito maduro de Betaglicano contiene los sitios de inicio de la transcripción, ubicamos las secuencias genómicas correspondientes en el Contig NT\_004686.5 del cromosoma 1, como se discutió en el capítulo VII. Con la disponibilidad de la secuencia nucleotídica de este Contig, extrajimos el texto de 2610 pares de bases río arriba de los sitios en los cuales se encuentran las secuencias genómicas correspondientes a RACE. La secuencia de 2610 pb se denominó PCRtemp y se ilustra en la Figura 25. En estas 2610 pares de bases esperamos encontrar el promotor del gene de Betaglicano.

PCR temp

10	20	30	40	50	60
aagtgatccaccacacctgg	cttcccaaattgctgagatt	ataggcatgagccacctcac			
70	80	90	100	110	120
ctggccgatgcattttaaattg	tgggattaatatattcacacc	tagttcaagtgttcacatct			
130	140	150	160	170	180
gtacttgaaactaaaaacttg	ttaacaaggcataaatactg	tattagtctgttttcatgct			
190	200	210	220	230	240
gctgataaagacatacccga	gactgggcaatttacaagg	aaaggtttattggacttaca			
250	260	270	280	290	300
gttccacatggctggggaga	cctcacaatcatggcagaag	gtaaggaggagcaagtcaca			
310	320	330	340	350	360
tatcacgtggatggcagcag	gcagagagagctggtacag	gagactcctgtttttaagc			
370	380	390	400	410	420
catcagatctcgtgagactt	attcactatcat	<del>tadaacaggtcacaggaaa</del>	actcatcccc		
430	440	450	460	470	480
aagattcaattacctccat	tgggttcctcctaggacaca	tggggattgtgggacttaca			
490	500	510	520	530	540
attcgagatgagatttgggtg	ggggacacagccaaaccata	tcaaatactgtttaatttga			
550	560	570	580	590	600
gtctaggagaaatcttaagg	ccttaaacttagactttata	aaagcagaggctagctccat			
610	620	630	640	650	660
ctggactgtttattccttca	gaatattatgcctacttgta	ttcccaatgcctgactcagg			
670	680	690	700	710	720
gctctcccatagtcagtgct	caatgaatattggttaaggaa	aaaaatggttcagtgaggag			
730	740	750	760	770	780
atthttgcagtaattcaggt	gactgatgacatttgacctt	tgtgaatgaatctccaattc			
790	800	810	820	830	840
tgcaattgggtgattctctc	gcctatagggtacaagaggaa	gaggacgagaaaatatttctt			
850	860	870	880	890	900
gttctctcaataactttcca	aataacatgtcataagaagc	ctcctcattctactctccaa			
910	920	930	940	950	960
catatataaagcctaggatt	gagctcaggtgcagcggcta	acgcctctaatacccagcact			
970	980	990	1000	1010	1020
ttgggagggccaaggcagggc	aatctcttgagctcaggagt	ttgagaccagcctggggcaac			
1030	1040	1050	1060	1070	1080
taagtgagacccccgcctcg	ccaacccccctctacaaaaaa	tacaaaaagtttagctgggag			
1090	1100	1110	1120	1130	1140
tgggtggcattttgtcagtaat	cccagctactcaggaggctg	agggtgggagaatcgcttgag			
1150	1160	1170	1180	1190	1200
cctgagaagcggagggttgta	gcgagctgagattgagattg	tgccactgcactccagcctg			
1210	1220	1230	1240	1250	1260
ggcgacagggcgagactcca	tctcaaaaaaaaaaaaaaaaa	aaaaaaaaaaaggcataggt			
1270	1280	1290	1300	1310	1320
ccatagccctttagaactgg	aagaagccttctaggccaag	cccttcatttgaactgagcc			
1330	1340	1350	1360	1370	1380
ttgacgggtattagccgtta	gccatttgggtgaaggctcga	cagtagtcaggactggaagt			
1390	1400	1410	1420	1430	1440
ctggctcttgaaacctgtca	gtccacggctcatttcacctg	cctgtttgtaaaacaatgtc			
1450	1460	1470	1480	1490	1500
ctgttacttaagtatacttt	ttaataagaaaaactgggtg	taagcaagagtatcagagtg			
1510	1520	1530	1540	1550	1560
gtgttctgggtgccagcagtt	caactctgcagcgtgtact	tcacaggcacaggctacttag			
1570	1580	1590	1600	1610	1620
gacgtgcacaggcacaggt	acttaggacgttgccccgc	gagactgtcgttcatatcgc			

1630	1640	1650	1660	1670	1680
gccacctccgggtaacgcc	ggtaa	<b>gctgacccagctcc</b>	<b>atctc</b>	ggtattaatttatt	
1690	1700	1710	1720	1730	1740
tagctagcagggggaacggc	tggtaatgcagctctcagct	gctggaggggaagaggcaag			
1750	1760	1770	1780	1790	1800
actccgccggcgcccacat	cagagcgtgacaacaagcgg	gaccggaggttgggaaaaag			
1810	1820	1830	1840	1850	1860
cacagggacagaatgaatgt	gtgcatgggggagctctgct	ggggagagggaagaggctg			
1870	1880	1890	1900	1910	1920
tgatcagattatctgaggtt	tcagaacttttatataaaga	ccaaggaaccaaataaaaaat			
1930	1940	1950	1960	1970	1980
ccaactcttccagggagccc	tggaagtccaaggacctc	gtttgagggagcaggtcccc			
1990	2000	2010	2020	2030	2040
gcagctcgtgccacgtctc	cctcgcgcttccccaggac	aggggcccgcaccaaccggg			
2050	2060	2070	2080	2090	2100
cacgctgagccccggccggc	ccctagccccttctctctcg	gcctcctcgtcccgaaggga			
2110	2120	2130	2140	2150	2160
agcacaggctcgagcagcat	tcggggctcatccatcagcc	gcggaggcacaagggggagga			
2170	2180	2190	2200	2210	2220
gagaggaggcgggaggcggg	aggcgggaggaggagtttct	ggcggccgagagcccccg			
2230	2240	2250	2260	2270	2280
tggggcccggcggaggagcgc	ccttccccccgcgctgat	tgctgtggcctcctgccgt			
2290	2300	2310	2320	2330	2340
ccccgccgcgtgtgtgcga	gggagggcagtgccgcccgg	tcgctctgatgggggtaatc			
2350	2360	2370	2380	2390	2400
gagggtttcggggacgccga	gcggcac	<b>ttctcccca</b>	<b>gccagctgaag</b>	gaggggcagtc	
2410	2420	2430	2440	2450	2460
ggcggctctcgcgccccggc	cactttccctgcgcgattcc	cggagctccctgcaggaggt			
2470	2480	2490	2500	2510	2520
gagagtcaccagcgggtccg	gatggcgtagttttgccgcg	gcgcagcagctgccggagct			
2530	2540	2550	2560	2570	2580
cgccgcgcggagcgcctggg	cgggaaacttgccgc	<b>gcttccctccaacttgcctg</b>	cggg		
2590	2600	2610			
tggatctccgctggacacac	cgctccgag				

Figura 25. Secuencia denominada PCRtemp, extraída del Contig NT\_004686.5 en la que se encuentran las secuencias RACE 8 y RACE 9. La base 2253 (en mayúscula y sombreada en gris) corresponde a la posición +1 de RACE 9; la base 2368 corresponde a la posición +1 de RACE 8. Sombreados en gris se indican los oligonucleótidos diseñados para amplificar por PCR a partir de DNA genómico humano.

TESIS CON  
FALLA DE ORIGEN

Sobre la secuencia PCRtemp se diseñaron 4 oligonucleótidos que flanquean la región donde proponemos que está el promotor, y que se sintetizaron para amplificar la región a partir de ADN genómico humano obtenido de leucocitos. Estos oligonucleótidos y su ubicación aproximada se ilustran en la Figura 26:

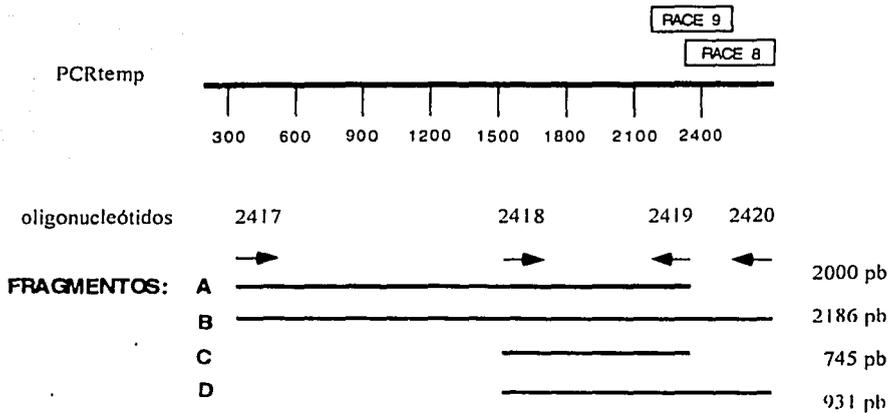


Figura 26. En la parte superior se esquematiza la secuencia PCRtemp y la ubicación de las secuencias RACE. Más abajo los números de identificación, la ubicación y orientación de los oligonucleótidos diseñados para amplificar los fragmentos A, B, C y D.

Los oligonucleótidos 2417, 2418, 2419 y 2420 tienen en sus extremos, sitios de reconocimiento para las enzimas de restricción Kpn I ó Hind III. Debido a que el vector pGL3-B que contiene el gene de luciferasa como reportero, tiene un sitio múltiple de clonación "corriente arriba" de la región codificadora de luciferasa. La presencia de estos sitios de restricción en los extremos amplificados facilitó la subclonación direccional de los productos de la PCR en pGL3-B.

Secuencias de los oligonucleótidos empleados en la reacción de amplificación de los fragmentos A, B, C y D:



Al terminar los ciclos de amplificación, se retiraron alícuotas (10µl) de cada tubo y se sometieron a electroforesis en gel de agarosa al 0.8 %, mostradas en la Figura 26.

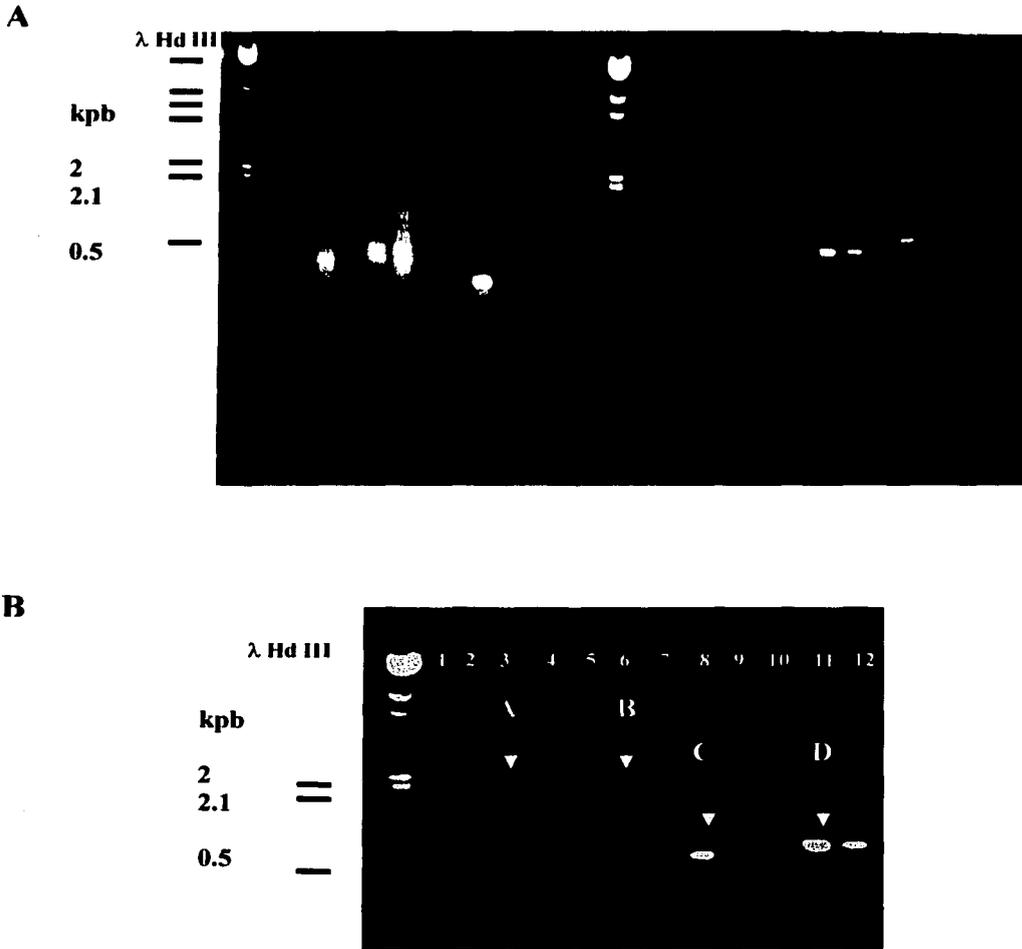


Figura 27. **A** muestras en azul y rojo corresponden a las rampas de tm 40-50°C. **B** Los fragmentos A y B se amplificaron preferencialmente en las concentraciones de 3 mM de MgCl<sub>2</sub> y en el sector de 60° C de tm. Mientras que los fragmentos C y D se amplificaron a concentraciones de 1.5 mM de MgCl<sub>2</sub> en el mismo sector de 60° C del gradiente de tm.

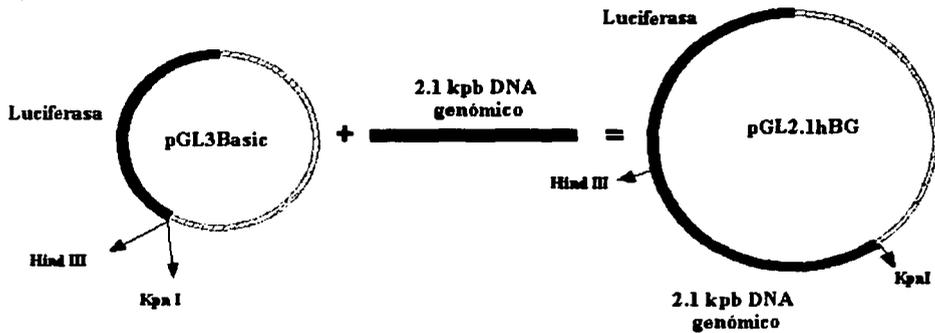


Figura 27. Estrategia de clonación. Los sitios Hind III y Kpn I del vector pGL3Basic, se aprovecharon para forzar la orientación correcta del fragmento genómico amplificado por PCR

### Estrategia de clonación

Los fragmentos amplificados se purificaron del gel y se clonaron primero en vectores pTOPO. Una vez identificados por secuenciación, los insertos A y B (2004 y 2191 pb respectivamente) se sacaron digiriendo con las enzimas Kpn I y Hind III para subclonar direccionalmente los fragmentos en el vector pGL3-B. Esta estrategia forzó y facilitó la orientación correcta del fragmento amplificado corriente arriba del gene reportero de luciferasa como se aprecia en el mapa de la figura 27.

### Ensayos de Transfección

La construcción pGL2.1hBG se transfectó en células C2C12 para medir la capacidad promotora del fragmento clonado. Para lograr esto las células se sometieron a un ensayo de diferenciación a miotubos por privación de factores de crecimiento, en el cual se ha establecido por Northernblot que hay un aumento del mRNA de Betaglicano (datos del laboratorio por publicarse). Como control interno para la normalización de los datos de expresión, se usó el sistema Dual-Luciferase Reporter System, el cual emplea además de la luciferasa de luciérnaga *Photinus pyralis*, la co-transfección de la luciferasa de *Renilla reniforme* como co-reportero. Esto permite eliminar variaciones experimentales tales como diferencias en el manejo de volúmenes o diferencias en la eficiencia de transfección entre las células y con ello precisar los datos de expresión del gene reportero. Brevemente: se sembraron  $6 \times 10^3$  células  $\times$   $\text{cm}^2$  y se cultivaron en medio de crecimiento por 48 hr, se transfectaron con lipofectamina, 5 hr después se les agregó SFB, y extracto de embrión de pollo a concentraciones finales de 10% y 0.5 % respectivamente, suplementos

Brevemente: se sembraron  $6 \times 10^3$  células  $\times \text{cm}^2$  y se cultivaron en medio de crecimiento por 48 hr, se transfectaron con lipofectamina, 5 hr después se les agregó SFB, y extracto de embrión de pollo a concentraciones finales de 10% y 0.5 % respectivamente, , suplementos que forman parte del medio de crecimiento de estas células. Al día siguiente se detonó la diferenciación cambiando los suplementos por suero de caballo al 10%. Las células se solubilizaron en los tiempos 0, 2, 4 y 6 días post-detonación y los lisados se almacenaron a  $-20^\circ \text{C}$  hasta el momento del ensayo enzimático. A partir del día 2 se cambió el medio respectivo cada 48 hrs, agregando AraC (citidilato arabinósido, GIBCO) para disminuir el efecto de las células que aún se estuviesen dividiendo. Las señales luminiscentes se detectaron y registraron con un luminómetro Lumat LB 9501. Cada punto temporal se hizo por triplicado y el análisis de los resultados se ilustran en la Figura 28 :

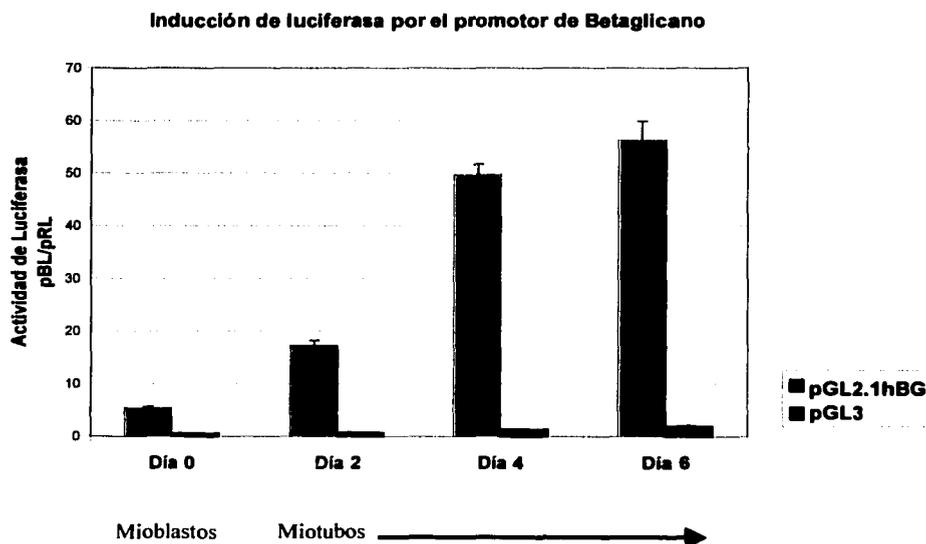


Figura 28. Actividad de luciferasa en el ensayo de diferenciación de células C2C12 a miotubos, transfectadas con el plásmido pGL2.1hBG

El fragmento genómico de 2.1 kpb indujo la expresión del gene reportero aproximadamente 10 veces más que el vector sin región promotora lo cual fuertemente sugiere que el fragmento de 2.1 kpb contiene el promotor del Betaglicano humano.

Considerando que este ensayo de diferenciación es un evento miogénico, y que la miogénesis se ha estudiado a un nivel de considerables detalles y mecanismos moleculares, se inició la caracterización del promotor humano explorando primeramente la presencia de elementos transcripcionales en cis involucrados en miogénesis. Para hacer esto, se probaron los efectos sobre la actividad del promotor de Betaglicano de diversos factores de

Los factores considerados inicialmente son :

<u>Factor</u>	<u>Efecto sobre miogénesis</u>	<u>Concentración empleada</u>
Acido Retinóico(AR)	inductor	1 $\mu$ M
TGF- $\beta$ 1	inhibidor	1 ng/ml
TGF- $\beta$ 2	inhibidor	"
TGF- $\beta$ 3	inhibidor	"
FGF 2	inhibidor	3 ng/ml
BMP	inductor	1 nM
HGF	inhibidor	3 ng/ml
Butirato	inhibidor de MyoD y Miogenina	2 mM

Las modificaciones de estos experimentos consistieron en que 24 hr después de la transfección con lipofectamina, las células se tripsinizaron y se resembraron en medio de crecimiento por 24 horas más, entonces se añadieron los factores mencionados y 24 horas después las células se solubilizaron y se realizó el ensayo enzimático. Los resultados se muestran en la gráfica de la figura 29

### Efecto de diversos factores sobre la inducción del promotor humano

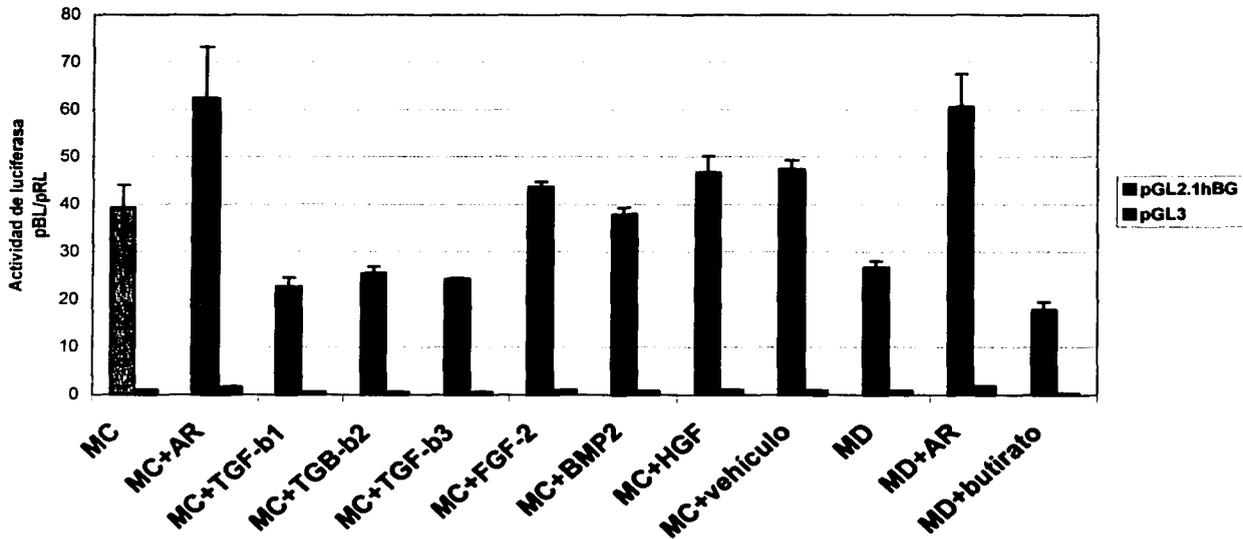


Figura 29. MC = Medio de Crecimiento. MD = Medio de Diferenciación

Los resultados que se muestran en la Figura 29, indican que el Acido Retinóico induce de modo importante la actividad promotora del Betaglicano lo cual es consistente

con el aumento del mRNA de Betaglicano observada durante la diferenciación de las células C2C12 a miotubos, y también es consistente la inhibición observada por el butirato que es un inhibidor de Miogenina y MyoD, genes maestros de la miogénesis. El efecto de los demás factores sobre la actividad del promotor no es tan definido como se ha establecido para la miogénesis, y estos resultados constituyen la primera aproximación experimental para estudiar detalladamente las señales que este promotor puede integrar estimulando o inhibiendo la transcripción del mRNA del Betaglicano.

### **Análisis e identificación de sitios reconocidos por diversos factores de transcripción**

Como se muestra en la Figura 30, la búsqueda de sitios de reconocimiento para factores de transcripción sobre la secuencia de 2.1 kpb que incluye el supuesto promotor, identificó elementos cuyos scores ( $>0.9$  en las matrices de comparación de las secuencias consenso) sugieren constituir los sitios de reconocimiento de los mediadores miogénicos involucrados en la transcripción del Betaglicano durante la diferenciación de las células C2C12. Es el caso de los sitios MEF2, MyoD, RAR (“Retinoic Acid Receptor”) y Miogenina. Es interesante destacar también la presencia de un elemento Nkx2-5 (score = 1) que es un factor de transcripción involucrado en cardiogénesis. Y no obstante presentar scores relativamente bajos  $-0.87$ ,  $0.69$  y  $0.65$ -, la presencia de 3 elementos SRF (“Serum Response Factor”) son importantes ya que han sido recientemente descritos como los mediadores de la miocárdina, un novedoso factor de transcripción involucrado también en cardiogénesis.

Es importante señalar la identificación de los sitios de reconocimiento de Smads3 y Smad4 mediadores de la señalización del TGF- $\beta$ , y el número importante de otros sitios con score alto no necesariamente involucrados en miogénesis, y pertenecientes a las familias de factores homeóticos como lo son POU, Ikaros2, octamer (no mostrados en la figura 30). El estudio detallado y la ponderación meticulosa de los fenómenos biológicos en los que estos elementos están involucrados, constituyen pistas importantes que permitirán discernir los eventos y las circunstancias en que este gene se enciende o se apaga.

### **Comentario y Discusión**

El esfuerzo inicial y crítico para definir el promotor humano de Betaglicano ha incluido tanto la ubicación del inicio de la transcripción como la identificación de potenciales elementos en *cis* involucrados en el control del inicio de la transcripción. Además la disponibilidad del ensayo funcional en células C2C12 constituye un sistema valioso que provee los elementos correspondientes en *trans* que definen al promotor y que permitirán la ampliación de estudios relacionados con la regulación de la transcripción. Aunque el ensayo de diferenciación transcurre en 6 días, es decir un tiempo considerado largo para ensayos de transfección transitoria, se puede argumentar que el proceso de diferenciación implica que las células no se están dividiendo y eso puede explicar la estabilidad y actividad sostenida de los plásmidos transfectados.

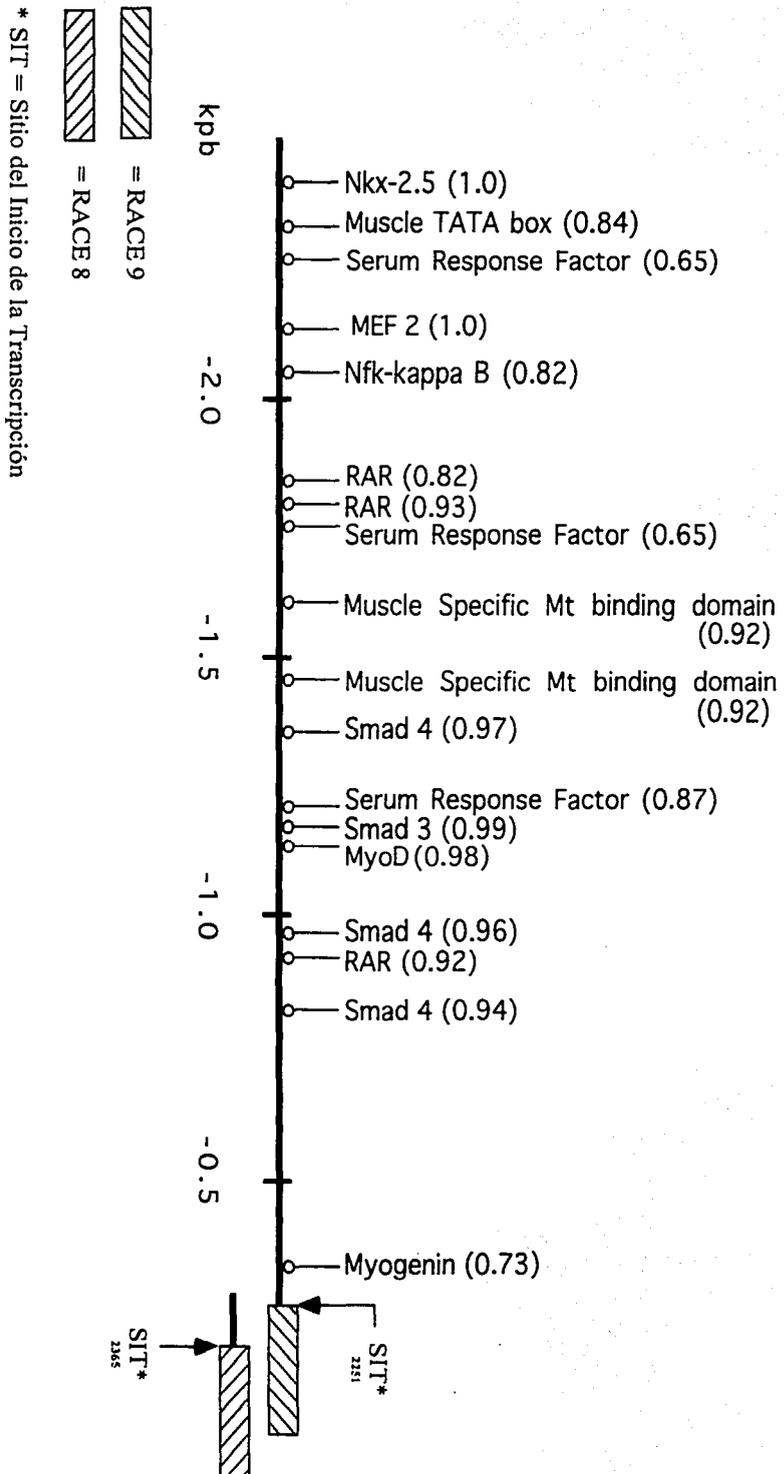


Figura 30. Posibles sitios de reconocimiento para diversos factores de la transcripción en el promotor humano de betagalactano

## IX. CONSTRUCCION DE TRANSGENES Promotor Alb-Betaglicano

Como último capítulo en esta tesis doctoral se reporta también la construcción de dos transgenes de Betaglicano bajo el promotor de albúmina. Una construcción que contiene el gene que codifica para la forma soluble y otra para la forma transmembranal del Betaglicano y las pruebas de expresión en cultivo de tejidos que demuestran su funcionalidad.

La idea que antecede a estas construcciones proviene de los ratones transgénicos llamados Alb- $\beta 1$  y mencionados en la introducción. Estos animales sobre-expresan la forma activa del TGF- $\beta 1$  aprovechando la potente inducción y la exclusiva actividad transcripcional hepática del promotor de albúmina [72]. Estos animales transgénicos tienen niveles circulantes altos de TGF- $\beta 1$  activo y los efectos sistémicos que se observan incluyen lesiones fibróticas en hígado, miocarditis, glomerulonefritis, arteritis, atrofia del páncreas, y atrofia testicular. El fenotipo de estos animales es una fuerte evidencia que apoya el papel etiológico del TGF- $\beta 1$  en una variedad de desórdenes inflamatorios y fibróticos [68-70]. Además como modelo transgénico constituye un sistema *in vivo* en el cual se pueden probar intervenciones terapéuticas para neutralizar los efectos dañinos del exceso de esta citocina. Tomando en cuenta que la función bioquímica propuesta para el Betaglicano soluble es la de un apagador molecular del TGF- $\beta$ , resulta atractivo desarrollar un modelo murino de sobreexpresión del Betaglicano mediada por el promotor de albúmina en el cual probar esta función bioquímica *in vivo*. Para generar estos animales transgénicos, usamos como estructura modelo el “transgene” que se usó para la generación de los ratones Alb- $\beta 1$  [52].

### Diseño de la Construcción

El primer paso para generar estos animales fue la construcción de los transgenes de Betaglicano bajo el mismo promotor murino de albúmina que se usó en la generación de los transgénicos Alb- $\beta 1$ . Para llevar a cabo estas construcciones, el Dr. Palmiter nos donó amablemente el plásmido Alb e/p (Figura 31) que contiene un elemento distal de activación (“enhancer”) y el elemento central del promotor de albúmina murina en un fragmento de 2.1 kpb [72].



Figura 31. El Plásmido Alb e/p tiene un polilinker en el extremo 3' y está flanqueado por varios sitios de restricción únicos

La Dra. Roberts nos donó amablemente el plásmido TGF $\beta$ -Alb (Figura 31) de donde extrajimos el extremo 3' UTR de la hormona de crecimiento humano (hGH), segmento que estabiliza y elonga la vida media del mensajero que produce el transgene [52].

Como insertos codificadores usamos las versiones soluble y transmembranal del Betaglicano cuyos sitios de inserción de GAGs están mutados. Esta decisión está basada en dos razones: por un lado se ha establecido que la unión del Betaglicano al TGF- $\beta$  es dependiente de la proteína central y no de los GAGs; y por otro lado a que los GAGs unen otra citocina importante: el FGF ("Fibroblast Growth Factor"), y aunque no se sabe qué significado o papel fisiológico tenga esta unión, el uso de una versión sin GAGs elimina la posibilidad de efectos mediados por la unión del FGF y permitirá correlacionar los efectos del Betaglicano soluble solo a su interacción con TGF- $\beta$ . Estas versiones de Betaglicano también contienen el epítipo de myc en el extremo amino lo que facilitará identificar en los ensayos funcionales la presencia de los productos proteicos de las construcciones.

### **Estrategia de construcción de los transgenes.**

Del plásmido Alb- $\beta$ 1 sacamos la región 3' no traducida de la hormona de crecimiento humana (hGH) [71] con las enzimas Sal I / Bgl II. El fragmento se purificó y se subclonó en pBSKS en los sitios Bam HI y Sal I. De ahí se extrajo con las enzimas Xho I / Xba I, sus extremos se rellenaron usando el fragmento Klenow y se subclonó en el plásmido e/p en un sitio Eco RV único. Esto permitió dejar entre el promotor de albúmina y el extremo 3' de hGH un sitio Bam HI único en el vector, en el cual se pudo introducir alguna de las dos versiones del Betaglicano. La construcción completa está flanqueada por dos sitios únicos: Sst II y Kpn I como se indica en la Figura 32, los que permiten separar la construcción del vector para la microinyección de embriones murinos.

Cada paso de la construcción se confirmó por secuenciación y una vez incluidos todos los fragmentos de interés en el lugar y orientación correctos, se ensayaron por transfección transitoria en células HepG2, un hepatoma humano que expresa marcadores de diferenciación hepática y que expresa constitutivamente los tres receptores del TGF- $\beta$ . Los productos de expresión se identificaron en ensayos de marcado por afinidad al TGF- $\beta$ 1 y por inmunoprecipitación usando anticuerpos anti-myc. Las muestras se sometieron a electroforesis en geles de poliacrilamida-SDS.

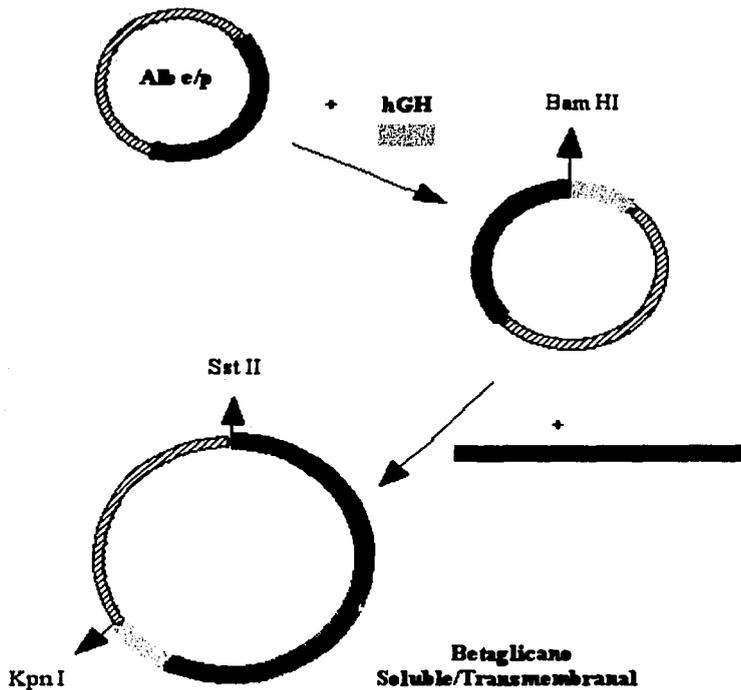


Figura 32. Estrategia para construir los transgenes de Betaglicano bajo el control del promotor murino de albúmina. Los sitios únicos Sst II y Kpn I flanquean la construcción completa. El sitio único Bam HI ubicado entre el promotor y el extremo 3' de la hormona humana de crecimiento (hGH), se usó para introducir las versiones de Betaglicano soluble o transmembranal.

Como se aprecia en la Figura 33, el método de transfección con fosfato de calcio, resultó más eficiente al permitir la entrada, la expresión y la identificación de la construcción transmembranal en mayores cantidades que la transfección con lipofectamina. Como era de esperarse los productos de las construcciones solubles no se detectaron en el extracto total. Para identificar la forma soluble, se recuperó el medio condicionado de éstas células y se realizó un marcado por afinidad en solución.

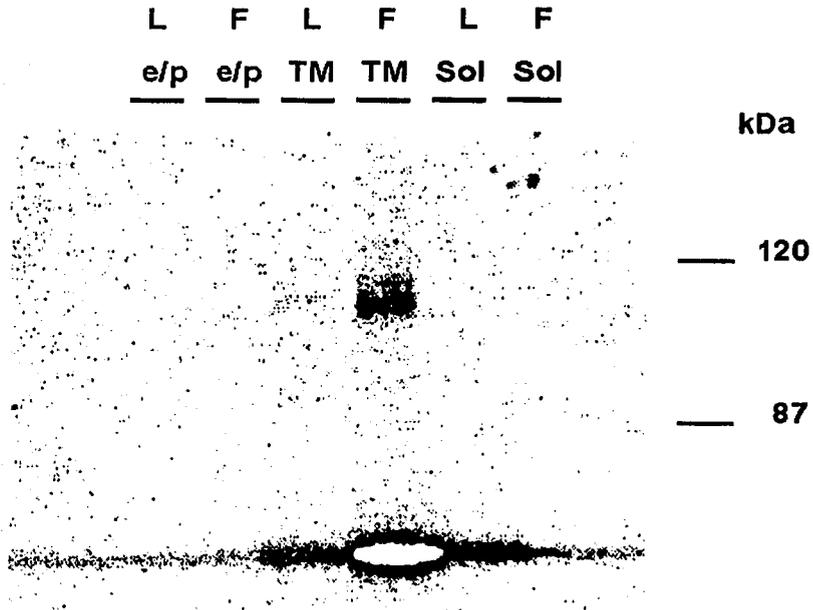


Figura 33. Marcado por afinidad con TGF- $\beta$ 1 e inmunoprecipitación con anti-myc del extracto total de células HepG2 transfectadas con las construcciones: **e/p** contiene el enhancer-promotor de albúmina sin Betaglicano, **TM** es la construcción e/p más la versión transmembranal de Betaglicano, y **Sol.** es la construcción e/p más la versión soluble del Betaglicano. Se ensayaron dos métodos de transfección; lipofectamina (**L**) y fosfato de calcio (**F**).

El marcado por afinidad del medio condicionado reveló la presencia de la forma soluble del Betaglicano con ambos métodos de transfección. Es interesante notar que en la transfección con lipofectamina de la forma transmembranal, se detectan rastros de la forma soluble, lo que indica que una porción de la forma transmembranal está siendo liberada también al medio. (Figura 34)

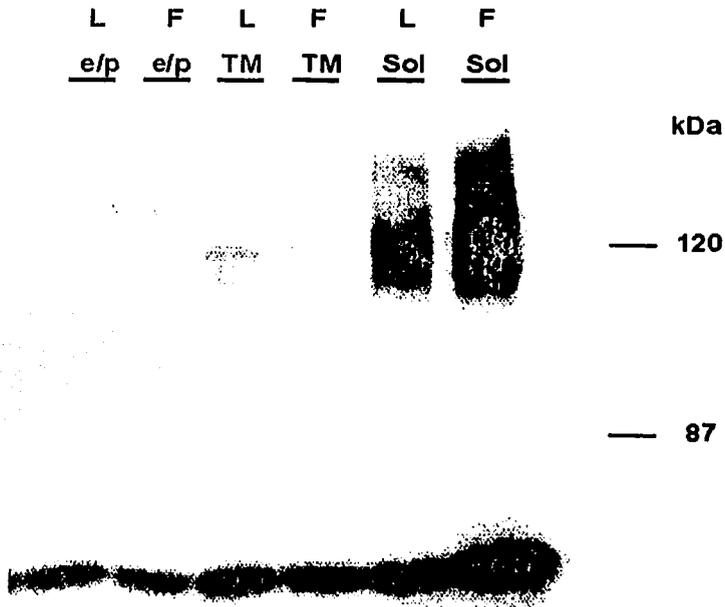


Figura 34. Marcado por afinidad e inmunoprecipitación con anti-myc del medio condicionado de células HepG2 transfectadas con las construcciones e/p, TM y Sol.

Una vez ensayadas estas construcciones *in vitro* se procedió a extraer la construcción linearizada y libre de secuencias del vector para lo cual, estos plásmidos se digirieron con los sitios únicos que flanquean la construcción: Sst II y Kpn I.

### **Comentario**

Estas construcciones se han usado ya en la microinyección de embriones en un trabajo colaborativo con la Universidad de Texas con el Dr. Héctor Martínez. Y ya hemos confirmado 6 animales que han integrado las construcciones. En el mediano plazo se procederá a la propagación de estos animales, para determinar si expresan los productos de las construcciones y si es así, continuar con la caracterización del fenotipo si es que presentan alguno. Si los transgenes se expresan y los animales no muestran un fenotipo obvio, lo siguiente será probar su respuesta a estímulos fibróticos, y probar si al cruzar estos animales con los transgénicos Alb-Beta 1 el fenotipo de este último se revierte.

## IX. PERSPECTIVAS Y CONSIDERACIONES

La clonación y la disponibilidad del cDNA de ratón nos aporta criterios concretos, respecto a las extrapolaciones que podremos hacer hacia el modelo humano cuando sigamos investigando y estudiando el Betaglicano en el modelo murino; lo mismo puede decirse de la caracterización del promotor humano. En la medida que el Betaglicano es una estructura química con potencial uso farmacológico como antagonista del TGF- $\beta$ , las comparaciones firmes entre el humano y los modelos de investigación disponibles se convierten en aspectos indispensables y su importancia sobrepasa el interés de sólo ampliar nuestra visión filogenética sobre este gene-proteína.

Los hallazgos en el extremo 5'UTR del mensajero humano y las diferencias con las secuencias murinas constituyen piezas importantes en el intento de caracterizar los mecanismos reguladores de la expresión de Betaglicano. En este momento no hay suficientes datos para definir si las dos especies del exón 1 reportadas aquí, son producto del procesamiento de empalme o de dos eventos iniciales de transcripción independientes. Aunque la evidencia producida con la metodología RACE apoya la idea de estar ante la presencia de dos sitios probables de inicio de la transcripción. Nuestros hallazgos son consistentes con un reporte en el que se han documentado los sitios de inicio de la transcripción (usando metodología emparentada con RACE) de más de 276 mRNAs y en el cual encontraron que la mayoría de estos genes tienen una distribución diversa de sitios de inicio de la transcripción, distribuidos en una media de 61.7 pb. Se trata particularmente de genes que no contienen caja TATA, sino un iniciador, como es probablemente el caso de Betaglicano. [72] La visión más actualizada es la de regiones, más que de sitios puntuales donde inicia la transcripción. En contraste la presencia de cajas TATA, es característica de los transcritos cuyo inicio se aglomera básicamente en un sitio, y se piensa que la torción y rigidez creada por la interacción de TBP ("TATA Binding Protein") y el DNA, ancla al complejo de preiniciación al DNA, determinando más estrictamente el lugar de inicio de la transcripción. En la base de datos de promotores eucarionetes de la EMBO, (EPD <http://222-epd.isb-sib.ch/> en la edición de enero del 2002, solo están registrados 258 genes humanos en los que experimentalmente se ha determinado el inicio de la transcripción, teniendo en mente que se estiman 30 000 genes en el humano, se pone en relieve la importancia que nuestros hallazgos tienen.

Por otra parte y aunque hay un trecho considerable para confirmar si estas especies ocurren *in vivo* así como descubrir su significado fisiológico, los hallazgos reportados en los extremos 5' UTR son además punto de partida para estudiar otros aspectos de la regulación de esta proteína, ya que ofrece sustratos estructurales sobre los cuales proponer y demostrar mecanismos reguladores de la expresión del Betaglicano a nivel de la traducción.

Considerando que los dos significados más ampliamente usados en la definición del promotor esto es, la identificación de las secuencias que se encuentran inmediatamente río arriba del inicio de la transcripción, y la identificación de elementos en *cis* que pueden mediar el inicio de la transcripción, han sido abordados para el Betaglicano, queda ahora abierta la puerta para extender y aprovechar estos conocimientos. Primero confirmar si la acumulación del mensajero en el sistema de las células C2C12 se debe efectivamente a un aumento en la transcripción o a un aumento en la estabilidad del transcrito, y segundo el contexto miogénico en el que parece inducirse este gene es una pista valiosa que nos

orienta sobre el donde y cuando buscar la inducción de Betaglicano durante el desarrollo embrionario. La generación de un nuevo transgene que incluya el promotor de Betaglicano controlando un gene reportero como Xgal en un ratón transgénico, revelaría los momentos críticos en los cuales se induce o reprime la expresión de este gene. Esta información serviría para aclarar las sospechas de esas potenciales funciones desconocidas.

La estructura exónica y su relación con las regiones protéicas que se presume tengan estructura secundaria (a nivel teórico), apoya la idea de la modularidad del Betaglicano, es decir la correspondencia exón-estructura secundaria o mejor aún exón-dominio protéico. El empleo de la distribución exónica para generar versiones protéicas mutantes con delección de exones específicos que no unan TGF- $\beta$ , aportarán información crítica sobre los residuos necesarios y estructura mínimas involucradas en esta interacción proteína-proteína.

La generación de ratones transgénicos (Alb-BG) que idealmente tendrán niveles circulantes altos de Betaglicano, nos permitirá probar el efecto de este aumento de Betaglicano circulante ante diversos estímulos fibrogénicos mediados por TGF- $\beta$ . Concretamente podremos probar esta hipótesis en el modelo de fibrosis hepática producida por intoxicación con tetracloruro de carbono (CCl<sub>4</sub>) y mediada por TGF- $\beta$ . Si el Betaglicano funciona como un apagador del TGF- $\beta$  *in vivo*, podremos observar una disminución del fenómeno fibrótico al desafiar a estos animales con CCl<sub>4</sub>. Así mismo, si el Betaglicano soluble circulante funciona como un antagonista del TGF- $\beta$ , al cruzar los animales transgénicos Alb-BG con los transgénicos Alb- $\beta$ 1, se revertirá de manera importante el fenotipo de estos últimos.

## XI. BIBLIOGRAFIA

1. Andres, J.L., et al., *Membrane-anchored and soluble forms of betaglycan, a polymorphic proteoglycan that binds transforming growth factor-beta*. J Cell Biol, 1989. **109**(6 Pt 1): p. 3137-45.
2. Lopez-Casillas, F., et al., *Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF-beta receptor system*. Cell, 1991. **67**(4): p. 785-95.
3. Blobbe, G.C., et al., *Functional roles for the cytoplasmic domain of the type III transforming growth factor beta receptor in regulating transforming growth factor beta signaling*. J Biol Chem, 2001. **276**(27): p. 24627-37.
4. Andres, J.L., et al., *Purification of the transforming growth factor-beta (TGF-beta) binding proteoglycan betaglycan*. J Biol Chem, 1991. **266**(34): p. 23282-7.
5. Moren A, I.H., Miyazono K, *Molecular cloning and characterization of the human and porcine transforming growth factor-beta type III receptors*. Biochem Biophys Res Commun, 1992. **189**(1): p. 356-62.
6. Barnett, J.V., et al., *Cloning and developmental expression of the chick type II and type III TGF beta receptors*. Dev Dyn, 1994. **199**(1): p. 12-27.
7. Massague, J., et al., *TGF-beta receptors*. Mol Reprod Dev, 1992. **32**(2): p. 99-104.
8. Massague, J. and F. Weis-Garcia, *Serine/threonine kinase receptors: mediators of transforming growth factor beta family signals*. Cancer Surv, 1996. **27**: p. 41-64.
9. Vivien, D., et al., *Signaling activity of homologous and heterologous transforming growth factor-beta receptor kinase complexes*. J Biol Chem, 1995. **270**(13): p. 7134-41.
10. Lopez-Casillas, F., et al., *Betaglycan can act as a dual modulator of TGF-beta access to signaling receptors: mapping of ligand binding and GAG attachment sites*. J Cell Biol, 1994. **124**(4): p. 557-68.
11. Lopez-Casillas, F., J.L. Wrana, and J. Massague, *Betaglycan presents ligand to the TGF beta signaling receptor*. Cell, 1993. **73**(7): p. 1435-44.
12. Wieser, R., J.L. Wrana, and J. Massague, *GS domain mutations that constitutively activate T beta R-I, the downstream signaling component in the TGF-beta receptor complex*. Embo J, 1995. **14**(10): p. 2199-208.
13. Wrana, J.L., et al., *TGF beta signals through a heteromeric protein kinase receptor complex*. Cell, 1992. **71**(6): p. 1003-14.
14. Lagna, G., et al., *Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways*. Nature, 1996. **383**(6603): p. 832-6.
15. Miyazono, K., *TGF-beta signaling by Smad proteins*. Cytokine Growth Factor Rev, 2000. **11**(1-2): p. 15-22.
16. Heldin, C.H., K. Miyazono, and P. ten Dijke, *TGF-beta signalling from cell membrane to nucleus through SMAD proteins*. Nature, 1997. **390**(6659): p. 465-71.

17. Whitman, M., *Smads and early developmental signaling by the TGF-beta superfamily*. Genes Dev, 1998. **12**(16): p. 2445-62.
18. Massague, J. and D. Wotton, *Transcriptional control by the TGF-beta/Smad signaling system*. Embo J, 2000. **19**(8): p. 1745-54.
19. Attisano, L.-H.S.T., *The Smads*. Genomebiology, 2001. **2**(8): p. 3010.1-3010.8.
20. Lewis, K.A., et al., *Betaglycan binds inhibin and can mediate functional antagonism of activin signalling*. Nature, 2000. **404**(6776): p. 411-4.
21. comunicacion personal del Dr. Robert Weinberg.
22. Massague, J., *The transforming growth factor-beta family*. Annu Rev Cell Biol, 1990. **6**: p. 597-641.
23. Paton, B.D.W.a.I.R., *Evolutionary origins of the transforming growth factor-beta gene family*. DNA Cell Biol, 1992. **11**(7): p. 497-510.
24. Roberts, A.B., et al., *Transforming growth factors: isolation of polypeptides from virally and chemically transformed cells by acid/ethanol extraction*. Proc Natl Acad Sci U S A, 1980. **77**(6): p. 3494-8.
25. Anzano, M.A., et al., *Synergistic interaction of two classes of transforming growth factors from murine sarcoma cells*. Cancer Res, 1982. **42**(11): p. 4776-8.
26. Massague, J., et al., *Multiple type-beta transforming growth factors and their receptors*. J Cell Physiol Suppl, 1987. **Suppl**(5): p. 43-7.
27. Jakowlew, S.B., et al., *Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid encoding transforming growth factor beta 4 from chicken embryo chondrocytes*. Mol Endocrinol, 1988. **2**(12): p. 1186-95.
28. Jakowlew, S.B., et al., *Complementary deoxyribonucleic acid cloning of a novel transforming growth factor-beta messenger ribonucleic acid from chick embryo chondrocytes*. Mol Endocrinol, 1988. **2**(8): p. 747-55.
29. Denhez, F., et al., *Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-beta 3*. Growth Factors, 1990. **3**(2): p. 139-46.
30. McDonald N.Q., W.A.H., *A structural superfamily of growth factors containing a cystine knot motif*. Cell, 1993. **73**(3): p. 421-4.
31. Roberts, A.B., et al., *Purification and properties of a type beta transforming growth factor from bovine kidney*. Biochemistry, 1983. **22**(25): p. 5692-8.
32. Derynck, R., et al., *Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells*. Nature, 1985. **316**(6030): p. 701-5.
33. Kondajiah, P., et al., *Identification of a novel transforming growth factor-beta (TGF-beta 5) mRNA in Xenopus laevis*. J Biol Chem, 1990. **265**(2): p. 1089-93.
34. Roberts, A.B., et al., *Transforming growth factors from neoplastic and nonneoplastic tissues*. Fed Proc, 1983. **42**(9): p. 2621-6.
35. Sporn, M.B. and A.B. Roberts, *Autocrine, paracrine and endocrine mechanisms of growth control*. Cancer Surv, 1985. **4**(4): p. 627-32.
36. Van Obberghen-Schilling, E., et al., *Transforming growth factor beta 1 positively regulates its own expression in normal and transformed cells*. J Biol Chem, 1988. **263**(16): p. 7741-6.
37. Flanders, K.C., et al., *Transforming growth factor-beta 1: histochemical localization with antibodies to different epitopes*. J Cell Biol, 1989. **108**(2): p. 653-60.

38. Thompson, N.L., et al., *Expression of transforming growth factor-beta 1 in specific cells and tissues of adult and neonatal mice*. J Cell Biol, 1989. **108**(2): p. 661-9.
39. Kehrl, J.H., et al., *Production of transforming growth factor beta by human T lymphocytes and its potential role in the regulation of T cell growth*. J Exp Med, 1986. **163**(5): p. 1037-50.
40. Sporn, M.B. and A.B. Roberts, *Peptide growth factors are multifunctional*. Nature, 1988. **332**(6161): p. 217-9.
41. Massague, J., *How cells read TGF-beta signals*. Nat Rev Mol Cell Biol, 2000. **1**(3): p. 169-78.
42. Letterio, J.J., et al., *Maternal rescue of transforming growth factor-beta 1 null mice*. Science, 1994. **264**(5167): p. 1936-8.
43. L.P., S., *TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes*. Development, 1997. **124**(13): p. 2659-70.
44. Proetzel G., e.a., *Transforming growth factor-beta 3 is required for secondary palate fusion*. Nat Genet, 1995. **11**(4): p. 409-14.
45. Lafyatis, R., et al., *Structural and functional characterization of the transforming growth factor beta 3 promoter. A cAMP-responsive element regulates basal and induced transcription*. J Biol Chem, 1990. **265**(31): p. 19128-36.
46. Roberts, A.B., et al., *Transcriptional control of expression of the TGF-betas*. Ann N Y Acad Sci, 1990. **593**: p. 43-50.
47. Roberts, A.B., et al., *Multiple forms of TGF-beta: distinct promoters and differential expression*. Ciba Found Symp, 1991. **157**: p. 7-15.
48. Bottinger, E.P., J.J. Letterio, and A.B. Roberts, *Biology of TGF-beta in knockout and transgenic mouse models*. Kidney Int, 1997. **51**(5): p. 1355-60.
49. Roberts, A.B., et al., *Transforming growth factor type beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro*. Proc Natl Acad Sci U S A, 1986. **83**(12): p. 4167-71.
50. Falanga, V., et al., *Transforming growth factor-beta: selective increase in glycosaminoglycan synthesis by cultures of fibroblasts from patients with progressive systemic sclerosis*. J Invest Dermatol, 1987. **89**(1): p. 100-4.
51. Connor, T.B., Jr., et al., *Correlation of fibrosis and transforming growth factor-beta type 2 levels in the eye*. J Clin Invest, 1989. **83**(5): p. 1661-6.
52. Sanderson, N., et al., *Hepatic expression of mature transforming growth factor beta 1 in transgenic mice results in multiple tissue lesions*. Proc Natl Acad Sci U S A, 1995. **92**(7): p. 2572-6.
53. Ji, C., et al., *Cloning the promoter for transforming growth factor-beta type III receptor. Basal and conditional expression in fetal rat osteoblasts*. J Biol Chem, 1999. **274**(43): p. 30487-94.
54. Norgaard, P., M. Spang-Thomsen, and H.S. Poulsen, *Expression and autoregulation of transforming growth factor beta receptor mRNA in small-cell lung cancer cell lines*. Br J Cancer, 1996. **73**(9): p. 1037-43.
55. Brown, C.B., et al., *Requirement of type III TGF-beta receptor for endocardial cell transformation in the heart*. Science, 1999. **283**(5410): p. 2080-2.
56. Seed, B., and Aruffo, A., Proc. Natl. Acad. Sci. U.S.A., 1987. **84**: p. 3365-3369.

57. Cheifetz, S., et al., *Distinct transforming growth factor-beta (TGF-beta) receptor subsets as determinants of cellular responsiveness to three TGF-beta isoforms*. J Biol Chem, 1990. **265**(33): p. 20533-8.
58. Graziano Pesole, S.L., Giorgio Grillo, Flavio Licciulli, Alessandra Larizza, Wojciech Makalowski, Cecilia Saccone, *UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs*. Nucleic Acids Res, 2000. **28**(1): p. 193-196.
59. Alberts B, B.D., Lewis J, Raff M, Roberts K, Watson J., *Chapter 6. Basic Genetic Mechanisms, in Molecular Biology of the Cell*. 1994, Garland Publishing Inc.: New York and London. p. 237-238.
60. Gray NK, W.M., *Control of translation initiation in animals*. Annu Rev Cell Dev Biol, 1998(14): p. 399-458.
61. Jang SK, K.H., Nickling MJH, Duke GM, Palmenberg AC, Wimmer E., *A segment of the 5' non-translated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translayion*. J, Virol., 1988. **62**: p. 2636-43.
62. Pelletier J, S.N., *Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA*. Nature, 1988(334): p. 320-325.
63. OH, S.S.P., *Gene regulation: translational initiation by internal ribosome binding*. Curr Opin.Genet. Dev., 1993. **3**: p. 295-300.
64. Bernstein J, S.O., Ie S-Y, Elroy-Stein O., *PDGF2/c-sis mRNA leader contains a differentiation-linked internal ribosomal entry site (D-IRES)*. J.Biol.Chem., 1997. **272**: p. 9356-62.
65. Ostareck DH, O.-I.A., Wilm M, Thiele BJ, Mann M and Hentze MW., *mRNA silencing in erythroid differentiation: hnRNP K and hnRNPE1 regulate 15-lipoxygenase translation from the 3' end*. Cell, 1997. **89**: p. 597-606.
66. Vagner S. G.M., Maret A, Bayard F, Amalric F, Prats H, Prats AC., *Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes*. Mol Cell Biol., 1995. **15**((1)): p. 35-44.
67. WJohnson, D.W., et al., *Assignment of human transforming growth factor-beta type I and type III receptor genes (TGFBRI and TGFBRII) to 9q33-q34 and 1p32-p33, respectively*. Genomics, 1995. **28**(2): p. 356-7.
68. Lee, D.K., et al., *The hepatitis B virus encoded oncoprotein pX amplifies TGF-beta family signaling through direct interaction with Smad4: potential mechanism of hepatitis B virus-induced liver fibrosis*. Genes Dev, 2001. **15**(4): p. 455-66.
69. Levi-Schaffer, F., et al., *Human eosinophils regulate human lung- and skin-derived fibroblast properties in vitro: a role for transforming growth factor beta (TGF-beta)*. Proc Natl Acad Sci U S A, 1999. **96**(17): p. 9660-5.
70. Munger, J.S., et al., *The integrin alpha v beta 6 binds and activates latent TGF beta 1: a mechanism for regulating pulmonary inflammation and fibrosis*. Cell, 1999. **96**(3): p. 319-28.
71. Low MJ, H.R., Goodman RH, Habener JF, Palmiter RD, Brinster RL., , and P.P.-i.f. MEDLINE], *Tissue-specific posttranslational processing of pre-prosomatostatin encoded by a metallothionein-somatostatin fusion gene in transgenic mice*. Cell., 1985. **41**(;1)): p. 211-9.

72. Yutaka Suzuki et. Al. *Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites*. EMBO 2001 reports, Vol 2 No. 5 pp 388-399

**ESTA TESIS NO SALF  
DE LA BIBLIOTECA**

Reprinted from

BIOCHIMICA ET BIOPHYSICA ACTA

**BBA**

---

Biochimica et Biophysica Acta 1384 (1998) 189–196

Short sequence-paper

## Murine betaglycan primary structure, expression and glycosaminoglycan attachment sites

M. Verónica Ponce-Castañeda, José Esparza-López, M. Magdalena Vilchis-Landeros,  
Valentín Mendoza R., Fernando López-Casillas \*

*Departamento de Biología Celular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, México City, México*

Received 9 January 1998; accepted 25 February 1998



BIOCHIMICA ET BIOPHYSICA ACTA  
International Journal of Biochemistry, Biophysics  
& Molecular Biology

Editor-in-Chief: Peter C. van der Vliet (Utrecht, The Netherlands)

BBA comprises the following ten sections:

<u>Subject area</u>	<u>Executive Editors</u>
• Bioenergetics (including Reviews on Bioenergetics)	M. Wikström (Helsinki, Finland) P. Rich (London, UK)
• Biomembranes • Reviews on Biomembranes	A.G. Lee (Southampton, UK) P.L. Yeagle (Storrs, CT, USA)
• Lipids and Lipid Metabolism	H. van den Bosch (Utrecht, The Netherlands) D.E. Vance (Edmonton, AB, Canada)
• Protein Structure and Molecular Enzymology	G.H. Lorimer (College Park, MD, USA) Y. Lindqvist (Stockholm, Sweden)
• Molecular Cell Research	T.E. Kreis (Geneva, Switzerland) J. Avruch (Charlestown, MA, USA)
• Gene Structure and Expression	P.C. van der Vliet (Utrecht, The Netherlands) J. Gralla (Los Angeles, CA, USA)
• Molecular Basis of Disease	J.R. Riordan (Scottsdale, AZ, USA)
• General Subjects	M. Fukuda (La Jolla, CA, USA)
• Reviews on Cancer	D.M. Livingston (Boston, MA, USA)

**Submission**

Authors should submit their manuscripts (four copies) to one of the appropriate Executive Editors at either of the following Editorial Offices:

BBA Editorial Secretariat  
P.O. Box 1345  
1000 BH Amsterdam  
The Netherlands

Tel. +31 20 4853510  
Fax +31 20 4853506  
E-mail: bba@elsevier.nl

BBA Editorial Secretariat  
275 Washington Street  
Newton, MA 02158  
USA

Tel. +1 (617) 630-2244  
Fax +1 (617) 630-2245  
E-mail: bba.elsevier@cahners.com

Although most published *Review* articles are submitted on the invitation of the Executive Editor(s), authors of suitable (mini) reviews may also submit their work unsolicited (four copies) by sending their papers(s) to one of the Executive Editors of the appropriate section.

Short sequence-paper

## Murine betaglycan primary structure, expression and glycosaminoglycan attachment sites

M. Verónica Ponce-Castañeda, José Esparza-López, M. Magdalena Vilchis-Landeros, Valentín Mendoza R., Fernando López-Casillas \*

*Departamento de Biología Celular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, México City, México*

Received 9 January 1998; accepted 25 February 1998

### Abstract

The primary structure of murine betaglycan, also known as transforming growth factor beta (TGF- $\beta$ ) type III receptor, was deduced from the nucleotide sequence of a cDNA clone isolated from a heart library. Murine betaglycan is a single spanning membrane polypeptide of 850 amino acids which is highly similar to betaglycan of other species. Transfection of this cDNA into COS1 cells resulted in the expression of a membrane proteoglycan that binds TGF- $\beta$  and is recognized by antibodies raised against rat betaglycan. COS1 cells transfected with the double mutant Ser533Ala; Ser544Ala of the murine betaglycan cDNA produced a TGF- $\beta$  type III receptor devoid of glycosaminoglycan chains. © 1998 Elsevier Science B.V. All rights reserved.

**Keywords:** TGF- $\beta$ ; TGF- $\beta$  receptor; Proteoglycan; Glycosaminoglycan

Transforming growth factor beta (TGF- $\beta$ ) is the prototype of a superfamily of growth factors involved in tissue repair and in the control of cell differentiation and proliferation [1,2]. TGF- $\beta$  signals through a heteromeric receptor complex formed by the type I and II receptors, which are membrane glycoproteins with cytoplasmatic serine/threonine kinase activities. Upon ligand binding and dimerization, the type II receptor kinase phosphorylates and activates the type I receptor kinase, which phosphorylates members of the *Smad* family, a novel type of proteins which transduce the TGF- $\beta$  signal to the nucleus [3,4].

Betaglycan, also known as the TGF- $\beta$  type III receptor, as well as endoglin, are accessory members

of the TGF- $\beta$  receptor system, that modulate the interaction of TGF- $\beta$  and the signaling receptors [5–7]. Betaglycan is a membrane proteoglycan which also exists in a soluble form that results from the shedding of its ectodomain [8–10]. The membrane and the soluble forms of betaglycan regulate the access of TGF- $\beta$  to the signaling receptors in opposite ways, i.e., while the membrane form potentiates it, the soluble forms antagonizes it, therefore, they differentially regulate the effects of this important growth factor [5,6,11]. Recently, betaglycan has been involved in the regulation of the autocrine secretion of TGF- $\beta$ 1 and in the tumorigenicity suppression of human breast cancer cells [12,13].

As the first step towards studying the structure and regulation of murine betaglycan gene, we resolved to clone and sequence a cDNA clone encoding murine betaglycan. For this purpose, we probed a mouse

\* Corresponding author. Apartado Postal 70-246, México City, D.F. 04510, México. Fax: +52-5-622-56-11; E-mail: fcasilla@ifcsun1.fisiol.unam.mx

GAATTCGCGGCCGCGTGCACCGGATGGCGCAGTTGAACTGCGCTGCTGAGCTCGCGGCCGCTGCGCGCGCTGGGGGACTCGTCTGGGC 90  
TACTCCCCACCTCCGCTGGACGACCGGTCTGGACACACCGCTTGGCAGGCAAGTTGAACATTGCAGAGAAGGATCTTAAGGCTACAC 180

M A V T S H H M V P V F V 13

CCGACTTGCCCACTTGCCATCCAGCTGAAGAACCAAAGGCTGTTGGAGAGATGGCAGTGACATCCCACCACATGGTCCCTGTGTTTGT 270

L M S A C L A T A G P E P S T R C E L S P I S A S H P V Q A 43  
CTGATGAGCGCCTGCCTGGCCACCGCAGGTCCAGAGCCAGCACC CGGTGTAACATGCACCGATCAGTGCCTCTCATCCAGTCCAGGCC 360

L M E S F T V L S G C A S R G T T G L P R E V H I L N L R S 73  
CTGATGAGAGCTTACCCTGTTCTGCTGGCTGTGCCAGCAGAGGCCACCACTGGGCTGCCAAGGGAGGTTACATCCTAAACCTCCGCGT 450

T D Q G L G Q P Q R E V T L H L N P I A S V H T H H K P V V 103  
ACAGACCAAGGACTAGGCCAGCCGAGAGAGAGGTTACCCCTGCATCTGAACCCCATTCGCTCCGTGCACACTCACCACAAGCCTGTTGTG 540

F L L N S P Q P L V W H V K T E R L A A G V P R L F L V S E 133  
TTCCTGCTCAACTCCCCACAGCCCTGGTGTGGCATGTGAAGACAGAGAGACTGGCTGCTGGTGTCCCCAGACTCTTCCTGGTTTCGGAG 630

G

G S V V Q F S S G N F S L T A E T E E R S F P O E N E H L L 163  
GGTCTGTGGTCCAGTTTTCATCAGGAAACTTCTCCTTGACAGCAGAAACAGAGAAAGGAGTTTCCTCAAGAAATGAGCATCTGCTA 720

H W A Q K E Y G A V T S F T E L K I A R N I Y I K V G E D O 193  
CACTGGGCCAAAAGGAATATGGAGCAGTGACTTCATTCACCGAACTTAAGATAGCAAGAAATATCTATATTAAGTGGGAGAAGATCAA 810

V F P P T C N I G K N F L S L N Y L A E Y L O P K A A E G C 223  
GTGTTCCCTCCACGTGTAACATAGGGAAAAATTTCTCTCGCTCAATTACCTTGGCGAGTACCTTCAACCCAAAGCCGCCGAAGGTTGT 900

V L A S O P H E K E V H I I E L I S P N S N P Y S T F Q V D 253  
GTCCTGGCCAGTCCAGCCCCAGAGAAGGAAGTGATATCATTTAGTAAATCTCCCCAACCTCAATCTTACAGCACCTTCCAGGTGGAT 990

I I I D I R P A R E D P E V V K N L V L I L K C K K S V N W 283  
ATAATAATTGACATACGACTGCTCGAGAGGATCTGAGTGGTCAAAAACCTCGTCTGATCTTGAAGTGCAAAAATCTGTCAACTGG 1080

V I K S F D V K G N L K V I A P D S I G F G K E S E R S M T 313  
GTGATCAAGTCTTTGACGTCAAGGGAAACTTGAAAGTTATGTCTCTGACAGTATTGGCTTGGAAAAGAGAGTGAACGATCCATGACA 1170

V T K L V R N D Y P S T Q E N L M K W A L D N G Y S P V T S 343  
GTGACAAAATTGGTAAGAAATGACTACCTTCCACCCAAAGAAATCTGATGAAGTGGGCATTGGCAATGGCTACAGCCCAGTGACGTCA 1260

Y T I A P V A N R F H L R L E N N E E M R D E E V H T I P P 373  
TACACCATAGCTCTGTGGCCAAATAGATTTTCATCTTGGCTTGAGAACACAGGAGATGAGAGATGAGGAAGTCCACACATTCTCTCT 1350

E L R I L L G P D H L P A L D S P P F Q G E I P N G G F P F 403  
GAGCTTCGGATCCTGCTGGCCCTGACCACCTGCTGCTGCGACAGCCACCTTCCAAGGGAAATCCCAATGGAGGTTTCCCTTTT 1440

P F P D I P R R G W K E G E D R I P R P K E P I I P R V Q L 433  
CCATTCGGGATATCCCCAGGAGAGGCTGGAAGGAGGAGAAGATAGGATCCCCCGGCCAAAAGAAACCCATCATTCCAGAGTTCAATTG 1530

L P D H R E P E E V Q G V N I A L S V K C D N E K M V V A 463  
CTTCAGACCACAGAGAGCCAGAAGTGAAGGGGCGTGAATATCGCCCTGTCAGTCAATGTGACAAATGAAAAGATGGTGGTAGCT 1620

V D K D S F Q T N G Y S G M E L T L L D P S C K A K M N G T 493  
 GTAGACAAAGATTCTTCCAGACCAATGGCTACTCGGGATGGAGCTCACCCCTGTGGATCCTCTCTGCAAAGCCAAGATGAATGGTACC 1710

H F V L E S P L N G C G T R H R R S A P D G V V Y Y N S I V 523  
 CACTTGTGTTCTGGAGTCTCCGCTGAAATGGCTGTGGTACTAGACATCGGAGGTCAGCCCCAGATGGTGTGGTTACTATAACTTATTGTG 1800

V Q A P S P G D S S G W P D G Y E D L E S G D N G F P G D T 553  
 GTGACGGCTCCATCCCTCGGGATAGCAGTGGCTGGCCAGCGGTACGAAGATTTGGAGTCGGGTGATAATGGATTCTCTGGAGACACA 1890

D E G E T A P L S R A G V V V F N C S L R Q L R S P S G F Q 583  
 GATGAAGGAGAACTGCCCCCTGAGCCGTGCTGGAGTGGTAGTGTTTAACTGCAGCTTGGCCGAGCTGAGGAGTCCCAGTGGCTCCAG 1980

D Q L D G N A T F N M E L Y N T D L F L V P S P G V F S V A 613  
 GACCAGCTCGATGGAAAATGCTACCTCAATATGGAGCTGTATAACACAGACCTCTTCTGGTGCCTCCCGGGGTCTTCTCTGTGGCA 2070

E N E H V Y V E V S V T K A D Q D L G F A I O T C F I S P Y 643  
 GAGAATGAGCATGTATATGTTGAGGTGTCTGTCACTAAGCGTACCAAGATCTGGATTTGCCATCCAAACCTGCTTTATCTCTCCATAC 2160

S N P D R M S D Y T I I E N I C P K D D S V K F Y S S K R V 673  
 TCAAACCCAGACAGAATGTCTGATTACACCATCATCGAGAATCTGTCCGAAGGATGACTCTGGAAGTCTACAGCTCCAAGAGAGTG 2250

H F P I P H A E V D K K R F S F V F K S V P N T S L L F L H 703  
 CACTTCCCCATCCCACATGCTGAAGTGGACAAGAAGCGGTTACGCTTGTGTTCAAGTCCGTGTTCAACACCTCCCTGCTCTTCTCGCAC 2340

C E L T L C S R N K G S Q K L P K C V T P D D A C T S L D A 733  
 TGCGAGCTGACCTGTGCTAGGAACAAGGCTCCAGAAAGCTGCCAAAGTGTGTACTCTGATGACGCTGCACCTCTCTAGATGCC 2430

T M I W T M M Q N K K T F T K P L A V V L Q V D Y K E N V P 763  
 ACCATGATCTGGACCATGATCGAGAATAAGAAGACATTCACCAAGCCCTGGCCGTGCTCCAGATGAGTATATAAAGAAAATGTTCCA 2520

N M K E S S P V P P P P Q I F H G L D T L T V M G I A F A A 793  
 AACATGAAGGAGTCCAGTCCGGTCTCTCTCCACACAGATTTCCACGGCCTGGACAGCTCACCGTATGGGCATTGGGTTTGCAGCA 2610

F V I G A L L T G A L W Y I Y S H T G E T A R R Q Q V P T S 823  
 TTTGTGATCGGAGCACTCTGACGGGGCTTGTGGTATATCTACTCCACACAGGGGAGAGCGGCAGGAAGGACGAAGTCCCTACCTCG 2700

P P A S E N S S A A H S I G S T Q S T P C S S S S T A stop 850  
 CCACCAGCTCGGAGAACAGCAGCGCAGCCACAGCATAGGGAGCACTCAGAGCACTCCCTGCTTAGCAGCAGCAGCGCCTAGGTGGAC 2790

AGATGGACCGCCTGCCTAGTGGACAGACGGATGCCTGCCACCGCCGGGGCAGGGCCAGATGCCGATGCTGGGTGCCAGACTCAGAAG 2880  
 GTTGTATCTGGTTCCTCTGTAAGAGAGAGTGAATTCAGTATACAGACAGCCTGCTCTACCCACCCCTCTCTTACCACGGCCACGTAA 2970  
 ATGTGACCCCTGACATCTGTACACGAAAGCTAAGCTGGTGGCCTCCCTGCCAGCCCTCGCAGGATGGGGGTTCAAATGTGAACATA 3060  
 TGCCAGTTTTGTCTTTTTTTTTTTTTATTTTTATGCTGCTTTGTCCAGGTGTCCAAACATCCATCATTGGGGTCTGATGTTTTAAA 3150  
 GAGTAAAGGAGACGGTGAAGGATGTCAGCTAGAGCGTAGAGCCAGGGAGAGACAGCCAGGGATPCTCGCCTAGCTGAACCAAGATGTA 3240  
 AATAGGAGAAATGCTCCCTCCATAGCCAATGTCATCTCGCTCTGGATGCCACTCCTGACCTGCCAGCCCTACATGGTCCCTGTGTA 3330  
 GCTTGTAGCCTTGGGGCCTGTCACTTGCCAAAGTAACATTCCTCTTGGGCCAAAATGAATGACTGTGATCTCCGATCCTGCCACAAG 3420  
 ACACCAACTCCAAAAGTGCCTTAAAGCCAGGTTTGTGCTCCTTTCCAGGAGAGAGCGTGAACGTCAGATTTGGCTGTCTCGAACACT 3510  
 CTGTCTAATCATACTTAAACATCAGGCAGGTTGGGGTTCAGTGGTAAGCGAAGGATTTATAGCAAGTAAATGCTGTGTATGGTCAAT 3600  
 AGGCACAAATGCAGAACCCAGCACATTTGGAGAGACGACCACTTCTCTTTAAAATGTCAATTCCTTCCAGAGTCGGAACCTGAGTTAT 3690  
 ATGTGGTATATATGACTCCAACCATCAGCCTTTCTGATAGCCTTTAGAGAAGCTAGAGAGCCCGGGATCTGATGATGTCCAG 3780  
 TGCTGCTCAGATGATACAGAGACAGGAGTTTCAGGTAACTTCTTACAGAAATGAAGTGTATGCCACTTGCCTTTGCTGAGTCCAG 3870  
 TGACTTTCAGTGTAAACTGAACCCAGAGGGATTCCTCATGCCAGATCCCGTGCAAATTAATAATACAAATTTCTAGAGTCCGTC 3960  
 GACCGCGCCGCAATTC 3977

heart cDNA library (Clontech's #ML5002b) with the rat betaglycan cDNA [9]. Our screening yielded several clones, one of them,  $\lambda$ CI12, had an insert of 3977 base pairs. The nucleotide (nt) sequence of  $\lambda$ CI12 insert revealed an open reading frame (ORF) of 2550 nt flanked by 231 and 1193 nt at the 5' and 3' untranslated regions, respectively. This ORF predicts a polypeptide of 850 amino acids with two hydrophobic stretches, one at the amino terminal end and the other close to the carboxyl end, corresponding to the signal peptide and the transmembrane regions, respectively, of the receptor core protein (Fig. 1). Thus, although the  $\lambda$ CI12 clone only accounts for approximately two thirds of betaglycan mRNA [9,15], it includes the complete coding region of mouse betaglycan.

With our reported mouse sequence, there are now a total of five species for which the betaglycan primary structure is known [16,17]. Amino acid sequence alignment with the reported betaglycan polypeptides reveals that mouse betaglycan is 94.3% identical to the rat, 81.8% to the human, 80.3% to the pig and 66.5% to the chicken betaglycan (Fig. 2). In all these species, betaglycan contains 17 cysteines that are 100% conserved. This is particularly relevant for the potential disulfide bonds arrangements of the 16 cysteines of the ectodomain of the receptor, since they may determine its proper folding and, therefore, its ligand binding properties. Importantly, the regions of maximal sequence similarity among the known

betaglycans are the transmembrane and the cytoplasmic regions, suggesting important, although still unknown, biological roles for these portions of the receptor. Comparison of the mouse betaglycan sequence against the GenBank database, revealed similarities already described between portions of betaglycan and endoglin ectodomains [16,18], and between betaglycan ectodomain and a group of extracellular proteins, which include uromodulin and zona pellucida glycoproteins [19]. However, the most remarkable similarity was found between the transmembrane and intracellular regions of mouse betaglycan and endoglin (Fig. 3), which further emphasizes the potential functional relevance of these regions in these TGF- $\beta$  binding cell surface molecules.

In order to test whether or not  $\lambda$ CI12 contains a bona fide functional murine betaglycan, we tested its TGF- $\beta$  binding ability by transient expression in COS1 cells. For this purpose, we subcloned its insert in the pcDNA3 expression vector (Invitrogen). The resulting construct (pcDNA3/CI12) has the  $\lambda$ CI12 insert under the control of the strong constitutive cytomegalovirus (CMV) promoter. COS1 cells transfected with pcDNA3/CI12 and the empty vector were subjected to affinity labeling with  $^{125}$ I-TGF- $\beta$ 1. The labeled receptors were then subjected to immunoprecipitation with rabbit serum #822, which contains antibodies raised against the rat betaglycan ectodomain [5], and then separated by SDS/PAGE and revealed by autoradiography (Fig. 4).

Fig. 1. Nucleotide sequence of  $\lambda$ CI12 cDNA clone and the predicted primary structure of murine betaglycan. The insert of  $\lambda$ CI12 was completely sequenced (GenBank Accession Number: AF039601) in both strands by the Sanger's chain termination method [14] using specific oligonucleotide primers and a commercially available kit (Thermosequenase, Amersham). Flanking *Eco*RI and *Not*I sites correspond to the vector cloning sites. The ORF, which starts with the ATG at nt 232 and finishes with the first in-frame stop codon at nt 2782, predicts a 850-amino acid polypeptide that is 94.3% identical to rat betaglycan. Hydrophobic regions corresponding to the putative signal peptide (residues 8–23) and transmembrane domain (residues 783–807) are indicated with a double underline. The putative sites for proteolytic cleavage of the ectodomain (Lys<sup>743</sup>Lys and Leu<sup>750</sup>AlaValVal) are shown with a broken underline. Canonical sites for Asn-linked glycosylation are indicated by @. Six potential sites for the attachment of the glycosaminoglycan chains (Ser–Gly dipeptides) were found (underlined); those that are utilized in the COS1 cells (see text and Fig. 5) are indicated by  $\$$ .

Fig. 2. Alignment of amino acid sequences of mouse, rat, human, pig and chicken betaglycans. The amino acid sequence of murine betaglycan was aligned with the rat (GenBank accession #M77809), human (#L07594), pig (#L07595) and chick (#L01121) betaglycan sequences. The particular numbering for each of these sequences are indicated at the right margin. Only the residues that differ from the murine sequence are shown. Gaps (–) were introduced in some places to optimize the alignment. All the cysteines (&), but only some of the predicted sites for Asn-linked glycosylation (@) and for glycosaminoglycan chains attachment (\$) in murine betaglycan were conserved in the species compared. Only one of the putative sites for proteolytic cleavage of the ectodomain, the Lys<sup>743</sup>Lys (double underline) is conserved in these species. The regions corresponding to the transmembrane and cytoplasmic domains (underlined) showed the highest degree of identity among the known betaglycans.

TESIS CON  
FALLA DE ORIGEN







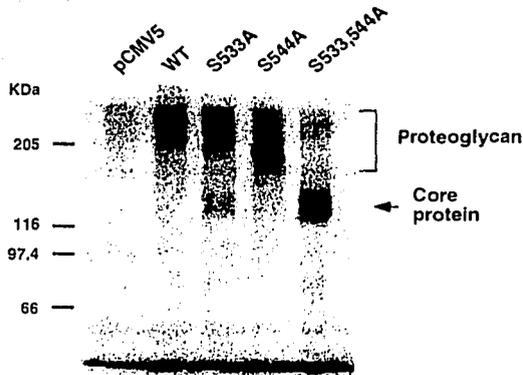


Fig. 5. Affinity labeling of the murine betaglycan mutants deficient in GAG chains. Murine betaglycan mutant cDNAs S533A, S544A and the double mutant S533,544A, were obtained by a Polymerase Chain Reaction-based site directed mutagenesis [26], confirmed by nucleotide sequencing and subcloned downstream the CMV promoter of pCMV5 [27]. COS1 cells were transiently transfected with the mutant, wild type betaglycan (WT), and empty pCMV5 plasmids, affinity labeled with 100 pM  $^{125}$ I-TGF- $\beta$ 2 (TGF- $\beta$ 2, Ciba Geigy, Switzerland), and the detergent cells extracts were separated by SDS/PAGE and revealed by exposure in a phosphor screen (Storm, Molecular Dynamics). Molecular weight markers in kilodaltons are shown.

proteoglycan form, showing only a small proportion of the core protein. Individually, both mutants S533A and S544A produced receptors with an increased proportion of the core protein; however, the level of conversion to the proteoglycan form, although reduced, was still high. On the other hand, the double mutant S533,544A, practically abolished the production of the proteoglycan form, yielding only the betaglycan core protein doublet at approximately 125–130 KDa. These results indicate that the major GAG attachment sites in the rat and the mouse betaglycan are structurally and functionally conserved.

We thank Dr. N. Cerletti for his kind gift of TGF- $\beta$ 2, and Dr. J. Aguirre and Dr. J.A. García-Sáinz for their critical reading of the manuscript. This work was supported in part by grants #IN200495 of DGAPA, UNAM and #2364P-N of CONACyT. F.L.-C. is an International Research Scholar of the Howard Hughes Medical Institute.

## References

- [1] A.B. Roberts, M.B. Sporn, *Growth Factors* 8 (1993) 1–9.
- [2] D.M. Kingsley, *Genes Dev.* 8 (1994) 133–146.
- [3] L. Attisano, J.L. Wrana, *Cytokine Growth Factor Rev.* 7 (1996) 327–339.
- [4] J. Massagué, F. Weis-Garcia, *Cancer Surv. Cell Signal.* 27 (1996) 41–64.
- [5] F. López-Casillas, J.L. Wrana, J. Massagué, *Cell* 73 (1993) 1435–1444.
- [6] F. López-Casillas, H.M. Payne, J.L. Andres, J. Massagué, *J. Cell Biol.* 124 (1994) 557–568.
- [7] P. Lastres, A. Latemendía, H. Zhang, C. Rius, N. Almendro, U. Raab, L.A. López, C. Langa, A. Fabra, M. Letarte, C. Bernabéu, *J. Cell Biol.* 133 (1996) 1109–1121.
- [8] J.L. Andres, K. Stanley, S. Cheifetz, J. Massagué, *J. Biol. Chem.* 109 (1989) 3137–3145.
- [9] F. López-Casillas, S. Cheifetz, J. Doody, J.L. Andres, W.S. Lane, J. Massagué, *Cell* 67 (1991) 785–795.
- [10] J. Arribas, F. López-Casillas, J. Massagué, *J. Biol. Chem.* 272 (1997) 17160–17165.
- [11] S. Sankar, N. Mahooti-Brooks, M. Centrella, T.L. McCarthy, J.A. Madri, *J. Biol. Chem.* 270 (1995) 13567–13572.
- [12] C. Chen, X.-F. Wang, L. Sun, *J. Biol. Chem.* 272 (1997) 12862–12867.
- [13] L. Sun, C. Chen, *J. Biol. Chem.* 272 (1997) 25367–25372.
- [14] F. Sanger, A.R. Coulson, B.G. Barrell, A.J.H. Smith, B.A. Roe, *J. Mol. Biol.* 143 (1980) 161–178.
- [15] X.-F. Wang, H.Y. Lin, E. Ng-Eaton, J. Downward, H.F. Lodish, R.A. Weinberg, *Cell* 67 (1991) 797–805.
- [16] A. Moren, H. Ichijo, K. Miyazono, *Biochem. Biophys. Res. Commun.* 189 (1992) 356–362.
- [17] J.V. Barnett, A. Moustakas, W. Lin, X.-F. Wang, H.Y. Lin, J.B. Galper, R.L. Maas, *Dev. Dyn.* 199 (1994) 12–27.
- [18] S. Cheifetz, T. Bellón, C.S.V. Calés, C. Bernabeu, J. Massagué, M. Letarte, *J. Biol. Chem.* 267 (1992) 19027–19030.
- [19] P. Bork, C. Sander, *FEBS Lett.* 300 (1992) 237–240.
- [20] S. St-Jacques, U. Cymerman, N. Pece, M. Letarte, *Endocrinology* 134 (1994) 2645–2657.
- [21] S. Cheifetz, A. Bassols, K. Stanley, M. Ohta, J. Greenberger, J. Massagué, *J. Biol. Chem.* 263 (1988) 10783–10789.
- [22] P.R. Segarini, S.M. Seyedin, *J. Biol. Chem.* 263 (1988) 8366–8370.
- [23] S. Cheifetz, J.L. Andres, J. Massagué, *J. Biol. Chem.* 263 (1988) 16984–16991.
- [24] J. Andres, D. DeFalcis, M. Noda, J. Massagué, *J. Biol. Chem.* 267 (1992) 5927–5930.
- [25] L. Zhang, J.D. Esko, *J. Biol. Chem.* 269 (1994) 19295–19299.
- [26] B. Cormack, in: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds.), *Current Protocols in Molecular Biology*, Vol. 1, Wiley, New York, 1997, pp. 8.5.1–8.5.10.
- [27] S. Andersson, D.N. Davis, H. Dahlback, H. Jornvall, D.W. Russell, *J. Biol. Chem.* 264 (1989) 8222–8229.

### Publication information

*Biochimica et Biophysica Acta (Protein Structure and Molecular Enzymology)* (ISSN 0167-4838). For 1998 volumes 1382–1388 are scheduled for publication. *Biochimica et Biophysica Acta* full subscription for 1998 consists of Volumes 1363–1408. Subscription prices are available upon request from the Publisher. Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by surface mail except to the following countries where Air delivery via SAL mail is ensured: Argentina, Australia, Brazil, Canada, Hong Kong, India, Israel, Japan, Malaysia, Mexico, New Zealand, Pakistan, PR China, Singapore, South Africa, South Korea, Taiwan, Thailand, USA. For all other countries airmail rates are available upon request. Claims for missing issues should be made within six months of our publication (mailing) date.

**Orders, claims, and product enquiries:** please contact the Customer Support Department at the Regional Sales Office nearest you:

**New York,** Elsevier Science, P.O. Box 945, New York, NY 10159-0945, USA. Tel: (+1) 212-633-3730, [Toll free number for North American Customers: 1-888-4ES-INFO (437-4636)], Fax: (+1) 212-633-3680, E-mail: usinfo-f@elsevier.com

**Amsterdam,** Elsevier Science, P.O. Box 211, 1000 AE Amsterdam, The Netherlands. Tel: (+31) 20-485-3757, Fax: (+31) 20-485-3432, E-mail: nlinfo-f@elsevier.nl

**Tokyo,** Elsevier Science, 9-15, Higashi-Azabu 1-chome, Minato-ku, Tokyo 106, Japan. Tel: (+81) 3-5561-5033, Fax: (+81) 3-5561-5047, E-mail: info@elsevier.co.jp

**Singapore,** Elsevier Science, No. 1 Temasek Avenue, #17-01 Millenia Tower, Singapore 039192. Tel: (+65) 434-3727, Fax: (+65) 337-2230, E-mail: asiainfo@elsevier.com.sg

### Enquiries concerning manuscripts and proofs

Questions arising after acceptance of the manuscript, especially those relating to proofs, should be directed to Issue Manager, *BBA*, Elsevier Science B.V., P.O. Box 2759, 1000 CT Amsterdam, The Netherlands. Courier address: Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands. Tel. (+31) 20 485 3273; Fax: (+31) 20 485 2775.

**Authors in Japan please note:** Upon request, Elsevier Science Japan will provide authors with a list of people who can check and improve the English of their paper (*before submission*). Please contact our Tokyo office: Elsevier Science Japan, 1-9-15 Higashi-Azabu, Minato-ku, Tokyo 106; Tel. (03)-5561-5032; Fax (03)-5561-5045.

BBA has no page charges and 50 offprints are provided free to the corresponding author of each published paper.

### Colour figures

Colour figures are published at the author's expense: 750 Dutch guilders for the first page and 500 Dutch guilders for each subsequent page, excluding sales tax.

This journal is in the *ADONIS* Service, whereby copies of individual articles can be printed out from CD-ROM on request. An explanatory leaflet can be obtained by writing to *ADONIS* B.V., P.O. Box 17005, 1001 JA Amsterdam, The Netherlands.

*US mailing notice – Biochimica et Biophysica Acta (Protein Structure and Molecular Enzymology)* (ISSN 0167-4838) is published monthly, semimonthly in April and September (total 14 issues) by Elsevier Science B.V. (Molenwerf 1, Postbus 211, 1000 AE Amsterdam, The Netherlands).

Annual subscription price in the USA US\$ 2292.00 (valid in North, Central and South America), including air speed delivery. Periodicals postage paid at Jamaica, NY 11431.

USA POSTMASTER: Send address changes to *Biochimica et Biophysica Acta (Protein Structure and Molecular Enzymology)*, Publications Expediting, Inc., 200 Meacham Avenue, Elmont, NY 11003.

AIRFREIGHT AND MAILING in the USA by Publications Expediting Inc., 200 Meacham Avenue, Elmont, NY 11003.