

00377



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

POSGRADO EN CIENCIAS BIOLOGICAS

FACULTAD DE CIENCIAS

AISLAMIENTO Y CARACTERIZACION DE LA REGION PROMOTORA DEL GEN DEL SARCOGLICANO EPSILON

TESIS

QUE PARA OBTENER EL GRADO ACADEMICO DE
MAESTRO EN CIENCIAS BIOLOGICAS
(BIOLOGIA EXPERIMENTAL)

PRESENTA:

ADRIAN FLORES SEGURA

DIRECTORA DE TESIS: **DRA. HAYDEE ROSAS VARGAS**

MEXICO, D. F.

ABRIL 2005

m343340



COORDINACIÓN



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.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS COORDINACIÓN

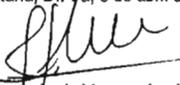
Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar, UNAM
Presente

Por medio de la presente me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 7 de marzo del 2005, se acordó poner a su consideración el siguiente jurado para el examen de grado de Maestría en Ciencias Biológicas (Biología Experimental) del(a) alumno(a) Flores Segura Adrián con número de cuenta 93073918 con la tesis titulada: "Aislamiento y caracterización de la región promotora del gen del sarcoglicano epsilon", bajo la dirección del(a) Dra. Haydeé Rosas Vargas.

Presidente:	Dr. Alejandro García Carrancá
Vocal:	Dr. Alejandro Zentella Dehesa
Secretario:	Dra. Haydeé Rosas Vargas
Suplente:	Dr. Ramón Mauricio Coral Vázquez
Suplente:	Dr. Guillermo Manuel Ceballos Reyes

Sin otro particular, quedo de usted.

Atentamente
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., 8 de abril del 2005


Dr. Juan José Morrone Lupi
Coordinador del Programa

c.c.p. Expediente del interesado

Agradezco el apoyo recibido durante la realización de este trabajo por parte de Consejo Nacional de Ciencia y Tecnología (**CONACyT**), a través de la beca **183196**.

Comité Tutoral:

Dra. Haydeé Rosas Vargas
Dr. Alejandro Zentella Dehesa
Dr. Ramón Mauricio Coral Vázquez

ÍNDICE

Abreviaturas	3
Figuras y tablas	4
Resumen	5
1. Introducción	6
1.1 Complejo Distrofina-Glicoproteínas	6
1.2 El complejo SG-SPN	8
1.3 Expresión del complejo SG-SPN	10
1.4 Regulación transcripcional de los sarcoglicanos	11
1.5 El Sarcoglicano Epsilon	13
2. Planteamiento del problema	16
3. Hipótesis	17
4. Objetivos	17
5. Estrategia experimental	18
6. Material y métodos	18
6.1 Tamizaje de una genoteca	18
6.2 Marcaje de la sonda	19
6.3 Extracción de ADN	20
6.4 Secuenciación	20
6.5 Diseño de oligonucleótidos	21
6.6 Amplificación de los fragmentos escalonados por PCR	22
6.7 Obtención de bacterias competentes	22
6.8 Generación de extremos cohesivos	23
6.9 Ligación y transformación	23
7. Resultados	26
7.1 Aislamiento y secuenciación de la región promotora del gen de <i>SGCE</i>	26
7.2 Construcción de los vectores de expresión	27
7.3 Análisis de la secuencia con los programas <i>Promoter Inspector</i> y <i>MatInspector</i>	32
7.4 Análisis comparativo de las regiones promotoras de α -SG y ϵ -SG	32
8. Discusión	37
9. Conclusiones	41
10. Referencias	42

ABREVIATURAS

aa	Aminoácido
ADN	Ácido desoxirribonucleico
DGC	Complejo distrofina glicoproteínas
DMD/DMB	Distrofia muscular Duchenne/Becker
h	Horas
kb	Kilobases
kDa	Kilodaltones
LGMD	Distrofia muscular de cinturas
min	Minutos
ml	Mililitros
NF1	Factor nuclear 1
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa
pmol	Picomoles
mARN	Ácido ribonucleico mensajero
rpm	Revoluciones por minuto
RT-PCR	Transcriptasa reversa acoplada a la reacción en cadena de la polimerasa
seg	Segundo
SG	Sarcoglicano
SGCE	Gen del sarcoglicano epsilon
SG-SPN	Complejo sarcoglicano sarcospán
SNC	Sistema nervioso central
SPN	Sarcospán
SRF	Factor de respuesta a Suero
TBP	Proteína de unión a caja TATA

FIGURAS Y TABLAS

Figura 1. Complejo Distrofina-Glicoproteínas	Pág. 7
Figura 2. Complejo Sarcoglicano-Sarcospán	Pág. 8
Figura 3. Mapas de los vectores utilizados	Pág. 25
Figura 4. Identificación de las clonas que hibridaron en el tamizaje radioactivo	Pág. 26
Figura 5. Secuencia del extremo 5' del gen <i>SGCE</i>	Pág. 28
Figura 6. Secuencia completa del extremo 5' del gen <i>SGCE</i>	Pág. 29
Figura 7. El promotor de ϵ -SG presenta un alto contenido de dinucleotidos C/G y tres islas CpG	Pág. 30
Figura 8. Amplificación de los fragmentos escalonados a partir de la secuencia de 2.4 kb del gen de SG- ϵ para la construcción de los vectores de expresión	Pág. 31
Figura 9. Amplificación con el oligonucleótido 2.F a partir de las construcciones generadas con el vector pDrive y el fragmento de 226 pb	Pág. 31
Figura 10. Análisis de las 2.4 kb analizadas hacia el extremo 5' gen <i>SGCE</i>	Pág. 33
Figura 11. Análisis comparativo de los elementos reguladores comunes, presentes en los promotores del gen de SG- α y el gen <i>SGCE</i> , con el programa DiAlign TF	Pág. 34
Tabla 1. Oligonucleótidos utilizados para generar los fragmentos con deleciones progresivas a partir del promotor del gen <i>SGCE</i>	Pág. 21

Resumen

El sarcoglicano epsilon es una proteína de expresión ubicua constituyente del complejo sarcoglicano-sarcospán. Se han identificado diversas mutaciones en el gen del sarcoglicano epsilon que causan distonia mioclónica, un desorden del sistema nervioso central. El patrón de expresión de epsilon sarcoglicano difiere de acuerdo al tipo de tejido y al estadio de diferenciación, además de estar sujeto al mecanismo de impronta como mecanismo de regulación transcripcional. En este trabajo se clonó y secuenció la región promotora del gen del sarcoglicano epsilon de ratón. Dentro de la secuencia obtenida, se logró identificar una región no reportada de 646 bases, la cual muestra un alto contenido de C/G, con lo cual se completaron 2.4 kb hacia el extremo 5' del sitio de inicio de la transcripción del gen del sarcoglicano epsilon. Mediante el uso del programa *MatInspector* se identificó una posible región promotora de 301 bases que coincide con la ubicación de la región rica en C/G, dentro de la cual se han localizado múltiples elementos de reconocimiento a factores de transcripción tales como SP1, NF1 y MAZ los cuales tienen una actividad reguladora relevante en otros promotores. Además se identificaron 156 dinucleótidos C/G, los cuales en su conjunto constituyen tres islas CpG que se encuentran contenidas dentro de la región promotora del gen de epsilon sarcoglicano. Finalmente, no se identificaron sitios de reconocimiento para factores de transcripción comunes en las regiones promotoras de los sarcoglicanos alfa y epsilon.

1. INTRODUCCIÓN

1.1 Complejo Distrofina-Glicoproteínas

El complejo distrofina-glicoproteínas (DGC, por sus siglas en inglés) está constituido por proteínas integrales de membrana y periféricas, que establecen una unión entre el citoesqueleto y la matriz extracelular (Ibraghimov-Beskrovnaya et al., 1992; Crosbie et al., 1999) (Figura 1). Por su ubicación, el DGC puede subdividirse en tres conjuntos: los componentes intracelulares integrados por las sintrofinas α , β -1, β -2, γ -1 y γ -2, la distrofina, α y β -distrobrevina; los componentes transmembranales entre los que se encuentran el distroglicano (DG) $-\beta$ y el complejo sarcoglicano-sarcospán (SG-SPN); y finalmente la región extracelular, donde se encuentra el DG- α (Campbell, 1995; Blake et al., 2002). El complejo DG es esencial dentro de este grupo de proteínas, el DG- α se une a la matriz extracelular a través de laminina- α 2 y establece interacción con el citoesqueleto por medio del DG- β . El extremo carboxilo terminal del DG- β se encuentra unido al dominio carboxilo terminal, rico en cisteína, de la distrofina y esta a su vez interactúa con los filamentos de actina a través de su dominio amino terminal (Blake et al., 2002). La interacción de los componentes del DGC es importante para el establecimiento de un puente de interacción entre la matriz extracelular y el citoesqueleto, así como para la transducción bioquímica y mecánica de señales. (Rando, 2001; Allikian et al., 2004; Lapidos et al., 2004).

La deficiencia de algunos de los componentes del DGC, como resultado de mutaciones en los genes que los codifican, se encuentra relacionada a diversos tipos de distrofias musculares (Emery, 2002). Un ejemplo representativo es el de las mutaciones en el gen de distrofina, que son causantes de la distrofia muscular de Duchenne/Becker (DMD/BMD) y en algunos casos de cardiomiopatía dilatada (Lapidos et al., 2004).

Si bien se ha caracterizado con mayor detalle la composición y posibles funciones del DGC en músculo, la expresión de la mayoría de los componentes

del DGC no está restringida a este tipo de tejido, ya que se ha observado la expresión de algunas proteínas del DGC en tejidos no musculares. Así ocurre en el sistema nervioso central (SNC) donde se ha observado la expresión de la distrofina y los DGs, entre otras (Imamura et al., 2000; Moukhles y Carbonetto, 2001; Lévi, et al., 2002). La función del DGC en el SNC no es clara, sin embargo, un tercio de los pacientes con DMD sufren un déficit cognitivo, aparentemente relacionado con alteraciones en la memoria a largo plazo (Moukhles y Carbonetto, 2001; Vaillend et al., 2004).

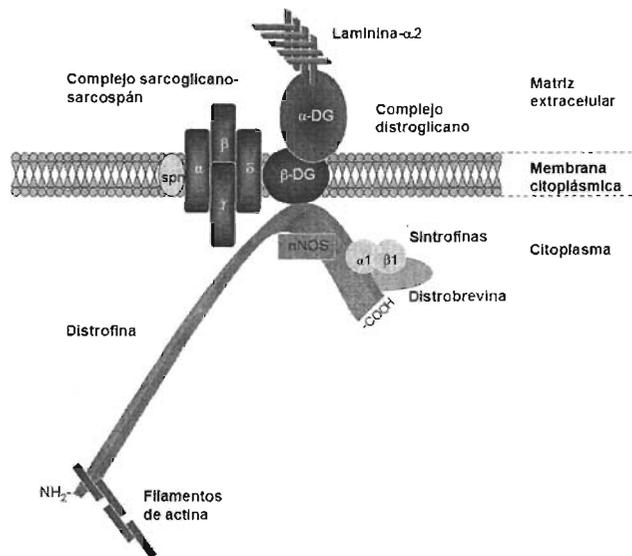


Figura 1. Complejo Distrofina-Glicoproteínas (DGC). Representación de los componentes del complejo distrofina glicoproteínas, entre los que se distinguen los componentes extracelulares como el DG- α ; los componentes transmembranales como son el complejo SG-SPN, compuesto por las cuatro subunidades de los SG's α , β , γ , y δ además de SPN, y DG- β ; y los componentes intracelulares como son distrofina, distrobrevina y α -1 y β -1 sintrofinas, además de la nNOS, proteína asociada al DGC (Modificado de: Cohn y Campbell, 2000).

1.2 El Complejo SG-SPN

El complejo SG-SPN (Figura 2) es un agrupamiento multimérico transmembranal constituido por una familia de proteínas denominadas sarcoglicanos (SG's), a las cuales se asocia la proteína sarcospán (SPN). El complejo SG-SPN expresado en la membrana muscular se ha descrito como un complejo tetramérico (sin considerar al SPN), el cual está constituido por los SG's α , β , γ , y δ con masas moleculares de 50, 43, 35 y 35 kDa respectivamente (Hack, et al., 2000). Los genes de los SG's α , β , γ , y δ se localizan en los loci 17q21, 4q12, 13q12 y 5q33 respectivamente (Bonneman et al., 1995; Lim et al., 1995; Nigro et al., 1996; Noguchi et al., 1995; Roberds et al., 1994). En 1997 se descubrió un homólogo del SG- α que fue designado como SG- ϵ , de 45 kDa, cuyo gen se localiza en el locus 7q21 (Ettinger et al., 1997; McNally et al., 1998). Wheeler et al. (2002) han descrito la presencia de un sexto SG denominado SG- ζ , de aproximadamente 40 kDa, cuyo gen se encuentra localizado en el locus 8p22 humano.

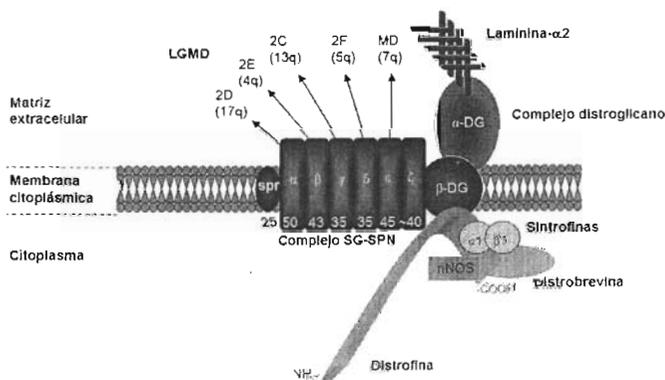


Figura 2. Complejo Sarcoglicano-Sarcospán. Representación de los seis SG's y SPN, señalándose el tipo de distrofia muscular de cinturas (LGMD2 C-F) que causa la deficiencia de α , β , γ y δ -SG's y Distrofia Mioclónica (MD) (DYT11) en el caso de ϵ -SG. El peso en kDa de cada uno de los SG's y SPN se señala con los números debajo de cada proteína. (Modificado de: Cohn y Campbell, 2000)

Sarcospán, por otra parte, es una proteína de 25 kDa con cuatro dominios transmembranales; cuyos extremos amino y carboxilo se ubican en la región citoplásmica y forma parte de la familia de proteínas tetraspan (Crosbie et al., 1997; Lebakken et al., 2000). Se asocia bioquímicamente a los SG's, y la deficiencia de alguno de éstos causa una disminución secundaria de SPN. El gen de esta proteína se encuentra en el locus 12p11.2 (Rando, 2001).

Con base en su localización subcelular, los SG's α y ϵ se clasifican como proteínas transmembranales tipo I (con el extremo amino terminal localizado en la región extracelular), mientras que los SG's β , γ , δ y ζ se clasifican como proteínas transmembranales tipo II (el extremo amino terminal se encuentra en la región intracelular) (Hack, et al., 2000). Es notable la alta homología entre los SG's α y ϵ (62%), γ y δ (69%) (Hack, et al., 2000); así como de los SG's γ y δ con el SG- ζ (74%) (Wheeler et al., 2002), en tanto que el SG- β guarda una baja homología con estos últimos (Hack, et al., 2000; Wheeler et al., 2002).

Se ha descrito la existencia de mutaciones en los genes de los SG's α , β , γ y δ que causan distrofias musculares de cinturas (LGMD, por sus siglas en inglés) de tipo LGMD2D, LGMD2E, LGMD2C y LGMD2F, respectivamente (Figura 2), denominadas en su conjunto como sarcoglicanopatías (Hack et al., 2000). Por otra parte, no se ha relacionado algún tipo de distrofia con mutaciones en el gen de SPN (Cohn y Campbell, 2000) ó SG- ϵ (Hack et al., 2000), mientras que en los modelos animales deficientes de γ y δ SG's la expresión del SG- ζ se reduce, observándose mayor decremento de esta proteína en el modelo deficiente del SG- δ (Wheeler et al., 2002), sin embargo, hasta la fecha, no se ha relacionado ninguna mutación en este gen directamente con algún tipo de distrofia.

La principal función adjudicada al complejo SG-SPN es la de estabilizar al DGC, particularmente en la interacción entre distrofina y β -DG, durante el proceso de contracción de la fibra muscular (Yoshida et al., 2000). Por otra parte, se cree que el complejo SG-SPN puede funcionar como receptor para algún ligando

extracelular a través de los SG's β , γ y δ , mientras que SG- α podría actuar como un efector secundario, en la posible participación del complejo SG-SPN en algún proceso de transducción de señales (Chan et al., 1998).

1.3 Expresión del complejo SG-SPN

Los componentes del complejo SG-SPN se expresan diferencialmente en un patrón espacio-temporal definido para cada uno de ellos. El SG- α se expresa exclusivamente en músculo estriado (Roberds et al., 1993), mientras que la expresión del SG- γ además de músculo estriado se ha observado también en músculo liso (Barresi et al., 2002). Por otra parte, los SG's β y δ además de observarse en músculo, se expresan en bajos niveles en tejidos no musculares, tales como endotelio, cerebro, y retina entre otros (Ramírez et al., 2004; Imamura et al., 2000; Estrada-Mena, com. pers.). De manera interesante la expresión del SG- ϵ se ha observado en todos los tejidos en los que se ha estudiado, por lo que se considera de expresión ubicua (Ettinger et al., 1997; McNally et al., 1998; Ramírez et al., 2004).

De acuerdo con estos patrones de expresión para cada una de las subunidades del complejo SG-SPN, se ha descrito la formación de dos complejos coexistentes en músculo estriado, uno constituido por los SG's α , β , γ y δ y otro formado por los SG's ϵ , β , γ y δ (Straub et al., 1999; Liu y Engvall, 1999; Cohn y Campbell, 2000). Mientras que en músculo liso las subunidades que forman el complejo son los SG' ϵ , β , γ y δ (Barresi et al., 2002). No obstante, Wheeler y McNally (2003) proponen que sea el SG- ζ y no γ -SG el que esté presente en músculo liso, argumentando que ambas proteínas puedan ser reconocidas por los mismos anticuerpos. Sin embargo, en nuestro laboratorio, la detección a nivel de transcrito del SG- ζ solo se ha logrado a través de RT-PCR anidado en diversos tejidos, en tanto que la expresión de γ -SG en músculo liso vascular se detecta mediante RT-PCR convencional (Ramírez et al., 2004).

Una vez que son sintetizados en el retículo endoplásmico, los diferentes componentes del complejo SG-SPN se asocian en un orden establecido; β -SG se asocia con δ -SG, una vez formado este heterodímero se asocia γ -SG, formando el heterotrímero β - δ - γ , al cual se une rápidamente α -SG (Noguchi, et al., 2001; Shi et al., 2004). La formación del heterodímero β - δ es necesario para el correcto ensamble y localización en la membrana del complejo SG-SPN (Shi et al., 2004). De hecho, mediante análisis de inmunofluorescencia en biopsias de pacientes con alguna sarcoglicanopatía, se ha observado que las mutaciones en alguno de los SG's no sólo disminuye la expresión del SG directamente afectado, sino que además causa deficiencia en el sarcolema de los otros SG's del complejo, sobre todo cuando se trata de mutaciones en los genes de β ó δ SG's (Liu y Engvall, 1999; McNally et al., 1998; Araishi et al., 1999; Crosbie et al., 2000). Todo esto resalta la importancia de la integridad del complejo para que pueda llegar a la membrana celular. Este control sobre el correcto ensamble y la estequiometría del complejo SG-SPN es similar a lo que se ha observado con otros complejos proteínicos de la membrana celular como receptores heteroméricos.

Las mutaciones en los SG's β , δ , y γ causan distrofia muscular asociada con cardiomiopatía, lo cual se ha observado tanto en pacientes como en ratones nulos para estos genes (Lapidos et al., 2004). Dentro de los modelos animales desarrollados para este grupo de proteínas existe el ratón deficiente para el SG- α , este ratón desarrolla distrofia muscular pero no cardiomiopatía, por lo que es posible que los altos niveles de expresión de ϵ -SG en tejido cardíaco compensen la deficiencia de α -SG, dada la homología observada entre ambas proteínas (Lapidos et al., 2004), hecho que si se ha observado en músculo esquelético mediante la sobre-expresión del SG- ϵ (Imamura et al., 2005).

1.4 Regulación transcripcional de los sarcoglicanos

Durante el desarrollo, la expresión de los sarcoglicanos sigue una progresión definida. Los SG's β y δ son los primeros en expresarse en el miotomo

en la cuarta semana de embriogénesis humana, seguidos por γ -SG que se expresa en la sexta semana de desarrollo, y por último, el SG- α aparece en las fibras musculares al final del periodo embrionario (Durand et al., 2002). En complemento Ettinger et al., (1997), observaron el transcrito de ϵ -SG en el día catorce de desarrollo en embriones de ratón. También se ha analizado la expresión de los SG's en células C2/4 durante su diferenciación y se ha observado que la expresión de α - y γ -SG's aumenta 20 y 50 veces, respectivamente, en miotubos con respecto a la expresión observada en mioblastos. De manera interesante la expresión de los SG's β y δ se mantiene sin cambio en ambos estadios de diferenciación (Noguchi et al., 1999). Estos cambios en la expresión a nivel de proteína correlacionan con lo observado por Wakabayashi et al., (2001) quienes mediante la caracterización parcial de los promotores de las subunidades del complejo SG-SPN en la misma línea celular, demostraron que hay un incremento en la transcripción de los SG's α y γ relacionado con la diferenciación celular, mientras que la expresión de un gen reportero bajo el control de los promotores de los SG's β y δ se mantuvo sin cambio a través del proceso de diferenciación. Estos resultados hacen evidente una regulación transcripcional específica para cada una de las subunidades del complejo SG-SPN.

En cuanto a la regulación de los promotores, Wakabayashi et al., (2001) detectaron en el promotor de γ -SG una secuencia rica en A/T y una caja E, que es sitio consenso para los factores de transcripción de la familia bHLH. Estas dos secuencias en conjunto constituyen un potenciador que controla la expresión de γ -SG en presencia del factor de transcripción MyoD. También se localizaron dos promotores para α -SG, uno de ellos localizado al extremo 5' del primer exón y el segundo dentro del primer intrón, de los cuales éste último demostró tener una mayor actividad asociada a la diferenciación de las células C2/4 (Wakabayashi et al., 2001).

Por otra parte, se ha observado que la sobre-expresión de γ -SG causa también un fenotipo distrófico, semejante al fenotipo distrófico del modelo deficiente para esta proteína (Zhu et al., 2001), lo que sugiere fuertemente que los genes de los SG's se encuentren sujetos a una regulación tanto negativa como positiva muy fina. Lo anterior se confirma en el caso del promotor localizado en el primer intrón de α -SG que es regulado por los factores de transcripción de la familia NF1, los cuales actúan alternativamente como activadores o reguladores negativos. En mioblastos NFI-A1.1 y NFI-B2 son activadores, mientras que NFI-C2 y NFI-X2 son reguladores negativos (Delgado et al., 2004a). En tanto que en miotubos todos actúan como activadores, mostrando que la regulación de α -SG depende del estadio de diferenciación. A partir de este mismo análisis, se tiene conocimiento de la presencia de dos cajas E dentro de la región promotora de α -SG, una de las cuales muestra alta afinidad por MyoD en ensayos *in vitro* (Delgado et al., 2004b) de manera semejante a otros promotores específicos de músculo.

1.5 El Sarcoglicano Epsilon

El gen del SG- ϵ (*SGCE*) se encuentra localizado en el locus humano 7q22, abarca 71 kb en las que se encuentran contenidos 12 exones. La alta homología entre α y ϵ SG's es indicativa de un posible evento de duplicación de genes, ya que presentan un 47% de homología a nivel de nucleótidos (43% de identidad) y un 62% a nivel de residuos de aminoácidos (aa). Además las fronteras intrón-exón de ambos genes son idénticas (Straub et al., 1999), sin embargo, ϵ -SG tiene regiones intrónicas más grandes con respecto a las que se observan en α -SG (McNally et al., 1998; Hack et al., 2000). En el ratón, el gen *SGCE* se localiza en el cromosoma 6, adyacente al gen *PEG10* (*Paternally Expressed Gene 10*), que se encuentra orientado en dirección opuesta. Ambos están incluidos dentro de un agrupamiento de varios genes que están sujetos a impronta ya sea materna o paterna (Ono et al., 2003). Esto sin embargo, no condiciona que ambos genes se

expresen de la misma manera, ya que Smallwood et al. (2004), han observado que son expresados diferencialmente en placenta.

El gen *SGCE* se encuentra sujeto a un mecanismo de regulación epigenética conocido como impronta, que consiste en silenciamiento de la expresión de uno de los alelos parentales. En el caso del gen *SGCE* se ha descrito que el mecanismo de impronta del alelo materno es por metilación de islas CpG localizadas en el extremo 5' del gen *SGCE* (Piras et al., 2000; Grabowsky et al., 2002). Recientemente, se describió que algunas mutaciones en el gen *SGCE* causan Distonia Mioclónica (DYT11), enfermedad del SNC que afecta principalmente los miembros superiores y el cuello, aunque ocasionalmente se extiende a los miembros inferiores, y como su nombre lo indica, se acompaña por contracciones mioclónicas (Zimprich et al., 2001).

A partir del gen *SGCE* se transcriben y maduran dos mRNA principales, uno con la secuencia codificante completa y otro mRNA que carece del exón 10, ya que sufre un proceso de empalme alternativo, éste último es expresado en mayor proporción que el mRNA que contiene la totalidad de la región codificante. Ambos son de aproximadamente 1.8 kb, con un marco de lectura abierto de 437 aa y se han encontrado en tejidos tanto musculares (músculo esquelético, liso y cardíaco) como no musculares (pulmón, páncreas, placenta, cerebro, corazón y en menor proporción en hígado y riñón) (Ettinger et al., 1997; McNally, et al., 1998). El mRNA de ϵ -SG muestra amplia expresión en diferentes regiones del cerebro, como el lóbulo temporal, lóbulo frontal, lóbulo occipital, médula, corteza cerebral, putamen, la médula espinal y el cerebelo (Zimprich et al., 2001; Xiao y LeDoux, 2003).

La proteína de ϵ -SG tiene una secuencia señal de 22 aa, seguida por el dominio extracelular de 262 aa que tiene un sitio de glicosilación que también está presente en α -SG, el dominio transmembranal de 23 aa y el dominio intracelular de 98 aa, mismo que contiene tres sitios potenciales de fosforilación (Ettinger et

al., 1997). También puede presentar un dominio transmembranal adicional que surge a partir de la inclusión del exón 10 (McNally et al., 1998).

La expresión de ϵ -SG ha sido detectada desde etapas tempranas del desarrollo en mioblastos, branquias alrededor del músculo liso en pulmón y en el endotelio vascular de embriones de ratón, mientras que en estas etapas no se ha detectado la expresión de α y β SG's, lo cual indica que la expresión de ϵ -SG antecede a la de esos otros SG's. Posteriormente los niveles de ϵ -SG se vuelven altos en músculo liso y endotelio vascular, pero decaen en tejido cardíaco y esquelético, sin observarse cambios post-natales en el patrón de expresión (Straub et al., 1999). Es interesante que ϵ -SG sea expresado en bajos niveles en músculo esquelético, ya que una posible sobre-regulación de ϵ -SG podría considerarse como alternativa en el tratamiento de las sarcoglicanopatías.

También es relevante el patrón de expresión de ϵ -SG, principalmente en sistema nervioso, donde ha sido co-localizado en nervio periférico de conejo (nervio siático, femoral y tibial) junto a β y δ SG's, el complejo DG, utrofina y la isoforma corta de distrofina Dp116, formando un complejo similar al reportado en músculo esquelético, sólo que en el nervio periférico no se asocian α ó γ SG's, ni SPN (Imamura et al., 2000).

2. PLANTEAMIENTO DEL PROBLEMA

El estudio del complejo SG-SPN se ha intensificado en los últimos años, pero aún se desconocen detalladamente los mecanismos que regulan la expresión de los componentes de este complejo y que resultan de suma importancia no solo para conocer con mayor detalle estas proteínas y sus genes, sino para el desarrollo de posibles terapias basadas en la sobre-expresión de genes homólogos, tal como se ha planteado en el caso de la distrofina y su homólogo utrofina, en el caso de pacientes con DMD (Dennis et al., 1996).

Uno de los abordajes de este tipo de estudios es la caracterización de las regiones promotoras, ya que en ellas recae uno de los puntos de regulación más importantes de la expresión génica. En particular la caracterización del promotor de ϵ -SG, resulta especialmente importante ya que ofrecerá información acerca de la regulación a la que se encuentra sujeto tanto en los tejidos musculares como no musculares, recordando que es muy posible que la sobre-expresión de ϵ -SG compense la deficiencia de α -SG en pacientes con sarcoglicanopatías. Además de la evidente importancia de la regulación del SG- ϵ en el SNC, manifiesta por su relación con el desarrollo de distonia mioclónica.

3. HIPÓTESIS

La expresión tejido específica descrita para α -SG frente a la expresión ubicua observada de ϵ -SG, debe ser resultado de una diferente estructura y presencia de elementos reguladores en las regiones promotoras de los genes que codifican para estas proteínas.

4. OBJETIVOS

Objetivo general:

- Aislar y caracterizar la región promotora del gen *SGCE*

Objetivos particulares:

- Aislar la región promotora del gen *SGCE* mediante el tamizaje de una genoteca de ratón.
- Secuenciar la región promotora aislada.
- Analizar *in silico* los elementos reguladores presentes en la secuencia promotora aislada.

5. ESTRATEGIA EXPERIMENTAL

Para aislar y caracterizar la región promotora del gen *SGCE* se realizó el tamizaje de una genoteca de ratón con una sonda marcada radioactivamente. Ya con la identificación de las clonas candidatas para la región promotora de ϵ -SG, se realizaron varios tamizajes para obtener clonas aisladas y extraer ADN de las mismas.

Una vez que se aisló la región candidata, fue secuenciada y se diseñaron oligonucleótidos dentro la región promotora para generar los fragmentos con deleciones progresivas, para ser clonados en el plásmido pDrive (Qiagen), además se realizaron diversos análisis de la secuencia para localizar la posible región promotora e identificar sitios de reconocimiento a factores de transcripción, dentro de la región hacia 5' del gen *SGCE*.

6. MATERIAL Y METODOS

6.1 Tamizaje de una genoteca

Para aislar las regiones candidatas se realizó el tamizaje de una genoteca de ratón (EMBL3 SP6/T7 Clontech) mediante hibridación con una sonda de ADN de 80 bases, diseñada y sintetizada *in vitro* a partir de la secuencia correspondiente al extremo 5' del primer exón del gen *SGCE* de ratón: **Epsilon5.F** 5'-GAG ACC CTT GTG CTT GGA CGG GAA AGG GTC GGG GGA CAC TCA AGA TGA GCC CCG CGA CCA CTG GCA CAT TCT TAC TGA CA-3' (Gibco).

Para realizar el tamizaje se infectaron 200 μ l de la bacteria *E. coli* (K802) a una D.O.₆₀₀ = 0.5 con la cantidad necesaria del fago recombinante, calculada a partir de la titulación de la genoteca, y se plaqueó en cajas Petri con 3 ml de agar suave, distribuyéndose uniformemente sobre la superficie de la caja, dejándose incubar a confluencia toda la noche a 37° C.

Posteriormente, se llevó a cabo la transferencia a las membranas de nitrocelulosa (Millipore), colocándolas sobre la superficie de la caja durante 1 min para la primera copia y 3 min para la réplica. Se hicieron marcas asimétricas con tinta indeleble en cada par de membranas con una misma orientación para poder compararlas e identificar de manera certera las clonas positivas. Posteriormente, las membranas se sumergieron en solución desnaturalizante (NaOH 0.5 M; NaCl 1.5 M) por 5 min con agitación, y enseguida se bañaron en solución neutralizante (Tris HCl 1 M pH 8.0; NaCl 1.5 M) por 2 min agitando ocasionalmente. Las membranas se dejaron secar sobre papel filtro y una vez que estaban completamente secas se fijó el ADN con luz ultravioleta a 1200 J por 1 min, empleando un crosslinker de luz ultravioleta UVC 500 (Hoeffer).

La hibridación de las membranas se realizó con la sonda Epsilon5.F marcada con $\gamma^{32}\text{P}$ de la siguiente manera: Se pre-hibridó en solución de hibridación (EDTA 2m M, NaCl 0.9 M, dextran sulfato 0.1 mM, polivinil pirrolidona 20 μM , ficoll 2 μM , Na_2HPO_4 9.3 mM, $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ 15.5 mM, formamida 40%, SDS 1%) en presencia de ADN de esperma de salmón (50 $\mu\text{g}/\mu\text{l}$), por al menos 2 h a 42° C con agitación. Posteriormente, se agregó la sonda marcada (25 ng) dejando hibridar las membranas toda la noche en la solución de hibridación a la misma temperatura en agitación.

6.2 Marcaje de la sonda

El marcaje de la sonda se realizó con $\gamma^{32}\text{P}$ bajo las condiciones de reacción que a continuación se describen: Se preparó una mezcla de 5 pmol de la sonda **Epsilon5.F**, 10 U de enzima T4 cinasa y 50 μCi de dATP en un volumen final de 25 μL y se incubó por 10 min a 37°C. Posteriormente, se llevó a 65°C por 10 min para inactivar la cinasa.

Después de la hibridación durante toda la noche, las membranas se lavaron con SSC (NaCl 3 M; citrato de sodio 0.3 M) 2X por 15 min y éstas fueron

colocadas sobre papel filtro cubierto con película plástica, posteriormente se colocaron dentro del casete con una película autorradiográfica, dejándose exponer a -70°C , por 4 h. Una vez revelada la placa autorradiográfica se identificaron las clonas que dieron señal positiva en la membrana original y el duplicado. Las clonas seleccionadas se amplificaron y posteriormente se realizaron, tamizajes subsecuentes, corroborando así que la sonda hibridara con las clonas seleccionadas. De esta manera, también se obtuvieron las clonas aisladas directamente de la caja de Petri para poder realizar la extracción de ADN.

6.3 Extracción de ADN

A partir de las clonas aisladas se infectó un cultivo de *E. coli*. de la cepa K802 y se crecieron a confluencia tal como se hizo para el tamizaje. Posteriormente, se aisló el ADN con el Lambda Midi Kit (Qiagen) de acuerdo con las indicaciones del producto. El ADN aislado fue analizado mediante PCR con los oligonucleótidos **ESGP1.F**: 5'-CTT GTG CTT GGA CGG GAA AG-3' y **ESGP2.R**: 5'-CCC TGA ACC TCC ATG TCG C-3' para corroborar que contenía el extremo 5' del gen SGCE. La reacción se llevo a cabo bajo las siguientes condiciones de amplificación: 94°C 5 min y 94°C 30 seg, 62°C 30 seg, 72°C 30 seg, durante 30 ciclos.

6.4 Secuenciación

La secuencia de la región candidata se obtuvo mediante secuenciación automatizada. La reacción de secuenciación se llevó a acabo con los kits de secuenciación Big Dye y Dicloro-rodamina (Applied Biosystems) en una reacción de 20 μl con los oligonucleótidos: **ESGP2.R**, **SeF**: 5'-GCG TCT GCA GCT CAT CGC T-3' y **SeR**: 5'-TCA ACC GTC CTC GAT CCT CC-3', bajo las siguientes condiciones: 94°C 5 min, y 94°C 5 seg, 50°C 10 seg y 62°C 4 min durante 25 ciclos.

6.5 Diseño de los oligonucleótidos

A partir de la secuencia de aproximadamente 2.4 kb, correspondiente al extremo 5' del gen *SGCE*, se diseñaron con el uso del programa OLIGO, 14 oligonucleótidos sentido y un antisentido (Tabla 1), para generar las construcciones con deleciones progresivas para los ensayos de actividad transcripcional frente a un gen reportero, a estos oligonucleótidos se les integró un sitio de reconocimiento para la enzima de restricción *XhoI*, como herramienta para su subclonación.

Tabla 1. Oligonucleótidos utilizados para generar los fragmentos con deleciones progresivas a partir del promotor del gen *SGCE*.

Nombre	Posición	Secuencia 5'-3'	Amplicon
STR	81	gcgcCTCGAGtggtcgcggggctcatcttg	-----
1.F	- 23	gcgcCTCGAGaggatcgaggacggttgagc	123 pb
1.5F	- 80	gcgcCTCGAGggtgcgcgatgctgaactgg	180 pb
2.F	- 126	gcgcCTCGAGactgctttggcccgcgcta	226 pb
2.5F	- 212	gcgcCTCGAGaaatcccagcgtgcttcgcg	312 pb
3.F	- 336	gcgcCTCGAGttctcggttggtgctctcc	446 pb
4.F	- 534	gcgcCTCGAGgcggtgacccgttttcttg	634 pb
5.F	- 750	gcgcCTCGAGgcgctctgcagctcatcgctg	850 pb
6.F	- 1056	gcgcCTCGAGcccatcagttcgtagcgtgt	1156 pb
7.F	- 1349	gcgcCTCGAGgcggttggttactctcctgc	1449 pb
8.F	- 1566	gcgcCTCGAGttaccaccagcaccaccacc	1666 pb
9.F	- 1743	gcgcCTCGAGccattcgtgtccccaagtcc	1843 pb
10.F	- 1977	gcgcCTCGAGcagggctgcagaatgactgc	2077 pb
11.F	- 2238	gcgcCTCGAGctctggccaccttgagga	2338 pb
12.F	- 2378	gcgcCTCGAGcaacagagaaaccgcagga	2478 pb

* El sitio de restricción para la enzima *XhoI* se indica en mayúsculas.

6.6 Amplificación de los fragmentos escalonados por PCR

Los productos de PCR de 123 a 850 pb fueron generados con los oligonucleótidos 1.F a 5.F en una reacción de PCR con la enzima Expand Long Template (Roche) bajo las siguientes condiciones: desnaturalización inicial a 94°C por 5 min; desnaturalización a 94°C por 30 seg, alineamiento a 65°C durante 30 seg y extensión a 72°C por un min, durante 32 ciclos, y siete min de extensión final a 72°C. Estas amplificaciones fueron visualizadas por electroforesis en geles de agarosa al 2%. Los fragmentos restantes, generados con los oligonucleótidos 6.F a 12.F, se amplificaron con la misma enzima bajo las condiciones de reacción que se describen a continuación: desnaturalización inicial a 94°C por 5 min; desnaturalización a 94°C por 30 seg, alineamiento a 65°C durante 30 seg y extensión a 72°C por tres min, también durante 32 ciclos y siete min de extensión final a 72°C. Los fragmentos generados con estas condiciones, de 1156 a 2478 pb, se visualizaron por electroforesis en geles de agarosa al 0.8%. Las amplificaciones por PCR fueron realizadas en un termociclador TGradient (Biometra).

6.7 Obtención de bacterias competentes

Las bacterias *E coli* de las diferentes cepas (DH5 α , JM109 y HB101) se hicieron competentes mediante cloruro de rubidio, siguiendo el método descrito a continuación: primero se inoculó con una colonia medio LB (3 ml) sin ampicilina, dejándose crecer a 37°C y 180 rpm durante toda la noche, posteriormente se agregaron 200 μ l de este cultivo en 20 ml de medio LB sin ampicilina, hasta alcanzar una D.O.₆₃₀= 0.4 - 0.6, una vez que el cultivo llegó esa densidad óptica, se colocaron en tubos de 50 ml y fueron centrifugados a 8000 rpm, durante cinco min a 4°C, se desechó el sobrenadante y la pastilla se resuspendió en 10 ml de la solución TBF1 (RbCl 100 mM, MnCl₂-4H₂O 50 mM, acetato de potasio 30 mM, CaCl₂-2H₂O, glicerol 15%), dejándose incubar por 20 min en hielo. A continuación se centrifugó nuevamente a 8000 rpm, durante cinco min a 4°C, se desechó el

sobrenadante y la pastilla se resuspendió en 2 ml de la solución TBF2 (MOPS 10 mM, RbCl 10 mM, CaCl₂-2H₂O 10 mM, glicerol 15%), dejándose incubar por 20 min en hielo. Las células competentes se colocaron en alícuotas de 200 µl en tubos eppendorf de 0.6 mL y se conservaron a -70°C.

6.8 Generación de los extremos cohesivos

Los productos de PCR utilizados para clonarse mediante ligación con los extremos cohesivos generados por la enzima *Xho*I, fueron purificados a partir de la reacción de PCR y posteriormente cortados con ésta enzima, bajo las siguientes condiciones de reacción: 500 ng de inserto, 1 U de enzima *Xho*I (New England Biolabs) y 1x de amortiguador NEB2, en un volumen total de 50 µl. Incubándose a 37°C durante toda la noche.

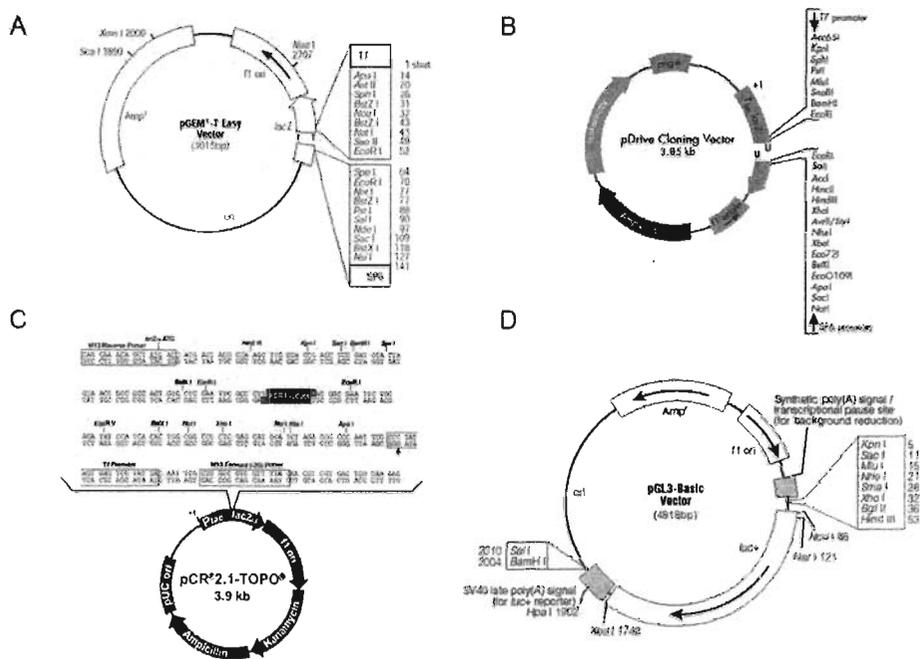
Por otra parte, a los fragmentos utilizados para clonarse directamente como productos de PCR se les agregaron adeninas en los extremos, siguiendo el protocolo de poli-adenilación descrito en el manual del vector pGEM-T easy, ya que la enzima Expand Long Template no tiene actividad de adenil-transferasa.

6.9 Ligación y transformación

Una vez generados los fragmentos, se realizó una reacción de ligación con los diferentes vectores con la enzima T4 DNA ligasa, en presencia de 50 ng de vector y en una relación molar vector:inserto 1:3 con un volumen final de 10 µl. Se emplearon los vectores pDrive (Qiagen) (Figura 3B) y pGL3-Basic (Promega) (Figura 3D) para clonar con los extremos cohesivos del corte con *Xho*I y los plásmidos pGEMT-easy (Promega) (Figura 3A) y PCR 2.1-TOPO (Invitrogen) (Figura 3C) para clonar los productos de PCR directamente.

La transformación de las células competentes de diferentes cepas (DH5α, JM109 y HB101) se realizó de la siguiente manera: la reacción de ligación se

colocó en un tubo de 1.5 ml y se le agregaron 50 μ l de células competentes incubándose por 2 min en hielo. Posteriormente, se dio un choque térmico a 42°C por 45 seg, incubándose nuevamente por 2 min en hielo. Después de este tiempo se agregaron 900 μ l de medio SOC a temperatura ambiente y se incubó en agitación (180 rpm) durante 90 min. Se plaquearon dos cajas de Petri (Agar LB con ampicilina [100 μ g/ml]) con 100 μ l de la reacción de transformación dejándose crecer a 37°C toda la noche. En el caso del vector pGL3-Basic, las clonas observadas en las cajas fueron probadas por PCR usando los mismos oligonucleótidos que se usaron para generar los fragmentos, mientras que en el caso de las clonas observadas en la transformación con los vectores pDrive, PCR 2.1-TOPO y pGEMT-easy se discriminó mediante la aplicación de X-gal, reactivo con el cual, si las bacterias contienen un plásmido con inserto son blancas, mientras que si poseen un plásmido sin inserto se tornan de color azul.



7. RESULTADOS

7.1 Aislamiento y secuenciación de la región promotora del gen *SGCE*

A partir del tamizaje de la genoteca de ratón EMBL3 SP6/T7 (Clontech) se obtuvieron tres clonas positivas que hibridaron con la sonda correspondiente al extremo 5' de la secuencia codificante del gen *SGCE*. En la figura 4 se muestra la identificación de la clona 12A en la membrana de nitrocelulosa con su réplica. A partir de éstas clonas se extrajo ADN, con el cual se realizaron reacciones de amplificación por PCR con los oligonucleótidos ESGP1.F y ESGP2.R para comprobar que contuvieran la región 5' del gen *SGCE*.

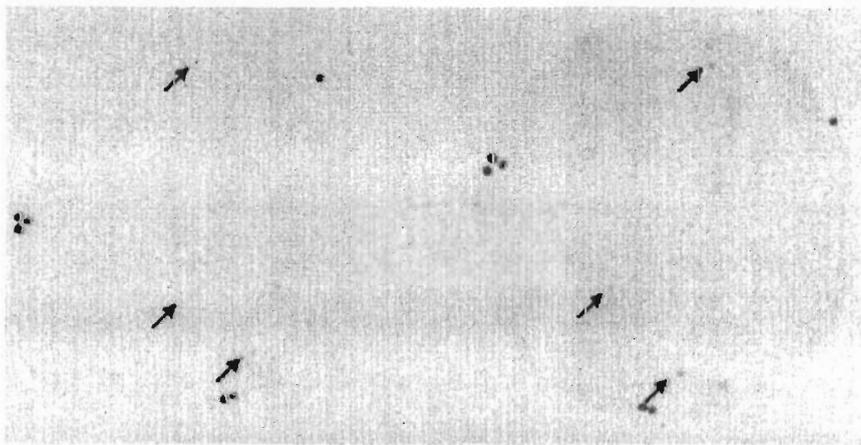


Figura 4. Identificación de las clonas que hibridaron en el tamizaje radioactivo. Autorradiografía de las membranas de nitrocelulosa correspondientes al tamizaje y aislamiento de la clona 12A. En cada una de las membranas se observan las marcas asimétricas de uno, dos y tres puntos en la circunferencia de las membranas. Las flechas señalan los sitios donde hibridó la sonda marcada radioactivamente y que aparecen tanto en el original (izquierda) y en la réplica (derecha), con los cuales se identificaron las clonas directamente en la caja de Petri.

A partir del ADN de las clonas aisladas se realizó la reacción de secuencia con la cual se obtuvo una lectura de 310 bases, de las cuales 173 se encuentran hacia 5' del sitio de inicio de la transcripción de ϵ -SG (Figura 5). Con la secuencia obtenida de la región hacia 5' de la región codificante de ϵ -SG se realizó un alineamiento con el programa BLAST (Altschul et al., 1997) y se detectó homología con la secuencia con número de acceso NW_000272 (GenBank), de la cual se muestra la región hacia 5' del gen *SGCE* de aproximadamente 2.4 kb, en la cual se observa una región de aproximadamente 494 nucleótidos no identificada (Figura 5).

Para obtener la secuencia no determinada se diseñaron los oligonucleótidos SeF y SeR flanqueando esta región (Figura 5). Como resultado se obtuvo una secuencia de 646 bases, en la que se observa un alto contenido de C/G (61 %) (Figura 6).

La secuencia completa de 2.4 kb contiene 156 dinucleótidos C/G (Figura 7A) que podrían metilarse, tal como se ha observado en el promotor del gen *SGCE* de humano, en el que se ha descrito la presencia de islas CpG. Mediante el análisis de esta secuencia con el programa CpG PLOT (Larsen et al., 1992), para identificar islas CpG, se identificaron tres islas presentes dentro del promotor del gen *SGCE*, las cuales se ubican entre los nucleótidos -1322 a -1011, -918 a -555 y -490 a -90 respectivamente (Figura 7B).

7.2 Construcción de los vectores de expresión

En la figura 8 se muestra la amplificación de los productos de PCR obtenidos con los oligonucleótidos diseñados dentro de la región promotora, cuya posición se ilustra en la figura 10, y que contienen en ambos extremos sitios de reconocimiento para la enzima de restricción *Xho*I. Estas amplificaciones se realizaron con la enzima Expand Long Template (Roche), que tiene la característica de amplificar fragmentos de gran tamaño mediante PCR.



Figura 5. Secuencia del extremo 5' del gen SGCE. El área sombreada de 310 pb es la secuencia obtenida con el oligonucleótido ESGP2.R (ubicado en la posición +137, indicado en negrita). A partir del alineamiento de esta secuencia en el programa BLAST, se obtuvo la secuencia NW_000272 con una longitud de 2352 nucleótidos hacia 5' del gen de ϵ -SG. Dentro de la secuencia NW_000272 se localizó una región no reportada estimada en 494 bases, indicada en *italica* y subrayada. Para secuenciar esta región se diseñaron los oligonucleótidos cuya secuencia se indica en **negrita** en las posiciones -578 (SeF) y -6 (SeR). Las flechas en la posición +1 y +85 indican los sitios de inicio de la transcripción y traducción, respectivamente.

-2400	GTGGACAAAA	AATAAACCCG	CGCAACGAG	AAACCCGAGG	GAGCGGGGGC	TGTGCAGCCA	GGCTCAGACA	GGCAGAGGGC	CACATGCGCA	CTGGAGGCAA
-2300	ACCTGAAGGG	AGGGGTTGCA	GCCCCGATTG	CGCAGGGCGA	AGCAGGCTGG	GGGAAGGGCG	GGCTCTGGGG	CACCTTGGAG	GAGCTCTGGG	AGGAGCTGAA
-2200	CAGTAAGAGA	GGCGGGGCGT	TGGGGGCGGG	GGCTAGGCAG	GCCACCTCAG	AGGAGGGCGG	CCTTGGGGGT	GGGGACCAGG	TAGGCTGGGT	CTAAGGCGTC
-2100	AGCACTTGGG	AGGTTGGGCC	TGTGGAAGCA	TGTTCTGGAA	AAGCGCGCGA	CAGGCGAGGC	GCGGGCCTC	TCTGCAGATG	AGGCAAGGCT	AGATGGGCGG
-2000	GGCAGGCTCT	AGACTGGCTA	CAACAGGGCT	CGAGAATGAC	TGATGGCTCT	CCCCCGCCCC	CCTACCTCCC	CAGACCCCTC	CCACCAGCTC	CCAACCCACA
-1900	GATCCCTTCC	GTCGAGCAAG	AGTGAATCCC	CCACCTCCCA	CCGTCCCATC	ATCTCCCACC	TCCCACCTC	TCCCACCTCC	CATCATCTCC	CACCTCCCAC
-1800	CTTGGCTTCG	TGCCCACCCA	GAAAACCGGA	CAGTTTCCAC	AAACTGTTGC	TGGACACCCA	TTCGTGTCCC	CAAGTCCAGT	GCACCGTCAC	CGTCACCTCA
-1700	ATCAATCACC	ACTAGACAAG	GCAGGCCATT	CCGCCATCTC	CAAATCCAG	CACCCCTCCA	CTCCACTGCC	ATAAGCAGGA	GTTACGTGGG	CAGAATGTTA
-1600	AAACCAATTT	TGGCCTCCCA	GGGCTCCATT	TTGTTTACCA	CCAGCACCC	CACCATCATG	CAGATAGCCG	TCCGACAGGA	AAACGGAGCC	AGTTTTGACC
-1500	CGAAAGGTGA	ACTGTGTTGT	CCCCTCACCA	AGGAGAGATA	CGT1CTCTCG	TCTTCCCAAC	GCTACTTTCT	GTGTCACTTA	AGAGGCTGCA	GGTCTTATGC
-1400	CTGTGCCGCA	GTTTGTAGCG	CATTTTCTAA	ATTCACTCTT	GTGCGCTTT	TGCGTTTGGT	TACTCTCTG	CAGCTTTCCA	AATTGAATAC	GTGAAACACG
-1300	AATCGGTTTT	GGGTATCACC	GTAGTTCGGG	CTGGTTTGTG	GCGCATGAAG	CGCCCCGTTT	GCCCCGTTCA	CAGGTTCCGC	TGTTCCGCCG	TTTCAATTGC
-1200	TCGTTCGTAC	GCTGAAGCGC	ATGAAGCGCC	CATTAGTACA	GTGAAGCGCA	TGGCGCGCCC	ATTAGTACAG	CGAAGCGCAT	TAAGCGCTCA	TTAGTATATC
-1100	GTAGCACATG	AAGCGCCCAT	TAGTACAGCG	AAGCGCATGA	AGCGCCCATC	AGTTCGTAGC	GTGTGAAGTG	CCTATGAGTA	TTTTGTAGGG	CATGAAGCGC
-1000	CTATGAGTAT	TTTGTAGCGC	ATGAAGCGCC	TATGAGTAT	TTGTAGCGCA	TGAAGCGCCT	ATGAGTATTT	TGTAGCGCAT	GAAGCGCCTA	TTAGGATTTT
-900	GTAGCSCATG	AAGCGCCCGT	TAGTTTATCA	CAGCGCAAGA	CGCGGCCGTC	TGCAGTGCAA	TCCGTTTCGA	GCGCGGCACT	TAAGAGTGCA	GAGCCAGTCA
-800	CTTTGCAGCG	TATTTGGGGA	GCTCATCACA	ATTCGGTCAG	TAGCTCATGT	GCGTCTGCAG	CTCATCGCTG	TCCGAAGCTC	ATGTGTCTCC	GCCCATTTCT
-700	GGACGCGAGT	CGTGA AAC	GAATTCGCGG	ACCCGAAGCC	GTGAAGACTT	GCTTTCACG	CTCTCGTCTC	AGGATCTGGC	CACTGAAGGC	TTGACAAAAT
-600	TTGCGAGGAG	ATATTTTGGG	GGGCGCTACT	TTTATCTCCC	GCACTCCCG	ATGTA AAC	GGGATTCGCT	GTGACCGTTT	TCTTGGAGTT	AAAAAAAAAA
-500	AAAAATCCTA	ACCATACTCA	CCACACGAGG	ATTCCTCTGC	GATGCCACGA	AGGATGAGGG	GTCCCGCGAC	GGCCCGGTCC	ACTGGGTGAC	CGCGGTCCGG
-400	AGAAGGCTCC	GGACCGCAGC	CCGCGCAGCA	GACCTAGGCG	AGGCTCGGTG	GACCTTCTCG	GTTGGCTGCT	CTCCCAGGT	CGTAGTTGGT	CAGGATTTCT
-300	CGACACACAC	CAACTCAAGT	TGGA AAC	TGTATCGGGA	GTCCGGACTG	CTTGCCCTTAT	ATAGAGAGAT	CCTAGGAGGG	GGGGCGCCAA	ATCCCAGCGT
-200	GCTTCGCGAG	CAGCGGGGGC	GGAGGGGGGC	GCGTGGGCGG	GCGCTTGGG	GGATGGGGTG	GGGGCGGGC	ATGCACTGCT	TTGGCCCGCC	GCTAGGGGCA
-100	GCATGGGCTC	AGGAAGGGGG	GGTGCAGCTG	GCTGAAGCTG	CCAAGCTGGG	AGGAAAAAAA	GAAAGGGGAG	GGAGGGGAGG	ATCGAGGACG	GTTGAGCTTA
	▷								▷	
+ 1	GGCAGGCCAA	GAATGCTATT	ATTCTGGTGG	TGGGAGCTGG	GAGACCCCTG	TGCTTGGACG	GGAAAGGGTC	GGGGGACACT	CAAGATGAGC	CCCAGGACCA
101	CTGGCACATT	CTTACTGACA	GGTTAGTGGC	GTCGCGCGGA	CATGGAGGTT	CAGGGGCCCA	CAGGCCGCGC	CCTCACACTC	CCCCACTCGG	GTTCGGGGGT
201	ACCAGGGCGT	GCCCAAAGGG	TGCTAGTCCC	GCTCTATGCC	CAATGCCTCC	CCCTCACAGC	ACCCCTCTTT	CAGGACACCC	TGCAGCCAC	CTTTTGGGCC
301	AAGTGGGTCC	CTCCCCAGGC	TTGACCCCTG	GCCTTCTCTT	TCTCTCTGCA	CCCCCACGGC	CCCTCCCTGT	CCAGCAGGGC	AGGAATTCAG	CTCCAGTTTC
401	CATTAAG---	-----	-----	-----	-----	-----	-----	-----	-----	-----

Figura 6. Secuencia completa del extremo 5' del gen SGCE. A partir de la reacción de secuencia con los oligonucleótidos SeF y SeR se obtuvo la lectura de 646 bases localizadas entre los nucleótidos -59 a -704 y señaladas en negrita, que corresponde a la región que no se había reportado. La secuencia subrayada entre las bases -1 y -301 corresponde a la posible región promotora basal, de aproximadamente 301 pb, identificada con el programa *PromoterInspector*. Las flechas en la posición +1 y +85 indican los sitios de inicio de la transcripción y traducción respectivamente.

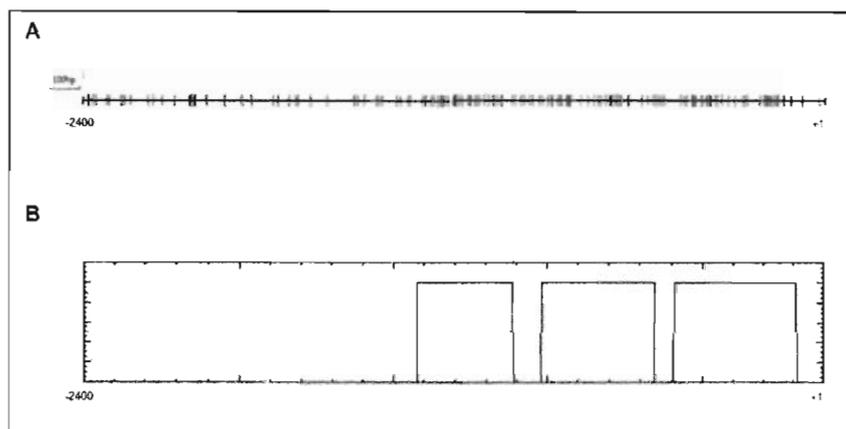


Figura 7. El promotor de ϵ -SG presenta un alto contenido de dinucleótidos C/G y tres islas CpG. A). Identificación de los 156 dinucleótidos CG presentes en la secuencia de 2.4 kb hacia 5' del sitio de inicio de la transcripción. B). Los cuadros dentro de este panel muestran la localización de las tres islas CpG dentro de la secuencia promotora de ϵ -SG.

Se realizaron ensayos de ligación de estos fragmentos con los vectores pGL3-Basic y pGEMT-easy para clonar con los extremos cohesivos del corte con *Xho*I, y de manera paralela en el plásmido PCR 2.1-TOPO para clonar los productos de PCR directamente, sin embargo, los resultados que en la placa eran positivos, resultaban sin inserto después de la extracción del plásmido, por lo que se intentaron los mismos ensayos a diferentes temperaturas de crecimiento (37 y 25 °C) y con diferentes cepas, como son DH5 α , JM109 y HB101, que son cepas con genotipo *recA*⁻, pues se sugería que el inserto pudiera estar recombinando con el ADN bacteriano durante el crecimiento en líquido previo a la extracción del plásmido.

Hasta el momento se han logrado clonar los productos de 226 pb en el vector de clonación pDrive (Figura 9), para posteriormente subclonarse en el vector de expresión o bien, subclonar en éste vector (pDrive+226) el gen reportero de luciferasa.

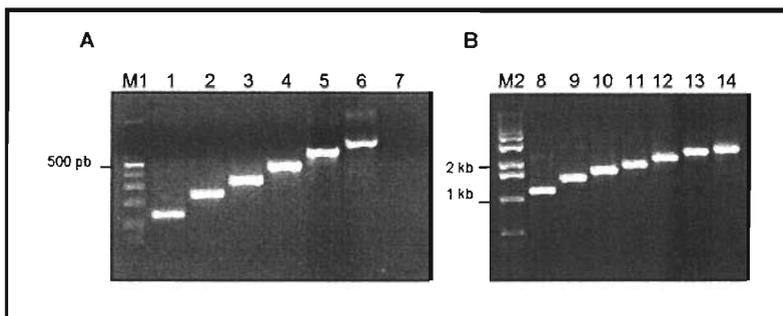


Figura 8. Amplificación de los fragmentos escalonados para la construcción de los vectores de expresión. A) 1-6 Amplificación de los productos de PCR generados con los oligonucleótidos localizados en las posiciones -23, -126, -212, -346, -534 y -750 respectivamente; 7: Control negativo. B) 8-14 Amplificación de los productos de PCR generados con los oligonucleótidos localizados en las posiciones -1056, -1349, -1566, -1743, -1977, -2238 y -2378 respectivamente. M1: Marcador de peso molecular escalera de 100 pb; M2: Marcador de peso molecular escalera de 1 kb.

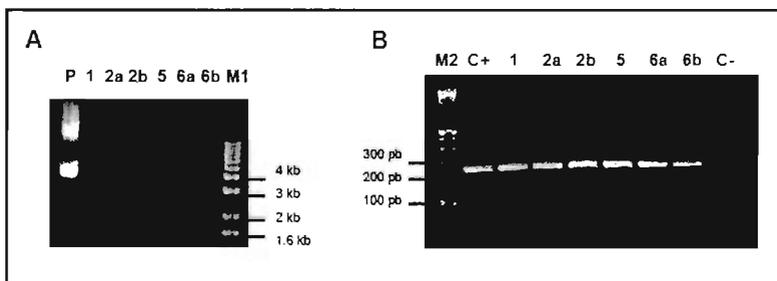


Figura 9. Amplificación con el oligonucleótido 2.F a partir de las construcciones generadas con el vector pDrive y el fragmento de 226 pb. A) Electroforesis en gel de agarosa (0.7 %) de los plásmidos purificados que contienen el inserto de 226 pb. P: plásmido sin inserto; 1, 2a, 2b, 5, 6a y 6b: plásmidos con el inserto de 226 pb. B) Electroforesis en gel de agarosa (2 %) de los productos amplificados con el oligonucleótido 2.F. C+: control positivo amplificado usando como molde el ADN purificado de las clonas aisladas en el tamizaje. 1, 2a, 2b, 5, 6a y 6b: amplificaciones usando como molde los respectivos plásmidos del panel A; C-: control negativo. M1: Marcador de peso molecular escalera de 1 kb. M2: Marcador de peso molecular escalera de 100 pb.

7.3 Análisis de la secuencia con los programas

Promoter Inspector, MatInspector y DiAlign

Mediante la utilización del programa *PromoterInspector* (Scherf et al., 2000; www.genomatix.de) se delimitó el promotor basal del gen *SGCE*, el cual tiene una longitud de 301 pb (Figura 6). La región promotora localizada con este programa se encuentra dentro de la región rica en C/G de 646 bases secuenciada previamente, sin identificarse una caja TATA en la región proximal al sitio de inicio de la transcripción del gen *SGCE*. Con la finalidad de identificar sitios consenso de unión a factores de transcripción se realizó el análisis por computadora de la secuencia de 2.4 kb con el programa *MatInspector* (Quandt et al., 1995; www.genomatix.de), con él se identificaron de los sitios de reconocimiento a factores de transcripción encontrándose nueve cajas GC, cuatro sitios potenciales de reconocimiento para SP1, dos para NF-1, dos para el factor de respuesta a suero (SRF) y dos cajas GATA entre otros (Figura 10).

7.4 Análisis comparativo de las regiones promotoras de α -SG y ϵ -SG

Mediante el uso del programa *DiAlign* se intentó identificar sitios consenso de unión a factores de transcripción en las regiones promotoras de los SG's α y ϵ . Dentro de la secuencia del promotor muscular del SG- α no se identificaron sitios consenso de unión a factores de transcripción comunes con el promotor de ϵ -SG descrito en este trabajo, si bien algunas regiones son similares a nivel de nucleótidos, como sitios de unión no se identificó alguno presente en ambos promotores.


```

Alfa1B 1 ggtgaagcac tgtttcttcc tcagggttca agggcagagt gttctgctgc
msge 1 cacctgtgga caaaaaataa aaccgcgcaa cagagaagacc gcag-----

Alfa1B 51 ctggttctcag agtttagacc tgcraagaa gqgtagatgc aggaat gaag
msge 45 -----

Alfa1B 101 ggccagaggt gtccaccatc aaaggcatca tggaaagacc GGACCGGGG
msge 45 ----- GGAGGCGGGG
*****

Alfa1B 151 CTGTGaactt tctgaacatt cagctggcat cagcatggct ttgtaggaag
msge 55 CTGTGcagcc -----
*****

Alfa1B 201 ttagaaaactg ccagccctca gcgat atcag ctgtaaaaca tgggctatgg
msge 65 -----

Alfa1B 251 gaat atctgt ctcat ggtac tgcctgtggg atgacatggg tcagagcaca
msge 65 -----

Alfa1B 301 taaatgcttg tgagcctttt ggcacagaga cggacagacc atcagtcctt
msge 65 -----

Alfa1B 351 ctccaccatca cacaaAGCCC CAGTCAGGCA Ggattttctc atctctgtat
msge 65 ----- AGGCT CAGACAGGCA Gaggccacac tgcgcactgg
*****

Alfa1B 401 ggtataggtg gctgaacatg tgcctgcttc agtatttttg tgagtcctct
msge 100 aggc aaacct gaaggagggg gttgcagccc qcattgcqca ggcgc aagca

Alfa1B 451 gcagagggcca gcacaaatcg tgcagcttat aat-----
msge 150 gqctggggga agggggggct ctggccacc ttggaggagc tctggaggga

Alfa1B 484 -----
msge 200 gctgaaccagt aagagagggc gggccttggg ggcgggggct aggcaggcca

Alfa1B 484 -----
msge 250 cctcagagga ggcgggcctt gggggtggg accaggtagg ctggctctaa

Alfa1B 484 ----- -TGTG GAAGCAgagg cctggagagg
msge 300 ggcgtcagca cttgggaggt tgggctGTG GAAGCAatgt ctggaaaaag
*****

Alfa1B 508 gctagttact taaggccacg tctgcaggaa ctgtcagact ggaactccaa
msge 350 ggcgcagcagg cgaaggcgcg ggcctc-----

Alfa1B 558 cccttatgga attagcccca gctatgtgag ctggagagag cctctctgct
msge 376 -----

Alfa1B 608 tacctgtccc catttggccg ataccgtctg gtttgtctcc agccggacct
msge 376 -----

Alfa1B 658 tcacatggga gatgttggte tgcggctcag gctggggaac agatacttct
msge 376 -----

Alfa1B 708 gcttgctgtt aaggcagttg tcagccagcc tctacactta ctagttgcta
msge 376 -----

Alfa1B 758 gacccaggt caggtgagtg gccacagaca cctgttccca caccatgatg
msge 376 -----

Alfa1B 808 ttctaaggct gggttggcac agggatggg ggggggtgag tcacaagtac
msge 376 -----

Alfa1B 858 ccagacagcc taggaagagg aggtacagga gaaaagactt cagagaggac
msge 376 -----

Alfa1B 908 gaaacacata tgacctgtc aaggTCTGCC CAAGGGACAA GGCAGGAGGG
msge 376 ----- TCTGCA GATGAGGCAA GGCTAGATGG
*****

Alfa1B 958 Gacaccagcc agagagagca gcttggatgg aggccaggaa actaagagga
msge 402 Gcgggqcaag tc-----

Alfa1B 1008 cacattaagc tagccagag tctagggatg ttggtgtgtg tggggtggg
msge 414 -----
    
```

Figura 11. Análisis comparativo de los elementos reguladores comunes, presentes en los promotores del gen de α -SG y el gen SGCE, con el programa DiAlign TF. No se observan similitudes en la estructura del promotor muscular del SG- α y la estructura de la región promotora del SG- ϵ .

Aislamiento y caracterización de la región promotora del sarcoglicano epsilon

```

Alfa1B 1058 aggtgccgtt caaattggcc acaaacatag aggaagggcc agaaagccat
msge 414 -----
Alfa1B 1108 cgactgggga tgagctggga gctcaagaac aattccagtt gtTGAGATTT
msge 414 -----TGAGACTG
*****
Alfa1B 1158 TAAACAACCTG AGCTGGGGAT GGGTTGgggt aggggtgta gagagatggc
msge 422 GCTACAACAG GGCTGCAGAA TGA CTGcatg cctccccccg cccccctacc
*****
Alfa1B 1208 tcagtgacct agagcacttg ctgctcttgc aaaggatctg ggtttagTTC
msge 472 -----TCC
*****
Alfa1B 1258 CCAGCACCCA CAGGGTGGCT CACAACCACC TGATactcca attccagagg
msge 475 CCAGACCCTT CCCACCAAGCT CCCAACCACC AGATcccttc cgtgcagcaa
*****
*****
*****
Alfa1B 1308 attccctctt tctgatcttt gtgggctcca gataagaag gcatacaata
msge 525 gagtgaatcc cccacctccc accgtcccac catctcccac ctcccctcat
Alfa1B 1358 acacacaggt aggccaccac ttgtacacag aaaataaaaa taaaccttaa
msge 575 ctcccacctc ccattcatctc ccacctccca ccttggcttc gtgccaccac
Alfa1B 1408 agccttaaga gagggcccaa gtagccagct ttggtggcttt gaaaacatac
msge 625 agaaaaccgg acagtttcca caaactgttg ctggacaccc att-----
Alfa1B 1458 gctgagggac tggagagatg gctcagtggt taggagtacg ggttgCTTTT
msge 668 -----CGTGT
*****
*****
Alfa1B 1508 CCTGAAAGTCC TGGGTTCTGT CCCAGCACCC ACAtggaggc taaaacagtc
msge 673 CCCCAAAGTCC AGTGCACCGT CACCGTCACC TCAatcaatc accactagac
*****
*****
Alfa1B 1558 aatctccgca aactttgtgt atccaataac ttctctctgt atgcagacaa
msge 723 aaggcaggcc attccgcatc ctccaaaatc cagcaccctc ccactccact
Alfa1B 1608 aaaaatgcaga gaagacatcc aaatgtacaa acacatgat -----
msge 773 gccat aagca ggagttacgt gggcagaatg ttaaaaccaa ttttggctc
Alfa1B 1648 -----
msge 823 ccagggtccc attttgttta ccaccagcac caccaccatc atgcagatag
Alfa1B 1648 -----
msge 873 ccgtccgaca gpaaaacgga gccagtttg acccgaaaagg tgaactgtgt
Alfa1B 1648 -----
msge 923 tgtcccctca ccaaggagag atacgttctc tctgttccc aacgtactt
Alfa1B 1648 -----
msge 973 tctgtgtcac ttaagaggct gcaggtctta tgcctgtgcc gcagtttga
Alfa1B 1648 -----
msge 1023 gcccattttc taaattcaat cttgttggcc tttggcttt ggttacttc
Alfa1B 1648 -----
msge 1073 ctgcagcttt ccaattgaa tacgtgaaac accaatcggg tttgggtatc

```

Figura 11. Continuación....

```

Alfa1B 1648 -----
msge 1123 aecgtagttc gggctggttt gtggccatg aagcgcgccg ttcgcccgt

Alfa1B 1648 -----
msge 1173 teacaggttc gcctgttcgc ccgttcatt tgctcgttcg tacgctgaag

Alfa1B 1648 -----
msge 1223 ccgatgaagc gccattagt acagtgaagc gcattggcgc cccattagta

Alfa1B 1648 -----
msge 1273 cagcgaagcg cattaaagcg tcattagtac atcgtagcac atgaagcgc

Alfa1B 1648 -----
msge 1323 cattagtaca ccgaagcga tgaagcgc atcagttcgt agcctgtgaa

Alfa1B 1648 -----
msge 1373 gtgcctatga gtattttga gcccatgaag ccctatgag tattttgtag

Alfa1B 1648 -----
msge 1423 ccgatgaagc gcctatgagt atttttagc gcataagcgc cctatgagta

Alfa1B 1648 -----
msge 1473 tttttagcg catgaagcg ctattaggat ttttagcgc atgaagcgc

Alfa1B 1648 -----
msge 1523 cgttagttta tcacagcga agacgcggc gtctcagtg caatccgttc

Alfa1B 1648 -----
msge 1573 tagcgcgc acttaagagt gcagagccag tcactttgca gcgtatttgg

Alfa1B 1648 -----
msge 1623 ggagctcatc acaattcggc cagtagctca tgtgcgtctg cagctcatcg

Alfa1B 1648 -----
msge 1873 cgtgtgaccg ttttTTGG GTAAAAAGAA AAACAAAAA CAAAactgga
          *****
          *****
          *****
          *****
          *****

Alfa1B 1683 gtggagcctg gtggtactga -----
msge 1923 tcaccacacg aggatcctc tccgatgcca cgaaggatga ggggtccgc

Alfa1B 1703 -----
msge 1973 gacggccgg tccactgggt gaccgcggtc cggagaagcg tccggaccgc

Alfa1B 1703 -----
msge 2023 agcccgccga ccagacctag gccaggtcgc gtggaccttc tcggttgcct

Alfa1B 1703 -----
msge 2073 gctctccca ggtcgtagtt ggtcaggatt cttcgcacaca caccaactca

Alfa1B 1703 -----
msge 2123 agttggaac gcgtgtatcg ggagtcggga ctgcttgcct tatatagaga

Alfa1B 1703 -----
msge 2173 gatcctagga gggggggc caaatccag -GAGCTCTGA GGGTAGAGGT
          ***** GTGCTTCGC GAGCAGCGGG

Alfa1B 1722 GT CCTAGGGT CCCC GG TGG tgt -----
msge 2223 GCGGAGGG GCGCGTGG cgggcctt ggggatgg gtgggggc

Alfa1B 1745 -----
msge 2273 gccatgcact gcttggccc gccgctaggg gcagcatgg ctcaggaaag

Alfa1B 1745 -----
msge 2323 ggggtgccc gatgctgAAC TGGCCAAGCT GGGAGGa-----

Alfa1B 1778 aatgagctgg ctgggatatt ctctgtatc caggccacat cagataccca
msge 2360 -----
    
```

Figura 11. Continuación....

8. DISCUSION

En este trabajo se aisló y secuenció la región promotora del gen *SGCE* de ratón de aproximadamente 2.4 kb. La secuencia obtenida corresponde al extremo 5' del gen *SGCE* y muestra una región con un alto contenido de C/G. Cuya presencia dentro del promotor del SG- ϵ es coherente con el mecanismo de impronta materna por metilación de dinucleótidos C/G, al que se encuentra sujeto (Grabowsky et al., 2002). Al comparar las secuencias de ratón y de humano se observa que esta región se conserva, lo que confirma su posible participación en el proceso de impronta.

La composición de esta secuencia con un alto contenido de C/G dentro de la región promotora del SG- ϵ puede ser la causa de que los fragmentos generados por PCR no hayan podido ser clonados en los diferentes vectores utilizados, ya que posiblemente adoptaron estructuras secundarias que han dificultado la clonación con los diferentes plásmidos usados y/o resultaron tóxicos para la célula y eran eliminados del plásmido.

Dentro de la secuencia de 2.4 kb hacia el extremo 5' del gen *SGCE* se localizaron numerosos sitios de reconocimiento a factores generales de transcripción, sin encontrarse alguna secuencia consenso correspondiente a caja TATA, esta secuencia es reconocida por la proteína de unión a caja TATA (TBP: *TATA binding protein*), una subunidad del factor de transcripción TFIID que al interaccionar con diversos componentes de la maquinaria transcripcional participa en el adecuado ensamble y orientación de ésta. Los promotores sin caja TATA (*TATA-less*) han sido asociados con genes que se expresan de manera constitutiva, lo cual corresponde al patrón de expresión observado para ϵ -SG ya que su expresión es ubicua; En el caso de este tipo de promotores se cree que la participación de TBP ocurra a través de la interacción directa o indirecta con otras proteínas de unión secuencia-específica a ADN (Locker, 2001).

Por otra parte también se localizaron sitios de reconocimiento a factores generales de transcripción como SP1 y NF1; La familia de factores de transcripción SP se unen a secuencias de ADN ricas en C/G, cuyos sitios de unión se encuentran presentes comúnmente en secuencias regulatorias como promotores, potenciadores e islas CpG. Además de la ausencia de un consenso para caja TATA, se localizaron nueve sitios consenso para caja GC, a los cuales se une el factor de transcripción SP1, a lo largo de la secuencia del promotor de ϵ -SG. En particular SP1 puede activar la transcripción en promotores proximales o a partir de potenciadores distales (Bouwman y Philipsen, 2002), también se ha observado que algunos promotores sin caja TATA requieren sitios de unión a SP1 para tener una actividad transcripcional significativa, mediante el reclutamiento de TFIID vía SP1 (Kollel y Crawford, 2002).

NF1 (Factor Nuclear 1, por sus siglas en inglés) es una familia de cuatro genes (NF1-A, NF1-B, NF1-C y NF1-X) que codifican para factores que son necesarios para la replicación viral y se ha observado que regulan positiva o negativamente la actividad transcripcional de algunos genes (Chaudhry et al., 1998). Esto es evidente en la regulación del promotor de α -SG donde los factores NF1 actúan alternativamente como activadores o como reguladores negativos. Delgado et al., (2004a) observaron que NF1-C2 y NF1-X2 actúan como reguladores negativos en mioblastos y como activadores en miotubos, mientras NF1-A1.1 y NF1-B2 son activadores en ambos estadios de diferenciación, con lo que se observa claramente la regulación tanto positiva, como negativa a la que se encuentra sujeto α -SG. Aunque a diferencia de α -SG, ϵ -SG es de expresión constitutiva, se han observado diferentes niveles de expresión a nivel de mRNA en los diversos tejidos donde se ha estudiado y también se ha localizado la proteína ϵ -SG en regiones particulares como zonas de sinapsis (Nishiyama et al., 2004) lo que sugiere una regulación muy fina, donde se puedan combinar la presencia o interacción de diversos factores de transcripción, tales como SP1 o NF1.

En la posición -170 a -190 del promotor de ϵ -SG se identificó una secuencia a la cuál según el análisis con el programa *MatInspector* se unirían tanto SP1 como MAZ. Este mismo tipo de elemento regulador se encuentra presente dentro del promotor del gen NR1 y es responsable de la inducción después de la diferenciación neuronal (Liu et al., 2004), por lo que se sugiere que existen mecanismos específicos neuronales que permiten altos niveles de expresión de ciertos genes a través de factores de transcripción ubicuos y que podría ser el caso de ϵ -SG (Okamoto et al., 2002). De acuerdo con el patrón de expresión, la localización y las consecuencias que tienen las mutaciones en el gen *SGCE*, la participación de esta proteína en el SNC parece tener un papel muy importante, y por lo tanto resulta relevante conocer la regulación que ésta secuencia pueda tener sobre la actividad transcripcional de ϵ -SG. Además los niveles que presenta ϵ -SG en cerebro en el organismo adulto son varias veces mas altos que los registrados en músculo (Xiao y LeDoux, 2003).

También se localizaron sitios consenso para SRF, que es un regulador importante de genes asociados con ciclo celular y diferenciación, como son *c-fos*, *fosB*, *junB*, *egr-1* y *-2*, genes neuronales como *nurr1* y *nurr77* y genes musculares como actinas y miosinas. Mediante la regulación de la expresión de esos genes SRF controla el crecimiento y diferenciación celular, la transmisión neuronal, así como el desarrollo y función musculares (Chai y Tarnawski, 2002).

Además se localizaron sitios consenso para E2F, que se encuentra involucrado en la regulación de ciclo celular en conjunto con otras proteínas y miembros de la familia KLF (Kruppel-like factor) que participan en diversos procesos como proliferación y diferenciación celular, regulación de ciclo celular y apoptosis (Chen et al., 2004).

Se identificaron dentro de esta secuencia algunos sitios de reconocimiento para WT1, producto del gen de Tumor de Wilms, que es un gen supresor de tumor y está involucrado con la regulación del desarrollo del sistema urogenital

(Nakagama et al., 1995). También se identificó un sitio de reconocimiento para EGR (early growth response) que es una proteína con dominios de dedos de zinc. WT1 y EGR muestran una alta homología en los dominios con dedos de zinc y se ha observado, *in vitro*, el reconocimiento de la secuencia consenso para EGR por WT1 con alta afinidad (Nakagama et al., 1995), por lo que la funcionalidad que pudieran tener estos sitios en particular en etapas tempranas del desarrollo puede ser relevante ya que se han detectado altos niveles de expresión de la proteína ϵ -SG desde etapas embrionarias tempranas (Ettinger et al., 1997).

Por otra parte, los resultados obtenidos por Xiao y LeDoux (2003) muestran niveles de expresión del mRNA de ϵ -SG es 10 veces más alto en etapas embrionarias e inmediatas postnatales de rata que lo observado en músculo de rata adulto.

Estos resultados sugieren una regulación espacio-temporal muy fina, en la que pudieran tener funcionalidad algunos de los elementos reguladores aquí descritos. Es por ello que la realización de los ensayos funcionales con la secuencia hacia 5' del gen *SGCE* ayudará a identificar los sitios de reconocimiento que tengan una importancia relevante en la regulación transcripcional de este gen. Por el momento los sitios consenso identificados para factores generales de transcripción concuerdan con los patrones de expresión observados para este gen.

No se identificaron posible sitios consenso comunes entre las secuencias promotoras de los SG's α y ϵ ; Mientras α -SG tiene 2 cajas E (Delgado et al., 2004b), que son sitios de reconocimiento para el factor de transcripción MyoD, que es un factor de transcripción que compromete a las células al linaje muscular, ϵ -SG, en contraste, no muestra algún sitio de unión probable para factores de transcripción que determinen algún compromiso sobre un linaje celular en particular dentro de la región promotora descrita en este trabajo.

9. CONCLUSIONES

Los resultados obtenidos a partir de este trabajo muestran información sobre la composición a nivel de nucleótidos de la posible región promotora del gen *SGCE*, la cual está caracterizada por un alto contenido de C/G y que corresponde al tipo de secuencias promotoras que están sujetas al proceso de impronta por metilación de islas CpG y que es diferente a la estructura observada en el promotor muscular de α -SG.

La identificación de sitios de reconocimiento para factores de restricción, las cajas GC, sitios para SP1, dos para NF-1, dos para SRF y dos cajas GATA, entre otros corresponden también al tipo de promotor que tienen los genes de expresión constitutiva, como es el caso de ϵ -SG, no obstante, la expresión diferencial que se ha observado en diversos tejidos para esta proteína nos hace suponer que la expresión de ϵ -SG esté regulada de una manera muy fina con respecto al estadio de diferenciación y al tipo celular, en los que posiblemente participen algunos sitios consenso para factores de transcripción identificados dentro de este promotor.

Por otra parte no se encontró ninguna similitud en cuanto a la presencia de factores de transcripción en la región promotora muscular de α -SG y el promotor del ϵ -SG, lo que a su vez nos hace suponer que la regulación transcripcional a la que se encuentran sujetos estos genes debe ser diferente para cada uno de ellos, que es elocuente de acuerdo al patrón de expresión descrito para cada proteína.

10. REFERENCIAS

1. Allikian, M.J., Hack, A.A., Mewborn, S., Mayer, U. y McNally, E.M. (2004) Genetic compensation for sarcoglycan loss by integrin $\alpha 7\alpha 1$ in muscle. *J Cell Sci.* 117(17):3821-3830.
2. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. y Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res.* 25:3389-3402.
3. Araiishi, K., Sasaoka, T., Imamura, M., Noguchi, S., Hama, H., Wakabayashi, E., Yoshida, M., Hori, T. y Ozawa, E. (1999) Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in β -sarcoglycan-deficient mice. *Hum Mol Genet.* 8(9): 1589-1598.
4. Barresi, R., Moore, S.A., Stollei, C.A., Mendell, J.R., y Campbell, K. (2000) Expression of γ -sarcoglycan in smooth muscle and its interaction with the smooth muscle sarcoglycan-sarcospan complex. *J Biol Chem.* 275(49):38554-38560.
5. Blake, D., Weir, A., Newey, S. y Davies, K. (2002) Function and genetics of dystrophin and dystrophin-related proteins in muscle. *Physiol Rev.* 82:291-329.
6. Bonnemann, C.G., Modi, R., Noguchi, S., Mizuno, Y., Yoshida, M., Gussoni, E., McNally, E.M., Duggan, D.J., Angelini, C. y Hoffman EP. (1995) Beta-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex. *Nat Genet.* 11(3): 266-273.
7. Bouwman, P. y Philipsen, S. (2002) Regulation of the activity of SP1-related transcription factors. *Mol Cell Endocrinol.* 195:27-38.
8. Campbell, K.P. (1995) Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. *Cell.* 80: 675-679.
9. Chai, J. y Tarnawski, A.S. (2002) Serum response factor: discovery, biochemistry, biological roles and implications for tissue injury healing. *J Physiol Pharm.* 53(2):147-157.
10. Chan, Y.M., Bonnemann, C.G., Lidov, H.G. y Kunkel, L.M. (1998) Molecular organization of sarcoglycan complex in mouse myotubes in culture. *J Cell Biol.* 143(7):2033-2044.
11. Chaudhry, A., Vitullo, A. y Gronostajski, R. (1998) Nuclear factor 1 (NF1) isoforms differentially activate simple versus complex NF1-responsive promoters. *J Biol Chem.* 273(29):18538-18546.
12. Chen, C., Zhou, Y., Zhou, Z., Sun, X., Otto, K.B., Uth, R.M. y Dong, J.T. (2004) Regulation of KLF5 involves the Sp1 transcription factor in human epithelial cells. *Gene.* 330: 133-142.
13. Cohn, R., y Campbell, K. (2000) Molecular basis of muscular dystrophies. *Muscle Nerve.* 23:1456-1471.
14. Crosbie, R.H., Heighway, J., Venzke, D.P., Lee, J.C. y Campbell, K. (1997) Sarcospan, the 25-kDa transmembrane component of the dystrophin-glycoprotein complex. *J Biol Chem.* 272(50):31221-31224.

15. Crosbie, R.H., Lebakken, C.S., Holt, K.H., Venzke, D.P., Straub, V., Lee, J.C., Grady, R.M., Chamberlain, J.S., Sanes J.R., y Campbell, K.P. (1999) Membrane targeting and stabilization of sarcospan is mediated by the sarcoglycan subcomplex. *J Cell Biol.* 145(1):153-165.
16. Crosbie, R., Lim, L.E., Moore, S.A., Hirano, M., Hays, A.P., Maybaum, S.W., Collin, H., Dovico, Stolle. C.A. Fardeau, M., Tomé, F. y Campbell, K.P. (2000) Molecular and genetic characterization of sarcospan: insights into sarcoglycan-sarcospan interactions. *Hum Mol Genet.* 9(13):2019-2027.
17. Delgado, P., Rosas, H., Recillas, F., Zentella, A., Bermúdez de León, M., Cisneros, B., Salamanca, F. y Coral, R. (2004a) NF1-C2 negatively regulates α -sarcoglycan promoter activity in C2C12 myoblasts. *Biochem Biophys Res Comm.* 319:1032-1039.
18. Delgado, P., Rosas, H., Recillas, F., Salamanca, F. y Coral, R. (2004b) Myo-D trans-activation potential modulation through an α -sarcoglycan promoter element over myogenic differentiation. *FEBS Letters.* En revision.
19. Dennis, C.L., Tinsley, J.M., Deconinck, A.E., y Davies, K.E. (1996) Molecular and functional analysis of the utrophin promoter. *Nucl Acids Res.* 24(9):1646-1652.
20. Durand, M., Suel, L., Barbet, J.P., Beckmann, J.S. y Fougerosse, F. (2002) Sequential expression of genes involved in muscular dystrophies during human development. *Morphologie.* 86(273): 9-12.
21. Emery, A. E. H. (2002) The muscular dystrophies. *Lancet.* 359: 687-695.
22. Ettinger, A.J., Feng, G., y Sanes, J.R. (1997) ϵ -sarcoglycan, a broadly expressed homologue of the gene mutated in limb-girdle muscular dystrophy 2D. *J Biol Chem.* 272:32534-32538.
23. Grabowsky, M., Zimprich, A., Lorenz-Deperieux, B., Kalscheuer, V., Asmus, F., Gasser, T., Meitinger, T. y Strom, T. (2002) The epsilon-sarcoglycan gene (SGCE), mutated in myoclonus-dystonia syndrome, is maternally imprinted. *Eur J Hum Genet.* 11:138-144.
24. Hack, A.A., Groh, M.E. y McNally, E.M. (2000) Sarcoglycans in muscular dystrophy. *Microsc Res Tech.* 48:167-180.
25. Ibragimov-Beskrovnaya, O., Ervasti, J.M., Leveille, C.J., Slaughter, C.A., Sernett, S.W. y Campbell, K.P. (1992) Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature.* 355:696-702.
26. Imamura, M., Araishi, K., Noguchi, S. y Ozawa, E. (2000) A sarcoglycan-dystroglycan complex anchors Dp116 and utrophin in the peripheral nervous system. *Hum Mol Genet* 9(20):3091-3100.
27. Imamura, M., Mochizuki, Y., Engvall, E., y Takeda, S. (2005) ϵ -sarcoglycan compensates for lack of α -sarcoglycan in a mouse model of limb girdle muscular dystrophy. *Hum Mol Genet.*
28. Kolllel, K.J. y Crawford, D.L. (2002) Evolution of Sp transcription factors. *Mol Biol Evol.* 19(3): 216-222.

29. Lapidos, K.A., Kakkar, R. y McNally, E. (2004). The dystrophin glycoprotein complex. Signalling strength and integrity for the sarcolemma. *Circ Res.* 94:1023-1031.
30. Larsen, F., Gundersen, G., Lopez, R., y Prydz, H. (1992) CpG islands as gene markers in the human genome. *Genomics* 13(4):1095-1107.
31. Lebakken, C.S., Venzke, D.P., Hrstka, R.F., Consolino, C.M., Faulkner, J.A., Williamson, R.A. y Campbell, K.P. (2000) Sarcospan-deficient mice maintain normal muscle function. *Mol Cell Biol.* 20(5):1669-1677.
32. Lévi, S., Grady, R. M., Henry, M. D., Campbell, K. P., Sanes, J. R. y Craig, A. M. (2002) Dystroglycan Is selectively associated with Inhibitory GABAergic synapses but is dispensable for their differentiation. *The Journal of Neuroscience.* 22(11):4274-4285.
33. Lim L.E., Duclos, F., Broux, O., Bourg, N., Sunada, Y., Allamand, V., Meyer, J., Richard, I., Moomaw, C. y Slaughter, C. (1995) Beta-sarcoglycan: characterization and role in limb-girdle muscular dystrophy linked to 4q12. *Nat Genet.* 11(3):257-65.
34. Liu, A.L. y Engvall, E. (1999) Sarcoglycan isoforms in skeletal muscle. *J Biol Chem.* 274(53): 38171-38176.
35. Liu, A., Hoffman, P.W., Lu, W. y Bai, G. (2004) NF-B site interacts with Sp factors and upregulates the NR1 promoter during neuronal differentiation. *J Biol Chem.* 279(17): 17449-17458.
36. Locker, J. (2001) Transcription Factors. Academic Press. U.K. 336 pp.
37. McNally, E.M., Ly, C.T. y Kunkel, L.M. (1998) Human ϵ -sarcoglycan is highly related to α -sarcoglycan (adhelin), the limb-girdle muscular dystrophy 2D gene. *FEBS Lett.* 422:27-32.
38. Moukhles, H. y Carbonetto, S. (2001) Dystroglycan contributes to the formation of multiple dystrophin-like complexes in brain. *J Neurochem.* 78:824-834.
39. Nakagama, H., Heinrich, G., Pelletier, J. y Housman, D.E. (1995) Sequence and structural requirements for high-affinity DNA binding by the WT1 gene product. *Mol Cell Biol.* 15(3): 1489-1498.
40. Nigro, V., de la Moreira, E., Piluso, G., Vainzof, M., Belsito, A., Politano, L., Puca, A.A., Passos-Bueno, M.R. y Zatz, M. (1996) Autosomal recessive muscular dystrophy, LGMD2F, is caused by a mutation in the δ -sarcoglycan gene. *Nat Genet.* 14:195-198.
41. Nishiyama, A., Endo, T., Takeda, S. y Imamura, M. (2004) Identification and characterization of ϵ -sarcoglycans in the central nervous system. *Mol Brain Res.* 125:1-12.
42. Noguchi, S., McNally, E.M., Ben Othmane, K., Hagiwara, Y., Mizuno, Y., Yoshida, M., Yamamoto, H., Bonneman, C.G., Gussoni, E., Denton, P.H., Kyriakides, T., Middleton, L., Hentai, F., Ben Hamida, M., Nonaka, I., Vance, J.M., Kunkel, L.M. y Ozawa, E. (1995) Mutations in the dystrophin-associated γ -sarcoglycan in chromosome 13 muscular dystrophy. *Science.* 270: 819-822.
43. Noguchi, S., Wakabayashi, E., Imamura, M., Yoshida, M. y Ozawa, E. (1999) Developmental expression of sarcoglycan gene products in cultured myocytes. *Biochem Biophys Res Comm.* 262:88-93.

44. Noguchi, S., Wakabayashi, E., Sasaoka, T. y Ozawa, E. (2001) Analysis of the spatial and tissue-specific transcription of γ -sarcoglycan gene using a transgenic mouse. *FEBS Letters*. 495:77-81.
45. Okamoto, S., Sherman, K., Bai, G. y Lipton, S. (2002) Effect of the ubiquitous transcription factors, SP1 and MAZ, on NMDA receptor subunit type 1 (NR1) expression during neuronal differentiation. *Mol Brain Res*. 107:89-96.
46. Ono, R., Shiura, H., Aburatani, H., Coda, T., Kaneko-Ishino, T. y Ishino, F. (2003) Identification of a large novel imprinted gene cluster on mouse proximal chromosome 6. *Genome Res*. 13:1696-1705.
47. Piras, G., El Kharroubi, A., Kozlov, S., Escalante-Alcalde, D., Hernández, L., Copeland, N., Gilbert, D., Jenkins, N. y Stewart, C. (2000) *Zac1* (*Lot1*), a potential tumor suppressor gene, and the gene for ϵ -sarcoglycan are maternally imprinted genes: identification by a subtractive screen of novel uniparental fibroblast lines. *Mol Cell Biol*. 20(9):3308-3315.
48. Quandt, K., Frech, K., Karas, H., Wingender, E. y Werner, T. (1995) MatInd and MatInspector - New fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucl Acids Res*. 23: 4878-4884.
49. Ramirez, I., Rosas, H., Ceballos, M., Salamanca, F. y Coral, R. (2005) Expression analysis of the sarcoglycan-sarcospan complex in the vascular smooth muscle and endothelium cells of human cord vessels. *J Vasc Res*. 42(1):1-7.
50. Rando, T.A. (2001) The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. *Muscle Nerve*. 24:1575-1594.
51. Roberds, S.L., Anderson, R.D., Ibraghimov-Beskrovnaya, O. y Campbell, K.P. (1993) Primary structure and muscle-specific expression of the 50-kDa dystrophin-associated glycoprotein (adhalin). *J Biol Chem*. 268(32):23739-23742.
52. Roberds, S.L., Leturcq, F., Allamand, V., Piccolo, F., Jeanpierre, M., Anderson, R.D., Lim, L.E., Lee, J.C., Tome, F.M., Romero, N.B., Fardeau, M., Beckmann, J.S., Kaplan, J. y Campbell, K.P. (1994) Missense mutations in the adhalin gene linked to autosomal recessive muscular dystrophy. *Cell*. 78(4): 625-33.
53. Scherf, M., Klingenhoff, A. y Werner, T. (2000) Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: A novel context analysis approach. *J Mol Biol* 297(3): 599-606.
54. Shi, W., Chen, Z., Schottenfeld, J., Stahl, R.C., Kunkel, L.M. y Chan, Y. (2004) Specific assembly pathway of sarcoglycans is dependent on beta- and delta- sarcoglycan. *Muscle Nerve*. 29:409-419.
55. Straub, V., Ettinger, A.J., Durbeej, M., Venzke, D.P., Cutshall, S., Sanes, J.R. y Campbell, K. (1999) ϵ -sarcoglycan replaces α -sarcoglycan in smooth muscle to form a unique dystrophin-glycoprotein complex. *J Biol Chem*. 274:27989-27996.

56. Vaillend, C., Billard, J.M. y Laroche, S. (2004) Impaired long-term spatial and recognition memory and enhanced CA1 hippocampal LPT in the dystrophin-deficient DMD (mdx) mouse. *Neurobiol Dis.* 17(1):10-20.
57. Wakabayashi-Takai, E., Noguchi, S. y Ozawa, E. (2001) Identification of myogenesis-dependent transcriptional enhancers in promoter region of mouse γ -sarcoglycan gene. *Eur J Biochem.* 268:948-957.
58. Wheeler, M.T., Zarnegar, S. y McNally E.M. (2002) ζ -Sarcoglycan, a novel component of the sarcoglycan complex, is reduced in muscular dystrophy. *Hum Mol Genet.* 11(18):2147-2154.
59. Wheeler, M.T. y McNally E.M. (2003) Sarcoglycans in vascular smooth and striated muscle. *Trends Cardiovasc Med.* 13:238-243.
60. Xiao, J. y LeDoux, M.S. (2003) Cloning, developmental regulation and neural localization of rat ϵ -sarcoglycan. *Mol Brain Res.* 119: 132-143.
61. Yoshida, M., Hama, H., Ishikawa-Sakurai, M., Imamura, M., Mizuno, Y., Araishi, K., Wakabayashi-Takai, E., Noguchi, S., Sasaoka, T. y Ozawa, E. (2000) Biochemical evidence for association of dystrobrevin with the sarcoglycan-sarcospan complex as a basis for understanding sarcoglycanopathy. *Hum Mol Genet.* 9:1033-1040.
62. Zhu, X., Hadhazy, M., Groh, M., Wheeler, M., Wollman, R. y McNally, E. (2001) Overexpression of γ -sarcoglycan induces severe muscular dystrophy. *J Biol Chem.* 276 (24):21785-21790.
63. Zimprich, A., Grabowsky, M., Asmus, F., Naumann, M., Berg, D., Bertram, M., Sheidtmann, K., Kern, P., Winkelmann, J., Müller-Myhsok, B., Riedel, L., Bauer, M., Müller, T., Castro, M., Meitinger, T., Strom, T. y Gasser, T. (2001) Mutations in the gene encoding ϵ -sarcoglycan cause myoclonus-dystonia syndrome. *Nat Genet.* 29:66-69.