



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL



CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



CENTRO DE INVESTIGACIÓN
DEL CÁNCER

UNIVERSIDAD DE SALAMANCA

**INSTITUTO DE BIOLOGÍA MOLECULAR Y CELULAR DEL CÁNCER
(CSIC-USAL)**

ESTUDIOS ESTRUCTURALES DE LAS PROTEÍNAS DE HEMIDESMOSOMAS: INTEGRINA $\alpha 6 \beta 4$ Y TETRASPANINA CD151

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR

Noelia Alonso García

**Bajo la dirección del doctor
José María de Pereda**

Salamanca, 2013

Dr. JOSÉ MARÍA DE PEREDA VEGA, Científico Titular del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Molecular y Celular del Cáncer (IBMCC, CSIC-Universidad de Salamanca).

CERTIFICA:

Que Dña. Noelia Alonso García, licenciada en Ciencias Químicas por la Universidad de Salamanca, ha realizado bajo su dirección el trabajo de Tesis Doctoral que lleva por título "***Estudios estructurales de las proteínas de hemidesmosomas: integrina $\alpha 6\beta 4$ y tetraspanina CD151***", y considera que éste reúne originalidad y contenidos suficientes para que sea presentada ante el Tribunal correspondiente y optar al Grado de Doctor por la Universidad de Salamanca.

Y para que así conste a los efectos oportunos, expide el presente certificado en Salamanca, a 22 de Agosto de 2013.

Dr. José María de Pereda Vega

Director de la tesis

Esta Tesis ha sido realizada siendo Noelia Alonso García beneficiaria de una ayuda Predoctoral del programa JAE-predoc del CSIC durante el periodo de Julio de 2008 a Junio de 2012.

Este trabajo se ha enmarcado dentro de los proyectos del Plan nacional I+D+i "Arquitectura de hemidesmosomas: estructura e interacciones macromoleculares" (BFU2006-01929/BMC) y "Bases estructurales de interacciones en hemidesmosomas: integrina $\alpha6\beta4$, BPAG1e y tetraspanina CD151" (BFU2009-08389), financiados por el Ministerio de Ciencia e Innovación, el Ministerio de Economía y Competitividad y el Fondo Europeo de Desarrollo Regional mediante los proyectos

Noelia Alonso García realizó parte del trabajo de esta tesis durante una estancia breve de cuatro meses en el laboratorio del Prof. Arnoud Sonnenberg en el Netherlands Cancer Institute (NKI) (Ámsterdam, Holanda) gracias a una ayuda del CSIC.

Agradecimientos.

Me gustaría dar las gracias a todas las personas que han sufrido conmigo tanto la etapa de trabajo en el laboratorio como la de redacción del texto final, aunque me va a resultar imposible nombraros uno por uno.

Es obligado comenzar los agradecimientos nombrando al Dr. José María de Pereda, director de la tesis, por darme la oportunidad de conocer el mundo de la investigación, así como por las correcciones y aportaciones al trabajo final.

Gracias, Dr. Julián Gómez, por su ayuda en el trabajo con células de insecto, agradecimiento que hago extensivo a la Dra. Sara Ortega por su tiempo y dedicación.

Tienen toda mi gratitud los doctores José Manuel Andreu y Germán Rivas Caballero por hacerme sentir como en casa en el Centro de Investigaciones Biológicas del CSIC.

Agradezco al Dr. Arnoud Sonnenberg que me acogiera en el NKI para completar este trabajo, así como al Pr. Patrick Celie sus explicaciones sobre sistemas de expresión eucariota.

Mi más sincero agradecimiento al Dr. Rubén Martínez Buey por su colaboración en la medida, análisis y corrección de la parte de SAXS, así como por distintas aportaciones e ideas que han contribuido a dar sentido a este trabajo.

Quiero agradecer a la Dra. Inés García Rubio su inestimable ayuda en la realización de las medidas de EPR-DEER.

No puedo olvidar a todos aquellos que han formado parte del L-17 y del L-19 durante estos 4 años. Gracias por todos los buenos momentos que hemos vivido, tanto dentro como fuera del CIC.

Muchas gracias a todos aquellos que durante el último año me habéis preguntado para cuándo y a los que me habéis acompañado durante tantas horas en la biblioteca. No puedo nombraros a todos, ni debo dejar de nombrar a la Dra. Marisa Alonso, que me dio la oportunidad de acercar este trabajo a mucha gente.

A mi familia.

Abreviaturas.

ABD	Dominio de unión a actina
Ac	Acetato
ADN	Ácido desoxirribonucleico
ADNc	ADN complementario
ADP	Parámetros de desplazamiento atómico
BPAG1e = BP230	Antígeno 1 del penfigoide buloso
BPAG2 = BP180	Antígeno 2 del penfigoide buloso
BSA	Albúmina de suero bovino
CDB	Dominio de unión a calcio
CH	Dominio de homología a calponina
CS	Segmento conector
C-terminal	Carboxilo terminal
CW-EPR	Resonancia paramagnética electrónica de onda continua
d	Distancia
DEER	Resonancia de doble electrón-electrón
Dmax	Distancia máxima intramolecular
DTNB	Ácido 5,5´ditio-bis(2-nitrobenzoico)
DTT	Ditiotreitol
E	Coefficiente de extinción másico
EB	Epidermólisis bulosa
EBD	Epidermólisis bulosa distrófica
EBJ	Epidermólisis bulosa juntural
EBJ-H	Epidermólisis bulosa juntural-Herlitz
EBJ-nH	Epidermólisis bulosa juntural-no Herlitz
EBJ-PA	Epidermólisis bulosa juntural con atresia pilórica
EBS	Epidermólisis bulosa simple
EC	Dominio extracelular
EDO	Etilenglicol
EDTA	Ácido etilendiamino tetraacético
EGF	Factor de crecimiento epidermal
EGTA	Ácido etilenglicol tetraacético
EMTS	Etilmercuriotiosalicilato sódico
Endo Hf	Endoglicosilasa H
EPR	Resonancia paramagnética electrónica
Erk	Quinasas reguladas por estímulos extracelulares
F	Señal de fluorescencia
F calc	Amplitud del factor de estructura calculada a partir de un modelo
F obs	Amplitud del factor de estructura observada experimentalmente
F _{hkl}	Factor de estructura
FLIM	Microscopía de fluorescencia de tiempo de vida (<i>Fluorescence-lifetime imaging microscopy</i>)
FnIII	Dominio fibronectina tipo III
FRET	Transferencia de energía de resonancia de Förster
GABEB	Epidermólisis bulosa benigna atrófica generalizada
GP67	Secuencia señal de la proteína gp67
GSR	Motivo de secuencia Gly-Ser-Arg
HDs	Hemidesmosomas
HEPES	Ácido 2-[4-(2-hidroxietil)-1-piperacilil-(1)] etanosulfónico
hGH	Hormona de crecimiento humana

HM	Melitina de abeja
I	Intensidad
ICP-OES	Espectroscopia de emisión óptica de plasma acoplado inductivamente
IP	Placa interna
IPTG	Isopropil- β -D-1-tiogalactopiranosido
Kd	Constante de disociación
LD	Lámina densa
LL	Lámina lúcida
MBP	Proteína de unión a maltosa
MTSSL	(2,2,5,5-tetrametilpirrolina-1-oxyl-3-metil) metanotiosulfonato
NaPi	Fosfato sódico
NCX	Intercambiador de sodio/calcio
NSD	Discrepancia espacial normalizada
N-terminal	Amino terminal
OG	Oregon Green 488
OP	Placa externa
P(r)	Función de distribución de distancias interatómicas
PCR	Reacción en cadena de la polimerasa (Polymerase Chain Reaction)
PDB	Banco de datos de proteínas (Protein Data Bank)
PE	Pentaetilenglicol
PEG	Polietilenglicol
Pen	Penicilina
PMSF	Fluoruro de fenilmetilsulfonilo (Phenylmethylsulfonyl Fluoride)
PNGasa F	N-glicosilasa F
PRD	Domino de repetición plakina
PTB	Dominio de unión a fosfotirosina
q	Vector de dispersión
r	Distancia
r.m.s.d.	Desviación cuadrática media (Root Mean Square Deviation)
R	Factor R de desacuerdo
Rg	Radio de giro
RMN	Resonancia magnética nuclear
RNAsa A	Ribonucleasa I pancreática
SAXS	Dispersión de rayos-X a bajo ángulo (Small Angle X-ray Scattering)
SDS	Dodecilsulfato sódico
SDSL	Marcaje paramagnético dirigido
SDS-PAGE	Electroforesis en gel de poliacrilamida con dodecilsulfato sódico
SIRAS	Reemplazo isomórfico simple con dispersión anómala
SOC	Medio super-óptimo con represión por catabolitos
SR	Repetición de espectrina (Spectrin Repeat)
SS	Secuencia señal
Strep	Estreptomina
TEV	Virus del grabado del tabaco
TLS	Traslación, libración y desplazamiento helicoidal (Translation /Libration/Screw)
TM	Segmento transmembrana
Tm	Temperatura de fusión o de desnaturalización
Tris	Tris(hidroximetil)-aminimetano
Trx	Tiorredoxina
UV	Ultravioleta
V	Volumen

V_M	Coeficiente de Matthews
X-gal	5-bromo-4-cloro-3-indolil- β -D-galactopiranosido
ϵ	Coeficiente de extincion molar
χ^2	Function de discrepancia normalizada

Índice general.

INTRODUCCIÓN	1
1.- Hemidesmosomas	3
2.- Componentes de los hemidesmosomas	5
2.1.- BP180.....	5
2.2.- Plakinas de HDs: Plectina y BPAG1e.....	6
2.3.- Integrina $\alpha 6\beta 4$	7
2.3.1.- Estructura de la región citoplasmática de la subunidad $\beta 4$	10
2.3.1.1.- El dominio Calx- β	10
2.3.1.2.- Los dominios FnIII	12
2.3.1.3.- Organización de otras regiones del dominio citoplasmático de $\beta 4$	13
2.4.- Tetraspanina CD151	14
2.4.1.- Estructura de CD151 y otras tetraspaninas.....	16
3.- Organización de los HDs.	19
3.1.- Interacciones de la región citoplasmática de $\beta 4$ con otros componentes de los HDs.	19
3.2.- Bases estructurales de la interacción entre $\alpha 6\beta 4$ y plectina	20
3.3.- Ensamblaje y regulación de HDs.	22
OBJETIVOS	25
MATERIALES Y MÉTODOS	29
1.- Clonación de fragmentos de $\beta 4$ en vectores de expresión en bacterias	31
2.- Mutagénesis dirigida	32
3.- Expresión de fragmentos de $\beta 4$ en bacterias y purificación	34
3.1.- Purificación del dominio Calx- β	35
3.2.- Purificación de fragmentos de la región FnIII-3,4 de $\beta 4$	35
4.- Cristalografía	36
4.1.- Cristalización de fragmentos de $\beta 4$	36
4.1.1.- Ensayos iniciales	36
4.1.2.- Cristalización del dominio Calx- β	37
4.1.3.- Cristalización del dominio FnIII-3	37
4.1.4.- Cristalización del dominio FnIII-4	37
4.2.- Toma de datos de difracción.....	38
4.3.- Procesado de los datos de difracción	38
4.4.- Resolución estructural: faseado.....	38
4.4.1.- Reemplazo molecular: faseado de los dominios Calx- β y FnIII-3 de $\beta 4$	39
4.4.2.- Reemplazo isomórfico simple con dispersión anómala (SIRAS): faseado del dominio FnIII-4 de $\beta 4$	40
4.5.- Refinado de las estructuras cristalográficas	40
4.6.- Análisis de las estructuras cristalográficas.....	41

5.- Dispersión de rayos-X a bajo ángulo (SAXS).....	42
5.1.- Adquisición de datos de SAXS	42
5.2.- Procesado y análisis de datos de SAXS	43
5.3.- Generación de modelos a partir de datos de SAXS	44
5.4.- Comparación de modelos atómicos con los datos de SAXS.....	45
6.- Medida de distancias intramoleculares mediante resonancia paramagnética electrónica (EPR).....	45
6.1.- Marcaje de proteínas con sondas paramagnéticas	45
6.2.- Valoración de grupos tiol libres con DTNB.....	46
6.3.- Experimentos de resonancia paramagnética electrónica (EPR).....	47
7.- Producción de CD151-EC2 en bacterias.....	49
7.1.- Clonación del EC2 de CD151 en vectores de expresión en bacterias.....	49
7.1.1.- Clonación de CD151-EC2 en los vectores pET15b, pETEV15b y pETEZZ	50
7.1.2.- Clonación de CD151-EC2 en el vector pET32aNco	51
7.1.3.- Clonación de CD151-EC2 en el vector pET22b	53
7.2.- Expresión de CD151-EC2 en bacterias y purificación	53
8.- Producción de CD151-EC2 en células de insecto	54
8.1.- Clonación de CD151-EC2 en vectores de transferencia para baculovirus	54
8.1.1.- Clonación de CD151-EC2 en el vector pAcGP67A.....	54
8.1.2.- Clonación de CD151-EC2 en el vector pFastBac-1 y derivados.	55
8.2.- Obtención de bácmidos en células DH10Bac.....	59
8.3.- Obtención de baculovirus por transformación en células Sf9	60
8.4.- Amplificación de baculovirus recombinantes	62
8.5.- Expresión del EC2 de CD151 en células de insecto y purificación.....	62
8.6.- Análisis de deglicosilación de CD151-EC2	63
9.- Métodos de caracterización de proteínas.....	64
9.1.- Cuantificación de la concentración de proteínas	64
9.2.- Ensayo de estabilidad térmica basado en fluorescencia	65
9.3.- Análisis de unión de calcio al dominio Calx- β mediante diálisis de equilibrio	65
9.4.- Cromatografía de exclusión molecular	66
9.5.- Dicroísmo circular	66
RESULTADOS.....	67
1.- Caracterización del dominio Calx-β de la integrina β4	69
1.1.- Estructura cristalográfica del dominio Calx- β de β 4.....	69
1.2.- Comparación del dominio Calx- β de β 4 con los dominios de unión a Ca^{2+} del NCX1	72
1.3.- Análisis del pseudo-sitio de unión a Ca^{2+} del dominio Calx- β de β 4.....	75
1.4.- Papel de R1003 en la estabilidad de dominio Calx- β de β 4	81
2.- Estructura de la región FnIII-3,4 de la integrina β4 y su interacción con BPAG1e.....	81
2.1.- Estructuras cristalográficas de los dominios FnIII-3 y FnIII-4 aislados.....	81
2.1.1.- Estructura cristalográfica del dominio FnIII-3	82
2.1.2.- Estructura cristalográfica del dominio FnIII-4	87

2.1.3.- Comparación de las estructuras de los cuatro dominios FnIII de la integrina β_4	91
2.2.- Estructura en solución de la región FnIII-3,4, análisis mediante dispersión de rayos-x a bajo ángulo (SAXS)	95
2.2.1.- Caracterización estructural de la región FnIII-3,4 en solución	95
2.2.2.- Estructura de la región FnIII-3, 4 en solución	97
2.3.- Análisis de la orientación relativa de los dominios FnIII-3 y FnIII-4 mediante SDSL-EPR-DEER	99
3.- Caracterización del dominio EC2 de la tretaspanina CD151	112
3.1.- Expresión de CD151-EC2 en bacterias	112
3.1.1.- Obtención de CD151-EC2 soluble replegado.	114
3.1.2.- Caracterización de CD151-EC2 replegado.....	116
3.1.2.1.- Análisis del estado de asociación del CD151-EC2 replegado	116
3.1.2.2.- Análisis del estado de oxidación de las Cys del CD151-EC2	117
3.1.2.3.- Contenido de estructura secundaria de CD151-EC2 replegado	118
3.1.2.4.- Ensayos de cristalización del CD151-EC2 replegado.....	119
3.2.- Expresión de CD151-EC2 en células de insecto.....	120
3.2.1.- Obtención de CD151-EC2 a partir de células de insecto.....	122
3.2.2.- Ensayos de cristalización del CD151-EC2 expresado en células de insecto.	125
DISCUSIÓN	127
1.- Estructura de la región citoplasmática de la integrina β_4	129
1.1.- El dominio Calx- β	129
1.1.1.- El dominio Calx- β de β_4 no une Ca^{2+}	129
1.1.2.- Papel de R1003 y R1077 en el dominio Calx- β	130
1.1.3.- Organización del dominio Calx- β en el contexto de la región citoplasmática de β_4	131
1.1.4.- Posible función del dominio Calx- β	132
1.2.- La región FnIII-3,4	133
1.2.1.- Estructura de la región FnIII-3,4	133
1.2.2.- Interacciones de FnIII-3,4 con proteínas de los HDs.....	134
1.2.3.- Interacciones de FnIII-3,4 con dominios PTB y SH2.....	136
1.2.4.- Contribución de FnIII-3,4 a la organización de otras regiones de β_4	139
1.2.5.- Potencial participación de FnIII-3,4 en cambios conformacionales en β_4	140
2.- CD151.	141
CONCLUSIONES	147
BIBLIOGRAFÍA	151
ANEXOS	167

INTRODUCCIÓN

1.- Hemidesmosomas

Los hemidesmosomas (HDs) son complejos multiproteicos especializados que median la adhesión estable de células epiteliales a la membrana basal subyacente, enlazando la matriz extracelular al sistema de filamentos intermedios del citoesqueleto (Koster et al., 2004a). En función de los componentes que forman estos complejos, se distinguen dos clases de HDs: tipo I y tipo II (Litjens et al., 2006). Epitelios complejos estratificados y pseudo-estratificados, como la piel, contienen HDs de tipo I que están formados por cinco proteínas, tres de ellas transmembrana: la integrina $\alpha 6\beta 4$ (Jones et al., 1991; Sonnenberg et al., 1991; Stepp et al., 1990), el antígeno 2 del penfigoide buloso (BPAG2 o BP180) (Giudice et al., 1992) y la tetraspanina CD151 (Sterk et al., 2000) y dos proteínas citoplasmáticas pertenecientes a la familia de las plakinas: plectina (McLean et al., 1996; Wiche et al., 1991) y BPAG1e (antígeno 1 del penfigoide buloso, también denominado BP230) (Stanley et al., 1988) (Figura I1). Epitelios simples como el intestino, contienen HDs de tipo II que son más rudimentarios y que están formados únicamente por la integrina $\alpha 6\beta 4$ y plectina (Nievers et al., 1999; Uematsu et al., 1994).

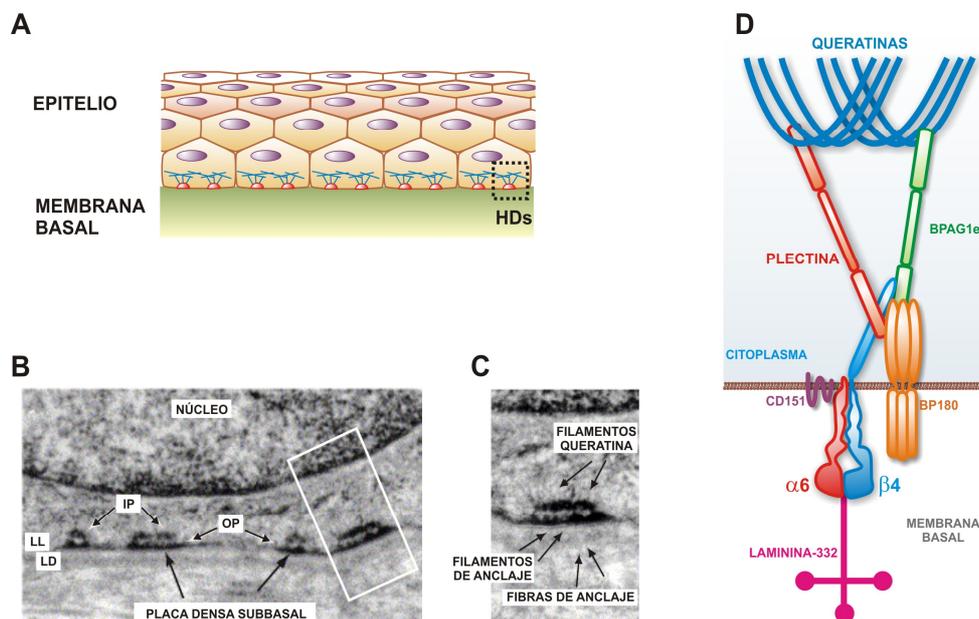


Figura I1.- Organización estructural de los hemidesmosomas (A) Representación esquemática de la localización de los HDs en la capa basal de células epiteliales donde median el anclaje a la membrana basal. (B) Micrografía electrónica de una célula epitelial con varios HDs. En función de la densidad electrónica en los HDs se distinguen varias zonas: lámina lúcida (LL), lámina densa (LD), placa interna (IP) y placa externa (OP) en la región intracelular (adaptado de (Litjens et al., 2006)). (C) Aumento del área dentro del rectángulo blanco del panel B (adaptado de (Litjens et al., 2006)). (D) Representación esquemática de los componentes de los HDs de tipo I y las interacciones que unen los receptores de adhesión al citoesqueleto de queratinas.

Alteraciones en los HDs están ligadas a varias enfermedades humanas caracterizadas por desórdenes y fragilidad de la piel. Estas patologías pueden ser hereditarias, inmunes o causadas por la combinación de factores genéticos y hereditarios.

Alteraciones en los genes que codifican proteínas de los HDs causan un grupo de genodermatosis denominadas epidermolisis bullosa o ampollar (EB) caracterizadas a la separación entre dermis y epidermis (Sawamura et al., 2010). La EB se puede clasificar en tres grandes grupos en función del punto de la zona de la membrana basal cutánea donde se produce la ruptura (Fine et al., 2008): en la EB simple (EBS) la separación se produce dentro de los queratinocitos basales; la EB juntural (EBJ) afecta a la membrana basal, produciéndose generalmente la ruptura en la lámina lúcida; en la EB distrófica (EBD) la ruptura ocurre en la dermis a nivel de las fibras de anclaje. Por último, el síndrome de Kindler constituye un cuarto tipo en que se observa mezcla de algunas de los anteriores (Brown et al., 1995).

Los diferentes tipos de EB se corresponden con alteraciones en proteínas específicas (Tabla I1). Así, la EBS basal está ligada a mutaciones en las queratinas K5 y K14 de queratinocitos basales, con mutaciones en plectina y BPAG1e que median la asociación de queratinas con HDs o con mutaciones específicas en la integrina $\beta 4$ en sitios donde interacciona con BP180 (Fine et al., 2008; Groves et al., 2010; Jonkman et al., 2002). Pacientes de EBJ presentan mutaciones en la laminina-332, BP180 o la integrina $\alpha 6\beta 4$. El subtipo más severo de EBJ, denominado Herlitz (EBJ-H), está causado por mutaciones que resultan en la pérdida de expresión de la laminina-332. Otras formas no Herlitz de EBJ (EBJ-nH) son el resultado de mutaciones sin sentido en las subunidades de laminina o en BP180. Mutaciones sin sentido o pérdida de expresión de la integrina $\alpha 6\beta 4$ causa un tercer subtipo de EBJ con atresia pilórica (EBJ-PA) (Fine et al., 2008; Sawamura et al., 2010). La EBD está causada principalmente por alteraciones en el colágeno tipo VII. Se ha descrito una mutación en el gen de la tetraspanina CD151 en una variante de EBD poco frecuente, EBD pretibial, asociada con nefritis hereditaria y β -talasemia menor (Pulkkinen et al., 1996).

Tabla I1. Principales tipos de EB, genes y proteínas afectadas.

Tipo de EB	Subtipo	Genes alterados	Proteínas alteradas
Simple (EBS)	Suprabasal	<i>PKP1</i> <i>DSP</i>	Plakofilina-1 Desmoplakina
	Basal	<i>KRT5</i> <i>KRT14</i> <i>PLEC1</i> <i>ITGA6, ITGB4</i> <i>DST</i>	Queratina-5 (K5) Queratina-14 (K14) Plectina Integrina $\alpha6\beta4$ BPAG1e
Juntural (EBJ)	Herlitz (EBJ-H)	<i>LAMA3, LAMB3, LAMC2</i>	Laminina-332
	No Herlitz (EBJ-nH)	<i>LAMA3, LAMB3, LAMC2</i> <i>COL17A1</i>	Laminina-332 BP180 o colágeno tipo XVII
	Con atresia pilórica (EBJ-PA)	<i>ITGA6, ITGB4</i>	Integrina $\alpha6\beta4$
Distrófica (EBD)	-	<i>COL7A1</i> <i>CD151</i>	Colágeno tipo VII Tetraspanina CD151
Síndrome de Kindler	-	<i>FERMT1 (KIND1)</i>	Kindlin-1

Los componentes de los HDs también están implicados en un grupo de enfermedades autoinmunes denominadas penfigoides que incluyen el penfigoide buloso y el penfigoide de membrana mucosa (Schmidt and Zillikens, 2013). El penfigoide buloso está ligado a autoanticuerpos frente a BP180 y BPAG1e, mientras que el penfigoide de membrana mucosa está asociado a autoanticuerpos frente a BP180, BPAG1e, laminina-332, integrina $\alpha6\beta4$ y colágeno tipo VII.

2.- Componentes de los hemidesmosomas

2.1.- BP180

BP180, también conocido como el antígeno 2 del penfigoide buloso (BPAG2) o colágeno tipo XVII, es una proteína transmembrana de tipo II formada por un dominio amino terminal intracelular de 466 residuos, un único segmento transmembrana y un dominio extracelular de 1008 aminoácidos (Giudice et al., 1992). BP180 se asocia con los filamentos de anclaje de los HDs en los queratinocitos basales (Franzke et al., 2003; Van den Bergh and Giudice, 2003; Zimina et al., 2005) y juega un papel fundamental en

adhesión epidermal. Mutaciones en el gen de BP180 (*COL17A*), que resultan en una carencia de esta proteína, causan EBJ no Herlitz generalizada (conocida anteriormente como EB benigna atrófica generalizada o GABEB), caracterizada por la formación de ampollas en articulaciones y huesos, alopecia, distrofia en uñas e hipoplasia en esmaltes dentales (Darling et al., 1997).

En los HDs el dominio intracelular de BP180 interacciona con la región citoplasmática de la subunidad $\beta 4$ de la integrina $\alpha 6\beta 4$ (Aho and Uitto, 1998b; Koster et al., 2003), con BPAG1e (Borradori et al., 1998; Hopkinson and Jones, 2000) y con plectina (Koster et al., 2003). El dominio extracelular contiene 15 dominios de colágeno y 16 segmentos no colagenosos (Franzke et al., 2005; Giudice et al., 1992). Este ectodominio media la formación de homotrímeros de BP180 e interacciona con la laminina-332 (Reddy et al., 1998; Tasanen et al., 2004) y con la subunidad $\alpha 6$ de la integrina $\alpha 6\beta 4$ (Hopkinson et al., 1995; Hopkinson et al., 1998).

2.2.- Plakinas de HDs: Plectina y BPAG1e.

Las plakinas son una familia de proteínas multifuncionales que median el anclaje de los filamentos intermedios a la membrana plasmática a través de complejos de adhesión asociados a la membrana como los desmosomas y los HDs (Jefferson et al., 2004). La expresión de plakinas es crítica para el crecimiento, desarrollo y supervivencia celular (Brown et al., 1995). En humanos, anomalías en plakinas causan cardiomiopatías, distrofias musculares y enfermedades de la piel (Pulkkinen et al., 1996; Ruhrberg et al., 1996).

Plectina y BPAG1e son proteínas de alta masa molecular que tienen una estructura tripartita formada por un dominio central alargado que separa las regiones N- y C-terminales que contienen sitios de unión a otras proteínas (Figura I2). La región N-terminal contiene una zona de ~1000 aminoácidos denominada dominio plakina que está conservada en casi todas las plakinas. El dominio plakina está formado por nueve dominios de repetición de espectrina (*spectrin repeats*, SR) en el caso de plectina y ocho SR en el caso de BPAG1; además contiene un dominio SH3 que se localiza en el quinto SR

(Jefferson et al., 2007; Ortega et al., 2011; Sonnenberg et al., 2007). Precediendo al dominio plakina plectina contiene un dominio de unión a actina (ABD) que está formado por dos dominios de homología a calponina (CH1 y CH2) (Garcia-Alvarez et al., 2003; Sevcik et al., 2004).

La región central tiene forma de varilla alargada, es rica en α -hélice y media la homodimerización mediante interacciones de tipo coiled-coil (Green et al., 1992). La región C-terminal, implicada en la interacción con filamentos intermedios, contiene varias copias del denominado "dominio de repetición plakina" (PRD). Los PRDs se clasifican en tres tipos (A, B y C) por su similitud con los tres PRDs presentes en desmoplakina. Plectina contiene además cerca de su extremo carboxilo un dominio GSR (Gly-Ser-Arg) implicado en la unión a microtúbulos.

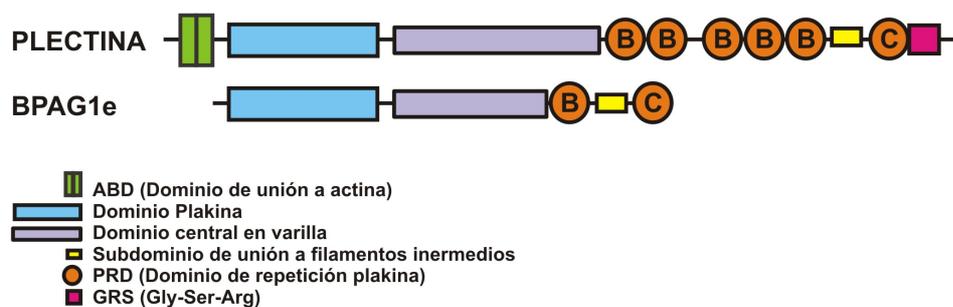


Figura 12.- Estructura modular de plectina y BPAG1e. Representación esquemática de los dominios que componen estas plakinas.

2.3.- Integrina $\alpha 6\beta 4$.

Las integrinas son una familia de receptores de adhesión de la superficie celular que transmiten señales bidireccionales entre el citoplasma y la matriz extracelular (Hynes, 2002; van der Flier and Sonnenberg, 2001). Las integrinas son receptores heterodiméricos de membrana formados por subunidades α y β , unidas de forma no covalente. Estas subunidades son glicoproteínas transmembrana de tipo I formadas por una región extracelular implicada en el reconocimiento de ligandos, una hélice- α transmembrana y colas citoplasmáticas generalmente cortas que contienen la región terminal y son responsables de la interacción con proteínas de señalización y del

citoesqueleto. (Arnaout et al., 2007; Campbell and Humphries, 2011; Wegener and Campbell, 2008).

En vertebrados, existen 18 subunidades α y ocho subunidades β que se combinan para formar 24 receptores $\alpha\beta$ (Barczyk et al., 2010; Hynes, 2002) (Figura I3). Los heterodímeros se expresan en diversos tipos celulares, como células del epitelio y del endotelio, fibroblastos, células hematopoyéticas, neuronas y células del músculo (Danen and Sonnenberg, 2003). Las integrinas modulan aspectos del comportamiento celular, como adhesión, movilidad, proliferación, señalización y diferenciación. La mayoría de las integrinas se unen a los componentes de la matriz extracelular, como lamininas, colágenos y fibronectina; si bien otras integrinas unen receptores celulares y median contactos célula-célula (van der Flier and Sonnenberg, 2001).

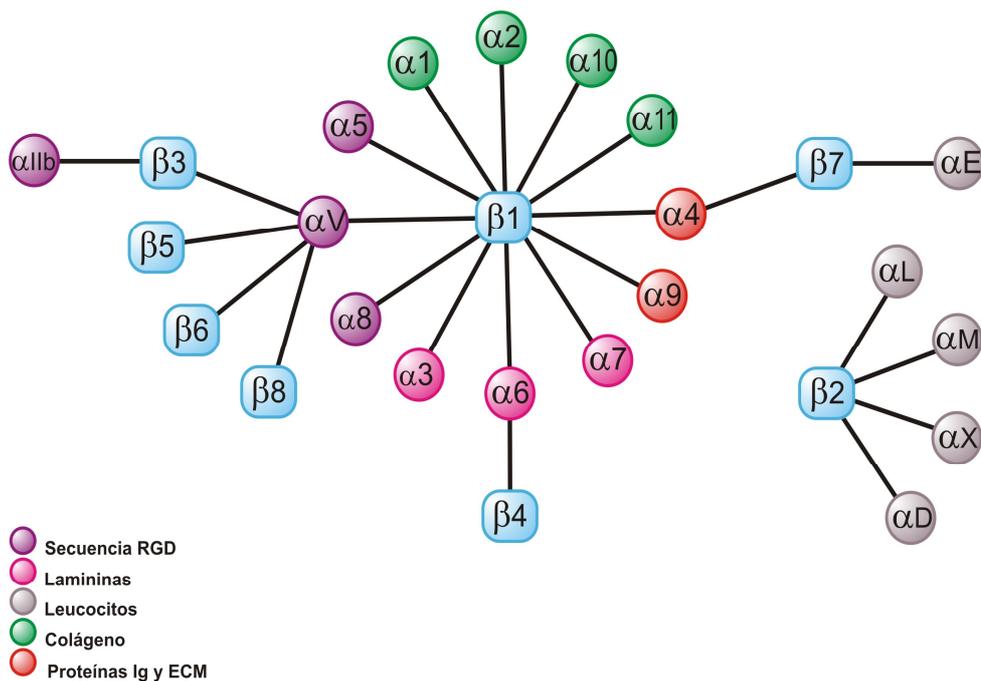


Figura I3.- Familia de receptores de integrinas. Representación esquemática de las 24 integrinas de mamíferos clasificadas por colores en función de los receptores específicos de cada una de ellas.

La subunidad $\beta 4$ (también conocida como el antígeno CD104 o GP150) sólo existe en combinación con $\alpha 6$ para formar la interina $\alpha 6\beta 4$. La región extracelular de $\alpha 6\beta 4$ se une

a lamininas de la membrana basal y muestra preferencia hacia la laminina-332 (laminina-5) que es un componente mayoritario de la membrana basal de la epidermis (Tsuruta et al., 2008). La región intracelular de la subunidad $\alpha 6$ es de pequeño tamaño (~50 aminoácidos), mientras que el dominio citoplasmático de $\beta 4$ es mucho más largo que en las demás subunidades β de otras integrinas y no tiene similitud de secuencia con éstas. En epitelios, $\alpha 6\beta 4$ se localiza en la capa basal de la membrana plasmática, donde forma parte de los HDs (Wilhelmsen et al., 2006). A diferencia de otras integrinas que se asocian a filamentos de actina, $\alpha 6\beta 4$ lo hace a filamentos de queratina a través de otros componentes de los HDs como plectina y BP230. Además de su papel en la formación y función de los HDs a través de la adhesión, $\alpha 6\beta 4$ favorece la migración de queratinocitos durante la cicatrización de heridas y promueve la invasión de carcinomas (Giancotti, 2007; Lipscomb et al., 2005), participa en el control del crecimiento celular (Mainiero et al., 1995) y en rutas de señalización implicadas en la proliferación y migración de queratinocitos y en invasión y supervivencia de células epiteliales y células tumorales (Wilhelmsen et al., 2006).

En humanos, mutaciones en los genes que codifican para la subunidad $\alpha 6$ (*ITGA6*) o $\beta 4$ (*ITGB4*) causan EBJ asociada con atresia pilórica (EBJ-PA) (Chung and Uitto, 2010). Esta enfermedad se caracteriza por fragilidad de la piel con formación de ampollas por separación entre la dermis y la epidermis; y la formación de HDs más rudimentarios (McMillan et al., 1998).

La inactivación en ratón del gen de la subunidad $\alpha 6$ o del de la subunidad $\beta 4$ causa un daño severo en la piel y otros epitelios caracterizado por la formación generalizada de ampollas. A nivel ultraestructural la piel de estos animales no contiene HDs y se observa una separación entre dermis y epidermis tanto a nivel de la membrana basal como en la cara basal del interior de los queratinocitos (Dowling et al., 1996; Georges-Labouesse et al., 1996; van der Neut et al., 1996). En resumen, la integrina $\alpha 6\beta 4$ es un componente esencial de los HDs y es necesaria para mantener la resistencia mecánica de epitelios.

2.3.1.- Estructura de la región citoplasmática de la subunidad $\beta 4$.

La integrina $\beta 4$, se caracteriza por una región citoplasmática inusualmente larga (~1000 residuos) comparada con las de otras subunidades β de integrinas que contienen ~50 aminoácidos (Hogervorst et al., 1990; Suzuki and Naitoh, 1990; Tamura et al., 1990). La región citoplasmática de $\beta 4$ presenta una estructura modular que contiene cinco dominios globulares (Figura I4): un dominio próximo a la membrana, denominado Calx- β (Schwarz and Benzer, 1997) y cuatro dominios fibronectina tipo III (FnIII). Los dominios FnIII se disponen en dos pares (FnIII-1,2 y FnIII-3,4). El FnIII-2 y el FnIII-3 están separados por una región de ~140 aminoácidos denominada "segmento conector" (CS). Tras el FnIII-4 se extiende una cola C-terminal de ~86 residuos.

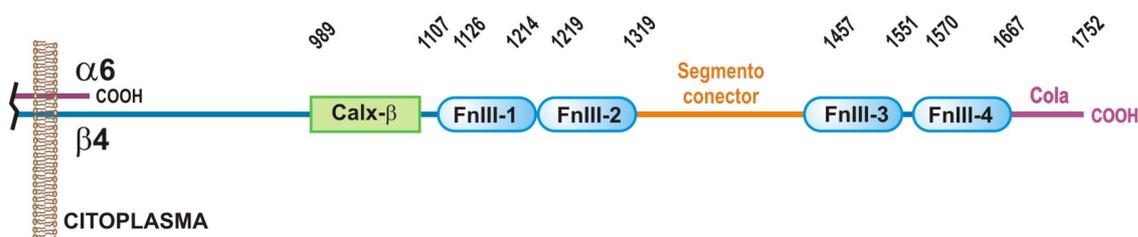


Figura I4.- Organización de la región citoplasmática de la integrina $\alpha 6\beta 4$.

2.3.1.1.- El dominio Calx- β

El dominio Calx- β se identificó primero en la región citoplasmática de los intercambiadores de $\text{Na}^+/\text{Ca}^{2+}$ (NCX) (Schwarz and Benzer, 1997), una familia de proteínas de membrana-plasma que contribuyen a la homeostasis de Ca^{2+} , exportando Ca^{2+} del citoplasma (Lytton, 2007). Los NCXs tienen un dominio intracelular grande que contiene dos copias en tandem del motivo Calx- β , que se corresponden con dos dominios de unión a Ca^{2+} (CBD1 y CBD2) (Hilge et al., 2006). Los motivos Calx- β , además de encontrarse en los NCXs y en $\beta 4$, están presentes en el ectodominio del receptor acoplado a proteína G VLGR1 (Nikkila et al., 2000), el factor de agregación MAFp3 de la esponja marina *Microciona prolifera* (Fernandez-Busquets et al., 1996), la proteína de matriz extracelular ECM3 del erizo de mar (Hodor et al., 2000), miembros de la familia 12-CSPG de proteínas de matriz extracelular (Kiyozumi et al., 2007), la glucosidasa- β

BgIM1 unida a la membrana de *Physarum polycephalum* (Hayase et al., 2008) y proteínas procedentes de la cyanobacteria *Synechocystis* sp. PCC6803 (Schwarz and Benzer, 1997).

Las estructuras de CBD1 y CBD2 del intercambiador 1 de $\text{Na}^+/\text{Ca}^{2+}$ (NCX1) se han determinado por NMR y por cristalografía de rayos-X (Figura I5) (Besserer et al., 2007; Hilge et al., 2006; Nicoll et al., 2006). La estructura del dominio Calx- β consiste en un sandwich- β formado por dos láminas- β ; este plegamiento se encuadra dentro de la superfamilia de las inmunoglobulinas. Los CBDs unen 2 ó 4 iones Ca^{2+} , próximos unos a los otros y localizados en uno de los extremos del sandwich- β , donde los iones Ca^{2+} están coordinados por residuos ácidos. A partir de la comparación de secuencias de $\beta 4$ con las de los CBDs de NCX1, los sitios de unión a Ca^{2+} no están totalmente conservados en el dominio Calx- β de $\beta 4$, sugiriendo que éste no une Ca^{2+} u otros cationes divalentes (Hilge et al., 2006).

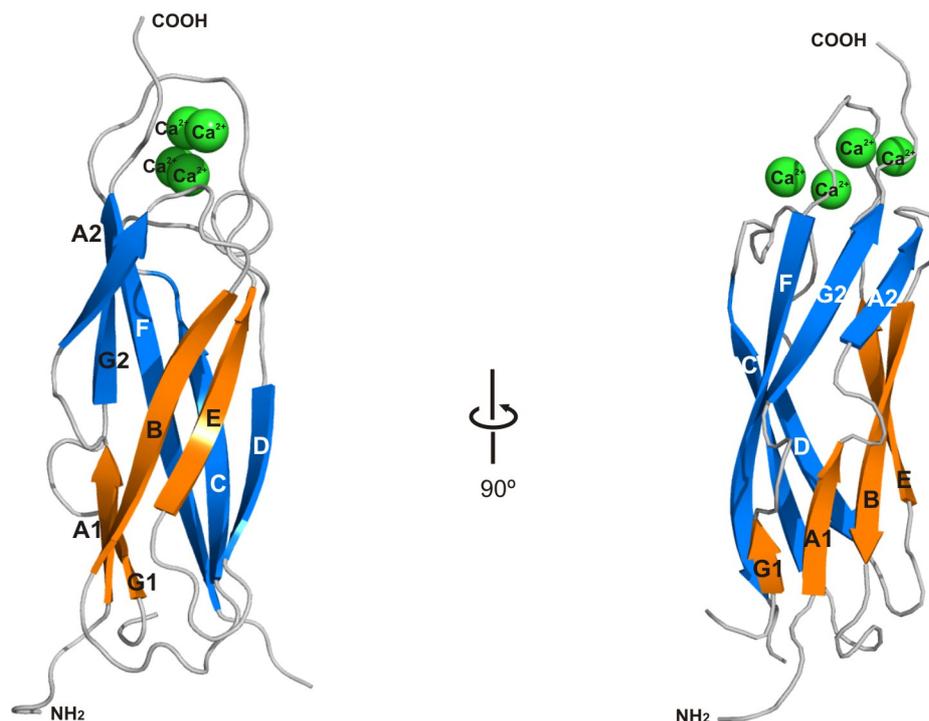


Figura I5.- Estructura del dominio Calx- β de unión a Ca^{2+} (CBD1) de NCX1. Representación de cintas de la estructura cristalográfica del dominio CBD1 de NCX1 (código del PDB: 2PDK). La estructura contiene cuatro iones Ca^{2+} que se muestran como esferas verdes.

2.3.1.2.- Los dominios FnIII

Los dominios FnIII, inicialmente descritos en fibronectina, de donde reciben su denominación, están presentes en gran cantidad de proteínas y especies desde bacterias a mamíferos. Al igual que el Calx- β , los FnIII tienen similitudes estructurales con dominios de la superfamilia de las inmunoglobulinas (Campbell and Spitzfaden, 1994). Los FnIII son dominios de ~100 aminoácidos formados por 7 hebras- β distribuidas en 2 láminas, A-B-E y C-C'-F-G, empaquetadas en un sandwich- β que encierra un núcleo hidrofóbico (Figura I6). Los extremos amino y carboxilo de cada FnIII se localizan en los polos opuestos del eje longitudinal, por lo que cuando aparecen múltiples FnIII en la secuencia de proteínas, éstos forman estructuras alargadas en cadena, como se ha visto en tenascina (Bisig et al., 1999; Hu et al., 2007) y en titina (Bucher et al., 2010), si bien la orientación relativa entre dominios contiguos no está conservada.

Se conoce la estructura cristalográfica del primer par de dominios FnIII de β 4 (código PDB: 1QG3) (de Pereda et al., 1999). El FnIII-1 y el FnIII-2 se disponen de forma consecutiva en la secuencia de β 4. En el cristal estos dos dominios adoptan una estructura alargada en la que la interfase entre el FnIII-1 y el FnIII-2 parece estabilizar la disposición relativa de los dos dominios (Figura I6). De hecho esta misma orientación aparece en otras dos estructuras cristalográficas de fragmentos de β 4 que contienen el FnIII-1,2 y secuencias crecientes del CS (códigos PDB: 3F7Q y 3F7R) (de Pereda et al., 2009b). Además, se ha observado mediante dispersión de rayos-x a bajo ángulo (SAXS) que el FnIII-1,2 adopta en solución una estructura alargada compatible con la observada en las estructuras cristalográficas (Chacon et al., 2000).

A diferencia del tándem FnIII-1,2, el FnIII-3 y el FnIII-4 están conectados por una secuencia de ~20 aminoácidos rica en Pro, la cual se cree que puede actuar como un conector flexible que permita movilidad entre los dos dominios y la adopción de conformaciones no extendidas.

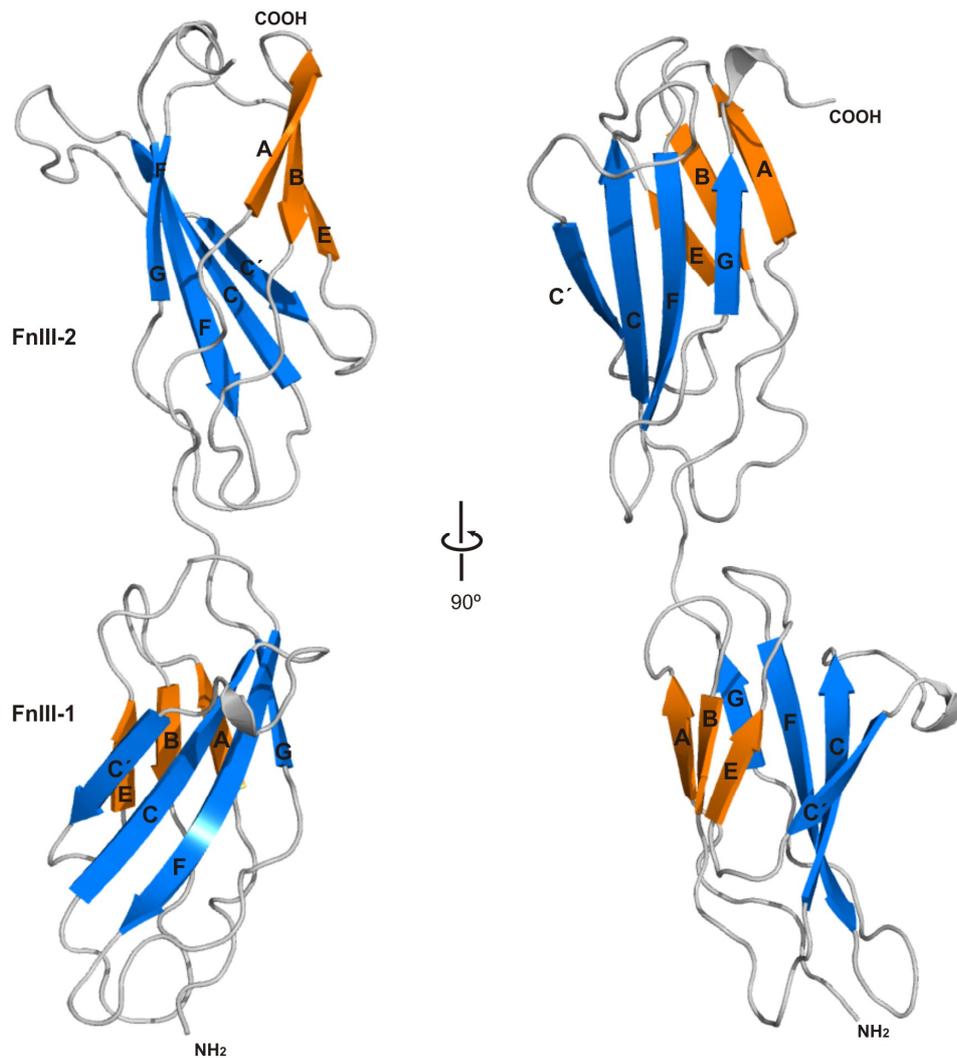


Figura 16.- Estructura de la región FnIII-1,2 de la subunidad $\beta 4$ de la integrina $\alpha 6\beta 4$. Representación en cintas de la estructura cristalográfica de los dominios FnIII-1 (inferior) y FnIII-2 (superior) (código PDB: 1QG3). Las hebras β se muestran en diferentes colores en función de la lámina β a la que pertenezcan.

2.3.1.3.- Organización de otras regiones del dominio citoplasmático de $\beta 4$

Experimentos de doble híbrido en levaduras y *far-western blot* usando fragmentos de $\beta 4$ sugieren que el CS y la cola C-terminal interactúan entre sí (Koster et al., 2004b; Reznicek et al., 1998). Recientemente, se ha mostrado usando transferencia de energía de resonancia de Förster (FRET) que la cola C-terminal se encuentra próxima al CS de la misma molécula en queratinocitos vivos. Esto sugiere que $\beta 4$ adopta una

estructura replegada en la cual la región rica en Pro situada entre FnIII-3 y FnIII-4 permite una conformación curvada (Frijns et al., 2012) (Figura I7).

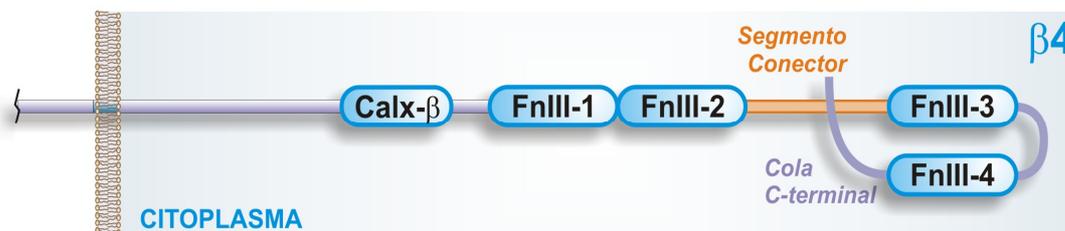


Figura I7.- Representación plegada de la región citoplasmática de la subunidad $\beta 4$ de la integrina $\alpha 6\beta 4$. La proximidad y posible interacción entre el CS y la cola C-terminal sugiere que la región FnIII-3,4 adopta una estructura en U que estaría favorecida por la flexibilidad de la secuencia que une los dominios FnIII-3 y FnIII-4.

2.4.- Tetraspanina CD151

Las tetraspaninas son una familia de proteínas integrales de membrana que se localizan en la superficie celular y regulan la morfología, movilidad, invasión, fusión y señalización celulares (Hemler, 2005). En humanos existen 32 tetraspaninas que se expresan en una amplia variedad de tipos celulares. La mayor parte de las tetraspaninas no realizan funciones típicas de receptores de superficie celular en el reconocimiento de ligandos; si bien algunas están implicadas en la unión a virus (Martin et al., 2005). En cambio la mayoría de las tetraspaninas actúan como organizadores de complejos multimoleculares de membrana asociándose entre si y con otras proteínas que incluyen integrinas, miembros de la superfamilia de las inmunoglobulinas, proteoglicanos, receptores de factores de crecimiento, reguladores del complemento, proteasas y cadherinas (Levy and Shoham, 2005). Estas agrupaciones de proteínas en la membrana plasmática reciben el nombre de red de tetraspaninas o microdominios enriquecidos en tetraspaninas (Berditchevski et al., 1996; Hemler, 2003).

CD151 es una tetraspanina con un amplio perfil de expresión siendo especialmente abundante en epitelio, endotelio, músculo cardiaco, músculo liso, células de Schwann,

megacariocitos y plaquetas (Sincock et al., 1997). En células epiteliales CD151 se localiza en la superficie de contacto con las membranas basales y forma parte de los HDs en queratinocitos (Sterk et al., 2000). En células endoteliales CD151 se localiza en zonas de interacción célula-célula (Sincock et al., 1999; Yanez-Mo et al., 1998).

CD151 se asocia con integrinas, siendo la principal tetraspanina que interacciona con estos receptores. CD151 regula funciones dependientes de integrinas, pero no modifica la afinidad de las integrinas por sus ligandos extracelulares. Específicamente CD151 forma complejos estables con las integrinas $\alpha3\beta1$, $\alpha6\beta1$, $\alpha6\beta4$ y $\alpha7\beta1$, que son receptores de lamininas (Sincock et al., 1999; Sterk et al., 2002; Yanez-Mo et al., 1998; Yauch et al., 2000). La asociación de CD151 con integrinas es altamente estable y específica, siendo resistente a condiciones relativamente agresivas como el detergente Triton X-100. Además esta asociación es estequiométrica, estando prácticamente la totalidad de la integrina $\alpha3\beta1$ asociada a CD151 *in vivo*. La asociación de CD151 con $\alpha3\beta1$ ocurre de forma temprana durante la biosíntesis, lo que sugiere que CD151 puede ser un ligando necesario para la correcta expresión de esta integrina en la superficie celular.

La relevancia de CD151 *in vivo* se ha puesto de manifiesto en estudios genéticos. Pacientes con una mutación que genera una forma truncada de CD151 desarrollan nefritis hereditaria, EBD pretibial, sordera sensorial y β -talasemia menor (Karamatic Crew et al., 2004). Estos defectos sugieren que CD151 regula la integridad de la membrana basal en riñón, oído interno, piel y médula ósea. Mutaciones en los genes que codifican las integrina $\alpha3$, $\alpha6$ y $\beta4$ producen fenotipos similares y sugiere que las asociaciones de CD151 con $\alpha3\beta1$ y $\alpha6\beta4$ son funcionalmente importantes.

La inactivación del gen *CD151* en ratón produce fallo renal severo; un efecto similar a la inactivación selectiva de la integrina $\alpha3$ de los podocitos renales (Sachs et al., 2006). Así mismo ratones deficientes en CD151 presentan anomalías en la función plaquetaria con retraso en la retracción de trombos y defectos en la agregación y extensión de plaquetas. Esto sugiere que la función de la integrina $\alpha_{IIB}\beta3$ se ve alterada ante la falta de CD151 y apoya la relevancia funcional de la interacción entre CD151 y $\alpha_{IIB}\beta3$. Queratinocitos de ratones CD151 *-/-* muestran defectos de migración en cultivos de

explantes de piel, lo cual sugiere alteraciones en la función de las integrinas $\alpha 3\beta 1$ y $\alpha 6\beta 4$ con las que se asocia CD151. Por último linfocitos T de estos animales son hiperproliferativos en respuesta a estímulos mitogénicos *in vivo* (Lau et al., 2004; Wright et al., 2004).

2.4.1.- Estructura de CD151 y otras tetraspaninas

CD151 (253 amino ácidos, 28 kDa) contiene elementos estructurales comunes a otras tetraspaninas, que incluyen cuatro segmentos transmembrana (TM1 a TM4) y motivos de secuencia específicos de esta superfamilia. Los dominios TM se orientan en la membrana citoplasmática de tal manera que los extremos amino y carboxilo y el bucle entre TM2- TM3 están situados en el citoplasma; mientras que las regiones entre TM1-TM2 y TM3-TM4 constituyen dos dominios extracelulares (EC1 y EC2) (Figura I8).

Los cuatro segmentos o hélices transmembrana son las regiones más conservadas dentro de la familia de las tetraspaninas. TM1, TM3 y TM4 contienen residuos polares conservados (Asn, Gln y Glu) que pueden participar en la formación de puentes de hidrógeno y por lo tanto estabilizar contactos intramoleculares entre las hélices transmembrana o interacciones intermoleculares, por ejemplo en la homo- o hetero-oligomerización de tetraspaninas. Se desconoce la estructura y orientación de las hélices transmembrana, pero se han propuesto modelos de empaquetamiento intra-e-intermolecular (Bienstock and Barrett, 2001; Kovalenko et al., 2005).

Las colas citoplasmáticas en los extremos amino y carboxilo son generalmente cortas (menos de 15 residuos) y muestran alto grado de conservación entre especies pero bajo grado de conservación entre miembros de la familia en una misma especie, lo que sugiere que están implicados en funciones específicas de cada tetraspanina (Stipp et al., 2003). La cola C-terminal es especialmente divergente, está implicada en la localización subcelular (Rous et al., 2002). Las regiones citoplasmáticas son zonas de modificación post-transduccional mediante la unión covalente de palmitatos a Cys conservadas próximas a los dominios transmembrana. La palmitoilización contribuye a asociación lateral con otras tetraspaninas y a la formación de microdominios enriquecidos en tetraspaninas (Charrin et al., 2002; Yang et al., 2002).

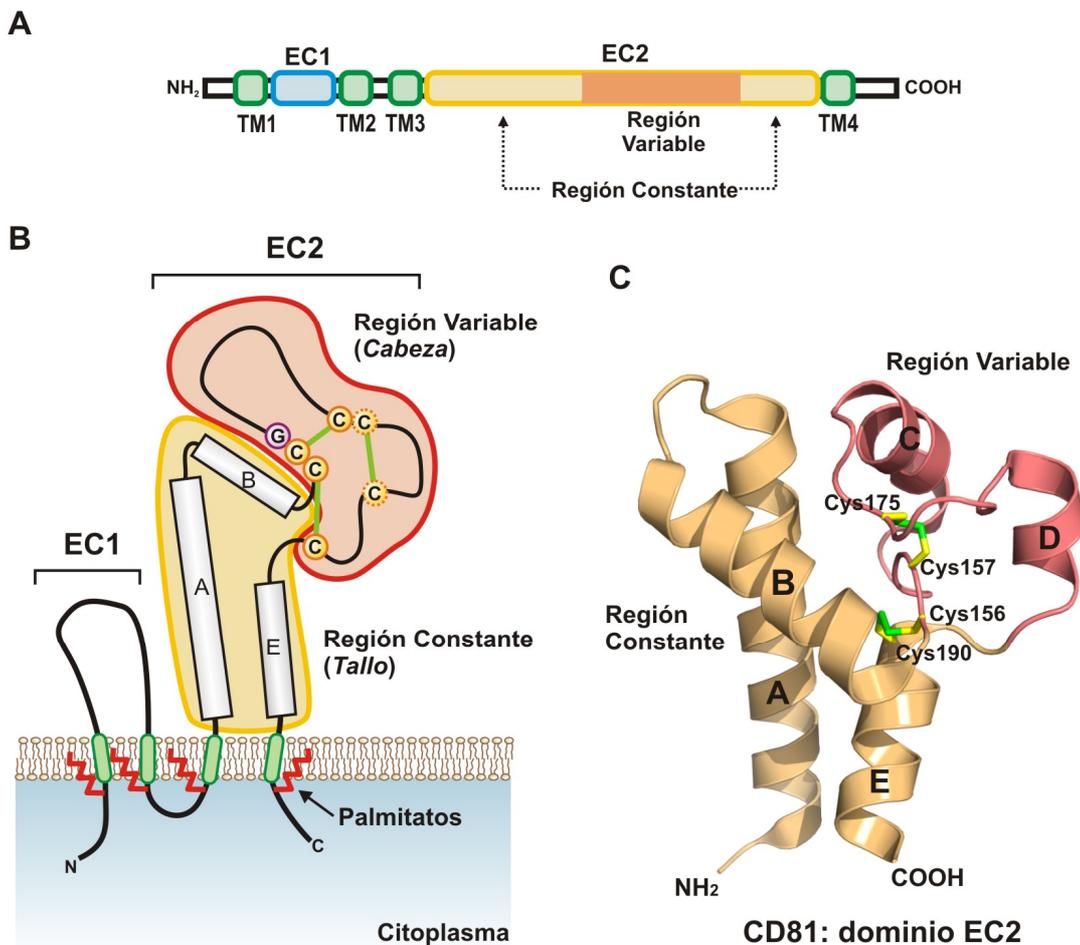


Figura 18. Estructura de las tetraspaninas. (A) Organización modular. Las tetraspaninas contienen cuatro segmentos transmembrana (TM1-4). Entre TM1 y TM2 se extiende un dominio extracelular de pequeño tamaño (EC1), mientras que entre TM3 y TM4 se localiza el dominio extracelular EC2 de entre 69 y 132 residuos. (B) Esquema de la estructura genérica de las tetraspaninas y su orientación en la membrana plasmática. (C) Representación de la estructura cristalográfica del dominio EC2 de CD81 (código PDB 1G8Q).

El primer dominio extracelular (EC1) es de pequeño tamaño con 13 a 31 residuos (21 aminoácidos en CD151). El segundo dominio extracelular (EC2) es de mayor tamaño, entre 69 y 132 residuos (105 aminoácidos en CD151), y en él se distingue una región conservada y otra variable que contiene la mayoría de los sitios de interacción con otras moléculas. El EC2 contiene residuos clave altamente conservados entre las tetraspaninas, éstos incluyen la secuencia CCG, otra Cys próxima al extremo C-terminal de EC2 y el motivo PXXC (donde X es cualquier amino ácido). El conocimiento estructural a alta resolución de las tetraspaninas se limita a la estructura cristalográfica del dominio EC2 de la tetraspanina CD81 (Kitadokoro et al., 2001a; Kitadokoro et al., 2002) (Figura

I8C). La estructura está formada por cinco hélices- α de las cuales las hélices A, B y E pertenecen a la región conservada mientras que las hélices C y D corresponden a la región variable. El dominio EC2 de CD81 tiene forma de champiñón en el cual las hélices A y E forman el tallo y las hélices B, C y D la cabeza. La región de cabeza o variable está estabilizada por dos puentes disulfuro formados por Cys conservadas en toda la familia.

El dominio EC2 de CD151 tiene un 12% de identidad de secuencia con CD81, lo cual sugiere que a pesar de compartir un plegamiento similar pueden existir diferencias estructurales entre ambas tetraspaninas. Dentro de la superfamilia de tetraspaninas CD81 pertenece al grupo 1, que presenta cuatro Cys conservadas en el EC2, mientras que CD151 pertenece al grupo 2b, con seis Cys conservadas (Seigneuret et al., 2001). Las dos Cys adicionales presentes en CD151 potencialmente forman un tercer puente disulfuro en la región variable. Además el dominio EC2 de CD151 es 28 residuos más largo que el de CD81, debido a la presencia de inserciones en el lazo que une las hélices C y D de la región variable.

La región variable del dominio EC2 es crucial para las interacciones de tetraspaninas con otras proteínas. Por ejemplo mutaciones puntuales en las hélices C y D de CD81 modifican la afinidad por proteínas del virus de la hepatitis C (Higginbottom et al., 2000). En CD151 sustitución de la secuencia $^{194}\text{QRD}^{196}$ por INF inhibe la interacción con las integrinas $\alpha 3\beta 1$ y $\alpha 6\beta 1$ (Kazarov et al., 2002). Esta secuencia se localiza en la región variable del dominio extracelular EC2 de CD151, lo que sugiere que esta región participa directamente en la unión lateral a la subunidad α de integrinas. Puesto que CD151 es la tetraspanina que más comúnmente se asocia a integrinas y esta función depende de su región variable dentro del EC2, es previsible que los determinantes estructurales de la interacción con las integrinas no estén conservados en otras tetraspaninas.

Por último, se desconoce la estructura y organización del dominio EC1. Debido a la interacción entre las hélices transmembrana es razonable suponer que el EC1 se localiza espacialmente muy próximo al EC2 sobre la superficie de la membrana y por lo tanto es posible que exista una interacción directa entre EC1 y EC2 que establezca la estructura del EC1.

3.- Organización de los HDs.

3.1.- Interacciones de la región citoplasmática de $\beta 4$ con otros componentes de los HDs.

La subunidad $\beta 4$ media la mayoría de las interacciones de la integrina $\alpha 6\beta 4$. En ratones, la delección del dominio citoplasmático de $\beta 4$ causan defectos severos en la piel, similares a los observados en la enfermedad humana EBJ-PA, e incapacita la formación de los HDs (Murgia et al., 1998).

Los dominios FnIII de $\beta 4$ participan en interacciones proteína-proteína en los HDs (Figura I9). La región FnIII-1,2 y la parte amino terminal del CS se unen al dominio de unión a actina (ABD) de plectina (de Pereda et al., 2009a; Geerts et al., 1999). La región C-terminal del CS, el FnIII-4 y la cola C-terminal se unen al dominio plakina de plectina (Frijns et al., 2012; Koster et al., 2004a; Koster et al., 2004b; Reznicek et al., 1998). La región FnIII-3,4 junto con la parte final del CS interactúan con la región N-terminal de BPAG1e (Koster et al., 2003). El dominio FnIII-3 también se une a BP180 (Koster et al., 2003).

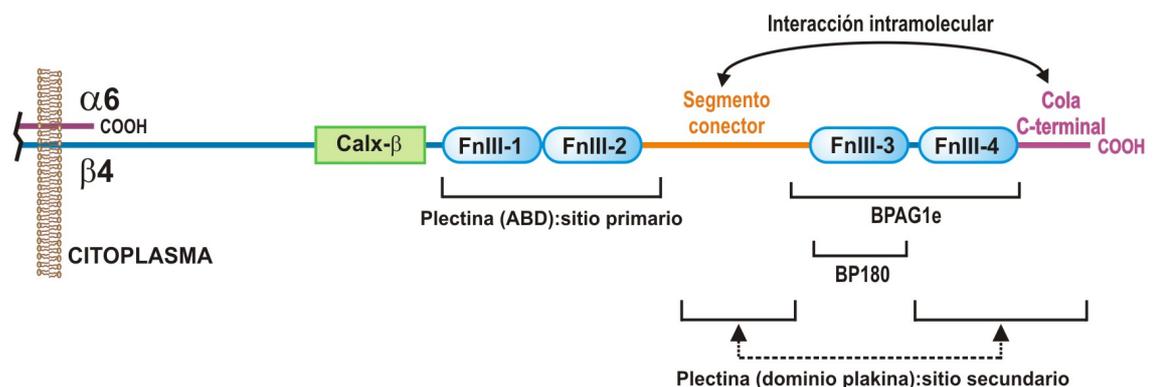


Figura I9.- Localización en la región citoplasmática de la integrina $\beta 4$ de sitios de interacción con otras proteínas de los HDs.

El dominio FnIII-3 también participa en procesos de señalización. El residuo Y1494 es necesario para la activación de la fosfatidilinositol 3-quinasa y para la estimulación de la invasión después de la ligación de $\alpha 6\beta 4$ (Shaw, 2001). La fosforilación de Y1494, junto

con Y1257 en el FnIII-2 y el Y1440 en el CS, da lugar a la unión de la fosfatasa Shp2 a $\beta 4$ y a la estimulación de la ruta de Erk (Bertotti et al., 2006). $\alpha 6\beta 4$ también se asocia a Erk por medio de la fosforilación de Y1526 en el FnIII-3, que es el sitio primario en $\beta 4$ de interacción con el dominio de unión a fosfotirosina (PTB) de Shc (Dans et al., 2001a).

3.2.- Bases estructurales de la interacción entre $\alpha 6\beta 4$ y plectina

De las interacciones que establece la región citoplasmática de $\beta 4$ con otras proteínas, la unión a plectina es la que se conoce con más detalle. $\beta 4$ interacciona con plectina en múltiples sitios. La interacción primaria se establece entre una región de $\beta 4$ que incluye los FnIII-1,2 y la región N-terminal del CS y el ABD de plectina (Geerts et al., 1999; Rezniczek et al., 1998). La interacción en solución en este sitio primario muestra una afinidad moderada, siendo la K_d $\sim 30 \mu M$ (de Pereda et al., 2009a; de Pereda et al., 2009b; García-Alvarez et al., 2003).

La estructura cristalográfica del complejo formado por el FnIII-1,2 y la región N-terminal del CS de $\beta 4$ unido al ABD de plectina, reveló con detalle cómo se produce esta interacción e identificó los elementos críticos de ésta (Figura I10) (de Pereda et al., 2009a). La mayor zona de contacto se establece entre el FnIII-2 de $\beta 4$ y el sub-dominio CH1 del ABD. Un pequeño fragmento de plectina que precede al CH1 también contacta con el FnIII-2; por el contrario, el sub-dominio CH2 no interacciona directamente con $\beta 4$. Parte del CS de $\beta 4$ también contribuye a la interfase de unión, situándose entre el CH1 y el FnIII-2. Por último, el FnIII-1 de $\beta 4$ proporciona contactos adicionales con el CH1.

La superficie de interacción en el FnIII-1 de $\beta 4$ se centra en los residuos básicos R1225, K1279 y R1281 del FnIII-2, que están enfrentados hacia una superficie ácida en el CH1 de plectina. Dos mutaciones puntuales en dos de estos residuos, R1225H y R1281W, causan formas no letales de epidermolisis bullosa (Nakano et al., 2001; Pulkkinen et al., 1998). Estas mutaciones destruyen puentes salinos intermoleculares esenciales y causan la inhibición de la interacción (de Pereda et al., 2009a; Koster et al., 2001).

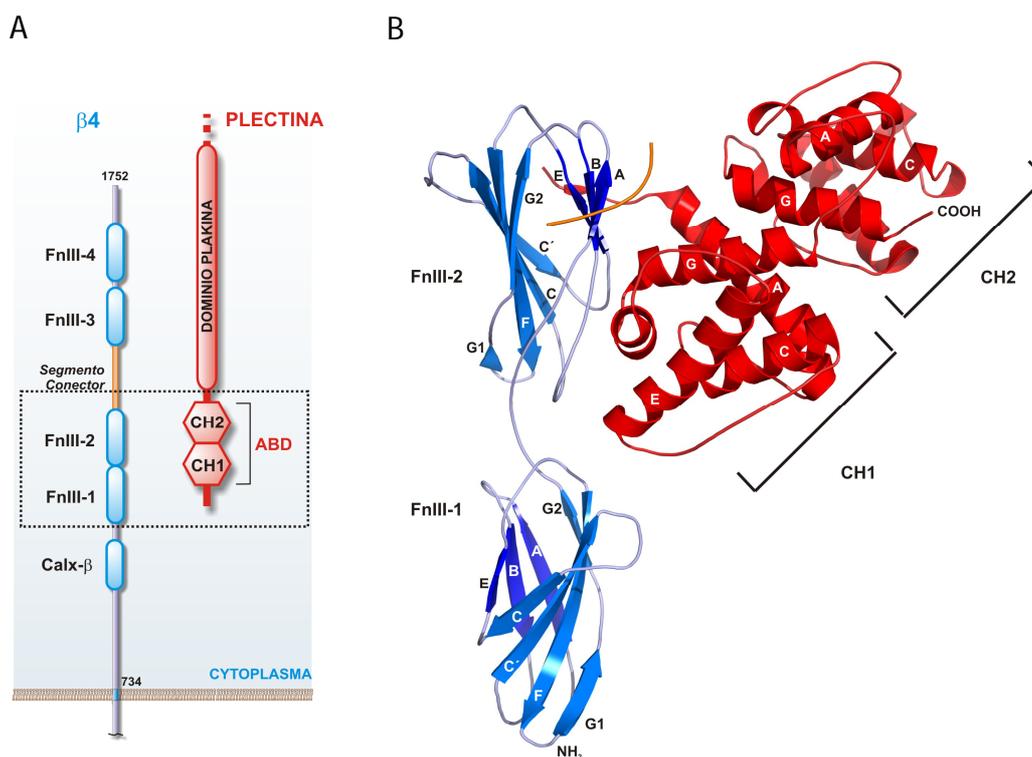


Figura I10.- Estructura del complejo β 4-plectina. (A) Esquema de la estructura de la región citoplasmática de β 4 y de la región N-terminal de plectina. Las regiones de β 4 y plectina implicadas en la principal interacción se indican mediante un cuadro de trazo discontinuo. (B) Representación de cintas de la estructura cristalográfica de la región de la integrina β 4, que incluye los dominios FnIII-1,2 (azul) y parte del CS (naranja), unida al ABD de plectina (rojo) (código PDB: 3F7P).

En las estructuras cristalográficas del FnIII-1,2-CS de β 4 en ausencia de plectina, la primera parte del CS se organiza sobre el FnIII-2 formando dos hebras- β adicionales, denominadas H e I (Figura I11A) (de Pereda et al., 2009a). En el complejo β 4-plectina el CS se localiza en otra posición, al extremo opuesto de la lámina ABE (Figura I11B). Por otro lado, no hay diferencias significativas en la estructura del FnIII-1 y el FnIII-2 entre los estados libre y unido a plectina. Este cambio conformacional del CS podría estar relacionado con un ligero aumento de la estabilidad de la integrina β 4 al unirse a plectina (García-Alvarez et al., 2003). La superficie de β 4 alrededor de la hebra E, en la que se sitúa la hebra H en el estado libre, esta ocupada en el complejo por el brazo N-terminal de plectina, que precede al CH1. El brazo N-terminal de plectina no es necesario para la unión a β 4 lo que sugiere que la reorganización del CS e β 4 es inducida por la unión del CH1. La reorganización del CS es necesaria para la unión a plectina, por

lo que la interacción β 4-pectina puede estar inhibida alostéricamente por factores que estabilicen la conformación no unida de β 4 (de Pereda et al., 2009a).

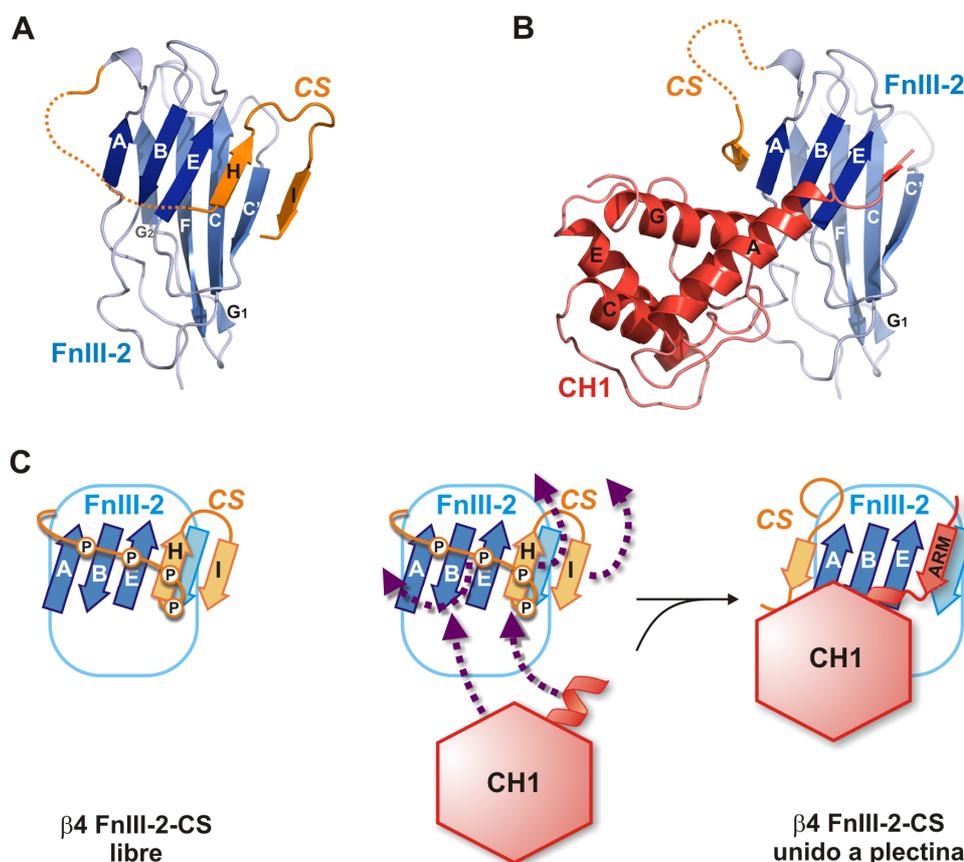


Figura I11.- Cambio conformacional del CS de β 4 ligado a la unión a plectina. (A) Representación de la estructura del FnIII-2 y el CS en ausencia de plectina (código PDB: 3F7P, cadena E). La hebra H se sitúa adyacente a la hebra E del FnIII-2. (B) Estructura del complejo β 4-pectina. El CS se sitúa junto a la hebra A del FnIII-2, mientras que la región junto a la hebra E está ocupada por la región de plectina que precede al CH1. Para simplificar la figura se han omitido el CH2 de plectina y el FnIII-1 de β 4. (C) Representación esquemática de las estructuras del FnIII-2 y el CS de β 4 en el estado libre (izquierda) y unido a plectina (derecha).

3.3.- Ensamblaje y regulación de HDs.

La interacción entre α 6 β 4 y plectina es esencial para el ensamblaje y la estabilidad de los HDs (Murgia et al., 1998; Nievers et al., 1998) y es un paso fundamental en la formación de estos complejos, ya que facilita la incorporación de otras proteínas, como BP180 y BPAG1e (Schaapveld et al., 1998).

Se ha propuesto un modelo secuencial o jerárquico del ensamblaje de HDs en el que α 6 β 4 juega un papel central (Wilhelmsen et al., 2006) (Figura I12). En una primera etapa

plectina se incorpora a los HDs mediante la interacción con $\beta 4$; este paso no requiere la unión de $\alpha 6\beta 4$ a lamininas de la matriz extracelular. La unión de plectina a $\beta 4$ se produciría inicialmente a través del ABD de plectina y el FnIII-1,2-CS de $\beta 4$, y se vería reforzada por contactos entre el dominio plakina de plectina y el CS y la cola C-terminal de $\beta 4$. La unión del dominio plakina a $\beta 4$ se considera secundaria ya que no es suficiente para que plectina se incorpore a los HDs. En el caso de HDs de tipo II, una vez que plectina se ha unido a $\beta 4$, se incorporaría BP180, que interacciona con el FnIII-3 de $\beta 4$ y con plectina, y BPAG1e, que se asocia con el FnIII-3,4 de $\beta 4$ y con BP180. Se ha propuesto que la unión de BP180 y BPAG1e a $\beta 4$ podría estar favorecida por un cambio conformacional en la cola citoplasmática de $\beta 4$ inducido por la unión de plectina, que expondría sitios de unión para estas proteínas. Por último, tanto plectina como BPAG1e tienen sitios de unión a filamentos intermedios en sus regiones C-terminal y conectan los HDs al citoesqueleto.

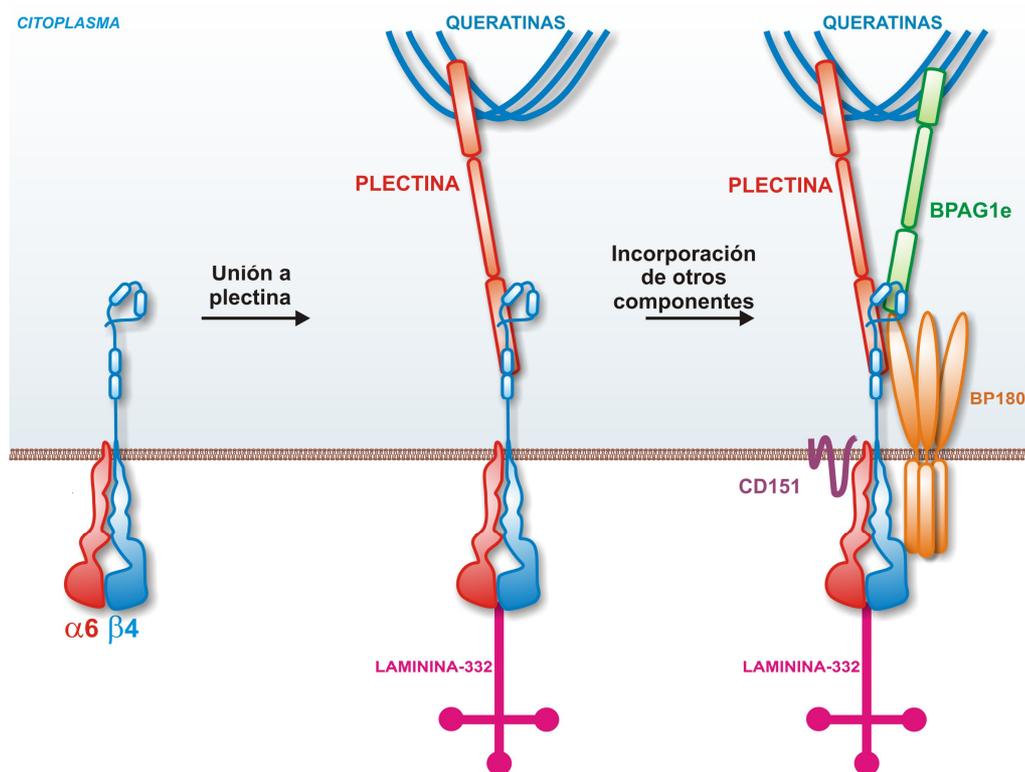


Figura I12.- Modelo secuencial o jerárquico del ensamblaje de los HDs. El dominio citoplasmático de $\beta 4$ adopta una conformación plegada en la que la cola C-terminal interacciona con el CS (izquierda). $\alpha 6\beta 4$ recluta plectina en los HDs mediante la interacción de plectina con la región citoplasmática de $\beta 4$ (centro). En una tercera etapa de maduración BPAG1e y BP180 se incorporan a los HDs interaccionando con el complejo $\alpha 6\beta 4$ -plectina (derecha).

Los HDs son estructuras altamente dinámicas y se pueden desensamblar rápidamente bajo condiciones en las que se requiere un desprendimiento parcial de la membrana basal; por ejemplo, durante la división celular, la diferenciación o la migración de queratinocitos o durante la invasión de células de carcinomas (Geuijen and Sonnenberg, 2002).

La regulación de la interacción entre $\beta 4$ y plectina juega un papel clave en el desensamblaje de los HDs. Varios factores de crecimiento como el factor de crecimiento epidermal (EGF), el factor de crecimiento de hepatocitos y la proteína estimuladora de macrófagos, se han implicado en el desensamblaje de HDs (Margadant et al., 2008). Estos factores activan la ruta de señalización de Ras/MAPK que resulta en la estimulación de la fosforilación de la región citoplasmática de $\beta 4$ en residuos de Ser (S1356, S1360 y S1364) que se localizan en el CS, adyacentes al sitio de unión al ABD de plectina. La fosforilación de estos residuos de Ser de $\beta 4$ da lugar al desensamblaje parcial de HDs ligado a una reducción de la interacción con plectina (Rabinovitz et al., 2004; Wilhelmsen et al., 2007). Las quinasas ERK1/2 y p90RSK1/2 también fosforilan $\beta 4$ en S1356 y S1364 respectivamente, desestabilizando la interacción del FnIII-1,2 de $\beta 4$ con el ABD de plectina (Frijns et al., 2010). La fosforilación de T1424, en el CS, es necesaria para el desensamblaje de HDs en respuesta a estimulación mediante EGF (Germain et al., 2009). Por último, fosforilación por PKD1 u otra quinasa de la familia CAMK de T1736, situada en la cola C-terminal de $\beta 4$, reduce la interacción con el dominio plakina de plectina (Frijns et al., 2012).

OBJETIVOS

El objetivo general de esta tesis es avanzar en el conocimiento de la organización de los hemidesmosomas, mediante la caracterización estructural de dos proteínas de hemidesmosomas, integrina $\alpha6\beta4$ y la tetraspanina CD151. Para ello se han abordado los siguientes objetivos específicos:

1. Caracterización de la estructura 3D de fragmentos de la región citoplasmática de la subunidad $\beta4$ de la integrina $\alpha6\beta4$.

1.1. Elucidación de la estructura del dominio Calx- β de $\beta4$ mediante cristalografía de rayos-x.

1.2. Elucidación y análisis de la estructura de la región FnIII-3,4 de $\beta4$ mediante métodos híbridos.

2. Expresión heteróloga y purificación del dominio extracelular de CD151 implicado en la interacción con integrinas, en cantidades adecuadas para su estudio estructural.

MATERIALES Y MÉTODOS

1.- Clonación de fragmentos de $\beta 4$ en vectores de expresión en bacterias

Las secuencias de ADNc que codifican diversos fragmentos de la integrina $\beta 4$ humana (código de Uniprot: P16144-2) se amplificaron mediante PCR empelando la polimerasa Pfu (Stratagene) y usando como molde en clon *IMAGE* 3640058 que contiene el ADNc que codifica la región 759-1752 de $\beta 4$ en el vector pOTB7. Los límites de los fragmentos de $\beta 4$, así como los oligonucleótidos cebadores específicos para cada uno de ellos, se detallan en la tabla M1. Los cebadores directos contienen un sitio reconocido por la enzima de restricción EcoRI y un sitio NdeI, mientras que los cebadores reversos contienen un codón de terminación, un sitio BamHI y un sitio EcoRI.

Tabla M1.- Cebadores utilizados para el clonaje de fragmentos de la integrina $\beta 4$.

FRAGMENTO		CEBADOR	
NOMBRE	REGIÓN	NOMBRE	SECUENCIA (sentido 5' a 3')
Calx- β	989-1107	B4-U989	TGA <u>GAA TTC CAT ATG</u> AGA GAC GTG GTG TCC TTT GAG C EcoRI NdeI R D V V S F E 989
		B4-L1107	GCC <u>GAA TTC GGA TCC CTA</u> TTC ATC TGG GTC CCT GAT G EcoRI BamHI Stop E D P D R I 1107
FnIII-3	1457-1548	B4-U4369	TGA <u>GAA TTC CAT ATG</u> GTG CCC GAC ACG CC EcoRI NdeI V P D T P 1457
		B4-L4644	GCC <u>GAA TTC GGA TCC CTA</u> AAT GGT GAT GAC ACC CTC EcoRI BamHI Stop I T I V G E 1549
FnIII-4	1572-1666	B4-U4714	TGA <u>GAA TTC CAT ATG</u> AGT GCC CCA GGC CCG CTG G EcoRI NdeI S A P G P L 1572
		B4-L4998	GCC <u>GAA TTC GGA TCC CTA</u> GGA CTC TAT GGT GAT GAT GC EcoRI BamHI Stop S E I T I I 1666
FnIII-3,4	1457-1666	B4-U4369	TGA <u>GAA TTC CAT ATG</u> GTG CCC GAC ACG CC EcoRI NdeI V P D T P 1457
		B4-L4998	GCC <u>GAA TTC GGA TCC CTA</u> GGA CTC TAT GGT GAT GAT GC EcoRI BamHI Stop S E I T I I 1666
CS-FnIII-3,4	1436-1666	B4-U1436	TGA <u>GAA TTC CAT ATG</u> CTG CCG AGG GAC TAC TCC ACC EcoRI NdeI L P R D Y S T 1436
		B4-L4998	GCC <u>GAA TTC GGA TCC CTA</u> GGA CTC TAT GGT GAT GAT GC EcoRI BamHI Stop S E I T I I 1666

deseadas se comprobó mediante secuenciación completa del ADN de la región codificante.

Tabla M2-A.- Mutaciones puntuales dentro del dominio Calx-β y cebadores empleados.

MUTACIÓN	CEBADOR	
	NOMBRE	SECUENCIA (sentido 5' a 3')
R1003A	B4-R1003A-F	GAG TTC TCG GTC AGC GCC GGG GAC CAG GTG GCC
	B4-R1003A-R	GGC CAC CTG GTC CCC GGC GCT GAC CGA GAA CTC
R1003E	B4-R1003E-F	GAG TTC TCG GTC AGC GAG GGG GAC CAG GTG GCC
	B4-R1003E-R	GGC CAC CTG GTC CCC CTC GCT GAC CGA GAA CTC

Tabla M2-B.- Mutaciones puntuales para la sustitución o introducción de Cys en el dominio FnIII-3 y cebadores empleados.

MUTACIÓN	CEBADOR	
	NOMBRE	SECUENCIA (sentido 5' a 3')
R1463C	B4- R1463C-F	CCC GAC ACG CCC ACC TGC CTG GTG TTC TCT GC
	B4- R1463C-R	GC AGA GAA CAC CAG GCA GGT GGG CGT GTC GGG
A1468C	B4-A1468C-F	CC CGC CTG GTG TTC TCT TGC CTG GGG CCC ACA TCT CTC
	B4-A1468C-R	GAG AGA TGT GGG CCC CAG GCA AGA GAA CAC CAG GCG GG
T1472C	B4-T1472C-F	G TTC TCT GCC CTG GGG CCC TGT TCT CTC AGA GTG AGC TGG C
	B4-T1472C-R	G CCA GCT CAC TCT GAG AGA ACA GGG CCC CAG GGC AGA GAAC
R1475C	B4-R1475C-F	GGG CCC ACA TCT CTC TGT GTG AGC TGG CAG GAG
	B4-R1475C-R	CTC CTG CCA GCT CAC ACA GAG AGA TGT GGG CCC
C1483A	B4-C1483A-F	GG CAG GAG CCG CGG GCC GAG CGG CCG CTG C
	B4-C1483A-R	G CAG CGG CCG CTC GGC CCG CGG CTC CTG CC
C1483S	B4-C1483S-F	CAG GAG CCG CGG TCC GAG CGG CCG CTG C
	B4-C1483S-R	G CAG CGG CCG CTC GGA CCG CGG CTC CTG
R1485C	B4- R1485C -F	G GAG CCG CGG TCC GAG TGC CCG CTG CAG GGC
	B4- R1485C R	GCC CTG CAG CGG GCA CTC GGA CCG CGG CTC C
L1497C	B4-L1497C-F	AGT GTG GAG TAC CAG CTG TGT AAC GGC GGT GAG CTG CAT C
	B4-L1497C-R	G ATG CAG CTC ACC GCC GTT ACA CAG CTG GTA CTC CAC ACT
R1504C	B4- R1504C-F	GGC GGT GAG CTG CAT TGC CTC AAC ATC CCC AAC
	B4- R1504C-R	GTT GGG GAT GTT GAG GCA ATG CAG CTC ACC GCC
A1511C	B4-A1511C-F	G CTC AAC ATC CCC AAC CCT TGC CAG ACC TCG GTG GTG G
	B4-A1511C-R	C CAC CAC CGA GGT CTG GCA AGG GTT GGG GAT GTT GAG C
N1523C	B4-N1523C-F	GAA GAC CTC CTG CCC TGC CAC TCC TAC GTG TTC
	B4-N1523C-R	GAA CAC GTA GGA GTG GCA GGG CAG GAG GTC TTC

Tabla M2-C.- Mutaciones puntuales para la sustitución o introducción de Cys en el segmento conector y en el dominio FnIII-4 y cebadores empleados.

MUTACIÓN	CEBADOR	
	NOMBRE	SECUENCIA (sentido 5' a 3')
C1559A	B4-C1559A-F	G CAG AGC CCA CTG GCT CCC CTG CCA GGC
	B4-C1559A-R	GCC TGG CAG GGG AGC CAG TGG GCT CTG C
P1576C	B4-P1576C-F	CCC AGT GCC CCA GGC TGT CTG GTG TTC ACT GCC
	B4-P1576C-R	GGC AGT GAA CAC CAG ACA GCC TGG GGC ACT GGG
N1598C	B4-N1598C-F	CGG CCA CGG AGG CCC TGT GGG GAT ATC GTC GGC
	B4-N1598C-R	GCC GAC GAT ATC CCC ACA GGG CCT CCG TGG CCG
C1608A	B4-C1608A-F	GC TAC CTG GTG ACC GCT GAG ATG GCC CAA G
	B4-C1608A-R	C TTG GGC CAT CTC AGC GGT CAC CAG GTA GC
C1608S	B4-C1608S-F	GGC TAC CTG GTG ACC TCT GAG ATG GCC CAA GG
	B4-C1608S-R	CC TTG GGC CAT CTC AGA GGT CAC CAG GTA GCC
A1619C	B4-A1619C-F	GGA GGG CCA GCC ACC TGT TTC CGG GTG GAT GGA
	B4-A1619C-R	TCC ATC CAC CCG GAA ACA GGT GGC TGG CCC TCC
N1639C	B4-N1639C-F	CCG GGC CTC AGC GAG TGC GTG CCC TAC AAG TTC
	B4-N1639C-R	GAA CTT GTA GGG CAC GCA CTC GCT GAG GCC CCG

3.- Expresión de fragmentos de $\beta 4$ en bacterias y purificación

Los fragmentos de $\beta 4$ se expresaron en la cepa BL21(DE3) de *E. coli*. Las bacterias con el plásmido correspondiente se crecieron en el medio Terrific Broth (Sambrook et al., 1989) suplementado con 100 mg l⁻¹ de ampicilina. La expresión de la proteína recombinante se indujo por incubación con 0.2 mM isopropil- β -D-1-tiogalactopiranosido (IPTG) durante 3 h a 37 °C o durante 12 a 14 h a 15 °C. Las bacterias se recogieron mediante centrifugación (rotor JLA 10500, 11800g, 4 °C, 20 min), se resuspendieron en 20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 5 mM imadazol, 1% Triton X-100 (~60 ml por cada litro de cultivo) y se congelaron a -75 °C hasta su utilización. La lisis celular se completó mediante sonicación, tras la cual las fracciones soluble e insoluble se separaron mediante centrifugaron a 40000g (rotor JA 25.50) durante 30 min a 4 °C.

3.1.- Purificación del dominio Calx-β

El dominio Calx-β aparece en la fracción insoluble tras la lisis celular. El pellet se resuspendió en 20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 5 mM imidazol, 6 M urea, durante 12 a 14 h a 4 °C con agitación, tras lo cual la muestra se clarificó mediante centrifugación a 40000g durante 30 min a 4 °C. El dominio Calx-β se purificó a partir del sobrenadante en condiciones desnaturalizantes mediante cromatografía de afinidad empleando una columna HisTrap HP quelante de Ni²⁺, de 5 ml (GE Healthcare). Tras cargar la muestra, la columna se lavó con 30 ml de 20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 20 mM imidazol, 6 M urea. Seguidamente el dominio Calx-β se eluyó con 20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 400 mM imidazol, 6 M urea. El Calx-β se replegó por dilución 1:20 en 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, que resultó en una concentración residual de 0.3 M urea. La muestra se clarificó por centrifugación (40000g, 30 min, 4 °C) y la fracción de proteína soluble se concentró hasta ~2 mg ml⁻¹ mediante ultrafiltración de flujo tangencial empleando un casete Centramate (Pall Life Sciences) con un tamaño de corte de 3 kDa. La cola de His se eliminó por digestión con la proteasa TEV y por posterior cromatografía de afinidad en condiciones nativas empleando una columna HisTrap HP de 5 ml. Finalmente, la muestra se dializó frente a 10 mM Tris-HCl (pH 7.5), 50 mM NaCl y se concentró mediante ultrafiltración empleando celdas Amicon y membranas YM3 (Millipore) hasta la concentración deseada.

Los mutantes del dominio Calx-β, que incorporan sustituciones puntuales, se purificaron de igual forma que el Calx-β silvestre.

3.2.- Purificación de fragmentos de la región FnIII-3,4 de β4

El dominio FnIII-3 se encontró en la fracción soluble y se purificó por cromatografía de afinidad utilizando una columna quelante de Ni²⁺ HisTrap de 5 ml (GE Healthcare) y un equipo ÄKTA Prime (GE Healthcare). La proteína de interés se eluyó con un gradiente de 5 a 500 mM imidazol en 20 mM Tris- HCl (pH 7.9), 500 mM NaCl. Las fracciones que contenían proteína se dializaron a 4 °C frente a 20 mM Tris-HCl (pH 7.5), 150 mM NaCl. La cola de octa-His se escindió por digestión con la proteasa TEV a 30 °C durante 10 h. Los restos de proteína sin digerir, la cola de poli-His y la proteasa TEV (que incorpora una

cola de poli-His) se eliminaron mediante una segunda cromatografía de afinidad. Finalmente, la proteína se dializó extensamente frente al buffer deseado y se concentró por ultrafiltración en celdas Amicon usando membranas YM3 o YM10 (Millipore).

El dominio FnIII-4 (1572-1666), el fragmento FnIII-3,4 (1457-1666) y sus mutantes y el fragmento CS-FnIII-3,4 (1436-1666) se purificaron de forma general de la misma manera que el dominio FnIII-3.

4.- Cristalografía

4.1.- Cristalización de fragmentos de $\beta 4$

4.1.1.- Ensayos iniciales

Los ensayos de cristalización iniciales se realizaron en algunos casos por el método de difusión de vapor y en otros casos por el método microbatch. En ambos casos se utilizaron las colecciones comerciales de soluciones de cristalización Crystal Screen, SALT Rx (Hampton Research), Wizard I y II, Cryo I y II (Emerald Biosystems) y PACT (Hanahan, 1983) que incluyen un total de 432 condiciones de cristalización. Los ensayos se realizaron a temperatura ambiente y a 4 °C.

Los experimentos de difusión de vapor se realizaron en placas MRC de 96 pocillos (Swissci). Para ello, se mezcló 1 μ l de solución de proteína con el mismo volumen de solución de cristalización y se equilibró esta mezcla frente a 80 μ l de la solución de cristalización. Los experimentos de microbatch se realizaron en placas de 384 pocillos ShallowWell (Nunc). En cada pocillo se mezcló 1 μ l de proteína con un volumen igual de solución de cristalización y se cubrieron con 5 μ l de una mezcla 1:2 de aceite de parafina (Fluka, referencia 76235) y de octametiltrisiloxano al 98% (Aldrich, referencia 235709). Las condiciones que dieron lugar a cristales se optimizaron en placas de 24 pocillos por el método de difusión de vapor en gota colgante utilizando placas XRL (Molecular Dimensions) o en gota sentada empleando placas Cryschem (Hampton Research). Brevemente, de 3 a 4 μ l de solución de proteína se mezclaron típicamente con un volumen igual de solución precipitante y la mezcla se equilibró frente a 1 ml de dicha solución.

4.1.2.- Cristalización del dominio Calx-β

Los cristales del dominio Calx-β se crecieron a temperatura ambiente por el método de difusión de vapor mezclando 3.5 μl de proteína a 24 mg ml⁻¹ en 10 mM Tris-HCl (pH 7.5), 50 mM NaCl con 4.5 μl de 50 mM Tris-HCl (pH 8.3), 26% PEG 1500. Antes de la toma de datos, los cristales se transfirieron a una solución 50 mM Tris-HCl (pH 8.3), 26% PEG 1500, 15% glicerol y se enfriaron por inmersión directa en nitrógeno líquido. También se crecieron cristales en presencia de Ca²⁺ usando como solución de cristalización 50 mM Tris-HCl (pH 7.9), 2 mM CaCl₂, 26% PEG 1500. Antes de congelar estos cristales en nitrógeno líquido, se transfirieron a una solución 50 mM Tris-HCl (pH 7.9), 2 mM CaCl₂, 27% PEG 1500, 15% glicerol. Finalmente, cristales crecidos en ausencia de Ca²⁺ se incubaron durante 1 h en una solución 50 mM Tris-HCl (pH 8.3), 10 mM CaCl₂, 25% PEG 1500 y se transfirieron a una solución crioprotectora conteniendo 10 mM CaCl₂ antes de congelarlos.

4.1.3.- Cristalización del dominio FnIII-3

Los cristales del dominio FnIII-3 se crecieron a 4 °C por el método de difusión de vapor mezclando 3 μl a 20 mg ml⁻¹ en 20 mM Tris-HCl (pH 7.5), 4 mM DTT con 3 μl 100 mM Tris-HCl (pH 7.0), 0.5 mM (NH₄)₂SO₄, 36% PEG 600, 12,5% glicerol. La solución de cristalización actuó como solución crioprotectora, por lo que los cristales se congelaron por inmersión directa en nitrógeno líquido.

4.1.4.- Cristalización del dominio FnIII-4

Los cristales del FnIII-4 se crecieron por el método de difusión de vapor mezclando 4 μl del dominio FnIII-4 a 27 mg ml⁻¹ en 10 mM Tris-HCl (pH 7.5), 2 mM DTT con 4 μl de 0.1 M acetato sódico (NaAc) (pH 4.2), 1.8 M fosfato K/Na (0.8 NaH₂PO₄ / 1.2 K₂HPO₄), 2 mM DTT. Antes de la recolección de datos, los cristales se transfirieron a una solución 0.1 M NaAc, 1.8 M fosfato K/Na, 5% glicerol y se congelaron por inmersión directa en nitrógeno líquido. Por otra parte, un cristal crecido en las mismas condiciones se incubó en 1 mM etilmercuritiosalicilato sódico (EMTS), 0.1 M NaAc (pH 4.2), 1.8 M fosfato Na/K

durante 6 h a 4 °C. Luego se transfirió a la solución crioprotectora 0.1 M NaAc (pH 4.2), 1.8 M fosfato Na/K, 5% glicerol y se congeló por inmersión directa en nitrógeno líquido.

4.2.- Toma de datos de difracción

Las imágenes de difracción de todos los cristales se midieron usando un generador de rayos-X de ánodo rotatorio de cobre Microstar-H (Bruker AXS) que produce radiación con una longitud de onda de 1.5418Å. Este difractorómetro dispone de un sistema de detección mar345dtb (Mar Research GmbH) que incluye un sistema goniométrico de un eje (ϕ) y un detector de tipo *image plate*. El equipo dispone además de un sistema criogénico Cryostream Serie 700 (Oxford Cryosystems) que mantiene constante la temperatura de la muestra (100 K) durante la toma de datos.

4.3.- Procesado de los datos de difracción

El análisis de las imágenes (posición e intensidad de las reflexiones) de cada cristal permite determinar la geometría de la celda cristalina y la simetría del grupo espacial al que pertenece el cristal. Los datos de intensidad de difracción recogidos se indexaron e integraron con el programa XDS y se escalaron y redujeron con XSCALE (Kabsch, 2010). Los valores de intensidad de las reflexiones se convirtieron en módulos de amplitud de factores de estructura con el programa XDSCONV que emplea el método de French y Wilson (French and Wilson, 1978). Finalmente, los datos se transformaron a archivos con el formato MTZ empleando los programas F2MTZ y CAD de la suite CCP4 (Winn et al., 2011).

4.4.- Resolución estructural: faseado

La resolución de la estructura cristalográfica de una proteína implica conocer la función de distribución de densidad electrónica a partir de los factores de difracción. Los factores de estructura (F_{hkl}) que forman el patrón de difracción se pueden expresar como:

$$F_{hkl} = |F_{hkl}| e^{i\phi_{hkl}}$$

Donde h, k, l son los índices de Miller de las reflexiones, $|F_{hkl}|$ es la amplitud, que es proporcional a la raíz cuadrada de la intensidad de difracción observada (I_{hkl}) y ϕ_{hkl} es la fase.

La función de densidad electrónica (ρ) es la transformada de Fourier del patrón de difracción, la cual se puede expresar como:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| \cos 2\pi(hx + ky + lz - \phi(hkl))$$

Donde V es el volumen de la celdilla unidad.

Por lo tanto el cálculo de la densidad electrónica requiere conocer la amplitud y la fase de los factores de estructura. Sin embargo en los experimentos de difracción no se puede medir la fase. Existen diversos métodos para resolver el problema de la fase. En el caso del dominio Calx- β y del FnIII-3 se recurrió al método de reemplazo molecular, mientras que para el dominio FnIII-4 se aplicó el método de reemplazo isomórfico simple con dispersión anómala (SIRAS).

4.4.1.- Reemplazo molecular: faseado de los dominios Calx- β y FnIII-3 de $\beta 4$

Las estructuras de los dominios Calx- β y FnIII-3 se resolvieron mediante reemplazo molecular empleando el programa PHASER (McCoy et al., 2005) dentro de la suite de programas CCP4 (Winn et al., 2011). En las búsquedas se emplearon modelos construidos a partir de estructuras cristalográficas de dominios homólogos. Para ello se partió de alineamientos de secuencia basados en perfiles, obtenidos con el servidor FFAS03 (Jaroszewski et al., 2005), a partir de los cuales se construyeron modelos híbridos utilizando el servidor SCWRL (Canutescu et al., 2003). En estos modelos híbridos se mantuvieron las conformaciones de las cadenas laterales de los residuos idénticos en ambas secuencias, mientras que los residuos no conservados se sustituyeron por Ser, con la excepción de los residuos de Ala y Gly, que se truncaron hasta los átomos de C^β y C^α , respectivamente.

4.4.2.- Reemplazo isomórfico simple con dispersión anómala (SIRAS): faseado del dominio FnIII-4 de β_4

La estructura del dominio FnIII-4 se resolvió mediante SIRAS empleando un derivado de Hg. La subestructura de átomos de Hg se resolvió empleando los programas SHELX C/D/E (Schneider and Sheldrick, 2002) y la interfaz gráfica HKL2MAP (Pape and Schneider, 2004). La subestructura de átomos de Hg y las distribuciones de probabilidad de sus fases se refinaron y calcularon con el programa autoSHARP (Vonrhein et al., 2007). Por último las fases se mejoraron usando los programas Solomon (Abrahams and Leslie, 1996) y DM (Cowtan, 1994). Las fases experimentales así obtenidas permitieron el cálculo de mapas de densidad electrónica interpretables. Empleando estos mapas se construyó de forma automática un modelo inicial con el programa ARP/WARP (Langer et al., 2008).

4.5.- Refinado de las estructuras cristalográficas

Las estructuras se refinaron utilizando el programa *PHENIX REFINE* (Afonine et al., 2005) empleando funciones de máxima probabilidad. Durante el refinado se monitorizó el factor R de desacuerdo (*Rwork*) que compara las amplitudes observadas experimentalmente (*Fobs*) con las calculadas a partir de los modelos (*Fcalc*) según la expresión:

$$R = 100 \left(\frac{\sum_{(hkl)} |Fobs| - |Fcalc|}{\sum_{(hkl)} |Fobs|} \right)$$

También se empleó un factor R (*Rfree*) calculado con un 5% de las reflexiones que se excluyeron del refinado.

Los pasos de refinado se alternaron con la inspección visual y en su caso la modificación manual de las estructuras usando el programa *COOT* (Emsley and Cowtan, 2004). Se emplearon mapas $2mFobs-DFcalc$ y $mFobs-DFcal$; también se emplearon mapas

promediados a partir de múltiples modelos en los que se introdujeron desplazamientos atómicos aleatorios (mapas *averaged kick*) (Praznikar et al., 2009).

El refinado se inició a partir de los modelos del Calx- β y del FnIII-3 obtenidos mediante reemplazo molecular o a partir del modelo parcial del FnIII-4 construido con ARP/WARP empleando los mapas obtenidos mediante SIRAS. En las primeras etapas de refinado se aplicaron métodos de refinado de cuerpos rígidos y métodos de recocido simulado. En etapas más avanzadas se refinaron las posiciones atómicas, los valores B o parámetros de desplazamiento atómico (ADP) isotrópicos individuales y los grupos ADP según un modelo de translación, libración y desplazamiento helicoidal (TLS) de cuerpos pseudo-rígidos (grupos TLS) (Schomaker and Trueblood, 1968; Winn et al., 2001). Los grupos TLS se identificaron mediante el análisis de los valores de ADP de las estructuras utilizando el servidor TLS Motion Determination (<http://skuld.bmsc.washington.edu/~tlsmd/>) (Painter and Merritt, 2006). Se incluyeron moléculas de solvente que se construyeron en picos positivos (típicamente de más de 3.5σ) de los mapas *mFobs-DFcal*, cuando podían formar potenciales puentes de hidrógeno. En determinados casos se incluyeron otras moléculas de bajo peso molecular presentes en el medio de cristalización. En las etapas finales del refinado se incluyeron átomos de hidrógenos de las moléculas de proteína. Las posiciones de los átomos de hidrógeno no se refinaron y éstos tomaron los valores B del átomo al que están unidos.

4.6.- Análisis de las estructuras cristalográficas

La geometría de las estructuras se analizó con el programa MOLPROBITY (Davis et al., 2007) (<http://molprobity.biochem.duke.edu/>) y con el servicio de validación ADIT (<http://validate.rcsb.org/>). La superposición de estructuras por pares se realizó con el servidor DALI Lite (Holm and Park, 2000) (<http://www.ebi.ac.uk/Tools/dalilite/>) o con el programa LSQKAB de CCP4 (Kabsch, 1976); estos dos métodos emplean ajustes de mínimos cuadrados. La superposición simultánea de múltiples estructuras se realizó con el programa THESEUS (Theobald and Steindel, 2012) que utiliza una función de máxima similitud. Las propiedades electrostáticas de las estructuras se analizaron con el

programa APBS (Baker et al., 2001). Las figuras de estructuras moleculares se crearon con el programa PyMOL (<http://www.pymol.org/>).

5.- Dispersión de rayos-X a bajo ángulo (SAXS)

La dispersión de rayos-X a bajo ángulo (SAXS) es una técnica ampliamente utilizada para estudiar sistemas no cristalinos como proteínas en solución. En un experimento de SAXS se hace incidir un haz de rayos-X sobre una disolución de proteína u otra macromolécula y se mide la radiación dispersada a ángulos muy bajos. SAXS es una técnica de contraste, por lo cual para conocer la dispersión producida por las macromoléculas es necesario sustraer la señal de dispersión debida al solvente. A partir del análisis de los perfiles de dispersión es posible estimar diversos parámetros estructurales de las proteínas objeto de estudio, como el radio de giro (R_g), la función de probabilidad de distribución de distancias ($P(r)$), la distancia máxima intramolecular (D_{max}) o el volumen de la partícula. Asimismo, a partir del perfil de dispersión es posible calcular modelos a baja resolución de la proteína utilizando métodos de modelado *ab initio*.

5.1.- Adquisición de datos de SAXS

Las medidas de SAXS de la región FnIII-3,4 (1457-1666) de β_4 se realizaron en la estación de dispersión coherente de rayos-X (cSAXS) de la fuente de radiación sincrotrón Swiss Light Source (SLS) del Paul Scherrer Institut (PSI, Villigen, Suiza). El detector (PILATUS 2M, Dectris Ltd.) se situó a una distancia de 2.15 metros de la muestra, de tal forma que se cubrió un rango de vectores de dispersión ($q=4\pi\sin\theta/\lambda$) desde aproximadamente 0.015 hasta 0.5 \AA^{-1} . La energía del haz de rayos-X se ajustó a 12 KeV ($\lambda=1.0 \text{\AA}$). Las muestras de proteína a diferentes concentraciones (desde 1 hasta 16 mg ml^{-1}) se midieron empleando capilares de borosilicato ($\varnothing = 1\text{mm}$, Hilgenberg GmbH) situados en un soporte de cobre termostatzado a $10 \pm 1 \text{ }^\circ\text{C}$. De la misma manera se midió el tampón en el que estaba disuelta la proteína (20 mM fosfato sódico (NaPi) (pH 7.5), 150 mM NaCl, 5% (v/v) glicerol, 3 mM DTT). Tanto el glicerol como el DTT minimizan los posibles daños por radiación de la muestra, secuestrando las especies reactivas de oxígeno generadas en la disolución acuosa por los rayos-X. Además, con el

fin de evitar o reducir al mínimo estos daños por radiación se tomaron datos en 10 posiciones diferentes de cada capilar (espaciados 0.5 mm entre sí), realizando un barrido a lo largo de estas posiciones con exposiciones de 0.5 s en cada punto y un tiempo total de exposición de 30 s. Los perfiles de dispersión se promediaron radialmente y, tras comprobar que no existía daño por radiación, los diversos perfiles tomados a lo largo del tiempo para una misma muestra se promediaron con objeto de maximizar la relación señal/ruido.

5.2.- Procesado y análisis de datos de SAXS

Las curvas de intensidad, I , en función del ángulo de dispersión, q , obtenidas para cada una de las concentraciones de proteína medidas, así como para el disolvente, se procesaron con el programa PRIMUS (Konarev et al., 2003). En primer lugar, se restó la contribución del disolvente de las muestras de proteína. Se normalizó por la concentración de proteína, se superpusieron los perfiles de las diferentes concentraciones de la muestra y se extrapolaron los datos a dilución infinita.

El radio de giro (R_g), definido como la distancia cuadrática media de los átomos de la proteína al centro de masas de la proteína, se calculó mediante la aproximación de Guinier, según la cual a bajo ángulo la dispersión se describe como:

$$\ln[I(q)] = \ln[I(0)] - \frac{q^2 R_g^2}{3}$$

El gráfico de Guinier, en el que se representa el $\ln(I(q))$ frente a q^2 , ha de ser lineal en el rango $0.7 < qR_G < 1.3$, para una proteína plegada, pura y monodispersa (Guinier and Fournet, 1955). El R_g se obtuvo de la pendiente del gráfico lineal de Guinier empleando el programa PRIMUS. Las posibles desviaciones del comportamiento lineal en la región de Guinier son una clara indicación de la presencia de interacciones inter-partícula, bien sean atractivas (agregación) o repulsivas, y es por tanto un buen punto de control de la calidad de las muestras.

La función de distribución de pares de distancia, $P(r)$, se obtuvo mediante métodos indirectos con el programa GNOM (Svergun, 1992). A partir de esta función se calcularon la distancia máxima en la partícula (D_{max}) y el R_g . Además, la forma de la función $P(r)$ permite distinguir entre proteínas globulares y alargadas: las proteínas globulares se caracterizan por presentar una función de distribución de pares de distancia de tipo gaussiano con un único pico cuyo valor máximo se corresponde con el radio de la esfera. Por el contrario, las moléculas muy alargadas (tipo prolato) presentan un pico con una cola que se extiende hacia valores altos de la distancia r .

En el gráfico de Kratky (Glatter and Kratky, 1982) se representan los datos de dispersión (q) frente a $I(q) \cdot q^2$. Esta representación se utiliza habitualmente para analizar o inspeccionar visualmente la conformación de las proteínas o el correcto plegamiento de éstas.

5.3.- Generación de modelos a partir de datos de SAXS

La envuelta a baja resolución de la región FnIII-3,4 de β_4 se modeló utilizando métodos *ab initio* que se basan exclusivamente en los datos experimentales de SAXS y no requieren de información externa adicional. Estos métodos utilizan la ecuación de Debye (Debye, 1915) para generar el perfil teórico de un conjunto de esferas, de un determinado radio, colocadas aleatoriamente al inicio. Los modelos van siendo optimizados mediante algoritmos genéticos en el programa DALAI_GA (Chacon et al., 2000; Chacon et al., 1998) o mediante "recocido simulado" en el programa DAMMIN (Svergun, 1999) hasta encontrar un modelo, cuyo perfil de dispersión teórico se ajusta lo mejor posible a los datos experimentales.

Para asegurar la diversidad del modelo, se realizaron 10 ejecuciones independientes de forma paralela con cada uno de los programas, DALAI_GA y DAMMIN. Todos los modelos generados se superpusieron por pares utilizando el programa SUPCOMB20 (Kozin and Svergun, 2001) y su similitud se cuantificó empleando el valor de discrepancia espacial normalizada (NSD), que es cero para modelos idénticos y que es mayor de uno para modelos que difieren sistemáticamente entre sí (Kozin and Svergun, 2001). Normalmente, se toma como referencia el modelo con el menor valor de NSD y se

descartan todos aquellos modelos cuyo NSD medio con respecto al resto de modelos fuese mayor que la NSD media de todos los modelos más 2 veces la desviación estándar. Los modelos resultantes se promediaron, con el programa DAMAVER (Volkov and Svergun, 2003) y se filtraron utilizando el programa DAMFILT (Volkov and Svergun, 2003), para reducir su volumen a la mitad, quedando sólo el volumen más poblado para cada grupo de estructuras.

Para la representación gráfica de las estructuras de baja resolución se empleó la suite de programas SITUS (Wriggers, 2010); primero se crearon envueltas tridimensionales a partir de los modelos pseudoatómicos usando el programa PDB2VOL, las cuales se convirtieron al formato de mapas de CCP4 con el programa MAP2MAP para su posterior representación con el programa PyMOL.

5.4.- Comparación de modelos atómicos con los datos de SAXS.

Las curvas teóricas de dispersión correspondientes a estructuras atómicas se calcularon y se compararon con los datos experimentales de SAXS con el programa CRY SOL (Svergun et al., 1995). Los modelos atómicos se superpusieron a las estructuras de baja resolución obtenidas a partir los datos de SAXS empleando el programa SUPCOMB20 (Kozin and Svergun, 2001).

6.- Medida de distancias intramoleculares mediante resonancia paramagnética electrónica (EPR).

Se determinaron distancias intramoleculares entre pares de residuos combinando marcaje paramagnético dirigido (SDSL) con resonancia paramagnética electrónica (EPR).

6.1.- Marcaje de proteínas con sondas paramagnéticas

Mutantes de la región FnIII-3,4 de $\beta 4$ (residuos 1457-1666) que contienen dos o tres Cys en posiciones específicas se marcaron con la sonda paramagnética (2,2,5,5-tetrametilpirrolina-1-oxyl-3-metil) metanotiosulfonato (MTSSL) (Enzo Life Science). Esta sonda reacciona específicamente con los grupos sulfhidrilo presentes en las Cys de la

proteína formando enlaces disulfuro a través de los cuales el grupo piperidin-oxyl queda unido (Figura M2).

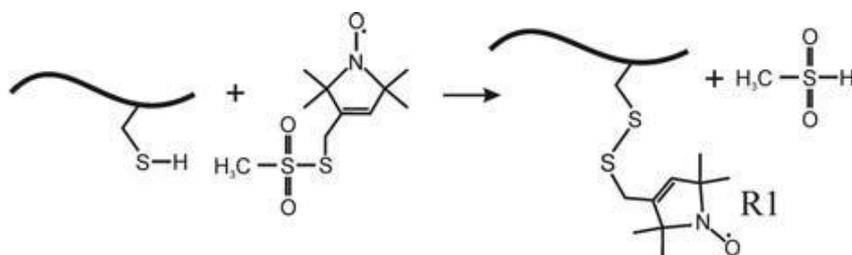


Figura M2.- Reacción de marcaje con MTSSL. La sonda paramagnética MTSSL reacciona con el grupo sulfhidri lo de una Cys de la proteína.

Los marcajes se realizaron típicamente incubando 500 μ l de proteína a 200 μ M en 20 mM NaPi (pH 7.5), 150 mM NaCl con un exceso molar 10x de MTSSL respecto a la concentración de proteína. La reacción se llevó a cabo con rotación durante 1 h a temperatura ambiente y seguidamente durante 12 h a 4 °C. El exceso de sonda libre se eliminó mediante 8 cambios de diálisis frente a 200 ml de 20 mM NaPi (pH 7.5), 150 mM NaCl.

6.2.- Valoración de grupos tiol libres con DTNB

El grado de marcaje con MTSSL de las proteínas β 4 se realizó mediante valoración de los grupos tiol libres con ácido 5,5´ditio-bis(2-nitrobenzoico) (DTNB) antes y después de la reacción de marcaje. El DTNB reacciona con grupos tiol libres de las Cys produciendo 2-nitro-5-tiobenzoato que es monitorizado espectrofotométricamente a 412 nm (Figura M3).

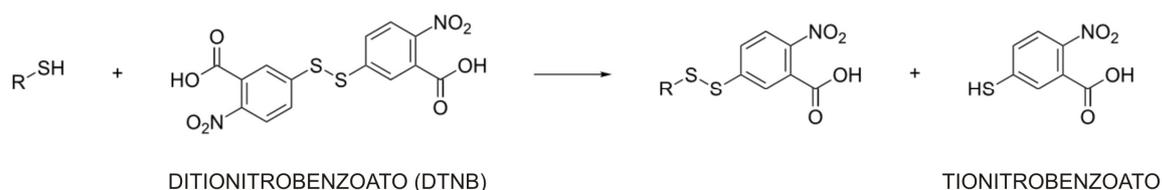


Figura M3.- Reacción de DTNB con grupos tiol libres. El DTNB reacciona con el grupo tiol de una Cys de la proteína; como producto de la reacción se genera tionitrobenzoato.

Se prepararon 300 μl de las proteínas marcadas y sin marcar a concentración conocida $\sim 10 \mu\text{M}$ en 20 mM NaPi (pH 8.0), 150 mM NaCl, 6 M cloruro de guanidinio. Se le añadieron 3 μl de 0.1 M DTNB (concentración final 1 mM) y la reacción se desarrolló en oscuridad a temperatura ambiente durante 30 min. La concentración de 2-nitro-5-tiobenzato, que corresponde a la de grupos tiol libres, se determinó espectrofotométricamente usando el coeficiente de extinción molar $13600 \text{ M}^{-1} \text{ cm}^{-1}$ a 412 nm.

El número de grupos tiol libres por molécula de proteína se calculó como el cociente entre la concentración de tionitrobenzoato y de proteína.

$$n^{\circ} \text{ de Cys libres por molécula} = \frac{[2\text{-nitro-5-tiobenzato}]}{[\text{proteína}]}$$

6.3.- Experimentos de resonancia paramagnética electrónica (EPR).

Los experimentos de resonancia de doble electrón-electrón (DEER) fueron realizados por la Dra Inés García Rubio en el centro ETH de Zurich (Suiza). Tras el marcaje con MTSSL los mutantes de $\beta 4$ se congelaron mediante inmersión en nitrógeno líquido y se enviaron al ETH en hielo seco. Una vez en el ETH, las muestras se descongelaron y se les añadió glicerol deuterado hasta llegar a una concentración de 30-50%. La deuteración disminuye la relajación de la sonda nitróxido, permitiendo obtener trazos de DEER más largos y mejorando la determinación de la línea base decreciente debida a interacciones espín-espín entre sondas de distintas moléculas. Por lo tanto, la distancia entre sondas paramagnéticas de una misma molécula es mayor para muestras deuteradas que sin deutar. Las muestras se transfirieron a tubos de EPR y se midieron a temperatura ambiente en onda continua mediante resonancia paramagnética electrónica (CW-EPR) con el fin de comprobar la eficiencia del marcaje. La distancia entre dos sondas de spin se midió usando experimentos de DEER (Jeschke and Polyhach, 2007) a unas frecuencias de microondas en rango de banda Q ($\sim 35 \text{ GHz}$), debido a su mayor sensibilidad. Los espectros obtenidos se corrigieron para eliminar la contribución de distancias entre diferentes dímeros y se analizaron usando procedimientos de regulación Tikonov para obtener distribuciones de distancias entre sondas de spin dentro del dímero.

Para cada una de las posiciones que han sido sustituidas por cisteína para ser marcadas con sonda paramagnética, se ha hecho un análisis con la librería de rotámeros integrada en el programa MMM (<http://www.epr.ethz.ch/software/index>). El resultado de este análisis consiste en la lista de los rotámetros posibles de la sonda paramagnética para cada una de las posiciones con su correspondiente probabilidad. Con esta lista de rotámetros, se obtuvo una posición del espín (supuesto localizado en el enlace N-O del grupo nitroxilo) promedio para cada posición.

$$(\bar{X}, \bar{Y}, \bar{Z}) = \left(\sum_i x_i p_i, \sum_i y_i p_i, \sum_i z_i p_i \right)$$

siendo i cada uno de los rotámetros posibles. Las coordenadas de este promedio se consideran la posición de la sonda paramagnética.

Para obtener la posición del residuo C1608 con respecto al dominio FnIII-3, se muestreó todo el espacio alrededor de este dominio con una malla de paso 1 Å. De todos los puntos de la malla, se seleccionó el de menor varianza (x, y, z) considerando todas las distancias de dobles mutantes con C1608 obtenidas experimentalmente (d) y que dieron como resultado distribuciones de distancias estrechas (un total de 5 cotas de distancia).

$$\text{varianza} = \sum_{j=1}^5 (x - \bar{X}_j)^2 + (y - \bar{Y}_j)^2 + (z - \bar{Z}_j)^2 - d_j^2$$

Así, se obtuvo para la sonda en la posición C1608, las coordenadas (43, -19, 9) identificando estas coordenadas con las obtenidas de la biblioteca de rotámetros

$$(\bar{X}_{C1608}, \bar{Y}_{C1608}, \bar{Z}_{C1608}).$$

Una vez obtenida la posición del residuo C1608 con respecto al dominio FnIII-3, todavía quedan tres grados de libertad de rotación del dominio FnIII-4 manteniendo el punto C1608 fijo. Para obtener cuál de las rotaciones corresponde con la orientación real del dominio 4 con respecto al 3, se utilizaron las cotas de distancia de los mutantes con tres sondas paramagnéticas, de los que podemos obtener las distancias entre el residuo C1598 del dominio 4 y los residuos C1523 y C1504 del dominio 3. De nuevo se

seleccionaron las posibles rotaciones que colocan al residuo C1598 en un punto (x, y, z) que minimiza la varianza

$$\text{varianza} = \sum_j (x - \bar{X}_j)^2 + (y - \bar{Y}_j)^2 + (z - \bar{Z}_j)^2 - d_j^2$$

con $j = 1, 2$ (cada una de las 2 cotas experimentales obtenidas para este residuo).

Fijando este punto, todavía queda un grado de libertad, la rotación con respecto al eje C1608-C1598. Se han generado modelos que difieren en este ángulo (un modelo cada 10°). Comparando los modelos generados con los datos de SAXS, se seleccionó el modelo final propuesto.

7.- Producción de CD151-EC2 en bacterias

7.1.- Clonación del EC2 de CD151 en vectores de expresión en bacterias

Las secuencia de ADNc que codifica el dominio EC2 (112-219) de la tetraspanina CD151 humana (código de Uniprot: P48509) se amplificó mediante PCR empleando la polimerasa Pfu. Como plantilla se utilizó el clon *IMAGE* 2958615, que contiene el ADNc de CD151 en el vector pOTB7. La región de ADNc que codifica para el dominio EC2 contiene un sitio de restricción BamHI que se destruyó mediante mutagénesis dirigida empleando los cebadores CD151-BamX-F y CD151-BamX-R (Tabla M3). Todas las construcciones de CD151 se crearon a partir de esta secuencia modificada.

Tabla M3.- Oligonucleótidos utilizados para destruir el sitio BamHI en la secuencia de CD151.

CEBADOR	
NOMBRE	SECUENCIA
CD151-BamX-F	5' GG CGA GAC AGT GAG TGG ATT CGC TCA CAG GAG GCC 3'
CD151-BamX-R	5' GGC CTC CTG TGA GCG AAT CCA CTC ACT GTC TCG CC 3'

El fragmento EC2 de CD151 se clonó en los vectores de expresión en bacterias pETEV15b, pET15b (Novagen), pET22b (Novagen), pETZZZ y una variante del vector pET32a (Novagen). También se intentó la clonación en el vector pEZZ18. En todos los casos la región que codifica para el EC2 se amplificó mediante PCR utilizando los

cebadores específicos que se indican en la tabla M4. Los nuevos sitios de restricción introducidos al comienzo y al final del ADNc amplificado son compatibles con los sitios de clonación de cada vector.

Tabla M4.- Cebadores utilizados para clonar del dominio EC2 de CD151 en los vectores pET empleando los sitios NdeI y BamHI.

CEBADOR	
NOMBRE	SECUENCIA
CD151-U112	5' TGA <u>GAA TTC</u> <u>CAT ATG</u> TAC GCC TAC TAC CAG CAG 3' EcoRI NdeI Y A Y Y Q Q
CD151-L219	5' GCC <u>GAA TTC</u> <u>GGA TCC CTA</u> GTG CTC CTG GAT GAA GGT C3' EcoRI BamHI Stop H E Q I F T

7.1.1.- Clonación de CD151-EC2 en los vectores pET15b, pETEV15b y pETEZZ

La región que codifica para el CD151-EC2 se clonó de igual manera en el vector pET15b y en dos derivados de éste creados en nuestro laboratorio, denominados pETEV15b (ver antes) y pETEZZ. Estos vectores codifican proteínas de fusión con secuencias adicionales en el extremo N-terminal. El vector pET15b codifica para una cola de 6 His y una secuencia reconocida por trombina (Figura M4-A). El vector pETEV15b codifica para una cola de 8 His, seguida de una secuencia reconocida por la proteasa del TEV (Figura M4-B). Por último en el vector pETEZZ se creó a partir del vector pET15b sustituyendo la cola de poli-His por la secuencia señal de la proteína A y dos copias del dominio Z obtenidas a partir del vector pEZZ-18 (GE Healthcare); tras esta región el vector pETEZZ codifica para un sitio de digestión por trombina (Figura M4-C y D). En todos los casos la región del EC2 se clonó empleando los sitios de restricción NdeI y BamHI.

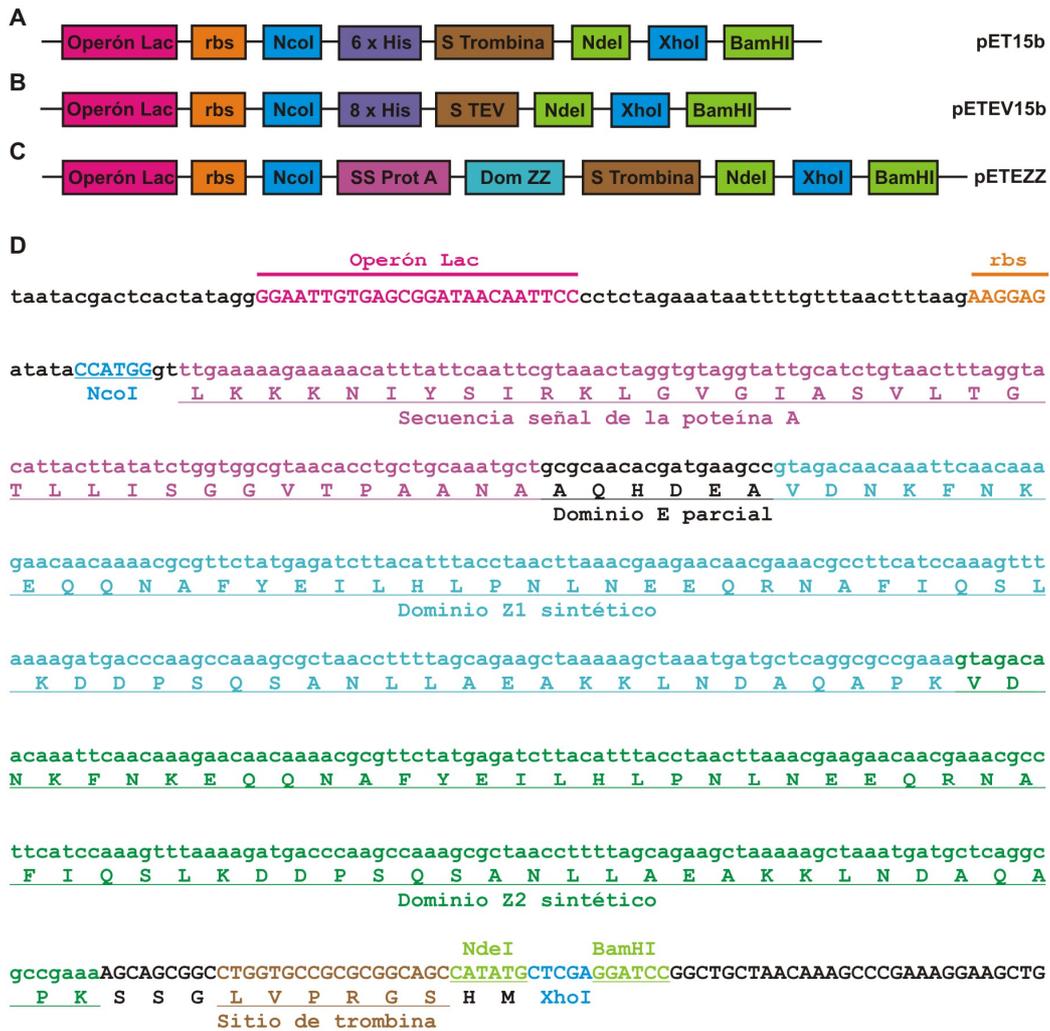


Figura M4.- Regiones de clonación y expresión de los vectores pET15b, pETEV15b y pETEZZ. Esquemas de la región de clonación en los vectores pET15b (A), pETEV15b (B) y pETEZZ (C). Todas las secuencias contienen el promotor del operón de la lactosa (Operón Lac), el sitio de unión al ribosoma (rbs) y varios sitios reconocidos por enzimas de restricción (se muestran en fondo verde aquellas que se usaron para la clonación de CD151-EC2). Algunos de los vectores contienen una cola de poli-His en la que se indica el número de His que contienen (6xHis u 8xHis). El sitio de restricción NdeI es precedido por una secuencia que codifica un sitio de reconocimiento de trombina (S trombina) o de la proteasa del virus del grabado del tabaco (S TEV). La secuencia señal de la proteína A (SS Prot A) y el doble dominio Z sintético (Dom ZZ) son específicos del vector pETEZZ. (D) Secuencia de la región de expresión de la hebra codificante del vector pETEZZ transcrita por la T7 RNA polimerasa.

7.1.2.- Clonación de CD151-EC2 en el vector pET32aNco

El vector pET32aNco es un derivado del vector comercial pET32a (Novagen), el cual codifica una proteína de fusión con tiorredoxina (Trx) en la región amino. La expresión de proteínas de fusión con Trx en *E. coli* deficiente en tiorredoxina reductasa, como la

cepa Origami, favorece la formación de puentes disulfuro en el citoplasma. El vector pET32aNco se construyó eliminando los dos sitios NdeI localizados en pET32a a ambos lados de la región de Trx e introduciendo un sitio NcoI tras el gen de Trx (Figura M5). A partir de la construcción del EC2 en el vector pET15b se obtuvo la región que codifica para la cola de poli-His, el sitio de trombina y el EC2, mediante digestión con las enzimas NcoI y BamHI. Este fragmento se insertó empleando los mismos sitios de restricción en el vector pET32aNco. La construcción resultante codifica para una proteína de fusión con la Trx seguida de una cola de 6 His, un sitio de trombina y el EC2 de CD151.



Figura M5.- Regiones de clonación y expresión del vector pET32a y del derivado pET32aNco. Esquemas de estas regiones en el vector pET32a (A) y pET32aNco (B). Ambas secuencias contienen el promotor del operón de la lactosa (Operón Lac), el sitio de unión al ribosoma (rbs), el gen de tiorredoxina (Trx), una cola de seis Histidinas (6 x His), un sitio de reconocimiento de trombina (S trombina) y la etiqueta S. Los sitios de restricción utilizados para la clonación de CD151-EC2 se muestran con fondo verde (NcoI y BamHI). (C) Secuencia de la región de expresión de la hebra codificante del vector pET32aNco transcrita por la T7 RNA polimerasa.

7.1.3.- Clonación de CD151-EC2 en el vector pET22b

El vector pET22b(+) (Novagen) codifica para proteínas de fusión que contienen en el extremo amino la secuencia señal pelB para la exportación de la proteína al espacio periplasmático (Figura M6). El CD151-EC2 se clonó en pET22b(+) a partir de la construcción del EC2 en pET15b. La región que codifica para la cola de poli-His, el sitio de trombina y el EC2 se escindió mediante digestión con NcoI y BamHI, y se clonó en pET22b(+) empleando los mismos sitios de restricción. La construcción resultante codifica para una proteína que contiene la secuencia pelB seguida de una cola de poli-His, el sitio de trombina y el CD151-EC2.



Figura M6.- Esquema de la región de clonación y expresión del vector pET22b(+). La secuencia contiene el promotor del operón de la lactosa (Operón Lac), el sitio de unión al ribosoma (rbs), una secuencia de secreción periplasmática de la pectato liasa B (pelB) y varios sitios reconocidos por enzimas de restricción.

7.2.- Expresión de CD151-EC2 en bacterias y purificación

El CD151-EC2 se expresó en la cepa BL21(DE3) de *E. coli* empleando la construcción creada en el vector pET15b. El CD151-EC2 se localizó en la fracción insoluble, se solubilizó en presencia de 6 M urea y se purificó en condiciones desnaturizantes mediante cromatografía de afinidad empleando columnas quelantes de Ni²⁺ de forma similar a como se purificó el dominio Calx-β. Como paso previo a ser replegada y con el fin de reducir las Cys, la proteína se incubó 1 h a 30 °C en presencia de 2 mM DTT. La proteína se renaturalizó mediante dilución 1:30 en 20 mM Tris-HCl (pH 8.0), 150 mM NaCl en la que se redujo la concentración de urea rápidamente de 6 M a 0.2 M. Se incluyeron adyuvantes como 50 mM arginina y 50 mM ácido glutámico. Seguidamente se indujo la formación de los puentes disulfuro mediante oxidación con 0.1 mM CuSO₄ y 0.3 mM o-fenantrolina durante 1 h a 4 °C. Una vez replegado el EC2, se concentró hasta un volumen de 50 ml en una celda de ultrafiltración Amicon de 400 mL (Millipore) empleando una membrana con un límite de permeabilidad de 3 kDa. La cola de poli-His

se eliminó mediante digestión con trombina en presencia de 1 mM de CaCl₂ a 30 °C durante 2 h. La reacción se detuvo por adición de 1 mM PMSF. Finalmente la muestra se concentró hasta un volumen menor de 5 ml y se purificó por cromatografía de exclusión molecular en una columna de alta resolución Sephacryl S-200 (26 x 600 mm) (GE Healthcare) equilibrada en 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 50 mM Arg, 50 mM Glu. Por último el CD151-EC2 se concentró mediante ultrafiltración en celdas Amicon hasta ~ 20 mg ml⁻¹ (~1.57 mM).

8.- Producción de CD151-EC2 en células de insecto

8.1.- Clonación de CD151-EC2 en vectores de transferencia para baculovirus

La región que codifica para el CD151-EC2 se ha clonado en el vector pAcGP67A (BD Biosciences), en el vector pFastBac-1 (Invitrogen) y en versiones modificadas de este último.

8.1.1.- Clonación de CD151-EC2 en el vector pAcGP67A.

A partir de la construcción del CD151-EC2 en el vector pET15b, se amplificó mediante PCR la región que codifica para la cola de 6-His, el sitio reconocido por trombina y el CD151-EC2. Para ello se emplearon los oligos PET-Bgl-F y PET-Eco-R (Tabla M5), que incluyen cerca de sus extremos 5' sitios BglIII y EcoRI respectivamente. El producto de la PCR se digirió con las enzimas BglIII y EcoRI y se clonó en el vector pAcGP67A previamente digerido con BamHI y EcoRI. Los extremos de cadena sencilla generados por BglIII (inserto) y BamHI (vector) son compatibles y tras la ligación se destruyen ambos sitios de restricción. La construcción resultante se denominó pAcGP67A-CD151-01 y codifica para la secuencia de secreción de la proteína gp67 seguida de la cola de 6-His, el sitio de trombina y el EC2.

Tabla M5.- Cebadores utilizados para el clonaje del dominio EC2 de CD151 en el vector pAcGP67A.

CEBADOR	
NOMBRE	SECUENCIA
PET-Bgl-F	5' GGA GAT ATA <u>GAT CTG</u> GGC AGC AGC CAT CAT C 3' Bgl II G S S H H
PET-Eco-R	5' GCC <u>GAA TTC</u> GCT TTG TTA GCA GCC <u>GGA TCC</u> 3' EcoRI A L L A A BamHI

8.1.2.- Clonación de CD151-EC2 en el vector pFastBac-1 y derivados.

A partir de la construcción pAcGP67A-CD151-01 se amplificó mediante PCR la región que codifica para la secuencia señal de gp67, la cola de 6-His, el sitio de trombina y el EC2. Se emplearon los oligonucleótidos GP67-F-BglIII y CD151-L219EcoRI (Tabla M6). El producto de la PCR se digirió con las enzimas BglIII y EcoRI y se ligó en el vector pFastBac-1 (Invitrogen) previamente digerido con BamHI y EcoRI. Esta construcción se denominó pFastBac1-gp67-His-CD151-EC2.

Tabla M6.- Cebadores utilizados para el clonaje del dominio EC2 de CD151 en el vector pFastBac-1 y sus derivados.

CEBADOR	
NOMBRE	SECUENCIA (sentido 5' a 3')
GP67-F-BglIII	GGG AGA TAT <u>AGA TCT</u> ATG CTA CTA GTA AAT CAG TCA CAC G R Y BglIII M L L V N Q S H
CD151-L219EcoRI	CC TTT <u>GAA TTC</u> CTA GTG CTC CTG GAT GAA GGT C EcoRI Stop H E Q I F T

A partir del vector pFastBac-1 se creó un derivado denominado pFastBac-1-CHis que codifica una proteína de fusión con una cola de 6-His en el extremo C-terminal. Este vector se creó introduciendo en el vector pFastBac-1 digerido con NotI y XhoI una secuencia de ADN formada por la combinación de los oligonucleótidos C-6His-Not-Xho-F y C-6His-Not-Xho-R (Tabla M7). De esta forma se destruyó el sitio XhoI original y se introdujo un nuevo sitio XhoI que precede a la secuencia que codifica una cola de 6-His tras la cual existe un codón de terminación.

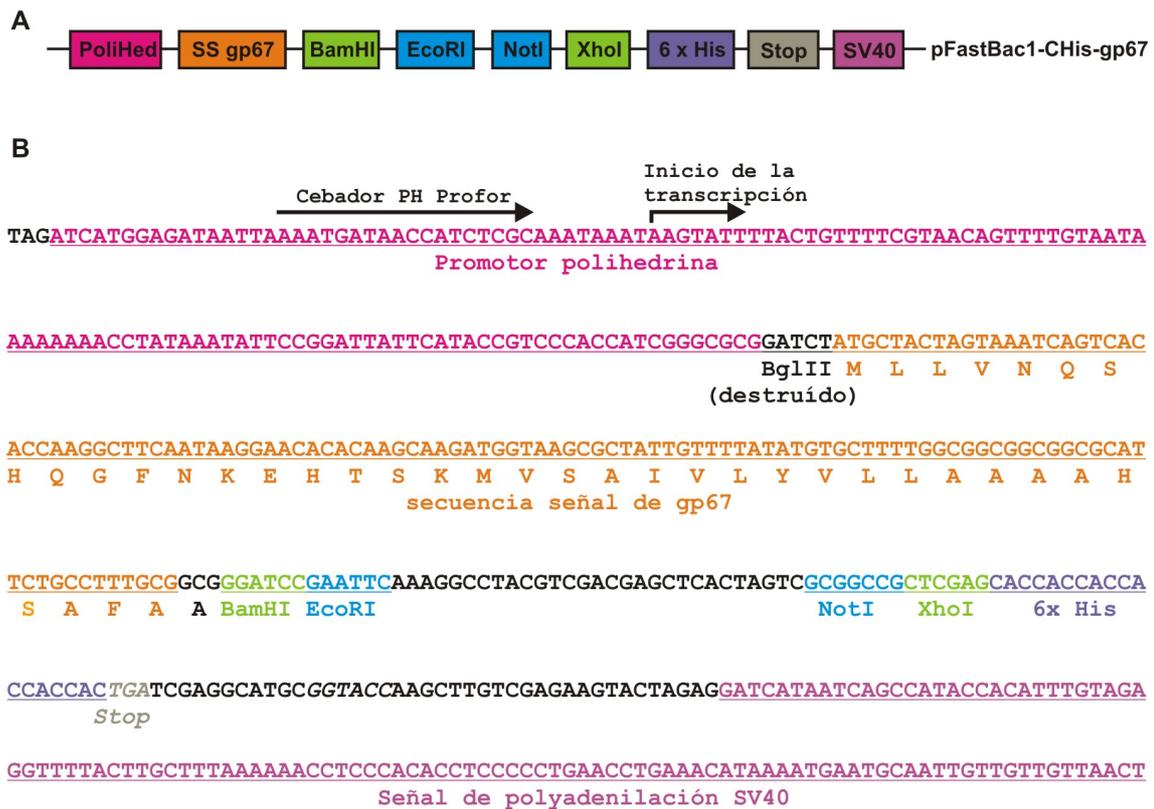


Figura M7.- Región de clonación y expresión del vector pFastBac1-CHis-gp67. (A) Esquema de la región de expresión del vector pFastBac1-CHis-gp67. La secuencia contiene un promotor de polihedrina (PoliHed) del nucleopolihedrovirus multicápside de la *Autographa californica* (AcMNPV), la secuencia señal gp67 (SSgp67), un sitio de clonación múltiple con varias secuencias reconocidas por enzimas de restricción y una señal de poliadenilación de SV40. (B) Secuencia de la hebra codificante de esta región.

Tabla M8.- Oligonucleótidos utilizados para amplificar la secuencia señal de gp67.

CEBADOR	
NOMBRE	SECUENCIA (sentido 5' a 3')
GP67-F-BglIII	GGG AGA TAT <u>AGA TCT ATG CTA CTA GTA AAT CAG TCA CAC</u> BglIII M L L V N Q S H
GP67-R-BamHI/ EcoRI	TTT <u>GAA TTC GGA TCC CGC CGC AAA GGC AGA ATG CGC</u> EcoRI BamHI A A F A S H A

pFastBac-1-CHis-HM (Figura M8) se construyó de forma análoga amplificando la región del vector pMelBacB (Invitrogen) que codifica para la secuencia señal de la melitina de abeja (HM), utilizando los cebadores HMSS-F-BglIII Y HMSS-R-BamHI/EcoRI (Tabla M9). El producto de la PCR

se insertó en el vector pFastBac-1-CHis tal como se ha descrito en la construcción del vector pFastBac-1-CHis-GP67.



Figura M8.- Región de clonación y expresión del vector pFastBac1-CHis-HM. (A) Esquema de la región de expresión del vector pFastBac1-CHis-HM. La secuencia contiene el promotor polihedrina (PoliHed), la secuencia señal de melitina de abeja (SSH M), un sitio de clonación múltiple con varias secuencias reconocidas por enzimas de restricción y una señal de poliadenilación de SV40. (B) Secuencia de la hebra codificante de esta región.

Tabla M9.- Oligonucleótidos utilizados para amplificar la secuencia señal de HM.

CEBADOR	
NOMBRE	SECUENCIA
HMSS-F-BglIII	GGG AGA TAT <u>AGA TCT</u> ATG AAA TTC TTA GTC AAC GTT GCC BglIII M F L Y V N V A
HMSS-R-BamHI/ EcoRI	TTT <u>GAA TTC</u> <u>GGA TCC</u> CGC ATA GAT GTA AGA AAT GTA TAC EcoRI BamHI A Y I Y S I T V

Las construcciones que codifican para el CD151-EC2 en los vectores pFastBac-1-CHis-GP67 y pFastBac-1-CHis-HM se construyeron amplificando mediante PCR la región del CD151-EC2 empleando los cebadores CD151-U112BamHI y CD151-L219XhoI (Tabla M10). El producto de PCR se digirió con BamHI y XhoI y se clonó en los vectores

pFastBac-1-CHis-GP67 y pFastBac-1-CHis-HM digeridos con estas mismas enzimas. Estas construcciones se denominaron gp67-EC2-poliHis y HM-EC2-poliHis.

Tabla M10.- Oligonucleótidos utilizados amplificar la secuencia de CD151-EC2 e insertarla en los vectores pFastBac-1-CHis-GP67 y pFastBac-1-CHis-HMSS.

CEBADOR	
NOMBRE	SECUENCIA
CD151-U112BamHI	5' TGG <u>GGA TCC</u> TAC GCC TAC TAC CAG CAG 3' BamHI Y A Y Y Q Q
CD151-L219XhoI	5' GAA <u>CTC GAG</u> GTG CTC CTG GAT GAA GGT C 3' XhoI H E Q I F T

8.2.- Obtención de b́acmidos en células DH10Bac

A partir de las construcciones de CD151-EC2 en el vector pFastBac-1 y sus derivados se generaron b́acmidos recombinantes empleando la cepa DH10Bac de *E. coli*. Esta cepa contiene un vector lanzadera de baculovirus (b́acmido) con un sitio mini-*attTn7*. Cuando se transforma el plásmido de expresión pFastBac-1 en células DH10Bac, tiene lugar la transposición entre el elemento mini-Tn7 del vector y el sitio mini-*attTn7* de b́acmido generándose el b́acmido recombinante.

50 μ l de células competentes DH10Bac se transformaron con 3 μ l de ADN (30 min en hielo y choque térmico de 45 s a 42 °C). Las células se recuperaron en 900 μ l de medio SOC (*Super Optimal broth with Catabolite repression*) (Hanahan, 1983) en agitación durante 90 min a 37 °C. Las células se recogieron mediante centrifugación (1 min a 850 g), se eliminó el sobrenadante y se resuspendió el pellet en 200 μ l de medio SOC. La inserción del mini-Tn7 en mini-*attTn7* perturba la expresión del gen *LacZ α* , por lo que se pudo hacer una selección blanco/azul de las colonias positivas. Para ello, las células se sembraron en placas de LB-agar suplementadas con 50 μ g ml⁻¹ kanamicina, 10 μ g ml⁻¹ tetraciclina, 7 μ g ml⁻¹ gentamicina, 100 μ g ml⁻¹ de sustrato cromogénico X-gal (5-bromo-4-cloro-3-indolyl- β -D-galactopyranosido) y 40 μ g ml⁻¹ del inductor IPTG. Los cultivos se crecieron a 37 °C durante 48 h. Se seleccionaron tres colonias blancas, que contienen el b́acmido recombinante, de cada construcción; con el fin de verificar el fenotipo se sembraron por agotamiento en otra placa de LB-agar de las mismas características y se

crecieron a 37 °C durante 48 h. Se tomaron 3 colonias de cada una de estas nuevas placas y con cada una se inocularon 2.5 ml de LB líquido suplementado con 50 µg ml⁻¹ kanamicina, 10 µg ml⁻¹ tetraciclina y 7 µg ml⁻¹ gentamicina. Estos cultivos se crecieron a 37 °C durante 16 h con agitación.

Las células se recogieron mediante centrifugación (2 min a 18000 g) y se resuspendieron en 300 µl de 15 mM Tris-HCl (pH 8.0), 10 mM EDTA y 100 µg ml⁻¹ ribonucleasa I pancreática (RNAsa A). A la suspensión anterior se le añadieron 300 µl de 200 mM NaCl, 1% SDS y las bacterias se lisaron a temperatura ambiente durante 5 min. Seguidamente la muestra se neutralizó añadiendo 300 µl de 3 M acetato de potasio (KAc) (pH 5.5) y se incubó 10 min a 4 °C. Los agregados macromoleculares se precipitaron por centrifugación (10 min a 18000 g). El sobrenadante, que contiene el ADN del báculo, se transfirió a un tubo que contenía 800 µl de isopropanol, se mezcló por inversión y se incubó a 4 °C durante 10 min. Seguidamente la muestra se centrifugó (15 min a 18000 g). El precipitado se lavó con 500 µl de 70% etanol y se volvió a centrifugar en las mismas condiciones. El sobrenadante se eliminó y el precipitado, que contiene el báculo, se resuspendió en 35 µl de 10 mM Tris-HCl (pH 8.0).

8.3.- Obtención de baculovirus por transformación en células Sf9

Los baculovirus se generaron por transfección en células de insecto Sf9 a partir de los báculos. Se emplearon placas de cultivo celular de 6 pocillos de fondo plano de 3.5 cm de diámetro y 1.6 cm de altura (BD Biosciences). En cada pocillo se añadió 1 ml de células Sf9 a una concentración de ~10⁶ células ml⁻¹ y se incubaron 15 min a 27 °C para permitir su adhesión a la superficie. Por otro lado, a 200 µl de medio Sf900 II SFM (GIBCO) se le añadieron 10 µl de báculo a ~30 ng µl⁻¹ y 5 µl de cellFECTIN (Invitrogen). La mezcla se incubó 15 min a temperatura ambiente. Pasado ese tiempo se añadió otro mililitro de medio a la mezcla de transfección, y todo ello se llevó a las placas, de las que previamente se había eliminado el medio. Los cultivos se incubaron durante 5 h a 27 °C. Pasado este tiempo, se cambió el medio de los pocillos con 3 ml de medio fresco suplementado con 200 unidades ml⁻¹ de penicilina y con 200 µg ml⁻¹ de estreptomicina (Pen Strep de GIBCO). Los cultivos se crecieron durante 3 días a 27 °C.

El sobrenadante de los cultivos contiene los virus recombinantes. Estos sobrenadantes se aislaron mediante centrifugación (5 min, 4 °C, 450 g) y eliminación de los precipitados. Estos primeros sobrenadantes se denominan P0 (virus del pase cero) y se almacenaron a 4 °C.

En la figura M9 se resume esquemáticamente el proceso de construcción de baculovirus recombinantes y expresión del EC2 de CD151 en células de insecto.

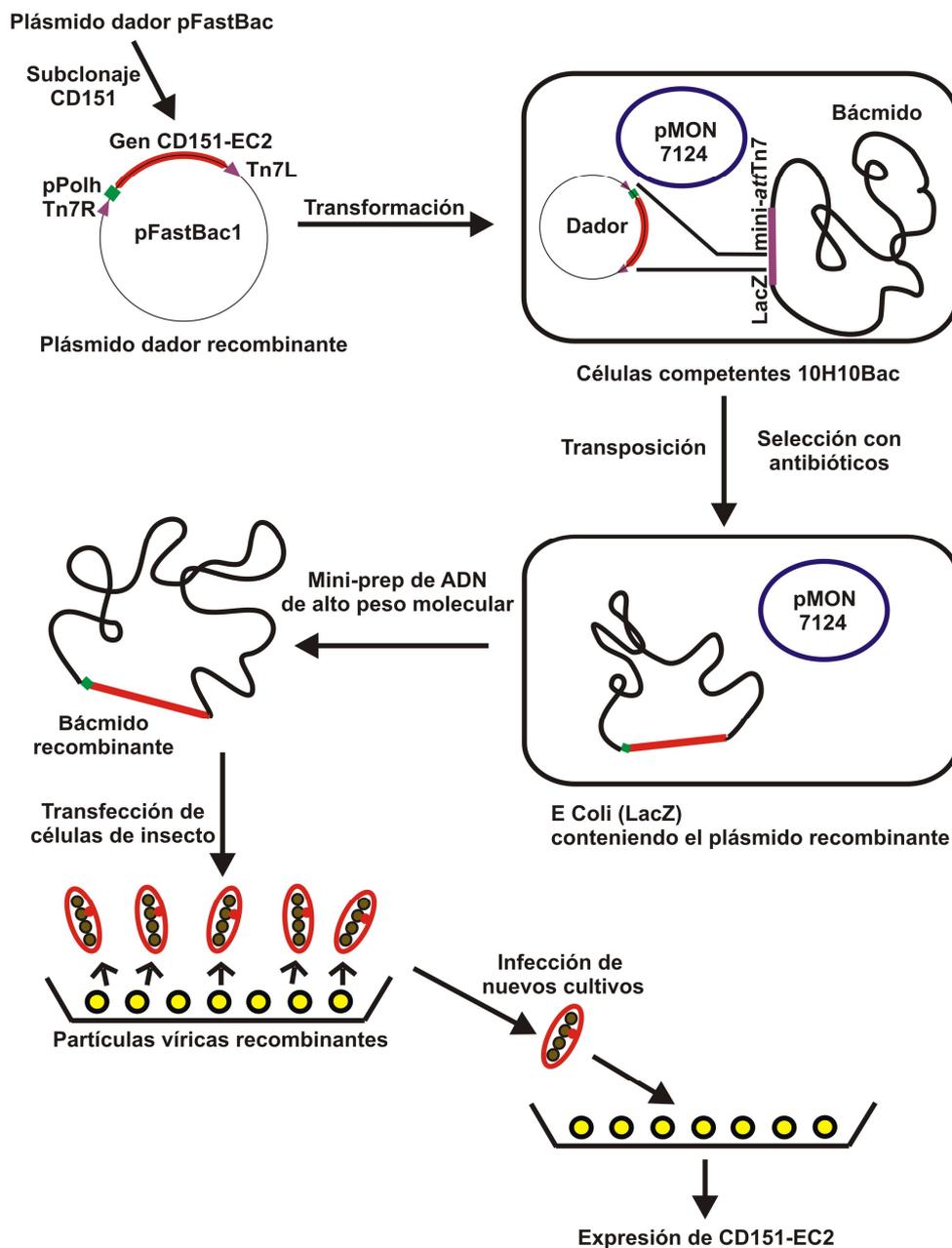


Figura M9.- Esquema general de la construcción de baculovirus recombinantes y expresión del gen de CD151 con el sistema "Bac to Bac". El gen de CD151-EC2 se clonó en el plásmido

pFastBac1. Así se generó un plásmido recombinante que se transformó en células competentes DH10Bac. Estas células contienen el elemento mini-*attTn7* en el báculo y un plásmido auxiliar (pMON 7124), de forma que el elemento Tn7 se transfiere desde el plásmido pFastBac1 al sitio de inserción mini-*attTn7* en el báculo y el plásmido pMON 7124 confiere las funciones de transposición en *trans*. Las colonias que contienen los báculos recombinantes se identificaron mediante interrupción del gen LacZ. El ADN de alto peso molecular (una mezcla de ADNs conteniendo el báculo) se extrajo mediante mini-prep. Con este ADN se transfectaron las células de insecto. Los virus así obtenidos se utilizaron para infectar nuevos cultivos con los que se expresó la proteína CD151-EC2.

8.4.- Amplificación de baculovirus recombinantes

Los virus se amplificaron a partir de la muestra P0. Para ello se infectaron 25 ml de células Sf9 a $\sim 10^6$ células ml⁻¹ con 1 ml de virus P0. La concentración y viabilidad de las células se midió utilizando un contador Innovatis CASY modelo TTC (Roche); para ello se diluyeron 12,5 μ l de cultivo en 5 ml de reactivo CASYton. Los cultivos se crecieron a 27 °C en un incubador sin control de humedad, con agitación orbital a 130 rpm, y con un tapón de goma que permitía la oxigenación del cultivo, hasta que la viabilidad de las células disminuyó por debajo del 90%. Los cultivos se centrifugaron y se guardaron los sobrenadantes, que se denominaron P1 (virus del pase 1).

8.5.- Expresión del EC2 de CD151 en células de insecto y purificación

Las condiciones de crecimiento y viabilidad de las células se optimizaron haciendo ensayos con diferentes concentraciones de virus (P1) en el medio de cultivo Sf900II SFM (GIBCO) suplementado con Pen Strep. En matraces con 25 ml de medio de cultivo se añadieron diferentes volúmenes de virus P1: 2.5 μ l, 3.7 μ l, 5.0 μ l, 7.5 μ l, 10.0 μ l, 12.5 μ l y 15.0 μ l. Con el fin de seleccionar las mejores condiciones para la expresión y purificación de proteína a gran escala. Cada 24 h se midió y comparó tanto el crecimiento como la viabilidad celular de cada uno de los cultivos.

La expresión y secreción de CD151-EC2 en los cultivos de células Sf9 se analizó enriqueciendo la muestra con una cromatografía de afinidad a pequeña escala. Típicamente 15 ml de medio de cultivo se incubaron 30 min a 4 °C con 250 μ l de resina Talon quelante de Co²⁺ equilibrada en 25 mM Tris-HCl (pH 8.0), 250 mM NaCl, 15 mM imidazol. Tras descartar el sobrenadante la resina se lavó dos veces con 2 ml de la misma

solución. Seguidamente las proteínas unidas se eluyeron con 100 μ l de 25 mM Tris-HCl (pH 8.0), 100 mM NaCl, 300 mM imidazol (incubación de 10 min). A 16 μ l de cada muestra se les añadió 4 μ l de tampón de muestra 4x y se analizaron mediante SDS-PAGE.

Para la producción a gran escala de CD151-EC2 se utilizaron 3 l de cultivo de células Sf9 a 1×10^6 células ml^{-1} en medio Sf900II SFM (GIBCO) suplementado con Pen Strep (en 6 matraces con 500 ml de medio de cultivo cada uno). El cultivo se infectó con un total de 1.2 ml de virus P1 (200 μ l de virus P1 para cada matraz de 500 ml de cultivo celular). Los cultivos se crecieron a 27 °C en agitación y se midió la viabilidad de las mismas. Cuando la viabilidad celular decreció por debajo del 90% los cultivos se centrifugaron a 2500 g 15 min a 4 °C para eliminar las células.

La proteína se purificó a partir del sobrenadante mediante cromatografía de afinidad empleando 2 ml de resina quelante Poros 20MC (Applied Biosystems) empaquetada manualmente en una columna XK16 (GE Healthcare). Tras cargar la muestra, la columna se lavó con 30 ml de 25mM Tris-HCl (pH 8.0), 250 mM NaCl y 15 mM imidazol. A continuación, la proteína se eluyó en 20mM Tris-HCl (pH 8.0), 100 mM NaCl, 30 mM imidazol. Las fracciones positivas se juntaron y se concentraron hasta ~2 ml empleando un filtro de centrifuga Amicon ultracel-3K (Millipore) por centrifugación a 3000g a 25 °C. Seguidamente la muestra se sometió a cromatografía de exclusión molecular utilizando una columna Superdex 75 (16 x 600 mm) (GE Healthcare) equilibrada en 20mM Tris-HCl (pH 7.5), 150 mM NaCl. Se empleó un flujo de 1 ml min^{-1} y se recogieron fracciones de 1 ml que se analizaron mediante SDS-PAGE. Por último, las fracciones que contenían CD151-EC2 se juntaron y se concentraron tal como se había hecho antes.

8.6.- Análisis de deglicosilación de CD151-EC2

La proteína CD151-EC2 expresada en células de insecto se trató con las endoglicosilasas: N-glicosilasa F (PNGasa F) y endoglicosilasa H (Endo Hf) (New Englan Biolabs).

Antes de realizar la digestión, se añadió 1 μ l de buffer de desnaturalización 10x (concentración final 1x: 0.5% SDS, 40 mM DTT) a 9 μ l de CD151-EC2 a 0.5 mg ml^{-1} en 20

mM Tris-HCl (pH 7.5), 150 mM NaCl, y se desnaturizó la proteína a 100 °C durante 10 min. Seguidamente se añadieron 2 µl de glicosidasa, 2 µl de 50 mM NaPi (pH 7.5) y se completó hasta un volumen final de 20 µl con 20 mM Tris-HCl (pH 7.5) 150 mM NaCl. La reacción se incubó a 37 °C durante 1 h. Los productos de reacción se analizaron mediante electroforesis en geles de poliacrilamida-SDS (15% acrilamida).

9.- Métodos de caracterización de proteínas

9.1.- Cuantificación de la concentración de proteínas

La concentración de las muestras de proteína se determinó por medidas espectrofotométricas. Los coeficientes de extinción molar ($\epsilon_{280\text{nm}}$) y másico ($E_{280\text{nm}}$) a 280 nm de cada proteína se calcularon a partir de su secuencia de aminoácidos (Pace et al., 1995) empleando la herramienta ProtParam (<http://web.expasy.org/protparam/>) del servidor ExPASy. En la tabla M11 se muestran los valores de $\epsilon_{280\text{nm}}$ y $E_{280\text{nm}}$ de las proteínas utilizadas, así como su masa molecular. Las mediciones de absorbancia se realizaron en un espectrofotómetro U-2001 UV/Visible (HITACHI) empleando cubetas de cuarzo de 10 mm de paso óptico (Hellma).

Tabla M11-. Coeficientes de extinción y masa molecular (MM) de los fragmentos de integrina $\beta 4$ y tetraspanina CD151.

PROTEÍNA	FRAGMENTO		$\epsilon_{280\text{nm}}$ ($\text{M}^{-1} \text{cm}^{-1}$) ^b	$E_{280\text{nm}}$ ($(\text{mg ml}^{-1})^{-1} \text{cm}^{-1}$)	MM (Da)
	NOMBRE	REGIÓN ^a			
Integrina $\beta 4$	Calx- β	989-1107	8480	0.608	13946.6
	FnIII-3	1457-1548	15470	1.429	10825.2
	FnIII-4	1572-1666	8480	0.801	10589.8
	CS-FnIII-3,4	1436-1666	25440	0.994	25605.7
	FnIII-3,4	1457-1666	24075	1.031	23360.2
CD151	EC2	112-219	18825	1.475	12766.1

^a Además de las regiones indicadas las proteínas contienen el tetrapéptido GSHM en el extremo amino.

^b Para los mutantes con sustituciones puntuales de amino ácidos se emplearon los mismos valores de coeficiente de extinción molar de las proteínas silvestres ya que no se alteró el número de residuos aromáticos.

9.2.- Ensayo de estabilidad térmica basado en fluorescencia

Los estudios de estabilidad térmica de proteínas se realizaron mediante la técnica Termofluor. Brevemente, en estos experimentos una solución de proteína que contiene el fluoróforo Sypro Orange (Molecular Probes, Invitrogen) se calienta a velocidad constante mientras se mide la intensidad de fluorescencia emitida por la sonda. En el estado nativo las regiones hidrofóbicas están mayoritariamente ocultas en el interior de las proteínas. Sin embargo, cuando las proteínas se desnaturalizan por el aumento de la temperatura, las zonas hidrofóbicas quedan expuestas al solvente y la sonda Sypro Orange se une a ellas, lo cual da lugar a un incremento de la señal de fluorescencia (Pantoliano et al., 2001). Estos experimentos se realizaron empleando placas Thermo-Fast 96 Non-Skirted (ABgene, Thermo Scientific). Se utilizó un volumen de reacción de 25 μl típicamente con 0.25 mg ml^{-1} de proteína y una concentración final del Sypro Orange 10x (unidades arbitrarias del fabricante) en 50 mM HEPES (pH 7.5), 150 mM NaCl, (en algunos casos se complementó con 5 mM EGTA o 10 mM CaCl_2). Los experimentos se realizaron en un termociclador de PCR a tiempo real IQ5 (BioRad). Las muestras se calentaron de 20 a 85 $^{\circ}\text{C}$ con una velocidad de calentamiento de 1 $^{\circ}\text{C min}^{-1}$. Las medidas de fluorescencia se realizaron empleando un filtro de excitación de 480 nm (ancho de banda de 20 nm) y uno de emisión a 530 nm (ancho de banda de 30 nm), y se recogieron datos cada 0.5 $^{\circ}\text{C}$. La temperatura de fusión, T_m , correspondiente al punto medio de la transición, se calculó como el máximo de la derivada (dF/dT) de la señal de fluorescencia en función de la temperatura.

9.3.- Análisis de unión de calcio al dominio Calx- β mediante diálisis de equilibrio

Muestras de 2.5 ml del dominio Calx- β de β_4 a 200 μM se dializaron a temperatura ambiente frente a cinco cambios de 400 ml de 20 mM Tris-HCl (pH 7.5), 150 mM NaCl y concentraciones variables de CaCl_2 . La concentración de proteína se determinó espectrofotométricamente. La concentración de Ca^{2+} en el tampón de diálisis y en las muestras de proteína se determinó por espectroscopia de emisión óptica de plasma acoplado inductivamente (ICP-OES). Las medidas se realizaron por el Servicio General de Análisis Químico Aplicado de la Facultad de Ciencias Químicas de la Universidad de

Salamanca. La concentración de Ca^{2+} unido al dominio Calx- β se determinó por diferencia entre la concentración de Ca^{2+} en el tampón (Ca^{2+} libre) y en la muestra de proteína (Ca^{2+} libre y unido). El número de iones Ca^{2+} unidos por dominio Calx- β se calculó como el cociente entre las concentraciones de Ca^{2+} unido y de proteína.

9.4.- Cromatografía de exclusión molecular

El CD151-EC2 se analizó en una columna Superdex 200 10/300 GL (10 x 300 mm) (GE Healthcare) equilibrada en 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, empleando un sistema de cromatografía ÄKTA Purifier (GE Healthcare). El flujo fue de 0.5 ml min^{-1} . Con el fin de determinar la masa molecular aparente de CD151-EC2, la columna se calibró empleando un juego de proteínas patrón que incluye tiroglobulina (670 kDa), γ -globulina (158 kDa), ovalbúmina (44 kDa) y mioglobina (17 kDa) (Bio-Rad).

9.5.- Dicroísmo circular

Las medidas de dicroísmo circular en la región del ultravioleta lejano se realizaron en el Centro de Investigaciones Biológicas (CSIC, Madrid) empleando un espectropolarímetro Jasco J-720. Se midieron espectros de dicroísmo circular de CD151-EC2 a 2 mg ml^{-1} en 20 mM NaPi (pH 7.5), 150 mM NaCl utilizando una cubeta de cuarzo de 0.1 mm de paso óptico (Hellma). Los espectros se obtuvieron como promedio de 4 medidas consecutivas entre 197 y 260 nm. El ancho de banda fue de 1 nm, la velocidad de barrido 20 nm min^{-1} y el tiempo de respuesta de 2 s. A partir de los espectros de dicroísmo circular se estimó el contenido de estructura secundaria empleando el servidor K2D2 (Perez-Iratxeta and Andrade-Navarro, 2008).

RESULTADOS

1.- Caracterización del dominio Calx- β de la integrina $\beta 4$

1.1.- Estructura cristalográfica del dominio Calx- β de $\beta 4$

Los límites del dominio Calx- β de $\beta 4$ se definieron en base a las estructuras de los dominios de unión a Ca^{2+} CBD1 y CBD2 del intercambiador de $\text{Na}^+/\text{Ca}^{2+}$ NCX1, resueltas mediante RMN (Hilge et al., 2006). El dominio Calx- β de $\beta 4$ (989-1107) se expresó insoluble en *E. Coli* como proteína de fusión con una cola de poli-His. Se purificó en condiciones desnaturizantes y se replegó por dilución rápida. La muestra así producida fue apta para análisis cristalográfico.

La estructura del dominio Calx- β se resolvió por reemplazo molecular empleando un modelo basado en la estructura del primer dominio de unión de Ca^{2+} del intercambiador NCX1 (código PDB: 2DPK) (Nicoll et al., 2006). La estructura se refinó frente a datos a 1.48 Å de resolución (Tabla R1). La unidad asimétrica de los cristales contiene dos moléculas de $\beta 4$ relacionadas por un eje binario no cristalográfico. El dímero observado en el cristal no parece tener relevancia funcional ya que el dominio Calx- β es monomérico en solución, como se determinó por cromatografía de exclusión molecular (resultados no mostrados). Las estructuras de las dos moléculas, que se refinaron de forma independiente, son prácticamente idénticas. Después de superponerlas, la r.m.s.d. en la posición de 106 C^α entre las dos moléculas es de 0.39 Å (la estimación del error en las coordenadas de la estructura es 0.17 Å). Una de las dos moléculas incluye siete residuos (1069-1073 y 1106-1107) que en la otra molécula se encuentran desordenados.

El dominio Calx- β está formado por 7 hebras β (A-G) (Figura R1). La hebra A presenta una discontinuidad en el patrón de enlaces de hidrógeno en torno a Q996 y P997, que la divide en las hebras A1 y A2. De forma similar, la hebra G presenta una distorsión alrededor de P1095, que forma un enlace peptídico en *cis* con Q1094, que da lugar a las hebras G1 y G2. Las hebras se organizan en dos láminas β , una formada por las hebras G1-A1-B-E y la otra por las hebras D-C-F-G2-A2, que en conjunto adoptan una estructura de sándwich- β característica del plegamiento de la superfamilia de las inmunoglobulinas. Los extremos amino y carboxilo se encuentran en lados opuestos del

eje longitudinal de la estructura, que tiene forma de cilindro de ~50 Å de altura y ~24 Å de diámetro.

Tabla R1.- Resumen del análisis cristalográfico del dominio Calx-β de la integrina β4.

	Estructura		
	Nativa	Cristalizada en presencia de 2 mM Ca ²⁺	Incubada en presencia de 10 mM Ca ²⁺
Colección de datos			
Grupo espacial	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Dimensiones de celda: <i>a, b, c</i> (Å)	50.3, 51.7, 87.5	50.3, 51.7, 87.4	50.3, 51.7, 87.4
Longitud de onda (Å)	1.5418	1.5418	1.5418
Resolución (Å)	1.48 (1.54 – 1.48) ^a	1.61 (1.69 – 1.61) ^a	1.40 (1.50 – 1.40) ^a
Reflexiones únicas	38089 (3644) ^a	30229 (4045) ^a	44715 (7607) ^a
Redundancia	13.4 (11.3) ^a	9.2 (8.3) ^a	12.7 (10.7) ^a
Compleitud (%)	98.3 (85.4) ^a	98.1 (88.1) ^a	98.1 (91.1) ^a
<i>R</i> meas ^b (%)	5.3 (31.2) ^a	9.3 (33.8) ^a	8.3 (24.5) ^a
< / σ >	29.5 (8.6) ^a	17.1 (6.6) ^a	20.5 (9.9) ^a
Estadísticas de refinado			
Rango de resolución (Å)	26 – 1.48 (1.52 – 1.48) ^a	20 – 1.60 (1.66 – 1.60) ^a	22 – 1.40 (1.44 – 1.40) ^a
Reflexiones únicas, <i>work / free</i>	36177 / 1911	28711 / 1515	42471 / 2229
<i>R</i> work (%)	15.5 (20.4) ^a	17.2 (21.7) ^a	18.2 (22.3) ^a
<i>R</i> free (%)	18.9 (24.8) ^a	21.3 (26.1) ^a	21.2 (23.9) ^a
Nº de átomos			
Proteína	3783 [1903] ^c	3742 [1878] ^c	3803 [1913] ^c
Solvente	334	296	308
Valor B promedio (Å ²)			
Gráfico de Wilson	14.8	11.4	9.9
Proteína ^d	19.5 / 20.7	17.4 / 18.6	18.6 / 18.8
Solvente	29.0	25.8	26.5
Desviaciones cuadráticas medias			
Longitudes de enlace (Å)	0.008	0.010	0.007
Ángulos de enlace (°)	0.960	1.273	0.909
Factores B cadena principal (Å ²)	1.67	2.14	1.54
Factores B cadenas laterales (Å ²)	4.47	4.49	4.01
Regiones mapa Ramachandran ^e			
favorecidas	246 (98.8%)	211 (97.7%)	214 (98.6%)
adicionalmente permitidas	3 (1.2%)	5 (2.3%)	3 (1.4%)
no permitidas	0	0	0
Código PDB	3FQ4	3H6A	3FSO

^a Valores para la zona de alta resolución

^b Factor R independiente de la multiplicidad (Diederichs and Karplus, 1997)

^c El número de átomos excluyendo los átomos de hidrógeno se muestra entre corchetes

^d Valores dados para cada cadena de la proteína

^e Valores calculados con MOLPROBITY (Lovell et al., 2003)

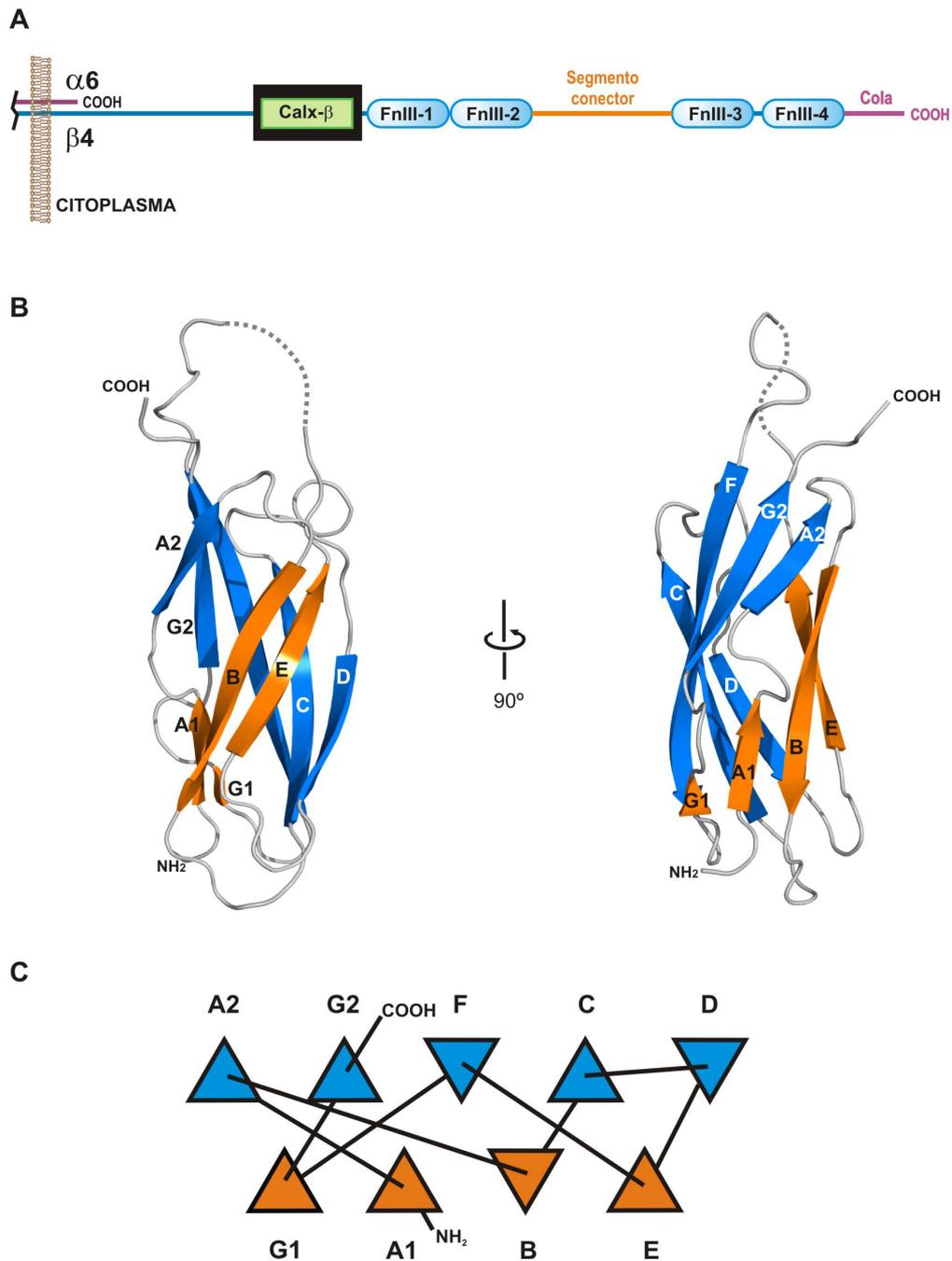


Figura R1.- Estructura cristalográfica del dominio Calx- β de β 4. (A) Representación esquemática de la región citoplasmática de β 4, el dominio Calx- β se resalta sobre fondo negro. (B) Dos vistas ortogonales de la representación en cintas de la estructura cristalográfica del Calx- β de β 4. Las hebras que forman cada una de las láminas β del dominio se muestran en el mismo color. La parte del lazo EF que está desordenada en la estructura se representa con trazo discontinuo. (C) Diagrama de topología del dominio Calx- β de β 4. Cada triángulo representa una hebra β que están coloreadas como en el panel B.

1.2.- Comparación del dominio Calx- β de $\beta 4$ con los dominios de unión a Ca^{2+} del NCX1

La superposición del dominio Calx- β de $\beta 4$ con las estructuras de los dominios CBD1 y CBD2 del intercambiador NCX1, obtenidas mediante NMR y cristalografía, revela un alto grado de conservación de las hebras que forman la estructura de sándwich- β (Figura R2). La r.m.s.d. para el conjunto de 100 C^α comunes usados en la superposición por parejas de los CBDs sobre el dominio Calx- β de $\beta 4$ varía entre 0.85 y 1.65 Å (Tabla R2). La similitud estructural está correlacionada con la conservación a nivel de secuencia. El dominio Calx- β de $\beta 4$ tiene un 33 % y un 27 % de identidad de secuencia con los dominios CBD1 y CBD2 de NCX1, respectivamente. La secuencia de $\beta 4$ contiene 18 residuos que están conservados en CBD1 y CBD2 de NCX1 (Figura R2C), los cuales incluyen: 11 residuos ocluidos en el interior del sándwich- β (F994, R1014, V1024, T1028, A1033, D1038, F1048, V1059, F1079, L1092 y I1102), tres glicinas que participan en giros del esqueleto polipeptídico (G1031 y G1093) o en una ligera distorsión de la hebra- β D (G1044) y tres residuos expuestos en la superficie de la estructura (T1032, E1052 y K1055) cuyas cadenas laterales participan en interacciones intramoleculares. Por tanto, este núcleo de residuos conservados contribuye a la estabilidad estructural del dominio Calx- β . Además, $\beta 4$ tiene otros 21 y 14 residuos conservados en CBD1 y CBD2, respectivamente. Por lo tanto, las similitudes no redundantes del dominio Calx- β de $\beta 4$ con el CBD1 y el CBD2 de NCX1 sugieren que el dominio de $\beta 4$ divergió durante la evolución antes de la duplicación de los CBDs en el NCX.

Tabla R2.- Comparación estructural del Calx- β de $\beta 4$ con los CBDs del intercambiador de $\text{Na}^+/\text{Ca}^{2+}$ NCX1.

Dominio	Método	Código PDB	r.m.s.d. ^a de C^α con el Calx- β de $\beta 4$ (Å)	
			Molécula A ^b	Molécula B ^b
CBD1, Ca^{2+} -unido	RMN	2fws	1.65	1.64
CBD1, Ca^{2+} -unido	Cristalografía	2dpk	0.88	0.85
CBD2, Ca^{2+} -unido	RMN	2fwu	1.33	1.45
CBD2, Ca^{2+} -unido	Cristalografía	2qvm	0.91	1.05
CBD2, Ca^{2+} -libre	Cristalografía	2qvk	1.01	1.10

^aUn conjunto de 100 átomos C^α equivalentes se usaron en todas las comparaciones.

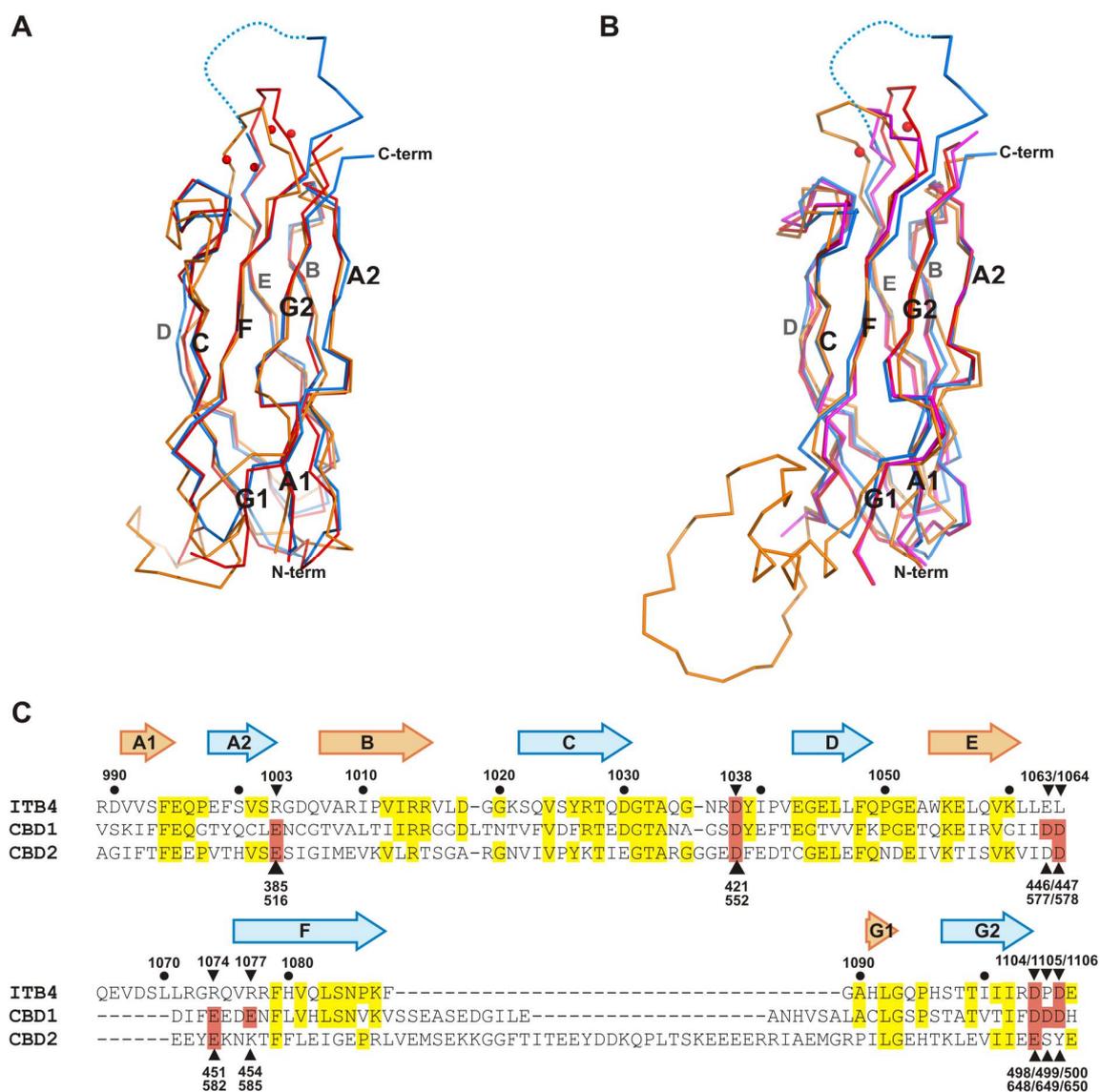
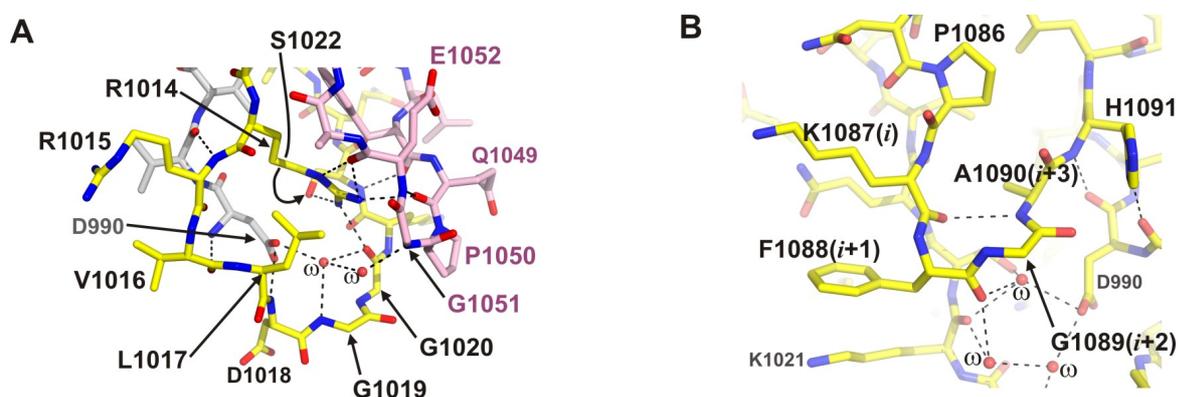


Figura R2.- Comparación del dominio Calx-β de β4 con los dominios de unión a Ca²⁺ del intercambiador de Na⁺/Ca²⁺ NCX1. (A) Superposición del trazado de C^α del dominio Calx-β de β4 (azul) y las estructuras determinadas mediante RMN (naranja; código PDB: 2FWS) y cristalografía (rojo; código PDB: 2DPK) del CBD1 de NCX1. (B) Superposición del trazado de C^α del dominio Calx-β de β4 (azul) y las estructuras por RMN (naranja; código PDB: 2FWU) y cristalográfica (rojo; código PDB: 2QVM) del CBD2 de NCX1 en presencia de Ca²⁺ y en ausencia de Ca²⁺ (morado; código PDB: 2QVK). Los iones Ca²⁺ presentes en las estructuras cristalográficas del CBD1 y del CBD2 se representan como esferas rojas. Todas las estructuras de A y B se superpusieron simultáneamente usando un método de máxima probabilidad. (C) Alineamiento de secuencias basado en el alineamiento de estructuras del dominio Calx-β de β4 (ITB4) y los dominios CBD1 y CBD2 de NCX1. Los elementos de estructura secundaria en la estructura de β4 se indican encima de las secuencias, utilizando los mismos colores que en la figura R1. Los residuos implicados en la coordinación de los iones Ca²⁺ en las estructuras de CBD1 y CBD2 están destacados en cajas rojas y sus números se muestran debajo de las secuencias; los números de los residuos equivalentes de β4 se muestran encima de las secuencias. Los residuos conservados en β4 y CBD1 o CBD2 se resaltan en cajas amarillas.

Las principales diferencias entre el dominio Calx- β de β 4 y los dominios de NCX1 se localizan en los lazos BC, EF y FG. En los tres dominios el lazo BC está estabilizado en sus extremos por un puente de hidrógeno entre el grupo guanidinio de una Arg conservada (R1014 en β 4, R396 en CBD1 y R527 en CBD2) y el carbonilo de un residuo (G1020 en β 4) al final del lazo (Figura R3-A). A pesar de esta característica en común, el lazo BC adopta conformaciones diferentes en estos tres dominios Calx- β . La estructura del lazo BC está condicionada en parte por las interacciones que establece con el extremo amino, que en β 4 incluye un enlace de hidrógeno entre el nitrógeno de D1018 (lazo BC) y el grupo carboxilo de D990 (extremo amino). En β 4 el lazo FG es mucho más corto que en CBD1 y CBD2 de NCX1, en los que este lazo contiene 20 y 39 residuos, respectivamente. El lazo FG de β 4 forma un giro β de tipo I' (Figura R3-B). La presencia de G1089, que ocupa la región γ_L del gráfico de Ramachandran ($\gamma=94^\circ$, $\psi=1^\circ$), en la posición $i+2$ favorece la formación de este giro de tipo I' (Hutchinson and Thornton, 1994). En el lado opuesto del sándwich- β , el lazo EF es seis residuos más largo en β 4 que en los CBDs de NCX1. El lazo EF no participa en contactos con otras moléculas de la celda del cristal de β 4 y está parcialmente desordenado, sugiriendo que tiene una gran variabilidad conformacional.



1.3.- Análisis del pseudo-sitio de unión a Ca^{2+} del dominio Calx- β de $\beta 4$

Los sitios de unión a Ca^{2+} de CBD1 y CBD2 en NCX1 están formados por residuos ácidos localizados en los lazos AB, CD y EF y en la región C-terminal tras la hebra G (Figura R4). D1038 y D1104 en $\beta 4$ ocupan posiciones equivalentes a D421 y D498 en CBD1, y D552 y E648 en CBD2, los cuales participan en la coordinación de iones Ca^{2+} . Por el contrario, otros residuos ácidos que contribuyen al sitio de unión de Ca^{2+} en el CBD1 (E385, D447 y E451) y CBD2 (E516, D578 y E582) de NCX1 están sustituidos en $\beta 4$ por residuos básicos o hidrofóbicos: R1003, L1064 y R1074.

La cadena lateral de R1003 se localiza en el centro de esta región y su grupo guanidinio ocupa posiciones similares a las de los iones Ca^{2+} en los CBDs de NCX1. La cadena lateral de R1003 adopta conformaciones ligeramente diferentes en las dos moléculas Calx- β presentes en la unidad asimétrica. En una de las moléculas el grupo guanidinio de R1003 forma un puente salino con D1104 y un puente de hidrógeno con el carbonilo de I1002. En la otra molécula R1003 forma un puente salino con D1038 y un enlace de hidrógeno con el carbonilo de L1062. D1038 forma a su vez un puente salino con R1077, que es similar al formado entre D552 y K585 en la estructura libre de Ca^{2+} del CBD2 de NCX1 (Figura R4-D). La cadena lateral de R1074, que en el alineamiento de secuencia equivale a E451 en CBD1 y a E582 en CBD2, esta orientada hacia afuera del pseudo-sitio de unión a Ca^{2+} y no participa en ningún contacto intramolecular. En resumen, los sitios de unión a Ca^{2+} de otros dominios Calx- β no están conservados en $\beta 4$, donde existen interacciones de compensación de carga que parecen favorecen la estabilidad en ausencia de cationes.

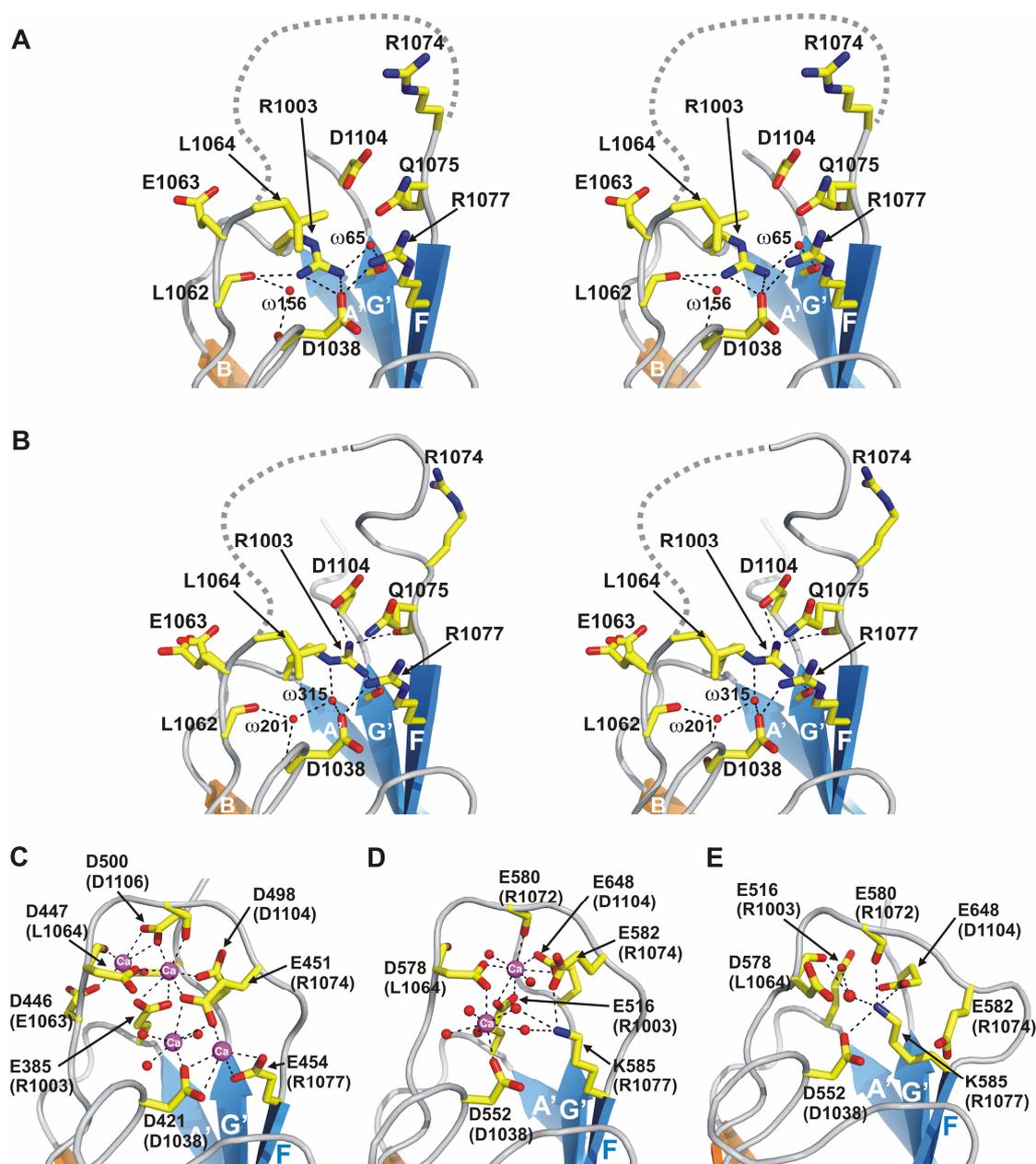


Figura R4.- Estructura de la región de $\beta 4$ equivalente a los sitios de unión de Ca^{2+} en NCX1. (A) Representación estereoscópica de los lazos A, B y EF y de residuos próximos de $\beta 4$ en la molécula A de la unidad asimétrica de los cristales. Se muestran las cadenas laterales de los residuos que ocupan posiciones equivalentes a los residuos de NCX1 implicados en la unión de Ca^{2+} . R1003 se modeló con ocupancia parcial (refinada a 0.91) y se asignó a una conformación alternativa distinta a la de la molécula de agua 156, la cual se encuentra a 1.8 Å del átomo NH1 de R1003. (B) Representación en estéreo de la misma región de la molécula B en la unidad asimétrica. En esta molécula R1003 forma un puente salino con D1104. Las estructuras cristalinas de los sitios de unión a Ca^{2+} de CBD1 y CBD2 de NCX1 se muestran en (C), (D) y (E). Éstos incluyen: (C) el CBD1 con Ca^{2+} unido; (D) el CBD2 con Ca^{2+} unido y (E) el CBD2 en ausencia de Ca^{2+} . Los residuos de $\beta 4$ equivalentes a aquellos de NCX1 se indican entre paréntesis en (C), (D) y (E). Los contactos polares se representan con líneas discontinuas.

La primera estructura cristalográfica del dominio Calx- β de $\beta 4$ se obtuvo en ausencia de Ca^{2+} . Por tanto, a pesar de la falta de conservación del sitio de unión a Ca^{2+} , no podíamos excluir que $\beta 4$ pueda unir Ca^{2+} mediante mecanismos no canónicos. Por ejemplo, el lazo EF contiene dos residuos ácidos, E1066 y D1068, que podrían participar en la unión a Ca^{2+} . Con el fin de estudiar esta posibilidad, determinamos dos estructuras adicionales del dominio Calx- β a 1.61 y 1.40 Å de resolución usando cristales crecidos en presencia de 2 mM Ca^{2+} y cristales crecidos sin calcio pero incubados en presencia de 10 mM de CaCl_2 , respectivamente. Sin embargo, ni los cristales obtenidos en presencia de Ca^{2+} ni los incubados con Ca^{2+} contienen cationes ordenados. No se apreciaron diferencias significativas en el pseudo-sitio de unión a Ca^{2+} , incluyendo las moléculas de solvente ordenadas, entre la estructura nativa, la estructura cristalizada en presencia de Ca^{2+} y la estructura incubada con CaCl_2 (Figura R5). Es poco probable que la ausencia de calcio ordenado en la estructura se deba al bloqueo del sitio de unión a calcio durante la cristalización. Esto se debe a tres razones. (i) La región de pseudo-sitio de unión a Ca^{2+} de $\beta 4$ es accesible desde los canales del solvente del cristal. (ii) El pseudo-sitio de unión no participa en contactos con otras moléculas del cristal. (iii) La libertad conformacional observada en el lazo EF sugiere que éste podría adaptarse para unir cationes sin perturbar el empaquetamiento del cristal. En resumen, la ausencia de Ca^{2+} ordenado en presencia de 2 o 10 mM de CaCl_2 sugiere que el dominio Calx- β de $\beta 4$ carece de un sitio funcional de unión a Ca^{2+} .

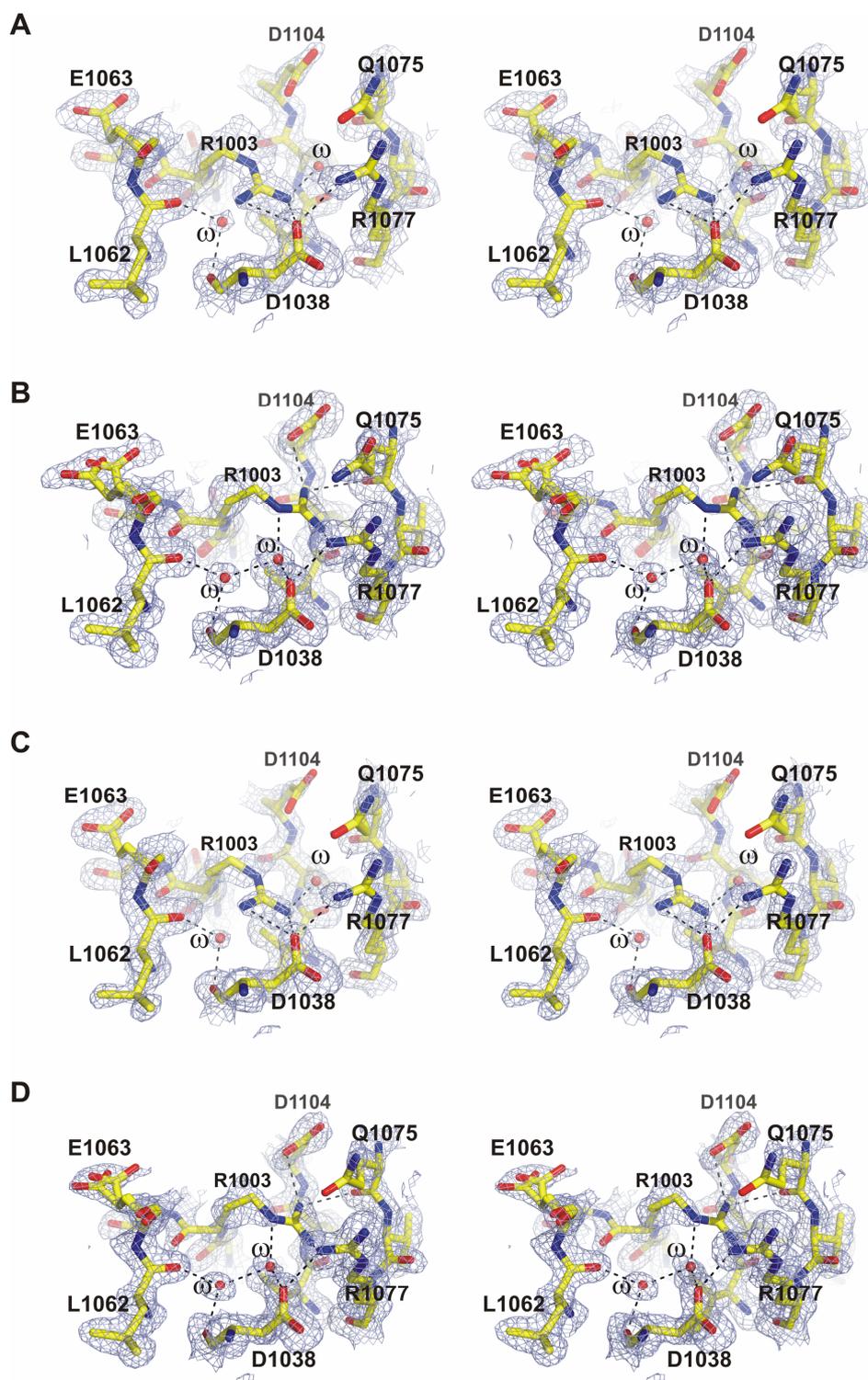


Figura R5.- Vistas estereoscópicas de los mapas OMIT de recocido simulado en los pseudositos de unión a Ca^{2+} de estructuras de $\beta 4$. (A) Estructura de la molécula A de la unidad asimétrica de los cristales nativos. (B) Estructura de la molécula B de la unidad asimétrica de los cristales nativos. (C) y (D) muestran la estructura de la molécula A y de la molécula B en presencia de 2 mM Ca^{2+} , respectivamente. Cada mapa ($2m\text{Fobs}-\text{DFcalc}$, contorno a 1σ) se calculó después de llevar a cabo una ronda de refinado usando recocido simulado (temperatura inicial de 5000 K) de modelos en los cuales las regiones mostradas en la figura se eliminaron. No se observaron diferencias significativas en la cadena de la proteína y del solvente entre moléculas equivalentes de las estructuras del nativo y co-cristalizada en presencia de Ca^{2+} .

Con el fin de comparar los resultados cristalográficos con datos obtenidos en solución, se analizó el efecto del Ca^{2+} en la estabilidad térmica del dominio Calx- β de $\beta 4$ empleando un ensayo basado en fluorescencia. El perfil de desnaturalización térmica del Calx- β de $\beta 4$ en ausencia de cationes exógenos se caracteriza por una transición a una T_m de 62 °C (Figura R6). Después del proceso de desnaturalización inicial, las muestras se enfriaron y se sometieron a un segundo ciclo de calentamiento. La T_m no varía entre la primera y la segunda curva de fusión, sugiriendo que la desnaturalización térmica de $\beta 4$ es reversible y se puede considerar como un proceso de equilibrio.

No se observaron cambios significativos en la T_m de $\beta 4$ en presencia de 10 mM CaCl_2 . Esto sugiere que el Calx- β no une Ca^{2+} ya que la unión de cationes al estado nativo lo estabilizaría elevando la T_m , tal como se observa en la estabilización térmica inducida por Ca^{2+} de proteínas que si lo unen, como el dominio C2 de la proteína quinasa C α , β II (Torrecillas et al., 2004) y los dominios extracelulares 1 y 2 de E-caderina (Volkov and Svergun, 2003). La T_m del Calx- β tampoco varió en experimentos control en presencia de 5 mM EDTA, por lo que se puede excluir que la estabilización se deba a Ca^{2+} residual unido durante la expresión o purificación.

Así mismo se analizó de forma directa la posible unión de Ca^{2+} al dominio Calx- β de $\beta 4$ mediante diálisis de equilibrio combinada con la determinación de Ca^{2+} por ICP-OES. El número de iones Ca^{2+} unidos por molécula de Calx- β a concentraciones de Ca^{2+} libre de 155, 390 y 816 μM fueron 0.03, 0.02 y 0.13 respectivamente. Por tanto, la ausencia de Ca^{2+} ordenado en presencia de 2 mM CaCl_2 es debida a la carencia de un sitio funcional de unión a Ca^{2+} . En resumen, los datos cristalográficos, de desnaturalización térmica y de diálisis de equilibrio apoyan que el dominio Calx- β de $\beta 4$ no une Ca^{2+} .

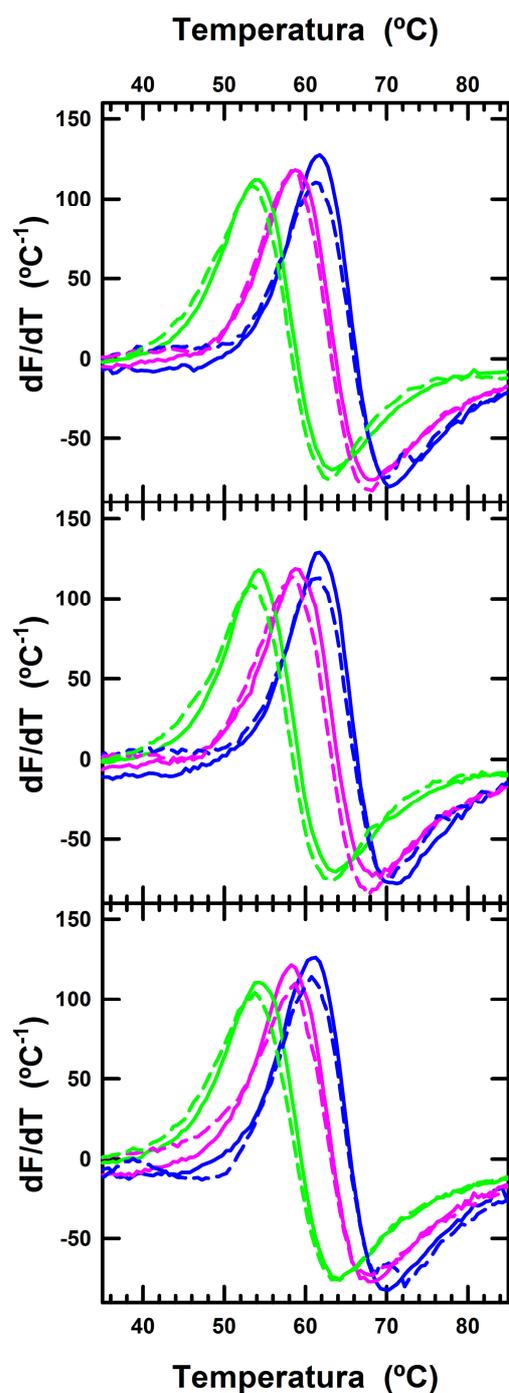


Figura R6.- Desnaturalización térmica del dominio Calx- β de β 4. (A) Curvas representativas de la desnaturalización del Calx- β silvestre (azul) y de los mutantes puntuales R1003A (rosa) y R1003E (verde) en 50 mM HEPES (pH 7.5), 150 mM NaCl. Las líneas continuas corresponden a la primera rampa de calentamiento, mientras que las curvas de trazo discontinuo corresponden a la segunda rampa de desnaturalización efectuada tras enfriar las muestras. (B) Desnaturalización en 50 mM HEPES (pH 7.5), 150 mM NaCl, 5 mM EGTA y (C) en 50 mM HEPES (pH 7.5), 150 mM NaCl, 10 mM CaCl₂.

Tabla R3.- Temperaturas de desnaturalización del dominio Calx- β y de mutantes puntuales.

Solución	T _m (°C)		
	WT ^a	R1003A	R1003E
50 mM HEPES (pH 7.5), 150 mM NaCl.	62	59	54
50 mM HEPES (pH 7.5), 150 mM NaCl, 5 mM EGTA	62	59	54
50 mM HEPES (pH 7.5), 150 mM NaCl, 10 mM CaCl ₂	62	59	54

^a secuencia silvestre

1.4.- Papel de R1003 en la estabilidad de dominio Calx- β de β 4

La estructura sugiere que R1003 es un residuo clave para mantener la estabilidad del pseudo-sitio de unión del Calx- β de β 4. Con el fin de evaluar la contribución de R1003 se comparó la estabilidad térmica de la proteína silvestre y de dos mutantes de β 4 en los que se sustituyó R1003 por Ala y Glu (Figura R6). La desnaturalización del mutante R1003A se produce con una T_m de 59 °C, que es 3 °C menor que la T_m de la proteína silvestre (Tabla R3). Como los CBDs de NCX1 tienen un Glu en la posición equivalente a R1003, se analizó el efecto de sustituir R1003 por un residuo ácido. La desnaturalización térmica del mutante R1003E presenta un valor de T_m de 54 °C, que es 8 °C menor que la proteína silvestre y 5 °C más baja que en el mutante R1003A. La T_m de R1003E no se modifica por la presencia de 10 mM de CaCl_2 , sugiriendo que esta sustitución simple no es suficiente para recrear el sitio de unión a Ca^{2+} debido a la falta de conservación de secuencia en el lazo EF entre β 4 y los CBDs de NCX1. En resumen, la presencia de un residuo básico en la posición 1003 contribuye a la estabilidad del dominio Calx- β de β 4.

2.- Estructura de la región FnIII-3,4 de la integrina β 4 y su interacción con BPAG1e

2.1.- Estructuras cristalográficas de los dominios FnIII-3 y FnIII-4 aislados.

Con el fin de elucidar la estructura de la región FnIII-3,4 de la integrina β 4 se generaron cuatro construcciones de esta región (Figura R7). El fragmento 1457-1666 contiene el tercer y cuarto dominios FnIII (FnIII-3,4). El fragmento 1436-1666 además de la región FnIII-3,4 contiene la parte final del segmento conector (CS) que precede al FnIII-3. El fragmento 1457-1548 corresponde al FnIII-3 aislado y el fragmento 1572-1666 al FnIII-4. La numeración de la secuencia corresponde a la isoforma beta-4A (Uniprot 16144-2) (Suzuki and Naitoh, 1990). Estas cuatro proteínas se expresaron solubles en *E. coli* como proteínas de fusión con una cola de poli-His en el extremo amino y se purificaron mediante cromatografía de afinidad. Las cuatro construcciones se sometieron a ensayos de cristalización obteniéndose cristales del FnIII-3 (1457-1548) y del FnIII-4 (1572-1666), los cuales permitieron elucidar las estructuras de estos dos dominios.

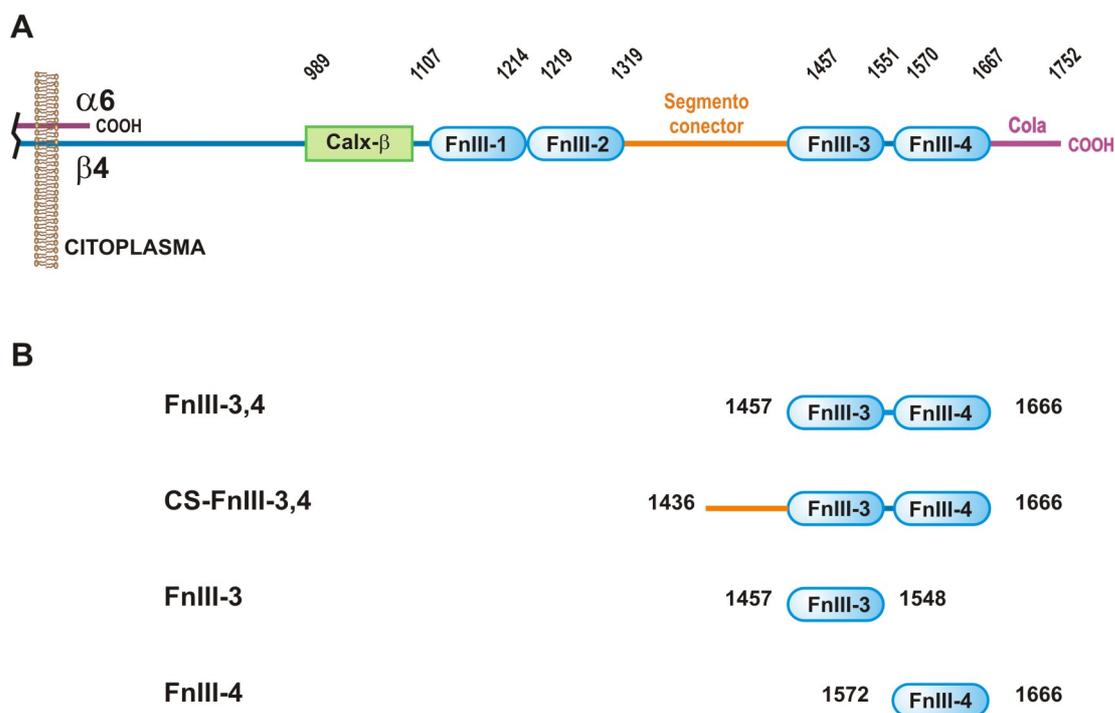


Figura R7.- Fragmentos de la interina $\beta 4$ empleados para el estudio de la región FnIII-3,4. (A) Organización de la región citoplasmática de $\beta 4$ y los límites de cada dominio. (B) Construcciones utilizadas para el estudio de $\beta 4$ y los límites de secuencia de cada una de ellas.

2.1.1.- Estructura cristalográfica del dominio FnIII-3

Los cristales del dominio FnIII-3 (residuos 1457-1548) difractaron a una resolución máxima de 1.60 Å de resolución. Los cristales del FnIII-3 pertenecen al grupo espacial $I2_12_12_1$ siendo las dimensiones de celda unidad: $a=66.7$ Å, $b=68.3$ Å y $c=88.4$ Å y $\alpha=\beta=\gamma=90^\circ$ (Tabla R4), que corresponden con un volumen de 403150 Å³. Considerando la masa molecular del FnIII-3 (10825.2 Da) y el número de operaciones de simetría correspondientes a este grupo espacial ($Z=8$), era razonable que la unidad asimétrica contenga dos moléculas de FnIII-3 que corresponden a un contenido de disolvente del 47 % y un valor del coeficiente de Matthews (V_M) de 2.33 Å³ Da⁻¹.

Tabla R4.- Resumen del análisis cristalográfico del dominio FnIII-3 (1457-1548).

Colección de datos	
Grupo espacial	I2 ₁ 2 ₁ 2 ₁
Dimensiones de celda: <i>a</i> , <i>b</i> , <i>c</i> (Å)	66.7, 68.3, 88.4
Longitud de onda (Å)	1.5418
Resolución (Å)	1.61 (1.69 – 1.61) ^a
Reflexiones únicas	26721 (3672) ^a
Redundancia	9.0 (8.1) ^a
Compleitud (%)	98.6 (91.2) ^a
R <i>meas</i> ^b (%)	3.9 (31.3) ^a
<I / σI>	41.7 (7.7) ^a
Estadísticas de refinado	
Rango de resolución (Å)	33.4 – 1.61 (1.66 – 1.61) ^a
Reflexiones únicas, <i>work</i> / <i>free</i>	25369 / 1340
R <i>work</i> (%)	15.0 (17.7) ^a
R <i>free</i> (%)	18.3 (21.7) ^a
Nº de átomos	
Proteína	3296 [1666] ^c
Solvente	221
Sulfatos	10
PEGs	38
Valor B promedio (Å ²)	
Gráfico de Wilson	14.2
Proteína ^d	13.7 / 15.1
Solvente	25.2
Sulfatos	17.3
PEGs	26.7
Desviaciones cuadráticas medias	
Longitudes de enlace (Å)	0.010
Ángulos de enlace (°)	1.333
Factores B cadena principal (Å ²)	0.98
Factores B cadenas laterales (Å ²)	3.73
Regiones mapa Ramachandran ^e	
favorecidas	241 (98.9 %)
adicionalmente permitidas	3 (1.1 %)
no permitidas	0

^a Los valores para la zona de alta resolución se indican entre paréntesis.

^b Factor R independiente de la multiplicidad (Diederichs and Karplus, 1997).

^c El número de átomos excluyendo los átomos de hidrógeno se muestra entre corchetes.

^d Valores dados para cada cadena de la proteína.

^e Valores calculados con MOLPROBITY (Lovell et al., 2003).

La estructura se faseó mediante reemplazo molecular. Se utilizó un modelo de homología creado a partir de la estructura cristalográfica del FnIII-2 de β4 (código PDB: 1QG3) (de Pereda et al., 1999). Se localizaron dos copias del modelo de búsqueda en el grupo I2₁2₁2₁, y no se obtuvo ninguna solución aceptable de la función de translación en el grupo espacial I222. La estructura se refinó frente a datos de difracción a 1.61 Å de

resolución. Las dos moléculas del FnIII-3 se refinaron de forma independiente. Durante el refinado se utilizaron dos grupos TLS que correspondían a cada una de las moléculas; la subdivisión en un mayor número de grupos TLS no mejoró el refinado. Dos densidades electrónicas con forma tetraédrica cuyos centros corresponden a picos en el mapa de diferencias anómalas se modelaron como iones sulfato. Una densidad en forma de herradura se modeló como una molécula de pentaetilenglicol (1PE). Adicionalmente se observaron dos segmentos de densidad electrónica que se modelaron como moléculas de polietilenglicol (PEG). Una de ellas junto a la molécula A, entre H1503 y R1504 y la otra próxima a la molécula B, cercana a la región entre P1486 y Q1488. Dos densidades electrónicas próximas a la molécula B se modelaron como etilenglicol (EDO). Nueve residuos en la molécula A (N1459, L1474, V1476, E1493, I1507, S1514, R1531, E1536 y E1541) y once en la molécula B (M1456, T1460, S1467, L1474, S1477, S1491, N1498, I1507, D1519, V1527 y E1541) se modelaron en dos conformaciones alternativas, y sus ocupancias se refinaron. En ambas moléculas se modelaron todos los residuos de la región 1457-1548. Además, en la molécula A se observa densidad electrónica en el extremo amino en la que se modelaron tres residuos adicionales (SHM) de la secuencia GSHM codificados por el vector. En la molécula B sólo se introdujo el residuo adicional M por la misma razón. El modelo final también contiene 221 moléculas de solvente. El modelo presenta una geometría excelente con un 98.8% de los ángulos de torsión de la cadena principal localizados en regiones favorables del gráfico Ramachandran y el 1.2% en regiones adicionalmente permitidas (Figura R8).

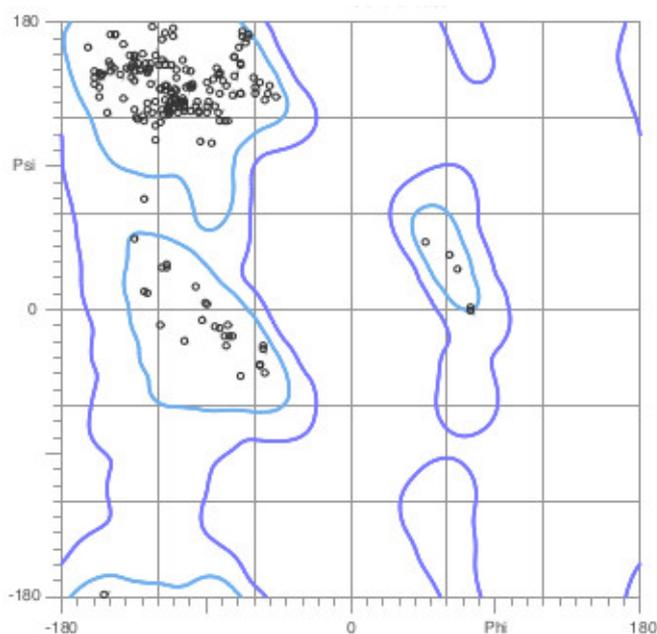


Figura R8.- Diagrama de Ramachandran de la estructura del dominio FnlII-3 de $\beta 4$.

Las dos moléculas de FnlII-3 presentes en la unidad asimétrica son prácticamente iguales. Tras superponer las dos copias del FnlII-3, la r.m.s.d. en la posición de 332 átomos de la cadena principal empleados en la superposición es 0.596 Å. El FnlII-3 presenta la estructura general característica del plegamiento de los dominios FnlII formado por dos láminas- β empaquetadas en una estructura de sándwich- β (Campbell and Spitzfaden, 1994). Una de las láminas- β está formada por 3 hebras (A, B y E) mientras que la otra lámina consiste en 4 hebras (C, C', F y G). La hebra G está dividida en dos segmentos: G1 y G2. Los extremos amino y carboxilo del dominio se encuentran en los polos opuestos del eje longitudinal de la estructura (Figura R9).

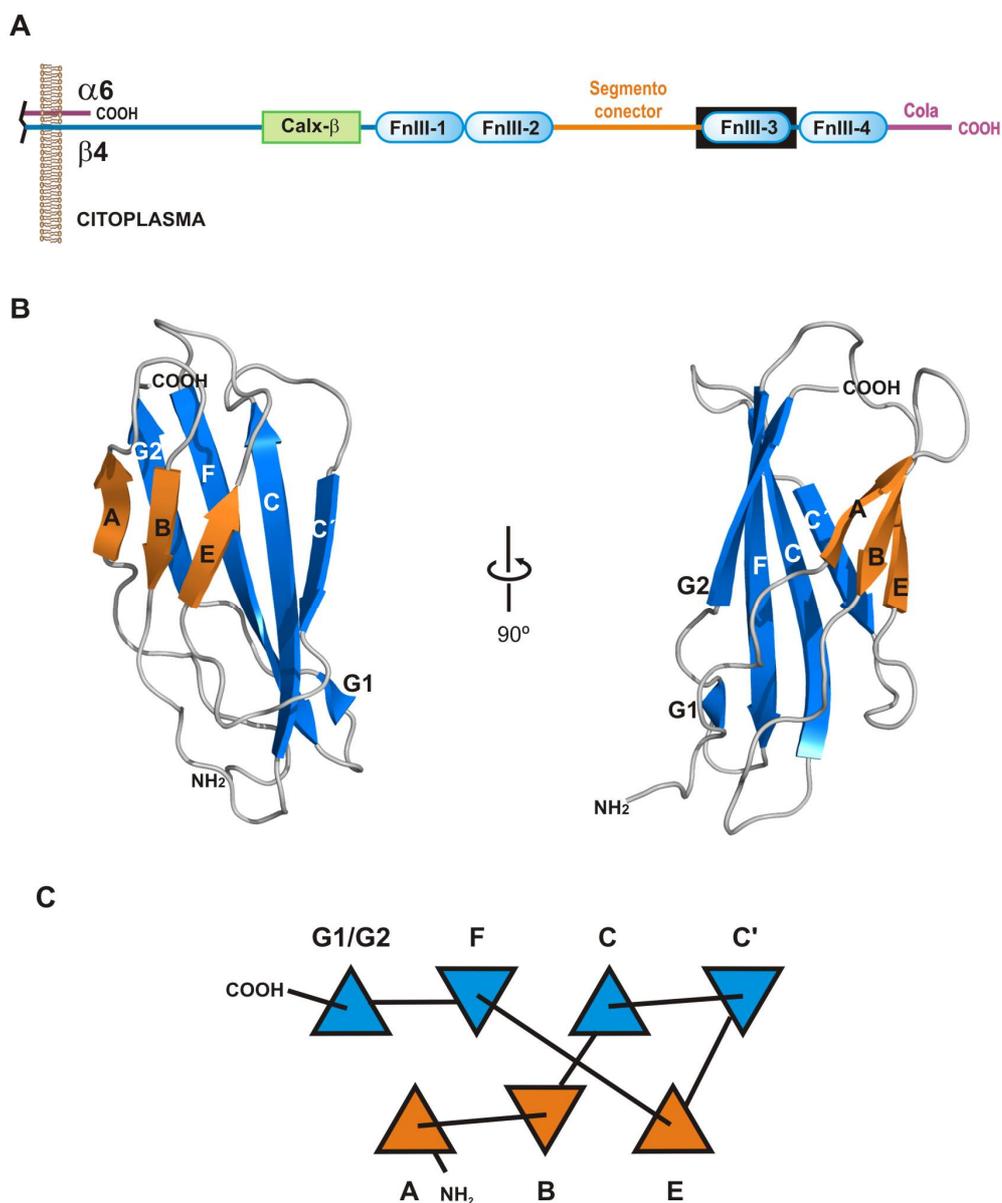


Figura R9.- Estructura cristalográfica del dominio FnIII-3 de $\beta 4$. (A) Representación esquemática de la región citoplasmática de $\beta 4$ en la que se resalta la posición del dominio FnIII-3. (B) Dos vistas ortogonales de la representación en cintas de la estructura cristalográfica de la región FnIII-3. Las hebras que forman cada una de las láminas β del dominio se muestran en el mismo color. (C) Diagrama de topología de la estructura del dominio FnIII. Los triángulos representan las hebras β que están coloreadas como en el panel B.

En la base de datos PDB se encuentra depositada una estructura del FnIII-3 de $\beta 4$ en solución resuelta mediante RMN (código 2YRZ). La estructura cristalográfica y de RMN son globalmente muy similares entre sí (Figura R10A). Tras superponer cada uno de los 20 modelos de la estructura de RMN sobre cada una de las dos moléculas de la unidad asimétrica de la estructura cristalográfica, la r.m.s.d. calculada para la posición de los Ca

de la región 1457-1548 oscila entre 0.68 Å y 0.93 Å. A pesar de la similitud global se observa una marcada diferencia en la cadena principal del segmento 1467-1469, que forma parte de la hebra A. En la estructura de RMN los grupos carbonilo de S1467, A1468 y L1489 están orientados en sentido opuesto a como se observa en la estructura cristalográfica (Figura R10B-C). Como resultado de la orientación de los enlaces peptídicos en esta región, en la estructura 2YRZ la hebra A no forma puentes de hidrógeno con la hebras B en la región 1466-1475.

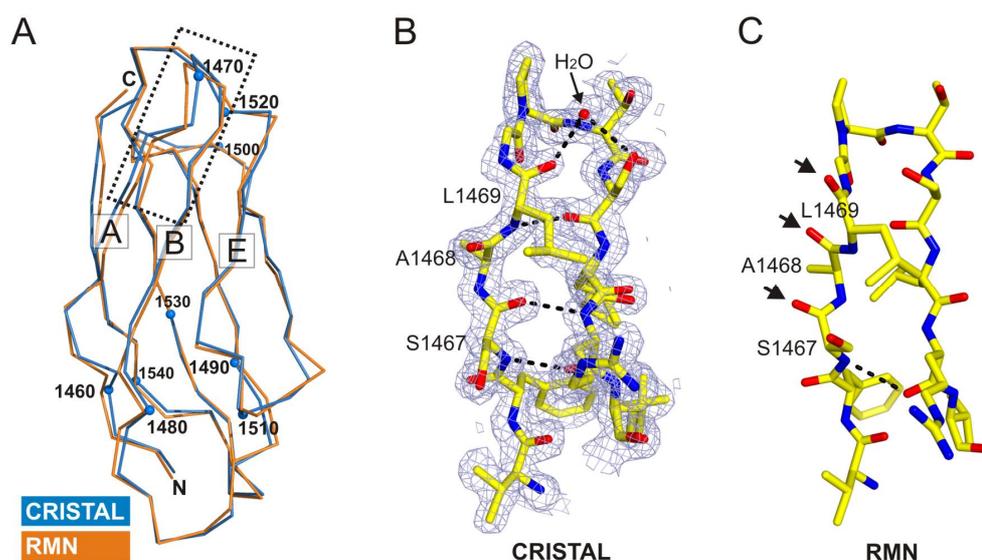


Figura R10.- Comparación de las estructuras del FnlII-3 determinadas mediante cristalografía y RMN. (A) Superposición de los trazados de C α de la estructura cristalográfica (azul) y de RMN (naranja). Se ha numerado uno de cada 10 residuos. La región indicada por un rectángulo de líneas discontinuas se muestra en detalle en los paneles B y C. (B) Detalle de la parte de las hebras A y B en la estructura cristalográfica. Los puentes de hidrógeno se indican mediante líneas discontinuas. Sobre el modelo se muestra un mapa de densidad electrónica 2mFobs-DFcalc, calculado tras refinar mediante recocido simulado un modelo en el que se había omitido la región 1466-1475. (C) Detalle de la sección equivalente a la mostrada en B en la estructura determinada mediante RMN. Se indican con flechas los grupos carbonilo que muestran una orientación marcadamente diferente a la observada en el cristal.

2.1.2.- Estructura cristalográfica del dominio FnlII-4

La estructura del dominio FnlII-4 (1572-1666) se resolvió mediante remplazo isomorfo simple con dispersión anómala (SIRAS) empleando datos de un cristal nativo y de un cristal derivatizado con el compuesto mercurial EMTS. El cristal nativo difractó a una resolución máxima de 1.80 Å, mientras que el cristal con mercurio difractó a 1.95 Å de

resolución (Tabla R5). Los cristales nativo y derivatizado son isomórficos y pertenecen al grupo espacial $P4_12_12$ con unas dimensiones de celda unidad $a=56.9 \text{ \AA}$, $b=56.9 \text{ \AA}$, y $c=78.3 \text{ \AA}$. Teniendo en cuenta el volumen de la celda unidad (253618 \AA^3), la masa molecular del FnIII-4 (10589.5 kDa) y el número de operaciones de simetría correspondientes al grupo espacial ($Z=8$) se estimó que la unidad asimétrica contenía probablemente una única molécula del FnIII-4. Esto corresponde a un contenido de solvente de ~59 % y a un coeficiente de Mathews de $2.99 \text{ \AA}^3 \text{ Da}^{-1}$.

Tabla R5.- Resumen del análisis cristalográfico del dominio FnIII-4 (1572-1666).

	Nativo	1 mM EMTS
Colección de datos		
Grupo espacial	$P4_12_12$	$P4_12_12$
Dimensiones de celda: a, b, c (Å)	56.9, 56.9, 78.4	56.9, 56.9, 78.4
Longitud de onda (Å)	1.5418	1.5418
Resolución (Å)	1.80 (1.90 – 1.80) ^a	1.95 (2.05 – 1.95) ^a
Reflexiones únicas	11858 (1632) ^a	17812 (2490) ^{a,b}
Redundancia	19.7 (20.1) ^a	15.0 (14.6) ^a
Completitud (%)	95.0 (89.5) ^a	99.6 (99.9) ^a
R_{meas} ^c (%)	5.2 (60.9) ^a	7.2 (42.0) ^a
$\langle I / \sigma \rangle$	45.1 (7.8) ^a	32.4 (8.8) ^a
R_{iso} ^d (%)		34.4
Potencia de faseado (<i>iso acent / iso cent / ano</i>)		2.38 / 2.06 / 2.36
Figura de mérito SHARP (<i>acent / cent</i>)		0.59 / 0.51
Estadísticas de refinado		
Rango de resolución (Å)	20 – 1.80 (1.98 – 1.80) ^a	
Reflexiones únicas, <i>work / free</i>	11284 / 571	
R_{work} (%)	19.2 (22.0) ^a	
R_{free} (%)	23.1 (27.0) ^a	
Nº de átomos		
Proteína	1521 [771] ^e	
Solvente	83	
Valor B promedio (Å ²)		
Gráfico de Wilson	23.8	
Proteína	34.7	
Solvente	34.6	
Desviaciones cuadráticas medias		
Longitudes de enlace (Å)	0.007	
Ángulos de enlace (°)	1.110	
Factores B cadena principal (Å ²)	2.78	
Factores B cadenas laterales (Å ²)	6.20	
Regiones mapa Ramachandran ^f		
favorecidas	105 (97.2 %)	
adicionalmente permitidas	3 (2.8 %)	
no permitidas	0	

^a Valores para la zona de alta resolución.

^b Número de reflexiones manteniendo los pares de Bijvoet separados.

^c Factor R independiente de la multiplicidad (Diederichs and Karplus, 1997).

^d $R_{iso} = \sum | |F_{der}| - |F_{nat}| | / \sum |F_{nat}|$, donde F_{der} es el factor de estructura del derivado con átomos pesados y F_{nat} el factor de estructura del cristal nativo de la proteína.

^e El número de átomos excluyendo los átomos de hidrógeno se muestra entre corchetes.

^f Valores calculados con MOLPROBITY (Lovell et al., 2003).

La subestructura de átomos de Hg del cristal derivatizado se resolvió inicialmente con los programas SHELX C/D/E (Schneider and Sheldrick, 2002) y se refinó seguidamente con SHARP (Bricogne et al., 2003). Dicha subestructura se caracteriza por un átomo de Hg con alta ocupación, que posteriormente se observó que se localiza junto al grupo tiol de la única Cys presente en el FnIII-4 (C1608). Además se detectaron otros tres Hg con baja ocupación (<10 % de la del primer Hg). Las fases calculadas con SHARP se mejoraron y extendieron hasta 1.8 Å usando los programas Solomon (Abrahams and Leslie, 1996) y DM (Cowtan, 1994). Utilizando estas fases se calculó un mapa de una calidad excepcional (Figura R11) que permitió la construcción automática de 94 residuos de los 99 de la molécula empleando el programa ARP/WARP (Langer et al., 2008).

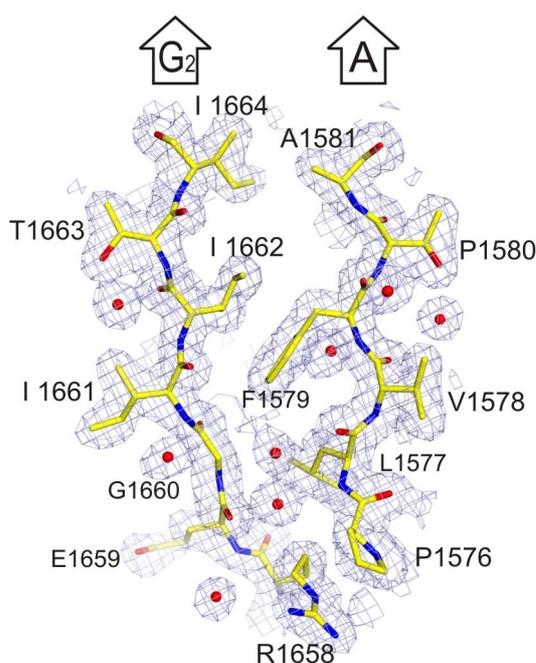


Figura R11.- Mapa de densidad electrónica del dominio FnIII-4 calculado con fases estimadas a partir de SIRAS. El mapa se calculó combinando las amplitudes observadas para el cristal nativo con las fases derivadas de SIRAS, tras mejorarlas con métodos de modificación de densidad. Se muestra una región del mapa, a nivel de 1σ , en torno a las hebras A y G2 del FnIII-4. Se muestra superpuesto el modelo refinado de esta región de la estructura.

La estructura se refinó frente a datos nativos a 1.80 Å de resolución de forma análoga a como se refinó la estructura del FnIII-3, con pequeñas diferencias. En este caso se refinó un único grupo TLS. Cinco residuos se modelaron en dos conformaciones alternativas (Q1588, E1609, M1610, T1618 y N1639) y sus ocupancias se refinaron. La unidad asimétrica contiene una molécula del FnIII-4, en la cual se modelaron todos los residuos de la región 1572-1666 y dos residuos adicionales (His-Met) en el extremo amino. El modelo incluye además 81 moléculas de agua. El refinado convergió a unos valores de R_{work} y R_{free} de 19.2 % y 23.1 % respectivamente. La estructura tiene una geometría excelente; los ángulos de torsión de la cadena principal del 98.9% de los residuos ocupa las regiones más favorecidas del mapa de Ramachandran, el 2.1% restante se encuentra en posiciones adicionalmente permitidas del gráfico (Figura R12). La estructura global del FnIII-4 es muy similar a la observada en el FnIII-3 y otros dominios FnIII (Figura R13).

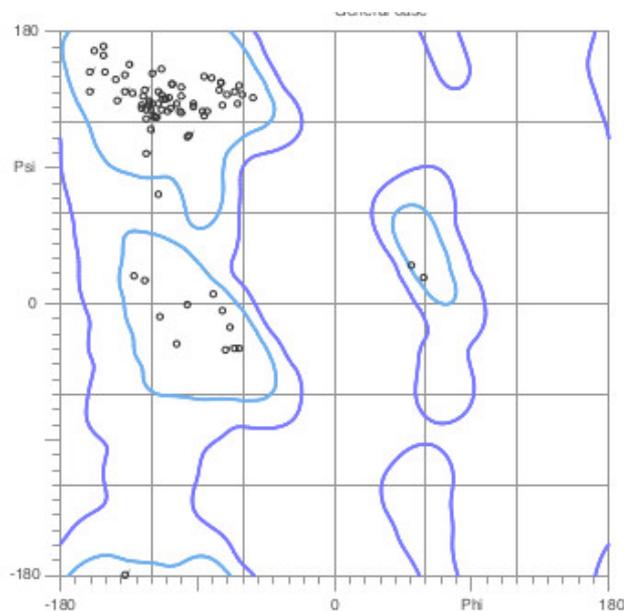


Figura R12.- Diagrama de Ramachandran de la estructura del dominio FnIII-4 de $\beta 4$.

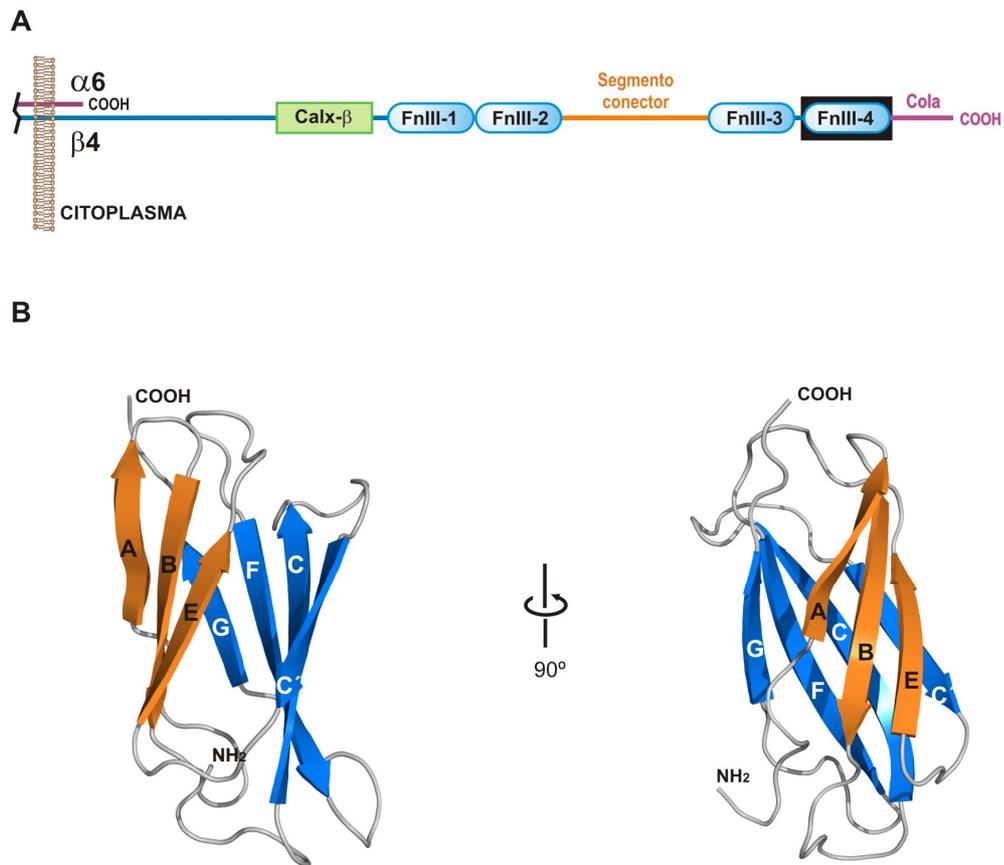


Figura R13.- Estructura cristalográfica del dominio FnIII-4 de $\beta 4$. (A) Representación esquemática de la región citoplasmática de $\beta 4$, el dominio FnIII-4 se resalta sobre fondo negro. (B) Dos vistas ortogonales de la representación en cintas de la estructura cristalográfica de la región FnIII-4. Las hebras que forman cada una de las láminas β del dominio se muestran en el mismo color.

2.1.3.- Comparación de las estructuras de los cuatro dominios FnIII de la integrina $\beta 4$

La superposición simultánea de las estructuras cristalográficas de los cuatro dominios FnIII de $\beta 4$ revela un alto grado de conservación en una amplia región del esqueleto polipeptídico que incluye las dos láminas- β y los lazos AB, EF y FG (Figura R14). Por el contrario los lazos BC, CC' y C'E, que presentan longitudes variables en los diferentes dominios FnIII, son las regiones en las que se observan las mayores diferencias en el trazado de la cadena polipeptídica.

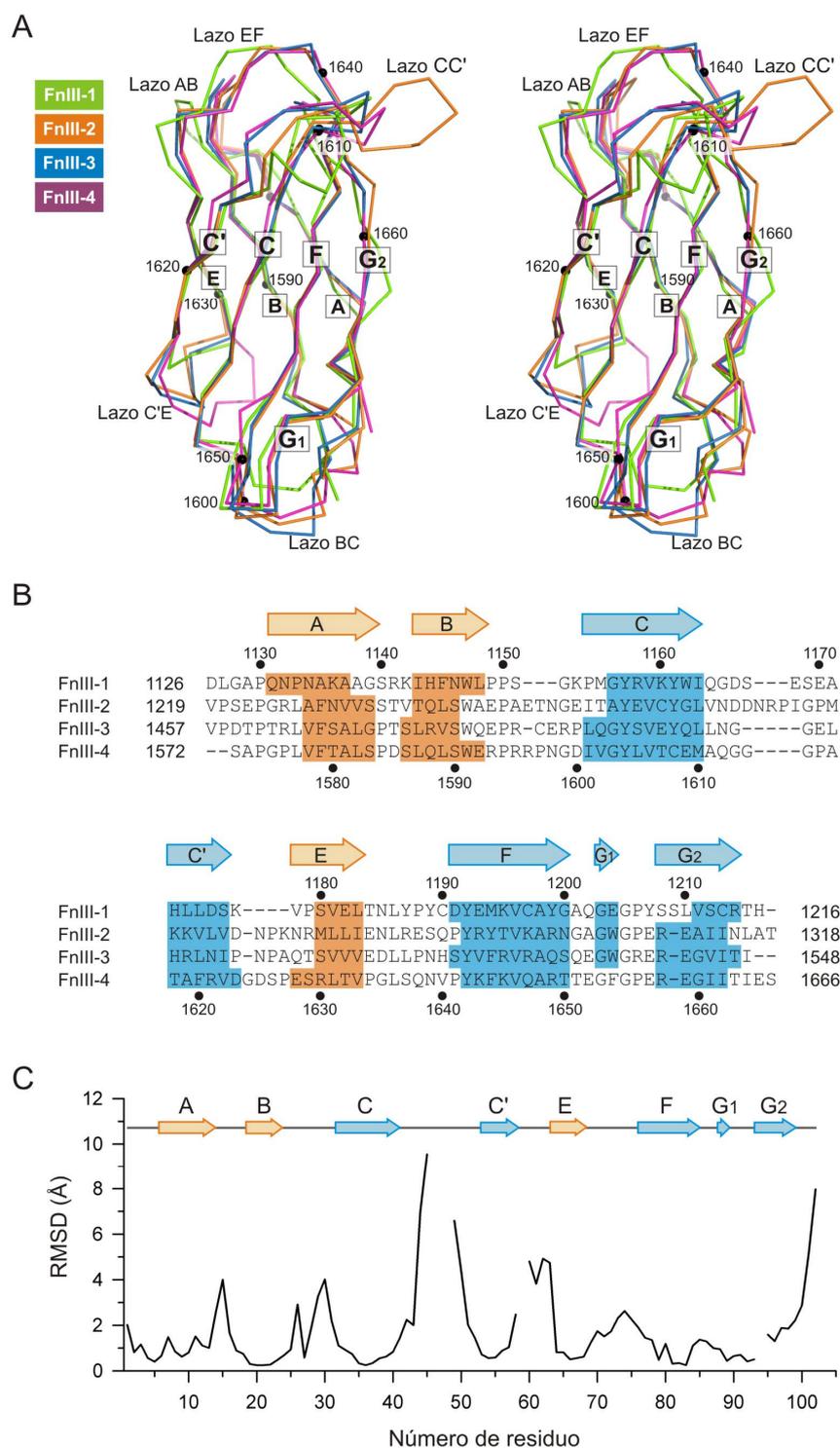


Figura R14.- Comparación de los cuatro dominios FnIII de $\beta 4$. (A) Par estereoscópico de los dominios FnIII de $\beta 4$ superpuestos empleando el programa THESEUS que utiliza un método de máxima similitud. El dominio FnIII-1 (verde) sobre los dominios FnIII-2 (naranja), FnIII-3 (azul) y FnIII-4(magenta). Se numeran los C α de cada décimo residuo del FnIII-4. (B) Alineamiento de secuencias basado en la superposición de las estructuras de los dominios FnIII. Los elementos de estructura secundaria se indican en la parte superior y mediante el sombreado da las secuencias correspondientes, siguiendo el mismo código de colores que en las figuras R9 y R11. (C) Representación de la r.m.s.d. por residuo. La numeración corresponde a la posición en el alineamiento múltiple de secuencia mostrado en B.

A pesar de que la estructura global de los cuatro dominios FnIII está conservada, la comparación por pares de estos dominios revela que los esqueletos polipeptídicos del FnIII-3 y el FnIII-4 son más similares entre sí que con el FnIII-1 o el FnIII-2 (Tabla R6). La similitud estructural está correlacionada con la conservación de secuencia ya que el grado de identidad de secuencia entre el FnIII-3 y el FnIII-4 (39%) es mayor que entre cualquier otra pareja de dominios FnIII de $\beta 4$.

Tabla R6.- Comparación por pares de los dominios FnIII de $\beta 4$.

	FnIII-1	FnIII-2	FnIII-3	FnIII-4	
FnIII-1		1.690 (Fn1A/Fn2A) 1.801 (Fn1A/Fn2B) 1.616 (Fn1B/Fn2A) 1.722 (Fn1B/Fn2B)	1.471 (Fn1A/Fn3A) 1.464 (Fn1A/Fn3B) 1.388 (Fn1B/Fn3A) 1.550 (Fn1B/Fn3B)	1.638 (Fn1A) 1.573 (Fn1B)	r.m.s.d. (Å) ^a
FnIII-2	17 %		1.291 (Fn2A/Fn3A) 1.310 (Fn2A/Fn3B) 1.317 (Fn2B/Fn3A) 1.307 (Fn2B/Fn3B)	1.491 (Fn2A) 1.541 (Fn2B)	
FnIII-3	24 %	33 %		1.033 (Fn3A) 1.217 (Fn3B)	
FnIII-4	20 %	37 %	39 %		
Identidad de secuencia					

^a Las estructuras se superpusieron empleando el programa DaliLite (Holm and Park, 2000). Se muestra la r.m.s.d. en la posición de los C α empleados para realizar la superposición.

El grado de conservación de secuencia entre los dominios no está distribuido homogéneamente en las estructuras FnIII. Los residuos orientados hacia el interior de la estructura están más conservados que los expuestos al solvente. Hay 11 residuos idénticos en los cuatro dominios FnIII. En el FnIII-3 estos residuos son: P1461, W1478, P1481, Y1490, V1492, L1520, Y1526, V1530, A1532, G1537 y G1539. De éstos, todos excepto G1537 y G1539, forman parte del núcleo hidrofóbico encerrado entre las dos láminas β . La diversidad en los residuos expuestos al solvente entre dominios se pone de manifiesto en las propiedades electrostáticas diferenciadas de cada dominio FnIII de $\beta 4$ (Figura R15).

En resumen, el esqueleto polipeptídico de los dominios FnIII de $\beta 4$ está altamente conservado pero las características de las superficies expuestas al solvente son específicas de cada dominio.

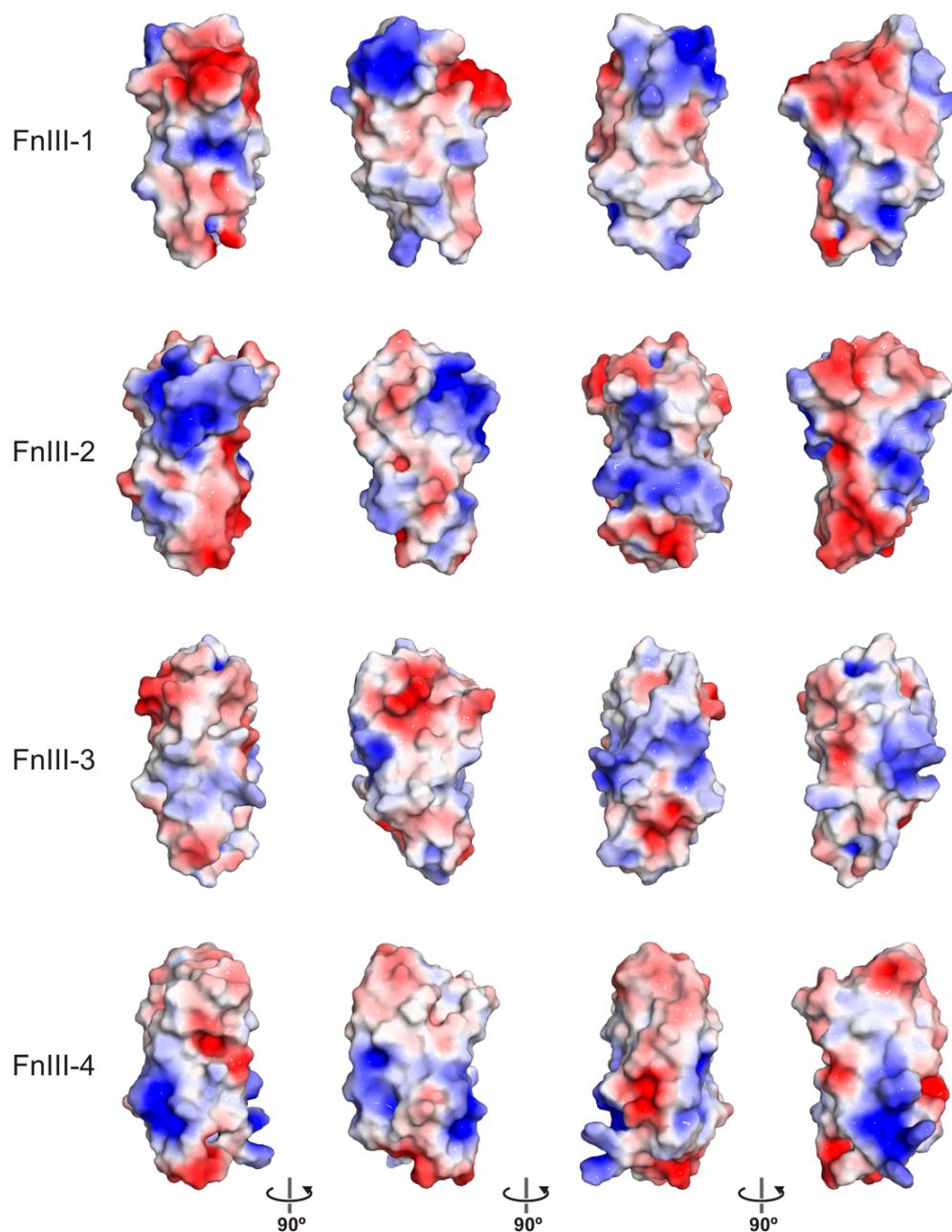


Figura R15.- Potencial electrostático molecular de los dominios FnIII de $\beta 4$. Representación de la superficie accesible al solvente de cada dominio FnIII coloreada según el potencial electrostático en dichas superficies. El potencial electrostático se calculó para cada dominio aislado con el programa APBS. Se utilizó el campo PARSE (Sitkoff et al., 1994) y se asignó carga neutra a los extremos amino y carboxilo de cada dominio. La escala de color varía entre -3 kTe^{-1} (rojo) y $+3 \text{ kTe}^{-1}$ (azul). Las representaciones de la izquierda corresponde a una orientación con la lámina- β ABE en el frente. Las otras 3 vistas corresponden a giros progresivos de 90° alrededor del eje vertical.

2.2.- Estructura en solución de la región FnIII-3,4, análisis mediante dispersión de rayos-x a bajo ángulo (SAXS)

La estructura en solución de la región FnIII-3,4 (1457-1666) de $\beta 4$ se caracterizó mediante SAXS. Esta técnica facilita información sobre el tamaño y la forma de macromoléculas en solución y permite obtener modelos estructurales a baja resolución. El análisis mediante SAXS, que normalmente es complementario a los estudios cristalográficos, ha sido de especial importancia en este caso debido a la imposibilidad de obtener cristales del fragmento FnIII-3,4.

El análisis mediante SAXS del fragmento FnIII-3,4 de $\beta 4$ se llevó a cabo en varios niveles. En primer lugar se estimaron los parámetros estructurales a partir de los perfiles experimentales de SAXS. Por otro lado se calcularon las estructuras a baja resolución mediante métodos *ab initio*. Finalmente los datos de SAXS se emplearon para generar y evaluar modelos atómicos de la región FnIII-3,4 creados empleando las estructuras cristalográficas de los dominios FnIII-3 y FnIII-4 aislados.

2.2.1.- Caracterización estructural de la región FnIII-3,4 en solución

Los perfiles de dispersión de la región FnIII-3,4 se midieron a cuatro concentraciones de proteína (1, 2, 8 y 16 mg ml⁻¹) (Figura R16A). Estos datos se normalizaron para cada concentración de proteína y se extrapolaron a dilución infinita (Figura R16B). Éste último es superponible con el perfil obtenido para cada una de las concentraciones (tras escalar por la concentración), indicando la ausencia de agregación en el rango de concentraciones de trabajo; además no se aprecian grandes diferencias en el radio de giro (R_g) calculado mediante la aproximación de Guinier para cada concentración, que es lineal para todas las concentraciones medidas, corroborando la buena calidad de los datos obtenidos.

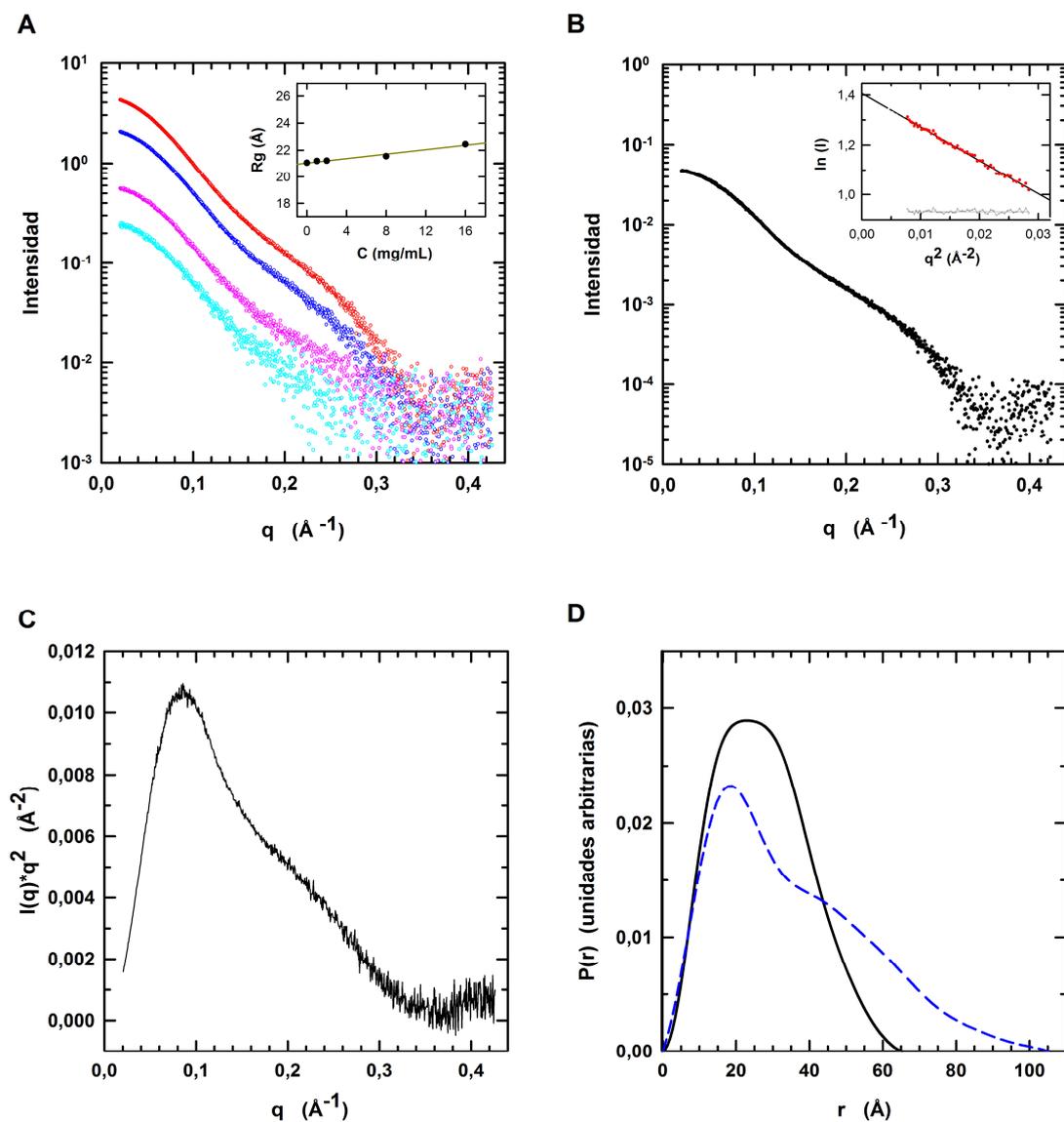


Figura R16.- Análisis de la región FnIII-3,4 mediante SAXS. (A) Perfiles de dispersión experimentales determinados a varias concentraciones de proteína; de mayor a menor intensidad: 16, 8, 2, y 1 mg ml⁻¹. En el inserto se representan los valores de R_g determinados mediante análisis de Guinier a cada concentración de proteína, así como el valor obtenido a partir de los datos extrapolados a dilución infinita. (B) Perfil de dispersión experimental extrapolado a dilución infinita. En el inserto se muestra el gráfico de Guinier. (C) Gráfico de Kratky calculado a partir de los datos de SAXS. (D) Función de distribución de distancias interatómicas. En negro se muestra la curva correspondiente a la región FnIII-3,4 y en azul con línea discontinua la de la región FnIII-1,2 (datos obtenidos de Chacon et al., 2000) para su comparación.

El gráfico de Kratky calculado para los datos de dispersión del FnIII-3,4 muestra una curva en forma de campana con un máximo bien definido que se aproxima a cero a valores altos de q (Figura R16C); esta forma es característica de proteínas globulares y

sugiere que el FnIII-3,4 en solución tiene una estructura compacta sin una presencia significativa de regiones desordenadas.

La masa molecular aparente que corresponde al volumen de exclusión estimado a partir de los datos de SAXS fue 23.2 kDa. Este valor es similar a la masa molecular del fragmento FnIII-3,4 calculada a partir de su secuencia (23591 Da), indicando que la proteína es un monómero en solución. Los datos de dispersión a muy bajo ángulo cumplen la aproximación de Guinier siendo el R_g 21.1 ± 0.12 Å; estimado con datos en el rango $0.56 < qR_g < 1.07$.

La función de distribución de pares, $P(r)$, tiene forma de campana característica de partículas compactas (Figura R16D). El R_g estimado en el cálculo de la $P(r)$ es 21.1 ± 0.5 Å. Este valor es igual al obtenido a partir del gráfico de Guinier. La dimensión máxima (D_{max}), posición donde la función $P(r)$ vuelve a ser cero para valores altos de r , es de 65 Å. El R_g y la D_{max} del FnIII-3,4 contrastan con los descritos para la región FnIII-1,2 (1126-1370) a partir del análisis de SAXS (R_g 29 Å y D_{max} 105 Å) (Chacon et al., 2000). La $P(r)$ de FnIII-1,2 presenta un máximo a distancias cortas y una cola que se extiende hasta valores altos de r ; este perfil es característico de una molécula alargada y concuerda con la estructura cristalográfica del FnIII-1,2 (de Pereda et al., 1999).

En resumen, los datos de SAXS indican que la región FnIII-3,4 es monomérica en solución y que presenta una estructura compacta que difiere significativamente de la estructura alargada observada para la región FnIII-1,2.

2.2.2.- Estructura de la región FnIII-3, 4 en solución

A partir de los datos de SAXS se obtuvieron modelos a baja resolución de la región FnIII-3,4 mediante la utilización de métodos *ab initio*. Para ello se utilizaron los programas DAMMIN (Svergun, 1999) y DALAI_GA (Chacon et al., 2000), que emplean algoritmos diferentes. Con cada programa se generaron 10 reconstrucciones independientes. El valor de promedio de NSD, que es una medida de la similitud entre modelos, para el conjunto de los 10 modelos generados con DAMMIN es 0.590 ± 0.013 (media \pm desviación estándar) y la NSD de los modelos obtenidos con DALAI_GA es 0.637 ± 0.018 .

Ningún modelo presenta un valor de NSD alejado del valor medio más de dos veces la desviación estándar. Los bajos valores de NSD indican que cada programa produce reconstrucciones estables que dan lugar a estructuras muy similares entre sí. A partir de cada conjunto de 10 modelos se calculó una estructura promedio. Las estructuras promedio calculadas con DAMMIN y DALAI_GA son muy similares entre sí, siendo el valor de NSD entre ellas de 0.617. En resumen, los modelos *ab initio* generados a partir de los datos obtenidos por SAXS son muy similares entre sí, independientemente del algoritmo de reconstrucción empleado.

La estructura del FnIII-3,4 en solución presenta una forma compacta y plana ligeramente alabeada con unas dimensiones de aproximadamente 65 Å x 45 Å x 30 Å (Figura R17). La reconstrucción promediada obtenida con el programa DAMMIF muestra un surco central que divide la estructura en dos lóbulos de tamaño similar. Sin embargo, este surco no está definido en la reconstrucción obtenida con el programa DALAI_GA. Las estructuras sugieren que existen interacciones laterales entre los dos dominios FnIII y que los ejes longitudinales de éstos dominios se encuentran aproximadamente en el mismo plano. La estructura compacta de FnIII-3,4 es marcadamente diferente de la región FnIII-1,2 que tiene forma alargada o de varilla, con una longitud de ~85 Å y una anchura de ~23 Å (de Pereda et al., 2009b; de Pereda et al., 1999).

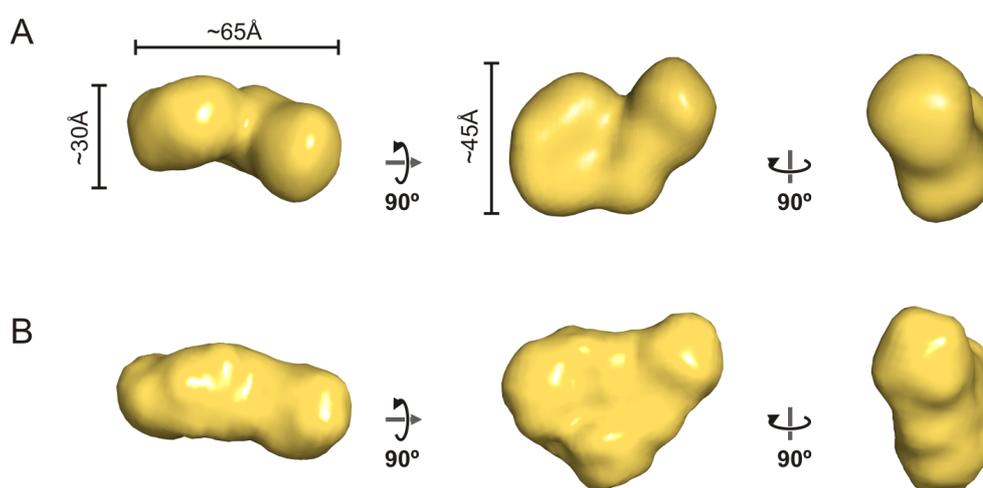


Figura R17.- Reconstrucciones de la estructura de la región FnIII-3,4 a partir de perfiles de SAXS. (A) Tres vistas ortogonales de la estructura promedio determinada a partir de diez modelos generados con el programa DAMMIN. (B) Estructura promedio correspondiente a diez modelos obtenidos con el programa DALAI_GA. Las tres representaciones en A y B se muestran en orientaciones equivalentes y en la misma escala.

Dado que se disponía de las estructuras cristalográficas de los dominios FnIII-3 y FnIII-4 en principio se podría pensar en elucidar la estructura de esta región haciendo un ajuste frente a los datos de SAXS en el que cada dominio FnIII se trate como un cuerpo rígido. Sin embargo, no es posible emplear este tipo de ajuste en este caso ya que no se puede diferenciar el FnIII-3 del FnIII-4. Además, a la resolución de los datos de SAXS, cada dominio FnIII se asemeja a un esferoide prolato, por lo que no sería posible discriminar entre las orientaciones relacionadas por la rotación a lo largo del eje longitudinal, ni entre las dos orientaciones posibles de los extremos longitudinales de cada dominio.

2.3.- Análisis de la orientación relativa de los dominios FnIII-3 y FnIII-4 mediante SDSL-EPR-DEER

Con el fin de conocer la orientación relativa de los dos dominios de la región FnIII-3,4 se midieron las distancias entre pares de residuos en el FnIII-3 y el FnIII-4. Para ello se combinó marcaje paramagnético dirigido (SDSL) con resonancia paramagnética electrónica de doble electrón (EPR-DEER). La espectroscopía DEER permite medir distancias entre centros paramagnéticos en un rango de 2 a 6 nm, pudiendo llegar en algunos casos a 8 nm (Jeschke, 2012). En el caso de proteínas que no contienen centros paramagnéticos en su estructura nativa, como es el caso de la integrina β_4 , éstos se pueden introducir mediante el marcaje con sondas paramagnéticas. En este trabajo se ha empleado la sonda MTSSL que contiene un grupo nitroxido como centro paramagnético y un grupo metanosulfonato que reacciona con grupos tiol de Cys formando un enlace disulfuro.

Los estudios de SDSL-DEER se realizaron empleando el mismo fragmento de β_4 , residuos 1457-1666, que se había empleado en la caracterización mediante SAXS. Con el fin de introducir la sonda MTSSL en pares de residuos específicos, uno en el FnIII-3 y el otro en el FnIII-4, se generó una colección de mutantes en los que se eliminaron algunas de las Cys de la proteína silvestre y se introdujeron otras nuevas. La secuencia silvestre de este fragmento contiene tres Cys. C1483 se localiza cerca del final de la hebra- β B del dominio FnIII-3 y su grupo SH se encuentra oculto mayoritariamente por los lazos BC y FG₁. C1559 se encuentra en centro de la secuencia que conecta los dominios FnIII-3 y

FnIII-4. C1608 forma parte de la hebra- β C del FnIII-4, su grupo SH se localiza en la hendidura formada por la hebra C' y el lazo EF y se encuentra expuesto al solvente.

La idea inicial era sustituir las tres Cys naturales y sobre ese triple mutante crear dobles mutantes, cada uno de ellos con una nueva Cys en cada dominio FnIII. En todos los casos se ha sustituido la Cys del conector por Ala (C1559A). Los dobles mutantes C1483A/C1559A y C1559A/C1608A se expresan solubles en *E. coli* (Tabla R7). Sin embargo el triple mutante C1483A/C1559A/C1608A se localiza en la fracción insoluble del lisado bacteriano. De igual modo la sustitución de C1483 o C1608 por Ser en triples mutantes da lugar a proteínas insolubles. Por lo tanto se crearon mutantes del fragmento FnIII-3,4 con una Cys de la secuencia silvestre en uno de los dominios FnIII y una nueva Cys en el otro FnIII. Las nuevas Cys se introdujeron en residuos de la superficie de los FnIII identificados a partir del análisis de las estructuras cristalográficas de los dominios FnIII-3 y FnIII-4.

Tabla R7.- Mutantes de β 4 FnIII-3,4 (1457-1666) en los que se han sustituido las Cys de la secuencia silvestre y resultados de la expresión en *E. coli*.

Nombre	Localización de las mutaciones			Expresión en <i>E. coli</i>
	FnIII-3	conector	FnIII-4	
CX1		C1559A		Soluble a 15°
CX2	C1483A	C1559A		Soluble a 15°
CX4		C1559A	C1608A	Soluble a 15°
CX5	C1483A	C1559A	C1608A	Insoluble
CX6	C1483S	C1559A	C1608A	Insoluble
CX7	C1483A	C1559A	C1608S	Insoluble
CX8	C1483S	C1559A		Soluble a 15°
CX9		C1559A	C1608S	Insoluble
CX10	C1483S	C1559A	C1608S	Insoluble

A partir del doble mutante C1483S/C1559A, denominado CX8, se crearon una serie de mutantes de FnIII-3,4 que contenían C1608 en el FnIII-4 y una segunda Cys en diversas posiciones del FnIII-3 (Tabla R8). Por sencillez, en los triples mutantes se indican sólo las dos Cys que contienen. Inicialmente se crearon cuatro mutantes: R1485C/C1608 (CX8A), N1523C/C1608 (CX8B), R1463C/C1608 (CX8C) y R1504C/C1608 (CX8D). R1485 y N1523 se localizan en extremos opuestos del eje longitudinal del FnIII-3; mientras que R1463 y R1504 se encuentran lados opuestos en la región ecuatorial de este dominio (Figura R18). Estos mutantes se expresaron solubles en bacterias y se marcaron con MTSSL.

Tabla R8.- Mutantes de $\beta 4$ FnIII-3,4 (1457-1666) de la serie CX8.

Nombre	Localización de las mutaciones			Expresión en <i>E. coli</i>
	FnIII-3	conector	FnIII-4	
CX8A	C1483S, R1485C	C1559A	-	Soluble a 15°
CX8B	C1483S, N1523C	C1559A	-	Soluble a 15°
CX8C	R1463C , C1483S	C1559A	-	Soluble a 15°
CX8D	C1483S, R1504C	C1559A	-	Soluble a 15°
CX8E	R1475C , C1483S	C1559A	-	Soluble a 15°
CX8F	A1468C , C1483S	C1559A	-	Soluble a 15°
CX8G	T1472C , C1483S	C1559A	-	Soluble a 15°
CX8H	C1483S, L1497C	C1559A	-	Soluble a 15°
CX8I	C1483S, A1511C	C1559A	-	Soluble a 15°

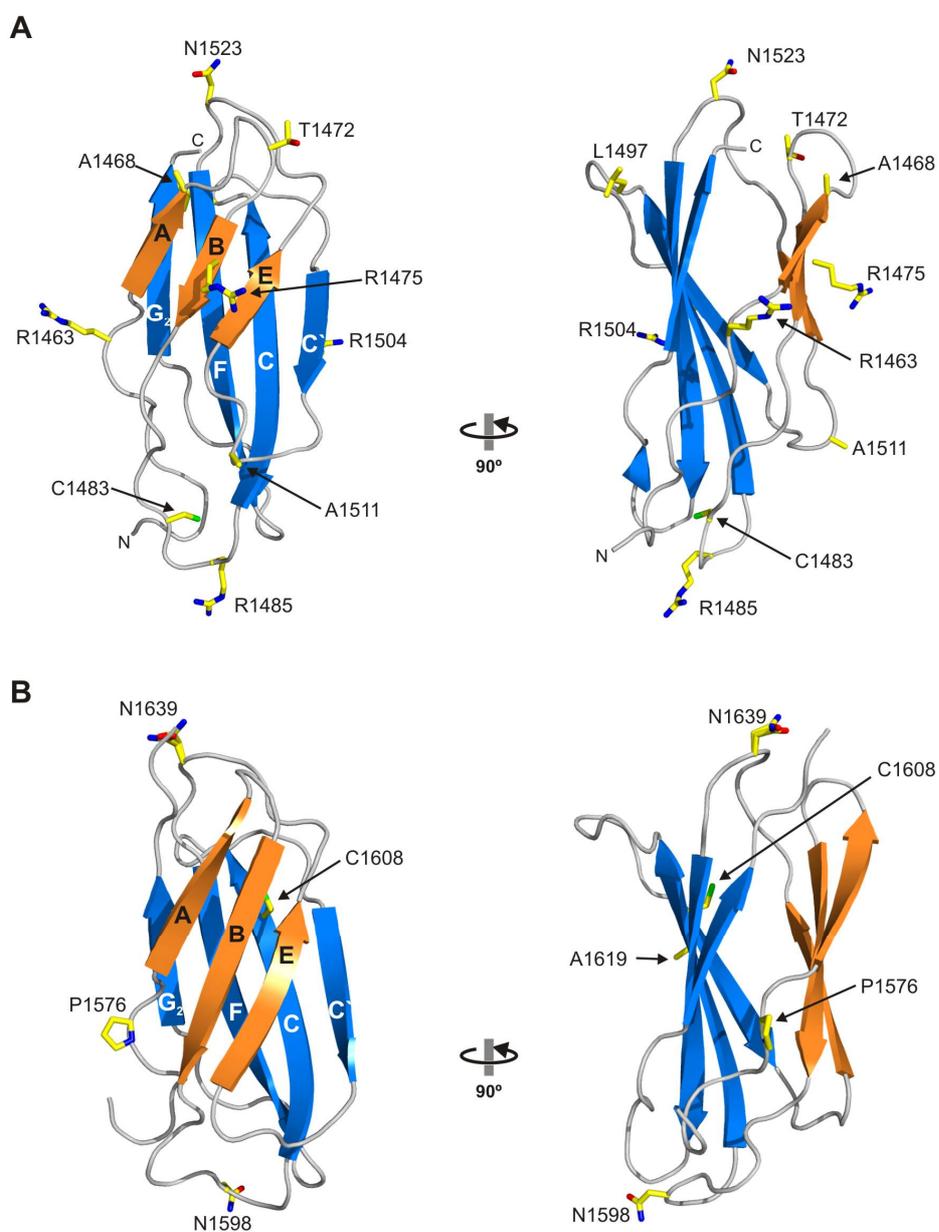


Figura R18.- Localización en la estructura del FnIII-3 y del FnIII-4 de los residuos sustituidos por Cys. Representación de cintas de las estructuras del FnIII-3 (A) y del FnIII-4 (B). Se muestran las cadenas laterales de los residuos que se han mutado a Cys. Además se indica la posición de C1483 en el FnIII-3 y C1608 en el FnIII-4, presentes en la secuencia silvestre. Todos los residuos sustituidos por Cys se localizan en la superficie de la estructura.

A partir de los espectros de DEER de los mutantes R1485C/C1608 (CX8A), R1463C/C1608 (CX8C) y R1504C/C1608 (CX8D) se obtuvieron distribuciones de distancias inter-spin con un pico estrecho mayoritario (Figura R19 y Tabla R9). Sin embargo el mutante N1523C/C1608 (CX8B) muestra una distribución de distancias bimodal. N1523 forma parte del lazo EF y está muy expuesto al solvente por lo que es posible que la distribución bimodal de distancias se deba a una elevada libertad conformacional de este lazo. A pesar de disponer de tres distancias bien definidas, no es posible asignar la posición de C1608 con precisión ya que R1463, R1485, R1504 y N1523 se encuentran aproximadamente en el mismo plano. Con el fin de resolver la ambigüedad sobre la posición de C1608 respecto al FnIII-3 se crearon cinco mutantes adicionales derivados del CX8: R1475C/C1608 (CX8E), A1468C/C1608 (CX8F), T1472C/C1608 (CX8G), L1497C/C1608 (CX8H) y A1511C/C1608 (CX8I). Estos cinco mutantes se expresaron solubles en bacteria. Las distancias inter-spin obtenidas con los mutantes R1475C/C1608 (CX8E), A1468C/C1608 (CX8F) y A1511C/C1608 (CX8I) muestran distribuciones anchas y con múltiples picos. Esto sugiere que las sustituciones R1475C, A1468C y A1511C o el marcaje con la sonda MTSSL en estas posiciones causan alteraciones en la estructura del FnIII-3 o en la interacción de éste con el FnIII-4. Por otro lado se obtuvieron distancias bien definidas empleando los mutantes T1472C/C1608 (CX8G) y L1497C/C1608 (CX8H). En conjunto a partir de los mutantes de esta serie CX8 se determinaron cinco distancias entre centros paramagnéticos que permitieron conocer la posición de C1608 respecto al FnIII-3.

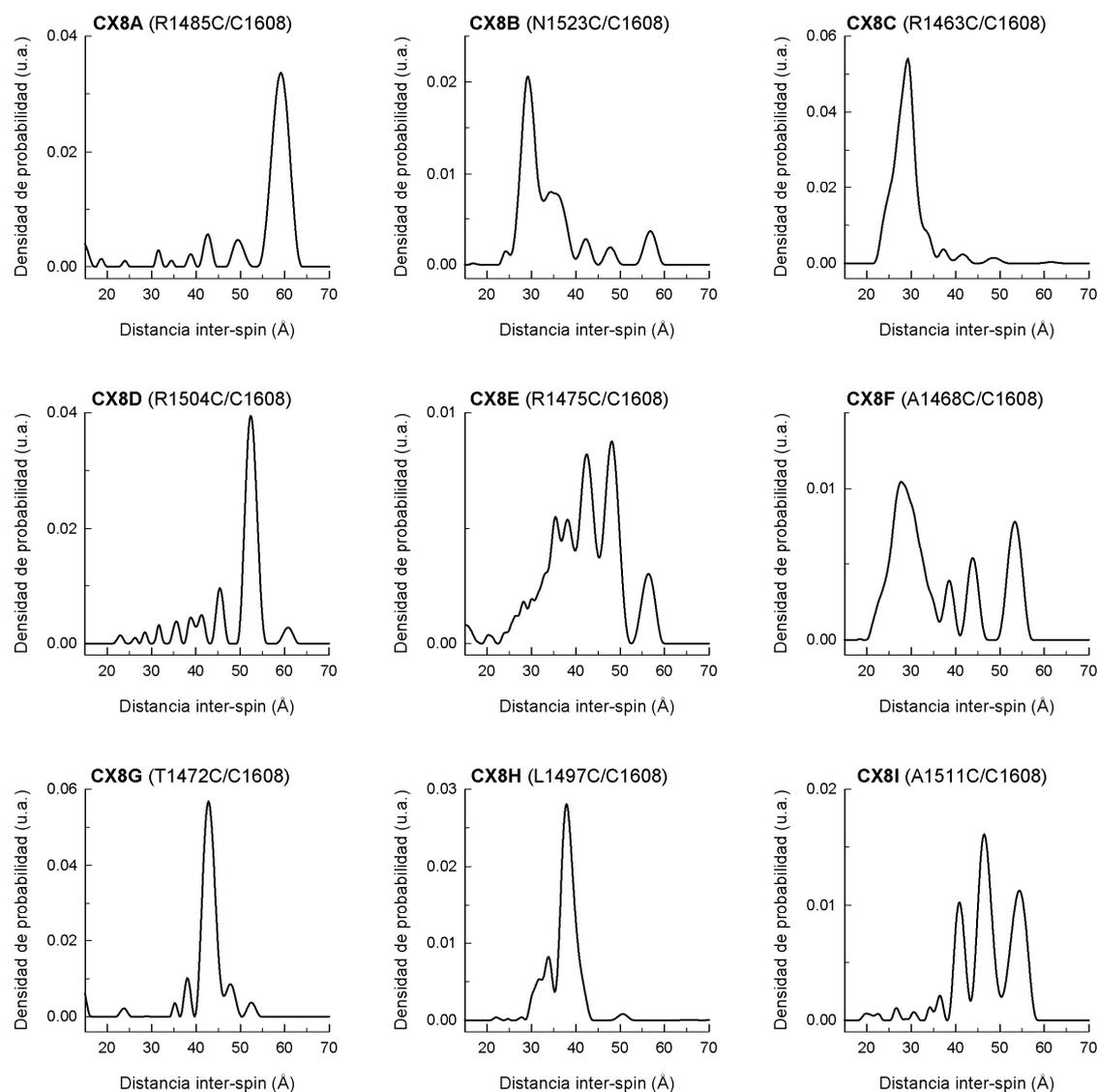


Figura R19.- Distribuciones de distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX8 calculadas a partir de los datos de EPR-DEER.

Tabla R9.- Distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX8.

Nombre	Cys en FnIII-3	Cys en FnIII-4	Distancia (Å)	Comentario
CX8A	R1485C	C1608	59	
CX8B	N1523C	C1608	29 / 35	Distribución bimodal
CX8C	R1463C	C1608	29	
CX8D	R1504C	C1608	53	
CX8E	R1475C	C1608	42 / 48 / 56	Distribución ancha, varios picos
CX8F	A1468C	C1608	28 / 44 / 53	Distribución ancha, varios picos
CX8G	T1472C	C1608	43	
CX8H	L1497C	C1608	38	
CX8I	A1511C	C1608	41 / 46 / 54	Distribución ancha, varios picos

De forma análoga se abordó la localización de C1483 del FnIII-3 con respecto al FnIII-4 empleando triples mutantes de $\beta 4$ FnIII-3,4 derivados de la proteína que contiene la doble sustitución C1559A/C1608A (CX4). Se crearon cuatro triples mutantes: C1483/N1598C (CX4A), C1483/N1639C (CX4B), C1483/P1576C (CX4C) y C1483/A1619C (CX4D) (Tabla R10). N1598, N1639C, P1576C y A1619 del FnIII-4 ocupan posiciones respectivamente equivalentes a R1485, N1523, R1463 y R1504 en el FnIII-3, los cuales se habían sustituido por Cys en los mutantes CX8A-D. Los cuatro mutantes de la serie CX4 se expresaron solubles, se marcaron con MTSSL y se analizaron mediante DEER. Las distribuciones de distancias entre centros paramagnéticos obtenidas con estos mutantes muestran en todos los casos picos múltiples y muy anchos (Figura R20 y Tabla R11), lo cual sugiere que bien las sustituciones introducidas o el marcaje con MTSSL causan alteraciones en la estructura de la región FnIII-3,4. Puesto que los cuatro residuos del FnIII-4 sustituidos por Cys se localizan en regiones separadas de la superficie del dominio y que la introducción de Cys en posiciones equivalentes del FnIII-3 no dio lugar a distribuciones anchas de distancias inter-spin; es razonable que la unión del MTSSL a C1483 sea la causa de alteraciones estructurales en el FnIII-3,4. En este sentido, la función de partición del MTSSL (que es una estimación de la accesibilidad del sitio de marcaje) calculada a partir del modelado de rotámeros de la sonda unido a C1483 es 0.041; un valor <0.05 se asocia a sitios poco favorables para el marcaje con MTSSL en los que la sonda puede introducir distorsiones en la estructura de la proteína.

Tabla R10. Mutantes de $\beta 4$ FnIII-3,4 (1457-1666) de la serie CX4.

Nombre	Localización de las mutaciones			Expresión en <i>E. coli</i>
	FnIII-3	conector	FnIII-4	
CX4A	-	C1559A	N1598C , C1608A	Soluble a 15°
CX4B	-	C1559A	C1608A, N1639C	Soluble a 15°
CX4C	-	C1559A	P1576C , C1608A	Soluble a 15°
CX4D	-	C1559A	C1608A, A1619C	Soluble a 15°

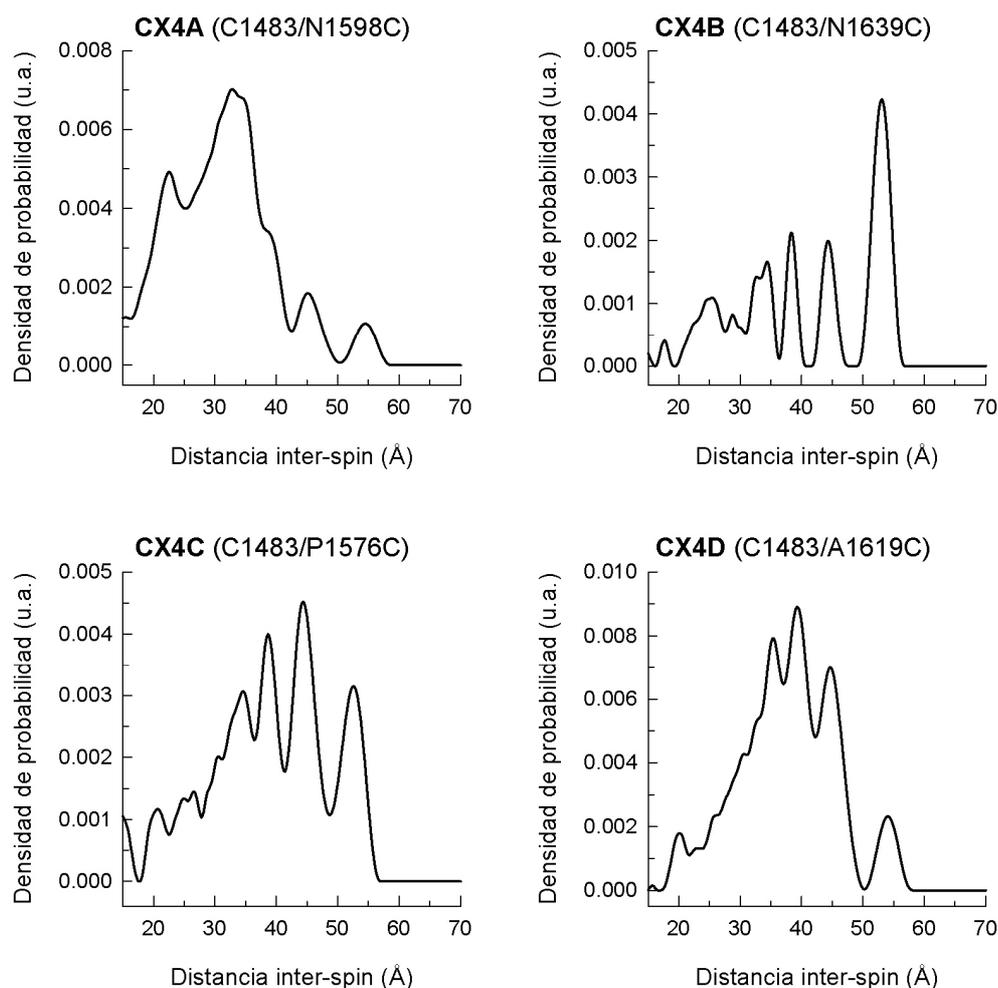


Figura R20.- Distribuciones de distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX4 calculadas a partir de los datos de EPR-DEER.

Tabla R11.- Distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX4.

Nombre	Cys en FnIII-3	Cys en FnIII-4	Distancia (Å)	Comentario
CX4A	C1483	N1598C	n.d.	Distribución ancha, varios picos
CX4B	C1483	N1639C	n.d.	Distribución ancha, varios picos
CX4C	C1483	P1576C	n.d.	Distribución ancha, varios picos
CX4D	C1483	A1619C	n.d.	Distribución ancha, varios picos

Con el fin de introducir el MTSSL en una posición cercana a C1483, a partir de los mutantes CX4A-D se crearon cuatro mutantes (serie CX11) en los que se sustituyeron simultáneamente C1483A y R1485C (Tabla R12): R1485C/N1598C (CX11A), R1485C/N1639C (CX11B), R1485C/P1576C (CX11C) y R1485C/A1619C (CX11D). La sustitución de C1483A y R1485C ya se había utilizado en el mutante R1485C/C1608 (CX8A) que no dio muestras de sufrir distorsiones en su estructura. Dentro de los

mutantes de la serie CX11, R1485C/N1639C (CX11B), R1485C/P1576C (CX11C) se expresaron insolubles en *E. coli*; esto concuerda con la expresión insoluble del triple mutante C1483A/C1559A/C1608A (CX5). Sin embargo, los mutantes quintuples R1485C/N1598C (CX11A) y R1485C/A1619C (CX11D) se expresaron solubles y se marcaron con MTSSL. El análisis mediante DEER de estos dos mutantes revela distribuciones de distancias inter-spin muy anchas y con múltiples picos (Figura R21 y Tabla R13), la cuales sugieren que en estas dos proteínas se han causado distorsiones estructurales.

Tabla R12.- Mutantes de $\beta 4$ FnIII-3,4 (1457-1666) de la serie CX11.

Nombre	Localización de las mutaciones			Expresión en <i>E. coli</i>
	FnIII-3	conector	FnIII-4	
CX11A	C1483A, R1485C	C1559A	N1598C , C1608A	Soluble a 15°
CX11B	C1483A, R1485C	C1559A	C1608A, N1639C	Insoluble
CX11C	C1483A, R1485C	C1559A	P1576C , C1608A	Insoluble
CX11D	C1483A, R1485C	C1559A	C1608A, A1619C	Soluble a 15°

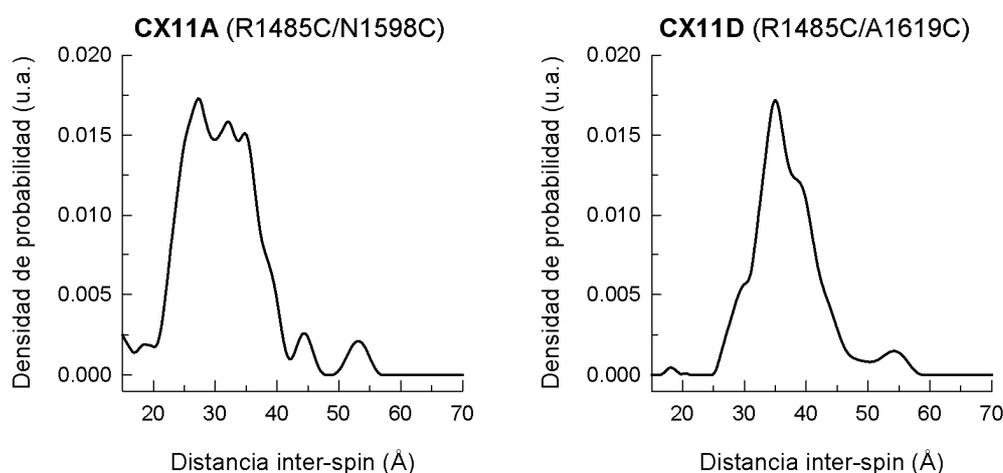


Figura R21.- Distribuciones de distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX11 calculadas a partir de los datos de EPR-DEER.

Tabla R13.- Distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX11.

Nombre	Cys en FnIII-3	Cys en FnIII-4	Distancia (Å)	Comentario
CX11A	R1485C	N1598C	27 / 32 / 35	Distribución ancha, varios picos
CX11D	R1485C	A1619C	35	Distribución ancha

Ante la falta de éxito en la determinación de distancias inter-dominio adicionales empleando mutantes con sólo dos Cys, se emplearon mutantes con tres Cys. Los

resultados obtenidos con los mutantes de las series CX4 y CX11 sugieren que la sustitución C1608A combinada con el marcaje de C1483 con MTSSL o con la sustitución C1483A causa alteraciones estructurales. Por lo tanto se crearon tres cuádruples mutantes (serie CX15) que incluyen la sustitución C1483S y que mantienen C1608 (Tabla R14): R1485C/C1608/A1619C (CX15A), N1523C/N1598C/C1608 (CX15B) y R1504C/N1598C/C1608 (CX15D). Las tres proteínas se expresaron solubles en bacteria y se analizaron mediante DEER.

Tabla R14.- Mutantes de $\beta 4$ FnIII-3,4 (1457-1666) de la serie CX15.

Nombre	Localización de las mutaciones			Expresión en <i>E. coli</i>
	FnIII-3	conector	FnIII-4	
CX15A	C1483S, R1485C	C1559A	A1619C	Soluble a 15°
CX15B	C1483S, N1523C	C1559A	N1598C	Soluble a 15°
CX15D	C1483S, R1504C	C1559A	N1598C	Soluble a 15°

El mutante R1485C/C1608/A1619C (CX15A) muestra tres picos en la distribución de distancias (Figura R22 y Tabla R15). Un pico a ~ 54 Å es similar a la distancia observada en el mutante R1485C/C1608 (CX8A) y se asignó al par 1485-1608. Los otros dos picos, centrados en 38 y 44 Å, se asignaron al par 1485-1619. Por último, la distancia estimada para el par 1608-1616 a partir de la estructura cristalográfica del FnIII-4 es ~ 12 Å, la cual se encuentra por debajo del rango de distancias observadas en experimentos de DEER.

La distribución de distancias determinada para el mutante N1523C/N1598C/C1608 (CX15B) muestra tres picos. Un pico a ~ 29 Å se asignó al par 1523-1608, que ya se había medido en el mutante N1523C/C1608 (CX8B). La distancia 1598-1608 estimada a partir de la estructura cristalográfica del FnIII-4 es ~ 32 Å (promedio para los rotámeros modelados) y corresponde aparentemente al máximo observado a 35 Å. Por lo tanto, el tercer pico a ~ 21 Å se asignó al par 1523-1598.

La distribución de distancias entre centros paramagnéticos determinada mediante DEER para el mutante R1504C/N1598C/C1608 (CX15D) muestra un máximo a ~ 52 Å que se asignó al par 1504-1608, anteriormente determinada empleando el mutante R1504C/C1608 (CX8D). Un segundo máximo ~ 35 Å se asignó al par 1598-1608 tal como se había observado en el mutante CX15B (ver antes). Un doblete a ~ 25 y ~ 29 Å parece corresponder al par 1504-1598.

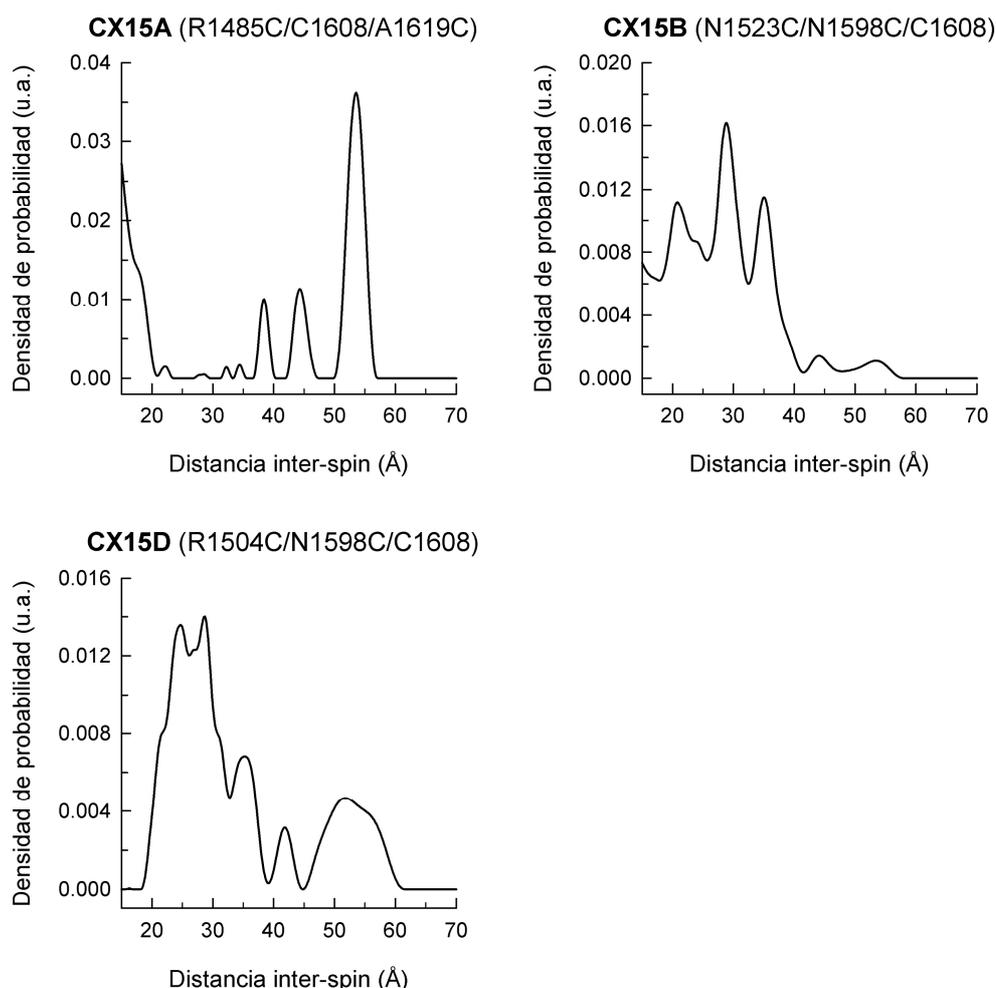


Figura R22.- Distribuciones de distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX15 calculadas a partir de los datos de EPR-DEER.

Tabla R15. Distancias entre centros paramagnéticos en los mutantes FnIII-3,4 de la serie CX15.

Nombre	Cys en FnIII-3	Cys en FnIII-4	Distancia (Å)	Comentario
CX15A	R1485C	C1608-A1619C	(12) ^a	No observada experimentalmente Distribución bimodal
	R1485C	A1619C	38 / 44	
	R1485C	C1608	54	
CX15B	N1523C	N1598C	22	Observada en el mutante CX8B
	N1523C	C1608	29	
		N1598C-C1608	35 (32) ^a	
CX15D	R1504C	N1598C	25 / 29	Observada en el mutante CX8D
		N1598C-C1608	35 (32) ^a	
	R1504C	A1608C	52	

^a Entre paréntesis se indican las distancias promedio modeladas a partir de la estructura cristalográfica del FnIII-4

En conjunto las distancias inter-dominio estimadas a partir del análisis de los mutantes de las series CX8 y CX15 permitieron conocer la posición de la sonda MTSSL unida a dos

residuos del FnIII-4, N1598C y C1608, respecto al FnIII-3. Al conocerse sólo la posición de dos residuos del FnIII-4, existe una ambigüedad en la orientación relativa de los dos dominios FnIII. Todas las posibles rotaciones del FnIII-4 alrededor del eje que pasa por los modelos de la sonda MTSSL unida a N1598 y C1608 cumplen las restricciones de distancias. Con el fin de resolver esta ambigüedad se analizó el ajuste de los modelos de la región FnIII-3,4 a los datos de SAXS. Para ello se generaron 18 modelos que corresponden a rotaciones alrededor del eje N1598-C1608 espaciadas 20°. Seguidamente se analizó el ajuste de cada modelo a los datos de SAXS analizando el valor de χ^2 , que refleja la discrepancia entre la curva de dispersión teórica calculada a partir de las coordenadas de cada modelo y el perfil de dispersión experimental; así mismo se calculó el R_g para cada modelo y se comparó con el determinado a partir de los datos de SAXS (Figura R23A). El mejor ajuste se observó para una única posición de la rotación del FnIII-4. El perfil de dispersión teórico de ese modelo reproduce el perfil de SAXS experimental (Figura R23B). Además la función de distribución de distancias interatómica, $P(r)$, calculada para el modelo se ajusta con tan sólo pequeñas diferencias a la $P(r)$ calculada a partir de los datos experimentales (Figura R23C). En resumen, la combinación de las estructuras cristalográficas de los dominios FnIII, los datos de SAXS y el estudio mediante SDSL-DEER ha permitido elucidar la estructura de la región FnIII-3,4 de $\beta 4$.

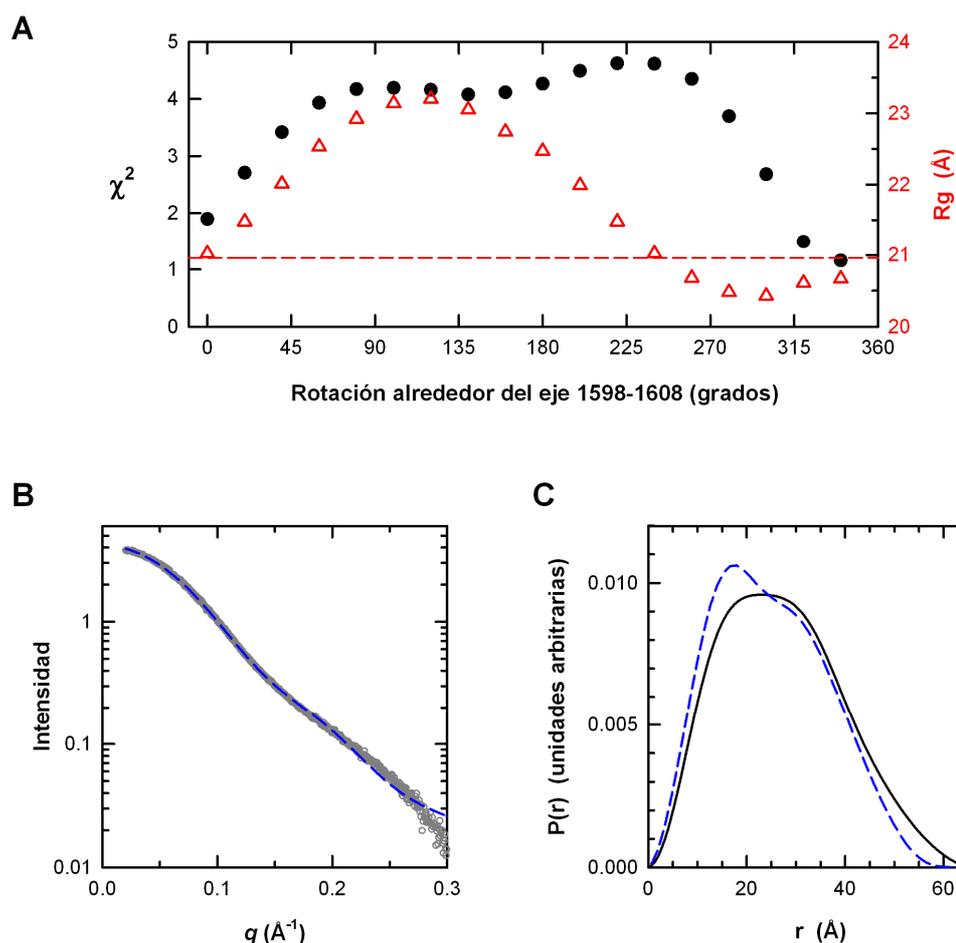


Figura R23.- Evaluación de modelos de FnIII-3,4 obtenidos mediante rotación alrededor del eje N1598-C1608 frente a datos de SAXS. (A) Representación del valor de χ^2 del ajuste del perfil de dispersión calculado para cada modelo a los datos de SAXS en el rango de q entre 0.02 y 0.30 Å⁻¹ (círculos negros) en función del ángulo de giro alrededor del eje que atraviesa la posición estimada de la sonda MTSSL unida a los residuos 1598 y 1608. También se muestra el R_g calculado para cada modelo (triángulos rojos vacíos). La línea discontinua corresponde al R_g determinado experimentalmente a partir de los datos de SAXS. El modelo que mejor se ajusta a los datos experimentales de SAXS corresponde a una rotación de 340° (el origen de rotación fue arbitrario). (B) Comparación del perfil de SAXS calculado para el modelo con mejor ajuste (línea azul discontinua) con el perfil de dispersión experimental (círculos grises). (C) Comparación de la función de distribución de distancias interatómicas del mejor modelo (línea azul discontinua) con la obtenida a partir de los datos de SAXS (línea negra continua).

En la figura R24 se muestra el modelo de la región FnIII-3,4 obtenido mediante la combinación de métodos híbridos y se compara con la estructura de la envuelta obtenida a partir de los datos de SAXS. La zona de contacto entre los dos dominios FnIII está centrada en las hebras A y B del FnIII-3, junto a las que se encuentra el extremo amino, el lazo BC y posiblemente parte del lazo FG₂ del FnIII-4. El extremo carboxilo del FnIII-3 (I1548) y el extremo amino del FnIII-4 (S1572) están relativamente próximos,

siendo la distancia entre ellos ~ 16 Å. La secuencia de 23 aminoácidos que enlaza el FnIII-3 con el FnIII-4 ($^{1549}\text{ESQVHPOSPLCPLPGSAFTLSTP}^{1571}$) es más larga que la secuencia mínima requerida para enlazar los dos dominios FnIII en la estructura. Por lo tanto es posible que esta secuencia conectora se localice en zonas de los modelos reconstruidos a partir del perfil de SAXS que no están ocupadas por los dominios FnIII, por ejemplo en la zona central entre ambos dominios. En resumen, es posible que el conector esté organizado junto o alrededor de la interfase de contacto entre el FnIII-3 y el FnIII-4 y que pueda contribuir a establecer contactos adicionales entre ambos dominios.

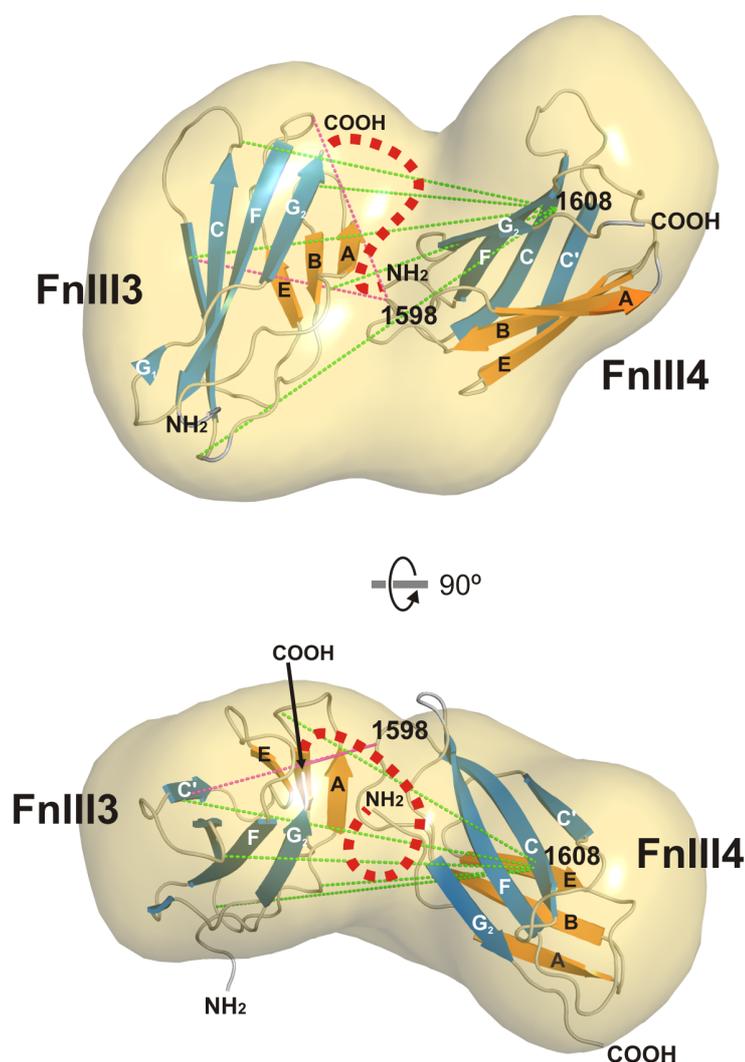


Figura R24.- Estructura de la región FnIII-3,4 de $\beta 4$ obtenida mediante métodos híbridos. Dos vistas ortogonales de la organización de los dominios FnIII-3 y el FnIII-4 superpuestas sobre la envuelta de la estructura promedio de las reconstrucciones obtenidas con el programa DAMMIN a partir del perfil de SAXS. Las líneas discontinuas representan de forma aproximada las restricciones de distancias empleadas para deducir la posición de los residuos del FnIII-4 1598 y 1608, con respecto al FnIII-3. El extremo carboxilo del FnIII-3 y el extremo amino del FnIII-4 se muestran conectados por una línea roja de trazo discontinuo.

El extremo amino del FnIII-3, que enlaza con el segmento conector de β 4, y el extremo carboxilo del FnIII-4, tras el que se extiende la cola C-terminal, se sitúan en el mismo lado de la estructura, en concreto en la parte cóncava de la misma. A pesar de que la distancia entre ambos extremos es ~ 40 Å, su ubicación en el misma cara de la estructura es compatible con una interacción entre el segmento conector y la cola C-terminal.

3.- Caracterización del dominio EC2 de la tetraspanina CD151

Los límites del dominio EC2 de la tetraspanina CD151, se identificaron por comparación con la estructura cristalográfica del dominio EC2 de CD81 (código PDB: 1G8Q) (Kitadokoro et al., 2001a). En la estructura de CD81 el primer y último residuo de CD181 ordenados son F113 y K201, que corresponden a Y112 y H219 en CD151 (Figura R25). Por lo tanto las construcciones del EC2 de CD151 empleadas en este trabajo corresponden a la región 112-219.

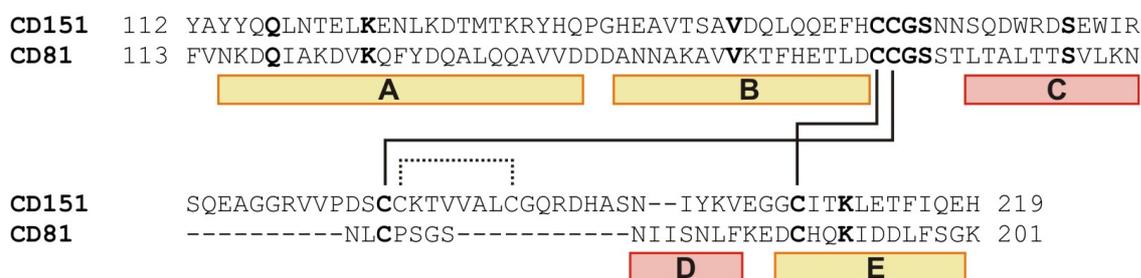


Figura R25.- Alineamiento de las secuencias de los dominios EC2 de CD151 y CD81. Las hélices- α observadas en la estructura de CD81 se indican bajo el alineamiento; las cajas amarillas corresponden a la región constante del EC2, mientras que las rojas a la región variable. Los dos puentes disulfuro que estabilizan la estructura de CD81 se indican por líneas continuas, el puente disulfuro adicional de CD151 se muestran en línea de puntos.

3.1.- Expresión de CD151-EC2 en bacterias

La estructura cristalográfica del EC2 de CD81 se había resuelto empleando proteína producida en *E coli*. Por lo tanto se abordó la producción del EC2 de CD151 en bacteria.

Para ello se crearon construcciones de este dominio en cinco vectores de expresión (Figura R26).

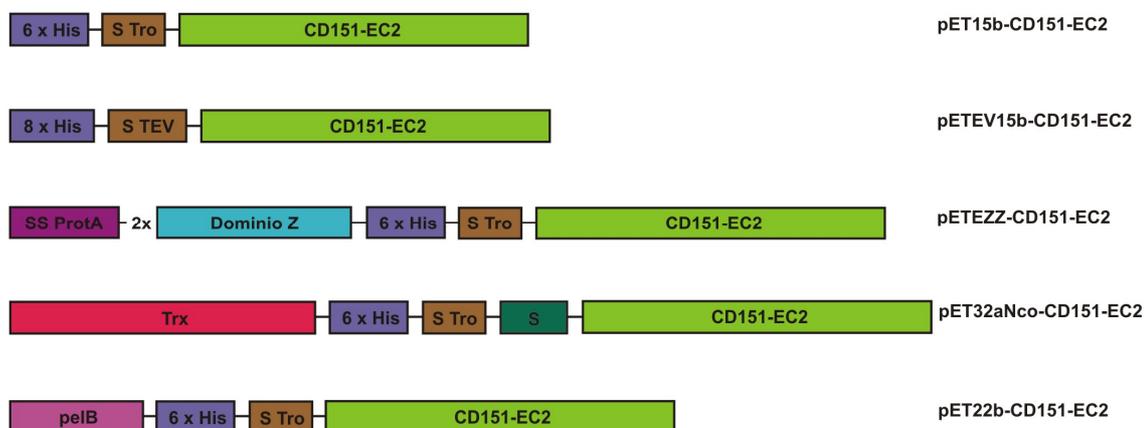


Figura R26.- Construcciones del EC2 de CD151 expresadas en bacteria. Representación esquemática de las proteínas de fusión generadas para la producción del CD151-EC2 en bacteria.

Inicialmente el EC2 se clonó en el vector pET15b y su derivado pETEV15b que codifican proteínas de fusión con una cola de poli-His en el extremo amino. En ambos casos el EC2 aparece en la fracción insoluble. Estos dos vectores dan lugar a proteínas que permanecen en el citoplasma bacteriano, el cual es un entorno reductor que previene la formación de los puentes disulfuro del EC2.

Con el fin de facilitar la formación de los tres puentes disulfuro del EC2 y por lo tanto favorecer la solubilidad de la proteína, se dirigió la secreción del EC2 de CD151 al espacio periplasmático. De hecho el EC2 de CD81 se había expresado empleando el vector de expresión pEZZ18 (GE Healthcare) que incluye una secuencia señal de la proteína A que media la secreción de la proteína de fusión al espacio periplasmático (Kitadokoro et al., 2001a). Repetidos intentos para clonar el EC2 de CD151 en el vector pEZZ18 resultaron infructuosos. La dificultad para obtener una construcción del EC2 de CD151 en pEZZ18 podría estar relacionada con la expresión constitutiva de las proteínas de fusión codificadas por este vector. Con el fin de evitar el posible efecto tóxico del EC2 de CD151 que pudiera impedir la obtención de clones empleando el vector pEZZ18, se construyó el vector pETEZZ, que tiene el gen controlado por un promotor inducible. Este vector codifica una proteína de fusión similar a la generada por pEZZ18 que consiste en una

secuencia señal de la proteína A, dos copias del dominio Z, una secuencia reconocida por trombina y el EC2 de CD151. La proteína de fusión de nuevo se localizó en la fracción insoluble.

Para dirigir la secreción del CD151-EC2 al espacio periplasmático también se clonó el cDNA del EC2 en el vector pET22b, que codifica una proteína de fusión con la secuencia señal pelB en el extremo amino, seguida de una cola de poli-His. La expresión del CD151-EC2 empleando el vector pET22b también da lugar a una proteína que se localiza en la fracción insoluble.

Como alternativa a la secreción al espacio periplásmico se intentó obtener la proteína soluble favoreciendo la formación de los puentes disulfuro del dominio EC2 en el citoplasma bacteriano. Para ello el cDNA del dominio EC2 se clonó en el vector pET32aNco. Este vector es una variante del vector pET32a creada en nuestro laboratorio para adaptar los sitios de restricción. La construcción del CD151-EC2 en pET32aNco codifica una proteína de fusión con Trx en el extremo amino seguida de una secuencia de poli-His, un sitio de trombina y el EC2. Esta construcción se utilizó en combinación con la cepa Origami de *E. coli*. Esta cepa contiene mutaciones en la tiorredoxina-reductasa B y en la glutatión-reductasa; de este modo el sistema permite la formación de puentes disulfuro en el citoplasma. La proteína de fusión producida mediante esta estrategia se encontró de nuevo en la fracción insoluble.

En resumen, todas las estrategias empeladas para la expresión de CD151-EC2 en bacteria dieron lugar a proteínas insolubles.

3.1.1.- Obtención de CD151-EC2 soluble replegado

Ante la imposibilidad de expresar el CD151-EC2 nativo en *E. coli*, se purificó este dominio en condiciones desnaturalizantes en presencia de 6 M urea. Para ello se empleó la construcción de CD151-EC2 en el vector pET15b (proteína de fusión con cola de poli-His en el extremo amino).

Como paso previo al replegado la muestra se incubó con DTT para asegurar que las seis Cys presentes en el dominio EC2 de CD151 estuvieran reducidas. A continuación, la

proteína se replegó por dilución rápida en presencia de los adyuvantes arginina y glutámico, que resultó en una concentración residual de urea 0.2 M. Seguidamente y con el fin de favorecer la formación de los puentes disulfuro del EC2 la proteína se incubó en presencia del par cobre/o-fenantrolina (proporción 1:3) que da lugar a un ambiente oxidante. Tras concentrar la muestra, la cola de poli-His se eliminó mediante digestión con trombina y se purificó mediante cromatografía de exclusión molecular. CD151-EC2 eluye mayoritariamente en un pico con una masa molecular aparente que corresponde a una especie monomérica (Figura R27). Las fracciones del pico monomérico se juntaron y la muestra se concentró hasta una concentración final de ~ 20 mg ml⁻¹. El rendimiento global del replegado y purificación fue de $\sim 20\%$.

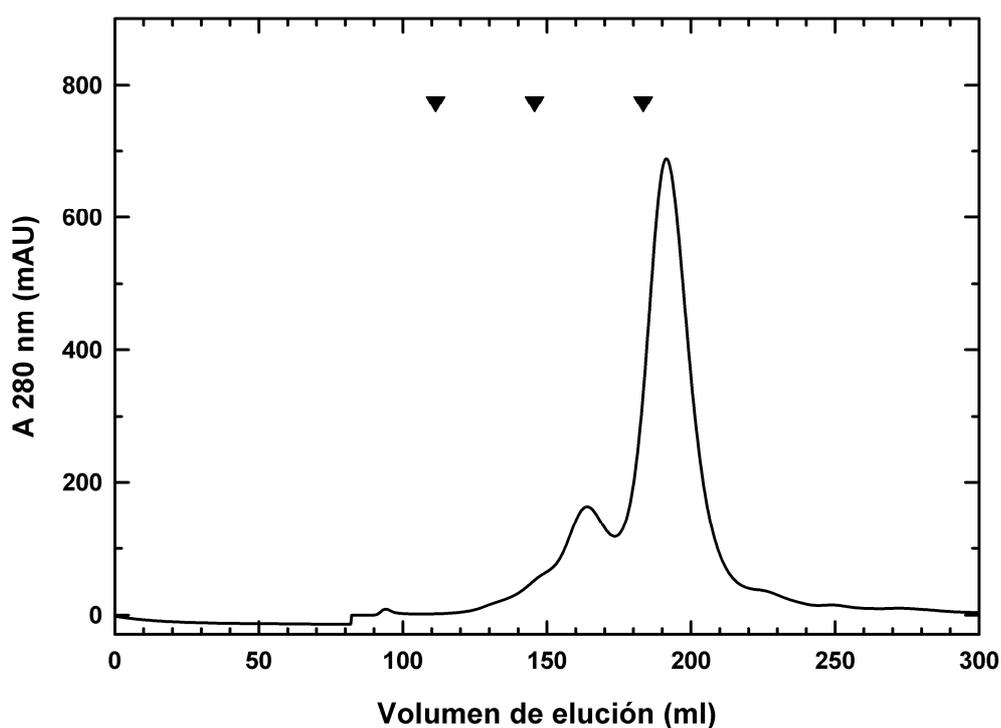


Figura R27.- Cromatografía preparativa de exclusión molecular de CD151-EC2 replegado. (A) Cromatograma correspondiente a la purificación del CD151-EC2 replegado en una columna Sephacryl S200 (26 x 600 mm) equilibrada en 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 50 mM Arg, 50 mM Glu. El flujo fue de 1.2 ml min⁻¹. Como paso previo a la cromatografía la muestra se concentró hasta un volumen de 4 ml, que se cargaron en su totalidad. El pico de elución mayoritario corresponde con una masa molecular aparente de 12 kDa. El pico minoritario corresponde a una masa molecular aparente de 29 kDa. La posición de los picos de elución de proteínas patrón, γ -globulina (158 kDa), ovoalbúmina (44 kDa) y mioglobina (17 kDa), se indica mediante triángulos invertidos.

3.1.2.- Caracterización de CD151-EC2 replegado

3.1.2.1.- Análisis del estado de asociación del CD151-EC2 replegado

El grado de asociación del EC2 de CD151 en solución se analizó con mayor precisión mediante cromatografía de exclusión molecular empleando una columna analítica. El EC2 eluye mayoritariamente en un pico que corresponde a una masa molecular aparente de 15.8 kDa, la cual es similar a la masa molecular teórica

(12766 Da) e indica que el EC2 replegado es un monómero. Este pico mayoritario presenta un pequeño hombro a menor volumen de elución correspondiente a una masa molecular aparente de ~24 kDa, que sugiere la presencia de una pequeña proporción de dímeros (Figura R28). La presencia de una pequeña proporción de CD151-EC2 dimérico puede estar relacionada con la homodimerización de CD151 y otras tetraspaninas *in vivo* (Kovalenko et al., 2004). Se ha propuesto que CD81 homodimeriza a través de una superficie hidrofóbica presente en la estructura de su EC2 (Drummer et al., 2005).

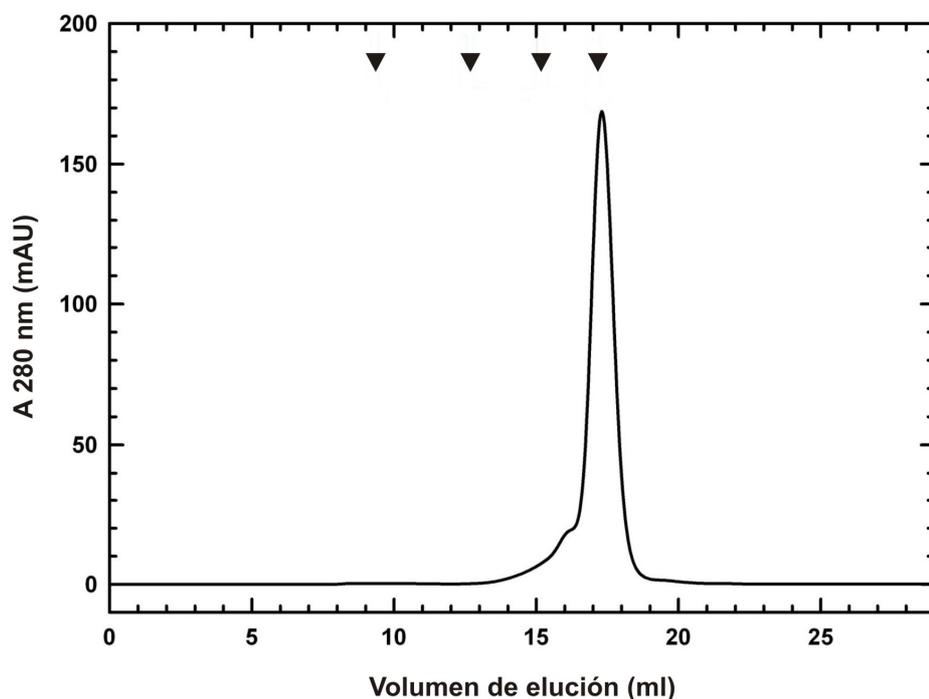


Figura R28.- Análisis del CD151-EC2 replegado mediante cromatografía de exclusión molecular. Se utilizó una columna Superdex 200 (10x300 mm) equilibrada en 20 mM Tris-HCl (pH 7.5), 150 mM NaCl a un flujo de 0.5 ml min⁻¹. Se cargaron 50 µl de CD151-EC2 a 5 mg ml⁻¹. El pico de elución corresponde con una masa molecular aparente de 15.8 kDa. Los triángulos situados en la parte superior indican la posición de elución de proteínas patrón: tiroglobulina (670 kDa), γ -globulina (158 kDa), ovoalbúmina (44 kDa) y mioglobina (17 kDa).

3.1.2.2.- Análisis del estado de oxidación de las Cys del CD151-EC2

El dominio EC2 de CD151 contiene seis Cys (C155, C156, C184, C185, C192 y C208). La valoración con DTNB de grupos tiol libres en condiciones desnaturalizantes (i.e. grupos totales) de la muestra de CD151-EC2 replegado reveló 0.03 ± 0.01 Cys libres por molécula de CD151-EC2. Esto indica que las seis Cys estaban oxidadas. Con el fin de analizar si las Cys forman puentes disulfuro se sometió al CD151-EC2 replegado a proteolisis limitada con tripsina y con quimotripsina. El análisis mediante SDS-PAGE en condiciones reductoras de las muestras digeridas reveló que el CD151-EC2 es sensible a tripsina, pero es relativamente resistente a la digestión con quimotripsina (Figura R29). Sin embargo, cuando la SDS-PAGE se realizó en condiciones no reductoras los fragmentos proteolíticos no se disocian. Esto sugiere que los productos de digestión de CD151-EC2 se mantienen enlazados mediante puentes disulfuro. No se abordó la identificación precisa de los puentes disulfuro por la gran complejidad debida a la proximidad de algunas de las seis Cys en la secuencia del EC2, en particular los pares de Cys contiguas C155-C156 y C184-C185.

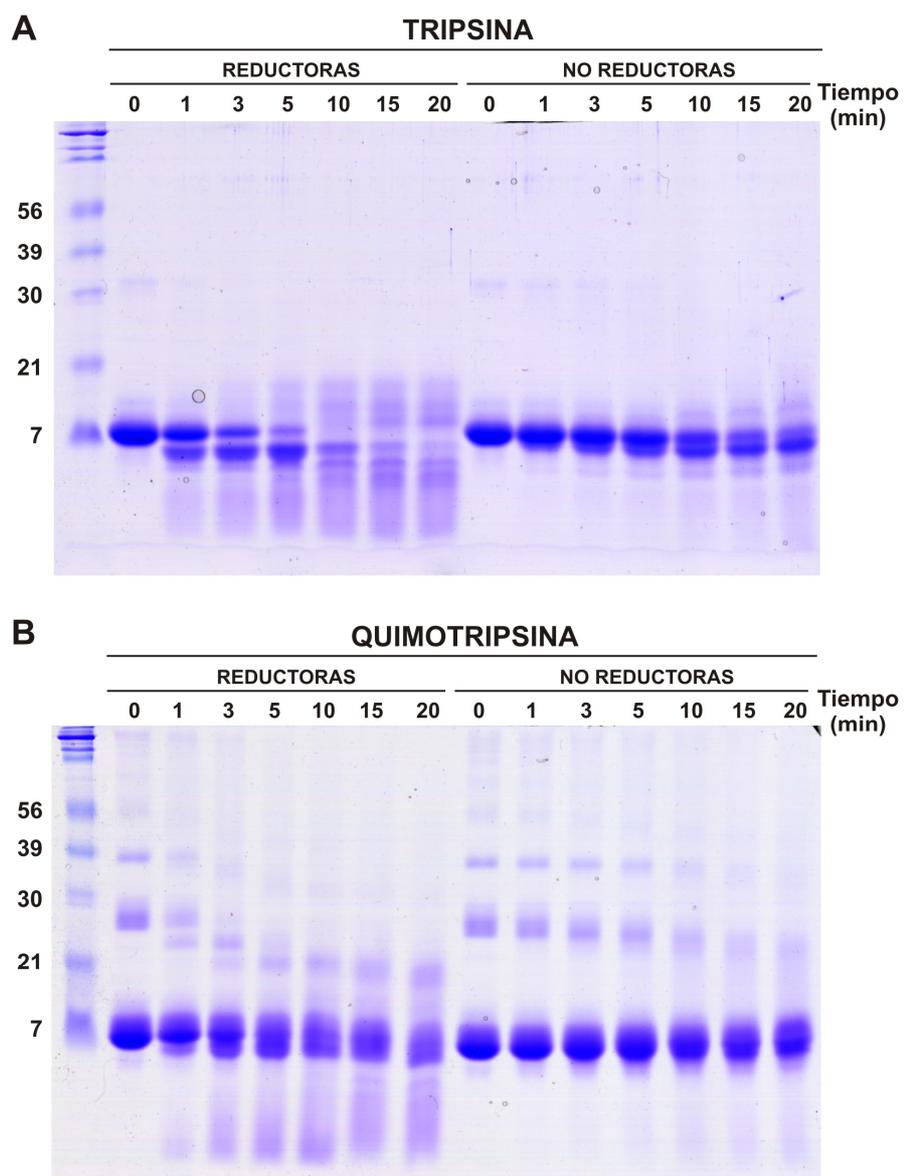


Figura R29.- Proteólisis limitada del CD151-EC2 replegado. El CD151-EC2 se digirió a temperatura ambiente con 0.5 % (p/p) tripsina (A) o quimotripsina (B). A varios tiempos de incubación se tomaron muestras que se analizaron mediante SDS-PAGE empleando geles del 13% de acrilamida tanto en condiciones reductoras como no reductoras.

3.1.2.3.- Contenido de estructura secundaria de CD151-EC2 replegado

La estructura secundaria del EC2-CD151 replegado se analizó mediante dicroísmo circular. El espectro de dicroísmo circular en el ultravioleta lejano del dominio EC2 se caracteriza por un mínimo a ~ 204 nm y un hombro negativo en la región 220-225 nm (Figura R30). El valle cerca de 204 nm sugiere una proporción significativa de estructura desordenada, si bien este mínimo se encuentra desplazado respecto al mínimo

característico de estructura desordenada que aparece a <200 nm, típicamente a ~198 nm. La estimación del contenido en estructura secundaria utilizando el servidor K2D2 indica ~7% α -hélice, ~34% hebra- β y ~59% estructura desordenada. La estimación del contenido de α -hélice es menor de lo esperado por comparación con el EC2 de CD81, el cual por otro lado no contiene láminas- β . Sin embargo, la similitud de secuencia entre CD81 y CD151 es muy baja en la región variable (que en CD81 contiene α -hélices C y D), lo que sugiere que la región variable del EC2 de CD151 podría adoptar un plegamiento con estructura β .

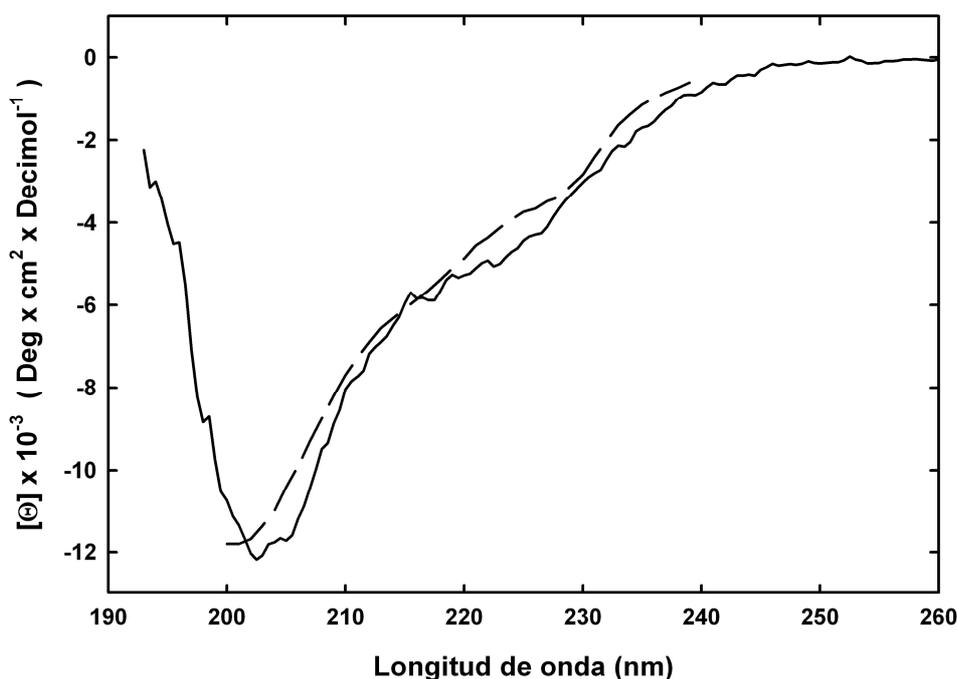


Figura R30.- Espectro de dicroísmo circular de CD151-EC2. La línea de trazo continuo corresponde al espectro de dicroísmo circular en la región del ultravioleta lejano de CD151-EC2 replegado a 2 mg ml⁻¹ en 20 mM NaPi (pH 7.5), 150 mM NaCl medido a 20° C. La línea de trazo discontinuo es el espectro modelo ajustado al espectro experimental, generado por el programa K2D2.

3.1.2.4.- Ensayos de cristalización del CD151-EC2 replegado

Se realizaron ensayos de cristalización del CD151-EC2 replegado mediante el método de difusión de vapor en gota sentada. En estos ensayos se empleó el CD151-EC2 a 17 mg ml⁻¹ en 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 50 mM Arg y 50 mM Glu. Se emplearon 432 soluciones de cristalización que se ensayaron a temperatura ambiente y a 4 °C. No se obtuvieron cristales en ninguna de las condiciones.

3.2.- Expresión de CD151-EC2 en células de insecto

A pesar de que el proceso de replegado del dominio EC2 e CD151 producido en bacteria da lugar a una proteína monomérica, en la que las Cys parecen formar puentes disulfuro y que es estable en solución a alta concentración, es posible que no se haya alcanzado una estructura nativa. Esto podría explicar la falta de éxito en la cristalización del EC2, así como el reducido contenido en estructura secundaria ordenada estimado a partir del espectro de dicroísmo circular.

Con el fin de producir EC2 en estado nativo y evitar el proceso de desnaturalización y replegado, se abordó la expresión heteróloga de CD151-EC2 en células de insecto utilizando baculovirus. Este sistema eucariótico de expresión permite modificaciones post-traduccionales y la formación de puentes disulfuro. Estas características son a priori importantes para lograr el correcto plegamiento de CD151-EC2, el cual contiene tres puentes disulfuro. Por otro lado, Asn159 en el EC2 forma parte de una secuencia consenso de N-glicosilación NX(ST) (Baldwin et al., 2008; Scholz et al., 2009). En resumen, el uso de sistemas de expresión eucariotas puede favorecer el plegamiento, la maduración y la estabilidad del EC2 de CD151.

Se emplearon dos sistemas de baculovirus para la expresión del EC2 en células de insecto: el sistema BaculoGold (BD Biosciences, Pharmingen) utilizando el vector de transferencia pAcGP67A y el sistema Bac-to-Bac (Invitrogen) empleando el vector pFastBac1. En el sistema BaculoGold se realiza una co-transfección en las células de insecto de ADN viral linearizado con el vector de transferencia que contiene el gen del EC2. De esta forma, dentro de las células de insecto se produce la recombinación entre regiones homólogas del vector de transferencia y el ADN viral que da lugar a virus que expresan el gen heterólogo. En cambio, en el sistema Bac-to-Bac la recombinación entre el vector que contiene el gen de interés y el bácmido tiene lugar antes de la transfección en las células de insecto.

Se crearon cuatro construcciones del CD151-EC2 en vectores de expresión con baculovirus utilizando derivados del vector pFastBac1 en el sistema Bac-to-Bac (Figura R31). En todos los casos las proteínas de fusión incorporan péptidos señal en la región N-

terminal que dirige la secreción del dominio EC2 al medio. Se han empleado dos secuencias señal: una derivada de la glicoproteína gp67 y la otra de melitina de abeja (HM). Por otro lado, las construcciones incorporan una secuencia de poli-His que se ha utilizado para la purificación mediante cromatografía de afinidad. Se han creado construcciones con la cola poli-His en el extremo amino o en el extremo carboxilo.

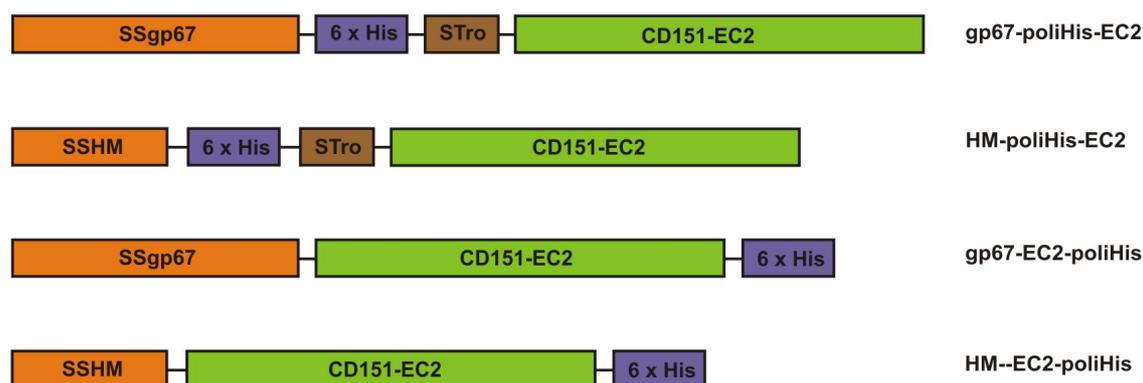


Figura R31.- Construcciones de CD151-EC2 empeladas para la expresión en células de insecto. En las cuatro construcciones se ha combinado la utilización de una secuencia señal (SS), de la proteína gp67 o de HM, con dos posiciones de una secuencia de poli-His (6xHis), en el extremo amino o carboxilo. Las construcciones con la secuencia de poli-His en la región amino incluyen una secuencia reconocida por trombina (STro), mientras que las colas de poli-His en el extremo carboxilo están fusionadas directamente tras el EC2. A la derecha se indica el nombre de cada proteína. Todas las construcciones se realizaron a partir del vector pFastBac1.

El análisis de la expresión de las cuatro construcciones de CD151-EC2 en células de insecto Sf9 se muestra en la 32. Se ha observado sobre-expresión de EC2 con las construcciones que incorporan la cola de poli-His en la región C-terminal: gp67-EC2-poliHis y HM-EC2-poliHis. No observándose diferencias significativas entre el uso de las secuencias señal de gp67 o HM. También se observa expresión, aunque en menor nivel, de la construcción gp67-poliHis-EC2 que emplea la secuencia señal de gp67 y contiene una cola de poli-His N-terminal. Por último, no se observa expresión de la construcción que contienen la secuencia de poli-His en el extremo amino y la secuencia señal HM.

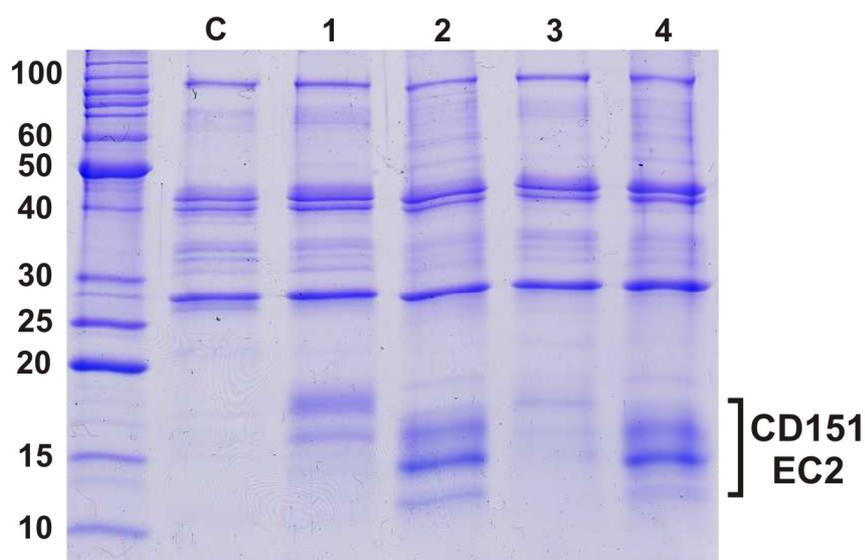


Figura R32.- Análisis de la expresión del dominio EC2 de CD151 en células Sf9 empleando diferentes variantes del vector pFastBac1. Para evaluar la expresión y secreción de CD151-EC2 recombinante se realizó un enriquecimiento de las muestras a partir de 15 ml de medio de cultivo empleando resina Talon quelate de Co^{2+} . Las proteínas unidas se eluyeron en 100 μl . A 16 μl de cada muestra se les añadieron 4 μl de tampón de muestra 4x y se analizaron mediante SDS-PAGE en un gel al 15% de acrilamida. Las proteínas se visualizaron con azul de Coomassie. (C) Control correspondiente a un cultivo de células Sf9 sin infectar. Células infectadas con baculovirus que expresan la construcción (1) gp67-poliHis-EC2, (2) gp67-EC2-poliHis, (3) HM-poliHis-EC2 o (4) HM-EC2-poliHis. La calle de la izquierda (MM) corresponde a los marcadores de peso molecular (Bench Mark, Invitrogen).

3.2.1.- Obtención de CD151-EC2 a partir de células de insecto.

Para la producción del CD151-EC2 a gran escala (3 litros de medio de cultivo) en células de insecto Sf9 se empleó la construcción gp67-EC2-poliHis. El EC2 se purificó en condiciones nativas mediante una primera cromatografía de afinidad en una columna Poros 20MC quelante de Ni^{2+} , tras la cual se realizó una cromatografía de exclusión molecular utilizando una columna Superdex 75 (16 x 600 mm) (Figura R33).

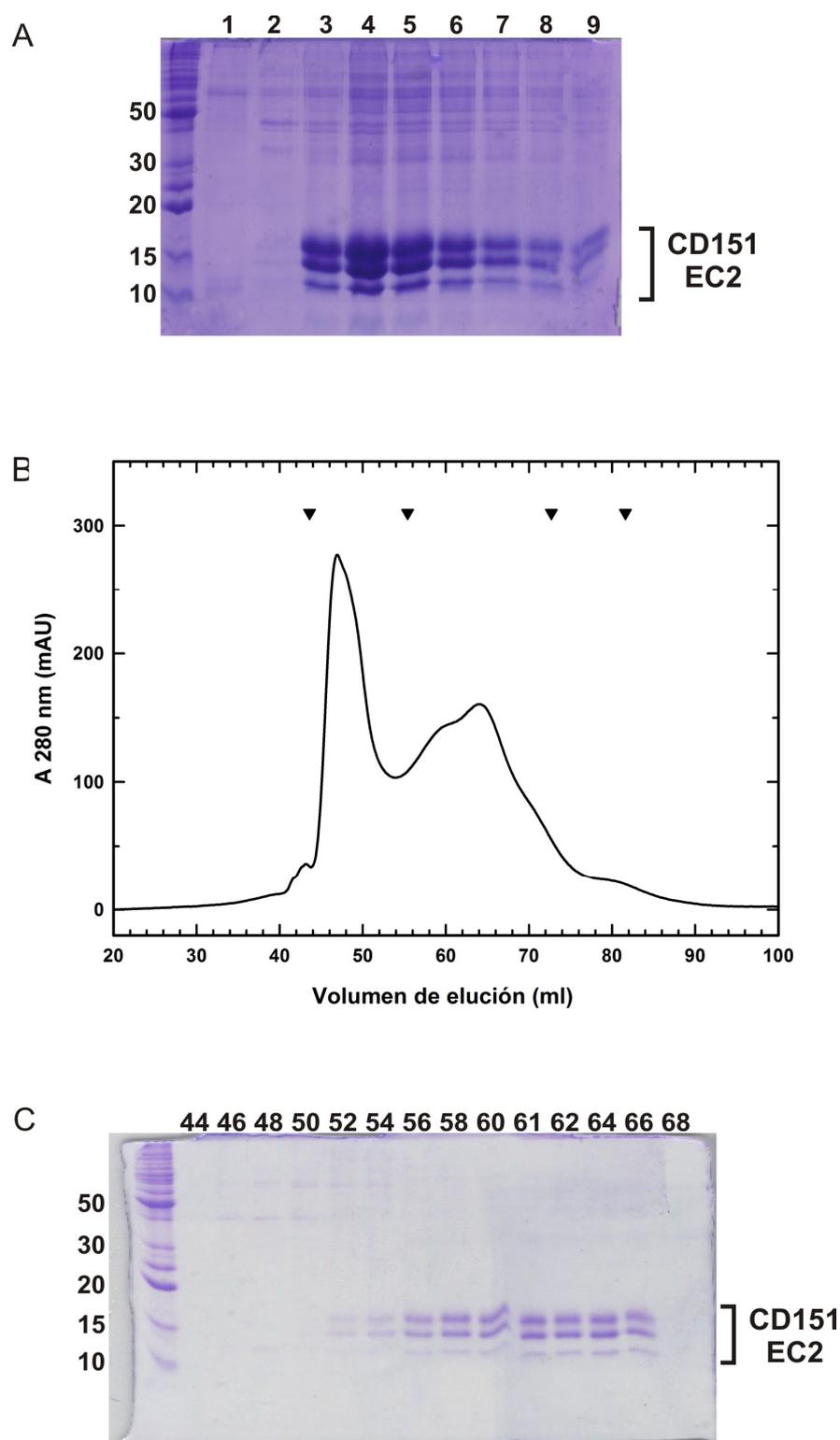


Figura R33.- Purificación del dominio EC2 de CD151 producido en células Sf9. (A) Análisis mediante SDS-PAGE (13% de acrilamida) de algunas de las fracciones recogidas de la cromatografía de afinidad en una columna Porus 20-MC. Se analizaron 9 fracciones eluidas con 20mM Tris-HCl (pH 8.0), 100 mM NaCl, 300 mM imidazol. Las proteínas se visualizaron con tinción de Coomassie. La calle de la izquierda corresponde a los marcadores BenchMark, cuya masa molecular se indican en kDa. (B) Cromatografía de exclusión molecular realizada tras la cromatografía de afinidad. Se utilizó una columna Superdex 75 (16x600 mm) equilibrada en 20

mM Tris-HCl (pH 7.5), 150 mM NaCl a un flujo de 1 ml min⁻¹. Se cargaron 1.5 ml de CD151-EC2 a 3 mg ml⁻¹. El volumen de elución de CD151 corresponde a una masa molecular aparente de 40 kDa. Los triángulos situados en la parte superior indican la posición de elución de los patrones: azul dextrano (2000 kDa), albúmina (67 kDa), quimotripsinógeno A (25 kDa), RNasa A (13.7 kDa). (C) Análisis mediante SDS-PAGE (13% de acrilamida) de fracciones de la cromatografía mostrada en B. Se recogieron fracciones de 1 ml y se analizaron 14 fracciones correspondientes a volúmenes de elución entre 44 y 69 ml. Sobre cada calle se indica el volumen de elución inicial de cada fracción analizada. La calle de la izquierda corresponde a los marcadores BenchMark.

El EC2 producido en células de insecto aparece como una triple banda en electroforesis (Figura R33A,C) que podría estar relacionada con modificaciones postraduccionales de la proteína. El EC2 contiene un residuo, Asn159, que forma parte de una secuencia consenso de glicosilación NX(ST). Por lo tanto es posible que esta heterogeneidad se deba a la glicosilación del EC2. Con el fin de analizar si el EC2 producido en células de insecto está glicosilado, se ha estudiado el efecto de dos endoglicosilasas: endoglicosidasa H (Endo Hf) y N-glicosidasa F (PNGasa F) (Figura R34). El tratamiento con Endo Hf no produce cambios en la movilidad electroforética del EC2 de CD151. Por el contrario, la incubación con PNGasa F reduce a dos el número de bandas de EC2 y aumenta la proporción de la banda de menor masa molecular aparente, la cual parece corresponder al EC2 totalmente deglicosilado.

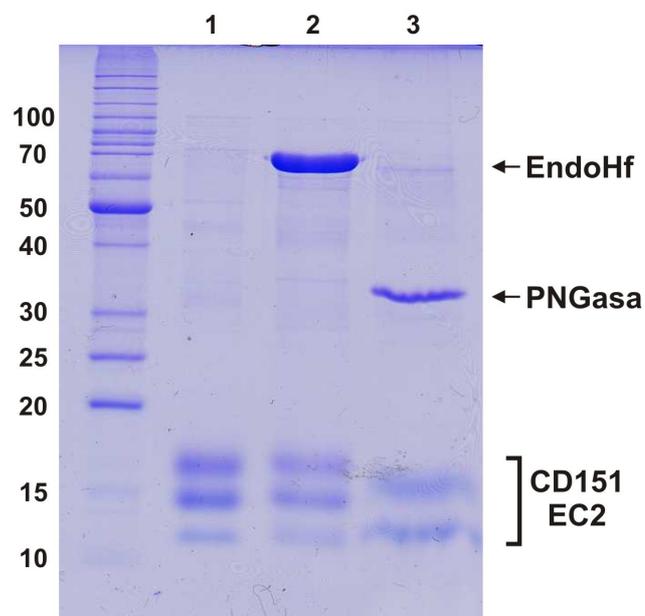


Figura R34.- Deglicosilación *in vitro* del dominio EC2 de CD151 producido en células Sf9. Análisis mediante SDS-PAGE (13% de acrilamida) de CD151-EC2 (1), CD151-EC2 tratado con la endoglicosilasa Endo Hf (2) y CD151-EC2 tratado con la endoglicosidasa PNGasa (3). En la calle de la izquierda se muestran los marcadores BenchMark, cuya masa en kDa se indica.

3.2.2.- Ensayos de cristalización del CD151-EC2 expresado en células de insecto

El CD151-EC2 purificado a partir de cultivos de células Sf9 se sometió a ensayos de cristalización. Se utilizó una solución del EC2 a 5 mg ml^{-1} en 20mM Tris-HCl (pH 7.5), 150 mM NaCl. Los ensayos de cristalización se han realizado en la *Protein Facility* del Nederlands Kanker Instituut (NKI, Ámsterdam) empleando un robot de cristalización Mosquito (TTP Labtech). Se empleó el método de difusión de vapor mezclando 200 nl de solución de CD151-EC2 con un volumen igual de solución de cristalización. Se utilizaron las colecciones de soluciones de cristalización JCSG, PACT (Hanahan, 1983; Newman et al., 2005), Classics y Cryos (Garman and Mitchell, 1996; Jancarik and Kim, 1991). Los ensayos de cristalización se realizaron a 4 y a 20 °C. Las placas se almacenaron y se documentaron periódicamente de forma automática empleando un equipo Rock Imager (Formulatrix) para los ensayos a 4 °C y un equipo Crystal Farm (Nexus Biosystems) para los ensayos a 20 °C. No se obtuvieron cristales del EC2 de CD151.

DISCUSIÓN

1.- Estructura de la región citoplasmática de la integrina $\beta 4$

El conocimiento de la estructura de $\alpha 6\beta 4$ se limitaba al primer par de dominios FnIII (FnIII-1,2) de $\beta 4$ (de Pereda et al., 2009a; de Pereda et al., 1999), una región que es esencial para la interacción con plectina y para la formación de los HD (Niessen et al., 1997). En este trabajo se ha caracterizado la estructura de otras regiones del dominio citoplasmático de $\beta 4$: el dominio Calx- β y el segundo par de dominios FnIII (FnIII-3,4).

1.1.- El dominio Calx- β

1.1.1.- El dominio Calx- β de $\beta 4$ no une Ca^{2+}

La estructura cristalográfica del dominio Calx- β de $\beta 4$ es globalmente similar a las estructuras de los CBDs, dominios de unión de Ca^{2+} , de NCX1. Sin embargo ninguna de las estructuras de $\beta 4$ descritas en este trabajo contienen Ca^{2+} , incluso cuando los cristales se han crecido en presencia de 2 mM CaCl_2 o se han incubado en soluciones con 10 mM CaCl_2 . La ausencia de unión de Ca^{2+} al dominio Calx- β de $\beta 4$ se ha confirmado mediante otros dos métodos: efecto de Ca^{2+} sobre la estabilidad térmica y medida directa de Ca^{2+} unido.

La temperatura de fusión del Calx- β de $\beta 4$ no varía en presencia de hasta 10 mM Ca^{2+} con respecto al valor de T_m observado en ausencia de cationes divalentes. En un sistema en equilibrio en el que un ligando se une al estado nativo pero no al estado plegado, se espera que el ligando estabilice la proteína, resultando en un incremento de T_m (Cimpmperman et al., 2008). Los sitios de unión a Ca^{2+} de los dominios Calx- β de NCX1 incluyen residuos de cuatro regiones distantes en la secuencia (el lazo A`B, el lazo CD, el lazo EF y la región C-terminal). Es razonable pensar que la unión a Ca^{2+} se pierde en el estado desplegado o al menos se reduce severamente la afinidad por Ca^{2+} debido a la desorganización estructural de los sitios de unión. Por lo tanto, la falta de estabilización de $\beta 4$ en presencia de Ca^{2+} sugiere que el estado nativo no une este catión y por lo tanto no es estabilizado. Por último y de forma complementaria, el análisis directo mediante diálisis de equilibrio reveló que el Calx- β no une iones Ca^{2+} de forma apreciable incluso a concentraciones milimolares.

La incapacidad de unión de Ca^{2+} del Calx- β de $\beta 4$ parece deberse a diferencias significativas con los CBDs en la región equivalente a los sitios de unión a Ca^{2+} de los CBDs de NCX1. El lazo EF es más largo en $\beta 4$ que en los CBDs y $\beta 4$ contiene residuos básicos o no cargados en las regiones equivalentes a los residuos ácidos que participan en la coordinación de los iones Ca^{2+} en los CBDs de NCX1. Por ello, nos referimos a esta región de $\beta 4$ como un pseudo-sitio de unión a Ca^{2+} . Este pseudo-sitio de unión a Ca^{2+} de $\beta 4$ se caracteriza por la presencia de dos residuos básicos, R1003 y R1077, que participan en puentes salinos intramoleculares.

1.1.2.- Papel de R1003 y R1077 en el dominio Calx- β

La presencia del residuo básico R1003, en el centro del pseudo-sitio de unión a Ca^{2+} de $\beta 4$ es única comparado con los CBDs de NCX1, que tienen residuos ácidos en posiciones equivalentes (E385 y E561) (Fig R2C). R1003 adopta conformaciones diferentes en cada una de las dos moléculas de la unidad asimétrica, formando puentes salinos con D1038 o con D1104 (Fig R4A-B). Combinando mutagénesis dirigida con análisis de desnaturalización térmica, hemos demostrado que R1003 contribuye a la estabilidad de $\beta 4$. Si se elimina la cadena lateral de R1003 a partir del C^β , en el mutante R1003A, se reduce la estabilidad térmica del dominio Calx- β . El efecto desestabilizador de esta sustitución parece estar relacionado con la pérdida de interacciones de compensación de cargas con D1038 y D1104 y de los enlaces de hidrógeno establecidos por el grupo guanidinio de R1003. La sustitución de R1003 por Glu resulta en una reducción adicional de la estabilidad térmica, que parece estar causada por una repulsión electrostática del nuevo grupo carboxilato con los de D1038 y D1104.

R1077 en $\beta 4$ juega un papel similar al de K585 en el CBD2 de NCX1 (Fig R4D-E). En ausencia de Ca^{2+} el CBD2 del NCX1 mantiene su estructura terciaria global (Hilge et al., 2006) y los sitios de unión a Ca^{2+} adoptan una conformación ordenada que facilita la unión rápida de Ca^{2+} , en la que K585 establece puentes salinos intramoleculares con D552 y E648 (Besserer et al., 2007). Cuando K585 se sustituye por Glu, el CBD2 mutante muestra una reducción moderada en la afinidad aparente por Ca^{2+} , pero experimenta una desorganización estructural general cuando se elimina el Ca^{2+} (Besserer et al., 2007;

Hilge et al., 2006). En contraste, la eliminación de Ca^{2+} del CBD1 de NCX1 induce grandes cambios estructurales globales y la desorganización de su estructura terciaria (Hilge et al., 2006; Johnson et al., 2008; Nicoll et al., 2007; Ottolia et al., 2004). El residuo equivalente a R1077 en CBD1 es E454. La sustitución E454K resulta en la estabilización global de la estructura terciaria del CBD1 en ausencia de Ca^{2+} y en una reducción en la afinidad aparente por Ca^{2+} (Hilge et al., 2006). En resumen, R1077 contribuye a la estabilidad del dominio Calx- β de $\beta 4$ por compensación de la carga negativa de D1038, de forma similar a como K585 estabiliza el CBD2 en ausencia de Ca^{2+} .

1.1.3.- Organización del dominio Calx- β en el contexto de la región citoplasmática de $\beta 4$

El dominio Calx- β y los cuatro dominios FnIII son los únicos dominios globulares identificados en la región citoplasmática de $\beta 4$ y en conjunto estos cinco dominios corresponden a la mitad de la secuencia de esta región. Estos dominios pertenecen a la superfamilia de las inmunoglobulinas, sus tamaños son similares y se caracterizan por tener los extremos N y C en los polos opuestos del eje longitudinal de la estructura. Esto sugiere que la región citoplasmática de $\beta 4$ podría tener una organización de tipo bolas en cordel, si bien las regiones FnIII-1,2 y el FnIII-3,4 adoptan estructuras aparentemente rígidas con reducida variabilidad en la orientación de los dominios individuales (ver abajo). Los dominios FnIII-1 y FnIII-2 están unidos por una secuencia muy corta y forman una estructura alargada aparentemente rígida, tal como se ha observado tanto en estructura cristalográfica (de Pereda et al., 2009a; de Pereda et al., 1999) como en solución (Chacon et al., 2000). El dominio Calx- β está conectado al FnIII-1 por una secuencia de ~20 residuos de longitud que a priori podría actuar como un lazo flexible. La longitud del segmento de unión entre el Calx- β y el FnIII-1 es similar a la que conecta el FnIII-3 y el FnIII-4. Puesto que la región FnIII-3,4 adopta una estructura aparentemente ordenada (ver abajo), no podemos descartar que el Calx- β también tenga una orientación aproximadamente estable con respecto al FnIII-1. En ese caso, la región del

pseudo-sitio de unión a Ca^{2+} de $\beta 4$ podría participar de la interfase de interacción con el FnIII-1.

1.1.4.- Posible función del dominio Calx- β

Hasta la fecha, no se ha asignado una función específica al dominio Calx- β de $\beta 4$.

Por un lado se ha implicado al dominio Calx- β en el procesamiento proteolítico de $\beta 4$. La región citoplasmática de $\beta 4$ es digerida *in vivo* de forma dependiente de Ca^{2+} y calpaína digiere a $\beta 4$ *in vitro* (Giancotti et al., 1992). En base a estos resultados se ha sugerido que el dominio Calx- β podría estar involucrado en la regulación del proceso proteolítico de la subunidad $\beta 4$ (May and Ponting, 1999). No obstante, calpaína digiere $\beta 4$ en dos sitios que probablemente se localizan en el FnIII-2 y en el segmento conector (Giancotti et al., 1992) pero no en el dominio Calx- β . Además nuestros datos indican que el dominio Calx- β no une Ca^{2+} ; por lo que no parece actuar como un sensor de Ca^{2+} . Por tanto, es improbable que el dominio Calx- β esté implicado en la regulación del procesamiento de $\beta 4$ inducido por Ca^{2+} .

La región citoplasmática de $\beta 4$ actúa como una plataforma de interacción proteína-proteína tanto en la señalización como en la adhesión celular. De los cinco dominios globulares de la región citoplasmática de $\beta 4$, los cuatro dominios FnIII median la asociación con otras proteínas, incluidos los componentes hemidesmosomales (Geerts et al., 1999; Koster et al., 2003; Koster et al., 2004b) y moléculas de señalización (Bertotti et al., 2006; Dans et al., 2001a). Por lo tanto, es posible que el dominio Calx- β de $\beta 4$, el cual está estructuralmente relacionado con los dominios FnIII, esté también implicado en las interacciones proteína-proteína. La quinasa Fyn de la familia Src se asocia con la integrina $\alpha 6\beta 4$ en respuesta a la estimulación por EGF, lo que da lugar a la fosforilación de la subunidad $\beta 4$. La asociación con Fyn requiere la región de $\beta 4$ próxima a la transmembrana, residuos 854 -1183, los cuales contienen el dominio Calx- β (residuos 989-1107) (Mariotti et al., 2001). Por tanto, el dominio Calx- β podría participar en interacciones proteína-proteína durante la señalización de $\alpha 6\beta 4$.

Es probable que los sitios potencialmente funcionales (ej: sitios de interacción con proteínas, intra o intermoleculares) en el dominio Calx- β de $\beta 4$ correspondan a áreas que no están conservadas en otros dominios Calx- β . El 45% de la secuencia del dominio de $\beta 4$ es idéntica a la secuencia del CBD1 o del CBD2 de NCX1. La mayoría de las regiones de $\beta 4$ que no están conservadas en ninguno de los dominios del NCX1 están en el lazo A'B, la primera mitad de la hebra B, el lazo EF y el lazo FG. Los lazos A'B y EF se localizan adyacentes uno al otro en uno de los extremos del sándwich- β y forman parte del pseudo-sitio de unión a Ca^{2+} . Es razonable que este extremo del dominio Calx- β haya evolucionado para realizar funciones diferentes en los NCXs y en $\beta 4$. La extensión del lazo EF de $\beta 4$ (¹⁰⁶²LELQEVDSLLRGRQ¹⁰⁷⁵), que es mayor que la de otros dominios Calx- β y no presenta similitud de secuencia con ellos, es de interés. En el plegamiento de inmunoglobulinas, las secuencias específicas de función se localizan frecuentemente en los lazos. Por ejemplo, el décimo dominio FnIII de fibronectina contiene un motivo RGD de unión a integrinas en el lazo flexible FG (Dickinson et al., 1994). Por tanto, proponemos que la superficie apical formada por los lazos A'B y EF del dominio Calx- β de $\beta 4$ puede constituir un sitio funcional.

1.2.- La región FnIII-3,4

1.2.1.- Estructura de la región FnIII-3,4

La estructura de la región FnIII-3,4 de $\beta 4$ obtenida mediante métodos híbridos ha revelado que esta región adopta una conformación compacta, aparentemente estable. La secuencia de 23 aminoácidos rica en Pro (¹⁵⁴⁹ESQVHPQSPLCPLPGSAFTLSTP¹⁵⁷¹) que conecta estos dos dominios FnIII, parece contribuir a la conformación compacta observada permitiendo la correcta orientación de los FnIIIs y posiblemente contribuyendo a la interfase de contactos inter-dominio. La estructura de la región FnIII-3,4 contrasta con la de la región FnIII-1,2, la cual muestra una organización lineal tanto en las estructuras cristalográficas (de Pereda et al., 2009a; de Pereda et al., 1999) como en solución (Chacon et al., 1998). La estructura alargada de FnIII-1,2 se debe, al menos

en parte, a que el FnIII-1 está conectado con el FnIII-2 por una secuencia muy corta de ~5 aminoácidos.

En resumen, a pesar de que los dominios Calx- β y FnIIIs de β 4 se encuentran dispuestos a lo largo de la secuencia unidos por regiones aparentemente no globulares, como las cuentas de un collar, la región citoplasmática de β 4 parece estar compuesta de segmentos multi-dominio con flexibilidad moderada. Es posible que globalmente la región citoplasmática de β 4 adopte un número limitado de conformaciones definidas tal como se ha observado en las estructuras de las regiones extracelulares de integrinas, las cuales están formadas por múltiples dominios (Campbell and Humphries, 2011).

1.2.2.- Interacciones de FnIII-3,4 con proteínas de los HDs

La región FnIII-3,4 de β 4 está implicada en la interacción con BPAG1e y BP180.

BPAG1e interacciona con la región citoplasmática de β 4 (Hopkinson and Jones, 2000). Mediante experimentos de doble-híbrido en levaduras se ha identificado que la interacción se produce entre el extremo amino de BPAG1e y la región 1436-1667 de β 4 (Koster et al., 2003), que corresponde a FnIII-3,4 y a la región C-terminal del CS (Figura I9). En nuestro grupo estamos analizando la interacción entre β 4 y BPAG1e *in vitro* mediante anisotropía de fluorescencia empleando fragmentos de BPAG1e marcados con sondas fluorescentes (Gómez, Manso, Alonso y de Pereda, resultados sin publicar). Nuestros datos indican que el CS de β 4 no es necesario para la unión a BPAG1e y que los dominios FnIII-3 o FnIII-4 aislados no se unen a BPAG1e. En resumen, la unión de BPAG1e a β 4 requiere la contribución simultánea del tercer y cuarto dominios FnIII. Esto sugiere que la disposición relativa del FnIII-3 y el FnIII-4 elucidada en este trabajo puede ser necesaria para la formación de la interfase de unión a BPAG1e.

Se pueden proponer varios modos de unión de β 4 a BPAG1e (Figura D1). La zona de unión a BPAG1e podría localizarse en la interfase de unión entre los dos dominios FnIII, la cual podría incluir la secuencia que une el FnIII-3 con el FnIII-4 (Figura D1A). Este modo de unión sería análogo a la forma en que la hormona de crecimiento humana (hGH) se une a su receptor (de Vos et al., 1992). La región extracelular del receptor de hGH

consiste en dos dominios FnIII unidos por una secuencia de cuatro aminoácidos. En la estructura del complejo una molécula de hGH se une a dos moléculas del receptor; aunque las zonas de hGH que contactan con cada receptor son diferentes, cada receptor se une de forma similar a hGH a través de una zona de bisagra en la interfase entre los dos dominios FnIII (Figura D1D). Alternativamente, el FnIII-3 y el FnIII-4 podrían albergar sitios independientes de baja afinidad que se unan a diferentes regiones de BPAG1e (Figura D1B). Una tercera posibilidad sería que BPAG1e se una principalmente a uno de los dominios FnIII de $\beta 4$, pero que el otro dominio FnIII sea necesario para la correcta organización del sitio de unión, por ejemplo estabilizando la estructura de la secuencia que une el FnIII-3 y el FnIII-4, la cual podría participar en la unión a BPAG1e (Figura D1C).

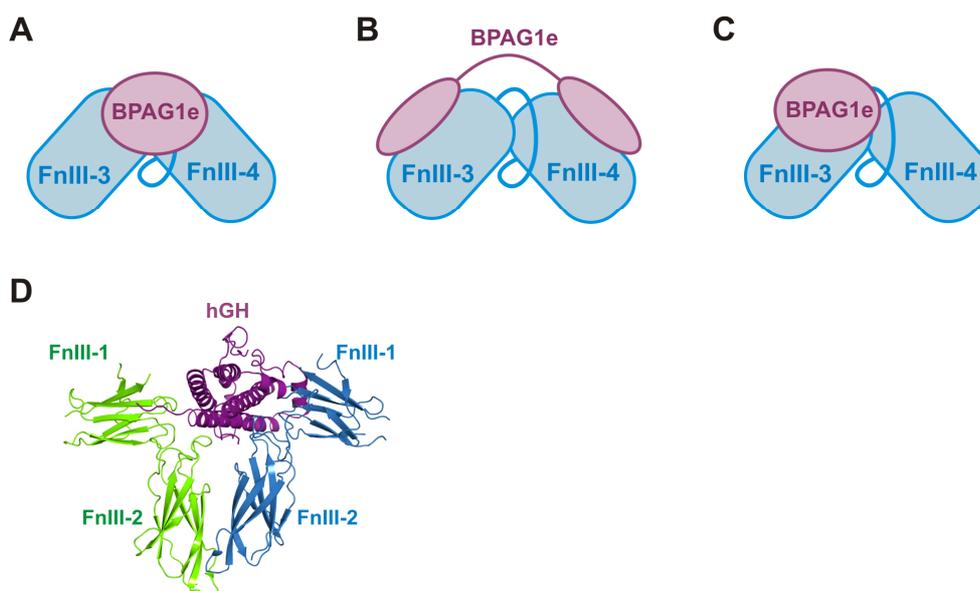


Figura D1.- Posibles modos de unión de BPAG1e a la región FnIII-3,4 de $\beta 4$. (A-C) Representación esquemática de tres posibles modos de unión en los que BPAG1e se une: en la zona de la interfase entre los dos dominios FnIII y establece contactos con ambos dominios (A), en dos subsitios situados en cada dominio FnIII (B), o a uno de los dominios FnIII y al conector que enlaza el FnIII-3 con el FnIII-4 (C). (D) Estructura del complejo formado por la hGH y dos receptores (código PDB 3HHR). hGH se une en la interfase entre los dos dominios FnIII de cada receptor, de forma similar al modelo propuesto en (A).

En un estudio de mutagénesis que se está realizando en nuestro grupo se han identificado cambios puntuales de residuos de la superficie del FnIII-3 que comprometen la unión de BPAG1e, pero hasta la fecha no se ha observado ninguna mutación puntual

en el FnIII-4 que comprometa esta interacción. Por lo tanto es probable que la unión de BPAG1e a $\beta 4$ se asemeje al tercer modelo propuesto, si bien no se puede descartar que el FnIII-4 contribuya a la unión mediante interacciones de largo alcance, por ejemplo electrostáticas.

La región citoplasmática de BP180 se une a la subunidad $\beta 4$ (Aho and Uitto, 1998a; Borradori et al., 1997; Schaapveld et al., 1998). El dominio FnIII-3 de $\beta 4$ es suficiente para la interacción con BP180 (Koster et al., 2003). Puesto que el FnIII-4 no es necesario para la interacción con BP180 es razonable que la zona de unión a BP180 en el FnIII de $\beta 4$ esté alejada tanto de la interfase entre el FnIII-3 y el FnIII-4 como de la secuencia que une ambos dominios.

1.2.3.- Interacciones de FnIII-3,4 con dominios PTB y SH2

Se ha descrito que la región FnIII-3,4 participa, a través de la fosforilación en tirosinas, en la interacción con dominios PTB y SH2.

El dominio PTB de la proteína adaptadora Shc se une a $\beta 4$ fosforilada en tirosinas. Y1526 en el FnIII-3 es el sitio primario de unión del dominio PTB, siendo Y1642 (en el FnIII-4) un sitio secundario (Dans et al., 2001b). Los dominios PTB reconocen motivos NPxY o de forma más general NxxY (donde x es cualquier aminoácido) en los que la fosforilación de la Tyr puede ser o no necesaria para la interacción (Uhlik et al., 2005); el dominio PTB de Shc reconoce secuencias con la Tyr fosforilada. En las secuencias con fosfo-Tyr unidas a dominios PTB de forma canónica se pueden distinguir dos partes: los aminoácidos que preceden al motivo NPxY forman una hebra- β que se sitúa en un lateral y prolonga la lámina- β central del dominio PTB, por otro lado la secuencia NPxY forma un giro reverso (Figura D2A). En $\beta 4$, Y1526 y Y1642 ocupan respectivamente posiciones equivalentes en el FnIII-3 y el FnIII-4. Estas Tyr se localizan al comienzo de la hebra F, su cadena lateral forma parte del núcleo hidrofóbico que estabiliza el sándwich- β (Figura D2B) y los residuos que preceden a estas Tyr forman el lazo E-F. La conformación de las secuencias de $\beta 4$ en torno a Y1526 y Y1642 es diferente a la observada en los complejos canónicos de ligando NPxY unidos a dominios PTB y la participación de estas secuencias de $\beta 4$ en la

estructura de los dominios FnIII es incompatible con una interacción canónica con dominios PTB.

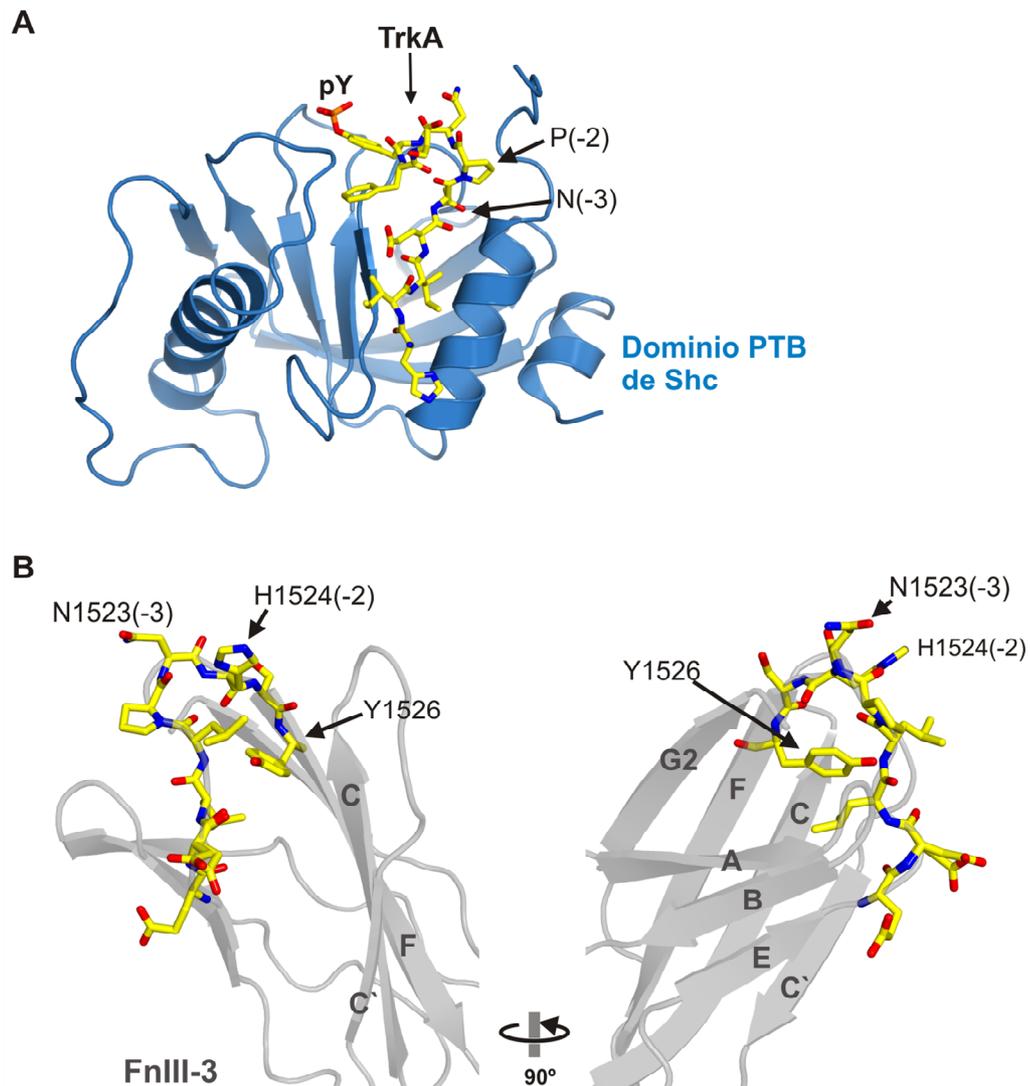


Figura D2.- Comparación de la estructura de Y1526 en el FnIII-3 con la estructura de un ligando con pTyr unido al dominio PTB de Shc. (A) Estructura de un fragmento del receptor TrkA del factor de crecimiento neuronal que contiene una secuencia NPxpY (representación de bastones) unido al dominio PTB de Shc (azul) (Código PDB: 1SHC) (Zhou et al., 1995). (B) Dos vistas ortogonales de la estructura del FnIII-3 de $\beta 4$. Se muestran como bastones Y1526 y varios residuos que le preceden, el resto del dominio FnIII-3 se muestra en representación de cintas (gris).

En respuesta a la activación del receptor del factor de crecimiento de hepatocitos Met, $\beta 4$ recluta la fosfatasa de tirosinas Shp2 (también conocida como PTPN11 o Syp). La unión a Shp2 requiere la fosforilación de $\beta 4$ en Y1257 (en el FnIII-2), Y1440 (en el CS) y Y1494 (en el FnIII-3). Shp2 contiene dos dominios SH2 (N-SH2 y C-SH2) y se ha propuesto

que Shp2 se une a través de estos dominios SH2 a las fosfo-tirosinas de $\beta 4$ (Bertotti et al., 2006). Las regiones de $\beta 4$ en torno a Y1257 y Y1494 se asemejan a la secuencia consenso de clase I de unión al N-SH2 de Shp2 y a la secuencia consenso de unión al C-SH2 (Imhof et al., 2006). En las estructuras de complejos del N-SH2 de Shp2 unido a ligando con fosfo-Tyr, el fosfo-péptido adopta una conformación extendida; además de las interacciones específicas de la pTyr, residuos adyacentes en las posiciones -1, +1, +3 y +5 establecen contactos adicionales con el SH2 (Lee et al., 1994). En $\beta 4$ Y1257 y Y1494 ocupan respectivamente posiciones equivalentes en los dominios FnIII-2 y FnIII-3. Estas Tyr se localizan en la hebra C y su cadena lateral forma parte del núcleo hidrofóbico encerrado en el sandwich- β (Figura D3). Por lo tanto, el contexto estructural de Y1257 y Y1494 en el FnIII-2 y el FnIII-3 no es compatible con la interacción canónica con dominios SH2.

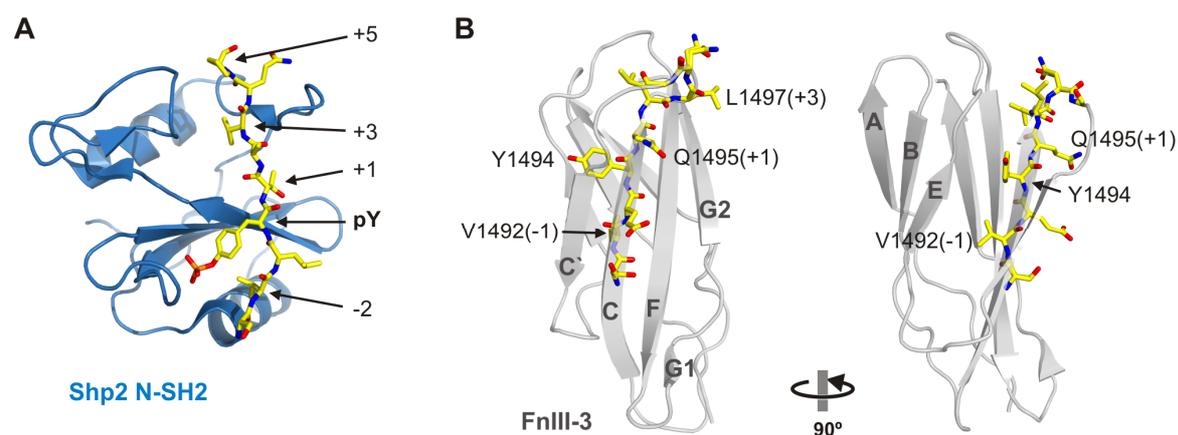


Figura D3.- Comparación de la estructura de Y1494 en el FnIII-3 con la estructura de un ligando con pTyr unido al dominio N-SH2 de Shp2. (A) Estructura de un fragmento del receptor del factor de crecimiento derivado de plaquetas (representación de bastones) unido al dominio N-SH2 de Shp2 (representación de cintas) (Código PDB: 1AYA) (Lee et al., 1994). (B) Dos vistas ortogonales de la estructura del FnIII-3 de $\beta 4$. Y1494 y residuos adyacentes se muestran como bastones, el resto del dominio se muestra en representación de cintas.

En resumen, los residuos de Tyr del FnIII-3 y el FnIII-4 de $\beta 4$ que se han implicado en la interacción con dominios PTB o SH2 se localizan en el interior del sándwich- β que forman los dominios FnIII. Las estructuras del FnIII-3 y el FnIII-4 determinadas en este trabajo sugieren que estas Tyr de $\beta 4$ carecen de la libertad conformacional necesaria para unirse a dominios PTB o SH2 de forma canónica. Por último, no podemos descartar

que la fosforilación de $\beta 4$ en Tyr desestabilice la estructura de los FnIII y cause un cambio conformacional que permita la unión de dominios PTB o SH2 a $\beta 4$.

1.2.4.- Contribución de FnIII-3,4 a la organización de otras regiones de $\beta 4$

La cola C-terminal de $\beta 4$ interacciona con el CS en ensayos de dos híbridos en levadura y en experimentos de *far-western blot* (Koster et al., 2004a; Reznicek et al., 1998; Schaapveld et al., 1998). Recientemente se ha observado mediante FRET y FLIM que la región C-terminal está próxima al CS tanto en lisados celulares como en queratinocitos vivos (Frijns et al., 2012). En conjunto se cree que existe un contacto directo entre el CS y la cola C-terminal, las cuales están separadas en la secuencia de $\beta 4$ por la región FnIII-3,4. En nuestra estructura de FnIII-3,4 el extremo amino del FnIII-3, que enlaza con el CS, se encuentra a ~ 40 Å del extremo carboxilo del FnIII-4 donde comienza la cola C-terminal. A pesar de esta separación, ambos extremos se localizan en el mismo lado de la estructura. Por lo tanto la estructura del FnIII-3,4 es compatible con una interacción directa entre el CS y la cola C-terminal. Es posible que FnIII-3,4 actúe como un soporte que aproxime el CS a la cola C-terminal y favorezca la interacción entre esas dos regiones de $\beta 4$ (Figura D4). También sería posible que el CS y la cola C-terminal formen un único dominio estructural en el que la región FnIII-3,4 está insertada. Se desconoce la estructura del CS y de la cola C-terminal de $\beta 4$. Además, ninguna de estas dos regiones muestra similitud con otras proteínas, por lo que se carece de estructuras homólogas que faciliten la predicción de la estructura y de su orientación relativa. No ha sido posible producir de forma soluble en bacteria fragmentos de $\beta 4$ que contengan el CS y la cola C-terminal (por ejemplo FnIII-1,2-CS-FnIII-3,4-C-terminal o CS-FnIII-3,4-C-terminal) (resultados no mostrados). Una posible solución para el estudio de esta amplia región de $\beta 4$ sería el uso de sistemas de expresión eucariotas.

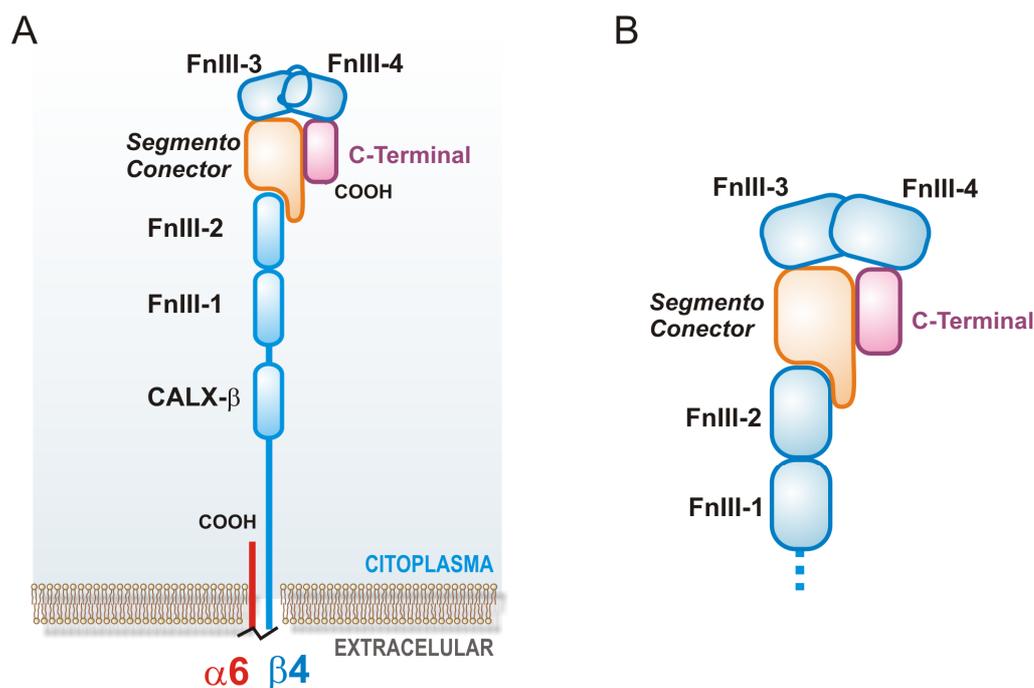


Figura D4.- Papel de la región FnIII-3,4 en la organización del dominio citoplasmático de β4. (A) Representación esquemática de la región citoplasmática de α6β4. (B) Detalle de la organización de la región FnIII-1,2-CS-FnIII-3,4-C-terminal de β4. La región FnIII-1,2 presenta una estructura alargada. La secuencia N-terminal del segmento conector (CS) está estructuralmente organizada sobre el FnIII-2 extendiendo las láminas-β del FnIII. La cola C-terminal se cree que interacciona directamente con el CS. La estructura compacta del FnIII-3,4 favorecería la proximidad entre el CS y la cola C-terminal.

1.2.5.- Potencial participación de FnIII-3,4 en cambios conformacionales en β4

La interacción entre β4 y plectina es un paso esencial en el ensamblaje de los HDs y facilita la posterior incorporación de otros componentes, incluyendo BPAG1e y BP180 (Wilhelmsen et al., 2006) (Figura I12). Se ha propuesto que la unión de plectina podría inducir un cambio conformacional en β4 que expondría sitios de unión a las otras proteínas. Inicialmente se creía que dicho cambio conformacional podría implicar una apertura de la interacción entre el CS y la cola C-terminal; sin embargo se ha observado que estas dos regiones permanecen próximas en queratinocitos vivos (Frijns et al., 2012). No se puede descartar que exista otro tipo de reorganización de β4 causada por plectina, la cual facilite la unión de BPAG1e.

Cuando el ABD de plectina se une a β4 induce un cambio en la estructura de la región inicial del CS (Figura I11) (de Pereda et al., 2009a). Puesto que el CS está en contacto con

la cola C-terminal y que la región FnIII-3,4 se sitúa entre ambos, es posible que el cambio en la organización de la región N-terminal del CS propague cambios en todo el segmento CS-FnIII-3,4-C-terminal, los cuales podrían inducir una conformación del FnIII-3,4 que reclute BPAG1e (Figura D5).

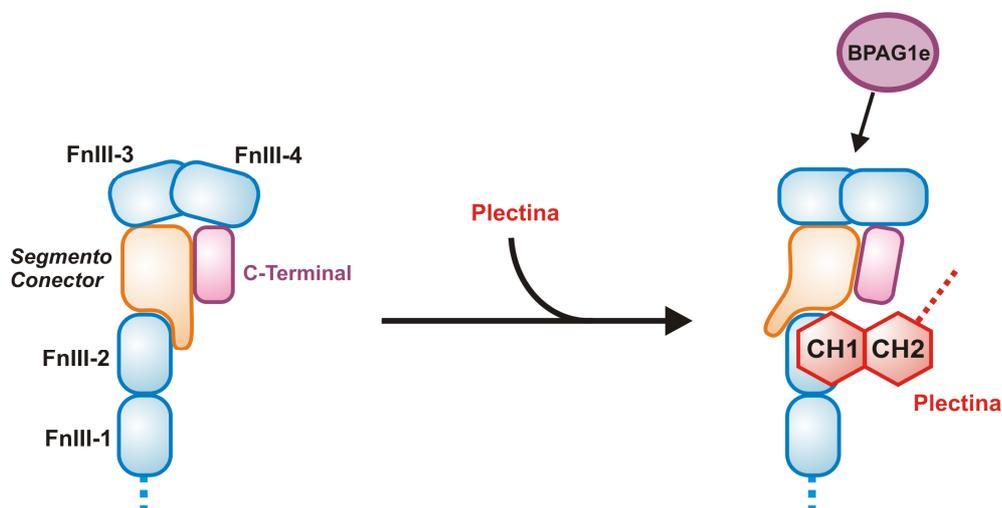


Figura D5.- Modelo hipotético de la propagación de cambios en $\beta 4$ ligados a la unión a plectina. La unión de plectina al FnIII-1,2 causa un cambio en la estructura de la región N-terminal del CS que podría transmitirse a otras regiones de $\beta 4$ favoreciendo la unión de BPAG1e o de otras proteínas de los HDs.

2.- CD151

Inicialmente se abordó la expresión de CD151-EC2 en bacterias. Para ello se emplearon diversas construcciones en las que el EC2 se expresó unido a diversas etiquetas o proteínas adyuvantes. En todos los casos el EC2 apareció en la fracción insoluble. Esto contrasta con la expresión soluble en *E. coli* del dominio EC2 de la tetraspanina CD81 (Kitadokoro et al., 2001b). Un paso limitante para la expresión soluble de CD151-EC2 podría ser la correcta formación de los tres puentes disulfuro que se espera que contenga el EC2 de CD151. Se ha intentado, sin éxito, expresar esta proteína soluble favoreciendo la formación de puentes disulfuro en *E. coli* incluyendo en las proteínas de fusión secuencias señal que dirigen la secreción al espacio periplasmático, tal como se había descrito para la expresión de CD81-EC2, o generando un ambiente oxidante en el citoplasma. En ambos casos la creación de un entorno que favorece la formación de puentes disulfuro es insuficiente para obtener CD151-EC2 soluble en bacterias. Esto

sugiere que existen diferencias significativas en el proceso de plegamiento de los dominios EC2 de CD151 y CD81.

Se ha descrito la expresión del EC2 de CD151 y de otras tetraspaninas en *E. coli* en forma de proteínas de fusión con GST, y estas proteínas recombinantes tienen capacidad para inhibir la infección de macrófagos por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1) (Ho et al., 2006). Dicho trabajo no analiza si el EC2 de CD151 fusionado a GST retiene una estructura nativa, por lo que no se puede descartar que el efecto observado sobre la infección por VIH-1 pueda deberse a motivos de secuencia cortos que no requieren de una correcta estructura tridimensional de este dominio. Por otro lado, recientemente se ha descrito la expresión soluble del dominio EC2 de la tetraspanina RDS (*Retinal degeneration slow*) (también conocida como periferina o Tspan-22) (Chakraborty et al., 2013). RDS se encuadra en el mismo subgrupo, 2b, que CD151 (Seigneuret et al., 2001). En dicho trabajo expresan en *E. coli* una proteína de fusión en la que el EC2 de RDS está unido al extremo carboxilo de la proteína de unión a maltosa (MBP) y no emplean mecanismos que favorezcan la formación de tres puentes disulfuro. Aunque la proteína de fusión contiene entre la MBP y el RDS-EC2 una secuencia reconocida por la proteasa *PreScission* no se indica que se haya separado el EC2 de la MBP ni que el EC2 sea estable y soluble en ausencia de la MBP. Por otro lado no se indica que se haya inducido de forma activa la formación de los tres puentes disulfuro predichos en el RDS-EC2. En resumen, no parece concluyente que el EC2 de RDS expresado en bacterias adopte una estructura nativa estable.

Ante la imposibilidad de producir el CD151-EC2 soluble en bacteria, la proteína se purificó en condiciones desnaturizantes y se replegó por dilución empleando un ambiente oxidante que favoreciera la formación de puentes disulfuro. De esta forma se obtuvo EC2 soluble a alta concentración en ausencia de agentes desnaturizantes; esta proteína tiene las seis Cys oxidadas, formando al menos algunas de ellas puentes disulfuro, y eluye en cromatografía de exclusión molecular mayoritariamente en un pico con una masa molecular aparente que se corresponde con la del monómero. Sin embargo la estimación del contenido de estructura secundaria estimada a partir del espectro de dicroísmo circular en la región del ultravioleta lejano (Figura R30) sugiere un bajo contenido en α -hélice (<10%). Las estructuras cristalográficas del EC2 de CD81

(códigos PDB: 1G8Q y 1IV5) (Kitadokoro et al., 2001a; Kitadokoro et al., 2002) revelan un contenido de α -hélice de ~65%. La región conservada del EC2 está formada por ~62 residuos de los cuales en CD81 el ~82% forman las hélices A, B y E. El EC2 de CD151 contiene ~108 residuos, por lo que el contenido en α -hélice de su región conservada supondría ~47%. Por lo tanto el bajo contenido en α -hélice estimado a partir del espectro de dicroísmo circular no parece ser compatible con la estructura helicoidal de la región conservada y sugiere que el CD151-EC2 replegado obtenido no adopta una estructura nativa.

Ante la falta de éxito en la producción soluble del EC2 de CD151 en *E. coli*, se expresó el EC2 en células de insecto empleando el sistema de baculovirus; de esta forma se ha obtenido proteína soluble. Esto sugiere que el EC2 de CD151 aislado es estable y que no requiere del resto de la proteína (por ejemplo los dominios transmembrana) para mantener su estructura. Por lo tanto la imposibilidad de obtener EC2 soluble en *E. coli* estaría causada por las limitaciones del sistema de expresión procariótico. El EC2 producido en células de insecto muestra heterogeneidad (Figura R33), la cual puede ser una limitación para su estudio estructural. Esta heterogeneidad parece estar ligada, al menos en parte, a procesos de glicosilación. La Asn159, al principio de la región variable del EC2, forma parte de una secuencia consenso de glicosilación NX(ST). Con el fin de reducir la heterogeneidad del EC2 recombinante, se podría prevenir la glicosilación sustituyendo Asn159. Si bien será necesario analizar el efecto de la eliminación de este potencial punto de glicosilación sobre los niveles de expresión y sobre la estabilidad del EC2.

Aunque este trabajo se ha centrado en la expresión y purificación del EC2, otra alternativa para el estudio de la estructura y función de CD151 sería producir la proteína completa. De hecho, ese sería el siguiente paso lógico tras el estudio del EC2 aislado. Hasta la fecha no se ha elucidado la estructura de ninguna tetraspanina completa a alta resolución, si bien se ha descrito la estructura a baja resolución, obtenida mediante criomicroscopía electrónica, de partículas de uroplakina que contienen las tetraspaninas uroplakina Ia y Ib (Min et al., 2006). Recientemente se ha descrito la producción de CD151 y otras tetraspaninas en *E. coli*, si bien CD151 obtenido de esta manera y solubilizado con detergentes tiene una masa molecular aparente de ~290 kDa, lo cual

sugiere que CD151 purificado de esta forma parece estar oligomerizando (Tarry et al., 2012). Es más razonable que la purificación de CD151 completo requiera el uso de sistemas eucarióticos de expresión como los empleados en este trabajo. Por último, una vez que se llegue a purificar CD151 completo sería interesante reconstituir CD151 en bicapas lipídicas, por ejemplo empleando nanodiscos lo cual permitiría su estudio en un ambiente similar a su entrono natural (Bayburt and Sligar, 2010; Borch and Hamann, 2009).

CONCLUSIONES

El dominio Calx- β de la integrina $\beta 4$ tiene una estructura globalmente similar a la de otros dominios Calx- β , como los de los intercambiadores de $\text{Na}^+/\text{Ca}^{2+}$, pero presenta diferencias significativas en la región supuestamente implicada en la unión de Ca^{2+} .

1. El dominio Calx- β de $\beta 4$ no une Ca^{2+} .
2. La R1003 y posiblemente la R1077, ambas en el pseudo-sitio de unión de Ca^{2+} , contribuyen a la estabilidad independiente de cationes del dominio Calx- β de $\beta 4$ mediante la formación de puentes salinos intramoleculares que compensan cargas de residuos ácidos.
3. El tercer y cuarto dominios FnIII de $\beta 4$ tienen estructuras globalmente similares a las de los otros dos dominios FnIII de esta integrina, pero existe poca conservación a nivel de los residuos de la superficie expuestos al solvente.
4. La región FnIII-3,4 de $\beta 4$ adopta en solución una conformación compacta aparentemente estable.
5. Los extremos amino y carboxilo de la región FnIII-3,4 se encuentran en el mismo lado de la estructura, sugiriendo que esta región favorece la proximidad entre el segmento conector y la cola C-terminal de $\beta 4$.
6. La expresión del dominio EC2 de la tetraspanina CD151 en forma de proteínas de fusión en *E. coli* resulta en su acumulación en la fracción insoluble.
7. El EC2 de CD151 expresado en *E. coli*, purificado en condiciones desnaturalizantes y replegado no parece adquirir una estructura nativa.
8. Se ha purificado el EC2 de CD151 soluble mediante expresión heteróloga en células de insecto combinada con cromatografía de afinidad; la proteína así obtenida muestra heterogeneidad relacionada con procesos de glicosilación.

BIBLIOGRAFÍA

- Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr* 52, 30-42.
- Afonine, P.V., Grosse-Kunstleve, R.W., and Adams, P.D. (2005). The Phenix refinement framework. *CCP4 Newsl* 42.
- Aho, S., and Uitto, J. (1998a). Direct interaction between the intracellular domains of bullous pemphigoid antigen 2 (BP180) and beta 4 integrin, hemidesmosomal components of basal keratinocytes. *Biochem Biophys Res Commun* 243, 694-699.
- Aho, S., and Uitto, J. (1998b). Direct interaction between the intracellular domains of bullous pemphigoid antigen 2 (BP180) and beta 4 integrin, hemidesmosomal components of basal keratinocytes. *Biochem Biophys Res Commun* 243, 694-699.
- Arnaout, M.A., Goodman, S.L., and Xiong, J.P. (2007). Structure and mechanics of integrin-based cell adhesion. *Curr Opin Cell Biol* 19, 495-507.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98, 10037-10041.
- Baldwin, G., Novitskaya, V., Sadej, R., Pochec, E., Litynska, A., Hartmann, C., Williams, J., Ashman, L., Eble, J.A., and Berditchevski, F. (2008). Tetraspanin CD151 regulates glycosylation of (alpha)3(beta)1 integrin. *J Biol Chem* 283, 35445-35454.
- Barczyk, M., Carracedo, S., and Gullberg, D. (2010). Integrins. *Cell and tissue research* 339, 269-280.
- Bayburt, T.H., and Sligar, S.G. (2010). Membrane protein assembly into Nanodiscs. *FEBS Lett* 584, 1721-1727.
- Berditchevski, F., Zutter, M.M., and Hemler, M.E. (1996). Characterization of novel complexes on the cell surface between integrins and proteins with 4 transmembrane domains (TM4 proteins). *Mol Biol Cell* 7, 193-207.
- Bertotti, A., Comoglio, P.M., and Trusolino, L. (2006). Beta4 integrin activates a Shp2-Src signaling pathway that sustains HGF-induced anchorage-independent growth. *J Cell Biol* 175, 993-1003.
- Besserer, G.M., Ottolia, M., Nicoll, D.A., Chaptal, V., Cascio, D., Philipson, K.D., and Abramson, J. (2007). The second Ca²⁺-binding domain of the Na⁺ Ca²⁺ exchanger is essential for regulation: crystal structures and mutational analysis. *Proc Natl Acad Sci U S A* 104, 18467-18472.
- Bienstock, R.J., and Barrett, J.C. (2001). KAI1, a prostate metastasis suppressor: prediction of solvated structure and interactions with binding partners; integrins, cadherins, and cell-surface receptor proteins. *Mol Carcinog* 32, 139-153.
- Bisig, D., Weber, P., Vaughan, L., Winterhalter, K.H., and Piontek, K. (1999). Purification, crystallization and preliminary crystallographic studies of a two fibronectin type-III domain segment from chicken tenascin encompassing the heparin- and contactin-binding regions. *Acta Crystallogr D Biol Crystallogr* 55, 1069-1073.
- Borch, J., and Hamann, T. (2009). The nanodisc: a novel tool for membrane protein studies. *Biol Chem* 390, 805-814.
- Borradori, L., Chavanas, S., Schaapveld, R.O., Gagnoux-Palacios, L., Calafat, J., Meneguzzi, G., and Sonnenberg, A. (1998). Role of the bullous pemphigoid antigen 180 (BP180) in the assembly of hemidesmosomes and cell adhesion--reexpression of BP180 in generalized atrophic benign epidermolysis bullosa keratinocytes. *Exp Cell Res* 239, 463-476.

- Borradori, L., Koch, P.J., Niessen, C.M., Erkeland, S., van Leusden, M.R., and Sonnenberg, A. (1997). The localization of bullous pemphigoid antigen 180 (BP180) in hemidesmosomes is mediated by its cytoplasmic domain and seems to be regulated by the beta4 integrin subunit. *J Cell Biol* *136*, 1333-1347.
- Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* *59*, 2023-2030.
- Brown, A., Bernier, G., Mathieu, M., Rossant, J., and Kothary, R. (1995). The mouse dystonia musculorum gene is a neural isoform of bullous pemphigoid antigen 1. *Nat Genet* *10*, 301-306.
- Bucher, R.M., Svergun, D.I., Muhle-Goll, C., and Mayans, O. (2010). The structure of the FnIII Tandem A77-A78 points to a periodically conserved architecture in the myosin-binding region of titin. *J Mol Biol* *401*, 843-853.
- Campbell, I.D., and Humphries, M.J. (2011). Integrin structure, activation, and interactions. *Cold Spring Harbor perspectives in biology* *3*.
- Campbell, I.D., and Spitzfaden, C. (1994). Building proteins with fibronectin type III modules. *Structure* *2*, 333-337.
- Canutescu, A.A., Shelenkov, A.A., and Dunbrack, R.L., Jr. (2003). A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci* *12*, 2001-2014.
- Chacon, P., Diaz, J.F., Moran, F., and Andreu, J.M. (2000). Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. *J Mol Biol* *299*, 1289-1302.
- Chacon, P., Moran, F., Diaz, J.F., Pantos, E., and Andreu, J.M. (1998). Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys J* *74*, 2760-2775.
- Chakraborty, D., Rodgers, K.K., Conley, S.M., and Naash, M.I. (2013). Structural characterization of the second intra-discal loop of the photoreceptor tetraspanin RDS. *FEBS J* *280*, 127-138.
- Charrin, S., Manie, S., Oualid, M., Billard, M., Boucheix, C., and Rubinstein, E. (2002). Differential stability of tetraspanin/tetraspanin interactions: role of palmitoylation. *FEBS Lett* *516*, 139-144.
- Chung, H.J., and Uitto, J. (2010). Epidermolysis bullosa with pyloric atresia. *Dermatologic clinics* *28*, 43-54.
- Cimpmperman, P., Baranauskiene, L., Jachimoviciute, S., Jachno, J., Torresan, J., Michailoviene, V., Matuliene, J., Sereikaite, J., Bumelis, V., and Matulis, D. (2008). A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys J* *95*, 3222-3231.
- Cowtan, K. (1994). 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* *31*, 34-38.
- Danen, E.H., and Sonnenberg, A. (2003). Integrins in regulation of tissue development and function. *J Pathol* *201*, 632-641.
- Dans, M., Gagnoux-Palacios, L., Blaikie, P., Klein, S., Mariotti, A., and Giancotti, F.G. (2001a). Tyrosine phosphorylation of the beta 4 integrin cytoplasmic domain mediates Shc signaling to extracellular signal-regulated kinase and antagonizes formation of hemidesmosomes. *J Biol Chem* *276*, 1494-1502.

- Dans, M., Gagnoux-Palacios, L., Blaikie, P., Klein, S., Mariotti, A., and Giancotti, F.G. (2001b). Tyrosine phosphorylation of the beta 4 integrin cytoplasmic domain mediates Shc signaling to extracellular signal-regulated kinase and antagonizes formation of hemidesmosomes. *J Biol Chem* 276, 1494-1502.
- Darling, T.N., Bauer, J.W., Hintner, H., and Yancey, K.B. (1997). Generalized atrophic benign epidermolysis bullosa. *Adv Dermatol* 13, 87-119; discussion 120.
- Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S., *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35, W375-383.
- de Pereda, J.M., Lillo, M.P., and Sonnenberg, A. (2009a). Structural basis of the interaction between integrin alpha6beta4 and plectin at the hemidesmosomes. *Embo J* 28, 1180-1190.
- de Pereda, J.M., Ortega, E., Alonso-Garcia, N., Gomez-Hernandez, M., and Sonnenberg, A. (2009b). Advances and perspectives of the architecture of hemidesmosomes: lessons from structural biology. *Cell Adh Migr* 3, 361-364.
- de Pereda, J.M., Wiche, G., and Liddington, R.C. (1999). Crystal structure of a tandem pair of fibronectin type III domains from the cytoplasmic tail of integrin alpha6beta4. *Embo J* 18, 4087-4095.
- de Vos, A.M., Ultsch, M., and Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 255, 306-312.
- Debye, P. (1915). Scattering from non-crystalline substances. *Ann Phys* 46, 809-823.
- Dickinson, C.D., Veerapandian, B., Dai, X.P., Hamlin, R.C., Xuong, N.H., Ruoslahti, E., and Ely, K.R. (1994). Crystal structure of the tenth type III cell adhesion module of human fibronectin. *J Mol Biol* 236, 1079-1092.
- Diederichs, K., and Karplus, P.A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* 4, 269-275.
- Dowling, J., Yu, Q.C., and Fuchs, E. (1996). Beta4 integrin is required for hemidesmosome formation, cell adhesion and cell survival. *J Cell Biol* 134, 559-572.
- Drummer, H.E., Wilson, K.A., and Pountourios, P. (2005). Determinants of CD81 dimerization and interaction with hepatitis C virus glycoprotein E2. *Biochem Biophys Res Commun* 328, 251-257.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
- Fernandez-Busquets, X., Kammerer, R.A., and Burger, M.M. (1996). A 35-kDa protein is the basic unit of the core from the 2 x 10(4)-kDa aggregation factor responsible for species-specific cell adhesion in the marine sponge *Microciona prolifera*. *J Biol Chem* 271, 23558-23565.
- Fine, J.D., Eady, R.A., Bauer, E.A., Bauer, J.W., Bruckner-Tuderman, L., Heagerty, A., Hintner, H., Hovnanian, A., Jonkman, M.F., Leigh, I., *et al.* (2008). The classification of inherited epidermolysis bullosa (EB): Report of the Third International Consensus Meeting on Diagnosis and Classification of EB. *Journal of the American Academy of Dermatology* 58, 931-950.
- Franzke, C.W., Bruckner, P., and Bruckner-Tuderman, L. (2005). Collagenous transmembrane proteins: recent insights into biology and pathology. *J Biol Chem* 280, 4005-4008.

- Franzke, C.W., Tasanen, K., Schumann, H., and Bruckner-Tuderman, L. (2003). Collagenous transmembrane proteins: collagen XVII as a prototype. *Matrix Biol* 22, 299-309.
- French, S., and Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Crystallographica Section A* 34, 517-525.
- Frijns, E., Kuikman, I., Litjens, S., Raspe, M., Jalink, K., Ports, M., Wilhelmsen, K., and Sonnenberg, A. (2012). Phosphorylation of threonine 1736 in the C-terminal tail of integrin beta4 contributes to hemidesmosome disassembly. *Mol Biol Cell* 23, 1475-1485.
- Frijns, E., Sachs, N., Kreft, M., Wilhelmsen, K., and Sonnenberg, A. (2010). EGF-induced MAPK signaling inhibits hemidesmosome formation through phosphorylation of the integrin {beta}4. *J Biol Chem* 285, 37650-37662.
- Garcia-Alvarez, B., Bobkov, A., Sonnenberg, A., and de Pereda, J.M. (2003). Structural and functional analysis of the actin binding domain of plectin suggests alternative mechanisms for binding to F-actin and to integrin alpha6beta4. *Structure* 11, 615-625.
- García-Alvarez, B., de Pereda, J.M., Calderwood, D.A., Critchley, D.R., Ulmer, T.S., Campbell, I.D., Ginsberg, M.H., and Liddington, R.C. (2003). Structural determinants of integrin recognition by Talin. *Mol Cell* 11, 49-58.
- Garman, E.F., and Mitchell, E.P. (1996). Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *Journal of applied crystallography* 29, 584-587.
- Geerts, D., Fontao, L., Nievers, M.G., Schaapveld, R.Q., Purkis, P.E., Wheeler, G.N., Lane, E.B., Leigh, I.M., and Sonnenberg, A. (1999). Binding of integrin alpha6beta4 to plectin prevents plectin association with F-actin but does not interfere with intermediate filament binding. *J Cell Biol* 147, 417-434.
- Georges-Labouesse, E., Messaddeq, N., Yehia, G., Cadalbert, L., Dierich, A., and Le Meur, M. (1996). Absence of integrin alpha 6 leads to epidermolysis bullosa and neonatal death in mice. *Nat Genet* 13, 370-373.
- Germain, E.C., Santos, T.M., and Rabinovitz, I. (2009). Phosphorylation of a novel site on the {beta}4 integrin at the trailing edge of migrating cells promotes hemidesmosome disassembly. *Mol Biol Cell* 20, 56-67.
- Geuijen, C.A., and Sonnenberg, A. (2002). Dynamics of the alpha6beta4 integrin in keratinocytes. *Mol Biol Cell* 13, 3845-3858.
- Giancotti, F.G. (2007). Targeting integrin beta4 for cancer and anti-angiogenic therapy. *Trends Pharmacol Sci* 28, 506-511.
- Giancotti, F.G., Stepp, M.A., Suzuki, S., Engvall, E., and Ruoslahti, E. (1992). Proteolytic processing of endogenous and recombinant beta 4 integrin subunit. *J Cell Biol* 118, 951-959.
- Giudice, G.J., Emery, D.J., and Diaz, L.A. (1992). Cloning and primary structural analysis of the bullous pemphigoid autoantigen BP180. *J Invest Dermatol* 99, 243-250.
- Glatter, O., and Kratky, O. (1982). *Small Angle X-ray Scattering*. Academic Press.
- Green, K.J., Virata, M.L., Elgart, G.W., Stanley, J.R., and Parry, D.A. (1992). Comparative structural analysis of desmoplakin, bullous pemphigoid antigen and plectin: members of a new gene family involved in organization of intermediate filaments. *Int J Biol Macromol* 14, 145-153.

- Groves, R.W., Liu, L., Dopping-Hepenstal, P.J., Markus, H.S., Lovell, P.A., Ozoemena, L., Lai-Cheong, J.E., Gawler, J., Owaribe, K., Hashimoto, T., *et al.* (2010). A homozygous nonsense mutation within the dystonin gene coding for the coiled-coil domain of the epithelial isoform of BPAG1 underlies a new subtype of autosomal recessive epidermolysis bullosa simplex. *J Invest Dermatol* *130*, 1551-1557.
- Guinier, A., and Fournet, G. (1955). *Small-angle scattering of X-rays* (Wiley).
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* *166*, 557-580.
- Hayase, M., Maekawa, A., Yubisui, T., and Minami, Y. (2008). Properties, intracellular localization, and stage-specific expression of membrane-bound beta-glucosidase, BgIM1, from *Physarum polycephalum*. *The international journal of biochemistry & cell biology* *40*, 2141-2150.
- Hemler, M.E. (2003). Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain. *Annu Rev Cell Dev Biol* *19*, 397-422.
- Hemler, M.E. (2005). Tetraspanin functions and associated microdomains. *Nat Rev Mol Cell Biol* *6*, 801-811.
- Higginbottom, A., Quinn, E.R., Kuo, C.C., Flint, M., Wilson, L.H., Bianchi, E., Nicosia, A., Monk, P.N., McKeating, J.A., and Levy, S. (2000). Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol* *74*, 3642-3649.
- Hilge, M., Aelen, J., and Vuister, G.W. (2006). Ca²⁺ regulation in the Na⁺/Ca²⁺ exchanger involves two markedly different Ca²⁺ sensors. *Mol Cell* *22*, 15-25.
- Ho, S.H., Martin, F., Higginbottom, A., Partridge, L.J., Parthasarathy, V., Moseley, G.W., Lopez, P., Cheng-Mayer, C., and Monk, P.N. (2006). Recombinant extracellular domains of tetraspanin proteins are potent inhibitors of the infection of macrophages by human immunodeficiency virus type 1. *J Virol* *80*, 6487-6496.
- Hodor, P.G., Illies, M.R., Broadley, S., and Ettensohn, C.A. (2000). Cell-substrate interactions during sea urchin gastrulation: migrating primary mesenchyme cells interact with and align extracellular matrix fibers that contain ECM3, a molecule with NG2-like and multiple calcium-binding domains. *Dev Biol* *222*, 181-194.
- Hogervorst, F., Kuikman, I., von dem Borne, A.E., and Sonnenberg, A. (1990). Cloning and sequence analysis of beta-4 cDNA: an integrin subunit that contains a unique 118 kd cytoplasmic domain. *The EMBO journal* *9*, 765-770.
- Holm, L., and Park, J. (2000). DaliLite workbench for protein structure comparison. *Bioinformatics* *16*, 566-567.
- Hopkinson, I., Evans, W., Chant, D., Hiscox, S., Berry, D., and Harding, K. (1995). Reverse transcription-polymerase chain reaction detection of collagen transcripts in healing human wounds. *Eur J Clin Invest* *25*, 539-542.
- Hopkinson, S.B., Findlay, K., deHart, G.W., and Jones, J.C. (1998). Interaction of BP180 (type XVII collagen) and alpha6 integrin is necessary for stabilization of hemidesmosome structure. *J Invest Dermatol* *111*, 1015-1022.
- Hopkinson, S.B., and Jones, J.C. (2000). The N terminus of the transmembrane protein BP180 interacts with the N-terminal domain of BP230, thereby mediating keratin cytoskeleton anchorage to the cell surface at the site of the hemidesmosome. *Mol Biol Cell* *11*, 277-286.

- Hu, X., Wang, H., Ke, H., and Kuhlman, B. (2007). High-resolution design of a protein loop. *Proc Natl Acad Sci U S A* *104*, 17668-17673.
- Hutchinson, E.G., and Thornton, J.M. (1994). A revised set of potentials for beta-turn formation in proteins. *Protein Sci* *3*, 2207-2216.
- Hynes, R.O. (2002). Integrins: bidirectional, allosteric signaling machines. *Cell* *110*, 673-687.
- Imhof, D., Wavreille, A.S., May, A., Zacharias, M., Tridandapani, S., and Pei, D. (2006). Sequence specificity of SHP-1 and SHP-2 Src homology 2 domains. Critical roles of residues beyond the pY+3 position. *J Biol Chem* *281*, 20271-20282.
- Jancarik, J., and Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of applied crystallography* *24*, 409-411.
- Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., and Godzik, A. (2005). FFAS03: a server for profile-profile sequence alignments. *Nucleic Acids Res* *33*, W284-288.
- Jefferson, J.J., Ciatto, C., Shapiro, L., and Liem, R.K. (2007). Structural analysis of the plakin domain of bullous pemphigoid antigen1 (BPAG1) suggests that plakins are members of the spectrin superfamily. *J Mol Biol* *366*, 244-257.
- Jefferson, J.J., Leung, C.L., and Liem, R.K. (2004). Plakins: goliaths that link cell junctions and the cytoskeleton. *Nat Rev Mol Cell Biol* *5*, 542-553.
- Jeschke, G. (2012). DEER distance measurements on proteins. *Annual review of physical chemistry* *63*, 419-446.
- Jeschke, G., and Polyhach, Y. (2007). Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Physical chemistry chemical physics : PCCP* *9*, 1895-1910.
- Johnson, E., Bruschiweiler-Li, L., Showalter, S.A., Vuister, G.W., Zhang, F., and Bruschiweiler, R. (2008). Structure and dynamics of Ca²⁺-binding domain 1 of the Na⁺/Ca²⁺ exchanger in the presence and in the absence of Ca²⁺. *J Mol Biol* *377*, 945-955.
- Jones, J.C., Kurpakus, M.A., Cooper, H.M., and Quaranta, V. (1991). A function for the integrin alpha 6 beta 4 in the hemidesmosome. *Cell regulation* *2*, 427-438.
- Jonkman, M.F., Pas, H.H., Nijenhuis, M., Kloosterhuis, G., and Steege, G. (2002). Deletion of a cytoplasmic domain of integrin beta4 causes epidermolysis bullosa simplex. *J Invest Dermatol* *119*, 1275-1281.
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* *32*, 922-923.
- Kabsch, W. (2010). Xds. *Acta Crystallogr D Biol Crystallogr* *66*, 125-132.
- Karamatic Crew, V., Burton, N., Kagan, A., Green, C.A., Levene, C., Flinter, F., Brady, R.L., Daniels, G., and Anstee, D.J. (2004). CD151, the first member of the tetraspanin (TM4) superfamily detected on erythrocytes, is essential for the correct assembly of human basement membranes in kidney and skin. *Blood* *104*, 2217-2223.
- Kazarov, A.R., Yang, X., Stipp, C.S., Sehgal, B., and Hemler, M.E. (2002). An extracellular site on tetraspanin CD151 determines alpha 3 and alpha 6 integrin-dependent cellular morphology. *J Cell Biol* *158*, 1299-1309.
- Kitadokoro, K., Bordo, D., Galli, G., Petracca, R., Falugi, F., Abrignani, S., Grandi, G., and Bolognesi, M. (2001a). CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs. *Embo J* *20*, 12-18.

- Kitadokoro, K., Galli, G., Petracca, R., Falugi, F., Grandi, G., and Bolognesi, M. (2001b). Crystallization and preliminary crystallographic studies on the large extracellular domain of human CD81, a tetraspanin receptor for hepatitis C virus. *Acta Crystallographica Section D* *57*, 156-158.
- Kitadokoro, K., Ponassi, M., Galli, G., Petracca, R., Falugi, F., Grandi, G., and Bolognesi, M. (2002). Subunit association and conformational flexibility in the head subdomain of human CD81 large extracellular loop. *Biol Chem* *383*, 1447-1452.
- Kiyozumi, D., Sugimoto, N., Nakano, I., and Sekiguchi, K. (2007). Frem3, a member of the 12 CSPG repeats-containing extracellular matrix protein family, is a basement membrane protein with tissue distribution patterns distinct from those of Fras1, Frem2, and QBRICK/Frem1. *Matrix Biol* *26*, 456-462.
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of applied crystallography* *36*, 1277-1282.
- Koster, J., Borradori, L., and Sonnenberg, A. (2004a). Hemidesmosomes: molecular organization and their importance for cell adhesion and disease. *Handb Exp Pharmacol*, 243-280.
- Koster, J., Geerts, D., Favre, B., Borradori, L., and Sonnenberg, A. (2003). Analysis of the interactions between BP180, BP230, plectin and the integrin alpha6beta4 important for hemidesmosome assembly. *J Cell Sci* *116*, 387-399.
- Koster, J., Kuikman, I., Kreft, M., and Sonnenberg, A. (2001). Two different mutations in the cytoplasmic domain of the integrin beta 4 subunit in nonlethal forms of epidermolysis bullosa prevent interaction of beta 4 with plectin. *J Invest Dermatol* *117*, 1405-1411.
- Koster, J., van Wilpe, S., Kuikman, I., Litjens, S.H., and Sonnenberg, A. (2004b). Role of binding of plectin to the integrin beta4 subunit in the assembly of hemidesmosomes. *Mol Biol Cell* *15*, 1211-1223.
- Kovalenko, O.V., Metcalf, D.G., DeGrado, W.F., and Hemler, M.E. (2005). Structural organization and interactions of transmembrane domains in tetraspanin proteins. *BMC Struct Biol* *5*, 11.
- Kovalenko, O.V., Yang, X., Kolesnikova, T.V., and Hemler, M.E. (2004). Evidence for specific tetraspanin homodimers: inhibition of palmitoylation makes cysteine residues available for cross-linking. *Biochem J* *377*, 407-417.
- Kozin, M.B., and Svergun, D.I. (2001). Automated matching of high- and low-resolution structural models. *Journal of applied crystallography* *34*, 33-41.
- Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* *3*, 1171-1179.
- Lau, L.M., Wee, J.L., Wright, M.D., Moseley, G.W., Hogarth, P.M., Ashman, L.K., and Jackson, D.E. (2004). The tetraspanin superfamily member CD151 regulates outside-in integrin alphaIIb beta3 signaling and platelet function. *Blood* *104*, 2368-2375.
- Lee, C.H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S.E., and Kuriyan, J. (1994). Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure* *2*, 423-438.
- Levy, S., and Shoham, T. (2005). Protein-protein interactions in the tetraspanin web. *Physiology (Bethesda)* *20*, 218-224.

- Lipscomb, E.A., Simpson, K.J., Lyle, S.R., Ring, J.E., Dugan, A.S., and Mercurio, A.M. (2005). The alpha6beta4 integrin maintains the survival of human breast carcinoma cells in vivo. *Cancer Res* 65, 10970-10976.
- Litjens, S.H., de Pereda, J.M., and Sonnenberg, A. (2006). Current insights into the formation and breakdown of hemidesmosomes. *Trends Cell Biol* 16, 376-383.
- Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by C α geometry: phi,psi and C β deviation. *Proteins* 50, 437-450.
- Lytton, J. (2007). Na⁺/Ca²⁺ exchangers: three mammalian gene families control Ca²⁺ transport. *The Biochemical journal* 406, 365-382.
- Mainiero, F., Pepe, A., Wary, K.K., Spinardi, L., Mohammadi, M., Schlessinger, J., and Giancotti, F.G. (1995). Signal transduction by the alpha 6 beta 4 integrin: distinct beta 4 subunit sites mediate recruitment of Shc/Grb2 and association with the cytoskeleton of hemidesmosomes. *EMBO J* 14, 4470-4481.
- Margadant, C., Frijns, E., Wilhelmssen, K., and Sonnenberg, A. (2008). Regulation of hemidesmosome disassembly by growth factor receptors. *Curr Opin Cell Biol* 20, 589-596.
- Mariotti, A., Kedeshian, P.A., Dans, M., Curatola, A.M., Gagnoux-Palacios, L., and Giancotti, F.G. (2001). EGF-R signaling through Fyn kinase disrupts the function of integrin alpha6beta4 at hemidesmosomes: role in epithelial cell migration and carcinoma invasion. *J Cell Biol* 155, 447-458.
- Martin, F., Roth, D.M., Jans, D.A., Pouton, C.W., Partridge, L.J., Monk, P.N., and Moseley, G.W. (2005). Tetraspanins in viral infections: a fundamental role in viral biology? *J Virol* 79, 10839-10851.
- May, A.P., and Ponting, C.P. (1999). Integrin alpha- and beta 4-subunit-domain homologues in cyanobacterial proteins. *Trends Biochem Sci* 24, 12-13.
- McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61, 458-464.
- McLean, W.H., Pulkkinen, L., Smith, F.J., Rugg, E.L., Lane, E.B., Bullrich, F., Burgeson, R.E., Amano, S., Hudson, D.L., Owaribe, K., *et al.* (1996). Loss of plectin causes epidermolysis bullosa with muscular dystrophy: cDNA cloning and genomic organization. *Genes Dev* 10, 1724-1735.
- McMillan, J.R., McGrath, J.A., Tidman, M.J., and Eady, R.A. (1998). Hemidesmosomes show abnormal association with the keratin filament network in junctional forms of epidermolysis bullosa. *J Invest Dermatol* 110, 132-137.
- Min, G., Wang, H., Sun, T.T., and Kong, X.P. (2006). Structural basis for tetraspanin functions as revealed by the cryo-EM structure of uroplakin complexes at 6-Å resolution. *J Cell Biol* 173, 975-983.
- Murgia, C., Blaikie, P., Kim, N., Dans, M., Petrie, H.T., and Giancotti, F.G. (1998). Cell cycle and adhesion defects in mice carrying a targeted deletion of the integrin beta4 cytoplasmic domain. *Embo J* 17, 3940-3951.
- Nakano, A., Pulkkinen, L., Murrell, D., Rico, J., Lucky, A.W., Garzon, M., Stevens, C.A., Robertson, S., Pfendner, E., and Uitto, J. (2001). Epidermolysis bullosa with congenital pyloric atresia: novel mutations in the beta 4 integrin gene (ITGB4) and genotype/phenotype correlations. *Pediatr Res* 49, 618-626.

- Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J.L., Stuart, D.I., and Perrakis, A. (2005). Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallographica Section D* *61*, 1426-1431.
- Nicoll, D.A., Ren, X., Ottolia, M., Phillips, M., Paredes, A.R., Abramson, J., and Philipson, K.D. (2007). What we know about the structure of NCX1 and how it relates to its function. *Annals of the New York Academy of Sciences* *1099*, 1-6.
- Nicoll, D.A., Sawaya, M.R., Kwon, S., Cascio, D., Philipson, K.D., and Abramson, J. (2006). The crystal structure of the primary Ca²⁺ sensor of the Na⁺/Ca²⁺ exchanger reveals a novel Ca²⁺ binding motif. *J Biol Chem* *281*, 21577-21581.
- Niessen, C.M., Hulsman, E.H., Oomen, L.C., Kuikman, I., and Sonnenberg, A. (1997). A minimal region on the integrin beta4 subunit that is critical to its localization in hemidesmosomes regulates the distribution of HD1/plectin in COS-7 cells. *J Cell Sci* *110*, 1705-1716.
- Nievers, M.G., Schaapveld, R.Q., Oomen, L.C., Fontao, L., Geerts, D., and Sonnenberg, A. (1998). Ligand-independent role of the beta 4 integrin subunit in the formation of hemidesmosomes. *J Cell Sci* *111*, 1659-1672.
- Nievers, M.G., Schaapveld, R.Q., and Sonnenberg, A. (1999). Biology and function of hemidesmosomes. *Matrix Biol* *18*, 5-17.
- Nikkila, H., McMillan, D.R., Nunez, B.S., Pascoe, L., Curnow, K.M., and White, P.C. (2000). Sequence similarities between a novel putative G protein-coupled receptor and Na⁺/Ca²⁺ exchangers define a cation binding domain. *Molecular endocrinology (Baltimore, Md)* *14*, 1351-1364.
- Ortega, E., Buey, R.M., Sonnenberg, A., and de Pereda, J.M. (2011). The structure of the plakin domain of plectin reveals a non-canonical SH3 domain interacting with its fourth spectrin repeat. *J Biol Chem* *286*, 12429-12438.
- Ottolia, M., Philipson, K.D., and John, S. (2004). Conformational changes of the Ca(2+) regulatory site of the Na(+)-Ca(2+) exchanger detected by FRET. *Biophys J* *87*, 899-906.
- Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* *4*, 2411-2423.
- Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* *62*, 439-450.
- Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B.A., Lane, P., *et al.* (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* *6*, 429-440.
- Pape, T., and Schneider, T.R. (2004). HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *Journal of applied crystallography* *37*, 843-844.
- Perez-Iratxeta, C., and Andrade-Navarro, M.A. (2008). K2D2: estimation of protein secondary structure from circular dichroism spectra. *BMC Struct Biol* *8*, 25.
- Praznikar, J., Afonine, P.V., Guncar, G., Adams, P.D., and Turk, D. (2009). Averaged kick maps: less noise, more signal... and probably less bias. *Acta Crystallogr D Biol Crystallogr* *65*, 921-931.
- Pulkkinen, L., Rouan, F., Bruckner-Tuderman, L., Wallerstein, R., Garzon, M., Brown, T., Smith, L., Carter, W., and Uitto, J. (1998). Novel ITGB4 mutations in lethal and nonlethal variants of epidermolysis bullosa with pyloric atresia: missense versus nonsense. *Am J Hum Genet* *63*, 1376-1387.

- Pulkkinen, L., Smith, F.J., Shimizu, H., Murata, S., Yaoita, H., Hachisuka, H., Nishikawa, T., McLean, W.H., and Uitto, J. (1996). Homozygous deletion mutations in the plectin gene (PLEC1) in patients with epidermolysis bullosa simplex associated with late-onset muscular dystrophy. *Hum Mol Genet* 5, 1539-1546.
- Rabinovitz, I., Tsomo, L., and Mercurio, A.M. (2004). Protein kinase C- α phosphorylation of specific serines in the connecting segment of the beta 4 integrin regulates the dynamics of type II hemidesmosomes. *Mol Cell Biol* 24, 4351-4360.
- Reddy, D., Muller, P., Tran, H., Nguyn, N., Schaecke, H., Bruckner-Tuderman, L., Giudice, G., and Marinkovich, P. (1998). The extracellular domain of BP180 binds laminin-5. *Journal of dermatological science* 16, S121.
- Rezniczek, G.A., de Pereda, J.M., Reipert, S., and Wiche, G. (1998). Linking integrin $\alpha 6 \beta 4$ -based cell adhesion to the intermediate filament cytoskeleton: direct interaction between the $\beta 4$ subunit and plectin at multiple molecular sites. *J Cell Biol* 141, 209-225.
- Rous, B.A., Reaves, B.J., Ihrke, G., Briggs, J.A., Gray, S.R., Stephens, D.J., Banting, G., and Luzio, J.P. (2002). Role of adaptor complex AP-3 in targeting wild-type and mutated CD63 to lysosomes. *Mol Biol Cell* 13, 1071-1082.
- Ruhrberg, C., Williamson, J.A., Sheer, D., and Watt, F.M. (1996). Chromosomal localisation of the human envoplakin gene (EVPL) to the region of the tylosis oesophageal cancer gene (TOCG) on 17q25. *Genomics* 37, 381-385.
- Sachs, N., Kreft, M., van den Bergh Weerman, M.A., Beynon, A.J., Peters, T.A., Weening, J.J., and Sonnenberg, A. (2006). Kidney failure in mice lacking the tetraspanin CD151. *J Cell Biol* 175, 33-39.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn (New York, Cold Spring Harbor Laboratory Press).
- Sawamura, D., Nakano, H., and Matsuzaki, Y. (2010). Overview of epidermolysis bullosa. *The Journal of dermatology* 37, 214-219.
- Schaapveld, R.Q., Borradori, L., Geerts, D., van Leusden, M.R., Kuikman, I., Nievers, M.G., Niessen, C.M., Steenbergen, R.D., Sniijders, P.J., and Sonnenberg, A. (1998). Hemidesmosome formation is initiated by the $\beta 4$ integrin subunit, requires complex formation of $\beta 4$ and HD1/plectin, and involves a direct interaction between $\beta 4$ and the bullous pemphigoid antigen 180. *J Cell Biol* 142, 271-284.
- Schmidt, E., and Zillikens, D. (2013). Pemphigoid diseases. *Lancet* 381, 320-332.
- Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* 58, 1772-1779.
- Scholz, C.J., Sauer, G., and Deissler, H. (2009). Glycosylation of tetraspanin Tspan-1 at four distinct sites promotes its transition through the endoplasmic reticulum. *Protein and peptide letters* 16, 1244-1248.
- Schomaker, V., and Trueblood, K.N. (1968). On the rigid-body motion of molecules in crystals. *Acta Crystallographica Section B* 24, 63-76.
- Schwarz, E.M., and Benzer, S. (1997). Calx, a Na-Ca exchanger gene of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 94, 10249-10254.
- Seigneuret, M., Delaguillaumie, A., Lagaudriere-Gesbert, C., and Conjeaud, H. (2001). Structure of the tetraspanin main extracellular domain. A partially conserved fold with a structurally variable domain insertion. *J Biol Chem* 276, 40055-40064.

- Sevcik, J., Urbanikova, L., Kost'an, J., Janda, L., and Wiche, G. (2004). Actin-binding domain of mouse plectin. Crystal structure and binding to vimentin. *Eur J Biochem* 271, 1873-1884.
- Shaw, L.M. (2001). Identification of insulin receptor substrate 1 (IRS-1) and IRS-2 as signaling intermediates in the $\alpha 6\beta 4$ integrin-dependent activation of phosphoinositide 3-OH kinase and promotion of invasion. *Mol Cell Biol* 21, 5082-5093.
- Sincock, P.M., Fitter, S., Parton, R.G., Berndt, M.C., Gamble, J.R., and Ashman, L.K. (1999). PETA-3/CD151, a member of the transmembrane 4 superfamily, is localised to the plasma membrane and endocytic system of endothelial cells, associates with multiple integrins and modulates cell function. *J Cell Sci* 112 (Pt 6), 833-844.
- Sincock, P.M., Mayrhofer, G., and Ashman, L.K. (1997). Localization of the transmembrane 4 superfamily (TM4SF) member PETA-3 (CD151) in normal human tissues: comparison with CD9, CD63, and $\alpha 5\beta 1$ integrin. *J Histochem Cytochem* 45, 515-525.
- Sitkoff, D., Sharp, K.A., and Honig, B. (1994). Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *The Journal of Physical Chemistry* 98, 1978-1988.
- Sonnenberg, A., Calafat, J., Janssen, H., Daams, H., van der Raaij-Helmer, L.M., Falcioni, R., Kennel, S.J., Aplin, J.D., Baker, J., Loizidou, M., *et al.* (1991). Integrin $\alpha 6/\beta 4$ complex is located in hemidesmosomes, suggesting a major role in epidermal cell-basement membrane adhesion. *J Cell Biol* 113, 907-917.
- Sonnenberg, A., Rojas, A.M., and de Pereda, J.M. (2007). The structure of a tandem pair of spectrin repeats of plectin reveals a modular organization of the plakin domain. *J Mol Biol* 368, 1379-1391.
- Stanley, J.R., Tanaka, T., Mueller, S., Klaus-Kovtun, V., and Roop, D. (1988). Isolation of complementary DNA for bullous pemphigoid antigen by use of patients' autoantibodies. *The Journal of clinical investigation* 82, 1864-1870.
- Stapp, M.A., Spurr-Michaud, S., Tisdale, A., Elwell, J., and Gipson, I.K. (1990). $\alpha 6\beta 4$ integrin heterodimer is a component of hemidesmosomes. *Proc Natl Acad Sci U S A* 87, 8970-8974.
- Sterk, L.M., Geuijen, C.A., Oomen, L.C., Calafat, J., Janssen, H., and Sonnenberg, A. (2000). The tetraspan molecule CD151, a novel constituent of hemidesmosomes, associates with the integrin $\alpha 6\beta 4$ and may regulate the spatial organization of hemidesmosomes. *J Cell Biol* 149, 969-982.
- Sterk, L.M., Geuijen, C.A., van den Berg, J.G., Claessen, N., Weening, J.J., and Sonnenberg, A. (2002). Association of the tetraspanin CD151 with the laminin-binding integrins $\alpha 3\beta 1$, $\alpha 6\beta 1$, $\alpha 6\beta 4$ and $\alpha 7\beta 1$ in cells in culture and in vivo. *J Cell Sci* 115, 1161-1173.
- Stipp, C.S., Kolesnikova, T.V., and Hemler, M.E. (2003). Functional domains in tetraspanin proteins. *Trends Biochem Sci* 28, 106-112.
- Suzuki, S., and Naitoh, Y. (1990). Amino acid sequence of a novel integrin $\beta 4$ subunit and primary expression of the mRNA in epithelial cells. *Embo J* 9, 757-763.
- Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of applied crystallography* 25, 495-503.
- Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of applied crystallography* 28, 768-773.

- Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76, 2879-2886.
- Tamura, R.N., Rozzo, C., Starr, L., Chambers, J., Reichardt, L.F., Cooper, H.M., and Quaranta, V. (1990). Epithelial integrin alpha 6 beta 4: complete primary structure of alpha 6 and variant forms of beta 4. *The Journal of cell biology* 111, 1593-1604.
- Tarry, M., Skaar, K., Heijne, G., Draheim, R.R., and Høgbom, M. (2012). Production of human tetraspanin proteins in *Escherichia coli*. *Protein Expr Purif* 82, 373-379.
- Tasanen, K., Tunggal, L., Chometon, G., Bruckner-Tuderman, L., and Aumailley, M. (2004). Keratinocytes from patients lacking collagen XVII display a migratory phenotype. *Am J Pathol* 164, 2027-2038.
- Theobald, D.L., and Steindel, P.A. (2012). Optimal simultaneous superpositioning of multiple structures with missing data. *Bioinformatics* 28, 1972-1979.
- Torrecillas, A., Laynez, J., Menendez, M., Corbalan-Garcia, S., and Gomez-Fernandez, J.C. (2004). Calorimetric study of the interaction of the C2 domains of classical protein kinase C isoenzymes with Ca²⁺ and phospholipids. *Biochemistry* 43, 11727-11739.
- Tsuruta, D., Kobayashi, H., Imanishi, H., Sugawara, K., Ishii, M., and Jones, J.C. (2008). Laminin-332-integrin interaction: a target for cancer therapy? *Current medicinal chemistry* 15, 1968-1975.
- Uematsu, J., Nishizawa, Y., Sonnenberg, A., and Owaribe, K. (1994). Demonstration of type II hemidesmosomes in a mammary gland epithelial cell line, BMGE-H. *Journal of biochemistry* 115, 469-476.
- Uhlik, M.T., Temple, B., Bencharit, S., Kimple, A.J., Siderovski, D.P., and Johnson, G.L. (2005). Structural and evolutionary division of phosphotyrosine binding (PTB) domains. *J Mol Biol* 345, 1-20.
- Van den Bergh, F., and Giudice, G.J. (2003). BP180 (type XVII collagen) and its role in cutaneous biology and disease. *Adv Dermatol* 19, 37-71.
- van der Flier, A., and Sonnenberg, A. (2001). Function and interactions of integrins. *Cell and tissue research* 305, 285-298.
- van der Neut, R., Krimpenfort, P., Calafat, J., Niessen, C.M., and Sonnenberg, A. (1996). Epithelial detachment due to absence of hemidesmosomes in integrin beta 4 null mice. *Nat Genet* 13, 366-369.
- Volkov, V.V., and Svergun, D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of applied crystallography* 36, 860-864.
- Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. *Methods Mol Biol* 364, 215-230.
- Wegener, K.L., and Campbell, I.D. (2008). Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (review). *Molecular membrane biology* 25, 376-387.
- Wiche, G., Becker, B., Lubert, K., Weitzer, G., Castanon, M.J., Hauptmann, R., Stratowa, C., and Stewart, M. (1991). Cloning and sequencing of rat plectin indicates a 466-kD polypeptide chain with a three-domain structure based on a central alpha-helical coiled coil. *J Cell Biol* 114, 83-99.

- Wilhelmsen, K., Litjens, S.H., Kuikman, I., Margadant, C., van Rheenen, J., and Sonnenberg, A. (2007). Serine phosphorylation of the integrin beta4 subunit is necessary for epidermal growth factor receptor induced hemidesmosome disruption. *Mol Biol Cell* 18, 3512-3522.
- Wilhelmsen, K., Litjens, S.H., and Sonnenberg, A. (2006). Multiple functions of the integrin alpha6beta4 in epidermal homeostasis and tumorigenesis. *Mol Cell Biol* 26, 2877-2886.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., *et al.* (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67, 235-242.
- Winn, M.D., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57, 122-133.
- Wriggers, W. (2010). Using Situs for the integration of multi-resolution structures. *Biophysical reviews* 2, 21-27.
- Wright, M.D., Geary, S.M., Fitter, S., Moseley, G.W., Lau, L.M., Sheng, K.C., Apostolopoulos, V., Stanley, E.G., Jackson, D.E., Ashman, L.K., *et al.* (2004). Characterization of mice lacking the tetraspanin superfamily member CD151
- Yanez-Mo, M., Alfranca, A., Cabanas, C., Marazuela, M., Tejedor, R., Ursa, M.A., Ashman, L.K., de Landazuri, M.O., and Sanchez-Madrid, F. (1998). Regulation of endothelial cell motility by complexes of tetraspan molecules CD81/TAPA-1 and CD151/PETA-3 with alpha3 beta1 integrin localized at endothelial lateral junctions. *J Cell Biol* 141, 791-804.
- Yang, X., Claas, C., Kraeft, S.K., Chen, L.B., Wang, Z., Kreidberg, J.A., and Hemler, M.E. (2002). Palmitoylation of tetraspanin proteins: modulation of CD151 lateral interactions, subcellular distribution, and integrin-dependent cell morphology. *Mol Biol Cell* 13, 767-781.
- Yauch, R.L., Kazarov, A.R., Desai, B., Lee, R.T., and Hemler, M.E. (2000). Direct extracellular contact between integrin alpha(3)beta(1) and TM4SF protein CD151. *J Biol Chem* 275, 9230-9238.
- Zhou, M.M., Ravichandran, K.S., Olejniczak, E.F., Petros, A.M., Meadows, R.P., Sattler, M., Harlan, J.E., Wade, W.S., Burakoff, S.J., and Fesik, S.W. (1995). Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature* 378, 584-592.
- Zimina, E.P., Bruckner-Tuderman, L., and Franzke, C.W. (2005). Shedding of collagen XVII ectodomain depends on plasma membrane microenvironment. *J Biol Chem* 280, 34019-34024

ANEXOS