



## Facultad de Ciencias

Departamento de Biología Molecular

“Desarrollo de una nueva metodología para el análisis sistemático y a gran escala de interacciones proteína-proteína mediante espectrometría de masas: aplicación al estudio del interactoma intracelular de los microdominios enriquecidos en tetraspaninas de linfocitos T humanos”

Memoria presentada para optar al grado de

Doctor en Ciencias por el licenciado

**Daniel Pérez Hernández**

Director de Tesis:

Prof. Jesús Vázquez Cobos

Co-directoras de Tesis:

Dra. María Yáñez Mo

Dra. Inmaculada Jorge Cerrudo

Madrid, 2012



## AGRADECIMIENTOS

Alguien dijo alguna vez que el trabajo científico era un 1% de inspiración y 99% de transpiración. Este trabajo no ha sido el trabajo de una persona, sino el trabajo y esfuerzo de mucha gente, y no sería justo quitar ni un ápice de mérito a todos los que han participado de un modo u otro en esta aventura tan genial que ha supuesto para mí realizar esta tesis.

Como no podía ser de otra manera, la primera persona que se me viene a la cabeza es Jesús. Todo el que le conoce sabe perfectamente de su rigor científico, de su entusiasmo contagioso por la ciencia y de su gran don para la docencia. Para mí ha sido mucho más que un jefe, me ha enseñado a pensar y a razonar, a dudar de todo, a tener tesón y constancia en los momentos más duros de la tesis. Él siempre ha estado ahí para aportar esa chispa que hace que todo sea mucho más fácil y apasionante: gracias.

Todos los sinsabores, frustraciones y malos ratos han sido más llevaderos gracias a Inma, una de las directoras de mi tesis. Desde hace más de 5 años me has acogido y cuidado, me has guiado y has trabajado hombro con hombro hasta el final dándolo todo y enseñándome con humildad y paciencia. Quedará para mí el recuerdo de sus grandes frases: "Dani, ¿no tienes trabajo? Yo te doy trabajo", y la que más me gusta: "estos colores no van bien en la gráfica". Gracias por ser tan buena compañera y mejor amiga.

A María, mi otra directora de tesis, que es una verdadera máquina en el conocimiento de toda la inmunología. Es un placer tener a alguien tan extraordinario como directora: has inculcado en mí la semilla de "las interacciones", y gracias a ti he logrado casar la proteómica y la biología (a veces harto difícil). Quiero agradecer también a Paco Sánchez Madrid (permite que te ponga así en los agradecimientos), la oportunidad que me ha brindado al poder trabajar con él y con su grupo (nombrarlos a todos aquí sería imposible). La sencillez y la humildad con la que explica conceptos tan distantes para mí es algo realmente magnífico. Compartir los momentos de seminarios y los "brainstorming" han sido para mí una lección que jamás olvidaré.

Quiero dar también un especial agradecimiento a mi tutora de tesis: Beatriz, muchas gracias por siempre estar disponible y tener siempre un rato. Siento una especial gratitud a la gente del laboratorio de proteómica. Recuerdo cuando empecé con ellos en los sótanos de la Facultad de Biológicas, siendo sólo 4 personas. Horacio, con ese ritmo tan caribeño que a todo puso su sabor: gracias por enseñarme todo sobre LTQs ("el secreto es darle champú de cariño y burbujitas de amor"). A Pablo, ese compi de tesis gabacho que ahora "europeíza" un poco las Floridas, ¡Viva Michael Jackson!. A Pedro, el otro componente de los Picapiedras, que nos volvía locos con sus teorías raras sobre cuantificación "automática". Después llegó Estefi, a la que agradezco ser siempre tan cercana y tan buena compi, los momentos desayunos nunca serán los mismos. Ahí acabaron nuestras penurias y nos mudamos al nuevo CBM. Ahí se incorporaron Marco y Elena, a los que agradezco su contribución al laboratorio en estos temas siempre esquivos del proyecto QuiXot. Margoth, una genial peruana-brasileña, que aportó un color muy especial en nuestro grupo (¡disidente!!). Llegó Raquel (siempre serás la Reichel, que lo sepas) y en pocos días desbancó a Jesús y pasó a tener el control y el poder del labo ("apunta el

protocolo, apúntalo”). Gracias por traer a mi vida a Jonesy (Alias “Chispas”). Después llegó Juan Carlos, el informático friki que nos ha hecho la vida muchiiiisimo más fácil (y sobre todo más divertida). A la nueva generación, compuesta por Fernando y Pili. Son un relevo que sin duda, a estas alturas, ya han superado con creces a los que nos vamos yendo. Y de ahí nos hemos vuelto a mudar al CNIC (si escribo un libro nos llamaré “los proteómicos errantes”), y hemos tenido la suerte y el privilegio de trabajar con gente de gran calidad científica y humana: Emilio, Enrique y Juan Antonio, provenientes, nada menos, de la mejor escuela de proteómica que se pueda imaginar (¡¡arriba los Ex-Méndez!!). Todo esto aliñado siempre por la gente del Servicio de Proteómica del CBM, nuestros vecinos. Gracias a Anabel, Espe e Inés (los congresos no hubieran sido los mismos sin montarla con tan buen rollo) y a Merche, Yoli, Nuria, y Carlos: ha sido un placer teneros de vecinos y enseñarme tantísimas cosas.

El grupo nunca ha estado cerrado a recoger y recibir a científicos para que nosotros aportáramos nuestro granito de arena a sus maravillosas investigaciones. Quiero agradecer a María, gente estupenda que se ha convertido en una proteómica de pro, a Marta, admirable portuguesa que con su amabilidad y su simpatía nos dejó una marca imborrable. Brian, esa noche irlandesa, Conchi, Jota Jota, Enrique, Carmen Aragón, María Pajares, Mariano, Lenka, Almudena e Irene (el futuro de la ciencia). A la gente del CBM, han sido durante un tiempo mi familia, especialmente esos momentos de relax jugando al fútbol.

Todo esto no habría sido posible sin el apoyo incondicional de mi familia. A pesar de haberles dado la tabarra contándoles siempre frikadas científicas siempre has estado ahí, aguantándome y queriéndome, a pesar de que no entendían la mitad de lo que decía. A mi padre, por haber educado mi parte intelectual y racional, siempre picándome con sus detalles culturitas: has hecho de mí una persona mejor y un motor para seguir adelante. A mi madre, que ha educado mi parte emocional. Gracias a ti he superado cosas que nunca creía que iba a superar. Nuestros momentos cocina son de los más felices de mi vida: eres un ejemplo para mí. A mi hermano Marcos: la admiración y afecto que siento no caben en esta tesis (y a la Patri por aguantarle....¡¡tienes el cielo ganado nena!!). A mis amigos, que siempre han sabido llevarme y soportarme, a pesar de ser como soy. Teo, has sido siempre un amigo como no hay otro (Austin Powers se sentiría orgulloso de nosotros), Mario, siempre serás un “carabanchelero”, y por extensión un grandísimo amigo (excepto cuando te contábamos el chiste de Bartolo-mé). Carlitos, eres la fiesta personificada, no cambies nunca. Lucia....¡¡ay mi Lucia!! Antonio Molina a tu lado es un asmático. A los compis de la facultad: Pelayo, Noel, Lolo, Pedro, Albañil, Fran, Yeti, Ali, Borjita.....¡¡El laruso está aquí!!.

Por último, quiero agradecer a la persona sin la cual todo esto no tendría sentido: Marta, gracias por los buenos momentos que hemos pasado, y por los que nos quedan por vivir. Esta tesis es tan tuya como mía. Por aguantar todos los momentos malos, por tener siempre una sonrisa y palabras de ánimo, por entenderme y, sobre todo, gracias por enseñarme lo bonito y sencillo que es disfrutar de la vida junto a ti.

**¡¡GRACIAS A TODOS!!**

# ÍNDICE

INTRODUCCIÓN.....	15
1. LOS DOMINIOS DE TETRASPANINAS: PROTEÍNAS ORGANIZADORAS DE LA MEMBRANA PLASMÁTICA .....	17
1.1. Papel de la tetraspanina CD9 en los procesos de adhesión leucocitaria .....	18
1.2. La tetraspanina CD81, organizadora de los TEMs.....	20
1.3. Asociación de la inmunoglobulina EWI-2 conCD9 y CD81 .....	20
1.4. La molécula de adhesión ICAM-1 es fundamental para el proceso de diapedesis.....	21
2. PAPEL DE LAS TETRASPANINAS EN EXOSOMAS.....	22
3. PROTEÓMICA Y ESPECTROMETRÍA DE MASAS .....	25
4. PROTEÓMICA DE INTERACCIÓN.....	27
5. PROTEÓMICA DE EXPRESIÓN DIFERENCIAL.....	30
5.1. Marcaje metabólico: SILAC .....	31
5.2. Marcaje químico: ICAT e iTRAQ .....	31
5.3. Marcaje enzimático con agua <sup>18</sup> O .....	33
OBJETIVOS.....	37
MATERIALES Y METODOS .....	41
Cultivos celulares y extracción de proteínas.....	43
Preparación de vesículas exocíticas (exosomas) .....	44
Anticuerpos .....	44
Microscopía electrónica.....	44
Diseño de los experimentos de interacción con los dominios intracelulares de las tetraspaninas ...	45
Digestión de proteínas en solución.....	47
Digestión de proteínas en gel de electroforesis .....	47
Digestión de proteínas en gel de electroforesis concentrador.....	48
Marcaje isotópico de péptidos con <sup>18</sup> O y fraccionamiento por isoelectroenfoque y por intercambio catiónico.....	49
Espectrometría de masas mediante LC-ESI-LIT e identificación de proteínas.....	50
Análisis de cuantificación de proteínas.....	51
Análisis estadístico de los datos de interacciones específicas.....	52
Análisis de biología de sistemas mediante herramientas bioinformáticas .....	53

<b>RESULTADOS</b> .....	55
1. DESARROLLO DE UN MÉTODO PARA LA IDENTIFICACIÓN SISTEMÁTICA DE PROTEÍNAS A GRAN ESCALA .....	57
2. VALIDACIÓN DEL NUEVO MÉTODO DE DIGESTIÓN TRÍPTICA EN GEL CONCENTRADOR MEDIANTE MARCAJE CON <sup>18</sup> O.....	60
2.1. Efecto de corte en la digestión en gel separador .....	61
2.2. Estrategia experimental para la cuantificación de proteomas con <sup>18</sup> O usando el nuevo método de digestión en gel concentrador .....	64
2.3. Efecto de la digestión parcial en la varianza a nivel de péptido .....	66
2.4. Aplicación del nuevo método de gel concentrador al estudio de la activación de las células T Jurkat.....	68
3. CARACTERIZACIÓN DEL INTERACTOMA DE LOS DOMINIOS INTRACELULARES DE PROTEÍNAS TEM EN LINFOCITOS T HUMANOS.....	73
3.1. Diseño experimental .....	73
3.2. Estudio de la especificidad de las proteínas interaccionantes con los distintos anzuelos .....	74
3.3. Validación estadística de las interacciones de proteínas por recuento de péptidos .....	78
3.4. Estudio de las interacciones de proteínas con CD81, EWI-2 e ICAM-1 por Biología de Sistemas.....	80
4. CARACTERIZACIÓN DEL INTERACTOMA INTRACELULAR DE CD81 Y EWI-2 EN EXOSOMAS DE LINFOCITOS T HUMANOS.....	81
4.1. Análisis de las proteínas interaccionantes con CD81 y EWI-2 en exosomas de linfocitos T humanos .....	82
4.2. Estudio de las interacciones de proteínas de exosoma con CD81 y EWI-2 por Biología de Sistemas .....	84
4.3. El papel de los TEMs en la composición proteica de los exosomas.....	86
5. ESTUDIO DE LOS CAMBIOS EN EL PROTEOMA DE EXOSOMA POR LA DELECIÓN DE LAS TETRASPANINAS CD9 Y CD81.....	87
5.1. Análisis de los cambios en la abundancia de proteínas en exosomas por la delección de CD9 en ratón.....	88
5.2. Análisis de los cambios en la abundancia de proteínas en exosomas por la delección de CD81 en ratón.....	89
<b>DISCUSIÓN</b> .....	95
1. DESARROLLO DE UNA METODOLOGÍA PARA EL ESTUDIO DE INTERACCIONES DE PROTEÍNAS A GRAN ESCALA MEDIANTE HERRAMIENTAS PROTEÓMICAS.....	97

2. CARACTERIZACIÓN DEL INTERACTOMA INTRACELULAR DE LOS DOMINIOS RICOS EN TETRASPANINAS.....	100
3. EL INTERACTOMA DE LAS TETRASPANINAS EN EXOSOMAS Y SU PAPEL EN LA COMPOSICIÓN PROTEICA .....	101
4. CONSIDERACIONES FINALES .....	103
CONCLUSIONES .....	105
BIBLIOGRAFÍA.....	109
ANEXOS I. TABLAS SUPLEMENTARIAS.....	123
ANEXOS II. FIGURAS SUPLEMENTARIAS.....	141





## RESUMEN

Las tetraspaninas son una superfamilia de proteínas transmembrana, que regulan la señalización mediada por moléculas de adhesión celular, receptores y proteínas estimuladoras y forman parte de unas plataformas organizadas en la membrana llamadas microdominios ricos en tetraspaninas (Tetraspanin Enriched Microdomains, TEM). La expresión ubicua de las tetraspaninas sugiere su relevancia en la estructura y el funcionamiento de la membrana plasmática. Las tetraspaninas también se encuentran en microvesículas secretadas como los exosomas, y regulan la internalización y el reciclaje de sus receptores asociados. Los exosomas son vesículas 30-100 nm de diámetro que se secretan mediante la fusión de cuerpos multivesiculares con la membrana plasmática. Estas microvesículas juegan un papel funcional en la presentación de antígeno y la inmunidad anti-tumoral, y la transmisión de patógenos. A pesar del creciente interés que éstas proteínas están adquiriendo, las interacciones intramembrana por las que están implicadas en varios procesos biológicos son poco conocidas.

Para profundizar en el estudio que desempeñan las proteínas que pertenecen a los TEMs en linfocitos T humanos y en sus correspondientes exosomas secretados, hemos desarrollado un nuevo método de digestión triptica basado en la concentración de proteínas en SDS-PAGE en una única banda para la identificación sistemática de interacciones entre proteínas por técnicas de EM a gran escala. La solidez y reproducibilidad de dicho método se ha validado mediante la técnica de marcaje isotópico estable con  $^{18}\text{O}$  de un conjunto de extractos de proteínas de diferente naturaleza, utilizando un modelo estadístico desarrollado en el laboratorio, demostrando la solidez y fiabilidad del método.

Este protocolo ha permitido el análisis de las muestras procedentes de los *pull-downs* realizados con 11 péptidos biotinilados con la secuencia de los dominios C-terminal de algunas proteínas relacionadas con los dominios TEM (CD9, CD81, CD151, EWI-2, ICAM-1, VCAM-1, CD147, ICAM-3, CD69, CXCR4 y CCR7), incubados con lisados de linfocitos T humanos de distintos donantes. A partir de un análisis estadístico de los resultados, se realizó un estudio por biología de sistemas de los grupos de proteínas interaccionantes identificadas con cada ligando, generando un mapa de interacciones para cada uno de ellos. En función de estos resultados, se eligieron los ligandos más interesantes desde el punto de vista biológico para estudiar las interacciones de proteínas analizando los *pull-downs* de los exosomas secretados por los linfocitos T humanos. Se ha realizado además un análisis estadístico de las interacciones a partir de las proteínas identificadas y un estudio por biología de sistemas de dichas proteínas en exosomas.

Finalmente y, como estudio complementario, se realizó un análisis comparativo del proteoma de exosomas de ratones CD9KO *-/-* ó CD81KO *-/-* con respecto a ratones *wild-type* usando el método de marcaje isotópico con  $^{18}\text{O}$ , que permitió un análisis de los cambios de abundancia de proteínas del exosoma debidos a las deleciones estudiadas.

Este trabajo ha permitido caracterizar la red de interacciones intracelulares de la que forman parte CD81, EWI-2 e ICAM-1. Los resultados derivados de los análisis del interactoma de dichas proteínas sugieren que los TEMs, además de su papel en la adhesión celular, actúan como plataformas especializadas en la membrana plasmática para la compartimentación selectiva de los receptores que van a ser secretados y su correcta exposición en la superficie de exosomas, amplificando así su funcionalidad. La supresión génica de CD81 perturba la inclusión de un repertorio selectivo de proteínas en los exosomas, como la proporción de *rac2* en exosomas, cuya inserción en los microdominios ricos en tetraspaninas está específicamente dirigida por la asociación con CD81.

## SUMMARY

Tetraspanins are a superfamily of transmembrane proteins that regulate signaling processes, cell adhesion, costimulatory receptors and are part of organized platforms in the plasma membrane called tetraspanins enriched microdomains (TEM). The ubiquitous expression of tetraspanins suggests the importance of these proteins in the structure and functionality in plasma membrane. The tetraspanins are also found in secreted microvesicles as exosomes and they may regulate the internalization and recycling of their associated receptors. Exosomes are vesicles of 30-100 nm in diameter, secreted through multivesicular bodies that merge with the plasma membrane. These microvesicles play a functional role in antigen presentation, anti-tumor immunity and pathogen transmission. In spite of the growing interest on these proteins, the interactions inside the membrane by which tetraspanins are involved in various biological processes are poorly known.

In order to study the role of proteins which belong to TEMs in human T lymphocytes and their corresponding secreted exosomes, we have developed a new method based on in-gel protein digestion in a single focused band and high-throughput identification of interacting proteins by large-scale MS. The robustness and reproducibility of this method has been validated analyzing a set of protein extracts by  $^{18}\text{O}$  isotope labeling. Using a statistical model developed in our laboratory, we could demonstrate the reliability of our method.

This protocol allowed for the analysis of several pull-downs performed with 11 biotinylated C-terminal peptide sequences of TEM associated-proteins (CD9, CD81, CD151, EWI-2, ICAM-1, VCAM-1, CD147, ICAM-3, CD69, CXCR4 and CCR7), incubated with human T lymphocytes from different donors. The interacting proteins were statistically validated and were studied by systems biology tools, clustered in groups of interacting proteins, generating a map of interactions with each ligand. Based on these results, we chose the ligands with most biological relevance, studying protein-protein interactions from pull-downs incubated with secreted exosomes of human T lymphocytes. These interactions were statistically validated and also analyzed by systems biology approaches.

Finally, as a complementary study, we performed a comparative analysis of exosome proteome from CD9KO  $-/-$  or CD81KO  $-/-$  mice against wild-type using  $^{18}\text{O}$  isotope labeling, these deletions triggered some changes in protein abundance in exosomes.

We characterized the network of intracellular interactions of CD81, EWI-2 and ICAM-1. The results derived from these analysis suggest that TEMs, in addition to its role in cell adhesion, act as specialized platforms in the plasma membrane and sort proteins that will be secreted, placing them in the correct exposure in exosome surface and amplifying their functionality. CD81 gene deletion disrupts the inclusion of a selective repertoire of proteins in the exosomes, like Rac2, whose insertion into the tetraspanin-rich microdomains is specifically directed by the association with CD81.

## INDICE DE TABLAS

<b>Tabla 1.</b> Eficiencias de digestión y varianzas calculadas a nivel de péptido.....	65
<b>Tabla 2.</b> Varianzas a nivel de espectro, péptido y proteína obtenidas a partir de las células T Jurkat.....	69
<b>Tabla 3.</b> Análisis comparativo de cambios de abundancia producidos por la activación de células T Jurkat .....	70
<b>Tabla 4.</b> Número de péptidos identificados en todos los ensayos de <i>pull-down</i> realizados con lisados de linfocitos T humanos.....	74
<b>Tabla 5.</b> Número de péptidos identificados en todos los <i>pull-down</i> realizados con extractos de exosomas.....	82
<b>Tabla 6.</b> Cambios de abundancia de proteínas detectados en exosomas de linfocitos T de ratones CD9 -/-.....	88
<b>Tabla 7.</b> Cambios de abundancia de proteínas detectados en exosomas de linfocitos T de ratones CD81 -/-.....	91



## INDICE DE FIGURAS

<b>Figura 1.</b> Estructura de tetraspaninas.....	15
<b>Figura 2.</b> La interacciones extracelular, intracelular y lateral en la red de tetraspaninas.....	17
<b>Figura 3.</b> Representación esquemática de la liberación extracelular de microvesículas exocíticas en el espacio extracelular.....	21
<b>Figura 4.</b> Representación gráfica de la composición proteica de exosomas clasificadas por la función realizada.....	22
<b>Figura 5.</b> : Esquema de las distintas aproximaciones usadas para la optimización del método de digestión usado para los pull-downs.....	56
<b>Figura 6.</b> Comparación de los distintos métodos de digestión.....	58
<b>Figura 7.</b> Análisis de los artefactos en la cuantificación por marcaje con <sup>18</sup> O asociados a la técnica de fraccionamiento de proteínas por SDS-PAGE.....	61
<b>Figura 8.</b> Esquema del protocolo de cuantificación propuesto para marcaje con <sup>18</sup> O.....	63
<b>Figura 9.</b> Efecto de la digestión parcial en la cuantificación relativa a nivel de péptido.....	66
<b>Figura 10.</b> Análisis de las normalidad de las cuantificaciones a nivel de proteínas.....	66
<b>Figura 11.</b> El análisis de rango dinámico en extractos citosólicos de Jurkat control y activado.....	68
<b>Figura 12.</b> La precisión de las cuantificaciones a niveles de espectro, péptido y proteína.....	71
<b>Figura 13.</b> Esquema del protocolo seguido para el análisis de alto rendimiento del interactoma de microdominios enriquecidos en tetraspaninas.....	73
<b>Figura 14.</b> Contaje del número de péptidos en donantes humanos para la identificación de proteínas que interaccionan con anzuelos peptídicos.....	75
<b>Figura 15.</b> Especificidad de algunas interacciones en lisados totales de linfoblastos humano.....	77
<b>Figura 16.</b> Número de proteínas que interaccionan específicamente para los anzuelos	

de CD81, EWI-2 y de ICAM-1.....	78
<b>Figura 17.</b> Redes de interacción de proteínas de los ligandos de CD81, EWI-2 e ICAM-1 identificados en linfocitos T humanos.....	79
<b>Figura 18.</b> : Análisis por microscopía electrónica de muestras de exosomas derivados de cultivos primarios de linfocitos humanos.....	80
<b>Figura 19.</b> Distribución del número de péptidos promedio identificados para CD81 y EWI-2 y respectivo control en los <i>pull-down</i> de exosomas de linfocitos T humanos.....	82
<b>Figura 20.</b> La especificidad de algunas interacciones en linfoblastos humanos se confirman en sus exosomas.....	83
<b>Figura 21.</b> Redes de proteínas que interaccionan con los ligandos CD81 y EWI-2 en extractos de exosoma de linfocitos T humanos.....	85
<b>Figura 22.</b> Proporción de proteínas de exosomas que pertenecen al interactoma de los ligandos CD81 y EWI-2. El porcentaje de proteínas identificadas.....	86
<b>Figura 23.</b> Efecto de la delección de CD9 en la composición de proteínas de exosomas en linfocitos T de ratón.....	87
<b>Figura 24.</b> Análisis por microscopía electrónica de exosomas provenientes de ratones “salvajes” (WT) y deficientes en CD81 (CD81-KO).....	89
<b>Figura 25.</b> Efecto de la delección de CD81 en la composición de proteínas de exosomas.....	90
<b>Figura 26.</b> La ausencia de CD81 produce cambios en la composición proteica de exosomas.....	92

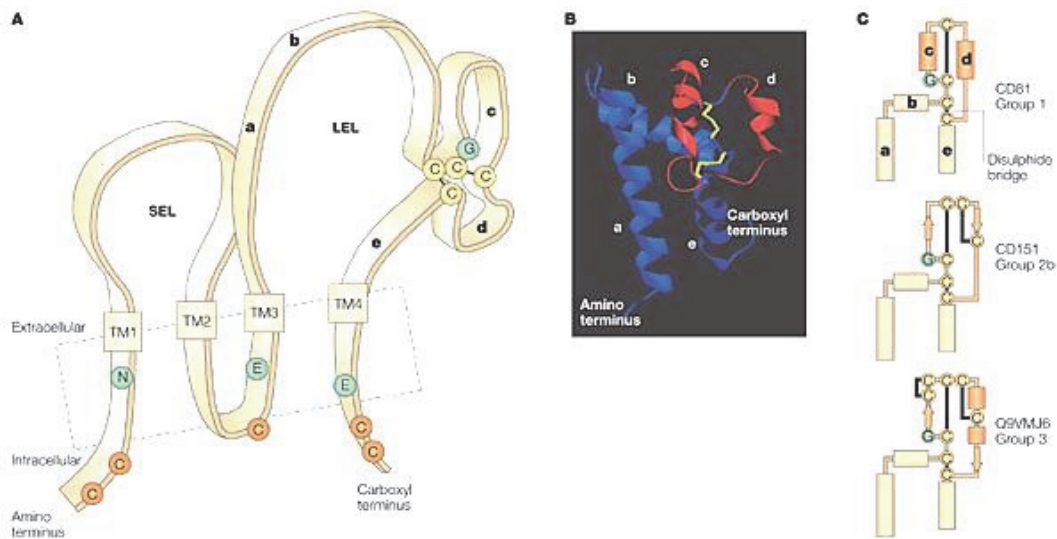
# INTRODUCCIÓN





## 1. LOS DOMINIOS DE TETRASPANINAS: PROTEÍNAS ORGANIZADORAS DE LA MEMBRANA PLASMÁTICA

Los receptores de adhesión presentes en la membrana plasmática presentan una localización subcelular exquisitamente regulada mediante sus asociaciones con el citoesqueleto cortical y su inclusión en distintos microdominios de membrana. Dentro de los microdominios presentes en la membrana plasmática se encuentran las balsas lipídicas (*lipid RAFTs*), las caveolas y los microdominios ricos en tetraspaninas (TEMs). Estos microdominios se generan mediante la asociación lateral de las tetraspaninas con distintos receptores transmembrana (Figura 1), incluyendo integrinas y receptores de la superfamilia de las inmunoglobulinas, entre otros [1]. Los TEMs desempeñan un papel crucial en los procesos de adhesión a la matriz extracelular y a otras células en varios escenarios biológicos [2], como la fusión esperma-óvulo, el desarrollo del embrión, la activación, la migración y la invasión de células del sistema inmune [3] o de las células tumorales [4]. Por otra parte, diferentes patógenos usan estos TEMs como puertas de entrada a la célula o como reservorios intracelulares o rutas de exportación. En este sentido, los diferentes miembros de la familia de las tetraspaninas regulan la infección por virus de hepatitis C [5], el VIH [6], el virus del papiloma [7] o del parásito que causa la malaria [8].



**Figura1: Estructura de tetraspaninas [1].** Dominios cuatro transmembrana de tetraspaninas, con un bucle pequeño y otro más grande, además de una cola citoplásmica pequeña (A). Los dominios conservados de las tetraspaninas (B) que las hacen pertenecer a dicha familia son unos motivos estructurales concretos, además de 4 a 6 residuos Cys extracelulares conservados (C).

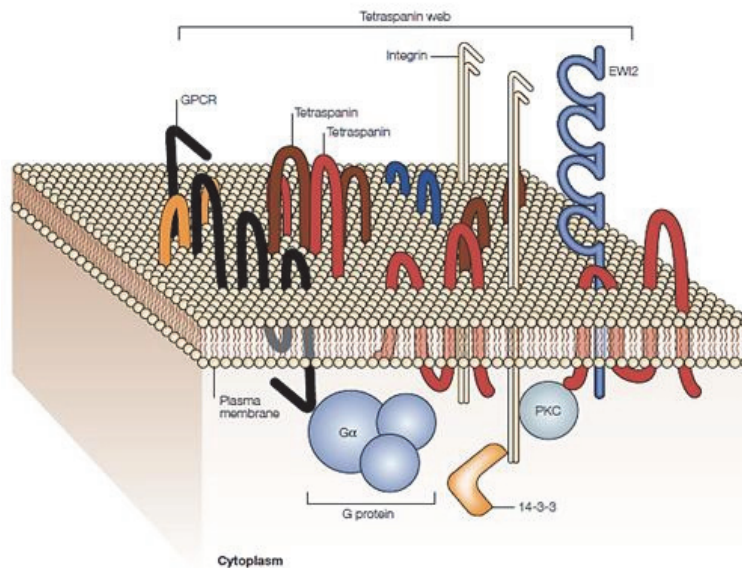
La expresión ubicua de las tetraspaninas, de las cuales 33 miembros se expresan en humanos, sugiere que tienen una gran relevancia en la estructura y la dinámica de la membrana

plasmática. Las tetraspaninas son bastante divergentes a nivel de secuencia, pero mantienen una estructura secundaria común, que comprende cuatro dominios transmembrana que flanquean dos bucles extracelulares: uno corto y otro más largo (LEL), el cual presenta una serie de residuos de cisteína conservados en toda la superfamilia que forman de 2 a 4 puentes disulfuro, y que sustentan la estructura tridimensional de dicho dominio. A través de los dominios LEL, las tetraspaninas interactúan con ellas mismas y con otras proteínas tales como factores de crecimiento, proteínas reguladoras de complemento, proteoglicanos, enzimas de señalización e integrinas [1, 9, 10]. Esta capacidad de interacción es la que permite organizar en la superficie celular los llamados TEMs [11-13] (Figura 2). De esta forma pueden modular la función de las proteínas que pertenecen a dichos microdominios, como la migración dependiente de integrinas [14] o la correcta unión de ICAM-1 y VCAM-1 durante la adhesión leucocitaria y la migración transendotelial [3]. Hay que destacar también que las tetraspaninas se unen directamente al colesterol [15] y a gangliósidos como GM3 [16], lo cual hace que las técnicas bioquímicas clásicas de centrifugación en un gradiente de sacarosa normalmente no sean capaces de discriminar entre los microdominios que se unen a los diferentes tipos de lípidos, que sin embargo tienen características funcionales y estructurales muy distintas [17].

Dentro de los TEMs, en este trabajo nos hemos centrado en el estudio de las tetraspaninas CD9 y CD81 y en los receptores asociados a ellas, concretamente la inmunoglobulina EWI-2 y la molécula de adhesión ICAM-1.

### 1.1. Papel de la tetraspanina CD9 en los procesos de adhesión leucocitaria

La tetraspanina CD9 fue descrita por primera vez como un factor relacionado con la motilidad, capaz de inhibir la migración de múltiples líneas de células cancerosas [18]. Posteriormente, se observó que CD9 está asociado con diversas integrinas, incluyendo  $\alpha3\beta1$ ,  $\alpha4\beta1$ ,  $\alpha5\beta1$ ,  $\alpha6\beta1$ ,  $\alpha6\beta4$  y  $\alpha11\beta3$  [19]. CD9 está además directamente asociado con proteínas de la familia EWI [9], el EGFR (receptor de factor de crecimiento epidérmico) [20] y DDR1 (dominio 1 del receptor de discoidina) [21]. También se han



**Figura 2: Las interacciones extracelular, intracelular y lateral en la red de tetraspaninas [1].** En un modelo hipotético, la red de tetraspaninas (TEM) forma transitoriamente grupos de proteínas de señalización transmembrana e intracelulares, facilitando así una respuesta a diversas señales extracelulares específica y modulada. En los leucocitos circulantes, las integrinas median exclusivamente en respuesta a señales de adhesión y a señales estimulantes endoteliales, como la quimiocinas. Las quimiocinas se unen a receptores de proteínas-G (GPCR) en la superficie celular de los leucocitos, mejorando su avidéz por integrinas y aumentando la adhesión. También son capaces de reclutar partículas de señalización intracelular, incluyendo la proteína quinasa C (PKC), que está involucrada en la fosforilación de integrinas, y a los componentes intracelulares de los complejos de GPCR. Además, CD81 se asocia con una isoforma que pertenece a la familia de proteínas 14-3-3. La familia 14-3-3 ha sido implicado en la regulación de las vías de señalización intracelular. Hay también asociaciones laterales entre tetraspaninas y otros miembros de la superfamilia de las inmunoglobulinas, como por ejemplo EW12.

descrito otras interacciones mediante ensayos de inmunoprecipitación con *syndecan* y CD44 [22] en queratinocitos, con CD36 en macrófagos y en plaquetas [23, 24], y con ICAM-1 [3] y *claudin-1* [25] en leucocitos, demostrándose que esta tetraspanina es fundamental para los procesos de adhesión focal en varios tipos celulares. Además, las interacciones de CD9 con proteínas EWI parecen regular la agrupación de integrinas en la membrana plasmática [26], un proceso crítico durante la migración celular. CD9 también desestabiliza la expresión de EGFR en la superficie [20]. DDR1 es una tirosina quinasa del receptor de colágenos y se ha demostrado que el *knockdown* de DDR1 ó de CD9 inhibe la migración en líneas celulares humanas de cáncer de mama sobre el colágeno [21]. Finalmente, se ha descrito la interacción del LEL de CD9 con fibronectina, posibilitando el aumento en la migración haptotáctica de CHO (células de ovario de hámster chino) [27].

## 1.2. La tetraspanina CD81, organizadora de los TEMs

La tetraspanina CD81 fue originalmente identificada como la diana de un anticuerpo antiproliferativo en células B humanas [28]. Se expresa en gran cantidad de tejidos y está involucrada en una gran variedad de respuestas fisiológicas [29], incluyendo funciones que están relacionadas con la asociación de CD81 con CD19 y moléculas MHC de clase II en las células B, con CD4 en células T y con las integrinas en las células B y T [30]. CD81 destaca por su capacidad de estabilizar y facilitar la organización de sus moléculas asociadas específicas en respuesta a diferentes estímulos [31]. De forma similar a otras tetraspaninas en leucocitos, CD81 tiende a asociarse con varias integrinas. Por ejemplo, se ha descrito la contribución de CD81 en la interacción de la integrina  $\alpha 4\beta 1$  (VLA4) con VCAM1 en el proceso de extravasación leucocitaria [32].

Otra característica de CD81 compartida por varias tetraspaninas es su función co-estimuladora en células T [33, 34]. En ratones, CD81 se expresa en todas las células B, pero sólo en células T activadas [35]. Durante la formación de la sinapsis inmune entre las células T y las B en la estimulación antigénica, CD81 se localiza en la zona de contacto entre los dos tipos celulares, lo cual remarca el papel de CD81 en la presentación antigénica [30]. Los ratones con deficiencia de CD81 no desarrollan la respuesta inmune mediada por los linfocitos T *helper* tipo 2 [36], y desarrollan una respuesta menor frente a alérgenos aéreos [37].

## 1.3. Asociación de la inmunoglobulina EWI-2 con CD9 y CD81

La proteína EWI-2 pertenece a una subfamilia de inmunoglobulinas, llamadas así debido a su motivo extracelular característico de Glu-Trp-Ile (EWI) [38-41]. Esta subfamilia incluye a los siguientes cuatro miembros: EWI-2 (también llamada PGRL en ratón), EWI-3, EWI-F/CD9P-1 y EWI-101. EWI-2 y EWI-F tienen colas citoplasmáticas de 10 y 27 aminoácidos y un dominio extracelular que consta de 4 y 6 dominios de Inmunoglobulinas, respectivamente [39-41]. Estas moléculas se asocian directamente con algunas tetraspaninas, como CD9 (Tspan29) y CD81 (Tspan28) [39-42]. EWI-2 también se asocia con la tetraspanina KAI1/CD82 (Tspan27) [43], cuya expresión se correlaciona inversamente con el potencial metastásico de una gran variedad de tumores [44]. Tanto EWI-2 y EWI-F se asocian además a CD151 (Tspan24), pero en este caso la interacción es indirecta, a través de otras tetraspaninas, como CD9 y CD81 [39, 41]. Se ha demostrado también que se asocia con integrinas a través de su asociación con tetraspaninas [10, 26].

Aunque las funciones biológicas de las proteínas EWI aún no se conocen en profundidad, las asociaciones con tetraspaninas indican que EWI podría influir en las funciones de las tetraspaninas

tales como la fusión celular, la adhesión y la migración [45, 46]. La regulación de la migración celular es característica de tetraspaninas, de ese modo EWI-2 puede regular la motilidad a través de ellos. De hecho, se ha demostrado que la proteína *heat-shock A8* sirve como un ligando para EWI-2 en la superficie de las células dendríticas [47]. Aunque las asociaciones EWI-2 y tetraspaninas están bien documentadas, los elementos estructurales que determinan las interacciones físicas entre EWI-2 y tetraspaninas siguen siendo controvertidos [40, 41]. El ectodominio [40] y la cola citoplasmática [26, 48] de EWI-2 y las regiones transmembrana de tetraspaninas son críticos para la interacción EWI-2-tetraspanina [38].

#### 1.4. La molécula de adhesión ICAM-1 es fundamental para el proceso de diapedesis

La molécula de adhesión intercelular-1 (ICAM-1) se expresa en varios tipos celulares, incluyendo los leucocitos y las células endoteliales [49]. La región extracelular presenta un dominio inmunoglobulina [49] y está unida a una región transmembrana hidrofóbica única (24 residuos) y a una cola corta citoplasmática (28 residuos), que tiene un residuo de tirosina importante para la señalización [50]. Se han identificado varios ligandos extracelulares de ICAM-1 en leucocitos, incluyendo las integrinas LFA-1 (CD11a/CD18) y Mac-1 (CD11b/CD18) [51], el fibrinógeno [52], y algunos rinovirus [53]. ICAM-1 juega un papel importante tanto en la respuesta inmune adaptativa como en la innata. Está implicado en la migración transendotelial de los leucocitos a sitios de inflamación, así como en las interacciones entre las células presentadoras de antígenos (APC) y las células T, durante la formación de la sinapsis inmunológica [54, 55].

Recientemente se ha identificado ICAM-1 en los TEMs [56]. Después de la adhesión leucocitaria, estos microdominios reclutan ICAM-1 y VCAM-1, aumentando la avidéz de estos ligandos de integrinas y, por tanto, favoreciendo la adherencia firme del leucocito. Para estas adhesiones, el vínculo con el citoesqueleto de actina es fundamental. Las proteínas de señalización que se ensamblan y se activan por acción de ICAM-1 son similares a las que se suelen encontrar en adhesiones focalizadas. Las consecuencias de la adhesión endotelial son la asociación de proteínas adaptadoras, como las proteínas ERMs, la *filamin*, la  *$\alpha$ -actinin*, y la *cortactin* (proteína de unión al complejo *Arp 2/3* que regula la dinámica de la actina). La *cortactin* precisa fosforilarse por *Src* para activarse [57]. Dicha fosforilación produce la activación de *RhoA* [58] y la polimerización de actina [59]. Al mismo tiempo, se activan las vías de señalización adicionales implicadas en la regulación del contacto célula-célula en células endoteliales [60]. Junto a las proteínas ERMs, ICAM-1 se une a la  $\alpha$ -

*actinin* a través de residuos básicos en la región yuxtamembrana de ICAM-1 [61, 62]. La asociación con  $\alpha$ -*actinin* -1 y -4 ha demostrado ser funcionalmente relevante, ya que la reducción de expresión de  $\alpha$ -*actinin* en células endoteliales humanas (HUVEC) bloquea parcialmente la migración transendotelial en neutrófilos [62]. La  $\alpha$ -*actinin* es una actina homodimérica que permite la actividad motora de la *myosin* en la contractibilidad. Además, la  $\alpha$ -*actinin* se localiza en las estructuras de acoplamiento endotelial, y probablemente regula la rigidez del citoesqueleto [63]. Recientemente se ha visto que *filamin B* se asocia directamente a los dominios intracelulares de ICAM-1 [64], pudiendo controlar la movilidad lateral de ICAM-1 en la membrana. La función de la *filamin B* es la de unir ICAM-1 al citoesqueleto cortical de actina, aunque lo más interesante es la conexión entre la *filamin* y la *caveolin-1*. En un principio se sabía que *caveolin-1* sólo se asociaba a *filamin A*, y no a *filamin B* [64]. Queda por determinar si la interacción *filamin A*-CAV1 se asocia selectivamente en las regiones TEMs del leucocito.

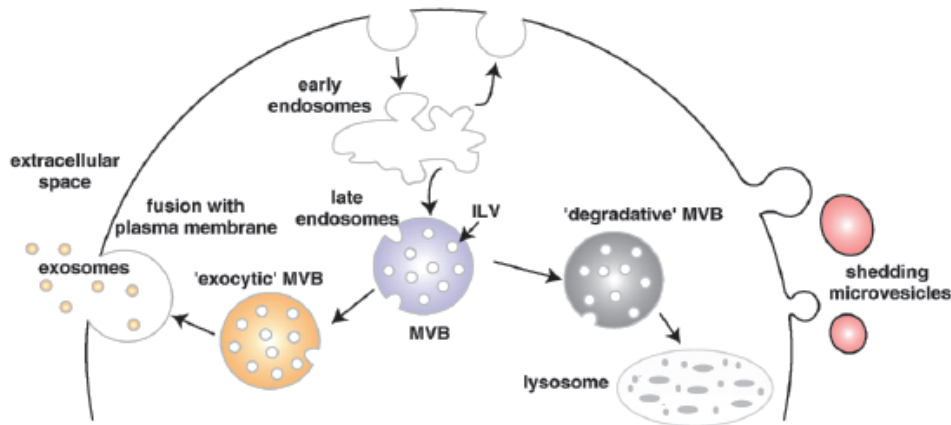
Otra interacción de ICAM-1 confirmada es al dominio SH3 que contiene *Rho-GEF* (SGEF). Los estudios bioquímicos mostraron que SGEF se asocia directamente, a través de su dominio SH3, a la parte C-terminal de ICAM-1, que contiene sitios de reconocimiento SH3 de clase I y de clase III [65, 66]. Cuando ICAM-1 interacciona, *RhoG GTPase* se activa, presumiblemente por SGEF [67], y se forma una estructura de anclaje en torno a los leucocitos adherentes [63, 65]. Aún no se conocen las implicaciones de la interacción ICAM-1-SGEF y RhoG en los TEMs de leucocitos. Además de SGEF, el trío de *Rho/Rac GEF* también puede ser activado por el acoplamiento de ICAM-1, ya que se asocia a *filamin* [68].

## 2. PAPEL DE LAS TETRASPANINAS EN EXOSOMAS.

Las tetraspaninas también regulan la internalización y reciclaje de sus receptores asociados [69] y se encuentran en alta concentración en microvesículas secretadas, como exosomas o ectosomas. Los exosomas son vesículas de 30-100 nm de diámetro que son secretadas por la fusión de cuerpos multivesiculares con la membrana plasmática [70].



La formación de exosomas comienza con la internalización de los endosomas, donde se van acumulando progresivamente las vesículas intraluminales (ILVs). Estas ILVs se almacenan dentro de los grandes cuerpos multivesiculares (MVBs). Los componentes proteicos de los ILVs se irán incorporando por diferente vía según su naturaleza: Las proteínas transmembrana desde la membrana invaginada y las proteínas citosólicas se engloban dentro de los ILVs [71]. Los MVBs pueden dirigirse hacia los lisosomas donde son degradados por el proteasoma, o pueden dirigirse a la membrana plasmática (MP) donde se fusiona con ella y liberan sus contenidos (ILVs) al espacio extracelular (Figura 3). Los ILVs que son liberados al espacio extracelular son nombrados “exosomas” [72]. La liberación de exosomas se ha asociado con la atenuación de importantes reguladores de señal, tales como  $\beta$ -catenin [73], implicada en el mecanismo de señalización de *Wnt*. Por otra parte, la expresión de ciertas tetraspaninas como CD82 (Tspan8) en exosomas son responsables de procesos tumorigénicos y angiogénicos [74]. Se ha demostrado que estas microvesículas llevan material genético, como ARNm y ARNm, que puede ser transferido a otras células, donde pueden



**Figura 3: Representación esquemática de la liberación extracelular de microvesículas exocíticas en el espacio extracelular [1].** Se muestra la liberación de exosomas y de MVB (cuerpos multivesiculares). En los endosomas tempranos las proteínas son recicladas a la membrana plasmática o reclutadas por ILVs (vesículas intraluminales) de las grandes MVBs. ILVs de MVBs se generan por gemación de la membrana en el lumen de los endosomas [3]. Debido a las propiedades biofísicas, los MVBs pueden ser degradativos (evolución a lisosomas) regulado por ESCRT o por ubiquitinación, o pueden ser exocítica (es decir, se fusionan con la membrana plasmática y se produce el lanzamiento de sus contenidos en exosomas). Los MVBs son liberados por el proceso de desprendimiento de la membrana plasmática.

ser traducidos y mediar el silenciamiento de ARN [75-80]. Este proceso es específico de la célula diana, de tal manera que el ARN sólo se transcribe en un tipo de célula [81]. De este modo, los exosomas constituyen un modo de comunicación intercelular que es importante en la respuesta inmune [82], y en la propagación de agentes de señalización célula-célula [83]. En particular, la abundancia relativa de las proteínas, ARNm y ARNm difiere entre los exosomas y las células

donantes. Esto implica que se produce una selección en los MVBs [84]. En el caso de la selección de proteínas, se puede lograr por monoubiquitinación, por su localización en microdominios de membrana ricos en colesterol o por una oligomerización de orden superior [85, 86].

Los exosomas contienen proteínas de distinta naturaleza, como Alix, TSG101, HSP70 y las tetraspaninas CD63, CD81 y CD9. La base de datos ExoCarta [87] contiene las proteínas y los ARN que se han identificado en exosomas provenientes de multitud de líneas celulares y fluidos biológicos que se han analizado por EM, Western blotting, microscopía electrónica y selección (“*sorting*”) celular. La composición proteica de los exosomas varía dependiendo del tejido o del tipo

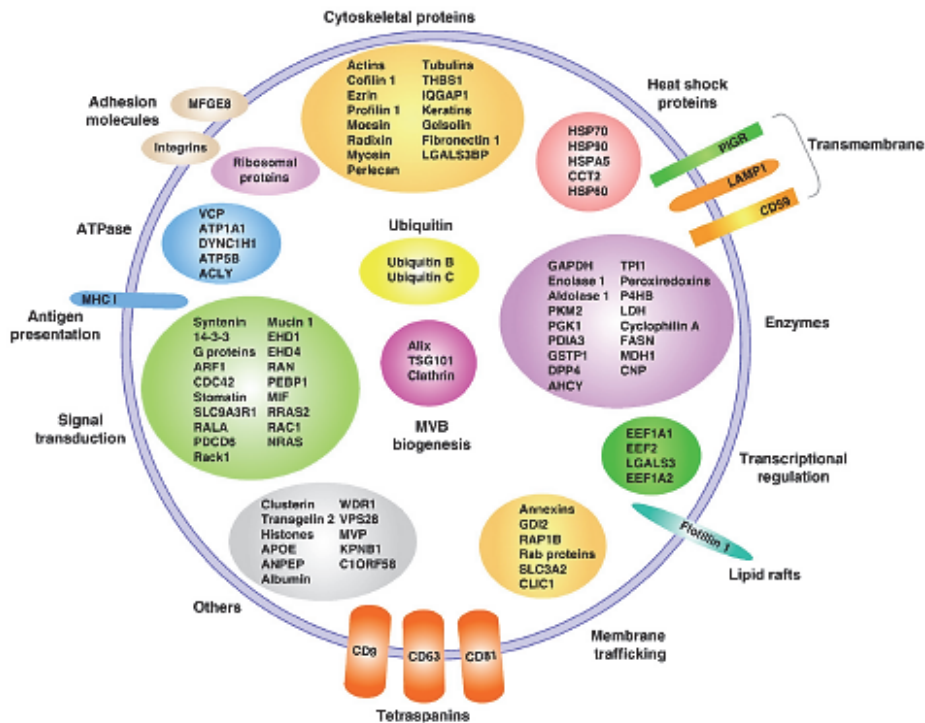


Figura 4: Representación gráfica de la composición proteica de exosomas clasificadas por la función realizada [2]. La base de datos ExoCarta [41] fue utilizada para descargar 19 estudios proteómicos de exosoma que habían identificado al menos 30 proteínas. Las proteínas que se identifican en más del 26% de estos estudios proteómicos están representadas en la figura como símbolos de genes o nombres de proteínas.

celular del que procedan [88], aunque existe un grupo de proteínas que se identifican en la mayoría de los análisis independientemente del origen celular (Figura 4). Los Rabs son unas proteínas citosólicas que se encuentran comúnmente en exosomas. Son de la familia de las GTPasas pequeñas, que regulan el acoplamiento del exosoma y la fusión de la membrana [89].

Además de Rabs, los exosomas son ricos en anexinas (Anexinas I, II, IV, V, VI, VII y X1), que ayudan en procesos de fusión [85]. Los exosomas están enriquecidos en algunas tetraspaninas (CD63, CD81 y CD9) [88], y en proteínas de choque térmico (HSP60, HSP70, HSPA5, CCT2 y HSP90), y



MHC II (moléculas presentadoras de antígenos) [86, 90]. También se encuentran proteínas con dominios RAS (como las Rabs) [91], enzimas metabólicas (*enolase 1*), proteínas ribosomales (RPS3), proteínas transmembrana (LAMP1), moléculas de señalización (*syntenin*, proteína 14-3-3), proteínas G (*cdc42*), moléculas de adhesión (integrinas), ATPasas (ATP1A1), y proteínas del citoesqueleto (proteínas ERM, tubulinas, *actin*) [88] (Figura 4).

Sin embargo, aunque la contribución de las tetraspaninas a la composición proteica de exosomas se conoce bien, su impacto en las funciones de exosomas no ha sido aún bien definido. Esto ha suscitado un elevado interés por las tetraspaninas y sus proteínas relacionadas. Hasta hace poco tiempo existían pocas herramientas para explorar de forma masiva y sistemática la relevancia funcional de estas moléculas, y la mayoría de los estudios se basaban en el empleo de anticuerpos monoclonales. En la actualidad, nuevas claves sobre la función de los TEMs están siendo elucidadas mediante el uso de ARNmi y ratones deficientes en la expresión de estas proteínas, así como con el uso de péptidos bloqueantes solubles que contienen la secuencias del LEL [2]. Aparte de las proteínas descritas anteriormente, tales como ERMs [92], la 14-3-3 (reguladora de *protein Kinase C*) [93] y la proteína *syntenin-1* [94], los posibles sitios de interacción citoplasmáticos de las tetraspaninas y sus proteínas asociadas no se conocen todavía. Esto se debe a que el tamaño de las secuencias intracelulares de estas proteínas es pequeño y, por tanto, se le ha prestado poca atención a la posibilidad de asociaciones a nivel citoplásmico. Nuestro grupo ha generado una serie de péptidos biotinilados con la secuencia de los dominios C-terminal de algunas tetraspaninas y de proteínas asociadas. Utilizando estos péptidos como anzuelos, en este trabajo nos hemos propuesto realizar un *screening* extensivo mediante técnicas de proteómica por espectrometría de masas con el fin de intentar identificar ligandos potenciales de proteínas asociadas a dominios tetraspaninas.

### 3. PROTEÓMICA Y ESPECTROMETRÍA DE MASAS

El término “Proteoma” fue introducido por Williams en 1994, y lo definió como “el equivalente proteico de un genoma”, concretando más tarde como “el conjunto de proteínas expresadas por el genoma de una célula, tejido u organismo completo” [95]. Con el tiempo, el concepto de proteoma ha sido matizado, y la disciplina que la estudia, la proteómica, ha sido también definida de varias formas, debido al alto dinamismo que presenta el sistema biológico a estudiar y también al extenso avance que este campo ha desarrollado en los últimos años. Un proteoma podría definirse como “el conjunto de proteínas que presenta un organismo, un tejido,

una célula, un orgánulo subcelular o fluido biológico en un momento determinado y en unas condiciones precisas”.

La proteómica se define, por tanto, como “la disciplina que comprende el grupo de metodologías orientadas al estudio del conjunto de proteínas expresadas por un sistema biológico, ya sea tejido, grupo de células, órganos subcelulares, o líquidos biológicos, en un tiempo determinado y bajo unas condiciones específicas”. Su objetivo fundamental es la completa descripción del conjunto de proteínas (proteómica cualitativa), el análisis comparativo de los cambios en los niveles de abundancia del conjunto de proteínas (proteómica cuantitativa), el análisis de los cambios producidos en las proteínas discretas modificando su actividad biológica (modificaciones post-traduccionales), y el análisis de las relaciones que tienen el conjunto de proteínas entre sí, es decir, el estudio de la interacción proteína-proteína. El camino trazado en los últimos años por los estudios en proteómica ya permite logros tan importantes como el análisis dinámico de los complejos proteicos de organismos [96] o el análisis cuantitativo del proteoma total de un organismo [97] lo que revela la importancia biomédica y la amplia aplicación que se aporta desde este campo al conocimiento de los mecanismos moleculares de los procesos biológicos, fisiológicos y patológicos, con una globalidad y perspectiva no alcanzada anteriormente.

La proteómica “clásica”, también conocida como Escuela Europea, es la aproximación experimental basada en el análisis de proteínas de forma aislada, separándolas mediante técnicas de electroforesis monodimensional (1-DE) o bidimensional (2-DE) en geles de poliacrilamida, aplicado directamente al proteoma de estudio. Los “spots” provenientes de geles 2-DE o las bandas de proteínas de los geles 1-DE se digieren con tripsina. De modo convencional se utiliza un MALDI-TOF para analizar por EM cada una de las proteínas aisladas. El MALDI-TOF es un espectrómetro de masas donde la muestra se incorpora junto con moléculas de matriz y se someten a irradiación por láser (MALDI). El láser promueve la formación de iones moleculares y se produce la separación de los iones formados en el espacio, por un tubo de vacío atendiendo a la relación masa/carga ( $m/z$ ) de los péptidos (TOF). Dicha combinación permite la identificación rápida y precisa de las proteínas a partir del mapa o “huella peptídica” (PMF) [98-100]. Cuando la identificación de la proteína no es inequívoca, la identificación se realiza a través de sus péptidos mediante fragmentación en una cámara de colisión. En la cámara de colisión, los péptidos ionizados son acelerados por un potencial eléctrico hacia niveles de alta energía cinética, y colisionan con las moléculas neutras de un gas (normalmente He, N o Ar) [101]. El método de fragmentación por CID (collision-induced dissociation) produce generalmente la fragmentación de los péptidos por el enlace peptídico. Dicha fragmentación genera los llamados espectros de fragmentación (también llamados espectros

MS/MS o MS<sup>2</sup>), que permitirán conocer la secuencia de aminoácidos por diferencias en la  $m/z$  entre los diferentes fragmentos. Esta identificación se realiza normalmente de manera automática comparando los espectros teóricos generados en las bases de datos mediante motores de búsqueda. Las aproximaciones de identificación a través de PMF y de espectros MS/MS son complementarias, y permiten la identificación de cualquier proteína de manera universal.

La metodología basada en 2-DE presenta una serie de limitaciones, como su falta de reproducibilidad, la gran dificultad técnica y el limitado rango dinámico inherente a la existencia de proteínas muy abundantes en relación a proteínas en muy baja concentración en la célula. Esto dificulta la identificación de las proteínas poco abundantes y, además, las proteínas con un punto isoelectrico ( $pI$ ) fuera de los rangos de enfoque, no se pueden identificar. Para salvar las limitaciones que esto suponía, se introdujo un método basado en el análisis a nivel de péptido, denominada “proteómica de segunda generación”, donde los extractos de proteínas son digeridos en solución y esta mezcla compleja de péptidos es la que se trata por medio de técnicas de separación cromatográficas (p.e. intercambio catiónico fuerte y posteriormente fase reversa), previo al análisis por EM [102]. El fraccionamiento del proteoma en 1DE-SDS-PAGE seguido por la digestión de fracciones del gel (que contienen potencialmente cientos de proteínas) ofrece una alternativa a la digestión en solución. Este método no sólo puede ser más eficaz para el análisis de proteínas hidrofóbicas gracias a la solubilización completa de las proteínas con SDS, sino que, además, permite la eliminación de detergentes y otros contaminantes que puedan dificultar la digestión con tripsina y puedan interferir en el análisis por EM. La digestión en gel de 1-DE-SDS-PAGE permite la cuantificación de los proteomas de cultivos celulares a una profundidad de varios miles de proteínas [103, 104]. Muy recientemente se ha descrito un método que combina las ventajas de la solubilización de las proteínas con SDS y de la digestión en solución en presencia de urea: es la digestión en filtros de centrifuga o FASP [105].

## 4. PROTEÓMICA DE INTERACCIÓN

Uno de los grandes retos de la proteómica actual es estudiar el mapa de uniones entre proteínas y la dinámica de los complejos macromoleculares, para dilucidar su composición proteica y la implicación funcional de estas interacciones bioactivas. Estos proyectos responden a la idea de que las proteínas no actúan de forma aislada sino a través de grandes complejos macromoleculares. Dependiendo del objeto de estudio, existen diferentes aproximaciones metodológicas para el análisis de las interacciones entre proteínas:

Una de las técnicas más usadas en la proteómica de interacción son las basadas en la afinidad, que incluyen los experimentos de *pull-down*, donde una molécula anzuelo se usa para unir

las proteínas de interés y sus proteínas asociadas [106-108]. Otro tipo de aproximaciones son las basadas en inmunoprecipitación [109]. La diferencia fundamental que existe entre una técnica y otra es que la inmunoprecipitación utiliza anticuerpos para la “pesca” de la proteína anzuelo, mientras que en las basadas en afinidad no se emplean anticuerpos. Por consiguiente, los anzuelos en la inmunoprecipitación serán proteínas o epítopos proteicos [110] mientras que en las estrategias basadas en afinidad pueden ser de diferente naturaleza, ya sea un fármaco, un péptido, un ácido nucleico, etc. [111, 112].

En ambas aproximaciones, los componentes proteicos reconocidos específicamente por el anzuelo, independientemente de su naturaleza, se eluyen y se fraccionan por SDS-PAGE. Las bandas de proteína son digeridas en gel y las mezclas de péptidos resultantes son analizadas mediante técnicas de LC-MS/MS, lo que conduce a la identificación de las proteínas mediante la búsqueda en la base de datos [113]. Este tipo de análisis es el que normalmente se prefiere a otros métodos existentes debido a que el SDS-PAGE, aunque tiene una pobre resolución, aumenta la recuperación de proteínas de baja solubilidad y es lo suficientemente sensible para permitir el análisis de pequeñas cantidades de complejos de proteínas. Por lo tanto, ya que las bandas de proteínas de SDS-PAGE a menudo contienen varias proteínas, las metodologías basadas en EM son capaces de manejar mezclas de péptidos muy complejos, y se emplean rutinariamente para la identificación de proteínas interaccionantes [114, 115].

En los métodos basados en la afinidad, el anzuelo se une covalentemente a un soporte insoluble (por lo general de agarosa/*Sepharose*). En una variante más generalizada se puede marcar el anzuelo con biotina. En este caso, la inmovilización del anzuelo por esta técnica se produce al recubrir con moléculas de *Streptavidin* la superficie inerte [116]. Estos sistemas proporcionan un efecto mínimo sobre la estructura terciaria y la actividad biológica del anzuelo, previniendo la inestabilidad del complejo [110].

Aunque estas técnicas son muy empleadas, la captura por afinidad adolece de un gran número de inconvenientes, ya que, durante la preparación del lisado celular, la arquitectura de los compartimentos subcelulares se rompe y las proteínas que normalmente están separadas en diferentes orgánulos pueden entrar en contacto, generando interacciones específicas pero no fisiológicas. Además, todas las interacciones tienen lugar en microesferas o superficies inertes donde las condiciones de la interacción son distintas a las condiciones en la célula, tales como la relación molar entre el anzuelo y sus proteínas asociadas.

Las técnicas de inmunoprecipitación se pueden realizar de manera directa, con los anzuelos unidos al anticuerpo, y éstos unidos a su vez a proteínas inmovilizadas (proteína A ó G

generalmente) y realizar posteriormente la incubación con el extracto proteico. También se puede realizar primeramente la incubación de la proteína anzuelo con el extracto proteico (o tener la proteína anzuelo de forma endógena en el extracto), e incorporar la forma soluble del anticuerpo, produciéndose la unión del anticuerpo a la proteína anzuelo, que ya tendrá incorporadas sus proteínas asociadas. Posteriormente se recuperarán los complejos extrayéndolos con proteínas inmovilizadas que unen anticuerpo [117]. En ambos casos, después de varios lavados, las proteínas interaccionantes y el anzuelo se eluyen a partir de las superficies inmovilizadas en condiciones desnaturizantes, y serán analizadas como se describe anteriormente. La inmunoprecipitación se realiza cuando la proteína anzuelo es suficientemente abundante como para permitir que una cantidad relativamente grande de complejos pueda ser aislada. Este método tiene una serie de ventajas, manteniendo los rendimientos de la inmunoprecipitación altos y la contaminación de la muestra bajo, disminuyendo la posible aparición de falsos positivos. Debido a que el anticuerpo está unido covalentemente a la superficie inmovilizada, la elución de los inmunoprecipitados puede realizarse en condiciones suaves. Por otra parte, la muestra final no está contaminada por las cadenas de anticuerpos, aumentando así la identificación de proteínas [114].

Al comparar experimentos de alto rendimiento de afinidad e inmunoprecipitación, apenas existe solapamiento entre los complejos de proteínas interaccionantes. Gavin et al. [118] y Krogan et al. [119], en sendos trabajos de caracterización de interacciones en levadura, fueron capaces de identificar más de 3000 proteínas diferentes en total (es decir, el 47% del proteoma de *Saccharomyces cerevisiae*) [120]. Sin embargo, cuando se compararon los resultados, sólo 1000 proteínas fueron comunes a ambos conjuntos de datos. Además de los efectos de las diferentes herramientas bioinformáticas y las diferencias posibles en la purificación, estos resultados podrían reflejar la dinámica de los procesos celulares. Las proteínas se asocian transitoriamente en complejos funcionales para ejecutar un proceso específico y luego se disocian rápidamente, y cada uno de ellos es libre para formar complejos que intervienen en diferentes mecanismos, representando un problema adicional para la proteómica de interacción [121].

La aplicación que se deriva de esta estrategia da la posibilidad de encontrar nuevas relaciones entre proteínas, antes desconocidas. El éxito de la metodología dependerá en gran medida de la especificidad de reconocimiento del “anzuelo-proteína asociada”. De este modo se irá conociendo la red de interacciones de proteínas, y se irá construyendo un mapa físico de las interacciones existentes en un sistema biológico. Un paso hacia esta meta ha sido la descripción del mapa completo del microorganismo *Helicobacter pylori* [122]. Esto se consiguió utilizando el método de doble híbrido de levadura (“*yeast two hybrid*”) para la detección de interacciones entre proteínas,

identificando 1200 conexiones entre las proteínas que cubren el 46.6% del genoma [122]. Otro amplio estudio fue el análisis realizado en todas las proteínas de la levadura *S. cerevisiae* [118], utilizando el mismo método de doble híbrido. La evidente complejidad de estos mapas sugiere que la funcionalidad de las proteínas no debe considerarse de forma aislada, sino dentro de una red global de componentes que interaccionan entre ellos. Por esta razón, se viene incluyendo en los últimos tiempos el término de “proteómica funcional” [123], que se refiere a las diversas aproximaciones proteómicas que permiten el estudio y caracterización de un grupo de proteínas determinado. Estos estudios permiten diferenciar entre la actividad de una proteína y su función biológica. La actividad de una proteína describe si la propia proteína es, por ejemplo, una enzima, un transportador, un receptor o un andamio estructural, pero esta actividad puede ser ejercida en procesos biológicos totalmente diferentes, según las necesidades de la célula. La función biológica de una proteína es entonces relacionada con el proceso específico en el que la proteína está realmente participando. En consecuencia, dado que la misma proteína puede ensamblarse con diferentes proteínas para formar diferentes complejos funcionales, aunque presente una sola actividad puede tener varias funciones biológicas. Gracias a la proteómica de interacción y a las herramientas bioinformáticas de biología de sistemas se pueden construir estos mapas completos de todas las interacciones de proteínas en la célula y mostrar la función biológica de las proteínas de un organismo a nivel global.

## 5. PROTEÓMICA DE EXPRESIÓN DIFERENCIAL

Las herramientas de proteómica cuantitativa basadas en la espectrometría de masas juegan un papel fundamental para entender los procesos fisiológicos y patológicos debidos a los cambios dinámicos del proteoma en la investigación biomédica. El desarrollo de nuevas técnicas de segunda generación y la incorporación de estrategias cuantitativas han sufrido un gran progreso en los últimos tiempos. Este tipo de aproximaciones se basa en la comparación de dos o más estados diferentes de un mismo proteoma, usando dilución isotópica estable, que les confiere una diferencia de masa en los espectros MS o MS/MS de cada uno de los péptidos analizados. La intensidad relativa evalúa la proporción exacta de cada proteína presente en las muestras originales. En estas técnicas mediante marcaje con isótopos estables, dicho isótopos pueden ser introducidos en la muestra mediante la incorporación metabólica de aminoácidos marcados (SILAC,[124]), mediante una reacción química con un reactivo de grupo marcado isobáricamente (iTRAQ, [125] e ICAT, [126]) o mediante un marcaje enzimático, mediado por la introducción de átomos de  $^{18}\text{O}$  en el aminoácido C-terminal de los péptidos trípticos [127].

## 5.1. Marcaje metabólico: SILAC

El marcaje metabólico SILAC (Stable Isotope Labeling by amino Acids in Cell culture) se basa en la incorporación de nutrientes por vía del metabolismo celular *in vivo*, incorporando al medio de cultivo aminoácidos marcados en uno o varios átomos pesados ( $^{13}\text{C}$  y/o  $^{15}\text{N}$ ) [128]. Después de varias generaciones de crecimiento, se incorporan todos los aminoácidos marcados a las proteínas de la célula. Posteriormente, se mezcla una misma cantidad de células provenientes de los dos estados biológicos a comparar, se extraen las proteínas y se digieren con tripsina. Los péptidos resultantes se analizan por espectrometría de masas para su identificación y cuantificación. Dicha cuantificación se realiza comparando las intensidades de los pares de picos correspondientes a péptidos iguales, que se encuentran separados por la diferencia de masa que corresponde al marcaje isotópico, sobre el espectro MS. El péptido se presenta como un doblete de envolturas isotópicas, calculando la diferencia relativa de las intensidades de estas envolturas isotópicas. Normalmente se utilizan residuos de lisina y arginina marcados, ya que la digestión con tripsina asegura que la mayoría de todos los péptidos tendrán su extremo carboxilo marcado. Este tipo de marcaje tiene la ventaja, además, de que presenta una variabilidad analítica muy baja, tras al procesamiento de la muestra, ya que la unión de ambas preparaciones es previa a la extracción de proteínas del lisado celular. Sin embargo, la técnica SILAC sólo es aplicable a cultivos celulares y no a fluidos biológicos o a tejidos. Recientemente se han generado unos modelos *in vivo* con ratones suplementados con dietas que incorporan lisina pesada [129], con pollos alimentados con dietas que incorporan valina pesada [130], e incluso con *Drosophila megalogaster* [131], pero tienen un elevado coste económico de generación y mantenimiento del modelo y, además, no se puede aplicar a modelos humanos.

## 5.2. Marcaje químico: ICAT e iTRAQ

El marcaje químico se realiza *in vitro* y se fundamenta en la incorporación de reactivos químicos marcados isotópicamente que se unen covalentemente a los péptidos o a las proteínas. No tiene restricciones en la aplicabilidad a todo tipo de muestras biológicas, pero existe una mayor variabilidad experimental, porque las muestras marcadas y no marcadas se mezclan sólo después de la digestión, pudiendo generar artefactos que afecten a la cuantificación. Existen diferentes estrategias de introducción de isótopos estables a proteínas o a péptidos por reacción química. La técnica basada en ICAT (Isotope Code Affinity Tag) [95] se compone de un grupo reactivo específico de residuos de cisteína (yodoacetamida), de biotina, y de un espaciador con 8 isótopos de hidrógeno ligero o con 8 isótopos de deuterio. Los extractos de proteínas se reducen previamente y se marcan

isotópicamente, una muestra con la versión “pesada” y la otra muestra a comparar, con la versión “ligera” del reactivo ICAT. Ambas muestras se mezclan y se digieren juntas usando una proteasa (tripsina o endoLys-C) para producir los fragmentos peptídicos. El digerido se somete a una cromatografía de afinidad con avidina, recuperando los péptidos que contienen residuos de cisteínas. Por un lado, esta particularidad simplifica el análisis ya que se obtienen fracciones enriquecidas en péptidos que contienen cisteínas (un residuo poco abundante en el proteoma, sólo un 1.4% de los aminoácidos presentes en las proteínas [132]) pero, a su vez, disminuye la cobertura de la secuencia de las proteínas, que conlleva una menor precisión en la cuantificación. La cantidad relativa de cada proteína se mide a partir de la intensidad de los espectros MS de los péptidos con cisteína, que difieren en 8 Da (8 hidrógenos intercambiables de reactivo) según procedan de la muestra marcada o de la muestra no marcada. El mayor inconveniente de esta técnica es que el reactivo genera reacciones secundarias con los extremos N-terminal de las proteínas y con los residuos de lisina. En el análisis por EM, puede ocurrir un efecto isotópico en la cromatografía líquida previa de los péptidos con deuterio, problemas de ionización de los péptidos y dificultades a la hora de interpretar los espectros de fragmentación por tener biotina. Recientemente se han generado unos reactivos basados en  $^{13}\text{C}$ , llamados cICAT, que evitan el efecto isotópico del deuterio y, además, presentan un extremo hidrolizable en medio ácido que permite separar la biotina de los péptidos antes del análisis por EM [133].

El marcaje isotópico con iTRAQ (Isobaric Tags for Relative and Absolute Quantification) [125] permite la cuantificación relativa de 8 proteomas en el mismo experimento, ya que son 8 posibles reactivos marcadores (aunque normalmente se comparan 4 proteomas a la vez). Otra diferencia respecto a los marcajes anteriores es que la cuantificación se realiza en el espectro MS/MS, y no en los espectros MS. A diferencia de los reactivos de SILAC e ICAT, en este método se marcan los péptidos y no las proteínas. El reactivo iTRAQ se compone de un grupo reportero, un grupo reactivo de unión a grupos amino del extremo N-terminal y de los residuos de lisina y un grupo de balance de masas. Los cuatro (u ocho) reactivos no se diferencian isobáricamente, así los péptidos marcados tendrán el mismo peso molecular y se fragmentarán a la vez. Durante la fragmentación, el enlace entre el grupo reportero y el grupo de balance de masas se rompe fácilmente y libera los iones reporteros monocargados de distinta masa ( $m/z$  de 114, 115, 116 y 117) que son específicos de cada una de las muestras. La intensidad de dichos iones permite la cuantificación relativa de las distintas especies marcadas. La ventaja de cuantificar los péptidos en el espectro de fragmentación es que aumenta la sensibilidad ya que la razón señal/ruido es muy alta, permitiendo usar una cantidad menor de muestra. Existe también el método TMT<sup>®</sup> [134], que funciona de la misma manera y permite comparar hasta 6 muestras a la vez en un mismo experimento.



### 5.3. Marcaje enzimático con agua $^{18}\text{O}$

El marcaje con  $^{18}\text{O}$  es un método enzimático realizado a nivel de péptido donde ocurre un intercambio de dos átomos de  $^{16}\text{O}$  por dos átomos de  $^{18}\text{O}$  durante una reacción catalizada por tripsina, aumentando así el peso molecular del péptido en 4 Da [135]. El mecanismo enzimático consiste en un ataque nucleofilo de  $\text{H}_2^{18}\text{O}$  sobre el grupo carbonilo del extremo C-terminal de los péptidos. Durante el proceso de marcaje se generan dos nuevas especies químicas que contienen uno o dos átomos de  $^{18}\text{O}$ , respectivamente, en el extremo C-terminal del péptido, además de la especie química no marcada, estableciéndose un equilibrio entre ellos. El marcaje se realiza incubando los extractos de proteínas con tripsina en condiciones que permiten la digestión, conteniendo el medio de reacción  $\text{H}_2^{18}\text{O}$  para la muestra a marcar y  $\text{H}_2^{16}\text{O}$  para la muestra control. Aunque originalmente el marcaje enzimático se realizaba conjuntamente durante la digestión triptica, existen diversos factores, como valores extremos de pH, que causan un intercambio de oxígenos con el medio [136]. Hoy en día los procesos de digestión y marcaje se llevan a cabo de forma separada y consecutiva, cada uno de ellos en sus condiciones óptimas [127, 135]. Después del marcaje, se mezclan las dos muestras, detectándose en el mismo espectro MS cada péptido marcado y sin marcar.

La cuantificación relativa de los péptidos se realiza a partir del cálculo de las intensidades de las envolturas isotópicas de las especies marcadas y sin marcar, procedentes de las dos muestras a comparar. Debido a que la incorporación de  $^{18}\text{O}$  puede ser incompleta (incorporándose 1 sólo átomo de  $^{18}\text{O}$ ), es necesario tener en cuenta el solapamiento de las envolturas isotópicas de las tres especies químicas (marcada, marcada incompleta y no marcada) cuando se utilizan espectrómetros de masas de baja/media resolución. Aunque existen algunos trabajos donde usan este tipo de equipos [137-140] no han sido utilizados de forma habitual en proteómica cuantitativa hasta la aparición de la trampa iónica lineal (LTQ), que presenta una velocidad de barrido superior a la trampa iónica tridimensional convencional (3D-IT) [141], y permite modos de barrido de alta resolución muy robustos en un rango de masas limitado (ZoomScan). En una trampa iónica lineal se pueden programar ciclos de barrido que realicen dos espectros, un ZoomScan centrado alrededor del doblete de envolturas isotópicas que se desea cuantificar y un espectro MS/MS para identificar el péptido [117, 142, 143].

El primer trabajo publicado donde se usó este método de marcaje a gran escala se realizó en el 2001, para estudiar el perfil proteico de dos tipos de adenovirus [144]. Este tipo de marcaje

presenta una serie de ventajas con respecto a los demás tipos de marcaje con isótopos estables. En primer lugar, es de aplicación universal, es decir, se puede usar con cualquier tipo de muestras biológicas, ya sea cultivo celular, tejido, fluidos (plasma, orina, etc.). En segundo lugar, se marcan todos los péptidos tripticos de la muestra, por lo que se amplía enormemente la cobertura de la secuencia de las proteínas, en comparación con el método de ICAT. Por otro lado, se puede usar en un rango amplio de cantidad de muestra (del orden de  $\mu\text{g}$  a  $\text{mg}$  de proteínas). No da lugar a reacciones secundarias que pudieran falsear las medidas de cuantificación y, al ser un marcaje enzimático, tiene una especificidad muy alta. Además, el reactivo ( $\text{H}_2^{18}\text{O}$ ) es muy estable y menos costoso que otros tipos de marcaje.

Sin embargo, y a pesar de todas las ventajas que se describen, no es un método tan extendido como los marcajes con iTRAQ y SILAC. Esto se debe, en gran medida, a que siempre se ha considerado un método muy complejo, delicado y difícil de estandarizar, ya que aún falta un protocolo estándar de aplicación general. La eficiencia de marcaje debe controlarse estrictamente porque, al ser un método enzimático, no todos los péptidos se marcan con la misma facilidad. En un trabajo previo realizado en nuestro laboratorio se describe un método para el cálculo de la eficiencia de marcaje. Mediante consideraciones cinéticas se demuestra que la concentración total del péptido marcado puede descomponerse en función de la concentración de cada una de las especies químicas (marcada, marcada incompleta y no marcada), utilizando un parámetro  $f$  que evalúa la eficiencia de marcaje [145]. El control riguroso de la eficiencia de marcaje permite determinar directamente las proporciones del péptido procedente de la muestra marcada y del mismo procedente de la muestra sin marcar, y llevar a cabo experimentos de cuantificación a gran escala al controlar el marcaje de cada uno de los péptidos cuantificados en un mismo análisis [145]. Existen una gran variedad de protocolos experimentales descritos para el método de marcaje con  $^{18}\text{O}$  [138, 144, 146, 147]. En general, la variedad metodológica de este tipo de marcaje reside fundamentalmente en la preparación de la muestra para obtener una digestión eficaz y reproducible, las condiciones en las que se produce el intercambio de oxígenos y el método de inactivación de la proteasa para que no se revierta el marcaje. Más recientemente se ha desarrollado en nuestro laboratorio un método robusto de cuantificación a gran escala de proteomas mediante marcaje con  $^{18}\text{O}$ , separación por punto isoeléctrico (sistema OffGel) y análisis mediante HPLC-LIT-MS [148]. Este protocolo controla perfectamente el marcaje de todos los péptidos cuantificados, demuestra que la separación por OffGel es compatible con el marcaje  $^{18}\text{O}$  en cualquier rango de pH y utiliza SDS para solubilizar las proteínas. Además, se han desarrollado una serie de algoritmos y modelos matemáticos para la cuantificación y el análisis estadístico de los datos obtenidos en experimentos de alto rendimiento de forma sencilla y automatizada,

implementados en la plataforma QuiXoT, que es parte de la tesis del Dr. Pedro José Navarro Álvarez [117, 148].



# OBJETIVOS



El objetivo general de este trabajo es profundizar en el estudio del papel biológico que desempeñan las proteínas que pertenecen a los TEMs en linfocitos T humanos y en sus correspondientes exosomas secretados mediante la identificación sistemática de los ligandos que interaccionan en la región intracelular de estas proteínas.

Para ello, planteamos los siguientes objetivos específicos:

1. Desarrollo de un método robusto para el análisis a gran escala de interacciones y de cambios de abundancia de proteínas mediante espectrometría de masas.
2. Caracterización del interactoma de la región intracelular de las proteínas de TEMs en linfocitos T humanos.
3. Caracterización del interactoma de la región intracelular de las proteínas de TEMs en vesículas exocíticas (exosomas) secretadas por linfocitos T humanos.
4. Estudio del papel de las tetraspaninas en la estructura proteica de exosomas secretados por linfocitos T de ratón.





## **MATERIALES Y METODOS**



### Cultivos celulares y extracción de proteínas

Los linfoblastos humanos se obtuvieron de sangre de donantes sanos por extracción en centrifugación en gradiente de Ficoll-Hypaque y estimulación durante 2 días con PHA ( 1 µg/ml) y posterior cultivo en presencia de IL-2 recombinante humano (50U/ml).

Los linfoblastos T de ratón se obtuvieron de células de bazo por activación con 2 µg/ml de concavalina A (Sigma) durante 36 h. El medio fue sustituido por medio base (RPMI 1640 (Sigma) suplementado con 10% de FBS (Invitrogen), 50 µM de 2-mercaptoetanol (Merck) y 100µ de piruvato de sodio (Lonza) suplementado con IL-2 a 200U/ml (Roche)).

Las líneas celulares humanas HepG2 y SK-N-MC se obtuvieron de American Type Culture Collection (ATCC no. HB-8065 y HTB-10TM). Las células fueron cultivadas en MEM suplementado con 10% de FBS, 2 mM de L-glutamina, 1 mM de piruvato sódico, aminoácidos no esenciales 0.1 mM y 0.05 mg/ml de gentamicina, a 37 °C y al 5% CO<sub>2</sub>. Después las células adherentes se extrajeron por tripsinización a partir de frascos de 175 mm<sup>2</sup> (Corning, Elmira, Nueva York) a 3x10<sup>6</sup> células/placa. Después de 24 h el sedimento celular se lavó dos veces con tampón fosfato salino helado, se resuspendieron y se incubaron durante 30 minutos y 350 µl de tampón fosfato salino helado con 1% de Triton X-100, 0.1mM de EDTA e inhibidor de proteasas (Roche Ciencias Aplicadas) durante 30 min a 4 °C. La suspensión se procesó en un homogeneizador Potter-Elvehjem y se centrifugó a 200×g durante 5 min para eliminar restos celulares. Se recogieron los sobrenadantes y la concentración de proteínas fue determinada por Bradford (Bio-Rad) utilizando BSA como estándar. Los lisados se obtuvieron después de completar la adición de tampón de Laemmli para SDS-PAGE (12.5% de glicerol, 2.5% SDS, 5% de β-mercaptoetanol, 0.01% azul de bromofenol, y 50 mM de Tris pH 6.8). Las mitocondrias cardíacas se donaron generosamente por el laboratorio del Dr David García Dorado, del Hospital Vall d'Hebrón, y fueron aisladas de corazones de rata de la raza Sprague-Dawley. Se sometieron a perfusión retrógrada durante 15 min a 37°C con una solución oxigenada de Krebs en un aparato de Languendorff. Después del equilibrado, los corazones fueron removidos de las cánulas y las mitocondrias cardíacas se obtuvieron por centrifugación diferencial y

ultracentrifugación en gradiente de Percoll, tal y como se describe [149]. La pureza de las preparaciones mitocondriales fue controlado por Western-blot utilizando anticuerpos para otros compartimentos celulares. La concentración de proteínas en extractos mitocondriales se midió utilizando el ensayo de proteínas Bradford.

### Preparación de vesículas exocíticas (exosomas)

Las microvesículas se prepararon a partir del sobrenadante de cultivo primario de linfoblastos tanto humanos como de ratones, cultivados con suero bobino fetal deplecionado de exosomas mediante ultracentrifugación a 100.000xg durante 12h. Las microvesículas se aislaron mediante etapas consecutivas de centrifugación y de filtración [150]: se centrifugaron las células (320xg durante 5 min) y el sobrenadante se filtró a través de las membranas de 0.22  $\mu\text{m}$ . Los exosomas se sedimentaron por ultracentrifugación a 100.000xg durante 60 min a 4 °C (Beckman Coulter Optima L-100 XP, Beckman Coulter). El enriquecimiento en exosomas de linfocitos se evaluó mediante tinción negativa con acetato de uranilo al 2% en un microscopio de transmisión electrónica Jeol JEM1010 como se describe.[150, 151]

### Anticuerpos

Las proteínas se separaron por SDS-PAGE y se transfirieron a membranas de nitrocelulosa, las cuales fueron incubadas con los distintos anticuerpos. El pAb anti-ERMs (90:3) se ha sido descrito previamente [92]. Los anticuerpos anti-filamina, anti- $\alpha$ -actinina, anti-nucleolina y anti- $\beta$ -actina fueron adquiridos de Santa Cruz Biotechnology (Santa Cruz, CA), anti-Rac1 / 2 y anti-EF-1 $\alpha$  fueron de Millipore (Billerica, MA ).

### Microscopía electrónica

Las células se sedimentaron y se fijaron *in situ* de manera convencional con 2% de glutaraldehído (Sigma) y 4% de paraformadehído en PBS durante 1 h a temperatura ambiente.

Las células se incorporaron a resina Epoxi (Taab Laboratories) y después se fijaron en 1% de tetróxido de osmio y 0.8% de ferrocianuro potásico durante 1 hora a 4°C. Después de lavar con agua bidestilada, las muestras se incubaron con 0.15% de ácido tánico durante 1 min. a temperatura ambiente, se lavaron, se incubaron con acetato de uranilo al 2% durante 60 min. a temperatura ambiente, y se deshidrataron con etanol. La infiltración de la resina se llevó a cabo en un gradiente de concentraciones de Epoxi-etanol a temperatura ambiente durante 24 h. La polimerización se realizó a 60 °C durante 48 h. Las secciones ultrafinas se observaron en un microscopio Jeol JEM-1010 de electrones, previa contratinción con acetato de uranilo al 2% y citrato de plomo.

Para el análisis de tinción negativa de los exosomas, las muestras secadas como anteriormente se describe se resuspendieron en PBS. El carbono activo y redes de colodión recubierto de cobre EM emergieron sobre una gota de la muestra, los lavados y teñidos se realizaron en gotas de acetato de uranilo al 2% durante 1 min.

### Diseño de los experimentos de interacción con los dominios intracelulares de las tetraspaninas

Los péptidos biotinilados correspondientes a las secuencias C-terminal intracelular de proteínas tetraspaninas y receptores asociados (o controles) se obtuvieron de Ray Biotech, Inc. (Norcross, GA) con las secuencias correspondientes:

Proteína	Péptido biotiniliado
CD9	biotina SGSG-CCAIRRNREMV
CD81	biotina SGSG- CCGIRNSSVY
CD151	biotina SGSG-YRSLKLEHY

EWI-2	biotina-SGSG-CCFMKRLRKR
ICAM-1	biotina-SGSG-RQRKIKKYRLQQAQKGTMPKPNTQATPP
VCAM1	biotina-SGSG-RKANMKGSYSLVEAQSKV
CD147	biotina-SGSG-KRRKPDQTLDEDDPGAAPLKGSGHHMNDKDKNRQRNAT
ICAM-3	biotina-SGSG-REHQRSYHVREESTYLPLTSMQPTEAMGEEPSRAE
CD69	MSENCFVAENSSLHPESGQENDATSPHFSTRHEGSFQ-GSGSK-biotina
CXCR4	biotina KFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS
CCR7	Biotina KFRNDLFKFLKDLGCLSQEQLRQWSSCRHIRRSSMSVEAETTTTFSP

Se realizaron 53 ensayos de *pull-down* en 7 donantes humanos diferentes para estudiar las interacciones en lisado total de linfocitos T. Para el análisis de las interacciones específicas en exosoma, se realizaron 18 ensayos de *pull-down* con extractos de exosomas procedentes de 6 donantes humanos de los linfoblastos T, incubándolos con los anzuelos CD81, EWI-2 y un control negativo (microesferas de *Streptavidin-Sepharose* sin anzuelo biotinilado unido).

Cada péptido biotinilado en una cantidad de 30 nmol se conjugó con 40µl de microesferas *Streptavidin-Sepharose* (Amersham Biosciences). Las células y las microvesículas de exosomas se lavaron con TBS (tampón salino de Tris) y se lisaron con 1% de Nonidet P-40, 2 mM de CaCl<sub>2</sub>, inhibidores de fosfatasa (1 mM de pirofosfato de sodio, 1 mM de ortovanadato sódico, 10 mM de fluoruro de sodio) e inhibidores de proteasas (Roche Applied Science) en TBS. Los extractos proteicos obtenidos tras la lisis se incubaron 60 min con las microesferas que contienen el anzuelo peptídico unido. Se lavaron 3 veces en tampón de lisis y 10 veces en 50 mM de bicarbonato de amonio pH 8.8, para el análisis por espectrometría de masas y por Western-blot.

### Digestión de proteínas en solución

Las muestras son lavadas con bicarbonato amónico 5 mM a pH 8.8. Posteriormente se utilizó un método convencional de extracción de proteínas por elución ácida usando una solución al 10% de ácido fórmico, incubándose durante 25 mins, aplicando filtros Wizzard (Promega, Madison, WI) y centrifugando 2 min a 13000*xrpm* obteniendo así una solución de proteínas libre de contaminantes que podrían dificultar el proceso de digestión o el análisis por EM. Se secó la muestra por centrifugación a vacío y se procede a la solubilización con bicarbonato amónico 25 mM pH 8.8. Se reducen los puentes disulfuro de las proteínas solubles con DTT 200mM durante 60 min y en agitación y posteriormente se produce la alquilación con iodoacetamida 50 mM incubando las muestras 60 min en agitación y oscuridad. Se procede a la digestión trípica de las proteínas con tripsina (Promega, Madison, WI) a una concentración de 12.5 mM en bicarbonato amónico. Las muestras se incubaron a 37°C 24 h en agitación. Los péptidos resultantes de la digestión trípica se resuspendieron en tampón de carga para el espectrómetro.

### Digestión de proteínas en gel de electroforesis

Los geles SDS-PAGE fueron preparados de acuerdo con el método convencional de Laemmli [151]) (al 4% de gel concentrador y al 10% de gel separador, a partir de una solución al 30% de acrilamida/bis-acrilamida (37:5:1)). Las proteínas se desnaturalizaron hirviéndolas durante 5 min en tampón de carga (50mM de Tris-HCl pH 8.8, 30% (v/v) de glicerol, 2% (p/v) de SDS, y 0.01% (p/v) de azul de bromofenol), incluyendo 25mM de DTT preparado en fresco. La electroforesis se realizó a 70V hasta que llegó al frente separador, donde se cambió el voltaje a 120V hasta su resolución completa. Se tiñó con solución de azul de Coomassie R-250 (Bio-Rad, CA, USA).

Se cortaron los carriles del gel y se fraccionaron en trozos pequeños. En cada uno de dichos fragmentos se reducen los puentes disulfuro de las proteínas solubles con DTT 200mM durante 60 min y en agitación y, posteriormente, se produce la alquilación con iodoacetamida

50 mM incubando las muestras 60 min en agitación y oscuridad. Se lavaron con H<sub>2</sub>O y acetonitrilo (calidad HPLC) alternativamente, eliminando los sobrenadantes en cada lavado. Se incorporó bicarbonato amónico 25 mM pH 8.8 y acetonitrilo en proporciones equivalentes. Se eliminó el sobrenadante y se añadió tripsina a 37°C. Después de 24 horas, se detiene la digestión con TFA hasta llegar a una concentración del 1%. Se lavó el sobrenadante con bicarbonato amónico 5mM pH=8.8 durante 1 hora en agitación. Posteriormente, se procede a la purificación de las muestras por columnas ZipTip (Millipore) de fase reversa (C18). Las muestras se secaron por centrifugación a vacío y se reconstituyeron en tampón de carga de espectrometría de masas (0.1% (v/v) de ácido fórmico en agua grado HPLC).

### Digestión de proteínas en gel de electroforesis concentrador

Aproximadamente 75µl de las muestras fueron resuspendidas en 25µl en el mismo tampón de carga que en gel de electroforesis convencional (ver apartado anterior). Las muestras se hirvieron durante 10 min y se aplicaron en pocillos de 2.8 cm de ancho en un gel convencional de SDS-PAGE (de 0.5 mm de espesor, al 4% de gel concentrador y al 10% de gel separador). La electroforesis se realizó a 30V y se detuvo tan pronto como el frente penetró 3 mm en el gel separador, de manera que todo el proteoma se concentró en la interfaz del gel. Las bandas de proteínas se visualizaron por tinción con Coomassie R-250. Se cortaron en cubos (2 x 2 cm) y se procedió a reducir las proteínas con DTT 200mM durante 60 min y alquilarlas con iodoacetamida 50mM en oscuridad durante 60 min. Se digirieron durante toda la noche a 37°C con 60 ng/µl de tripsina en bicarbonato amónico 50 mM, pH 8.8 con 10% (v/v) de ACN y el 0,01% (p/v) 5-ciclohexil-1-pentil-β-D-maltoside [148]. Los péptidos trípticos resultantes fueron extraídos después de 1 h de incubación en bicarbonato de amonio 12.5 mM, pH 8.8. Después de acidificar las muestras con TFA a una concentración final de 1% para detener la reacción enzimática, se desalaron con cartuchos Oasis C18 (Waters). Los péptidos resultantes de cada corte de gel se secaron por centrifugación a vacío y fueron analizados por separado mediante RP-HPLC-LIT para la identificación de proteínas o fraccionamiento por IEF para la cuantificación de proteínas por el método de marcaje con <sup>18</sup>O.



### Marcaje isotópico de péptidos con $^{18}\text{O}$ y fraccionamiento por isoelectroenfoque y por intercambio catiónico

Los péptidos resultantes de la digestión trípica de las proteínas en gel concentrador se marcaron diferencialmente con  $^{16}\text{O}/^{18}\text{O}$ , tal como se describe anteriormente [117], con ligeras modificaciones. Los péptidos provenientes de 3 ratones WT fueron marcados con  $^{18}\text{O}$  (95%; Isotec, Miamisburg, OH), y péptidos provenientes de 3 ratones knock-out CD9 -/- se marcaron con  $^{16}\text{O}$  en citrato de sodio 500 mM a pH 6.0, y 20% (v/v) de acetonitrilo, usando tripsina inmovilizada en una proporción de 1:200 (v: p) con respecto a la cantidad de proteína inicial. En el experimento con ratones CD81 knock-out, los péptidos provenientes de 3 ratones WT fueron marcados con  $^{16}\text{O}$  y los péptidos provenientes de 3 ratones knock-out CD81 +/- se marcaron con  $^{18}\text{O}$ . Después del marcaje, se detuvo la actividad de la tripsina eliminando las bolas de tripsina con un filtro físico (Wizard minicolumnas, Promega) y se inhibió la actividad remanente mediante la adición de TLCK (Tosyl-L-Lysine Chloromethyl Ketone), un inhibidor irreversible de tripsina, a una concentración final de 1mM, durante 1 h a 37°C. Las dos muestras marcadas se mezclaron y se diluyeron al 2% (v/v) de acetonitrilo. El pH se ajustó a 3 con formato amónico 1 M para evitar el desmarcaje por hidrólisis ácida. Las muestras se desalaron con cartuchos Oasis C18 (Waters), eluyendo con 50% (v/v) de acetonitrilo en formato amónico 5 mM, pH 3.0 y se secaron por centrifugación a vacío.

Para realizar el intercambio catiónico se utilizaron cartuchos OASIS MCX (Waters). La mezcla peptídica se resuspendió en formato amónico 5 mM a pH 3.0, y se cargaron en los cartuchos ya equilibrados. Los péptidos se separaron añadiendo secuencialmente y por duplicado formato amónico 500mM, formato amónico 1 M, formato amónico 1,5 M, todos al 25% (v/v) de acetonitrilo y formato amónico 2M. Todas las preparaciones fueron ajustadas a pH 3.0. Posteriormente se secaron y se desalaron con cartuchos OMIX C18 (Varian, Inc, Agilent, USA). Finalmente, los péptidos se secaron por centrifugación a vacío antes del análisis por RP-HPLC-LIT.

Para la realización del isoelectroenfoque en solución, la mezcla peptídica se resuspendió en tampón de enfoque (5% de glicerol y 2% de IPG buffer de pH 3.0-10 (GE Healthcare)) y se cargó

en tiras de isoelectroenfoque de 24 pocillos (24 cm Immobiline DryStrip, pH 3.0-10 (GE Healthcare)). Los péptidos se separaron por IEF en el sistema OFFGEL modelo 3100 (Agilent, Santa Clara, CA), utilizando el protocolo recomendado por el fabricante. Las fracciones recuperadas se acidificaron con formato amónico 1M, pH 3.0 y se desalaron con cartuchos OMIX C18 (Varian, Inc, Agilent, USA), eluyendo los péptidos con 50% (v/v) de acetonitrilo en formato de amonio 5 mM, pH 3.0. Finalmente, los péptidos se secaron por centrifugación a vacío antes del análisis por RP-HPLC-LIT.

### Espectrometría de masas mediante LC-ESI-LIT e identificación de proteínas

Cada muestra fue analizada por LC-MS/MS utilizando un sistema Surveyor LC acoplado a un espectrómetro de masas de iones lineal trampa (modelo de LTQ, Thermo Finnigan, San Jose, CA) [143, 152]. Los péptidos se concentran y se desalan en una pre-columna RP (0.32 x 30 mm, Biobasic C18, Thermo Electron) y se eluyen en línea con una columna RP analítica (0.18 x 150 mm, Biobasic C18, Thermo Electron) que opera a 2  $\mu$ l/min, utilizando el siguiente gradiente: 100% de A durante 10 min, 5-40% de B en 180 min, 40-95% de B en 5 min, y 95% de B durante 3 min (disolvente A: 0.1 % (v/v) de ácido fórmico; disolvente B: 0.1% de ácido fórmico (v/v) y 80 % (v/v) de acetonitrilo).

Para la identificación de los péptidos el LTQ fue programado en un modo de análisis dependiente de datos usando el método de barrido *Fullscan* entre 400 a 1600 m/z, seguido de fragmentación en CID de los 15 precursores más intensos detectados. Para la cuantificación de péptidos marcados con  $^{16}\text{O}/^{18}\text{O}$  el LTQ fue programado en un modo de análisis dependiente de datos con un barrido *Fullscan* entre 400 a 1600 m/z, seguido de *ZoomScan* y fragmentación en modo CID en conmutación de los 6 precursores más intensos. Los demás parámetros (exclusión dinámica, numero de microscans, etc.) fueron ajustados según trabajos previos [117, 142].

La identificación de proteínas se llevó a cabo utilizando el algoritmo SEQUEST (3.2 BioWorks, Thermo Finnigan), que calcula la función *Xcorr* para evaluar la correlación entre el espectro de fragmentación y la información de la secuencia de aminoácidos presente en la base de datos, y

el parámetro  $\Delta Cn$ , que es una puntuación normalizada que se obtiene de la diferencia del  $Xcorr$  de las primera y segunda secuencia candidatas, previamente descrito [153, 154]. A partir de los espectros MS/MS de un archivo RAW se realiza una búsqueda frente a la base de datos de proteínas humanas *Swissprot* (*Uniprot* versión 14.0, 19.929 entradas para la secuencia humana) complementada con tripsina porcina y estreptavidina bacteriana. Para los experimentos de identificación de proteínas a gran escala se utilizaron los siguientes parámetros de búsqueda: oxidación del residuo de metionina (+15,9949 Da) como modificación variable y carboxiamidometilación del residuo de cisteína (+57,034 Da) como modificación fija. Para los experimentos de cuantificación de proteínas a gran escala con  $^{18}\text{O}$  se incluyeron además la incorporación de  $^{18}\text{O}$  a los residuos de lisina y arginina (+4 Da). En ambos casos se permitieron el salto dos sitios de corte trípticos, con una tolerancia de 2 Da para el ion precursor y de 1.2 Da para los iones fragmentos. Los resultados obtenidos en SEQUEST se validaron utilizando el método robusto y no paramétrico del cociente de probabilidades desarrollado en nuestro laboratorio llamado *pRatio* [154], realizando una búsqueda frente a la base de datos normal y otra frente a la base de datos invertida. La tasa de falsos posibles candidatos (*FDR*) en las identificaciones obtenidas se calculó utilizando un método refinado del *pRatio* desarrollado en nuestro grupo [153], considerando el resultado de la identificación con una *FDR* igual o inferior al 5%.

### Análisis de cuantificación de proteínas

La cuantificación de proteínas por marcaje isotópico con  $^{18}\text{O}$  se realizó a partir de espectros *ZoomScan*, según un algoritmo descrito previamente [142, 143] y utilizando *QuiXoT*, un programa desarrollado en el laboratorio que permite calcular automáticamente los datos de cuantificación a partir de los espectros *ZoomScan* pertenecientes a las identificaciones de péptidos con una *FDR* menor o igual al 5%. Los espectros se ajustan a una curva teórica a partir del cálculo de la eficiencia de marcaje, permitiendo la determinación de la cantidad de las proteínas en las dos muestras, expresada como el  $\log_2 (A/B)$ , siendo *A* la cantidad de proteína proveniente de la muestra tratada con  $^{16}\text{O}$  y *B* la cantidad de la proteína de la muestra tratada

con  $^{18}\text{O}$ . Cuando no hay diferencias en la cantidad de proteína entre las dos muestras, el  $\log_2(A/B) = 0$ . La cuantificación de cada proteína se determina a partir de la cuantificación de varios péptidos y cada péptido se cuantifica a partir de varios espectros, mediante una media ponderada para cada uno de ellos, con un peso estadístico en cada uno de los tres niveles. Este modelo estadístico permite determinar la varianza local de cada uno de los espectros, péptidos y proteínas, que viene dada por la inversas de los respectivos pesos estadísticos. Estas varianzas locales sirven para estimar si un valor determinado se desvía significativamente de la media correspondiente, denominados valores atípicos. Los valores atípicos a nivel de proteína con una  $FDR_q$  menor del 5% se consideran cambios significativos en la abundancia de las proteínas, correspondientes a la proporción de valores promedios de las proteínas que se desvían de la distribución normal.

### Análisis estadístico de los datos de interacciones específicas

Para determinar las interacciones específicas entre las proteínas identificadas en los distintos *pull-downs* con lisado total de linfocitos T se contabilizó el número total de péptidos identificados con todos los anzuelos carboxi-terminal biotinilados y se analizaron las interacciones en cada donante. Se calculó la mediana de la cantidad de péptidos identificados en todos los anzuelos para cada proteína y se estimó la desviación estándar con la desviación absoluta media (MAD), a excepción del donante 6, donde la mediana no es confiable porque el cálculo con únicamente 3 anzuelos no es suficiente para determinar estos parámetros de manera eficiente, así que utilizamos el control negativo como media. El número de péptidos identificados en el mismo donante con distintos anzuelos siguen una distribución normal para cada donante. Se consideró que la interacción de una proteína con el ligando era significativamente específica al valor de  $p < 0.05$ . En el caso de las proteínas identificadas con un péptido y que no se encontraron en otros ligandos también se consideraron como interacción específica. Todas las interacciones específicas se validaron cuando las proteínas daban interacción significativa en, al menos, la mitad de los experimentos, teniendo en cuenta 3

experimentos para validar los ligandos ICAM-1 y EWI-2 y dos experimentos para el ligando CD81.

Los resultados de la identificación de proteínas obtenidas en cada *pull-down* con extractos de exosomas se analizaron estadísticamente mediante la prueba de *t* de *Student* en parejas de una cola, en cada proteína. Se consideró que la interacción era significativamente específica al valor de  $p < 0.05$ .

### **Análisis de biología de sistemas mediante herramientas bioinformáticas**

Las proteínas identificadas en los *pull-downs* y estadísticamente validadas como interacciones específicas con uno o varios anzuelos en los lisados de linfoblastos T y en los extractos de exosomas se agruparon y se analizaron por biología de sistemas mediante el análisis de Ingenuity Pathway (IPA; Ingenuity Systems, Mountain View, CA; <http://www.ingenuity.com>). IPA se utiliza para agrupar e interpretar estas proteínas en los términos de una red de interacción. La Base de Ingenuity Pathways Knowledge (IKB) se actualiza periódicamente, y las interacciones entre diferentes proteínas provienen de la literatura científica. IPA utiliza esta base de datos para la construcción de grupos de interacción de proteínas que implican interacciones directas e indirectas, interacciones físicas y de enlace, catalogando la naturaleza de las interacciones de forma sucesiva. Se han superpuesto colores en una red molecular global desarrollada a partir de la información en el IKB, agrupando las proteínas basándonos en la similitud ontológica e incluyendo las relaciones directas e indirectas en el análisis de redes. Se han seleccionado todas las fuentes de datos, las especies, los tejidos, las líneas celulares y los tipos de relaciones para construir las redes de interacción.



# RESULTADOS





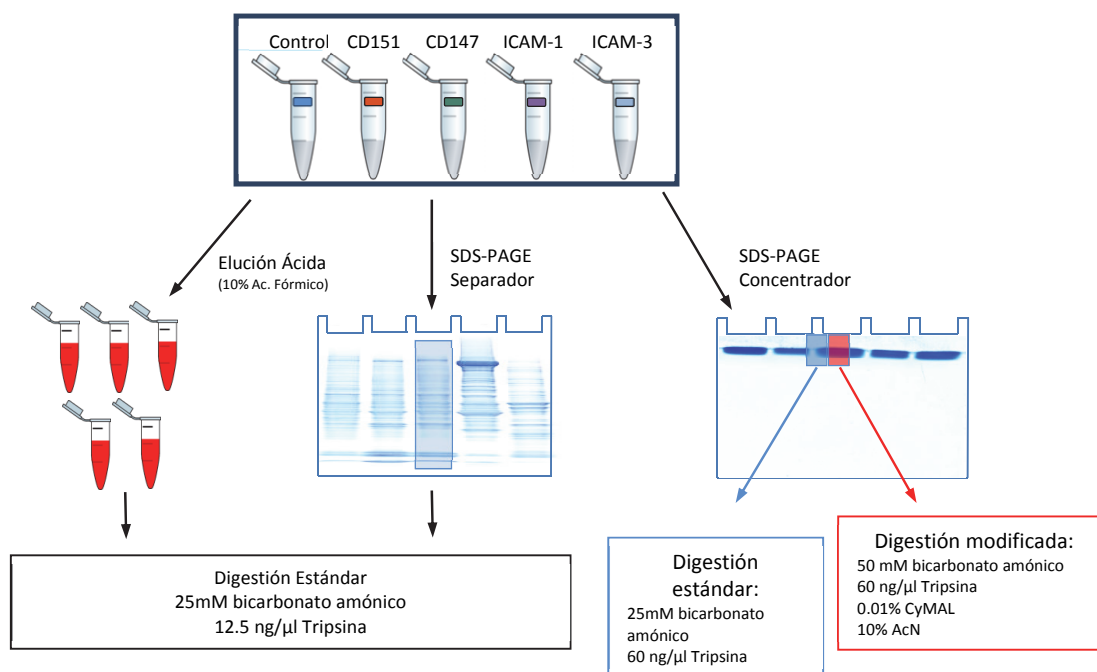
Los TEMs están implicados en numerosos procesos en los linfocitos T, como la señalización, la migración transendotelial, la adhesión celular o la activación del sistema inmune, y conectan receptores con cascadas de señalización específicas y con el citoesqueleto [1]. Conocer las interacciones que se llevan a cabo en la parte intracelular de las proteínas pertenecientes a los TEMs en linfocitos T es fundamental para profundizar en el entendimiento de todos estos procesos biológicos. Las proteínas que interactúan con estos dominios TEMs aún no se han caracterizado en profundidad. Para realizar este estudio hemos diseñado unos experimentos de *pull-down*, que permiten determinar una interacción física entre dos o más proteínas y que servirán para confirmar la existencia de una interacción descrita por otras técnicas experimentales, y también como un ensayo de cribado para la identificación de interacciones aún no conocidas [155]. Al llevar a cabo el estudio mediante procedimientos basados en la afinidad, un solo experimento de proteómica funcional podría conducir a una larga lista de posibles proteínas interaccionantes. A pesar de la gran cantidad de información que existe en las bases de datos de proteínas y de interacción proteína-proteína y de la disponibilidad de herramientas bioinformáticas de análisis, la inferencia de un significado biológico de una lista de proteínas interaccionantes sigue siendo un problema difícil de resolver. Al igual que en otras aproximaciones experimentales de proteómica, es necesario desarrollar un método reproducible de análisis mediante espectrometría de masas, que permita la identificación inequívoca y a gran escala de las proteínas interaccionantes con sus anzuelos, además de un sistema de validación estadística robusto de dichas interacciones específicas [121].

## 1. DESARROLLO DE UN MÉTODO PARA LA IDENTIFICACIÓN SISTEMÁTICA DE PROTEÍNAS A GRAN ESCALA

Para caracterizar el interactoma mediante técnicas de *pull-down* y análisis masivo de proteínas por espectrometría de masas es necesario controlar la variabilidad técnica que implica este tipo de aproximaciones metodológicas. La digestión trípica de proteínas, necesaria para el posterior análisis por EM, debe ser muy reproducible para poder comparar distintos ensayos de *pull-down*, ya sean en un mismo experimento o entre réplicas biológicas [121]. Además, el problema de la detección por EM de proteínas con un bajo número de copias no resulta de fácil solución y, en el estudio de interacciones proteína-proteína, son las interacciones con proteínas poco abundantes en muestras complejas las que suelen ser más relevantes a nivel biológico [3]. Con el objetivo de desarrollar un protocolo reproducible para la identificación masiva y sistemática de péptidos se compararon distintas estrategias de digestión trípica descritas en la bibliografía [138, 144, 146, 147] para minimizar los efectos que interfieren de forma crítica en la eficiencia y la reproducibilidad de la reacción enzimática, con el propósito de optimizar el rango dinámico de las proteínas en el análisis por EM.

La optimización del método de digestión se llevó a cabo a partir de una batería de experimentos de *pull-down* de lisados de linfoblastos T humanos. Como anzuelos se utilizaron péptidos biotinilados del extremo carboxi-terminal de las proteínas CD147, CD151, ICAM-1 e ICAM-3. Dichos péptidos se unieron a microsferas de *Streptavidin-Sepharose* por unión no covalente a la biotina y se incubaron con los lisados celulares. Se incluyó además un experimento control que se

realizó con microesferas de *Streptavidin-Sepharose* sin péptido biotinilado unido. Se procedió al lavado de las muestras para minimizar las interacciones no específicas y se digirieron con tripsina en solución y en geles de poliacrilamida. Para estos últimos se probaron distintas geometrías de corte del gel y diferentes tampones de reacción de la digestión. Las distintas estrategias abordadas se resumen en la Figura 5.



**Figura 5: Esquema de las distintas aproximaciones empleadas para la optimización del método de digestión usado para los pull-downs.** Los lisados celulares de linfocitos T provenientes de donantes humanos fueron incubados con ligandos peptídicos de ICAM-1, CD147, CD151 e ICAM-3 unidos a microesferas de *Streptavidin-Sepharose* y un control consistente en microesferas de *Streptavidin-Sepharose* sin péptido unido. Las proteínas unidas se sometieron a digestión triptica y los péptidos resultantes se analizaron por EM. La comparación de los distintos métodos se realizó contabilizando el número de péptidos identificados cuya FDR < 5% por el método del pRatio y el porcentaje de digestiones parciales.

En una primera aproximación, las muestras obtenidas de los diferentes *pull-downs* se digirieron con tripsina en solución. Para ello, previamente se realizó una elución ácida de las proteínas unidas a las microesferas. Para conseguir un rendimiento óptimo en la extracción de proteínas y eliminar las interacciones no covalentes entre proteínas y la matriz de *Seapharose* se usó una solución de ácido fórmico al 10% (v/v), en la que se incubaron las muestras durante 25 minutos. Posteriormente, se filtraron las muestras y se recuperaron las proteínas interaccionantes tal y como se describe en Materiales y Métodos, eliminando las microesferas de *Streptavidin-Sepharose* que podrían interferir en la eficiencia de la digestión o en el análisis por EM. Finalmente, se procedió a la digestión triptica de las proteínas, usando las condiciones establecidas en trabajos previos de nuestro grupo [117] y los péptidos tripticos resultantes se analizaron por LC-MS/MS.

En segundo lugar, procedimos a separar las proteínas de los *pull-downs* sometiendo las microesferas que contenían las proteínas unidas a electroforesis monodimensional en SDS-PAGE convencional. Cada carril se cortó en 20 trozos, que se digirieron de forma separada para favorecer

la entrada de la tripsina al gel. La digestión triptica de cada trozo de gel se realizó usando un protocolo estándar de digestión en gel, descrito en Materiales y Métodos. Finalmente, los péptidos obtenidos en cada fracción se mezclaron y se analizaron por LC-MS/MS como una única muestra, para poder comparar el resultado con las demás estrategias de digestión.

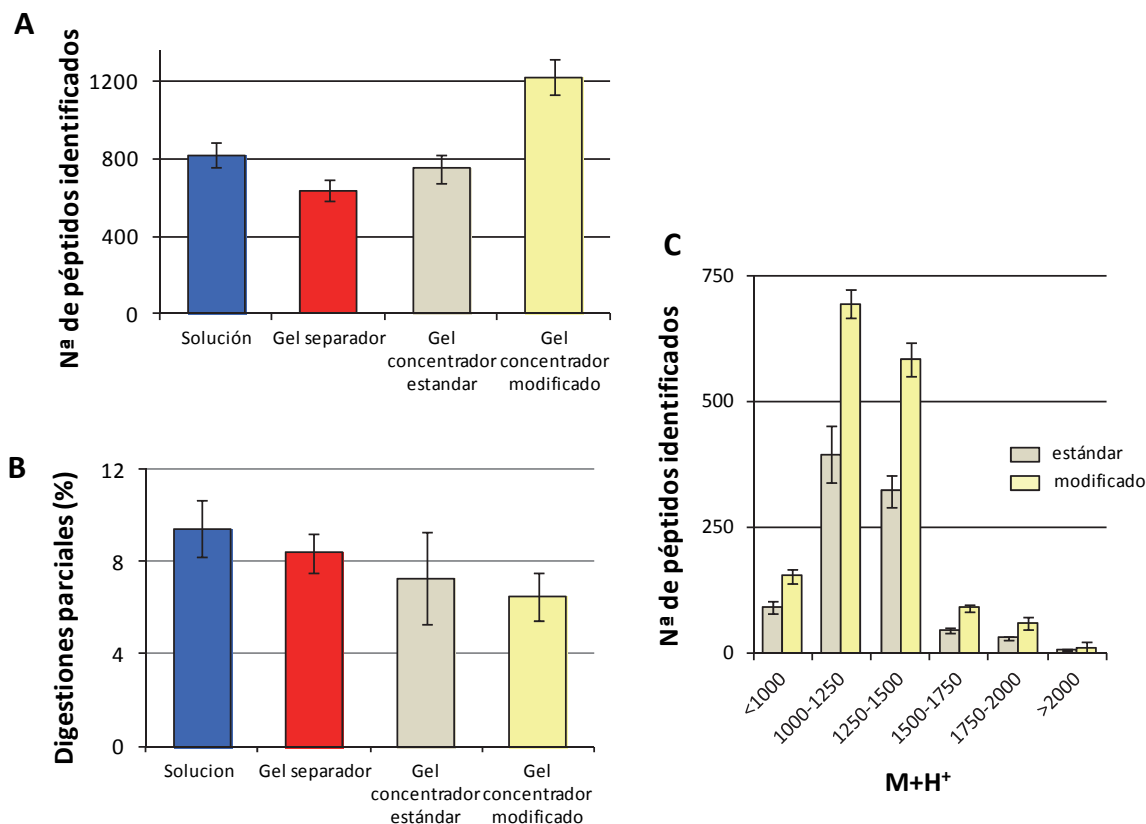
El último método de digestión que se probó consistió en aplicar las muestras de los *pull-downs* a un gel de SDS-PAGE convencional, que se dejó correr hasta la interfase entre la parte concentradora y la parte separadora del gel, concentrando todo el extracto de proteínas en una única banda. En este experimento se utilizó el doble de cantidad de muestra que la utilizada en los anteriores métodos, digiriendo una mitad de la banda con el tampón de reacción estándar (bicarbonato amónico 25 mM pH 8.8). La digestión en gel tiene la accesibilidad restringida de la tripsina a las proteínas que más íntimamente están ligadas a la matriz del gel y suele haber problemas en la digestión de proteínas de membrana, que poseen un fuerte componente hidrofóbico. Por ello se introdujeron cambios en el protocolo de digestión: se duplicó la capacidad de tamponamiento aumentando la concentración del tampón bicarbonato amónico a 50 mM pH 8.8 y se añadió 10% (v/v) de acetonitrilo y 0.1% (p/v) de detergente CyMAL, para favorecer la accesibilidad de la tripsina a las proteínas y posterior salida de los péptidos digeridos del entramado del gel [156]. Con estas nuevas condiciones se digirió la otra mitad de la banda. Además, con este método de gel concentrador todas las proteínas de cada *pull-down* se encuentran concentradas en una única banda, por lo que se utilizó más cantidad de tripsina para la reacción, en una proporción proteína:proteasa de 25:1 (p/p), que corresponde a una concentración de 60 ng/ $\mu$ l. Previamente se cuantificó la cantidad de proteína en cada *pull-down* mediante el método de Bradford [157], obteniéndose un total de aproximadamente 50  $\mu$ g en cada uno de ellos. Tras el proceso de digestión, los péptidos resultantes se analizaron por LC-MS/MS en las mismas condiciones de trabajo que los abordajes anteriores.

Para comparar cada una de las estrategias empleadas, se calculó el número de péptidos únicos identificados con una tasa de falsos positivos o FDR menor del 5% en todos los ensayos (Figura 6A y Figura Suplementaria 1). Los resultados indicaron que el rendimiento de la digestión fue similar en los tres primeros métodos, tanto en solución como en gel separador o en gel concentrador. Sin embargo, el resultado fue muy superior en la digestión realizada en gel concentrador con el protocolo modificado, obteniéndose un mayor número de identificaciones en 3 de los 5 *pull-downs* analizados en comparación con los resultados de los otros métodos, y con un total de más de 1200 péptidos.

Por otro lado, estudiamos la eficiencia de la digestión comparando el porcentaje de péptidos parcialmente digeridos en todos los ensayos (Figura 6B y Figura Suplementaria 2). La digestión parcial se expresó como el porcentaje de péptidos identificados conteniendo en su secuencia uno o más sitios de corte de tripsina que no estuvieran flanqueados por un residuo de prolina. Los resultados indicaron claramente que el porcentaje de digestión parcial en los ensayos realizados en solución era mucho mayor que los realizados en gel y que el método del gel concentrador era el que producía el menor porcentaje, reflejando una mayor eficiencia de digestión triptica.

Otra ventaja del método modificado en gel concentrador se observó al estudiar el tamaño de los péptidos identificados en los *pull-downs* (Figura 6C y Figura Suplementaria 3). Al representar el número de péptidos identificados en función del tamaño del precursor ( $M+H^+$ ) en cada método de

digestión, se apreció que en el método modificado había una proporción relativa de péptidos superior en todos los rangos estudiados, incluso con un tamaño mayor, debido probablemente a que la presencia del detergente CyMAL (5-ciclohexil-1-pentil-β-D-maltósido) en el tampón de reacción aumenta el rendimiento de la digestión de péptidos con mayor número de aminoácidos.



**Figura 6: Comparación de los distintos métodos de digestión.** La eficacia del método de digestión se evalúa comparando el número de péptidos únicos identificados en preparaciones de *pull-down* con CD147, CD151, ICAM-1 e ICAM-3, y un control realizado con microesferas sin ligando peptídico unido, digiriendo los extractos en solución, en gel SDS-PAGE separador, y en gel SDS-PAGE concentrador de la muestra. Esta última prueba se realizó con tampón estándar de digestión y también con tampón conteniendo detergente CyMAL, 10% de acetonitrilo y 50 mM de bicarbonato amónico a pH 8.8. (A) Número de péptidos identificados con todos los ligandos para cada método de digestión. (B) Porcentaje de digestión parcial obtenido para los distintos métodos. (C) Número de péptidos identificados en gel concentrante con tampón de digestión estándar (gris) y tampón de digestión modificado (amarillo), en función de la masa del precursor.

## 2. VALIDACIÓN DEL NUEVO MÉTODO DE DIGESTIÓN TRÍPTICA EN GEL CONCENTRADOR MEDIANTE MARCAJE CON <sup>18</sup>O

En cualquier tipo de análisis comparativo mediante herramientas de proteómica, ya sean de marcaje isotópico estable o mediante técnicas libres de marcaje (*label-free*), es fundamental que la etapa de digestión de proteínas sea homogénea y reproducible. Para validar la solidez y reproducibilidad de la digestión del nuevo método de digestión optimizada basada en gel

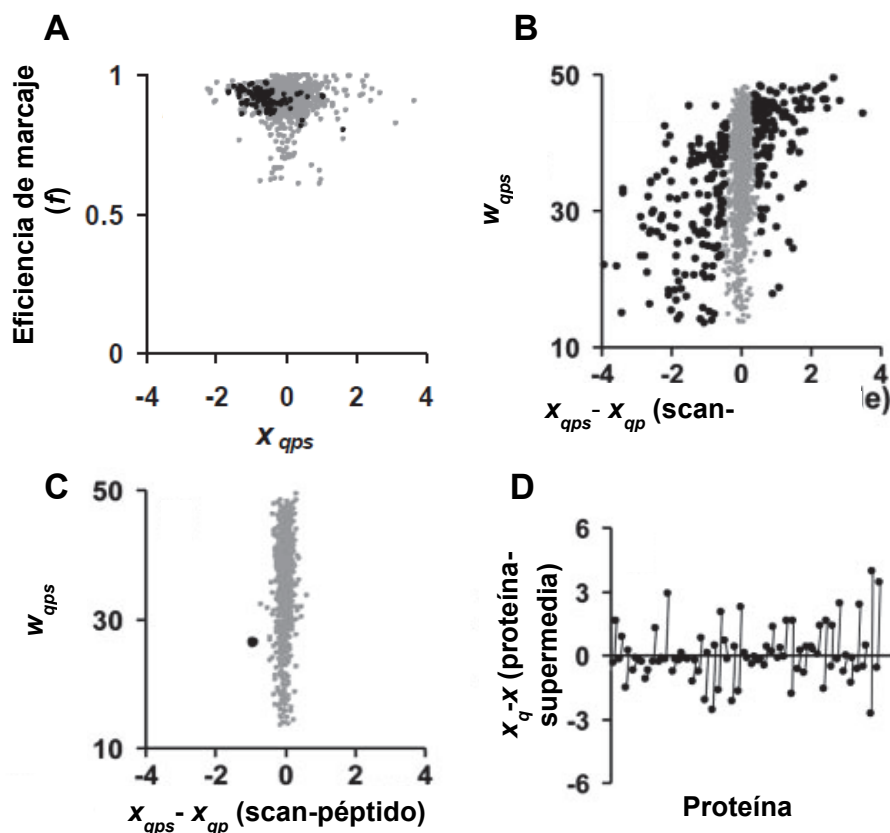
concentrador, se analizaron distintos proteomas mediante marcaje enzimático diferencial con  $^{18}\text{O}$ . Esta técnica permite estudiar con mayor profundidad los artefactos debidos a la digestión diferencial entre muestras, mediante el uso de un modelo estadístico de efectos aleatorios. Dicho modelo ha sido desarrollado para el análisis de resultados obtenidos del marcaje peptídico con  $^{18}\text{O}$  y espectrometría de masas en trampa iónica lineal [117]. El modelo estadístico asigna una varianza diferente a cada medición cuantitativa individual y descompone las fuentes de error a nivel de espectro, péptido y proteína, permitiendo que se puedan estudiar de forma separada. La cuantificación relativa de cada péptido se realiza a partir de la modelización y ajuste de un perfil teórico de isótopos a los espectros *ZoomScan* [142]. Las medias de péptidos y las varianzas de cada uno se calculan a partir de las diferentes medidas con las que se cuantifica cada péptido y, finalmente, las medias de proteína y las varianzas se calculan a partir de los diferentes péptidos que pertenecen a la misma proteína [117]. La varianza a nivel de espectro depende principalmente del análisis por EM y la varianza a nivel de proteína de la preparación de los extractos de proteínas. Sin embargo, la varianza asociada a cada péptido mide la dispersión de los valores de cuantificación obtenidos de diferentes péptidos que pertenecen a la misma proteína y depende principalmente de los métodos de digestión triptica y de marcaje usados. Por eso, el cálculo de la varianza a nivel de péptido cuando dos muestras iguales digeridas por separado se comparan entre sí (hipótesis nula) permite detectar errores sistemáticos asociados al proceso de generación de péptidos a partir de las proteínas y validar si los métodos de digestión analizados en este trabajo resultan suficientemente reproducibles a la hora de llevar a cabo una identificación masiva de proteínas [142].

## 2.1. Efecto de corte en la digestión en gel separador

Se utilizaron dos extractos diferentes de proteínas mitocondriales de corazón de rata preparados en las mismas condiciones. Las dos muestras se separaron por SDS-PAGE en dos carriles adyacentes y cada carril se cortó horizontalmente en 10 trozos. Las 20 piezas de gel resultantes se sometieron a digestión con tripsina individualmente y los péptidos resultantes se desalaron, y se sometieron al marcaje enzimático con  $^{18}\text{O}$  catalizado por tripsina, tal y como se describe en Materiales y Métodos. Los péptidos de cada pieza de un carril se marcaron con  $\text{H}_2^{16}\text{O}$  y los correspondientes a las piezas del segundo carril se marcaron con  $\text{H}_2^{18}\text{O}$ . Para asegurar que se eliminaba completamente la actividad enzimática residual de la tripsina y evitar el intercambio de oxígenos de los péptidos marcados en presencia de agua  $^{16}\text{O}$  [148], el protocolo dispone de dos pasos de inactivación de la tripsina. En primer lugar, las microesferas de tripsina inmovilizada usadas para el marcaje se separaron mediante filtración y, en segundo lugar, se inhibió la posible actividad residual de la enzima mediante un proceso de reducción seguido de alquilación de las muestras. Los dos pares de muestras de péptidos que correspondían a cada carril del gel se mezclaron y los péptidos resultantes marcados con  $^{16}\text{O}/^{18}\text{O}$  se analizaron por LC-MS/MS en una trampa iónica lineal (LTQ), analizando un total de 10 carreras. El LTQ fue programado para realizar un barrido en modo *ZoomScan* y un espectro *MS/MS* en los seis iones más intensos detectados en un barrido tipo *FullScan*, tal y como ya se ha descrito [117]. La primera exploración se utiliza para la cuantificación del par marcado  $^{16}\text{O}/^{18}\text{O}$ , y la segunda para la identificación de los péptidos. La eficiencia de marcaje de cada uno de los pares de péptidos se calculó de forma automática mediante un algoritmo descrito anteriormente [145].

Con esta aproximación metodológica, la mayoría de los péptidos se marcaron con una eficiencia ( $f$ ) de 0.9, y sólo una pequeña cantidad presentó una eficiencia de marcaje inferior a 0.8 (Figura 7A). Este resultado es representativo de varios experimentos diferentes realizados con el mismo protocolo. Dado que la eficiencia de marcaje se define como la fracción de átomos de  $^{18}\text{O}$ , en estas condiciones de trabajo, la fracción de péptidos no marcados pertenecientes a la muestra marcada es menor del 4%. Por tanto, el efecto de la eficiencia del marcaje en los ratios calculados ( $X_{spq} = \log_2 \frac{^{16}\text{O}}{^{18}\text{O}}$ ) no fue superior al 4% [145]. En la misma figura se puede observar que la representación de la nube de las cuantificaciones tenía una mayor dispersión en la escala de  $\log_2$ -*ratio* que las observadas en análisis previos realizados mediante digestión en solución [117]. Cuando se analizaron por separado las cuantificaciones de las diferentes fracciones de gel, observamos que la nube de puntos de cada fracción se desplazaba ligeramente en relación a los otros (Fig. 7A, puntos negros), lo que explica el aumento de la dispersión general de las cuantificaciones en relación con la media general. Este efecto podría reducirse, en parte, restando a las cuantificaciones en cada fracción su propia media general, pero este procedimiento produce varianzas más altas que las observadas cuando todas las proteínas se digieren juntas en solución.





**Figura 7: Análisis de los artefactos en la cuantificación por marcaje con  $^{18}\text{O}$  asociados a la técnica de fraccionamiento de proteínas por SDS-PAGE.** (A) Distribución de las eficiencias de marcaje en función de la relación por el log<sub>2</sub> en un experimento representativo (puntos en gris). Los espectros correspondientes a una fracción del gel están indicados con puntos negros. (B y C) El análisis de valores atípicos en el nivel de análisis. El peso estadístico asociado a cada cuantificación se representa frente a la diferencia entre la medida de los eventos y el promedio de su péptido correspondiente, a fin de destacar la presencia de valores atípicos. Los puntos negros indican valores atípicos en el nivel de espectro ( $\text{FDR}_{qps} < 5\%$ ). En (C) los péptidos y las proteínas que se cuantificaron en gel separador fueron considerados como una especie diferente. En (D) se analizan las proteínas cuantificadas en más de una fracción de gel. La gráfica muestra los promedios de proteína corregida por la media general, suponiendo que las proteínas que se cuantificaron en fracciones de gel diferentes son entidades diferentes. Las líneas unen los promedios de la misma proteína cuantificadas en diferentes fracciones. Para todas las proteínas, las medias con los pesos más altos de proteínas se representan en primer lugar. Hay que tener en cuenta que las líneas cruzan el eje horizontal cuando el log<sub>2</sub>-ratio tiene distinto signo en las fracciones de gel.

Una de las ventajas del método estadístico utilizado para analizar estos datos es que las varianzas a nivel de espectro se pueden utilizar para detectar la presencia de valores atípicos, es decir, los espectros que se desvían de la media del péptido más de lo esperado por su varianza estimada. Con este protocolo de digestión en gel separador se observó una proporción notable de valores atípicos (un 15% del número total de espectros en el ejemplo que se presenta en la Figura 7B, véanse los puntos negros). Un análisis detallado de estos valores extremos reveló que fueron producidos por los péptidos procedentes de proteínas que se cuantificaron en diferentes fracciones del gel, es decir, las proteínas que se encuentran en la frontera entre dos trozos de gel adyacentes. Por consiguiente, cuando los datos de las proteínas identificadas en las diferentes fracciones fueron

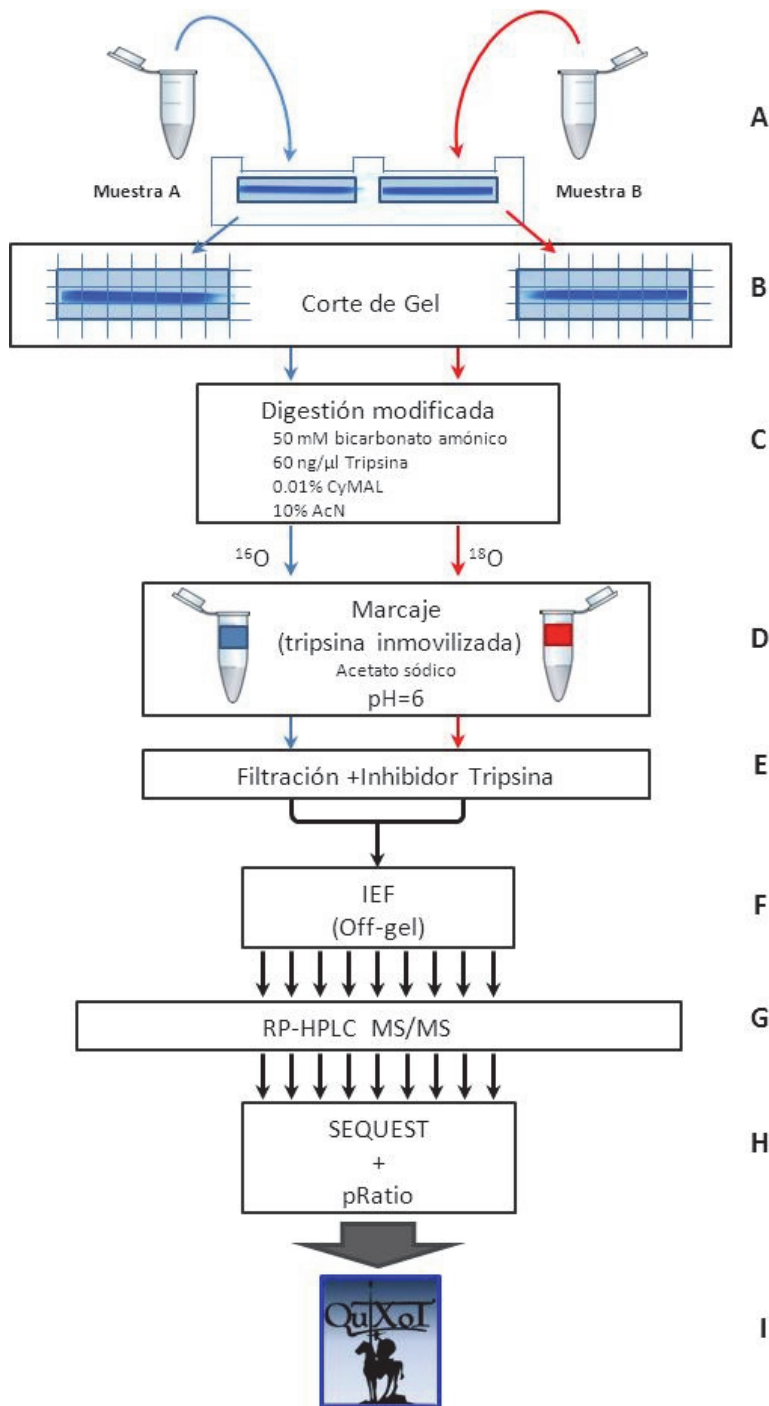
tratados estadísticamente como si fueran entidades diferentes, el número de valores atípicos en los espectros analizados se hizo despreciable, tal y como se observó en un trabajo anterior [117] (Figura 7C), indicando que los valores extremos detectados en la Figura 7B no fueron consecuencia de errores de medición, sino debidos al “efecto de corte” del gel. Otra evidencia de este efecto se observó al representar el  $\log_2$ -ratio de proteínas que se cuantificaron por separado en fracciones adyacentes. Como se observa en la Figura 7D, en la mayoría de los casos las abundancias relativas de las proteínas cuantificadas en una fracción eran claramente diferentes a las de la fracción adyacente en el gel. Este mismo efecto se observó en varios experimentos realizados siguiendo este protocolo de digestión, incluso teniendo un extremo cuidado en que las dos muestras se comportaran igual electroforéticamente y haciendo incluso los cortes de los carriles en zonas donde no se apreciaban visualmente bandas de proteínas, usando la tinción de Coomassie.

Por tanto, a la vista de los resultados obtenidos en la validación de este método de digestión en gel separador, llegamos a la conclusión de que esta aproximación experimental no era en suficientemente reproducible e introducía una gran variabilidad en la medida de cuantificación de péptidos al comparar dos muestras marcadas con  $^{18}\text{O}$  tras la digestión trípica.

### 2.2. Estrategia experimental para la cuantificación de proteomas con $^{18}\text{O}$ usando el nuevo método de digestión en gel concentrador

Una estrategia alternativa libre de artefactos debidos al “efecto de corte” del gel es usar la digestión trípica en gel concentrador. Este nuevo método permitió el diseño de una aproximación experimental fiable, reproducible y universal para el análisis cuantitativo de proteomas mediante  $^{18}\text{O}$  a gran escala [148] y fue desarrollado en colaboración con el Dr. Pablo Martínez-Acedo, quien optimizó la etapa de marcaje de péptidos con  $^{18}\text{O}$  y fraccionamiento en solución por punto isoelectrónico. El nuevo protocolo, esquematizado en la Figura 8, se basa en la digestión trípica de un extracto de proteínas concentrado en la interfase de la parte concentradora y separadora de un gel de SDS-PAGE. Las dos muestras a comparar A y B, se visualizan en el gel como una única banda mediante tinción con Coomassie; las dos bandas se recortan en porciones exactamente iguales y cada porción se corta, a su vez, en cuadraditos de 2x2 mm, digiriendo los cuadraditos de cada porción en un mismo tubo. El tampón de reacción contiene bicarbonato amónico 50 mM pH 8.8, 10% (v/v) de acetonitrilo y 0.1% (p/v) de CyMAL, para optimizar la eficiencia de la digestión. Todo el proceso se realiza en condiciones controladas de geometría del gel, usando la misma cantidad de proteína, el mismo volumen de reacción y la misma cantidad de enzima, según se describe en Materiales y Métodos, para que la digestión fuera exactamente igual en las dos muestras a comparar. La matriz del gel sirve como cámara de reacción para eliminar el SDS del tampón de electroforesis y otros contaminantes que puedan contener los extractos de proteínas, y para llevar a cabo la reducción y alquilación de las proteínas previas a la digestión.





**Figura 8: Esquema del protocolo de cuantificación propuesto para marcaje con  $^{18}\text{O}$ .** (A) Los extractos proteicos se concentran en un gel de poliacrilamida. (B) El trozo de gel se corta y (C) se procede a la digestión modificada en gel. (D) Después de controlar la eficiencia de digestión los péptidos en solución se marcan diferencialmente con tampón para mantener el pH a 6.0. (E) Se inhibe la tripsina en dos vías: filtración de la tripsina inmovilizada y con Inhibidor TLCK. (F) Una vez que se comprueba que el marcaje ha sido efectivo se mezclan las dos muestras, se desalan, y se separan los péptidos por IEF-OFFGel. Se desalan de nuevo los péptidos y, manteniendo el pH a 3.0 (G) se analiza cada una de las fracciones extraídas del gel por LTQ acoplado on-line a cromatografía en fase reversa. (H) Se realiza la identificación peptídica por SEQUEST y el control de la tasa de error en la identificación con el pRatio. (I) A continuación se cuantifican las proteínas por el programa bioinformático QuiXoT.

Para poner a punto el nuevo protocolo se utilizaron extractos de proteínas de distinta naturaleza: líneas celulares de hepatocitos humanos (HepG2) y de neuroblastoma (SK-N-MC), mitocondrias aisladas de tejido de corazón de rata, y fracción secretada y fracción citosólica de células T Jurkat en cultivo. Se analizó la eficiencia de la digestión calculando la proporción de péptidos identificados que contenían uno o más sitios de corte, obteniéndose menos de un 10% de digestión parcial en todos los proteomas analizados, incluyendo algunos ricos en proteínas de membrana, lo que demostró la buena reproducibilidad del método (Tabla 1). Este resultado coincide con los obtenidos en los ensayos previos al comparar los distintos métodos de digestión triptica, descritos en este capítulo, donde la digestión en solución resultó ser el método menos eficiente y se obtuvo el mejor resultado de digestión completa con el método modificado de digestión en gel concentrador. Además, coincide con el resultado obtenido en un trabajo anterior del análisis del

proteoma de células HUVEC en el que la digestión de las proteínas se llevó a cabo en solución y fue menos eficiente [117].

Los péptidos resultantes de las dos muestras A y B se marcaron diferencialmente con  $^{16}\text{O}/^{18}\text{O}$  y tripsina inmovilizada, tal y como se describe en Materiales y Métodos. Tras el marcaje, las microesferas de tripsina inmovilizada se eliminaron por filtración y se añadió TLCK para inhibir algún posible residuo de actividad enzimática. A continuación, se mezclaron las dos muestras, se desalaron a pH 3.0 y se procedió al fraccionamiento de los péptidos en solución por isoelectroenfoque (IEF) en 24 fracciones, usando el sistema de OffGel. Finalmente, los pares de péptidos  $^{16}\text{O}/^{18}\text{O}$  de cada fracción de OffGel se analizaron por espectrometría de masas en una trampa iónica lineal. El análisis de identificación de los péptidos dio como resultado una redundancia peptídica del fraccionamiento de péptidos, definida como el *ratio* de péptidos identificados en fracciones consecutivas en función del número de péptidos únicos identificados en total, menor de 2. Más del 65% de los péptidos identificados se encontraron en una única fracción y más del 75% en una o dos fracciones.

### 2.3. Efecto de la digestión parcial en la varianza a nivel de péptido

Para evaluar las fuentes de variabilidad del nuevo método se aplicó nuevamente el modelo estadístico para la cuantificación de  $^{18}\text{O}$  en trampa lineal previamente descrito [145]. La proporción de valores atípicos a nivel de espectro fue despreciable (en general, no más de 3 valores atípicos en un total de 1300 espectros cuantificados). Además, las varianzas calculadas a nivel de espectro fueron muy similares en todos los proteomas analizados y similares a los obtenidos en un estudio anterior del análisis del proteoma de células HUVEC [117] (datos no mostrados). En cuanto a la varianza a nivel de proteína, no fue significativamente diferente a cero en ninguno de los proteomas analizados (datos no mostrados), indicando que los distintos cultivos celulares y subproteomas fraccionados utilizados en el estudio se procesaron de forma que no introdujeron ninguna fuente de error apreciable que afectara a la comparación relativa de los niveles de proteína. El análisis de la varianza a nivel de péptido es el parámetro más importante para evaluar la idoneidad del método propuesto, ya que esta varianza depende fundamentalmente del procesamiento de la muestra en la etapa de digestión y el posterior marcaje de los péptidos con  $^{18}\text{O}$ . Las varianzas calculadas a este nivel fueron muy similares en todos los proteomas estudiados y esencialmente idénticas a las obtenidas con el método de digestión en solución y posterior marcaje [117] (Tabla 1).

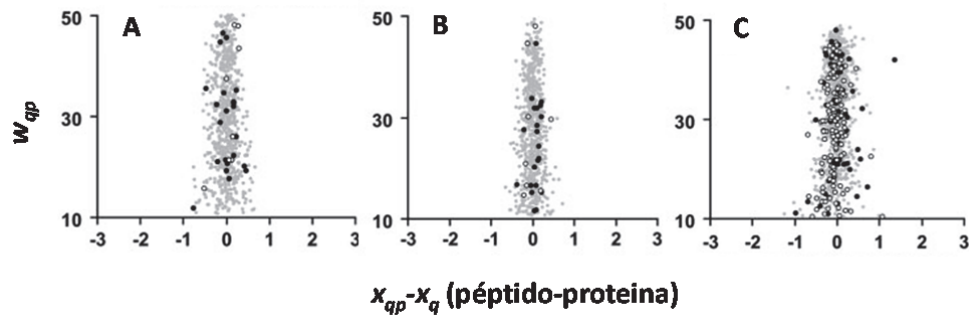
**Tabla 1.** Eficiencias de digestión y varianzas calculadas a nivel de péptido, obtenidas en los análisis cuantitativos de marcaje con <sup>18</sup>O de varios proteomas usando el protocolo propuesto.

Proteomas	% digestión parcial <sup>1</sup>	Varianza de péptido ( $\sigma_p^2$ ) (95% C.I.)
HepG2 (1)	5.3	0.024 (0.019 - 0.031)
HepG2 (2)	4.2	0.018 (0.015 - 0.021)
SK-N-MC (1)	2.8	0.018 (0.007 - 0.026)
SK-N-MC (2)	4.6	0.019 (0.015 - 0.021)
Secretoma	6.7	0.023 (0.014 - 0.029)
Mitocondria de Rata	8	0.031 (0.020 - 0.040)
HUVEC <sup>2</sup>	12.9-39.8	0.014-0.021

<sup>1</sup> Porcentaje de péptidos conteniendo uno o más sitios de corte de tripsina en el interior de la secuencia no flanqueados por residuos de prolina

<sup>2</sup> HUVEC, células endoteliales humanas vasculares de cordón umbilical.

Las cuantificaciones de péptidos digeridos que contenían sitios de corte no procesados por la tripsina (Figura 9, puntos negros) o la de los “subpéptidos” correspondientes (cuya secuencia pertenece a la de un péptido no digerido completamente en el mismo experimento) completamente digeridos (Figura 9, puntos blancos) no mostraron una desviación apreciable del resto de los péptidos, indicando que, en todos los casos, la digestión fue reproducible en las dos muestras que se compararon.



**Figura 9:** Efecto de la digestión parcial en la cuantificación relativa a nivel de péptido. Extractos de proteínas de las células de hepatoma (HepG2) (A), células de neuroblastoma (SK-N-MC) (B) y mitocondrias de corazón de rata (C) fueron utilizados como muestras de prueba. Se indican con puntos negros los espectros correspondientes a los péptidos que contienen sitios de corte internos, excepto aquellos con prolina después de un residuo básico; los “subpéptidos” totalmente digeridos derivados de un péptido identificado con digestión parcial se indican con círculos blancos.

Se obtuvieron resultados similares en todos los proteomas analizados, listados en la Tabla 1. En coherencia con estos resultados, los valores atípicos a nivel de péptido presentaban un  $\log_2$ -ratio desplazado de la media de la proteína de forma significativa. A juzgar por el análisis del parámetro  $FDR_p$  [148] (ver Materiales y Métodos), el número de valores atípicos a nivel de péptido fue

despreciable (en general, menos de 4 péptidos en un total de 1800 péptidos cuantificados) en todos los proteomas analizados.

Finalmente, se estudió la precisión de la hipótesis nula asociada al modelo estadístico. Al representar la distribución del  $\log_2$ -ratio estandarizado a nivel de proteína de acuerdo a sus varianzas estimadas (valores de  $Z_q$ ) y ajustando la curva a la función de Gauss, la desviación estándar de la curva no fue, en ningún caso, significativamente diferente de uno, según lo esperado (Figura 10). Además, no se detectó desviación de la normalidad en ninguno de los proteomas estudiados (Figura 10, recuadro). Por lo tanto, los errores producidos por el método propuesto fueron reproducibles, bien controlados y de acuerdo con la hipótesis nula del modelo estadístico utilizado.

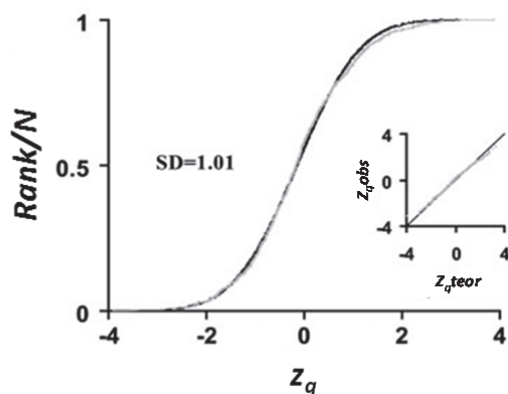


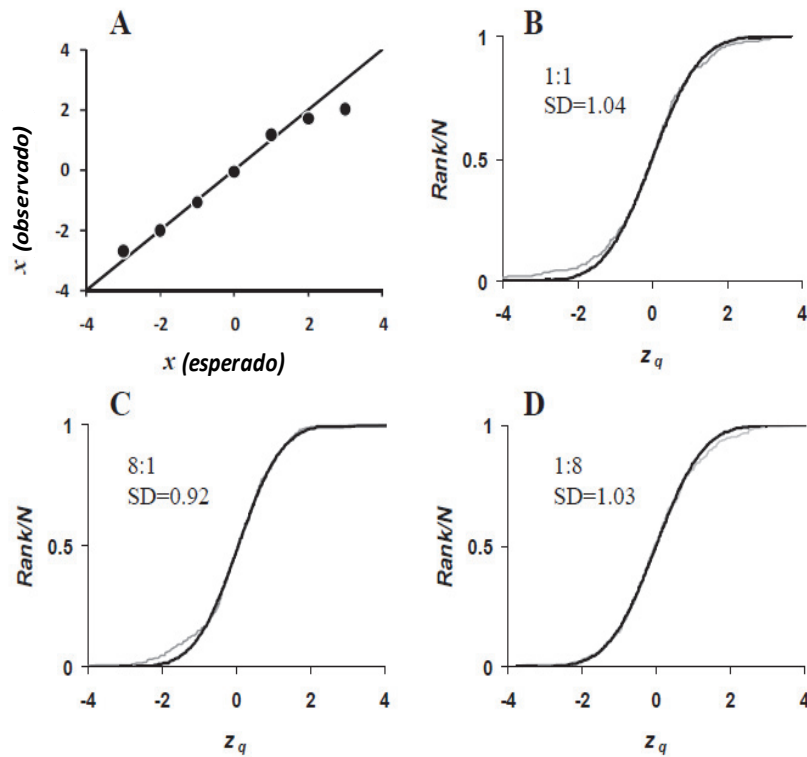
Figura 10: Análisis de la normalidad de las cuantificaciones a nivel de proteínas. El gráfico muestra la distribución de frecuencia acumulada de la variable tipificada  $Z_q$  (puntos grises), así como los gráficos de probabilidad normal correspondiente (recuadro). Los datos fueron obtenidos del análisis comparativo de los extractos de proteínas de las células de hepatoma HepG2. La curva se ajustó por mínimos cuadrados a una distribución normal (línea en negro); las cifras indican la desviación estándar (SD).

## 2.4. Aplicación del nuevo método de gel concentrador al estudio de la activación de las células T Jurkat

Para determinar el rango dinámico de cuantificación y, al mismo tiempo, demostrar la utilidad y fiabilidad del nuevo método de digestión y cuantificación para detectar los cambios de expresión de proteínas dentro de un contexto fisiológico, hemos utilizado como modelo extractos de proteínas del citosol de células T Jurkat control y estimuladas con PMA y Ionomicina. Esta parte del trabajo ha sido desarrollada en colaboración con la Dra. Elena Bonzón Kulichenko. Los extractos de proteínas fueron digeridos con el nuevo protocolo de digestión en gel concentrador y los péptidos resultantes de las células control y activadas se marcaron con  $^{16}\text{O}$  y  $^{18}\text{O}$ , respectivamente [148]. Las dos muestras se mezclaron entonces en siete proporciones diferentes (1:1, 2:1, 4:1, 8:1, 1:2, 1:4, 1:8;  $^{16}\text{O}$ : $^{18}\text{O}$ ), se fraccionaron por IEF y se analizaron por LC-MS/MS, de forma que los resultados cuantitativos de estos siete experimentos independientes se podrían comparar en términos de la varianza y la significación estadística de los cambios de expresión.

En cada experimento la "gran media" ( $X$ ) se calculó como la media ponderada de los cocientes de los  $\log_2$ -ratios de todas las proteínas cuantificadas [117]. Cuando se compararon las estimaciones de la "gran media" con los valores esperados, se obtuvieron unos resultados correctos en todos los experimentos, a excepción de una ligera subestimación del valor esperado en el experimento de la proporción 8:1 (Figura 11A). Este efecto era el esperado, ya que en esta proporción las envolturas isotópicas de los péptidos marcados con  $^{16}\text{O}$  interfieren significativamente con sus homólogos marcados con  $^{18}\text{O}$ , disminuyendo la precisión de los ratios estimados. Sin embargo, en todos los casos hubo un ajuste con las curvas esperadas para una distribución normal

con una varianza de uno, incluso en los dos experimentos de proporciones más extremas (Figura 11B-D), lo que demuestra la validez del modelo de hipótesis nula en prácticamente todo el rango dinámico analizado.



**Figura 11: Análisis del rango dinámico en extractos citosólicos de Jurkat control y activado.** Los lisados de las células T se concentraron en gel, se digirieron las proteínas, marcando con  $^{15}\text{O}$  y  $^{18}\text{O}$  los péptidos resultantes, respectivamente, se mezclaron en varias proporciones (1:1, 1:2, 1:4, 1:8, 2:1, 4:1, 8:1) y se sometieron a cuantificación relativa. (A) Comparación de los valores de la “gran media” observados y esperados forman una línea recta que representa la identidad entre los valores  $x$  (observado) e  $y$  (esperado). (B-D) Distribuciones de frecuencias acumuladas de la variable estandarizada  $Z_q$  (puntos grises) de proporciones relativas de muestra 1:1, 8:1 y 1:8 ( $^{15}\text{O}$ :  $^{18}\text{O}$ ). Las curvas fueron ajustadas por mínimos cuadrados a una distribución normal (líneas negras). En todos los casos se muestra la desviación estándar (SD) correspondiente al mejor ajuste.

Las varianzas a nivel de espectro, péptido y proteína se mantuvieron razonablemente estables y fueron similares a las obtenidas con los otros proteomas analizados en la proporción 1:1, excepto en la varianza a nivel de péptido en el caso de las proporciones 4:1 y 8:1, donde, por la misma razón de interferencia de las envolturas isotópicas, mostraron una leve tendencia a aumentar (Tabla 2).

**Tabla 2.** Varianzas a nivel de espectro, péptido y proteína obtenidas a partir de las células T Jurkat cuando las muestras de control y tratada se mezclan en diferentes proporciones.

Sample ratio	Scan variance ( $\sigma_s^2$ ) (95% C.I.)	Peptide variance ( $\sigma_p^2$ ) (95% C.I.)	Protein variance ( $\sigma_d^2$ ) (95% C.I.)
1:1	0.019 (0.016-0.024)	0.020 (0.019-0.044)	0.004 (0.000-0.010)
1:2	0.010 (0.008-0.015)	0.026 (0.022-0.031)	0.008 (0.000-0.010)
2:1	0.027 (0.024-0.029)	0.045 (0.037-0.053)	0.019 (0.008-0.032)
1:4	0.022 (0.005-0.024)	0.031 (0.030-0.039)	0.008 (0.000-0.010)
4:1	0.015 (0.011-0.017)	0.084 (0.071-0.098)	0.027 (0.013-0.039)
1:8	0.027 (0.023-0.031)	0.054 (0.042-0.071)	0.024 (0.006-0.044)
8:1	0.027 (0.020-0.034)	0.19 (0.16-0.24)	0.026 (0.000-0.060)

Para confirmar la validez del método en todo el rango dinámico de la cuantificación, se compararon los cambios en la abundancia de proteínas inducidos por la activación de células T Jurkat en los siete experimentos. Como se muestra en la Tabla 3, se observó una buena correlación en la magnitud y el signo de los cambios de expresión más importantes detectados en todos los experimentos. Además, aunque la significatividad estadística de los cambios de expresión (expresada en términos de la variable estandarizada  $Z_q$  en la Tabla 3) mostró una tendencia a disminuir en los resultados de los experimentos con los *ratios* más extremos, como se esperaba, los resultados fueron muy consistentes en todos los experimentos.



Tabla 3. Análisis comparativo de cambios de abundancia producidos por la activación de células T Jurkat. Las muestras tratadas y los controles son mezclados en diferentes proporciones.

Num Acceso	Nombre de Proteína <sup>a</sup>	Log <sub>2</sub> -ratio corregido <sup>b</sup>								Cambio <sup>d</sup>	z <sub>q</sub> <sup>c</sup>								FDR <sub>q</sub> <sup>d</sup>
		1:1	1:2	2:1	1:4	4:1	1:8	8:1	All		1:1	1:2	2:1	1:4	4:1	1:8	8:1	All	
		0	0.2	0.4	0.6	0.8	1	1.2	1.4		0	1	2	3	4	5	6	7	
P11021	GRP78 78 kDa glucose-regulated protein precursor	-0.54	-0.54	-0.41	-0.03	-0.40	-0.77	-0.58	-0.56	1.47	up	-5.23	-5.37	-2.51	-5.99	-1.74	-3.77	-1.90	-10.8
P08670	VIME Vimentin	-0.55	-0.49	-0.47	-0.58	-0.38	-0.65	-0.20	-0.52	1.43	up	-5.99	-4.78	-2.84	-5.28	-1.44	-3.24	-0.74	-10.2
P17661	DESM Desmin	-0.60	-0.64	-0.47	-0.66	-0.42	-0.67	-0.01	-0.58	1.50	up	-3.36	-3.30	-1.57	-3.17	-0.97	-2.26	-0.02	-6.23
P43490	NAMPT Nicotinamide phosphoribosyltransferase	-0.75	-0.75	-0.65			0.09		-0.60	1.52	up	-3.77	-2.75	-3.01		0.29			-5.04
Q16891	IMMT Mitochondrial inner membrane protein	-0.49	-0.35	-0.72	-0.26				-0.39	1.31	up	-2.74	-2.95	-2.55	-1.46				-4.74
P30520	PURA2 Adenylosuccinate synthetase isoform 2	-0.78	-0.49						-0.69	1.61	up	-4.16	-1.86						-4.47
Q9UJZ1	STML2 Stomatin-like protein 2	-0.53	-0.28	-0.34	-0.35				-0.36	1.28	up	-3.01	-2.28	-1.49	-2.08				-4.42
P21796	VDAC1 Voltage-dependent anion-selective channel protein 1	-0.53							-0.49	1.40	up	-4.18		-1.61					-4.41
Q9UQ80	PA2G4 Proliferation-associated protein 2G4	-0.84	0.01	-0.71	-0.16	-0.91	-0.85	-0.93	-0.33	1.26	up	-4.60	0.09	-2.49	-0.73	-2.12	-2.57	-1.78	-4.15
P34932	HSP74 Heat shock 70 kDa protein 4	-0.60	-0.14	-0.09	-0.23	-0.34	-0.29	-0.21	-0.27	1.21	up	-4.37	-1.30	-0.44	-1.06	-0.83	-1.41	-0.73	-4.12
Q00325	MPCP Phosphate carrier protein	-0.50	-0.23	-0.29	-0.45	-0.35		0.03	-0.35	1.27	up	-2.84	-1.40	-1.04	-2.06	-0.78		0.06	-3.72
P41219	PER1 Peripherin	-0.60	-0.59						-0.60	1.51	up	-2.99	-2.10						-3.66
Q9Y4L1	HYOU1 Hyposia up-regulated protein 1 precursor	-0.15	-0.22	-0.49	-0.15	0.07	-0.49	-0.15	-0.22	1.16	up	-1.31	-2.17	-2.13	-1.10	0.18	-2.04	-0.29	-3.57
P45880	VDAC2 Voltage-dependent anion-selective channel protein 2	-0.50	-0.27						-0.37	1.30	up	-2.75	-1.70						-3.09
P42025	ACTY Beta-contractin	-0.46	-0.76						-0.59	1.50	up	-1.78	-2.62						-3.07
P62899	RL31 60S ribosomal protein L31	-0.28			-0.67				-0.48	1.40	up	-1.20		-3.02					-3.02
P25705	ATPA ATP synthase subunit alpha	-0.23	-0.14	-0.02	-0.16	-0.23	-0.51	0.03	-0.17	1.13	up	-1.79	-1.21	-0.09	-1.32	-0.83	-2.02	0.08	-2.83
P62988	UBIQ Ubiquitin	0.284	0.173	0.18	0.27	-0.02	0.256	-0.09	0.215	1.16	down	1.643	1.132	0.678	1.66	-0.04	0.799	-0.17	2.61
P60174	TPIS Triosephosphate isomerase	0.347	-0.04	0.228	-0.08	0.51	0.038	0.552	0.166	1.12	down	3.46	-0.32	1.231	-0.5	1.664	0.119	1.464	2.7
Q60234	GMFG Glna maturation factor gamma	0.485	-0.12				0.349	-1.12	0.349	1.27	down	2.805	-0.42			2.979	-2.11	2.73	
Q9NV11	FANCI Fanconi anemia group 1 protein	0.21		0.71					0.36	1.29	down	1.38		2.99					2.80
Q9P215	SYLC Leucyl-tRNA synthetase	0.20	-0.10	0.44	0.51		0.31		0.26	1.20	down	1.62	-0.32	2.57	1.24		0.56	2.84	
Q16643	DREB Drebrin	0.47	0.18	0.27	0.21				0.27	1.21	down	2.58	1.27	0.94	0.99				2.9
Q8TCS8	PNPT1 Polyribonucleotide nucleotidyltransferase 1	0.05		0.91		0.57			0.41	1.33	down	0.25		3.53		1.81		2.92	
P17252	KPCA Protein kinase C alpha type	0.53	0.40						0.48	1.40	down	2.60	1.39					2.93	
Q00767	ACOD Acyl-CoA desaturase	0.49	0.44						0.47	1.39	down	2.45	1.61					2.93	
P54819	KAD2 Adenylate kinase isoform 2	0.13	-0.16	0.34		0.81			0.35	1.28	down	0.67	-0.55	3.31		2.56		2.94	
P62917	RL8 60S ribosomal protein L8	0.37	0.24			0.68			0.31	1.24	down	1.94	1.87			1.68		2.98	
Q9NR28	DBLH1 Diablo homolog	0.46	0.12						0.54	1.46	down	3.62	0.44					3.11	
Q6UXN9	WDR82 WD repeat-containing protein 82	0.42	0.34	0.31			0.48		0.38	1.30	down	2.25	1.74	1.08			0.95	3.15	
Q14165	K0152 Uncharacterized protein KIAA0152 precursor	0.63	-0.32			1.11			0.52	1.43	down	2.77	-1.07			4.18		3.27	
P20700	LMNB1 Lamin-B1	0.34	0.54						0.43	1.34	down	1.96	2.75					3.29	
P16949	STMN1 Stathmin	0.47	0.23	0.29	0.13	0.38	0.12	0.20	0.27	1.20	down	2.86	1.55	1.35	0.77	0.95	0.37	0.40	
P06400	RB Retinoblastoma-associated protein	0.37	0.47				0.39		0.39	1.31	down	2.85	1.64				1.05	3.44	
P58546	MTPN Myotrophin	0.47				0.79	0.96		0.57	1.48	down	2.44				2.23	1.61	3.54	
P60981	DEST Dextrin	0.36			0.78				0.54	1.45	down	1.84		3.36				3.58	
Q7L9L4	MOL1A Mps one binder kinase activator-like 1A	0.29	0.45	0.66	0.33	0.53	0.42	0.58	0.41	1.33	down	1.70	2.36	2.43	1.58	1.32	1.43	1.17	
P31350	RIR2 Ribonucleoside-diphosphate reductase subunit M2	0.63	0.36	0.26	0.52	-0.12	0.42	-0.13	0.40	1.32	down	3.39	2.59	1.18	2.45	-0.29	1.27	-0.26	
P62805	H4 Histone H4	0.40	0.26	0.59	0.18	0.91	0.07	0.78	0.35	1.27	down	3.04	1.83	2.98	1.09	2.57	0.28	1.99	
P01733	TVB1 T-cell receptor beta chain V region YT35 precursor	0.99	0.52			0.99	0.81		0.81	1.75	down	4.01	1.78			2.15	1.82	1.09	
P61604	CH10 10 kDa heat shock protein	0.41	0.32	0.63	0.06	0.58	0.78		0.37	1.29	down	3.16	2.51	3.22	0.37	1.33	2.55	3.21	
P37268	FDFT Squaleen synthetase	1.01	0.81						1.01	2.03	down	7.01	2.91					8.11	

<sup>a</sup> Las proteínas que muestran los cambios de abundancia estadísticamente significativas están numeradas como en la figura 20.

<sup>b</sup> El log<sub>2</sub>-ratio se corrigió restando el valor de la gran media.

<sup>c</sup> Los valores normales estandarizados (Z<sub>q</sub>) se obtienen dividiendo el log<sub>2</sub>-ratio de la proteína por su varianza. Se consideran válidas las proteínas cuantificadas con más de 1 scan. Los datos se colorean de acuerdo con las escalas de color en la parte superior.

<sup>d</sup> La tasa de posibles falsos errores de cuantificación a nivel de proteínas (ver referencia 148 del texto)

Al representar las dispersiones de los log<sub>2</sub>-ratios en los tres niveles (espectro, péptido y proteína), se observa que las medidas con un peso estadístico alto presentaban menos dispersión (Figura 12). Incluso las medidas menos precisas rara vez se desviaban más del doble del valor esperado (Figura 12A). Las dispersiones de las medidas de los espectros individuales respecto a su media del péptido (Figura 12B) y de péptidos respecto a su media de proteína (Figura 12C) son muy estrechas y, debido al efecto de promediar, se produce una distribución aún más estrecha de las proteínas en torno a la “gran media” (Figura 12D). En estas condiciones, los cambios de expresión tan pequeños como en el caso de la proteína *vimentina* (con un aumento de 1.4 veces en células activadas, Tabla 3) pueden verse claramente como un aumento estadísticamente significativo (Figura 12, puntos negros).

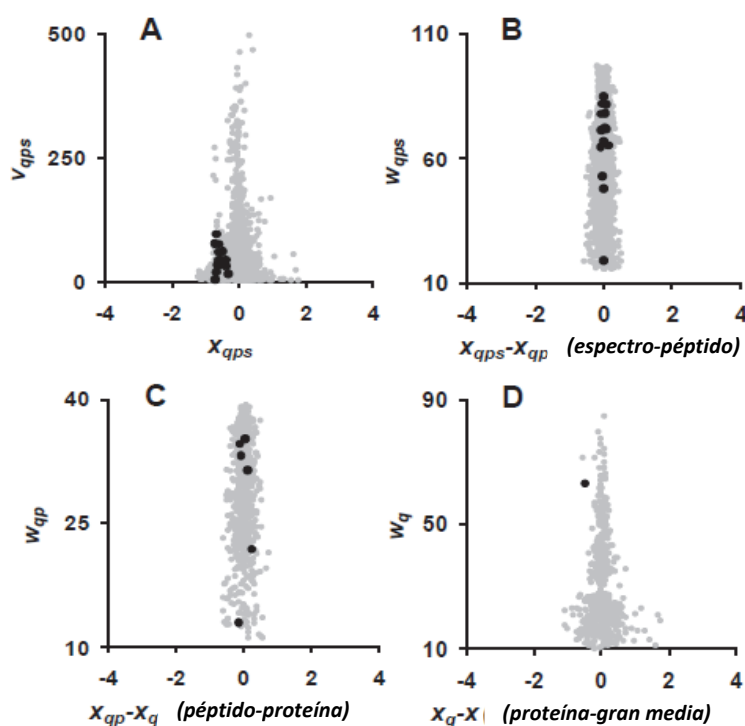


Figura 12: La precisión de las cuantificaciones a niveles de espectro, péptido y proteína. Los cuatro gráficos muestran la dispersión total de los  $\log_2$ -ratios (A), la dispersión de espectros de su media de péptido (B), la dispersión de péptidos a partir de su media de proteína (C) y la dispersión de las proteínas respecto a la “gran media” (D). El eje Y es el peso estadístico de este matemático (A) o los pesos estadísticos en cada uno de los tres niveles (B, C, D), cuya inversa se puede utilizar como estimación de las varianzas locales. Estos resultados se toman a partir del análisis cuantitativo de los extractos citosólicos de células T Jurkat control y activadas en proporción 1:1 ( $^{15}\text{O}$ : $^{18}\text{O}$ ). Los puntos negros corresponden a cuantificaciones asociadas a la proteína *vimentin*.

Para determinar qué proteínas mostraban cambios de abundancia significativos tras la estimulación de las células T Jurkat se analizaron los resultados obtenidos de todos los experimentos en conjunto. En términos estadísticos, los cambios en la abundancia de proteínas son las medidas de las cuantificaciones que se desvían del valor esperado de la “gran media”. Usando un corte de  $FDR_q$  del 10%, 17 proteínas aumentaron y 25 proteínas disminuyeron debido a la activación de las células T con PMA y Ionomicina (Tabla 3). Entre ellas, muchas proteínas han sido descritas previamente implicadas en la activación de células T, como *nicotinamide phosphoribosyltransferase* [158], *stomatin-like protein 2* [159], *proliferation-associated protein 2G4* [160] y *adenylosuccinate synthetase isozyme 2* [161]. Otras proteínas que también aumentaron significativamente en este análisis participan en la respuesta celular en condiciones de estrés, como la inducción de apoptosis con la activación de células T con PMA [162], tales como las proteínas sensores de estrés *78KDa glucose-regulated protein precursory hypoxia up-regulated protein 1 precursor* [163], *ATP synthase subunit alpha* [164], las proteínas de apoptosis mediada en mitocondria *voltage-dependent anion-selective channel protein 1* (VDAC1) [165], VDAC2 [166], y *mitochondrial inner membrane protein* [167]. En relación a las proteínas que disminuyeron en el análisis, se encontraron proteínas del citoesqueleto (*dextrin*, *lamin-B1*, *stathmin* y *myotrophin*), proteínas implicadas en el metabolismo energético (*triosephosphate isomerase*) y de lípidos (*acyl-CoA desaturase* y *squalene syntethase*) y en la síntesis y degradación de proteínas (*ubiquitin*, *leucyl-tRNA synthetase* y *60S ribosomal protein L8*). La *histone H4* y la *protein kinase C alpha type* también se observan disminuidas en la fracción citosólica de las células T activadas, debido posiblemente a su relocalización en la membrana plasmática en una situación de apoptosis [168].



Todos estos resultados indicaron que, aunque se pierda un poco de precisión en la detección de los cambios de proteínas en ratios altos de rango dinámico, el método propuesto en este trabajo permite un control total de las fuentes de error y una estimación precisa de las significancias estadísticas para la detección de los cambios de expresión de proteínas que tienen una relevancia biológica en un rango dinámico cuantitativo que abarca cerca de dos órdenes de magnitud.

### 3. CARACTERIZACIÓN DEL INTERACTOMA DE LOS DOMINIOS INTRACELULARES DE PROTEÍNAS TEM EN LINFOCITOS T HUMANOS.

#### 3.1. Diseño experimental

Con el fin de profundizar en el papel biológico de las interacciones intracelulares de las proteínas asociadas a los dominios TEMs, realizamos una batería de ensayos de *pull-downs* usando péptidos biotinilados como anzuelos unidos a microesferas de *Streptavidin-Sepharose*. Las secuencias de los péptidos comprendían la región C-terminal de las tetraspaninas CD9, CD81 y CD151, y de las proteínas asociadas a tetraspaninas ICAM-1, VCAM-1 y EWI-2 [3], miembros de la superfamilia de las inmunoglobulinas [38, 41]. Como ensayos control, se emplearon péptidos que contienen las secuencias intracelulares de otros receptores diferentes, tales como los miembros de la Ig CD147 e ICAM-3, la lectina C-CD69, y los receptores de quimioquinas CCR7 y CXCR4, así como microesferas sin péptido anzuelo. Las microesferas se incubaron con los lisados totales procedentes de cultivos de linfoblastos T humanos de siete donantes. Las proteínas unidas a las microesferas, tras una serie de lavados para eliminar interacciones inespecíficas, se procesaron siguiendo el método modificado puesto a punto de digestión triptica de proteínas en gel concentrador, validado mediante marcaje con  $^{18}\text{O}$  (ver Apartados 1 y 2 de este Capítulo). Los péptidos resultantes de la digestión se analizaron mediante LC-MS/MS en una trampa iónica lineal, como se describe en Materiales y Métodos. Un esquema detallado de la estrategia experimental seguida en todos los ensayos de *pull-down* para el estudio de la interacción de proteínas realizados se describe en la Figura 13.

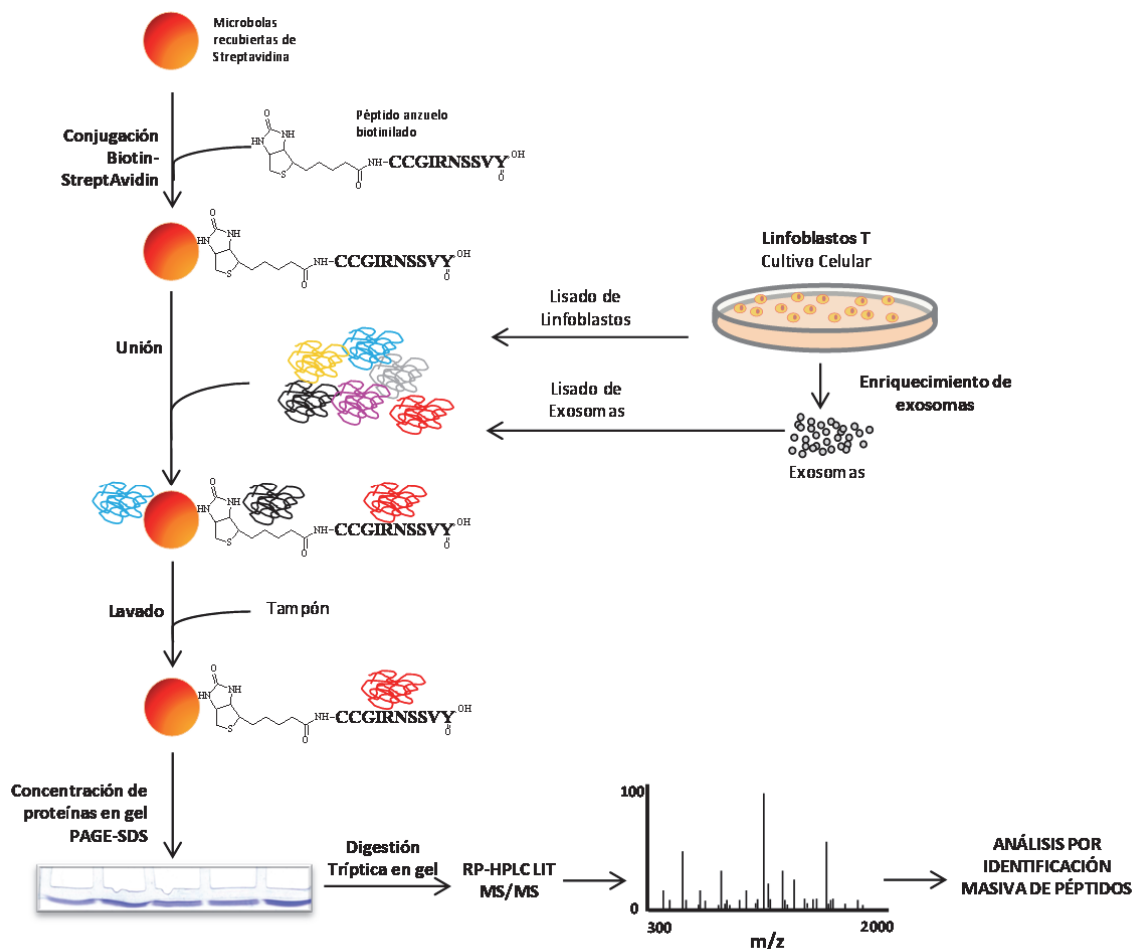


Figura 13: Esquema del protocolo seguido para el análisis de alto rendimiento del interactoma de microdominios enriquecidos en tetraspaninas. Péptidos sintéticos biotinilados, que contienen las regiones intracelulares (C-terminal en todos los casos a excepción de CD69) de tetraspaninas CD9, CD81 (representado en la figura) y CD151, sus receptores asociados EWI-2, ICAM-1, VCAM-1 y el control de diferentes receptores (CD147, CD69, ICAM-3, CCR7 y CXCR4) se unieron a microesferas de *Streptavidin-Sepharose* y se incubaron con extractos de proteínas de linfoblastos humanos. Los exosomas producidos por estas células se incubaron con los anzuelos peptídicos de CD81 y el EWI-2. Las microesferas se lavaron para eliminar los contaminantes y se aplicaron directamente a un gel concentrador SDS-PAGE. Las bandas de proteínas se concentraron y se digirieron con tripsina. Los péptidos resultantes se identificaron mediante cromatografía líquida en combinación con espectrometría de masas de trampa iónica lineal.

## 3.2. Estudio de la especificidad de las proteínas interaccionantes con los distintos anzuelos

La especificidad de las interacciones con cada ligando fue comprobada comparando el número de péptidos identificados para cada proteína en un donante determinado, sirviendo los demás anzuelos como controles negativos. En total, se realizaron 53 *pull-downs* con lisados de linfoblastos T humanos de siete donantes, probando hasta once anzuelos diferentes. En conjunto, 4918 péptidos únicos, correspondiente a 1957 proteínas diferentes, fueron identificados con una FDR menor del 5% en esta serie de experimentos (Tabla 4). La diferencia en el número de péptidos identificados en cada uno de los donantes se debe a que el proceso de digestión se fue optimizando

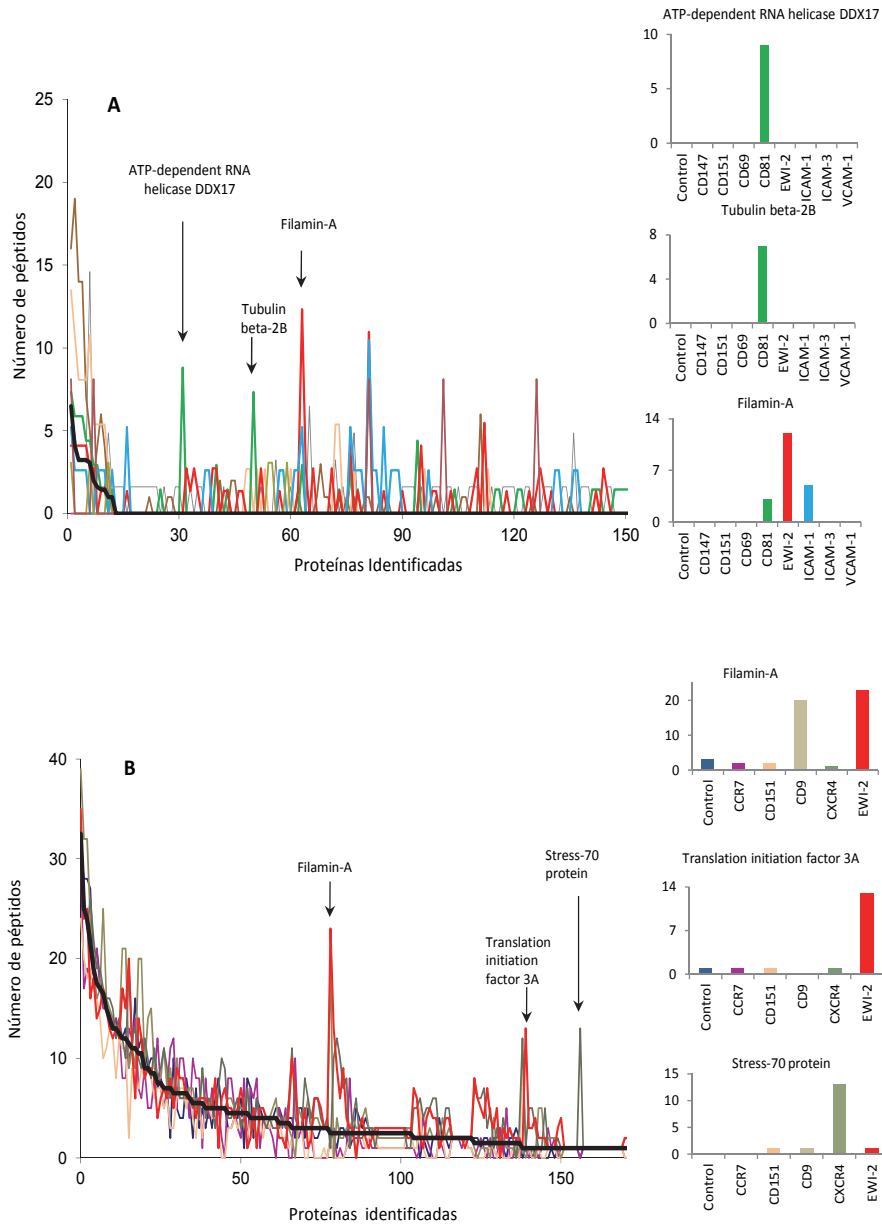
durante el desarrollo de estos experimentos, con la excepción del donante 7, en el que el número de péptidos identificados fue menor de lo esperado.

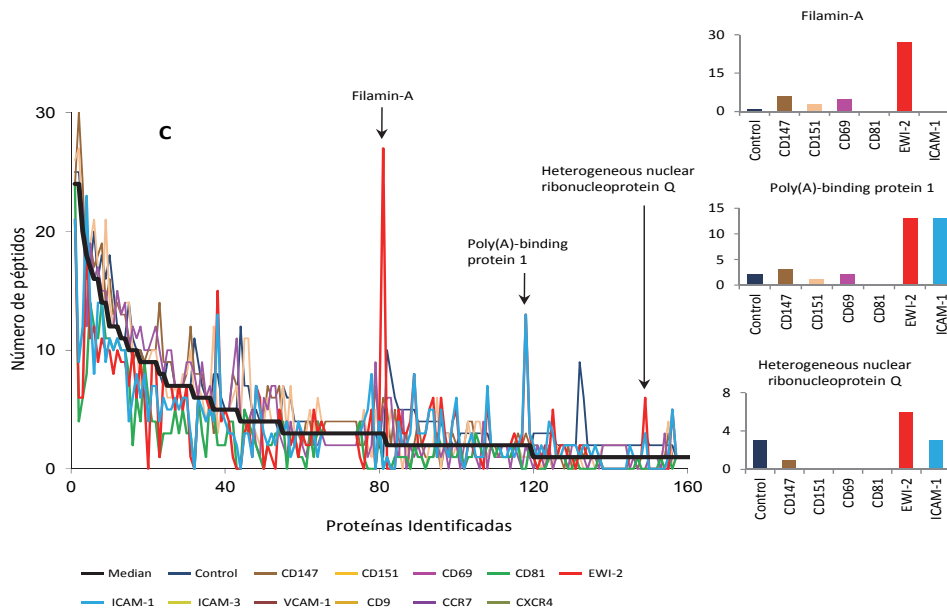
**Tabla 4.** Número de péptidos identificados en todos los ensayos de *pull-down* realizados con lisados de linfocitos T humanos, con una FDR<5%.

Donante	Número de péptidos identificados											
	CONTROL	CD147	CD69	CD151	CD81	EWI-2	ICAM-1	VCAM-1	ICAM-3	CD9	CCR7	CXCR4
1	173	248	297	374			264	120	129			
2	251	251	137	361		375	358	144	140			
3	293	879	97	249	523	441	179	82	129			
4	440	714	14	265	482	521	269	88	232			
5	1060			748		1680				1235	913	1544
6	679				2096	2162	1602					
7	243	520	348	273	570			337	269	352	144	456
<b>Nº ensayos</b>	<b>7</b>	<b>5</b>	<b>5</b>	<b>6</b>	<b>4</b>	<b>5</b>	<b>5</b>	<b>5</b>	<b>5</b>	<b>2</b>	<b>2</b>	<b>2</b>

La distribución del número de péptidos por cada proteína ordenados en un mismo experimento por su promedio fue muy similar en los distintos anzuelos utilizados para el mismo donante. La Figura 14 representa esta distribución en 3 de los 7 donantes analizados. Observamos que existe “un fondo” de proteínas identificadas con un número relativamente constante de péptidos en casi todos los anzuelos, que corresponden a las interacciones inespecíficas de proteínas. Pero las interacciones específicas fueron fácilmente detectables, destacándose claramente de la distribución de fondo no específica.

En algunas proteínas el número de péptidos destaca sólo con determinados anzuelos en el mismo donante. Tal es el caso de la *ATP dependent RNA helicase DDX17* y la *tubulin beta-2B* en el experimento 3 (Figura 14A), donde se identificaron con 7 y 6 péptidos, respectivamente, con el anzuelo CD81 y no se detectó ningún péptido de estas proteínas con otros anzuelos. Esto también ocurre en otros experimentos de otros donantes, como la proteína *translation initiation factor 3A*, identificada con 13 péptidos sólo con el anzuelo EWI-2, y la *stress 70 protein* identificada con 14 péptidos sólo con el anzuelo CXCR4, en el experimento 5 (Figura 14B). Otros ejemplos serían la *Heterogeneous nuclear ribonucleoprotein Q* o la *poly (A)-binding protein 1*, que se identificaron con un destacado número de péptidos con los anzuelos EWI-2 e ICAM-1 en el experimento 4 (Figura 14C). La proteína *filamin-A* se identifica con los anzuelos CD81, ICAM-1 y EWI-2 en el experimento 3 (Figura 14A), mientras que en los experimentos 4 (Figura 14C) y 5 (Figura 14B) sobresale con EWI-2 y CD9, y con EWI-2, respectivamente. La coherencia del resultado se debe a la solidez del protocolo utilizado, que solubiliza completamente las proteínas añadiendo SDS a los extractos proteicos y evita la pérdida de proteínas realizándose en un entorno bien controlado para la digestión tróptica [148].





**Figura 14: Distribución del número de péptidos de las proteínas identificadas que interactúan con los distintos anzuelos peptídicos en diferentes donantes.** Las gráficas representan el número de péptidos identificados por cada proteína en orden descendente según su mediana. Los colores representan los ligandos peptídicos. Las gráficas de la izquierda representan la distribución de este parámetro para las proteínas más abundantes en el experimento 3 (A), experimento 4(C) y experimento 5 (B). Los histogramas de la derecha representan el número de péptidos de las proteínas ejemplo que se indican en las flechas, mostrando las interacciones específicas con uno o más anzuelos.

La mayoría de las interacciones específicas observadas en un experimento también se detectaron en otras réplicas biológicas (Figura 15, sección derecha). La especificidad de las interacciones de algunas proteínas se confirmó por análisis de Western-blot (Figura 15, sección izquierda). Algunas proteínas, como la  $\alpha$ -actinin, filamin y la nucleolin (Figura 15A-C) mostraron interacción específica con los anzuelos de CD81, ICAM-1 y EWI-2, y se identificaron con un gran número de péptidos, mientras que su interacción con otros anzuelos apenas se detectó. Mediante Western-blot se observó que las proteínas ERMs (ezrin, radixin y moesin) se asociaban con la región citoplasmática de VCAM-1 y, en menor medida, con los anzuelos de EWI-2, ICAM-1, ICAM-3 y CD81, de acuerdo con estudios previos [63, 92, 149] (Figura 15G). Se ha observado, además, que algunas de las proteínas interactuaban específicamente con CD81, como las proteínas relacionadas con la familia ras, el ras-related C3 botulinum toxin substrate 2 (Rac2) (Figura 15D) o el elongation factor 1-alpha ( $\alpha$ -EF1) (Figura 15E). Las proteínas que interactuaron de manera no específica, como la actin, se detectaron con un número bajo de péptidos en la mayoría de los anzuelos en diferentes experimentos (Figura 15F).

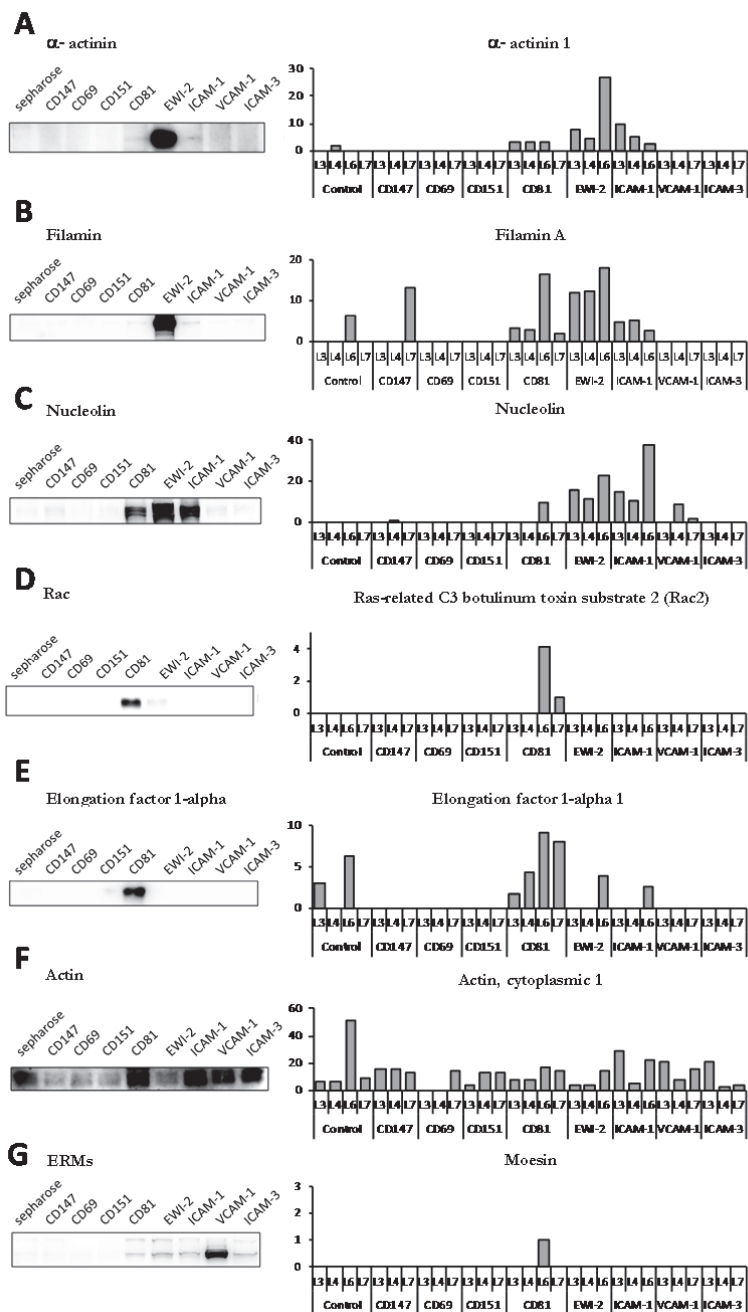


Figura 15: Especificidad de algunas interacciones en lisados totales de linfoblastos humanos. Las figuras muestran los datos correspondientes a la *alfa-actinin* (A), *filamin* (B), *nucleolin* (C), el *Rac* (D), *elongation factor 1-alpha* (E), *actin* (F) y ERM (G). Se muestran los resultados de los análisis proteómicos de pull-downs con los ligandos de CD147, CD69, CD151, CD81, EWI-2, ICAM-1, VCAM-1 e ICAM-3 en lisados de linfoblastos humanos de cuatro donantes independientes (que aparecen como L3, L4, L6 y L7).

### 3.3. Validación estadística de las interacciones de proteínas por recuento de péptidos

Para determinar las proteínas que interactúan específicamente con cada uno de los anzuelos, se realizó un análisis estadístico suponiendo que el número de péptidos identificados en una proteína con los distintos anzuelos sigue una distribución normal en cada donante. Para cada una de las proteínas, se calculó la mediana del número de péptidos identificados entre todos los anzuelos probados en cada donante y su desviación estándar (estimada de forma robusta mediante el método MAD –*Median Absolute Deviation*-). En el donante 6, la cantidad de anzuelos probados no fue suficientemente significativa (sólo 3 anzuelos) para dar una medida fiable en el cálculo de estos parámetros, por lo que en este caso se eligió el número de péptidos únicos identificados en el control como media. Para tener en cuenta las proteínas poco abundantes y ganar en sensibilidad, se dieron por válidas interacciones de proteínas con un solo péptido identificado con anzuelo en cuestión y ninguno en el resto de los anzuelos. Una vez calculados dichos parámetros, se consideró que una proteína interactuaba específicamente si la probabilidad de que la interacción fuera al azar era menor del 5% ( $p < 0.05$ ). Por último, se impuso un segundo filtro a los datos basado en la consistencia con que las diferentes interacciones se detectaron en las diferentes réplicas biológicas. Se consideró que una proteína interactuaba de forma significativa con un anzuelo si la interacción se reproducía en al menos la mitad de los experimentos realizados para dicho anzuelo (2 experimentos para el anzuelo CD81 (5 réplicas en total) y 3 experimentos para los anzuelos ICAM-1 y EWI-2 (6 réplicas en total)). En los ensayos de *pull-down* realizados con los demás anzuelos apenas se identificaron proteínas que interactuaban de forma significativa, por lo que no se usaron en estudios posteriores (Tabla Suplementaria 1). Un total de 207 proteínas mostraron interacción específica con al menos uno de los anzuelos CD81, EWI-2 e ICAM-1. De estas interacciones, 106 proteínas interactuaron con EWI-2, 54 con ICAM-1 y 110 con CD81 (Figura 16, Tabla Suplementaria 1). Los anzuelos EWI-2 e ICAM-1 presentaron un gran número de proteínas interaccionantes en común, en total 38 proteínas y los 3 anzuelos sólo presentaron en común 7 proteínas interaccionantes.

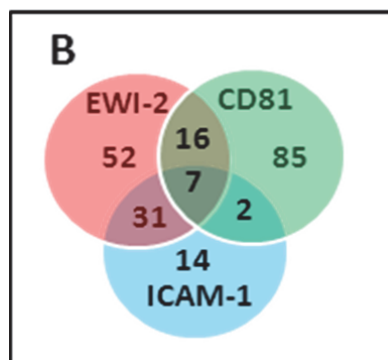
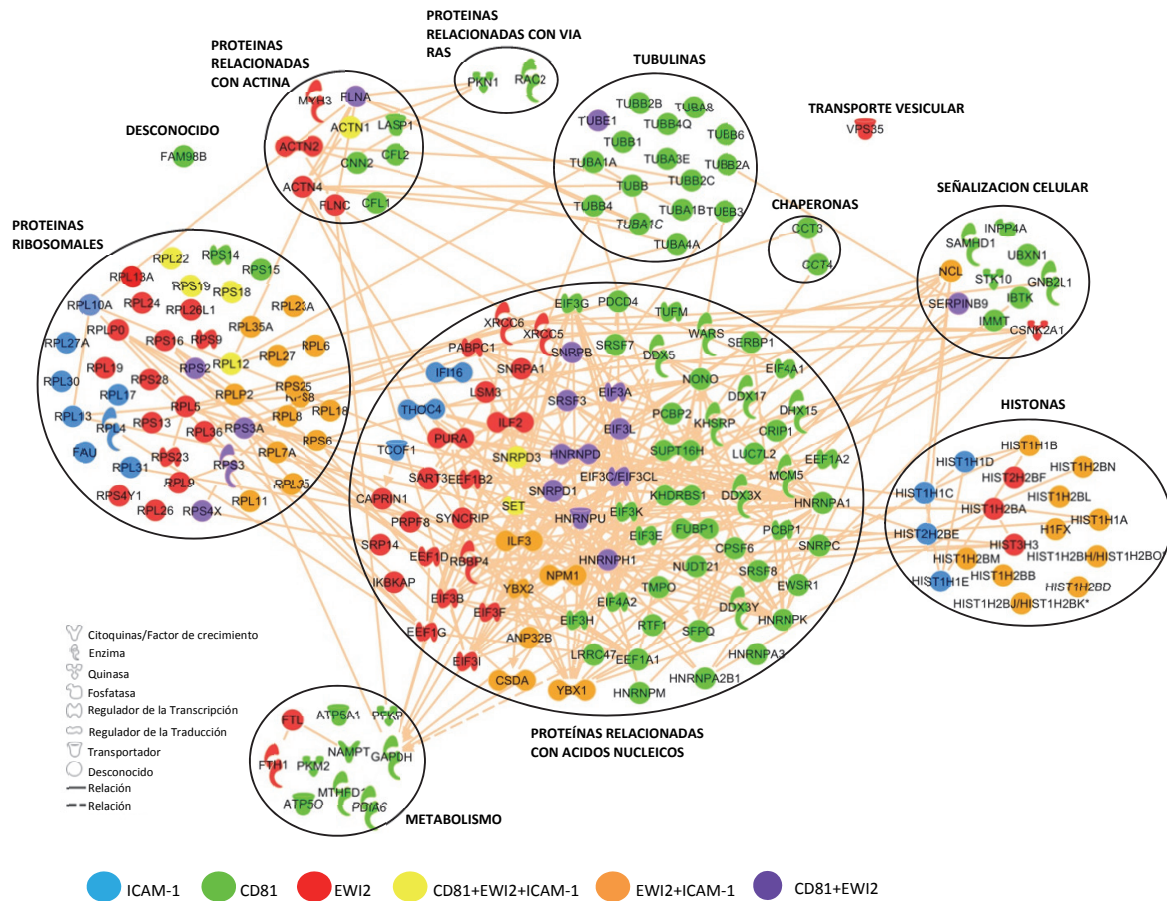


Figura 16: Número de proteínas que interactúan específicamente para los anzuelos de CD81, EWI-2 e ICAM-1. En el diagrama de Venn se muestra el grado de solapamiento en la interacción específica de las proteínas obtenidas de linfoblastos lisados con los ligandos peptídicos de CD81, EWI-2 e ICAM-1.



### 3.4. Estudio de las interacciones de proteínas con CD81, EWI-2 e ICAM-1 por Biología de Sistemas

Todas las interacciones específicas se analizaron mediante el programa *Ingenuity Pathway Analysis* (IPA, <https://analysis.ingenuity.com>), que reveló algunas pautas emergentes al agrupar las proteínas por su funcionalidad común en la red de interacciones (Figura 17). En conjunto, el grupo mayoritario compuesto por 82 proteínas estaba relacionado con ácidos nucleicos, donde aparecían un gran número de interacciones con CD81 (51 proteínas) y con EWI-2 (29 proteínas). Otros grupos funcionales numerosos fueron las proteínas ribosomales (en total 46 proteínas) y las histonas (en total 17 proteínas), que interactuaban principalmente con ICAM-1 y EWI-2, así como un grupo



**Figura 17: Redes de interacción de proteínas de los ligandos de CD81, EWI-2 e ICAM-1 identificados en linfocitos T humanos.** Los ligandos intracelulares de CD81, EWI-2 o ICAM-1 se agruparon de forma manual en grupos según su función ontológica con el programa de análisis de IPA. Las líneas representan interacciones validadas proteína-proteína, como se indica en el programa IPA. Las proteínas son de color de acuerdo a su interacción con ICAM-1 (azul), CD81 (verde), EWI-2 (rojo), con dos diferentes anzueltos (violeta y naranja) o tres anzueltos (amarillo).

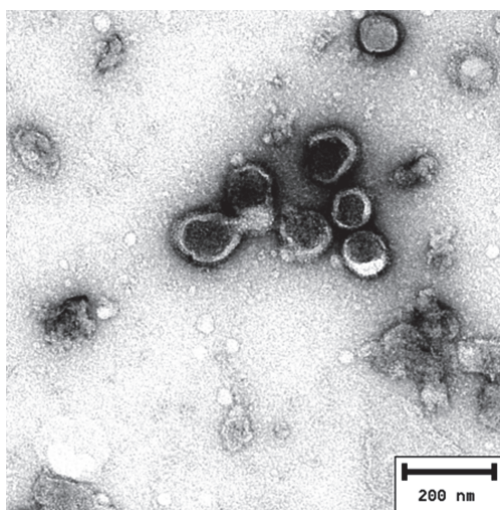


minoritario de proteínas relacionadas con la actina (10 proteínas), donde también solapaban interacciones con CD81 y EWI-2 (ver Tabla Suplementaria 1).

De forma destacada se observaron grupos de proteínas que interaccionaban con CD81 principalmente, tales como un grupo de 16 tubulinas, y otros grupos minoritarios pero no menos importantes como proteínas de señalización celular, proteínas del metabolismo, proteínas relacionadas con la familia *ras* y chaperonas (ver Tabla Suplementaria 1).

#### 4. CARACTERIZACIÓN DEL INTERACTOMA INTRACELULAR DE CD81 Y EWI-2 EN EXOSOMAS DE LINFOCITOS T HUMANOS.

Las interacciones de los componentes de los TEMs con proteínas ribosomales y proteínas de unión y procesamiento de ácidos nucleicos puede tener consecuencias funcionales en la adhesión y desarrollo de los linfocitos T [169, 170]. Además, se ha descrito que los exosomas transportan ARNm y ARNm y es por ello que las proteínas relacionadas con ácidos nucleicos se están empleando como marcadores moleculares de estos orgánulos [171-173]. Para explorar el papel de los TEMs en exosomas, se purificaron microvesículas secretadas a partir de los cultivos primarios de linfoblastos T humanos. El análisis por microscopía electrónica de estas preparaciones confirmó la presencia de



**Figura 18:** Análisis por microscopía electrónica de muestras de exosomas derivados de cultivos primarios de linfocitos humanos. Los exosomas se enriquecieron de los medios de cultivo de linfoblastos como se describe en Materiales y Métodos. Se visualizaron por tinción negativa en un microscopio electrónico Jeol JEM-1010.

estructuras de membrana de tamaño heterogéneo, que en su mayoría presentaron la clásica forma de copa, que es la morfología típica de los exosomas cuando se observan mediante tinción negativa (Figura 18).

En primer lugar, aplicamos la metodología puesta a punto para la identificación a gran escala de proteínas, previamente descrita (ver Apartados 1 y 2 de este Capítulo) para determinar el grado de profundidad de análisis que podíamos alcanzar en lo que respecta al proteoma del exosoma. Los

extractos de proteínas procedentes de las preparaciones de exosomas se digirieron en gel concentrador con el protocolo establecido y los péptidos resultantes se analizaron en una trampa iónica lineal. El resultado del análisis de alto rendimiento permitió la identificación de un total de 3488 péptidos únicos (FDR <5%) correspondientes a 1389 proteínas, procedentes de extractos de exosomas de cinco donantes humanos, de las cuales, 489 proteínas fueron identificadas en tres o más donantes. El 74% de los péptidos identificados pertenecían a proteínas que se encuentran publicadas en EXOCARTA (<http://www.exocarta.org>), que es una base de datos *on-line* que contiene las proteínas identificadas en los exosomas de multitud de tipos celulares y multitud de especies, con más de 4500 entradas. Estos resultados confirmaron la pureza de las preparaciones de exosomas y la alta cobertura de proteínas identificadas. Entre estas proteínas se incluyen una proporción relativamente grande de tetraspaninas y moléculas asociadas a tetraspaninas, tales como las isoformas de MHC, CD4, CD8, ADAM-10, CD98, integrinas, ERM y *syntenin*. El listado de todas las proteínas identificadas se muestran en la Tabla Suplementaria 2 y es, hasta la fecha, el estudio más profundo en la identificación de la composición proteica de exosomas de linfocitos T posterior a trabajos anteriores [174].

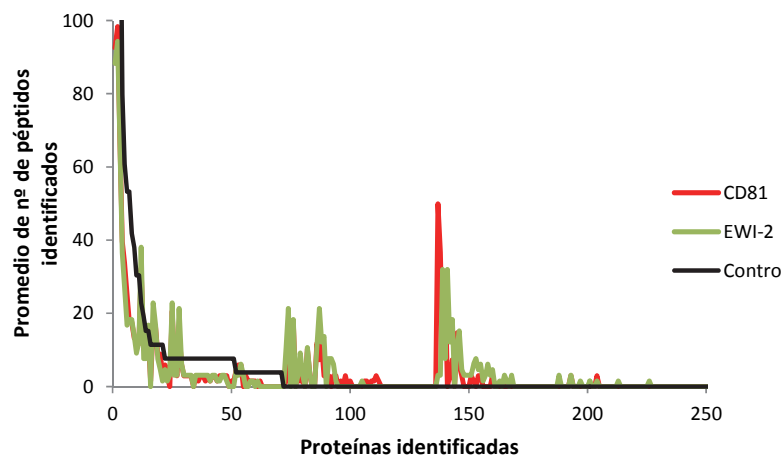
### 4.1. Análisis de las proteínas interaccionantes con CD81 y EWI-2 en exosomas de linfocitos T humanos

El estudio del interactoma de los dominios TEMs en exosomas se llevó a cabo utilizando como péptidos anzuelo los dominios C-terminal biotinilados de CD81 y EWI-2, ya que fueron los anzuelos más efectivos en la recuperación de proteínas interaccionantes, según los resultados obtenidos en el estudio del interactoma en los lisados celulares de linfocitos T. Los ensayos de *pull-down* se llevaron a cabo siguiendo la misma estrategia experimental, descrita en la Figura 13. Los péptidos biotinilados unidos a las microesferas de *Streptavidin-Sepharose* se incubaron con los extractos proteicos de exosomas procedentes de linfocitos T humanos de 5 donantes diferentes. Las microesferas con las proteínas interaccionantes unidas se digirieron con tripsina según el método optimizado de digestión en gel concentrador y los péptidos resultantes fueron identificados por espectrometría de masas de alto rendimiento en un LTQ. En total, se identificaron 2135 péptidos únicos (FDR<5%) correspondientes a 992 proteínas entre todos los ensayos realizados (Tabla 5).

**Tabla 5.** Número de péptidos identificados en todos los *pull-down* realizados con extractos de exosomas de linfocitos T humanos, a una FDR<5%.

Número de péptidos identificados			
Donante	CONTROL	CD81	EWI-2
E1	335	644	701
E2	242	645	720
E3	223	596	638
E4	209	808	816
E5	250	569	530

La especificidad de las interacciones se determinó mediante el recuento del número de péptidos identificados para cada proteína con cada uno de los anzuelos. El número de péptidos identificados en los ensayos de *pull-down* en los que se usaron microesferas sin péptido biotinilado unido se utilizó como control negativo. La distribución del número de péptidos por proteínas ordenados por su mediana fue muy similar en los dos anzuelos utilizados con un mismo donante. Las interacciones específicas fueron fácilmente detectables, destacándose claramente de la distribución de fondo no específica (Figura 19).



**Figura 19:** Distribución del número de péptidos promedio identificados para CD81 y EWI-2 y respectivo control en los *pull-down* de exosomas de linfocitos T humanos. La gráfica muestra el valor promedio de péptidos identificados en los experimentos realizados a los 5 donantes, ordenando las proteínas en orden descendente según el valor de la mediana en el control. Los colores representan los ligandos peptídicos CD81 (verde), EWI-2 (rojo) y control (negro), que muestra el "fondo" de proteínas que interactúan con las bolas de sepharose.

La validación estadística utilizada para distinguir las proteínas que interactúan específicamente con cada uno de los anzuelos se realizó mediante el análisis estadístico de la *t de Student* de una cola, considerando las varianzas de ambas poblaciones semejantes. Para cada una de las proteínas identificadas se calculó la probabilidad de que la interacción se debiera al azar, y se consideró que una proteína interactuaba significativamente si el *p-value* era menor del 5%. De todas las proteínas identificadas, 189 proteínas dieron interacción específica significativa, de las cuales 105 fueron específicas de CD81 y 148 lo fueron de EWI-2; 64 proteínas interactuaron con ambos anzuelos (Tabla Suplementaria 3 y Figura 20).

Las proteínas exosomales interaccionantes de cada anzuelo se repiten con un patrón aún más definido en las diferentes réplicas biológicas y la especificidad de estas interacciones fue confirmada por análisis de Western-blot (Figura 20). Algunas de las proteínas fueron detectadas previamente en lisados

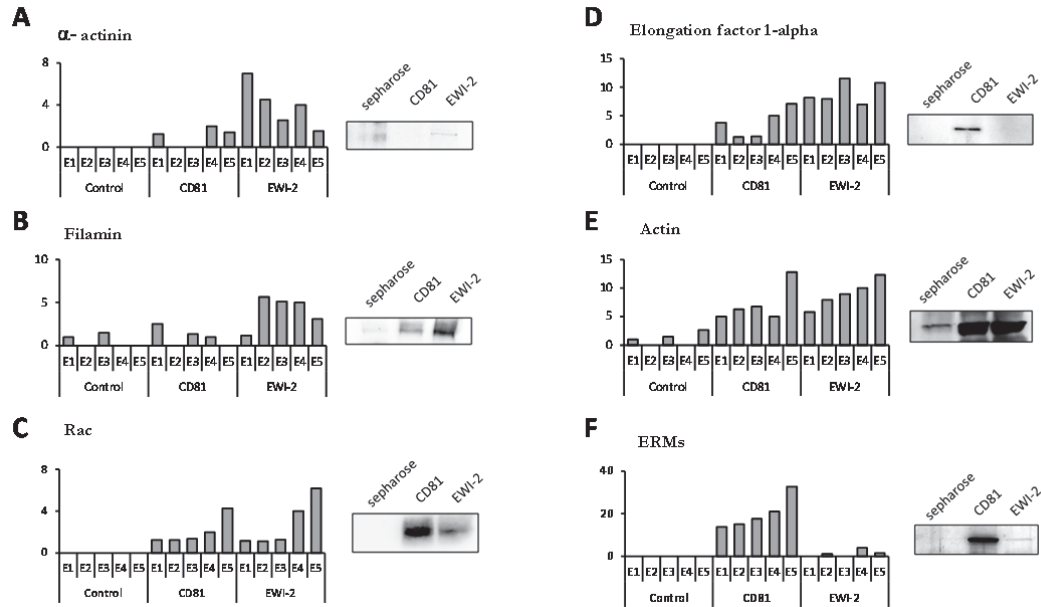


Figura 20: La especificidad de algunas interacciones en linfoblastos humanos se confirman en sus exosomas (ver figura 15). Las figuras muestran los datos correspondientes a las mismas proteínas que la figura 10:  $\alpha$ -actinin (A), filamin (B), Rac2 (C), elongation factor 1-alpha (D), actin (E) y ERM (F). Se muestra el número de péptidos identificados en los pull-downs con los anzuelos de CD81 y EWI-2 en los extractos de exosoma provenientes de linfocitos T humanos de 5 donantes diferentes (que aparece como E1-E5). Los Western Blot confirman estas interacciones.

linfocitarios (Figura 15), como  $\alpha$ -actinin o filamin A (Figura 20 A-B), donde se observó que interaccionaban con los anzuelos CD81 y EWI-2, aparentemente con más afinidad con EWI-2, confirmándose este dato por Western-blot. Además, se observó una interacción con ambos anzuelos de las proteínas relacionadas con la familia ras, como el ras-related C3 botulinum toxin substrate 2 (Rac2) (Figura 20C), donde previamente sólo se había detectado interacción con CD81 en los lisados celulares, y el elongation factor 1-alpha ( $\alpha$ -EF1) (Figura 20D). La proteína nucleolin no se identificó en ningún ensayo de pull-down en exosomas, aunque si se observó su interacción en lisados totales (Figura 15C). Se ha observado además que algunas de las proteínas interaccionan específicamente con CD81, como las ERMs, las cuales no se obtuvo interacción específica de manera significativa por espectrometría de masas en lisados celulares (Figura 15).

## 4.2. Estudio de las interacciones de proteínas de exosoma con CD81 y EWI-2 por Biología de Sistemas

El análisis de las interacciones identificadas en exosomas mediante el programa bioinformático de biología de sistemas *Ingenuity Pathway Analysis* (IPA) reveló una vez más una red coherente entre las proteínas interaccionantes (Figura 21). La mayoría de los grupos de proteínas que interaccionan con CD81 y EWI-2 en lisados celulares totales también fueron detectados en los exosomas. Así, las proteínas

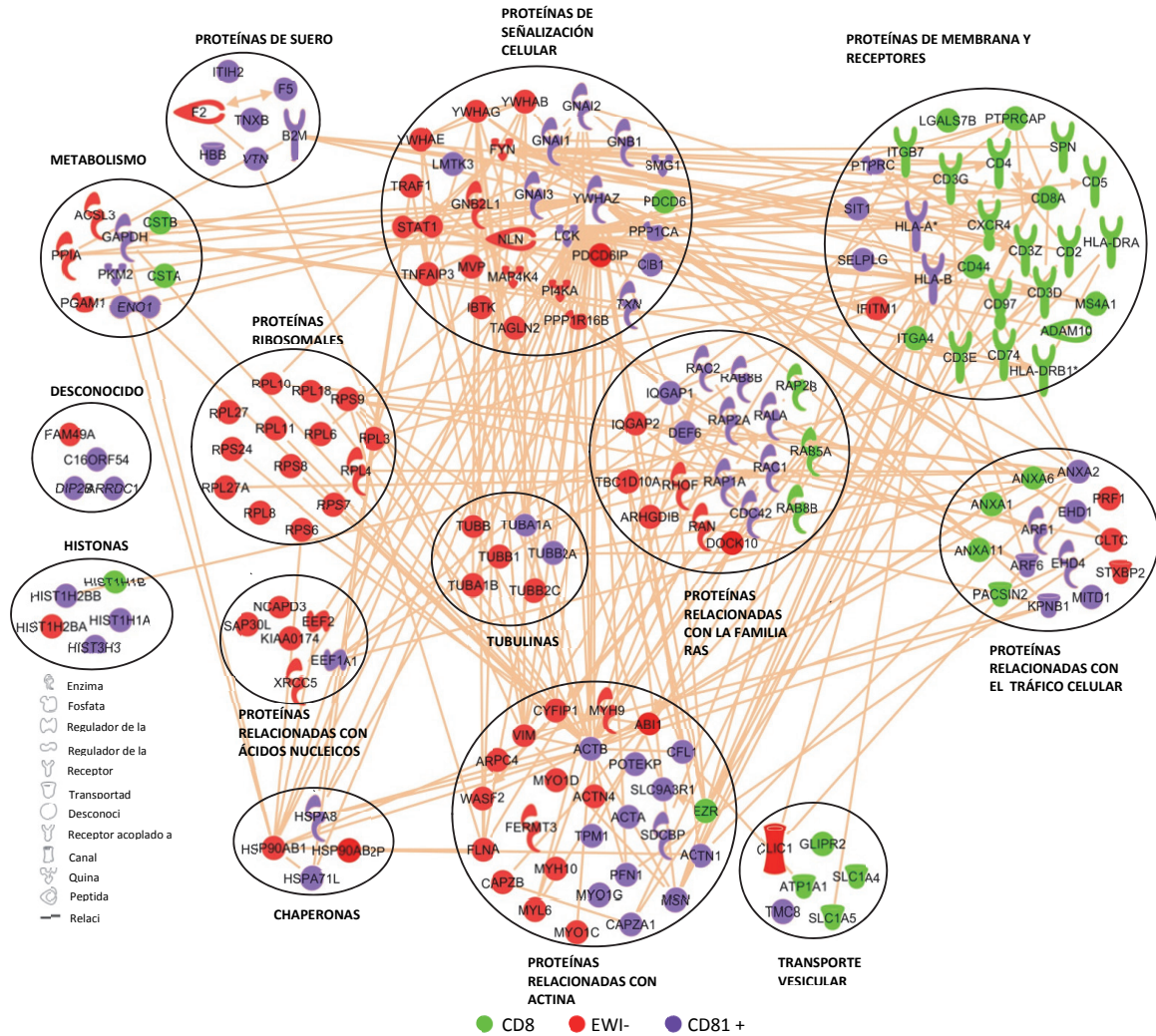


Figura 21: Redes de proteínas que interaccionan con los anzuelos CD81 y EWI-2 en extractos de exosoma de linfocitos T humanos. Las líneas representan interacciones proteína-proteína validadas, como se indica en el programa IPA. Las proteínas que interaccionan están agrupadas por el proceso ontológico en que están involucradas, y su color es de acuerdo a su interacción con CD81 (verde), EWI-2 (rojo), y con los dos anzuelos (violeta).



como  $\alpha$ -actinin, filamin A, y los grupos de proteínas como las histonas (5 proteínas), las tubulinas (6 proteínas), las proteínas ribosomales (14 proteínas), las chaperonas (4 proteínas), las proteínas relacionadas con ácido nucleicos (6 proteínas), con el metabolismo (8 proteínas), con la señalización celular (27 proteínas) y proteínas de la familia *ras* (18 proteínas) también fueron validadas como interacciones en exosomas. Además, se confirmaron interacciones con algunas proteínas que están altamente enriquecidas en exosomas, como el caso las proteínas ERMs, cuya interacción con CD81 y EWI-2 ya se había descrito en un trabajo anterior [92] (Figura 20F). Por otro lado, se observaron nuevas interacciones en exosomas, tales como algunas proteínas del grupo de los receptores de membrana y proteínas transmembrana asociadas a tetraspaninas, como CD44, CD20 y HLA (27 proteínas en total), interaccionando principalmente con CD81. Este resultado sugiere que las interacciones intracelulares entre estas proteínas y CD81 reforzaría la unión lateral existente en los TEMs en la membrana plasmática, pudiendo ser relevantes en la regulación de la composición de los propios TEMs. Sólo dos nuevos grupos de proteínas interaccionantes aparecieron en exosomas y no se observaron en lisado total: las proteínas relacionadas con el tráfico celular (14 proteínas) y las proteínas de transporte vesicular (6 proteínas) (Figura 21).

### 4.3. El papel de los TEMs en la composición proteica de los exosomas

La composición proteica de los exosomas, así como los resultados de interacción obtenidos de los experimentos de *pull-down* indicaron que los anzuelos de la tetraspanina CD81 y su molécula asociada EWI-2 abarcan una proporción considerable del proteoma de exosoma. El interactoma intracelular de EWI-2 y CD81 en lisados celulares totales representa el 11% del número de proteínas de exosomas (Figura 22A) identificadas, y el 45% del componente peptídico identificado en las dichas microvesículas (Figura 22B). Puesto que el número de péptidos identificados en experimentos de espectrometría de masas es proporcional a la concentración en peso de las proteínas [113], estos resultados sugieren que la mayoría de las de las proteínas dentro de los exosomas interaccionan con TEM y, por tanto, que las tetraspaninas podrían regular la inclusión de proteínas en la estructura exosomal.

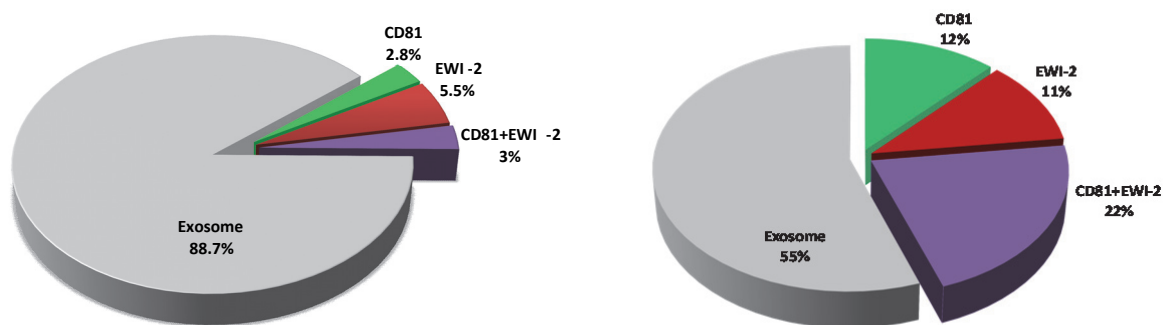


Figura 22: Proporción de proteínas de exosomas que pertenecen al interactoma de los ligandos CD81 y EWI-2. El porcentaje de proteínas identificadas (A) y de péptidos identificados (B) en extractos de exosoma se calcula sumando el número de proteínas o de péptidos identificados, respectivamente, en exosomas, que pertenecen a las proteínas que interactúan con los anzuelos de EWI-2 y de CD81.

## 5. ESTUDIO DE LOS CAMBIOS EN EL PROTEOMA DE EXOSOMA POR LA DELECCIÓN DE LAS TETRASPANINAS CD9 Y CD81

Para estudiar de forma directa la participación de las tetraspaninas en la composición proteica de los exosomas, se comparó la composición del proteoma de las microvesículas excretadas por linfoblastos de ratón “salvajes” (WT) con ratones *knock out* CD81  $-/-$  (CD81KO) mediante proteómica cuantitativa a gran escala mediante marcaje isotópico estable de los péptidos con  $^{18}\text{O}$ . A modo comparativo se procedió a hacer un análisis de la composición del proteoma de exosomas de ratones WT frente a ratones *knock out* CD9  $-/-$  (CD9KO). CD9 es una tetraspanina íntimamente relacionada con CD81 y EWI-2 [4, 6, 92].

Las muestras se procesaron con el protocolo optimizado de digestión triptica de las proteínas en gel concentrador. Los péptidos resultantes de la digestión de los extractos proteicos de cada muestra (ratones WT y ratones *knock out*) se marcaron de forma diferencial con  $^{16}\text{O}/^{18}\text{O}$  con tripsina inmovilizada. Tras el marcaje, se inhibió la actividad enzimática residual de tripsina que pudiera haber en las muestras. Para el experimento comparativo de ratones WT frente a ratones CD9KO se procedió a fraccionar los péptidos en solución usando cartuchos de MCX (relleno de retención mixta (intercambio catiónico-fase reversa), mientras que para el experimento de ratones WT frente a ratones CD81KO se utilizó el isoelectroenfoque (IEF) como método de fraccionamiento en solución de los péptidos. Los pares de péptidos resultantes marcados con  $^{16}\text{O}/^{18}\text{O}$  se analizaron por espectrometría de masas en una trampa iónica lineal (LTQ), siendo un total de 8 carreras (número de fracciones del MCX) y de 24 carreras (número de fracciones de IEF) (Figura 8). El control de la eficiencia de marcaje así como la cuantificación relativa de los péptidos se calculó de forma

automática y con una posterior revisión manual mediante la plataforma bioinformática QuiXoT, desarrolla en nuestro laboratorio [117].

## 5.1. Análisis de los cambios en la abundancia de proteínas en exosomas por la delección de CD9 en ratón

El análisis cuantitativo se realizó a partir de un *pool* de extractos de exosomas secretados por linfoblastos de 3 ratones CD9  $-/-$  y se comparó con un *pool* procedente de exosomas de 3 ratones WT, utilizando el marcaje con  $^{18}\text{O}$  para los ratones WT y con  $^{16}\text{O}$  para los ratones CD9KO. En total, se cuantificaron 396 proteínas ( $FDR < 5\%$ ). La distribución de los  $\log_2$ -ratios a nivel de proteína estandarizados según sus varianzas estimadas ( $Z_q$ ) se ajustaba perfectamente a una función gaussiana, siguiendo la tendencia de la hipótesis nula descrita por el modelo estadístico desarrollado para el análisis de datos cuantitativos [117, 148] (Figura 23, recuadro). El efecto de la delección de CD9 en el proteoma de exosoma provocó una disminución significativa en la abundancia de 16 de las proteínas cuantificadas ( $FDR_q < 5\%$ ), no llegándose a observar un aumento significativo de ninguna proteína (Figura 23).

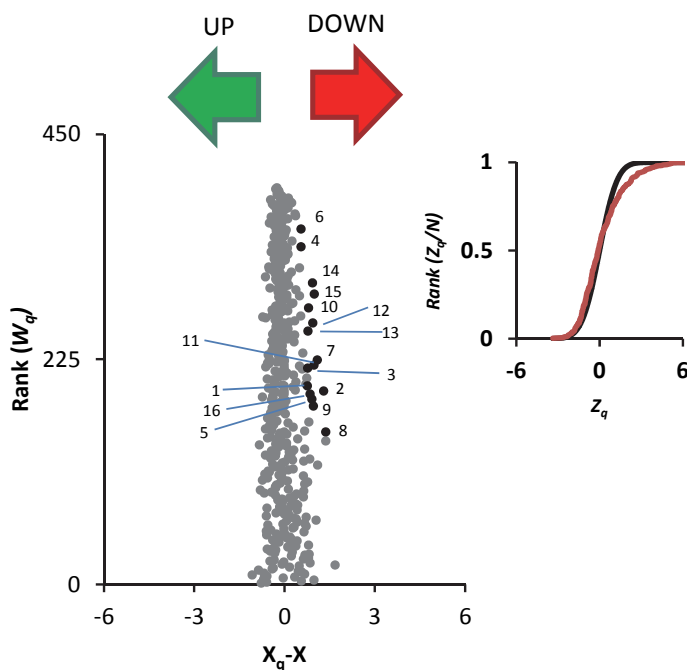


Figura 23: Efecto de la delección de CD9 en la composición de proteínas de exosomas en linfocitos T de ratón. Análisis cuantitativo de la composición de proteínas de exosomas de linfoblastos de ratones WT y linfoblastos/esplenocitos delecionados de CD9 y marcaje con el isótopos estables  $^{18}\text{O}$ , seguido de análisis de espectrometría de masas de alto rendimiento. Los puntos negros indican las proteínas con los cambios de abundancia estadísticamente significativos con respecto a la mayoría de los que no cambian las proteínas (puntos grises), y están numerados para ser identificados en la tabla 9. Los números corresponden a las 16 proteínas con diferencias de abundancia significativa de las 396 proteínas cuantificadas, nombradas en la tabla 9. El recuadro muestra la distribución de la frecuencia acumulada de la variable estandarizada  $z_q$  (línea de negro), que expresa los datos cuantitativos en términos de unidades de la varianza, y se muestra que cumple con la distribución de hipótesis nula esperada y con la desviación debida a las proteínas que cambian su abundancia.

Esta subpoblación de proteínas que disminuyó significativamente su abundancia se observa en una desviación de la hipótesis nula en la parte derecha de la distribución  $\log_2$ -ratio (Figura 23, recuadro). Entre las 16 proteínas que disminuyeron su abundancia, la mayoría eran proteínas de la membrana plasmática, tales como CD147, CD48, CD5 y MHCs de clase I, clásicamente relacionadas con las tetraspaninas [2]. También se observaron proteínas relacionadas con la familia *ras*, como el *Rab GDP dissociation inhibitor alpha* y proteínas de señalización celular, tales como las proteínas 14-



3-3 epsilon, 14-3-3 eta y guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2. Otra proteína afectada estaba relacionada con el transporte de metabolitos como la *choline transporter-like protein 2*, así como la proteína *ribosomal 60S acidic ribosomal protein P0*. En su conjunto, las proteínas afectadas en el ratón CD9  $-/-$  eran, en su mayoría, específicas de superficie celular. A pesar del bajo número de proteínas cuantificadas, los resultados obtenidos presentan bastante similitud con las proteínas que interaccionaban con CD81, EWI-2 o ICAM-1 en lisado total (Apartado 3 de este Capítulo). La relación de proteínas que cambian su abundancia debido a la depleción de CD9 en exosoma se muestra en la Tabla 6.

**Tabla 6.** Cambios en la abundancia de proteínas detectados en exosomas de linfocitos T de ratones CD9  $-/-$  Knock-out respecto a los exosomas de animales sin deplecionar.

Número <sup>a</sup>	Número de Acceso	Proteína	log <sub>2</sub> -ratio corregido <sup>b</sup>	Cambio de abundancia	Z <sub>q</sub> <sup>c</sup>	FDR <sub>q</sub> <sup>d</sup>
1	P18572	Basigin (CD147)	-0.72	1.65 DOWN	3.02	4.48E-02
2	Q3TED2	Programmed cell death6-interacting protein	-0.77	1.70 DOWN	3.20	2.50E-02
3	P14869	60S acidic ribosomal protein P0	-0.78	1.71 DOWN	3.28	2.13E-02
4	Q8BY89	Choline transporter-like protein 2	-0.56	1.47 DOWN	3.38	1.79E-02
5	Q9D8N0	Elongation factor 1-gamma	-0.85	1.81 DOWN	3.50	1.14E-02
6	P12763	Alpha-2-HS-glycoprotein	-0.55	1.47 DOWN	3.63	8.76E-03
7	P68510	14-3-3 protein eta	-0.79	1.73 DOWN	3.70	6.55E-03
8	P18181	CD48 antigen	-0.92	1.89 DOWN	3.73	5.81E-03
9	Q3SZH5	Angiotensinogen (Serpin peptidase inhibitor, clade A, member 8)	-0.97	1.96 DOWN	3.89	3.92E-03
10	P13379	T-cell surface glycoprotein CD5	-0.81	1.75 DOWN	3.97	2.79E-03
11	P56528	ADP-ribosyl cyclase 1	-0.99	1.99 DOWN	4.10	2.30E-03
12	Q99LX0	Protein DJ-1	-0.95	1.93 DOWN	4.49	4.03E-04
13	P62259	14-3-3 protein epsilon	-1.10	2.14 DOWN	4.55	3.07E-04
14	P01901	H-2 class I histocompatibility antigen, K-B alpha chain	-0.93	1.91 DOWN	4.88	1.07E-04
15	P62880	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	-0.99	1.99 DOWN	4.99	6.05E-05
16	P50396	Rab GDP dissociation inhibitor alpha	-1.30	2.47 DOWN	5.18	8.66E-05

<sup>a</sup> Las proteínas que muestran los cambios de abundancia estadísticamente significativas están numeradas como en la figura 23.

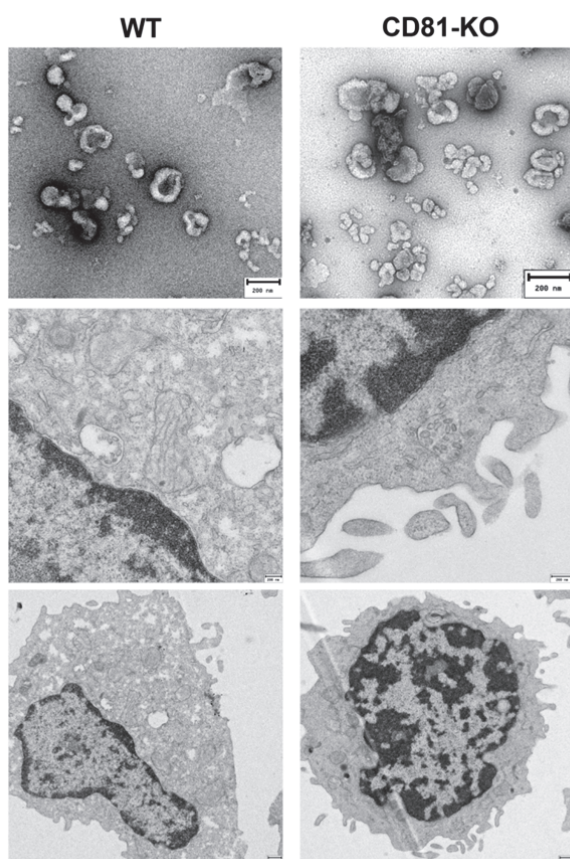
<sup>b</sup> El log<sub>2</sub>-ratio se corrigió restando el valor de la gran media.

<sup>c</sup> Los valores normales estandarizados se obtienen dividiendo el log<sub>2</sub>-ratio de la proteína por su varianza. La magnitud del cambio de la variable estandarizada (Z<sub>q</sub>) se muestra de acuerdo a la escala de colores indicada en la parte superior de la tabla.

<sup>d</sup> Tasa de posibles falsos errores de cuantificación a nivel de proteína (ver ref. 148)

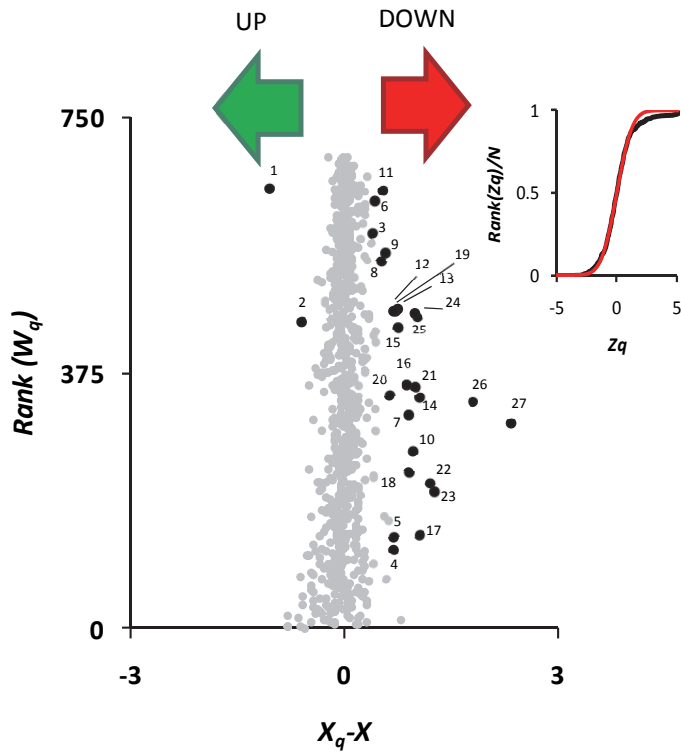
## 5.2. Análisis de los cambios en la abundancia de proteínas en exosomas por la delección de CD81 en ratón

Con el fin de explorar aún más la hipótesis de que las tetraspaninas podrían dirigir el reclutamiento de proteínas hacia el exosoma, se llevó a cabo el análisis comparativo de la composición del proteoma de exosomas de ratones al deleccionar CD81 por proteómica cuantitativa de alto rendimiento utilizando marcaje con <sup>18</sup>O. Se prepararon microvesículas excretadas por linfoblastos de 3 ratones deficientes en CD81 y se comparó con las microvesículas de 3 ratones WT. La microscopía electrónica no reveló diferencias apreciables en la morfología de los exosomas, células o cuerpos multivesiculares (Figura 24).



**Figura 24:** Análisis por microscopía electrónica de exosomas provenientes de ratones “salvajes” (WT) y deficientes en CD81 (CD81-KO). Exosomas (fotografías superiores) se enriquecieron desde los medios de cultivo de linfoblastos como se describe en Métodos y tefidas negativamente. Las células (fotografías inferiores) se sedimentaron y se fijaron con glutaraldehído al 2% y 4% de paraformaldehído y embebido en resina epoxi. Se observaron las secciones ultrafinas con un microscopio Jeol JEM-1010 de electrones, previa contratinación con acetato de uranilo y citrato de plomo. Un aumento mayor de un cuerpo multivesicular se muestra en las fotografías del medio.

Utilizando el mismo desarrollo experimental descrito anteriormente, los péptidos resultantes de la digestión triptica de los *pools* de los extractos de proteínas procedentes de CD81KO y de WT se marcaron diferencialmente con  $^{16}\text{O}/^{18}\text{O}$ . El análisis comparativo reveló un total de 692 proteínas cuantificadas ( $FDR < 5\%$ ) y la distribución de los  $\log_2$ -ratios estandarizados se ajustaba a una curva gaussiana, siguiendo la hipótesis nula descrita por el modelo estadístico para cuantificación. La existencia de una subpoblación de proteínas que disminuyó significativamente su abundancia se reflejó claramente en la desviación de la hipótesis nula en la parte derecha de la distribución  $\log_2$ -ratio (Figura 25, recuadro), siendo un total de 25 proteínas significativamente menos abundantes ( $FDR_q < 5\%$ ) en los exosomas de los ratones CD81KO (Figura 25), indicando que la ausencia de CD81 podría afectar a la inclusión de estas proteínas en los exosomas.



**Figura 25: Efecto de la delección de CD81 en la composición de proteínas de exosomas.** Análisis cuantitativo de la composición de proteínas de exosomas de linfoblastos de ratones WT y linfoblastos/esplenocitos delecionados de CD81 y marcaje con el isótopos estables  $^{18}\text{O}$ , seguido de análisis de espectrometría de masas de alto rendimiento. Los puntos negros indican las proteínas con los cambios de abundancia estadísticamente significativos con respecto a la mayoría de los que no cambian las proteínas (puntos grises), y están numerados para ser identificados en la tabla 10. Como se muestra, CD81 (proteína 27) es la proteína con más reducción en su abundancia entre las 692 proteínas cuantificadas. Los números corresponden a las proteínas enumeradas en la tabla 10. El recuadro muestra la distribución de la frecuencia acumulada de la variable estandarizada  $z_q$  (línea de negro), que expresa los datos cuantitativos en términos de unidades de la varianza, y se muestra que cumple con la distribución de hipótesis nula esperada y con la desviación debida a las proteínas que cambian su abundancia.

La relación de las proteínas que disminuyeron su abundancia debido a la depleción de CD81 en exosoma se muestra en la Tabla 7. Entre ellas, cabe destacar algunas proteínas de membrana asociadas a tetraspaninas como componentes del complejo BCR, CD20, ICAM-1 y distintos isotipos de MHC. Las interacciones de éstas proteínas con CD81 ya han sido descritas en otros trabajos [175-177]. Otras proteínas de membrana aún no descritas como la *interleukin-2 receptor subunit alpha* y el *transient receptor potential cation channel subfamily V member* también se vieron atenuadas. Algunas proteínas afectadas fueron las proteínas relacionadas con la familia *ras* como la *ras-related C3 botulinum toxin substrate 2* (Rac2), la *transforming protein RhoA* y la *cell división control protein 42 homolog* (cdc42), y proteínas ribosomales como la *ribosomal protein S5* y la *40S ribosomal protein S3a*. Según lo previsto, la proteína más disminuida en su abundancia fue la proteína CD81.

**Tabla 7.** Cambios en la abundancia de proteínas detectados en exosomas de linfocitos T de ratones CD81 *-/- Knock-out* respecto a los exosomas de animales sin deplecionar.

Número <sup>a</sup>	Número de Acceso	Proteína	log <sub>2</sub> -ratio corregido <sup>b</sup>	Cambio de abundancia	Z <sub>q</sub> <sup>c</sup>	FDR <sub>q</sub> <sup>d</sup>
1	Q64514	Tripeptidyl-peptidase 2	-1.05	2.07 UP	-9.35	0.00E+00
2	Q05CR0	LOC72520 protein (Fragment)	-0.61	1.53 UP	-4.22	8.82E-04
3	P19437	B-lymphocyte antigen CD20	0.39	1.31 DOWN	3.17	0.04
4	Q9WTR1	Transient receptor potential cation channel subfamily V member 2	0.69	1.61 DOWN	3.23	0.04
5	P62814	V-type proton ATPase subunit B, brain isoform	0.69	1.62 DOWN	3.34	0.02
6	P14483	H-2 class II histocompatibility antigen, A beta chain	0.40	1.32 DOWN	3.50	0.01
7	Q6MGB8	RT1 class I, A2	0.64	1.56 DOWN	3.65	0.01
8	Q05144	Ras-related C3 botulinum toxin substrate 2	0.51	1.42 DOWN	3.90	3.03E-03
9	Q9QU10	Transforming protein RhoA	0.55	1.47 DOWN	4.35	4.88E-04
10	P31254	Ubiquitin-like modifier-activating enzyme 1 Y	0.89	1.85 DOWN	4.66	1.17E-04
11	P01872	Ig mu chain C region secreted form	0.53	1.44 DOWN	4.70	1.10E-04
12	P01590	Interleukin-2 receptor subunit alpha	0.67	1.60 DOWN	4.74	9.18E-05
13	Q9QZ19	Serine incorporator 3	0.70	1.63 DOWN	4.92	3.70E-05
14	Q91V55	Ribosomal protein S5	0.89	1.85 DOWN	4.99	3.20E-05
15	P01898	H-2 class I histocompatibility antigen, Q10 alpha chain	0.74	1.67 DOWN	5.03	2.56E-05
16	Q64467	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	0.87	1.83 DOWN	5.08	1.99E-05
17	Q9Z0N1	Eukaryotic translation initiation factor 2 subunit 3, X-linked	1.05	2.08 DOWN	5.10	2.09E-05
18	Q80UL9	Junctional adhesion molecule-like	0.96	1.94 DOWN	5.14	1.75E-05
19	P13597	Intercellular adhesion molecule 1	0.74	1.67 DOWN	5.19	1.81E-05
20	A8DUN2	Beta-globin	1.00	2.00 DOWN	5.74	8.12E-07
21	P14046	alpha-1-inhibitor 3	1.05	2.06 DOWN	5.96	2.13E-07
22	P16045	Galectin-1	1.20	2.30 DOWN	6.27	4.12E-08
23	P97351	40S ribosomal protein S3a	1.26	2.40 DOWN	6.51	8.48E-09
24	P62880	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	0.97	1.96 DOWN	6.81	2.28E-09
25	P60766	Cell division control protein 42 homolog	1.02	2.03 DOWN	7.08	3.41E-10
26	P01882	Ig delta chain C region membrane-bound form	1.79	3.47 DOWN	10.15	0
27	P35762	CD81 antigen	2.34	5.07 DOWN	12.95	0

<sup>a</sup> Las proteínas que muestran los cambios de abundancia estadísticamente significativas están numeradas como en la figura 25.

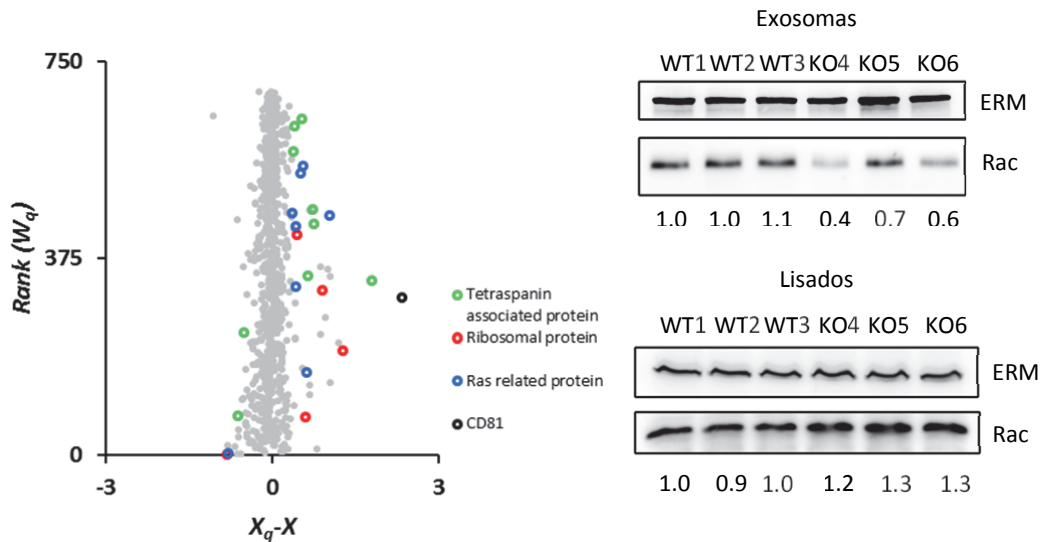
<sup>b</sup> El *log<sub>2</sub>-ratio* se corrigió restando el valor de la gran media.

<sup>c</sup> Los valores normales estandarizados se obtienen dividiendo el *log<sub>2</sub>-ratio* de la proteína por su varianza. La magnitud del cambio de la variable estandarizada (Z<sub>q</sub>) se muestra de acuerdo a la escala de colores indicada en la parte superior de la tabla.

<sup>d</sup> Tasa de posibles falsos errores de cuantificación a nivel de proteínas (ver ref . 148)

Al relajar el criterio de corte en la significatividad del cambio de abundancia a una  $|Z_q| > 2.4$  ( $FDR_q < 22\%$ ) se observó un enriquecimiento en la disminución de algunos grupos de proteínas muy relevantes. Tal es el caso de un grupo que forman 4 proteínas ribosomales (Figura 26A, puntos rojos), que se ven significativamente disminuidas en su abundancia. Asimismo se observó un grupo de 6 proteínas relacionadas con tetraspaninas (Figura 26A, puntos verdes) y un grupo de 7 proteínas relacionadas con la familia *ras* (Figura 26A, puntos azules) en el que también se apreciaba una tendencia a disminuir su abundancia relativa. Estas categorías se han observado previamente como parte del interactoma de CD81 y EWI-2 en los ensayos de *pull-down* en linfocitos T humanos y también en sus exosomas. En la categoría de proteínas relacionada con la familia *ras* en particular, se realizó un análisis por Western-blot que confirmó que la proteína Rac era menos abundante en

los exosomas deficientes en CD81 (Figura 26B), a pesar de que en los linfocitos T deficientes en CD81 la expresión total de *Rac* no se modificó (Figura 26B).



**Figura 26: La ausencia de CD81 produce cambios en la composición proteica de exosomas.** Comparación en la composición de proteínas de exosomas derivados de ratones delecionados para CD81 y ratones control, mediante el marcaje con isótopos estables seguido por el análisis de proteómica cuantitativa. Mezclas de tres extractos de exosoma murinos de linfoblastos delecionados de CD81 y WT fueron digeridos con tripsina y los péptidos resultantes marcados, ya sea con  $^{16}\text{O}$  o con  $^{18}\text{O}$ . Se mezclaron, se fraccionaron y cuantificaron por análisis LC-MS/MS. La gráfica muestra la distribución de los  $\log_2$ -ratio de las proteínas, clasificados por su peso estadístico (que es la inversa de la varianza y las medidas de la exactitud de la cuantificación). (A) Las proteínas cuantificadas a partir de las categorías ontológicas muestran un comportamiento similar en los cambios de su abundancia. Los puntos coloreados son proteínas ribosomales, proteínas relacionados con Ras, y proteínas de la membrana y asociadas a tetraspaninas con  $Zq \geq |2.4|$ . Las proteínas marcadas están nombradas en la tabla suplementaria 4. Estas categorías muestran una alteración de su abundancia coordinada de una parte del proteoma. (B) Análisis por Western-blot del contenido de Rac en exosomas o en lisados celulares de los distintos ratones separados, los salvaje (WT1-3) o los esplenocitos delecionados en CD81 (KO4-6). Las proteínas ERM se muestran como control de carga. Los números corresponden a los análisis densitométricos de la señal de Rac y corregido por el control de carga de la proteína ERM.



## DISCUSIÓN





## 1. DESARROLLO DE UNA METODOLOGÍA PARA EL ESTUDIO DE INTERACCIONES DE PROTEÍNAS A GRAN ESCALA MEDIANTE HERRAMIENTAS PROTEÓMICAS

En este trabajo hemos desarrollado un método de identificación sistemática y a gran escala de interacciones entre proteínas basado en la técnica de *pull-down* y el análisis por EM. En primer lugar, nos planteamos la necesidad de optimizar las condiciones experimentales de la digestión trípica de proteínas, con el fin de minimizar los artefactos producidos en dicho proceso, y conseguir un protocolo sólido y reproducible. El método más comúnmente usado en proteómica es la digestión en solución después de la desnaturalización de las proteínas con urea [117, 178, 179]. Tras la digestión, para eliminar contaminantes de la muestra que puedan interferir en el análisis por EM se suelen usar filtros [180, 181]. En muchos trabajos de proteómica, las proteínas se digieren en solución y los péptidos resultantes se fraccionan mediante cromatografía de intercambio catiónico, previo al análisis de cromatografía líquida de alta presión en fase reversa (RP-HPLC)-MS [103]. Una alternativa a la digestión en solución seguida del fraccionamiento de péptidos es la separación de proteínas por SDS-PAGE seguida de un troceado del gel, la digestión de cada trozo por separado y el análisis de los péptidos de cada uno de ellos por RP-HPLC-MS. La digestión en gel resulta ser más efectiva porque se usa SDS para la solubilización de las proteínas. Además, la matriz del gel permite eliminar de manera efectiva el SDS y otros detergentes y/o contaminantes de la muestra que dificultan el análisis de EM. Por otro lado, este tipo de aproximación experimental basada en la solubilización de las proteínas con SDS es compatible con la mayoría de los diferentes tipos de proteomas, incluso para el análisis de proteínas hidrofóbicas como las proteínas de membrana. Esta estrategia se ha utilizado en muchos trabajos de proteómica cuantitativa, usando SILAC como método de marcaje isotópico [104]. Sin embargo, en esta estrategia el marcaje isotópico se lleva a cabo metabólicamente, en una etapa anterior a la separación de las proteínas y a la digestión de las mismas. Hasta el momento, no se había demostrado que esta estrategia sea útil en la práctica cuando se emplean marcajes isotópicos postdigestión, como es el caso de  $^{18}\text{O}$  ó iTRAQ, en los que la reproducibilidad en la separación y digestión de las proteínas es crítica para que la cuantificación sea fiable.

En este trabajo hemos diseñado un nuevo protocolo de digestión trípica basado en la concentración de proteínas en SDS-PAGE en una única banda, para aprovechar las ventajas de la digestión en gel y evitar así el efecto de “corte de gel”. Para demostrar que el nuevo método de digestión en gel concentrador es más sólido y reproducible que los métodos de digestión en gel

separador y en solución, marcamos con  $^{18}\text{O}$  los péptidos trípticos resultantes de la digestión de dos réplicas técnicas del mismo proteoma. Los resultados obtenidos en varios proteomas de distinta naturaleza se analizaron utilizando un modelo estadístico publicado previamente [103, 104, 117] con el fin de detectar posibles artefactos debidos a la digestión. Mediante esta validación, hemos demostrado claramente que el protocolo de digestión en gel separador genera un aumento en la varianza de la cuantificación de las proteínas, debido a que cada fracción se digiere y se marca por separado. Esta manipulación de la muestra introduce una serie de efectos sistemáticos que son imposibles de controlar exactamente en la misma medida en todas las fracciones del gel. Incluso cuando el gel de electroforesis se corta en zonas donde las proteínas no son detectadas por tinción, es muy difícil, sino imposible, cortar todas las proteínas en las fracciones con la misma proporción exactamente en las dos muestras, especialmente en el caso de las proteínas que se localicen en el punto de corte entre dos trozos de gel adyacentes. Esto hace poco fiable y reproducible la cuantificación de una proteína situada en fracciones adyacentes del gel. Nuestros datos revelan por primera vez el efecto de "corte de gel" asociado a la separación de proteínas en 1D-SDS-PAGE, lo que demuestra los problemas encontrados cuando se utiliza este protocolo en experimentos donde se lleva a cabo una cuantificación mediante marcaje isotópico post-digestión. Por otro lado, mejorando las condiciones de la reacción enzimática en el nuevo método de digestión en gel concentrador, el rendimiento de la digestión es mayor y más reproducible que en los otros métodos probados. Incluso se recuperaron péptidos de mayor tamaño al añadir el detergente CYMAL a la mezcla de reacción. Este detergente ha sido descrito que mejora la recuperación de grandes péptidos hidrofóbicos, que pueden quedarse atrapados en la matriz del gel [182, 183].

La idoneidad del método de digestión en gel concentrador descrito en este trabajo para la identificación de proteínas a gran escala se pudo comprobar mediante el análisis estadístico de la precisión con que se cuantifican las proteínas usando el método de marcaje con  $^{18}\text{O}$ . Se han analizado un gran número de proteomas de distinta procedencia y el método también se ha aplicado al estudio de los cambios de abundancia de proteínas en muestras tratadas y muestras control mezcladas en diferentes proporciones relativas, demostrando que la varianza a nivel de espectro es similar en todos ellos.

Las dos fuentes principales de error a nivel de péptido son las digestiones parciales y la oxidación de metioninas, que pueden ocurrir en diversos grados en las dos muestras a comparar, haciendo que la cuantificación de las proteínas sea poco fiable. La oxidación de metioninas es un proceso que ocurre al azar, generando potencialmente unos valores de cuantificación diferentes para un mismo péptido en diferentes estados de oxidación. La variabilidad debida a la digestión

parcial residual puede también introducir cambios de expresión artefactuales porque si la digestión no tiene lugar exactamente con la misma eficiencia en las dos muestras los péptidos generados no reflejan de forma fidedigna la concentración de la proteína de la que provienen. En este trabajo, y en conjunto con la tesis doctoral de Pablo Martínez Acedo, demostramos que estas fuentes potenciales de error no aumentan la variabilidad en la cuantificación de péptidos individuales cuando se utiliza el nuevo método desarrollado durante este trabajo, lo que hace prácticamente idénticas las varianzas a nivel de péptido para todos los proteomas analizados.

A nivel de proteína, la fuente de varianza está relacionada con el método usado en la preparación de las muestras y con la extracción de proteínas y no depende de la técnica de espectrometría de masas usada para la cuantificación de proteínas. Las varianzas a nivel de proteína son similares entre todos los proteomas analizados y próximas a cero, lo que indica que no hay sesgos introducidos durante la preparación de las muestras de cultivos de células o el fraccionamiento subcelular. Finalmente, la distribución de las proteínas cuantificadas, normalizada en función de su varianza estimada, sigue una función gaussiana con una desviación estándar de uno. Este resultado indica que la integración de las varianzas de los tres niveles de cuantificación (espectro, péptido y proteína) determina correctamente la varianza final de cada proteína, de acuerdo con el modelo estadístico empleado.

Este método de digestión triptica de proteomas puesto a punto en este trabajo de tesis constituye la tecnología proteómica de identificación sistemática y a gran escala con la validación estadística más estricta publicada hasta la fecha. Esta estrategia experimental se ha aplicado a varios modelos biológicos [148], proporcionando una varianzas de referencia para futuros análisis proteómicos de cuantificación a gran escala que permitirían detectar fuentes de error. Además, hemos demostrado la utilidad de este nuevo protocolo mediante el estudio de los cambios relativos de abundancia en el proteoma de células T en su proceso de activación.

Todo ello hace que esta estrategia experimental sea también la adecuada para el estudio sistemático y en profundidad de preparaciones basadas en la técnica de *pull-down*, donde las proteínas interaccionan diferencialmente con distintos ligandos peptídicos, permitiendo detectar de forma fiable las interacciones específicas entre un extenso fondo de interacciones inespecíficas, que se detecta de forma reproducible en todas las muestras.

## 2. CARACTERIZACIÓN DEL INTERACTOMA INTRACELULAR DE LOS DOMINIOS RICOS EN TETRASPANINAS

Los microdominios ricos en tetraspaninas o TEMs juegan un papel fundamental en multitud de procesos biológicos que implican la adhesión a la matriz extracelular y a otras células. La inserción en estos dominios de membrana especializados asegura un correcto nivel de agregación y agrupamiento de los receptores de adhesión asociados, necesario para que tenga lugar la migración y adhesión celular [2]. Sin embargo, existen evidencias que sugieren que los componentes de los TEMs también tienen un papel activo en la conexión de estos receptores específicos con vías de señalización, así como para el enlace de las proteínas con el citoesqueleto. El enfoque basado en el análisis a gran escala mediante LC-MS/MS desarrollado en este trabajo nos ha permitido llevar a cabo una exploración sistemática de las conexiones intracelulares del extremo C-terminal de varias tetraspaninas y algunas de sus proteínas asociadas en linfoblastos humanos y en sus exosomas secretados.

Las regiones intracelulares de las tetraspaninas, tanto en el extremo N-terminal como en el C-terminal, suelen ser bastante cortas y no presentan regiones conservadas entre los diferentes miembros de las distintas superfamilias. Las regiones C-terminales de algunas tetraspaninas tienen carga neta básica [92], otras muestran un motivo de unión PDZ [94, 184], mientras que otras presentan secuencias de internalización clásica [10, 185]. En el caso de los receptores asociados a tetraspaninas como EWI-2 [92], ICAM-1 y VCAM-1 [63]), interactúan con proteínas ERM y con proteínas que se asocian con actina [92]. Dado que la proteína ICAM-3 también se puede asociar al ERM [149], en este estudio se incluyó ICAM-3 como control interno de la especificidad de las interacciones. También se sabe que la tetraspanina CD81, se une directamente a ERMs [92]. La naturaleza de las interacciones de las proteínas pertenecientes a TEM y sus proteínas asociadas en los dominios C-terminales no pueden atribuirse a meras asociaciones por interacciones electrostáticas o asociaciones que se produjeran exclusivamente por medio de las ERM, ya que en nuestros resultados no vimos prácticamente ninguna interacción con nuestro anzuelo control que fue ICAM-3, mientras que con los anzuelos arriba mencionados se detectaron interacciones con proteínas de diferente naturaleza. Por otro lado, el diseño experimental utilizado en este trabajo no nos permite determinar si las interacciones detectadas mediante una aproximación proteómica se deben a contactos directos proteína-proteína o están mediadas a través de moléculas intermedias. Por ejemplo, se sabe que las proteínas *α-actinin* y *filamin* colocalizan con factores de elongación y con polisomas en los puntos de ramificación de los filamentos de actina [63]. Por tanto, estas proteínas podrían actuar de intermediarias en la interacción entre ICAM-1, EWI-2 y CD81 y las

proteínas de unión de ácidos nucleicos. También podrían mediar en la interacción con las proteínas de señalización tales como *protein RhoA*, y otras proteínas de la familia *ras* [68]. Esta red de conexiones entre proteínas explicaría la presencia focalizada de los dominios TEM en las zonas donde se produce el contacto entre el leucocito y el endotelio vascular en el proceso de extravasación leucocitaria [63, 65]. Es importante reseñar aquí que algunas de estas interacciones han sido ya descritas mediante técnicas de coinmunoprecipitación entre moléculas endógenas [61, 64, 186], reforzando la especificidad de los resultados obtenidos utilizando nuestra aproximación *in vitro*. También hemos demostrado interacciones con proteínas no descritas previamente, como es el caso de la *nucleolin*, que interacciona con ICAM-1 y EWI-2. Esta proteína está íntimamente relacionada con proteínas ribosomales e histonas [187, 188], que también hemos visto como parte del interactoma de ICAM-1 y EWI-2. Esta conexión sugiere que el papel de los TEMs en el reclutamiento de histonas y proteínas ribosomales podría conducir a estas proteínas hacia el lugar donde ejercerán su función.

La conexión entre los microdominios enriquecidos en tetraspaninas con las proteínas de unión a ARN tiene una potencial relevancia en el proceso de migración y motilidad celular. Se ha visto que los focos iniciales de adhesión incorporan ARN y proteínas de unión a ARN para formar el llamado Centro de Iniciación de Motilidad [170]. Esta localización es compatible con el papel de las tetraspaninas en las etapas tempranas de adhesión [189]. Además, el acoplamiento con integrinas y la tensión mecánica en la membrana plasmática produce el reclutamiento de ribosomas hacia sitios de adhesión focal [169]. La adhesión a la matriz extracelular regula la transcripción [190] y se ha visto que algunos factores de iniciación de la transcripción interaccionan directamente con integrinas [191]. Por otra parte, esta conexión con proteínas que unen ácidos nucleicos podría ser relevante para el reclutamiento de ARN mensajero y ARN microsatélite hacia el interior de microvesículas secretadas como los exosomas [70, 192].

### 3. EL INTERACTOMA DE LAS TETRASPANINAS EN EXOSOMAS Y SU PAPEL EN LA COMPOSICIÓN PROTEICA

Está bien descrito en la literatura que las tetraspaninas forman parte de los cuerpos multivesiculares y de los exosomas [150, 193]. En este trabajo, el análisis proteómico de alto rendimiento confirmó que los exosomas contienen una gran número de tetraspaninas y de receptores asociados a tetraspaninas, incluyendo moléculas MHC, inmunoglobulinas (EWI-2), moléculas de adhesión (ICAM-1), integrinas (*Integrin beta-7*, *Integrin beta-1*, *integrin alpha M*),

anexinas (*annexin A1, A4, A7*), proteínas de la familia, proteínas del metabolismo (*peroxiredoxin, glyceraldehyde-3-phosphate dehydrogenase*), y otros ligandos asociados a tetraspaninas como como CD44 [194], CD98 [195] y ADAM-10 [196, 197]. Además, el interactoma intracelular de CD81 y EWI-2 abarca casi la mitad del contenido proteico del exosoma, hecho que plantea la hipótesis de que estas plataformas de adhesión actúen seleccionando el repertorio de proteínas que son dirigidas a los exosomas.

El papel de las tetraspaninas en la formación de los exosomas es poco conocido. Se sabe que la liberación de exosomas en células dendríticas CD9 *knock-out* es deficiente [73], basándose en mediciones de Western-blot de la proteína *flotillin*. Nuestros datos sugieren que la delección de CD9 dificulta la inclusión de un grupo concreto de proteínas en los exosomas. Este repertorio incluye algunas proteínas transmembrana asociadas a CD9, tales como las moléculas MHC, CD5 y CD147. Por otra parte, la ausencia de CD9 también disminuye la abundancia de algunas proteínas de señalización celular, como la proteína 14-3-3, que también se asocia a CD81. CD81 y CD9 se asocian comúnmente [93], lo que sugiere que CD9 puede jugar un papel en la interacción entre CD81 y la proteína 14-3-3, que regula la función de muchas isoformas de la *protein kinase C* (PKC), así como con *phosphatidylinositol 4,5-bisphosphate 3-kinase* (PI3K) [198]. Se ha visto en otros estudios que las tetraspaninas CD9, CD53, CD81, CD82 y CD151 se asocian con la forma activada de PKC [199]. También hay estudios que asocian la integrina  $\alpha 3\beta 1$  con CD81 y con PI4K [200]. Una asociación controlada entre CD81 y la proteína 14-3-3 podría proporcionar un mecanismo para la regulación de la actividad de PKC por CD81, lo que a su vez permitiría modular la función de la integrina.

A pesar de que la cantidad de CD81 es normalmente mayor que la de CD9 en exosomas, no hemos detectado graves defectos en la secreción de exosomas o en la morfología de los cuerpos multivesiculares en linfocitos con deficiencia de CD81. Por otra parte, aunque sí se observa una menor expresión de *flotillin*, los exosomas deficientes en CD81 no mostraron una reducción en el contenido total de proteínas o en la expresión de otras proteínas abundantes en exosomas como las proteínas ERM. Recientemente se ha publicado que la tetraspanina CD63 es fundamental para la biogénesis de los orgánulos relacionados con derivados de las vesículas intraluminales (melanosomas) [201]. Sin embargo, las rutas de internalización y la cinética de incorporación a exosomas pueden diferir entre las distintas tetraspaninas y sus moléculas asociadas [202, 203].

En cuanto a la selección de las moléculas que componen los exosomas, se ha descrito que el reclutamiento de CD9P-1 (un ligando directo de las tetraspaninas CD9 y CD81) a exosomas se produce en menor medida después del silenciamiento de las tetraspaninas CD9 y CD81 [204]. Otro caso similar se observa con la presencia de *E-cadherin* y  *$\beta$ -catenin* en la fracción exosomal, que

depende de su asociación con la tetraspanina CD82 [73]. Es interesante constatar aquí que la pauta de proteínas que disminuyen de expresión a causa de la depleción de CD9 y de CD81, que incluye las proteínas relacionadas con la familia *ras*, algunas proteínas ribosomales y proteínas de MHC, muestra una gran semejanza entre sí. Esto refuerza la idea de que CD9 y CD81 están relacionadas [204], y sugiere que estas tetraspaninas se asocian para ejecutar muchas de sus funciones que son posiblemente mediadas por rutas de señalización muy similares.

A pesar del alto grado de redundancia funcional entre las tetraspaninas y del gran nivel de solapamiento entre el interactoma intracelular de CD81 y el de sus proteínas asociadas EWI-2 e ICAM-1, nuestros datos sugieren que la supresión génica de CD81 perturba la inclusión de un repertorio selectivo de proteínas en los exosomas. Este repertorio incluye algunas proteínas transmembrana específicamente asociadas a CD81, tales como las moléculas MHC, CD20, BCR e ICAM-1. Por otra parte, la delección de CD81 disminuye la proporción de *rac2* en exosomas, cuya inserción en los microdominios ricos en tetraspaninas está específicamente dirigida por la asociación con CD81.

#### 4. CONSIDERACIONES FINALES

En este trabajo se ha desarrollado un nuevo método experimental basado en la concentración de proteínas en SDS-PAGE, que por medio de la digestión tróptica en gel en una sola fracción y el marcaje de los péptidos por  $^{18}\text{O}$ , junto con el algoritmo computacional desarrollado previamente en el laboratorio [117], podría convertirse en un método universal para llevar a cabo estudios comparativos a gran escala mediante espectrometría de masas en trampa iónica lineal. Con estos avances técnicos es ahora posible explotar de forma sistemática las ventajas de la identificación y la cuantificación masiva de proteomas, cualquiera que sea su procedencia (tejido, cultivo celular, fluidos biológicos, etc.) y organismo, dada la aplicabilidad universal del método. La nueva estrategia, además, no produce reacciones químicas secundarias, es muy reproducible y está validada estadísticamente. Aún quedaría por analizar si el método de digestión en gel concentrador es superior a las nuevas metodologías basadas en digestión en solución con filtros (FASP), que podrían minimizar el coste de material y el tiempo de procesamiento de muestra, pero cuya aplicación a la cuantificación de proteomas usando técnicas de marcaje isotópico postdigestión no ha sido demostrada todavía.

Gracias a estos avances hemos podido llevar a cabo el análisis sistemático de proteomas que se han seleccionado por afinidad con varios anzuelos peptídicos de proteínas relacionadas con TEMs.

Este trabajo ha permitido caracterizar la red de interacciones intracelulares de la que forman parte CD81, EWI-2 e ICAM-1. Los resultados derivados de los análisis del interactoma de dichas proteínas sugieren que los TEMs, además de su papel en la adhesión celular, actúan como plataformas especializadas en la membrana plasmática para la compartimentación selectiva de los receptores que van a ser secretados y su correcta exposición en la superficie de exosomas, amplificando así su funcionalidad, tal y como se ha demostrado en las funciones inmunes como la presentación de antígeno [70]. Sin embargo, aún no se ha definido el papel biológico y funcional de la mayoría de las proteínas que interactúan con los dominios TEMs. La definición biológica de todas estas interacciones aportaría información sobre las diferentes rutas de señalización o medios de acción que los TEMs utilizan para ejecutar sus funciones. Esto ampliaría enormemente el conocimiento de las actividades que llevan a cabo estos dominios. Además contribuiría a dilucidar el papel que desempeñan los exosomas en la comunicación celular, que aún queda mucho por describir.

Por último, la interacción de los TEMs con las proteínas relacionadas con ácidos nucleicos abre la posibilidad de que estos microdominios sean factores determinantes en la selección de ARNm y ARNm<sub>i</sub> en estas vesículas. Gracias a la metodología puesta a punto en este trabajo, se pueden determinar las interacciones de los ARNm<sub>i</sub> con proteínas de exosomas y conocer en detalle qué proteínas reclutan estos ARNm<sub>i</sub>, que permitiría conocer en profundidad el mecanismo de selección de ARNm<sub>i</sub> a exosomas. La inserción de proteínas dirigidas por tetraspaninas en exosomas podría tener un papel en el diseño de terapias que guíen fármacos hacia estos nanotransportadores naturales.



## **CONCLUSIONES**



1. La tecnología de preparación de muestras y marcaje isotópico desarrollada en nuestro laboratorio permite la caracterización a gran escala de proteínas interaccionantes y la cuantificación diferencial sistemática de proteomas.
2. Los dominios intracelulares de tetraspaninas forman redes coherentes de interacciones en linfocitos T humanos, conectando distintos grupos funcionales de proteínas, incluyendo proteínas ribosomales, histonas, proteínas del citoesqueleto, proteínas relacionadas con ácidos nucleicos, chaperonas, proteínas del metabolismo, proteínas de señalización celular y proteínas de la familia *ras*.
3. El interactoma intracelular de las tetraspaninas CD81 y EWI-2 en exosomas de linfocitos T humanos comprende, además de los grupos descritos en el punto anterior, nuevos grupos funcionales de proteínas relacionadas con el tráfico celular y proteínas y receptores de membrana.
4. Las proteínas pertenecientes al interactoma intracelular de tetraspaninas constituyen un 45% de la composición proteica de los exosomas de linfocitos T humanos.
5. La tetraspanina CD81 interacciona selectivamente con proteínas relacionadas con la vía Ras. La depleción de esta tetraspanina disminuye la presencia de estas proteínas en exosomas de linfocitos T de ratón.
6. Nuestros resultados permiten plantear la hipótesis de que las tetraspaninas desempeñan un papel como plataformas especializadas de la membrana que estabilizan el proteoma de los exosomas.



## **BIBLIOGRAFÍA**



1. Yanez-Mo, M., et al., *Tetraspanin-enriched microdomains: a functional unit in cell plasma membranes*. Trends Cell Biol, 2009. **19**(9): p. 434-46.
2. Hemler, M.E., *Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain*. Annu Rev Cell Dev Biol, 2003. **19**: p. 397-422.
3. Barreiro, O., et al., *Endothelial tetraspanin microdomains regulate leukocyte firm adhesion during extravasation*. Blood, 2005. **105**(7): p. 2852-61.
4. Longo, N., et al., *Regulatory role of tetraspanin CD9 in tumor-endothelial cell interaction during transendothelial invasion of melanoma cells*. Blood, 2001. **98**(13): p. 3717-26.
5. Pileri, P., et al., *Binding of hepatitis C virus to CD81*. Science, 1998. **282**(5390): p. 938-41.
6. Gordon-Alonso, M., et al., *Tetraspanins CD9 and CD81 modulate HIV-1-induced membrane fusion*. J Immunol, 2006. **177**(8): p. 5129-37.
7. Spoden, G., et al., *Clathrin- and caveolin-independent entry of human papillomavirus type 16--involvement of tetraspanin-enriched microdomains (TEMs)*. PLoS One, 2008. **3**(10): p. e3313.
8. Aird, W.C., *Endothelium in health and disease*. Pharmacol Rep, 2008. **60**(1): p. 139-43.
9. Charrin, S., et al., *Lateral organization of membrane proteins: tetraspanins spin their web*. Biochem J, 2009. **420**(2): p. 133-54.
10. Stipp, C.S., T.V. Kolesnikova, and M.E. Hemler, *Functional domains in tetraspanin proteins*. Trends Biochem Sci, 2003. **28**(2): p. 106-12.
11. Szollosi, J., et al., *Supramolecular complexes of MHC class I, MHC class II, CD20, and tetraspan molecules (CD53, CD81, and CD82) at the surface of a B cell line JY*. J Immunol, 1996. **157**(7): p. 2939-46.
12. Schick, M.R. and S. Levy, *The TAPA-1 molecule is associated on the surface of B cells with HLA-DR molecules*. J Immunol, 1993. **151**(8): p. 4090-7.
13. Yauch, R.L., et al., *Direct extracellular contact between integrin alpha(3)beta(1) and TM4SF protein CD151*. J Biol Chem, 2000. **275**(13): p. 9230-8.
14. Chambrion, C. and F. Le Naour, *The tetraspanins CD9 and CD81 regulate CD9P1-induced effects on cell migration*. PLoS One, 2010. **5**(6): p. e11219.
15. Charrin, S., et al., *A physical and functional link between cholesterol and tetraspanins*. Eur J Immunol, 2003. **33**(9): p. 2479-89.
16. Mitsuzuka, K., *A Specific Microdomain ("Glycosynapse 3") Controls Phenotypic Conversion and Reversion of Bladder Cancer Cells through GM3-mediated Interaction of  $\alpha 3 \beta 1$  Integrin with CD9*, in *Journal of Biological Chemistry*. 2005. p. 35545-35553.
17. Regina Todeschini, A. and S.I. Hakomori, *Functional role of glycosphingolipids and gangliosides in control of cell adhesion, motility, and growth, through glycosynaptic microdomains*. Biochim Biophys Acta, 2008. **1780**(3): p. 421-33.
18. Miyake, M., et al., *Identification of the motility-related protein (MRP-1), recognized by monoclonal antibody M31-15, which inhibits cell motility*. J Exp Med, 1991. **174**(6): p. 1347-54.
19. Berditchevski, F., *Complexes of tetraspanins with integrins: more than meets the eye*. J Cell Sci, 2001. **114**(Pt 23): p. 4143-51.
20. Murayama, Y., et al., *The tetraspanin CD9 modulates epidermal growth factor receptor signaling in cancer cells*. J Cell Physiol, 2008. **216**(1): p. 135-43.
21. Castro-Sanchez, L., et al., *Native type IV collagen induces cell migration through a CD9 and DDR1-dependent pathway in MDA-MB-231 breast cancer cells*. Eur J Cell Biol, 2010. **89**(11): p. 843-52.
22. Jones, P.H., L.A. Bishop, and F.M. Watt, *Functional significance of CD9 association with beta 1 integrins in human epidermal keratinocytes*. Cell Adhes Commun, 1996. **4**(4-5): p. 297-305.

23. Huang, W., M. Febbraio, and R.L. Silverstein, *CD9 tetraspanin interacts with CD36 on the surface of macrophages: a possible regulatory influence on uptake of oxidized low density lipoprotein*. PLoS One, 2011. **6**(12): p. e29092.
24. Miao, W.M., et al., *CD36 associates with CD9 and integrins on human blood platelets*. Blood, 2001. **97**(6): p. 1689-96.
25. Kovalenko, O.V., X.H. Yang, and M.E. Hemler, *A novel cysteine cross-linking method reveals a direct association between claudin-1 and tetraspanin CD9*. Mol Cell Proteomics, 2007. **6**(11): p. 1855-67.
26. Kolesnikova, T.V., et al., *EWI-2 modulates lymphocyte integrin alpha4beta1 functions*. Blood, 2004. **103**(8): p. 3013-9.
27. Longhurst, C.M., et al., *Chinese hamster ovary cell motility to fibronectin is modulated by the second extracellular loop of CD9. Identification of a putative fibronectin binding site*. J Biol Chem, 2002. **277**(36): p. 32445-52.
28. Oren, R., et al., *TAPA-1, the target of an antiproliferative antibody, defines a new family of transmembrane proteins*. Mol Cell Biol, 1990. **10**(8): p. 4007-15.
29. Levy, S., S.C. Todd, and H.T. Maecker, *CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system*. Annu Rev Immunol, 1998. **16**: p. 89-109.
30. Mittelbrunn, M., et al., *Cutting edge: dynamic redistribution of tetraspanin CD81 at the central zone of the immune synapse in both T lymphocytes and APC*. J Immunol, 2002. **169**(12): p. 6691-5.
31. Bradbury, L.E., et al., *The CD19/CD21 signal transducing complex of human B lymphocytes includes the target of antiproliferative antibody-1 and Leu-13 molecules*. J Immunol, 1992. **149**(9): p. 2841-50.
32. Feigelson, S.W., et al., *The CD81 tetraspanin facilitates instantaneous leukocyte VLA-4 adhesion strengthening to vascular cell adhesion molecule 1 (VCAM-1) under shear flow*. J Biol Chem, 2003. **278**(51): p. 51203-12.
33. Witherden, D.A., R. Boismenu, and W.L. Havran, *CD81 and CD28 costimulate T cells through distinct pathways*. J Immunol, 2000. **165**(4): p. 1902-9.
34. Sagi, Y., et al., *Complementary costimulation of human T-cell subpopulations by cluster of differentiation 28 (CD28) and CD81*. Proc Natl Acad Sci U S A, 2012. **109**(5): p. 1613-8.
35. Maecker, H.T., et al., *Differential expression of murine CD81 highlighted by new anti-mouse CD81 monoclonal antibodies*. Hybridoma, 2000. **19**(1): p. 15-22.
36. Maecker, H.T., M.S. Do, and S. Levy, *CD81 on B cells promotes interleukin 4 secretion and antibody production during T helper type 2 immune responses*. Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2458-62.
37. Deng, J., et al., *Allergen-induced airway hyperreactivity is diminished in CD81-deficient mice*. J Immunol, 2000. **165**(9): p. 5054-61.
38. Charrin, S., et al., *EWI-2 is a new component of the tetraspanin web in hepatocytes and lymphoid cells*. Biochem J, 2003. **373**(Pt 2): p. 409-21.
39. Charrin, S., et al., *The major CD9 and CD81 molecular partner. Identification and characterization of the complexes*. J Biol Chem, 2001. **276**(17): p. 14329-37.
40. Clark, K.L., et al., *PGRL is a major CD81-associated protein on lymphocytes and distinguishes a new family of cell surface proteins*. J Immunol, 2001. **167**(9): p. 5115-21.
41. Stipp, C.S., T.V. Kolesnikova, and M.E. Hemler, *EWI-2 is a major CD9 and CD81 partner and member of a novel Ig protein subfamily*. J Biol Chem, 2001. **276**(44): p. 40545-54.
42. Stipp, C.S., D. Orlicky, and M.E. Hemler, *FPRP, a major, highly stoichiometric, highly specific CD81- and CD9-associated protein*. J Biol Chem, 2001. **276**(7): p. 4853-62.



43. Zhang, X.A., et al., *EWI2/PGRL associates with the metastasis suppressor KAI1/CD82 and inhibits the migration of prostate cancer cells*. *Cancer Res*, 2003. **63**(10): p. 2665-74.
44. Boucheix, C., et al., *Tetraspanins and malignancy*. *Expert Rev Mol Med*, 2001. **2001**: p. 1-17.
45. Hemler, M.E., *Tetraspanin functions and associated microdomains*. *Nat Rev Mol Cell Biol*, 2005. **6**(10): p. 801-11.
46. Levy, S. and T. Shoham, *The tetraspanin web modulates immune-signalling complexes*. *Nat Rev Immunol*, 2005. **5**(2): p. 136-48.
47. Kettner, S., et al., *EWI-2/CD316 is an inducible receptor of HSPA8 on human dendritic cells*. *Mol Cell Biol*, 2007. **27**(21): p. 7718-26.
48. Stipp, C.S., T.V. Kolesnikova, and M.E. Hemler, *EWI-2 regulates alpha3beta1 integrin-dependent cell functions on laminin-5*. *J Cell Biol*, 2003. **163**(5): p. 1167-77.
49. van de Stolpe, A. and P.T. van der Saag, *Intercellular adhesion molecule-1*. *J Mol Med (Berl)*, 1996. **74**(1): p. 13-33.
50. Stassen, F.R., T. Vainas, and C.A. Bruggeman, *Infection and atherosclerosis. An alternative view on an outdated hypothesis*. *Pharmacol Rep*, 2008. **60**(1): p. 85-92.
51. Marlin, S.D. and T.A. Springer, *Purified intercellular adhesion molecule-1 (ICAM-1) is a ligand for lymphocyte function-associated antigen 1 (LFA-1)*. *Cell*, 1987. **51**(5): p. 813-9.
52. Languino, L.R., et al., *Fibrinogen mediates leukocyte adhesion to vascular endothelium through an ICAM-1-dependent pathway*. *Cell*, 1993. **73**(7): p. 1423-34.
53. Greve, J.M., et al., *The major human rhinovirus receptor is ICAM-1*. *Cell*, 1989. **56**(5): p. 839-47.
54. Miyasaka, M. and T. Tanaka, *Lymphocyte trafficking across high endothelial venules: dogmas and enigmas*. *Nat Rev Immunol*, 2004. **4**(5): p. 360-70.
55. Hosseini, B.H., et al., *Immune synapse formation determines interaction forces between T cells and antigen-presenting cells measured by atomic force microscopy*. *Proc Natl Acad Sci U S A*, 2009. **106**(42): p. 17852-7.
56. Barreiro, O., et al., *Endothelial adhesion receptors are recruited to adherent leukocytes by inclusion in preformed tetraspanin nanoplateforms*. *J Cell Biol*, 2008. **183**(3): p. 527-42.
57. Durieu-Trautmann, O., et al., *Intercellular adhesion molecule 1 activation induces tyrosine phosphorylation of the cytoskeleton-associated protein cortactin in brain microvessel endothelial cells*. *J Biol Chem*, 1994. **269**(17): p. 12536-40.
58. van Buul, J.D., E. Kanters, and P.L. Hordijk, *Endothelial signaling by Ig-like cell adhesion molecules*. *Arterioscler Thromb Vasc Biol*, 2007. **27**(9): p. 1870-6.
59. Ley, K., et al., *Getting to the site of inflammation: the leukocyte adhesion cascade updated*. *Nat Rev Immunol*, 2007. **7**(9): p. 678-89.
60. Orlova, V.V. and T. Chavakis, *Regulation of vascular endothelial permeability by junctional adhesion molecules (JAM)*. *Thromb Haemost*, 2007. **98**(2): p. 327-32.
61. Carpen, O., et al., *Association of intercellular adhesion molecule-1 (ICAM-1) with actin-containing cytoskeleton and alpha-actinin*. *J Cell Biol*, 1992. **118**(5): p. 1223-34.
62. Celli, L., et al., *Evidence of a functional role for interaction between ICAM-1 and nonmuscle alpha-actinins in leukocyte diapedesis*. *J Immunol*, 2006. **177**(6): p. 4113-21.
63. Barreiro, O., et al., *Dynamic interaction of VCAM-1 and ICAM-1 with moesin and ezrin in a novel endothelial docking structure for adherent leukocytes*. *J Cell Biol*, 2002. **157**(7): p. 1233-45.
64. Kanters, E., et al., *Filamin B mediates ICAM-1-driven leukocyte transendothelial migration*. *J Biol Chem*, 2008. **283**(46): p. 31830-9.
65. van Buul, J.D., et al., *RhoG regulates endothelial apical cup assembly downstream from ICAM1 engagement and is involved in leukocyte trans-endothelial migration*. *J Cell Biol*, 2007. **178**(7): p. 1279-93.

66. Alcaide, P., S. Auerbach, and F.W. Lusinskas, *Neutrophil recruitment under shear flow: it's all about endothelial cell rings and gaps*. *Microcirculation*, 2009. **16**(1): p. 43-57.
67. Ellerbroek, S.M., et al., *SGEF, a RhoG guanine nucleotide exchange factor that stimulates macropinocytosis*. *Mol Biol Cell*, 2004. **15**(7): p. 3309-19.
68. Bellanger, J.M., et al., *The Rac1- and RhoG-specific GEF domain of Trio targets filamin to remodel cytoskeletal actin*. *Nat Cell Biol*, 2000. **2**(12): p. 888-92.
69. Silvie, O., et al., *Hepatocyte CD81 is required for Plasmodium falciparum and Plasmodium yoelii sporozoite infectivity*. *Nat Med*, 2003. **9**(1): p. 93-6.
70. Thery, C., M. Ostrowski, and E. Segura, *Membrane vesicles as conveyors of immune responses*. *Nat Rev Immunol*, 2009. **9**(8): p. 581-93.
71. van Niel, G., et al., *Exosomes: a common pathway for a specialized function*. *J Biochem*, 2006. **140**(1): p. 13-21.
72. Simpson, R.J., et al., *Exosomes: proteomic insights and diagnostic potential*. *Expert Rev Proteomics*, 2009. **6**(3): p. 267-83.
73. Chairoungdua, A., et al., *Exosome release of beta-catenin: a novel mechanism that antagonizes Wnt signaling*. *J Cell Biol*, 2010. **190**(6): p. 1079-91.
74. Nazarenko, I., et al., *Cell surface tetraspanin Tspan8 contributes to molecular pathways of exosome-induced endothelial cell activation*. *Cancer Res*, 2010. **70**(4): p. 1668-78.
75. Skog, J., et al., *Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers*. *Nat Cell Biol*, 2008. **10**(12): p. 1470-6.
76. Balaj, L., et al., *Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences*. *Nat Commun*, 2011. **2**: p. 180.
77. Deregibus, M.C., et al., *Endothelial progenitor cell derived microvesicles activate an angiogenic program in endothelial cells by a horizontal transfer of mRNA*. *Blood*, 2007. **110**(7): p. 2440-8.
78. Ratajczak, J., et al., *Embryonic stem cell-derived microvesicles reprogram hematopoietic progenitors: evidence for horizontal transfer of mRNA and protein delivery*. *Leukemia*, 2006. **20**(5): p. 847-56.
79. Zernecke, A., et al., *Delivery of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection*. *Sci Signal*, 2009. **2**(100): p. ra81.
80. Pegtel, D.M., et al., *Functional delivery of viral miRNAs via exosomes*. *Proc Natl Acad Sci U S A*, 2010. **107**(14): p. 6328-33.
81. Corbeel, L. and K. Freson, *Rab proteins and Rab-associated proteins: major actors in the mechanism of protein-trafficking disorders*. *Eur J Pediatr*, 2008. **167**(7): p. 723-9.
82. Stenmark, H., *Rab GTPases as coordinators of vesicle traffic*. *Nat Rev Mol Cell Biol*, 2009. **10**(8): p. 513-25.
83. Ji, H., et al., *Difference gel electrophoresis analysis of Ras-transformed fibroblast cell-derived exosomes*. *Electrophoresis*, 2008. **29**(12): p. 2660-71.
84. Hurley, J.H., *ESCRT complexes and the biogenesis of multivesicular bodies*. *Curr Opin Cell Biol*, 2008. **20**(1): p. 4-11.
85. Futter, C.E. and I.J. White, *Annexins and endocytosis*. *Traffic*, 2007. **8**(8): p. 951-8.
86. Denzer, K., et al., *Follicular dendritic cells carry MHC class II-expressing microvesicles at their surface*. *J Immunol*, 2000. **165**(3): p. 1259-65.
87. Mathivanan, S. and R.J. Simpson, *ExoCarta: A compendium of exosomal proteins and RNA*. *Proteomics*, 2009. **9**(21): p. 4997-5000.
88. Mathivanan, S., et al., *Proteomics analysis of A33 immunoaffinity-purified exosomes released from the human colon tumor cell line LIM1215 reveals a tissue-specific protein signature*. *Mol Cell Proteomics*, 2010. **9**(2): p. 197-208.

89. Mears, R., et al., *Proteomic analysis of melanoma-derived exosomes by two-dimensional polyacrylamide gel electrophoresis and mass spectrometry*. *Proteomics*, 2004. **4**(12): p. 4019-31.
90. Raposo, G., et al., *Accumulation of major histocompatibility complex class II molecules in mast cell secretory granules and their release upon degranulation*. *Mol Biol Cell*, 1997. **8**(12): p. 2631-45.
91. Raiborg, C., et al., *FYVE and coiled-coil domains determine the specific localisation of Hrs to early endosomes*. *J Cell Sci*, 2001. **114**(Pt 12): p. 2255-63.
92. Sala-Valdes, M., et al., *EWI-2 and EWI-F Link the Tetraspanin Web to the Actin Cytoskeleton through Their Direct Association with Ezrin-Radixin-Moesin Proteins*. *J Biol Chem*, 2006. **281**(28): p. 19665-75.
93. Clark, K.L., et al., *CD81 associates with 14-3-3 in a redox-regulated palmitoylation-dependent manner*. *J Biol Chem*, 2004. **279**(19): p. 19401-6.
94. Latysheva, N., et al., *Syntenin-1 is a new component of tetraspanin-enriched microdomains: mechanisms and consequences of the interaction of syntenin-1 with CD63*. *Mol Cell Biol*, 2006. **26**(20): p. 7707-18.
95. Wilkins, M.R., et al., *From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis*. *Biotechnology (N Y)*, 1996. **14**(1): p. 61-5.
96. Ho, Y., et al., *Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry*. *Nature*, 2002. **415**(6868): p. 180-3.
97. Skiba, M., T.C. Mettenleiter, and A. Karger, *Quantitative whole-cell proteome analysis of pseudorabies virus-infected cells*. *J Virol*, 2008. **82**(19): p. 9689-99.
98. Henzel, W.J., C. Watanabe, and J.T. Stults, *Protein identification: the origins of peptide mass fingerprinting*. *J Am Soc Mass Spectrom*, 2003. **14**(9): p. 931-42.
99. Mann, M., P. Hojrup, and P. Roepstorff, *Use of mass spectrometric molecular weight information to identify proteins in sequence databases*. *Biol Mass Spectrom*, 1993. **22**(6): p. 338-45.
100. Pandey, A., J.S. Andersen, and M. Mann, *Use of mass spectrometry to study signaling pathways*. *Sci STKE*, 2000. **2000**(37): p. pl1.
101. Jones, A.W. and H.J. Cooper, *Dissociation techniques in mass spectrometry-based proteomics*. *Analyst*, 2011. **136**(17): p. 3419-29.
102. Yates, J.R., 3rd, et al., *Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database*. *Anal Chem*, 1995. **67**(8): p. 1426-36.
103. de Godoy, L.M., et al., *Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast*. *Nature*, 2008. **455**(7217): p. 1251-4.
104. de Godoy, L.M., et al., *Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system*. *Genome Biol*, 2006. **7**(6): p. R50.
105. Wisniewski, J.R., et al., *Universal sample preparation method for proteome analysis*. *Nat Methods*, 2009. **6**(5): p. 359-62.
106. Wright, C.W. and C.S. Duckett, *The aryl hydrocarbon nuclear translocator alters CD30-mediated NF-kappaB-dependent transcription*. *Science*, 2009. **323**(5911): p. 251-5.
107. Markillie, L.M., et al., *Simple protein complex purification and identification method for high-throughput mapping of protein interaction networks*. *J Proteome Res*, 2005. **4**(2): p. 268-74.
108. Wu, F., et al., *Proteome-wide identification of novel binding partners to the oncogenic fusion gene protein, NPM-ALK, using tandem affinity purification and mass spectrometry*. *Am J Pathol*, 2009. **174**(2): p. 361-70.
109. Zada, A.A., et al., *Proteomic discovery of Max as a novel interacting partner of C/EBPalpha: a Myc/Max/Mad link*. *Leukemia*, 2006. **20**(12): p. 2137-46.

110. Terpe, K., *Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems*. Appl Microbiol Biotechnol, 2003. **60**(5): p. 523-33.
111. Jonker, N., et al., *Recent developments in protein-ligand affinity mass spectrometry*. Anal Bioanal Chem, 2011. **399**(8): p. 2669-81.
112. Aye, T.T., et al., *Selectivity in enrichment of cAMP-dependent protein kinase regulatory subunits type I and type II and their interactors using modified cAMP affinity resins*. Mol Cell Proteomics, 2009. **8**(5): p. 1016-28.
113. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics*. Nature, 2003. **422**(6928): p. 198-207.
114. Meier, N., et al., *Novel binding partners of Ldb1 are required for haematopoietic development*. Development, 2006. **133**(24): p. 4913-23.
115. Henkin, J.A., et al., *Mass processing--an improved technique for protein identification with mass spectrometry data*. J Biomol Tech, 2004. **15**(4): p. 230-7.
116. Haun, M. and S. Wasi, *Biotinylated antibodies bound to streptavidin beads: a versatile solid matrix for immunoassays*. Anal Biochem, 1990. **191**(2): p. 337-42.
117. Jorge, I., et al., *Statistical model to analyze quantitative proteomics data obtained by 18O/16O labeling and linear ion trap mass spectrometry: application to the study of vascular endothelial growth factor-induced angiogenesis in endothelial cells*. Mol Cell Proteomics, 2009. **8**(5): p. 1130-49.
118. Gavin, A.C., et al., *Functional organization of the yeast proteome by systematic analysis of protein complexes*. Nature, 2002. **415**(6868): p. 141-7.
119. Krogan, N.J., et al., *Global landscape of protein complexes in the yeast Saccharomyces cerevisiae*. Nature, 2006. **440**(7084): p. 637-43.
120. Goll, J. and P. Uetz, *The elusive yeast interactome*. Genome Biol, 2006. **7**(6): p. 223.
121. Monti, M., et al., *Puzzle of protein complexes in vivo: a present and future challenge for functional proteomics*. Expert Rev Proteomics, 2009. **6**(2): p. 159-69.
122. Rain, J.C., et al., *The protein-protein interaction map of Helicobacter pylori*. Nature, 2001. **409**(6817): p. 211-5.
123. Pennisi, E., *Genomics. New genomes shed light on complex cells*. Science, 2001. **292**(5520): p. 1280-1.
124. Conrads, K.A., et al., *Quantitative proteomic analysis of inorganic phosphate-induced murine MC3T3-E1 osteoblast cells*. Electrophoresis, 2004. **25**(9): p. 1342-52.
125. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*. Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
126. Wasinger, V.C. and G.L. Corthals, *Proteomic tools for biomedicine*. J Chromatogr B Analyt Technol Biomed Life Sci, 2002. **771**(1-2): p. 33-48.
127. Fenselau, C. and X. Yao, *Proteolytic labeling with 18O for comparative proteomics studies: preparation of 18O-labeled peptides and the 18O/16O peptide mixture*. Methods Mol Biol, 2007. **359**: p. 135-42.
128. Chen, X., L.M. Smith, and E.M. Bradbury, *Site-specific mass tagging with stable isotopes in proteins for accurate and efficient protein identification*. Anal Chem, 2000. **72**(6): p. 1134-43.
129. Kruger, M., et al., *SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function*. Cell, 2008. **134**(2): p. 353-64.
130. Hayter, J.R., et al., *The subunit structure and dynamics of the 20S proteasome in chicken skeletal muscle*. Mol Cell Proteomics, 2005. **4**(9): p. 1370-81.
131. Sury, M.D., J.X. Chen, and M. Selbach, *The SILAC fly allows for accurate protein quantification in vivo*. Mol Cell Proteomics, 2010. **9**(10): p. 2173-83.



132. Gevaert, K., et al., *Stable isotopic labeling in proteomics*. Proteomics, 2008. **8**(23-24): p. 4873-85.
133. Han, D.K., et al., *Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry*. Nat Biotechnol, 2001. **19**(10): p. 946-51.
134. Thompson, A., et al., *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS*. Anal Chem, 2003. **75**(8): p. 1895-904.
135. Yao, X., C. Afonso, and C. Fenselau, *Dissection of proteolytic 18O labeling: endoprotease-catalyzed 16O-to-18O exchange of truncated peptide substrates*. J Proteome Res, 2003. **2**(2): p. 147-52.
136. Schnolzer, M., P. Jedrzejewski, and W.D. Lehmann, *Protease-catalyzed incorporation of 18O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry*. Electrophoresis, 1996. **17**(5): p. 945-53.
137. Heller, M., et al., *Trypsin catalyzed 16O-to-18O exchange for comparative proteomics: tandem mass spectrometry comparison using MALDI-TOF, ESI-QTOF, and ESI-ion trap mass spectrometers*. J Am Soc Mass Spectrom, 2003. **14**(7): p. 704-18.
138. Zang, L., et al., *Proteomic analysis of ductal carcinoma of the breast using laser capture microdissection, LC-MS, and 16O/18O isotopic labeling*. J Proteome Res, 2004. **3**(3): p. 604-12.
139. Sakai, J., et al., *18O-labeling quantitative proteomics using an ion trap mass spectrometer*. Proteomics, 2005. **5**(1): p. 16-23.
140. Hicks, W.A., et al., *Simultaneous quantification and identification using 18O labeling with an ion trap mass spectrometer and the analysis software application "ZoomQuant"*. J Am Soc Mass Spectrom, 2005. **16**(6): p. 916-25.
141. Mayya, V., et al., *Systematic comparison of a two-dimensional ion trap and a three-dimensional ion trap mass spectrometer in proteomics*. Mol Cell Proteomics, 2005. **4**(2): p. 214-23.
142. Lopez-Ferrer, D., et al., *Quantitative proteomics using 16O/18O labeling and linear ion trap mass spectrometry*. Proteomics, 2006. **6** Suppl 1: p. S4-11.
143. Jorge, I., et al., *High-sensitivity analysis of specific peptides in complex samples by selected MS/MS ion monitoring and linear ion trap mass spectrometry: application to biological studies*. J Mass Spectrom, 2007. **42**(11): p. 1391-403.
144. Yao, X., et al., *Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus*. Anal Chem, 2001. **73**(13): p. 2836-42.
145. Ramos-Fernandez, A., D. Lopez-Ferrer, and J. Vazquez, *Improved method for differential expression proteomics using trypsin-catalyzed 18O labeling with a correction for labeling efficiency*. Mol Cell Proteomics, 2007. **6**(7): p. 1274-86.
146. Stewart, R., M. Prince, and A. Mann, *Age, vascular risk, and cognitive decline in an older, British, African-Caribbean population*. J Am Geriatr Soc, 2003. **51**(11): p. 1547-53.
147. Staes, A., et al., *Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18*. J Proteome Res, 2004. **3**(4): p. 786-91.
148. Bonzon-Kulichenko, E., et al., *A robust method for quantitative high-throughput analysis of proteomes by 18O labeling*. Mol Cell Proteomics, 2011. **10**(1): p. M110 003335.
149. Serrador, J.M., et al., *Moesin interacts with the cytoplasmic region of intercellular adhesion molecule-3 and is redistributed to the uropod of T lymphocytes during cell polarization*. J Cell Biol, 1997. **138**(6): p. 1409-23.
150. Thery, C., et al., *Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles*. J Immunol, 2001. **166**(12): p. 7309-18.
151. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 1970. **227**(5259): p. 680-5.

152. Lopez-Ferrer, D., et al., *Statistical model for large-scale peptide identification in databases from tandem mass spectra using SEQUEST*. Anal Chem, 2004. **76**(23): p. 6853-60.
153. Navarro, P. and J. Vazquez, *A refined method to calculate false discovery rates for peptide identification using decoy databases*. J Proteome Res, 2009. **8**(4): p. 1792-6.
154. Martinez-Bartolome, S., et al., *Properties of average score distributions of SEQUEST: the probability ratio method*. Mol Cell Proteomics, 2008. **7**(6): p. 1135-45.
155. Kool, J., et al., *Studying protein-protein affinity and immobilized ligand-protein affinity interactions using MS-based methods*. Anal Bioanal Chem, 2011. **401**(4): p. 1109-25.
156. Shevchenko, A., et al., *In-gel digestion for mass spectrometric characterization of proteins and proteomes*. Nat Protoc, 2006. **1**(6): p. 2856-60.
157. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 1976. **72**: p. 248-54.
158. Feske, S., et al., *Gene regulation mediated by calcium signals in T lymphocytes*. Nat Immunol, 2001. **2**(4): p. 316-24.
159. Rongvaux, A., et al., *Pre-B-cell colony-enhancing factor, whose expression is up-regulated in activated lymphocytes, is a nicotinamide phosphoribosyltransferase, a cytosolic enzyme involved in NAD biosynthesis*. Eur J Immunol, 2002. **32**(11): p. 3225-34.
160. Kirchhof, M.G., et al., *Modulation of T cell activation by stomatin-like protein 2*. J Immunol, 2008. **181**(3): p. 1927-36.
161. Sun, H., et al., *Molecular cloning and characterization of a novel muscle adenylosuccinate synthetase, AdSSL1, from human bone marrow stromal cells*. Mol Cell Biochem, 2005. **269**(1-2): p. 85-94.
162. Rodriguez-Tarduchy, G., et al., *Apoptosis but not other activation events is inhibited by a mutation in the transmembrane domain of T cell receptor beta that impairs CD3zeta association*. J Biol Chem, 1996. **271**(48): p. 30417-25.
163. Lopez-Anton, N., et al., *The marine product cephalostatin 1 activates an endoplasmic reticulum stress-specific and apoptosome-independent apoptotic signaling pathway*. J Biol Chem, 2006. **281**(44): p. 33078-86.
164. Singh, S. and A. Khar, *Differential gene expression during apoptosis induced by a serum factor: role of mitochondrial F0-F1 ATP synthase complex*. Apoptosis, 2005. **10**(6): p. 1469-82.
165. Godbole, A., et al., *VDAC is a conserved element of death pathways in plant and animal systems*. Biochim Biophys Acta, 2003. **1642**(1-2): p. 87-96.
166. Roy, S.S., et al., *VDAC2 is required for truncated BID-induced mitochondrial apoptosis by recruiting BAK to the mitochondria*. EMBO Rep, 2009. **10**(12): p. 1341-7.
167. John, G.B., et al., *The mitochondrial inner membrane protein mitofilin controls cristae morphology*. Mol Biol Cell, 2005. **16**(3): p. 1543-54.
168. Solstad, T., et al., *Quantitative proteome analysis of detergent-resistant membranes identifies the differential regulation of protein kinase C isoforms in apoptotic T cells*. Proteomics, 2010. **10**(15): p. 2758-68.
169. Chicurel, M.E., et al., *Integrin binding and mechanical tension induce movement of mRNA and ribosomes to focal adhesions*. Nature, 1998. **392**(6677): p. 730-3.
170. de Hoog, C.L., L.J. Foster, and M. Mann, *RNA and RNA binding proteins participate in early stages of cell spreading through spreading initiation centers*. Cell, 2004. **117**(5): p. 649-62.
171. Zoller, M., *Tetraspanins: push and pull in suppressing and promoting metastasis*. Nat Rev Cancer, 2009. **9**(1): p. 40-55.
172. Hu, G., K.M. Drescher, and X.M. Chen, *Exosomal miRNAs: Biological Properties and Therapeutic Potential*. Front Genet, 2012. **3**: p. 56.

173. Mittelbrunn, M. and F. Sanchez-Madrid, *Intercellular communication: diverse structures for exchange of genetic information*. Nat Rev Mol Cell Biol, 2012. **13**(5): p. 328-35.
174. Pleil, J.D., et al., *Observing the human exposome as reflected in breath biomarkers: heat map data interpretation for environmental and intelligence research*. J Breath Res, 2011. **5**(3): p. 037104.
175. Mongini, P.K. and J.K. Inman, *Cytokine dependency of human B cell cycle progression elicited by ligands which coengage BCR and the CD21/CD19/CD81 costimulatory complex*. Cell Immunol, 2001. **207**(2): p. 127-40.
176. Shoham, T., et al., *The tetraspanin CD81 regulates the expression of CD19 during B cell development in a postendoplasmic reticulum compartment*. J Immunol, 2003. **171**(8): p. 4062-72.
177. Rohlena, J., et al., *Endothelial CD81 is a marker of early human atherosclerotic plaques and facilitates monocyte adhesion*. Cardiovasc Res, 2009. **81**(1): p. 187-96.
178. Blonder, J., et al., *Quantitative profiling of the detergent-resistant membrane proteome of iota-b toxin induced vero cells*. J Proteome Res, 2005. **4**(2): p. 523-31.
179. Wang, N., et al., *Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis*. J Proteome Res, 2007. **6**(1): p. 263-72.
180. Manza, L.L., et al., *Sample preparation and digestion for proteomic analyses using spin filters*. Proteomics, 2005. **5**(7): p. 1742-5.
181. Liebler, D.C. and A.J. Ham, *Spin filter-based sample preparation for shotgun proteomics*. Nat Methods, 2009. **6**(11): p. 785; author reply 785-6.
182. Villar, M., et al., *Systematic characterization of phosphorylation sites in NFATc2 by linear ion trap mass spectrometry*. Proteomics, 2006. **6 Suppl 1**: p. S16-27.
183. Katayama, H., et al., *Efficient in-gel digestion procedure using 5-cyclohexyl-1-pentyl-beta-D-maltoside as an additive for gel-based membrane proteomics*. Rapid Commun Mass Spectrom, 2004. **18**(20): p. 2388-94.
184. Pan, Y., et al., *The developmental regulation of CD81 in the rat retina*. Mol Vis, 2007. **13**: p. 181-9.
185. Liu, L., et al., *Tetraspanin CD151 promotes cell migration by regulating integrin trafficking*. J Biol Chem, 2007. **282**(43): p. 31631-42.
186. Mazzocca, A., F. Liotta, and V. Carloni, *Tetraspanin CD81-regulated cell motility plays a critical role in intrahepatic metastasis of hepatocellular carcinoma*. Gastroenterology, 2008. **135**(1): p. 244-256 e1.
187. Storck, S., M. Thiry, and P. Bouvet, *Conditional knockout of nucleolin in DT40 cells reveals the functional redundancy of its RNA-binding domains*. Biol Cell, 2009. **101**(3): p. 153-67.
188. Gaume, X., et al., *In vivo Study of the Histone Chaperone Activity of Nucleolin by FRAP*. Biochem Res Int, 2011. **2011**: p. 187624.
189. Berditchevski, F. and E. Odintsova, *Characterization of integrin-tetraspanin adhesion complexes: role of tetraspanins in integrin signaling*. J Cell Biol, 1999. **146**(2): p. 477-92.
190. Gorrini, C., et al., *Fibronectin controls cap-dependent translation through beta1 integrin and eukaryotic initiation factors 4 and 2 coordinated pathways*. Proc Natl Acad Sci U S A, 2005. **102**(26): p. 9200-5.
191. Sanvito, F., et al., *The beta4 integrin interactor p27(BBP/eIF6) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly*. J Cell Biol, 1999. **144**(5): p. 823-37.
192. Simons, M. and G. Raposo, *Exosomes--vesicular carriers for intercellular communication*. Curr Opin Cell Biol, 2009. **21**(4): p. 575-81.

193. Escola, J.M., et al., *Selective enrichment of tetraspan proteins on the internal vesicles of multivesicular endosomes and on exosomes secreted by human B-lymphocytes*. J Biol Chem, 1998. **273**(32): p. 20121-7.
194. Kuhn, S., et al., *A complex of EpCAM, claudin-7, CD44 variant isoforms, and tetraspanins promotes colorectal cancer progression*. Mol Cancer Res, 2007. **5**(6): p. 553-67.
195. Dominguez, F., et al., *Human endometrial CD98 is essential for blastocyst adhesion*. PLoS One, 2010. **5**(10): p. e13380.
196. Arduise, C., et al., *Tetraspanins regulate ADAM10-mediated cleavage of TNF-alpha and epidermal growth factor*. J Immunol, 2008. **181**(10): p. 7002-13.
197. Xu, D., C. Sharma, and M.E. Hemler, *Tetraspanin12 regulates ADAM10-dependent cleavage of amyloid precursor protein*. Faseb J, 2009.
198. van Hemert, M.J., H.Y. Steensma, and G.P. van Heusden, *14-3-3 proteins: key regulators of cell division, signalling and apoptosis*. Bioessays, 2001. **23**(10): p. 936-46.
199. Zhang, X.A., A.L. Bontrager, and M.E. Hemler, *Transmembrane-4 superfamily proteins associate with activated protein kinase C (PKC) and link PKC to specific beta(1) integrins*. J Biol Chem, 2001. **276**(27): p. 25005-13.
200. Berditchevski, F., et al., *A novel link between integrins, transmembrane-4 superfamily proteins (CD63 and CD81), and phosphatidylinositol 4-kinase*. J Biol Chem, 1997. **272**(5): p. 2595-8.
201. van Niel, G., et al., *The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis*. Dev Cell, 2011. **21**(4): p. 708-21.
202. Berditchevski, F. and E. Odintsova, *Tetraspanins as regulators of protein trafficking*. Traffic, 2007. **8**(2): p. 89-96.
203. Rana, S., et al., *Activation-induced internalization differs for the tetraspanins CD9 and Tspan8: Impact on tumor cell motility*. Int J Biochem Cell Biol.
204. Abache, T., et al., *The transferrin receptor and the tetraspanin web molecules CD9, CD81, and CD9P-1 are differentially sorted into exosomes after TPA treatment of K562 cells*. J Cell Biochem, 2007. **102**(3): p. 650-64.



## **ANEXOS**



## **ANEXOS I. TABLAS SUPLEMENTARIAS**



# ANEXOS I. TABLAS SUPLEMENTARIAS

**Tabla Suplementaria 1.** Lista de proteínas que interaccionan específicamente con CD81, EWI-2 e ICAM-1 en linfocitos T humanos.

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	Número de experimentos <sup>c</sup>						
			CD147	CD69	CD151	CD81 <sup>d</sup>	EWI-2 <sup>d</sup>	ICAM-1 <sup>d</sup>	VCAM-1
<b>Proteínas relacionadas con Actina</b>									
P12814	ACTN1	Alpha-actinin-1				3	3	3	
P35609	ACTN2	Alpha-actinin-2					3		
O43707	ACTN4	Alpha-actinin-4				1	3		1
Q99439	CNN2	Calponin-2					2		
P23528	CFL1	Cofilin-1	1				3	2	
Q9Y281	CFL2	Cofilin-2	1				3	1	
P21333	FLNA	Filamin-A (Alpha-filamin)	1				3	4	2
Q14315	FLNC	Filamin-C						4	
Q14847	LASP1	LIM and SH3 domain protein 1					2		
P11055	MYH3	Myosin-3	1		1			3	1
<b>Señalización Celular</b>									
P68400	CSNK2A1	Casein kinase II subunit alpha						3	1
P63244	GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1					2		
Q9P2D0	IBTK	Inhibitor of Bruton tyrosine kinase					2	1	
Q16891	IMMT	Mitochondrial inner membrane protein					2		
P19338	NCL	Nucleolin (Protein C23)	1				1	4	5
Q9Y3Z3	SAMHD1	SAM domain and HD domain-containing protein 1		1			4		
Q04323	UBXN1	SAPK substrate protein 1					2		
O94804	STK10	Serine/threonine-protein kinase 10			1		2		
P50453	SERPINB9	Serpin B9					2	3	
Q96FE3	INPP4A	Type I inositol-3,4-bisphosphate 4-phosphatase					2		
<b>Chaperonas</b>									
P50991	CCT4	T-complex protein 1 subunit delta					2	1	1
P49368	CCT3	T-complex protein 1 subunit gamma	1		2		2	2	1
<b>Histonas</b>									
P58876	HIST1H2BD	Histone 2B type 1-D				1	4	3	
O60814	HIST1H2BJ/HIST1H2BK	Histone 2B type 1-K				1	4	3	
Q02539	HIST1H1A	Histone H1.1			1	2	1	4	
P16402	HIST1H1D	Histone H1.3 (Histone H1c)				1	1	3	2
P10412	HIST1H1E	Histone H1.4 (Histone H1e)				1	2	4	
P16401	HIST1H1B	Histone H1.5 (Histone H1a)			1		3	4	2
Q92522	H1FX	Histone H1x					4	3	2
Q96A08	HIST1H2BA	Histone H2B type 1-A				1	3	1	
P33778	HIST1H2BB	Histone H2B type 1-B				1	3	3	
Q93079	HIST1H2BH/HIST1H2BO	Histone H2B type 1-H				1	4	3	
P06899	HIST1H2BJ/HIST1H2BK	Histone H2B type 1-J				1	3	3	
Q99880	HIST1H2BL	Histone H2B type 1-L				1	4	3	
Q99879	HIST1H2BM	Histone H2B type 1-M				1	4	3	
Q99877	HIST1H2BN	Histone H2B type 1-N				1	4	3	
P23527	HIST1H2BH/HIST1H2BO	Histone H2B type 1-O				1	3	3	
Q5QNW6	HIST2H2BF	Histone H2B type 2				1	4	2	
Q16778	HIST2H2BE (includes others)	Histone H2B type 2-E					2	3	
P62807		Histone H2B.a/g/h/k/l				1	4	3	
P68431		Histone H3.1				1	3	1	
Q16695	HIST3H3	Histone H3.1t				1	3	1	
Q71D13		Histone H3.2				1	3	1	
P84243		Histone H3.3				1	3	1	
P62805		Histone H4	2		1	1	3	1	2
Q09028	RBBP4	Histone-binding protein RBBP4					3		
P16403	HIST1H1C	HistoneH1.2				1	1	3	2
<b>Metabolismo</b>									
Q01813	PFKP	6-phosphofruktokinase type C					3		
P14618	PKM2	Pyruvate kinase isozymes M1/M2	1	1			2	1	
P11586	MTHFD1	C-1-tetrahydrofolate synthase, cytoplasmic					2		
P02794	FTH1	Ferritin heavy chain	1					4	
P02792	FTL	Ferritin light chain						3	
P25705	ATP5A1	ATP synthase alpha chain, mitochondrial precursor	2	1			2		
P04406	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	2		4		3		2
P43490	NAMPT	Nicotinamide phosphoribosyltransferase					2		
Q15084	PDIA6	Protein disulfide-isomerase A6 precursor	1				2	1	
P48047	ATP5O	ATP synthase subunit O					2		
<b>Proteínas relacionadas con ácidos nucleicos</b>									
Q92688	ANP32B	Acidic leucine-rich nuclear phosphoprotein 32				1	3	3	
P12956	XRCC6	ATP-dependent DNA helicase 2 subunit 1					4	1	
P13010	XRCC5	ATP-dependent DNA helicase 2 subunit 2					4	1	
O00571	DDX3X	ATP-dependent RNA helicase DDX3X				4	1	1	
O15523	DDX3Y	ATP-dependent RNA helicase DDX3Y				4	1		
Q43809	NUDT21	Cleavage and polyadenylation specificity factor subunit 5				2			
Q16630	CPSF6	Cleavage and polyadenylation specificity factor subunit 6				2			
P50238	CRIP1	Cysteine-rich protein 1				2			
P33992	MCM5	DNA replication licensing factor MCM5				2	1		
P16989	CSDA	DNA-binding protein A					4	3	

# ANEXOS I. TABLAS SUPLEMENTARIAS

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	Número de experimentos <sup>c</sup>							
			CD147	CD69	CD151	CD81 <sup>d</sup>	EWI-2 <sup>d</sup>	ICAM-1 <sup>d</sup>	VCAM-1	ICAM-3
Q05639	EEF1A2	Elongation factor 1				1	2	1		
P68104	EEF1A1	Elongation factor 1-alpha 1					3			
P24534	EEF1B2	Elongation factor 1-beta						3		
P29692	EEF1D	Elongation factor 1-delta					1	3		
P26641	EEF1G	Elongation factor 1-gamma						3		
P49411	TUFM	Elongation factor Tu, mitochondrial precursor					3			
Q95163	KBKAP	Elongator complex protein 1						3		
P60842	EF4A1	Eukaryotic initiation factor 4A-I					2			
Q14240	EF4A2	Eukaryotic initiation factor 4A-II					2			
Q14152	EIF3A	Eukaryotic translation initiation factor 3 subunit A					2	4	1	1
P55884	EIF3B	Eukaryotic translation initiation factor 3 subunit B					1	4	1	
Q99613	EIF3C/EIF3CL	Eukaryotic translation initiation factor 3 subunit C					2	3	1	
Q9Y262	EIF3L	Eukaryotic translation initiation factor 3 subunit E					2	3	1	
P60228	EIF3E	Eukaryotic translation initiation factor 3 subunit E					2	2	1	
O00303	EIF3F	Eukaryotic translation initiation factor 3 subunit F					1	4	2	
O75821	EIF3G	Eukaryotic translation initiation factor 3 subunit G					2	1	1	
O15372	EIF3H	Eukaryotic translation initiation factor 3 subunit H					2	1	1	
Q13347	EIF3I	Eukaryotic translation initiation factor 3 subunit I					1	4	1	
Q9UBQ5	EIF3K	Eukaryotic translation initiation factor 3 subunit K					2	2	1	
Q9Y5B9	SUPT16H	FACT complex subunit SPT16					2	1	1	
Q96AE4	FUBP1	Far upstream element-binding protein 1					2			
Q92945	KHSRP	Far upstream element-binding protein 2					2			
Q16666	IFI16	Gamma-interferon-inducible protein Ifi-16						2	4	
Q14444	CAPRIN1	GPI-anchored protein p137						3		
P09651	HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	1				3	2	1	
P51991	HNRNPA3	Heterogeneous nuclear ribonucleoprotein A3		1			2	1	1	
Q14103	HNRNPD	Heterogeneous nuclear ribonucleoprotein D0					2	3	2	
P31943	HNRNPH1	Heterogeneous nuclear ribonucleoprotein H					3	3	1	
P61978	HNRNPK	Heterogeneous nuclear ribonucleoprotein K		1			2	1	1	
P52272	HNRNPM	Heterogeneous nuclear ribonucleoprotein M					2	1		
O60506	SYNCRIP	Heterogeneous nuclear ribonucleoprotein Q					1	4	1	
P22626	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1		2			4	2	1	
Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U					4	3	2	1
Q12905	ILF2	Interleukin enhancer-binding factor 2						4	1	
Q12906	ILF3	Interleukin enhancer-binding factor 3						3	3	
Q8N1G4	LRR047	Leucine-rich repeat-containing protein 47	1				2			
Q15233	NONO	Non-POU domain-containing octamer-binding protein					4			
P67809	YBX1	Nuclease sensitive element-binding protein 1						4	3	
P06748	Npm1	Nucleophosmin					1	4	3	
Q8NCS1	SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein					2			
Q15365	PCBP1	Poly(rC)-binding protein 1		1			2			
Q15366	PCBP2	Poly(rC)-binding protein 2					2	1		
P11940	PABPC1	Polyadenylate-binding protein 1					1	3	2	
Q6P2Q9	PRFF8	Pre-mRNA-processing-splicing factor 8						3	1	
Q92841	DDX17	Probable ATP-dependent RNA helicase DDX17					4	1	1	
P17844	DDX5	Probable ATP-dependent RNA helicase DDX5					4	1		
Q53EL6	PDCD4	Programmed cell death protein 4					3	2		
Q01105	SET	Protein SET					2	4	4	
O43143	DHX15	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase					3	1		
Q9Y383	LUC7L2	Putative RNA-binding protein Luc7-like 2					2			
Q92541	RTF1	RNA polymerase-associated protein RTF1 homolog					2			
Q01844	EWSR1	RNA-binding protein EWS					4			
P37108	SRP14	Signal recognition particle 14 kDa protein						3	2	
Q07666		signal transduction-associated protein 1					2			
P62314	SNRPD1	Small nuclear ribonucleoprotein Sm D1					2	3	2	
P62318	SNRPD3	Small nuclear ribonucleoprotein Sm D3					2	4	3	
P14678	SNRFB	Small nuclear ribonucleoprotein-associated proteins B and B'					2	3	1	
Q9BRJ6	SRSF8	Splicing factor, arginine/serine-rich 2B					2	1		
P84103	SRSF3	Splicing factor, arginine/serine-rich 3					3	3	1	
Q16629	SF3B3	Splicing factor, arginine/serine-rich 7					3	2		
P23246	SFPQ	Splicing factor, proline- and glutamine-rich					4			
Q15020	SART3	Squamous cell carcinoma antigen recognized by T-cells 3						5	2	
Q86V81	THOC4	THO complex subunit 4		1		1	1	2	3	
Q00577	PURA	Transcriptional activator protein Pur-alpha						4	1	
Q13428	TOOF1	Treacle protein						2	3	
P23381	WARS	Tryptophanyl-tRNA synthetase, cytoplasmic					2			
P09234	SNRPC	U1 small nuclear ribonucleoprotein C					2			
P09661	SNRPA1	U2 small nuclear ribonucleoprotein A'						3	1	
P62310	LSM3	U6 snRNA-associated Sm-like protein LSM3						3	1	
Q9Y2T7	YBX2	Y-box-binding protein						3	3	
<b>Proteínas relacionadas con vía ras</b>										
Q16512	PKN1	Serine/threonine-protein kinase N1					2			
Q9H0B2		RANBP2-like and GRIP domain-containing protein 7	1				2			1 1

# ANEXOS I. TABLAS SUPLEMENTARIAS

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	Número de experimentos <sup>c</sup>							
			CD147	CD69	CD151	CD81 <sup>d</sup>	EWI-2	ICAM-1 <sup>d</sup>	VCAM-1	ICAM-3
P15153	RAC2	Ras-related C3 botulinum toxin substrate 2 precursor				2				
<b>Proteínas ribosomales</b>										
P62277	RPS13	40S ribosomal protein S13				1	3	1		
P62263	RPS14	40S ribosomal protein S14	1	4	2	2	2	2		
P62841	RPS15	40S ribosomal protein S15		2	2	1				
P62249	RPS16	40S ribosomal protein S16		1	3	1				
P62269	RPS18	40S ribosomal protein S18	1	2	3	3				
P39019	RPS19	40S ribosomal protein S19		2	3	3				
P15880	RPS2	40S ribosomal protein S2 (S4) (LLRep3 protein)		2	3	2				
P62266	RPS23	40S ribosomal protein S23		1	3	2				
P62851	RPS25	40S ribosomal protein S25	1	2	2	3	3			
Q5JNZ5		40S ribosomal protein S26-like 1				3	1			
P62857	RPS28	40S ribosomal protein S28	1	1	3	1				
P23396	RPS3	40S ribosomal protein S3		3	3	1				
P62861	FAU	40S ribosomal protein S30		1	2	3				
P61247	RPS3A	40S ribosomal protein S3a		2	4	2				
P62701	RPS4X	40S ribosomal protein S4, X isoform		2	5	2				
P22090	RPS4Y1	40S ribosomal protein S4, Y isoform 1			1	3	2			
P62753	RPS6	40S ribosomal protein S6		1	3	3				
P62241	RPS8	40S ribosomal protein S8		1	3	4	1			
P46781	RPS9	40S ribosomal protein S9		1	3	2				
P05388	Gm8730/Rplp0	60S acidic ribosomal protein P0		1	4	2				
P05387	RPLP2	60S acidic ribosomal protein P2	1	1	4	4	1			
P62906	RPL10A	60S ribosomal protein L10a	1		4	2				
P62913	RPL11	60S ribosomal protein L11		1	4	3				
P30050	EG241053	60S ribosomal protein L12		2	5	4	1			
P26373	RPL13	60S ribosomal protein L13		1	2	3				
P40429	RPL13A	60S ribosomal protein L13a			3	2				
P18621	RPL17	60S ribosomal protein L17			1	2	3			
Q07020	RPL18	60S ribosomal protein L18	1		3	3	1			
P84098	RPL19	60S ribosomal protein L19			4	1				
P35268	RPL22	60S ribosomal protein L22		3	4	3	2			
P62750	RPL23A	60S ribosomal protein L23a		1	3	3				
P83731	RPL24	60S ribosomal protein L24		1	3	1				
P61254	RPL26	60S ribosomal protein L26			3	1				
Q9UNX3	RPL26L1	60S ribosomal protein L26-like 1			4	2				
P61353	RPL27	60S ribosomal protein L27		1	5	3	1			
P46776	RPL27A	60S ribosomal protein L27a	1	1	2	3				
P62888	RPL30	60S ribosomal protein L30			2	4	1			
P62899	RPL31	60S ribosomal protein L31		1	2	4				
P42766	RPL35	60S ribosomal protein L35		1	3	3				
P18077	RPL35A	60S ribosomal protein L35a			4	3				
Q9Y3U8	RPL36	60S ribosomal protein L36			3	1				
P36578	RPL4	60S ribosomal protein L4 (L1)		1	2	3				
P46777	RPL5	60S ribosomal protein L5		1	4	1				
Q02878	RPL6	60S ribosomal protein L6		1	4	3				
P62424	Rpl7a (includes others)	60S ribosomal protein L7a			3	4	1			
P62917	RPL8	60S ribosomal protein L8			3	3	1			
P32969	RPL9	60S ribosomal protein L9		1	3	2				
<b>Proteínas relacionadas con el tráfico celular</b>										
P42166		Lamina-associated polypeptide 2 isoform alpha				2				
<b>Tubulinas</b>										
Q71U36	TUBA1A	Tubulin alpha-1A chain		2	4	1				
Q9BQE3	TUBA1C	Tubulin alpha-1C chain		2	4					
Q13748		Tubulin alpha-3C/D		2	4					
Q6PEY2	TUBA3E	Tubulin alpha-3E chain		2	2					
P68366	TUBA4A	Tubulin alpha-4A chain		1	2					
Q9NY65	TUBA8	Tubulin alpha-8 chain	1		2	2			1	
P68363	TUBA1B	Tubulin alpha-ubiquitous chain		2	4	1				
Q07437	TUBB	Tubulin beta chain			4	1	1			
Q9H4B7	TUBB1	Tubulin beta-1 chain	1		4	1	1			
Q13885	TUBB2A	Tubulin beta-2A chain	1		4					
Q9BVA1	TUBB2B	Tubulin beta-2B chain	1		4					
P68371	TUBB2C	Tubulin beta-2C chain	1		4					1
Q13509	TUBB3	Tubulin beta-3 chain	1		4					1
P04350	TUBB4	Tubulin beta-4 chain	1		4	1	1			1
Q99867	TUBB4Q	Tubulin beta-4q chain	1		4	1	1			
Q9BUIF5	TUBB6	Tubulin beta-6 chain		2	1	1				
Q9LUT0	TUBE1	Tubulin epsilon chain	1		2	3				
<b>Desconocido</b>										
Q52LJ0	FAM98B	Protein FAM98B				2				
<b>Transporte vesicular</b>										
Q96QK1	VPS35	Vacuolar protein sorting-associated protein 35					3			

<sup>a</sup> Referencia de las proteínas se tomó del software del IPA (<http://www.ingenuity.com>).

<sup>b</sup> Proteínas que interactúan con CD81, EWI-2 e ICAM-1.

<sup>c</sup> Número de experimentos donde son observadas las interacciones específicas.

<sup>d</sup> Las proteínas que interactúan específicamente con los ligandos de CD81, EWI-2 and ICAM-1 están en rojo; la intensidad del color depende del número de experimentos en los que la proteína interactúa.

# ANEXOS I. TABLAS SUPLEMENTARIAS

**Tabla suplementaria 2.** Proteínas identificadas en extractos de exosomas de linfocitos  
T en más de tres experimentos.

Entrada	Nº de acceso	Proteína	Nº peptidos	Nº Experimentos
CD45	P08575	Leukocyte common antigen	53	5
ANXA6	P08133	Annexin A6	49	5
MOES	P26038	Moesin	49	5
MYO1G	B011T2	Myosin-Ig	43	5
EZR1	P15311	Ezrin	39	5
EH1	Q9H4M9	EH domain-containing protein 1	31	5
CLH1	Q00610	Clathrin heavy chain 1	30	5
KPYM	P14618	Pyruvate kinase isozymes M1/M2	29	5
ITAL	P20701	Integrin alpha-L	28	5
ITB2	F05107	Integrin beta-2	27	5
DPP4	P27487	Dipeptidyl peptidase 4	26	5
ANXA2	F07355	Annexin A2	25	5
ITA4	P13612	Integrin alpha-4	25	5
LCK	P06239	Proto-oncogene tyrosine-protein kinase LCK	25	5
IQGA1	R46940	Ras GTPase-activating-like protein IQGAP1	25	5
PLSL	P13796	Plastin-2 OS	22	5
ACTA	F62736	Actin, aortic smooth muscle	21	5
ENOA	F06733	Alpha-enolase	21	5
FLOT1	O75955	Flotillin-1	21	5
FLOT2	Q14254	Flotillin-2	21	5
HSP7C	P11142	Heat shock cognate 71 kDa protein	19	5
1A01	F30443	HLA class I histocompatibility antigen, A-1 alpha chain	19	5
DOC10	Q96BY6	Dedicator of cytokinesis protein 10	17	5
PGK1	P00558	Phosphoglycerate kinase 1	17	5
4F2	F08195	4F2 cell-surface antigen heavy chain	16	5
ANXA5	F08758	Annexin A5	16	5
G3P	F04406	Glyceraldehyde-3-phosphate dehydrogenase	15	5
NHRF1	O14745	Na(+)/H(+) exchange regulatory cofactor NHE-RF1	15	5
TFR1	P02786	Transferrin receptor protein 1	15	5
FLNA	F21333	Filamin-A	14	5
1A02	F01892	HLA class I histocompatibility antigen, A-2 alpha chain	14	5
ICAM1	F05362	Intercellular adhesion molecule 1	14	5
RASL3	Q86YV0	RAS protein activator like-3	14	5
AT1A1	P05023	Sodium/potassium-transporting ATPase subunit alpha-1	14	5
TBA1A	Q71U36	Tubulin alpha-1A chain	14	5
TBB2A	Q13885	Tubulin beta-2A chain	14	5
ACTB	F60709	Actin, cytoplasmic 1	13	5
ANX11	F50995	Annexin A11	13	5
CLIC1	O00299	Chloride intracellular channel protein 1	13	5
ARF1	F84077	ADP-ribosylation factor 1	12	5
A2MG	F01023	Alpha-2-macroglobulin	12	5
CD44	P16070	CD44 antigen OS	12	5
EF1A1	F68104	Elongation factor 1-alpha 1	12	5
URP2	Q86LX7	Fermitin family homolog 3	12	5
ALDOA	F04075	Fructose-bisphosphate aldolase A	12	5
GNAI2	F04899	Guanine nucleotide-binding protein G(i), alpha-2 subunit	12	5
GBB1	F62873	Guanine nucleotide-binding protein G(i)/G(s)/G(t) subunit	12	5
ITB1	F05556	Integrin beta-1	12	5
ITB7	P26010	Integrin beta-7	12	5
GDIR2	F52566	Rho GDP-dissociation inhibitor 2	12	5
ANXA1	F04083	Annexin A1	11	5
ANXA4	F09525	Annexin A4	11	5
CD97	R48960	CD97 antigen OS	11	5
COR1A	P31146	Coronin-1A	11	5
EH4	Q9H223	EH domain-containing protein 4	11	5
H4	F62805	Histone H4	11	5
TNAP3	P21580	Tumor necrosis factor, alpha-induced protein 3	11	5
1433B	P31946	14-3-3 protein beta/alpha	10	5
SATT	R43007	Neutral amino acid transporter A	10	5
AAAT	Q15758	Neutral amino acid transporter B(0)	10	5
PROF1	F07737	Profilin-1	10	5
RAP1A	F62834	Ras-related protein Rap-1A	10	5
PP1A	F62136	Serine/threonine-protein phosphatase PP1-alpha catalytic s	10	5
ALBU	P02768	Serum albumin	10	5
CD2	F06729	T-cell surface antigen CD2	10	5
RHOA	F61586	Transforming protein RhoA	10	5
CP054	Q6UWD8	Transmembrane protein C16orf54	10	5
CD82	P27701	CD82 antigen	9	5
COF1	P23528	Cofilin-1	9	5
CYFP1	Q7L576	Cytoplasmic FMR1-interacting protein 1	9	5
STOM	F27105	Erythrocyte band 7 integral membrane protein	9	5
PP1A	F62937	Peptidyl-prolyl cis-trans isomerase A	9	5
SDCB1	O00560	Syntenin-1	9	5
CD3Z	P20963	T-cell surface glycoprotein CD3 zeta chain OS	9	5
1433Z	F63104	14-3-3 protein zeta/delta	8	5



# ANEXOS I. TABLAS SUPLEMENTARIAS

Entrada	Nº de acceso	Proteína	Nº péptidos	Nº Experimentos
CB1	Q99828	Calcium and integrin-binding protein 1	8	5
CYFP2	Q96F07	Cytoplasmic FMR1-interacting protein 2	8	5
GNA11	P63096	Guanine nucleotide-binding protein G(i), alpha-1 subunit	8	5
1A25	P18462	HLA class I histocompatibility antigen, A-25 alpha chain	8	5
2DRA	P01903	HLA class II histocompatibility antigen, DR alpha chain	8	5
2B11	P04229	HLA class II histocompatibility antigen, DRB1-1 beta chain	8	5
RAB7A	P51149	Ras-related protein Rab-7a	8	5
RAB8A	P61006	Ras-related protein Rab-8A	8	5
RAP2A	P10114	Ras-related protein Rap-2a	8	5
SIT1	Q9Y3P8	Signaling threshold-regulating transmembrane adapter 1	8	5
1433E	P62258	14-3-3 protein epsilon	7	5
ARF6	P62330	ADP-ribosylation factor 6	7	5
CD48	P09326	CD48 antigen	7	5
CDC42	P60953	Cell division control protein 42 homolog	7	5
IF4A1	P60842	Eukaryotic initiation factor 4A-1	7	5
LEUK	P16150	Leukosialin	7	5
LDHA	P00338	L-lactate dehydrogenase A chain	7	5
NCKFL	P55160	Nck-associated protein 1-like	7	5
GDB	P50395	Rab GDP dissociation inhibitor beta	7	5
RAB35	Q15286	Ras-related protein Rab-35	7	5
CD3D	P04234	T-cell surface glycoprotein CD3 delta chain	7	5
CD3E	P07766	T-cell surface glycoprotein CD3 epsilon chain	7	5
TMC8	Q8IU68	Transmembrane channel-like protein 8	7	5
TPS	P60174	Triosephosphate isomerase	7	5
SYWC	P23381	Tryptophanyl-tRNA synthetase, cytoplasmic	7	5
ARRD1	Q8NSI2	Arrestin domain-containing protein 1	6	5
HBA	P69905	Hemoglobin subunit alpha	6	5
H2B1B	P33778	Histone H2B type 1-B	6	5
ICAM3	P32942	Intercellular adhesion molecule 3	6	5
MYL6	P60660	Myosin light polypeptide 6	6	5
FA49B	Q9NUQ9	Protein FAM49B	6	5
RAC1	P63000	Ras-related C3 botulinum toxin substrate 1	6	5
RB11A	P62491	Ras-related protein Rab-11A	6	5
GTR14	Q8TDB8	Solute carrier family 2, facilitated glucose transporter	6	5
ACTN1	P12814	Alpha-actinin-1	5	5
CD20	P11836	B-lymphocyte antigen CD20 OS	5	5
CALM	P62158	Calmodulin	5	5
GBB2	P62879	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit	5	5
H2A1A	Q96QV6	Histone H2A type 1-A	5	5
1B07	P01889	HLA class I histocompatibility antigen, B-7 alpha chain	5	5
DFB1	P04440	HLA class II histocompatibility antigen, DP beta 1 chain	5	5
PRDX1	Q06830	Peroxisiredoxin-1	5	5
PLS3	Q9NRY6	Phospholipid scramblase 3	5	5
PTCA	Q14761	Protein tyrosine phosphatase receptor type C-associated protein	5	5
FYN	P06241	Proto-oncogene tyrosine-protein kinase Fyn	5	5
AT1B3	P54709	Sodium/potassium-transporting ATPase subunit beta-3	5	5
TBB5	P07437	Tubulin beta chain	5	5
FETUA	P02765	Alpha-2-HS-glycoprotein	4	5
B2MG	P61769	Beta-2-microglobulin	4	5
CAB39	Q9Y376	Calcium-binding protein 39	4	5
CD81	P60033	CD81 antigen	4	5
PRKDC	P78527	DNA-dependent protein kinase catalytic subunit	4	5
GAPR1	Q9H4G4	Golgi-associated plant pathogenesis-related protein 1	4	5
1B13	P30461	HLA class I histocompatibility antigen, B-13 alpha chain	4	5
HB21	P01918	HLA class II histocompatibility antigen, DQ(1) beta chain	4	5
ITA5	P08648	Integrin alpha-5	4	5
CD37	P11049	Leukocyte antigen CD37	4	5
CD53	P19397	Leukocyte surface antigen CD53	4	5
RAC2	P15153	Ras-related C3 botulinum toxin substrate 2	4	5
RAP1B	P61224	Ras-related protein Rap-1b	4	5
CD4	P01730	T-cell surface glycoprotein CD4 OS	4	5
TBB1	Q9H4B7	Tubulin beta-1 chain	4	5
TBB2C	P68371	Tubulin beta-2C chain	4	5
UBIQ	P62988	Ubiquitin	4	5
BST2	Q10589	Bone marrow stromal antigen 2	3	5
CXCR4	P61073	C-X-C chemokine receptor type 4	3	5
H31T	Q16695	Histone H3.1t	3	5
ITIH2	P19823	Inter-alpha-trypsin inhibitor heavy chain H2	3	5
FM1	P13164	Interferon-induced transmembrane protein 1	3	5
LSR	Q86X29	Lipolysis-stimulated lipoprotein receptor	3	5
NEUL	Q9BYT8	Neurolysin, mitochondrial	3	5
FDC10	Q9BUL8	Programmed cell death protein 10	3	5
SELPL	Q14242	P-selectin glycoprotein ligand 1	3	5
GDIR1	P52565	Rho GDP-dissociation inhibitor 1	3	5

# ANEXOS I. TABLAS SUPLEMENTARIAS

Entrada	Nº de acceso	Proteína	Nº peptidos	Nº Experimentos
APOA1	P02647	Apolipoprotein A-I	2	5
ACTBM	Q9BYX7	Beta-actin-like protein 3	2	5
HBB	P68871	Hemoglobin subunit beta	2	5
H2A1B	P04908	Histone H2A type 1-B/E	2	5
H2B1C	P62807	Histone H2B type 1-C/E/F/G/I	2	5
2DA2	P01906	HLA class II histocompatibility antigen, DQ alpha 2 chain	2	5
2B1B	P20039	HLA class II histocompatibility antigen, DRB1-11 beta chain	2	5
TRFL	P02788	Lactotransferrin	2	5
ASH1L	Q9NR48	Probable histone-lysine N-methyltransferase ASH1L	2	5
S10A9	P06702	Protein S100-A9	2	5
RAB10	P61026	Ras-related protein Rab-10	2	5
TCPE	P48643	T-complex protein 1 subunit epsilon	2	5
A1A1T	P01009	Alpha-1-antitrypsin	1	5
BRSK1	Q8TDC3	BR serine/threonine-protein kinase 1	1	5
CAD20	Q9HBT6	Cadherin-20 OS	1	5
G3PT	O14556	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	1	5
HS71L	P34931	Heat shock 70 kDa protein 1L	1	5
P3H4	P12577	Large structural protein	1	5
A1ATR	P20848	Putative alpha-1-antitrypsin-related protein	1	5
SPTN2	O15020	Spectrin beta chain, brain 2	1	5
STIM2	Q9P246	Stromal interaction molecule 2	1	5
TENX	P22105	Tenascin-X	1	5
TLN1	Q9Y490	Talin-1	16	4
FMNL	Q95466	Formin-like protein 1	14	4
FDC6I	Q8WUM4	Programmed cell death 6-interacting protein	14	4
IQGGA2	Q13576	Ras GTPase-activating-like protein IQGAP2	14	4
DIF2B	Q9P265	Disco-interacting protein 2 homolog B	13	4
MYH9	P35579	Myosin-9	13	4
PECA1	P16284	Platelet endothelial cell adhesion molecule	13	4
CD38	P28907	ADP-ribosyl cyclase 1	11	4
DEF6	Q9H4E7	Differentially expressed in FDCP 6 homolog	11	4
P4KA	P42356	Phosphatidylinositol 4-kinase alpha	11	4
WDR1	Q75083	WD repeat-containing protein 1	10	4
ADA10	O14672	Disintegrin and metalloproteinase domain-containing protein 10	9	4
CD1A	P06127	T-cell surface glycoprotein CD1a	9	4
GNAI3	P08754	Guanine nucleotide-binding protein G(k) subunit alpha	8	4
IGSF8	Q969P0	Immunoglobulin superfamily member 8	8	4
LDHB	P07195	L-lactate dehydrogenase B chain	8	4
PAG1	Q9NWO8	Phosphoprotein associated with glycosphingolipid-enriched	8	4
EFR3A	Q14156	Protein EFR3 homolog A	8	4
STK10	Q94804	Serine/threonine-protein kinase 10	8	4
TMC6	Q7Z403	Transmembrane channel-like protein 6	8	4
VIME	P08670	Vimentin	8	4
ARF2	P61160	Actin-related protein 2	7	4
ARF3	P61158	Actin-related protein 3	7	4
ANXA7	P20073	Annexin A7	7	4
EF2	P13639	Elongation factor 2	7	4
TB10A	Q9BXI6	TBC1 domain family member 10A	7	4
CD3G	P09693	T-cell surface glycoprotein CD3 gamma chain	7	4
BASI	P35613	Basigin	6	4
CO3	P01024	Complement C3	6	4
CAZA1	P52907	F-actin-capping protein subunit alpha-1	6	4
G6PI	P06744	Glucose-6-phosphate isomerase	6	4
M4K4	Q95819	Mitogen-activated protein kinase kinase kinase 4	6	4
RHOG	P84095	Rho-related GTP-binding protein RhoG	6	4
GTR1	P11166	Solute carrier family 2, facilitated glucose transporter	6	4
CD8A	P01732	T-cell surface glycoprotein CD8 alpha chain	6	4
TSN14	Q8NG11	Tetraspanin-14	6	4
TAGL2	P37802	Transgelin-2	6	4
CAP1	Q01518	Adenylyl cyclase-associated protein 1	5	4
CTL2	Q8MWA5	Choline transporter-like protein 2	5	4
GSTP1	P09211	Glutathione S-transferase P	5	4
GNA13	Q14344	Guanine nucleotide-binding protein subunit alpha-13	5	4
2DA1	P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	5	4
2B18	Q30134	HLA class II histocompatibility antigen, DRB1-8 beta chain	5	4
ITAX	P20702	Integrin alpha-X	5	4
RALA	P11233	Ras-related protein Ral-A	5	4
SNF23	O00161	Synaptosomal-associated protein 23	5	4
ARPC2	O15144	Actin-related protein 2/3 complex subunit 2	4	4
ESYT1	Q9BSJ8	Extended synaptotagmin-1	4	4
RASH	P01112	GTPase HRas	4	4
HSP72	P54652	Heat shock-related 70 kDa protein 2	4	4
H12	P16403	Histone H1.2	4	4
H15	P16401	Histone H1.5	4	4
H2B1A	Q96A08	Histone H2B type 1-A	4	4

# ANEXOS I. TABLAS SUPLEMENTARIAS

Entrada	Nº de acceso	Proteína	Nº péptidos	Nº Experimentos
H90B2	Q58FF8	Putative heat shock protein HSP 90-beta 2	4	4
H90B3	Q58FF7	Putative heat shock protein HSP 90-beta-3	4	4
GDIA	P31150	Rab GDP dissociation inhibitor alpha	4	4
RFTN1	Q14699	Raftin	4	4
RAP2B	P61225	Ras-related protein Rap-2b	4	4
1433T	P27348	14-3-3 protein theta	3	4
RL18	Q07020	60S ribosomal protein L18	3	4
GRF78	P11021	78 kDa glucose-regulated protein	3	4
CD59	P13987	CD59 glycoprotein	3	4
CPNE3	O75131	Copine-3	3	4
1B35	P30685	HLA class I histocompatibility antigen, B-35 alpha chain	3	4
HYCC1	Q9BY13	Hyccin	3	4
PGAM1	P18669	Phosphoglycerate mutase 1	3	4
FLS1	O15162	Phospholipid scramblase 1	3	4
STXB2	Q15833	Syntaxin-binding protein 2	3	4
TCA	P01848	T-cell receptor alpha chain C region	3	4
TCB	P01850	T-cell receptor beta chain C region	3	4
VTNC	P04004	Vitronectin	3	4
ARPC4	P59998	Actin-related protein 2/3 complex subunit 4	2	4
IF5A1	P63241	Eukaryotic translation initiation factor 5A-1	2	4
ALDOC	P09972	Fructose-bisphosphate aldolase C	2	4
GBG2	P59768	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit	2	4
2B1F	P01911	HLA class II histocompatibility antigen, DRB1-15 beta chain	2	4
ICOS	Q9Y6W8	Inducible T-cell costimulator	2	4
ITHEL	Q6UXX5	Inter-alpha-trypsin inhibitor heavy chain H5-like protein	2	4
CD47	Q08722	Leukocyte surface antigen CD47	2	4
CAPG	P40121	Macrophage-capping protein	2	4
MITD1	Q8WV92	MIT domain-containing protein 1	2	4
MYL6B	P14649	Myosin light chain 6B	2	4
NICA	Q92542	Nicastrin	2	4
PLSI	Q14651	Plastin-1	2	4
PKHA4	Q9H4M7	Pleckstrin homology domain-containing family A member 4	2	4
RAB8B	Q92930	Ras-related protein Rab-8B	2	4
SH3L3	Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	2	4
S12A7	Q9Y666	Solute carrier family 12 member 7	2	4
TTC27	Q6P3X3	Tetratricopeptide repeat protein 27	2	4
THIO	P10599	Thioredoxin	2	4
THBG	P05543	Thyroxine-binding globulin	2	4
TBA1B	P68363	Tubulin alpha-1B chain	2	4
PSMD1	Q99460	26S proteasome non-ATPase regulatory subunit 1	1	4
ARPC5	O15511	Actin-related protein 2/3 complex subunit 5	1	4
ANXA3	P12429	Annexin A3	1	4
APIB1	Q10567	AP-1 complex subunit beta-1	1	4
CASC1	Q6TDU7	Cancer susceptibility candidate protein 1	1	4
CD9	P21926	CD9 antigen	1	4
CORO7	P57737	Coronin-7	1	4
CY24B	P04839	Cytochrome b-245 heavy chain	1	4
DGKI	O75912	Diacylglycerol kinase iota	1	4
ENPL	P14625	Endoplasmic	1	4
GFPT1	Q06210	Glucosamine-fructose-6-phosphate aminotransferase	1	4
GNL3L	Q9NVN8	Guanine nucleotide-binding protein-like 3-like protein	1	4
H11	Q02539	Histone H1.1	1	4
1B15	P30464	HLA class I histocompatibility antigen, B-15 alpha chain	1	4
IP3KA	P23677	Inositol-trisphosphate 3-kinase A	1	4
MATRX	P27019	Matrix protein	1	4
VL2	P36750	Minor capsid protein L2	1	4
PCBP1	Q15365	Poly(rC)-binding protein 1	1	4
PABP4	Q13310	Polyadenylate-binding protein 4	1	4
PZP	P20742	Pregnancy zone protein	1	4
CC049	Q96BT1	Putative uncharacterized protein C3orf49	1	4
RGPA1	Q6GYQ0	Ral GTPase-activating protein alpha subunit 1	1	4
RAB1A	P62820	Ras-related protein Rab-1A	1	4
RAB3D	Q95716	Ras-related protein Rab-3D	1	4
SLAF7	Q9NQ25	SLAM family member 7	1	4
TRAT1	Q6PZ9	T-cell receptor-associated transmembrane adapter 1	1	4
K1109	Q2LD37	Uncharacterized protein KIAA1109	1	4
K1529	Q9PIZ9	Uncharacterized protein KIAA1529	1	4
CN159	Q7Z3D6	UPF0317 protein C14orf159, mitochondrial	1	4
UL114	P16769	Uracil-DNA glycosylase	1	4
UACA	Q9BZF9	Uveal autoantigen with coiled-coil domains	1	4
MVP	Q14764	Major vault protein	20	3
ITAM	P11215	Integrin alpha-M	9	3
FA65B	Q9Y4F9	Protein FAM65B	7	3
STAT1	P42224	Signal transducer and activator of transcription 1-alpha/beta	7	3
GNAS1	Q5JWF2	Guanine nucleotide-binding protein G(s) subunit alpha isoforms Xlas	6	3

# ANEXOS I. TABLAS SUPLEMENTARIAS

Entrada	Nº de acceso	Proteína	Nº péptidos	Nº Experimentos
ACSL4	Q60488	Long-chain-fatty-acid--CoA ligase 4	6	3
RRAS2	P62070	Ras-related protein R-Ras2	6	3
TCPB	P78371	T-complex protein 1 subunit beta	6	3
RS2	P15880	40S ribosomal protein S2	5	3
RS8	P62241	40S ribosomal protein S8	5	3
FA5	P12259	Coagulation factor V OS	5	3
DOCK8	Q8NF50	Dedicator of cytokinesis protein 8	5	3
IGSF2	Q93033	Immunoglobulin superfamily member 2	5	3
MFGM	Q08431	Lactadherin	5	3
PERF	P14222	Perforin-1	5	3
AT2B4	P23634	Plasma membrane calcium-transporting ATPase 4	5	3
41	P11171	Protein 4.1	5	3
TNIP1	Q15025	TNFAIP3-interacting protein 1	5	3
CSK	P41240	Tyrosine-protein kinase CSK	5	3
1433F	Q04917	14-3-3 protein eta	4	3
RL3	P39023	60S ribosomal protein L3	4	3
RL4	P36578	60S ribosomal protein L4	4	3
RL7	P18124	60S ribosomal protein L7	4	3
RL7A	P62424	60S ribosomal protein L7a	4	3
DOCK2	Q92608	Dedicator of cytokinesis protein 2	4	3
HPT	P00738	Haptoglobin	4	3
PP16B	Q96T49	Protein phosphatase 1 regulatory inhibitor subunit 16B	4	3
S10AB	P31949	Protein S100-A11	4	3
RHOE	Q9HBD0	Rho-related GTP-binding protein RhoF	4	3
SAV	P22629	Streptavidin	4	3
STX4	Q12846	Syntaxin-4	4	3
ACTN4	Q43707	Alpha-actinin-4	3	3
ANT3	P01008	Antithrombin-III	3	3
BAF	O75531	Barrier-to-autointegration factor	3	3
CD40L	P29965	CD40 ligand	3	3
AMPL	P28838	Cytosol aminopeptidase	3	3
DFP2A	Q14689	Disco-interacting protein 2 homolog A	3	3
EVL	Q9UI08	Ena/VASP-like protein	3	3
GNLY	P22749	Granulysin	3	3
HS90A	P07900	Heat shock protein HSP 90-alpha	3	3
HS90B	P08238	Heat shock protein HSP 90-beta	3	3
HEMO	P02790	Hemopexin	3	3
HG2A	P04233	HLA class II histocompatibility antigen gamma chain	3	3
HB22	P01919	HLA class II histocompatibility antigen, DQ(W1.1) beta chain	3	3
GBP1	P32455	Interferon-induced guanylate-binding protein 1	3	3
MINK1	Q8N4C8	Msshapen-like kinase 1	3	3
SHIP1	Q92835	Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase 1	3	3
PRR8	Q6P2Q9	Pre-mRNA-processing-splicing factor 8	3	3
PDOD6	O75340	Programmed cell death protein 6	3	3
TTYH3	Q9C0H2	Protein tyrosine homolog 3	3	3
PNFH	P00491	Purine nucleoside phosphorylase	3	3
RAB5A	P20339	Ras-related protein Rab-5A	3	3
RAB5C	P51148	Ras-related protein Rab-5C	3	3
PTPRA	P18433	Receptor-type tyrosine-protein phosphatase alpha	3	3
MST4	Q9P289	Serine/threonine-protein kinase MST4	3	3
FNBP2	O75044	SLIT-ROBO Rho GTPase-activating protein 2	3	3
CD6	P30203	T-cell differentiation antigen CD6	3	3
TXNIP	Q9H3M7	Thioredoxin-interacting protein	3	3
TRAF1	Q13077	TNF receptor-associated factor 1	3	3
PTN6	P29350	Tyrosine-protein phosphatase non-receptor type 6	3	3
PSD11	O00231	26S proteasome non-ATPase regulatory subunit 11	2	3
RSSA	P08865	40S ribosomal protein SA	2	3
RLA0L	Q8NHV5	60S acidic ribosomal protein P0-like	2	3
RL14	P50914	60S ribosomal protein L14	2	3
RL27A	P46776	60S ribosomal protein L27a	2	3
RL6	Q02878	60S ribosomal protein L6	2	3
FETA	P02771	Alpha-fetoprotein	2	3
ACTBL	Q562R1	Beta-actin-like protein 2	2	3
BPAEA	Q94833	Bullous pemphigoid antigen 1, isoforms 6/9/10	2	3
CD226	Q15762	CD226 antigen	2	3
CD63	P08962	CD63 antigen	2	3
CLIC5	Q9NZA1	Chloride intracellular channel protein 5	2	3
CLDN1	Q9NY35	Claudin domain-containing protein 1	2	3
COTL1	Q14019	Coactosin-like protein	2	3
S29A1	Q99808	Equilibrative nucleoside transporter 1	2	3
LEG1	P09382	Galectin-1	2	3
GGT1	P19440	Gamma-glutamyltranspeptidase 1	2	3
RASN	P01111	GTPase NRas	2	3
HS105	Q92598	Heat shock protein 105 kDa	2	3
HNRPK	P61978	Heterogeneous nuclear ribonucleoprotein K	2	3

# ANEXOS I. TABLAS SUPLEMENTARIAS

Entrada	Nº de acceso	Proteína	Nº péptidos	Nº Experimentos
H2AV	Q71U9	Histone H2A.V	2	3
HB23	P05537	HLA class II histocompatibility antigen, DQ(W3) beta chain	2	3
ITA1	P56199	Integrin alpha-1	2	3
LAT	O43561	Linker for activation of T-cells family member 1	2	3
LUM	P51884	Lumican	2	3
PIIB	P23284	Peptidyl-prolyl cis-trans isomerase B	2	3
PRDX2	P32119	Peroxiredoxin-2	2	3
PICAL	Q13492	Phosphatidylinositol-binding clathrin assembly protein	2	3
HS905	Q58FG0	Putative heat shock protein HSP90-alpha A5	2	3
RADI	P35241	Radixin	2	3
RAP2C	Q9Y3L5	Ras-related protein Rap-2c	2	3
RN213	Q63HN8	RING finger protein 213	2	3
SH21A	O60880	SH2 domain-containing protein 1A	2	3
S38A5	Q8WUX1	Sodium-coupled neutral amino acid transporter 5	2	3
STX3	Q13277	Syntaxin-3	2	3
TCPD	P50991	T-complex protein 1 subunit delta	2	3
TYPH	P19971	Thymidine phosphorylase	2	3
TRAF2	Q12933	TNF receptor-associated factor 2	2	3
TM149	Q9H665	Transmembrane protein 149	2	3
TBE	Q9LUT0	Tubulin epsilon chain	2	3
TNR6	P25445	Tumor necrosis factor receptor superfamily member 6	2	3
JAK3	P52333	Tyrosine-protein kinase JAK3	2	3
UBP5	P45974	Ubiquitin carboxyl-terminal hydrolase 5	2	3
UBS3A	P57075	Ubiquitin-associated and SH3 domain-containing protein A	2	3
WASF2	Q9Y6W5	Wiskott-Aldrich syndrome protein family member 2	2	3
RS5	P46782	40S ribosomal protein S5	1	3
RS7	P62081	40S ribosomal protein S7	1	3
THIL	P24752	Acetyl-CoA acetyltransferase, mitochondrial	1	3
ARC	Q7LC44	Activity-regulated cytoskeleton-associated protein	1	3
APC	P25054	Adenomatous polyposis coli protein	1	3
TTPA	P49638	Alpha-tocopherol transfer protein	1	3
APOB	P04114	Apolipoprotein B-100	1	3
CALX	P27824	Calnexin	1	3
CLAT	P28329	Choline O-acetyltransferase	1	3
CSTF3	Q12996	Cleavage stimulation factor 77 kDa subunit	1	3
CCD18	Q5T9S5	Coiled-coil domain-containing protein 18	1	3
CRY1	Q16526	Cryptochrome-1	1	3
EMB	Q6PCB8	ErbB3	1	3
EFNA4	P52798	Ephrin-A4 OS	1	3
FSIP1	Q8NA03	Fibrous sheath-interacting protein 1	1	3
POLG	Q82122	Genome polyprotein	1	3
HSFX1	Q9UBD0	Heat shock transcription factor, X-linked	1	3
H2A2B	Q8JUE6	Histone H2A type 2-B	1	3
1A11	P13746	HLA class I histocompatibility antigen, A-11 alpha chain	1	3
1A80	Q09160	HLA class I histocompatibility antigen, A-80 alpha chain	1	3
DQB2	P05538	HLA class II histocompatibility antigen, DQ beta 2 chain	1	3
2B14	P13760	HLA class II histocompatibility antigen, DRB1-4 beta chain	1	3
HB2S	P01916	HLA class II histocompatibility antigen, SB beta chain	1	3
HXA4	Q00056	Homeobox protein Hox-A4	1	3
IL15	P40933	Interleukin-15	1	3
KIF3C	O14782	Kinesin-like protein KIF3C	1	3
LAMB4	A4D0S4	Laminin subunit beta-4	1	3
LRGUK	Q96M69	Leucine-rich repeat and guanylate kinase domain-containing protein	1	3
ACSL3	Q95573	Long-chain-fatty-acid--CoA ligase 3	1	3
LAMP6	Q13571	Lysosomal-associated transmembrane protein 5	1	3
LAMP1	P11279	Lysosome-associated membrane glycoprotein 1	1	3
LAMP2	P13473	Lysosome-associated membrane glycoprotein 2	1	3
KIME	Q03426	Mevalonate kinase	1	3
MYH11	P35749	Myosin-11	1	3
MYO1E	Q12965	Myosin-1e	1	3
NLRX1	Q86UT6	NLR family member X1	1	3
NUP93	Q8N1F7	Nuclear pore complex protein Nup93	1	3
PDZD7	Q9H5P4	PDZ domain-containing protein 7	1	3
PHRF1	Q9P1Y6	PHD and RING finger domain-containing protein 1	1	3
PRICK4	Q2TBC4	Prickle-like protein 4	1	3
FAM3C	Q92520	Protein FAM3C	1	3
S10A6	P06703	Protein S100-A6	1	3
TANC1	Q9COD5	Protein TANC1	1	3
PLP2	Q04941	Proteolipid protein 2	1	3
PCD17	O14917	Protocadherin-17	1	3
CD8BL	A6NJV9	Putative T-cell surface glycoprotein CD8 beta-2 chain	1	3
VE1	Q02049	Replication protein E1	1	3
RTTN	Q86VV8	Rotatin	1	3
PRSL1	Q6UWY2	Serine protease 1-like protein 1	1	3
SMG1	Q96Q15	Serine/threonine-protein kinase SMG1	1	3



# ANEXOS I. TABLAS SUPLEMENTARIAS

**Tabla suplementaria 3.** Lista de proteínas que interactúan con CD81 y EWI-2 en exosomas de linfocitos T humanos

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	CD81 <sup>c</sup>	EWI-2 <sup>c</sup>
<b>Proteínas relacionadas con Actina</b>				
Q8IZP0	ABI1	Abl interactor 1	0.09	0.01
P62736	ACTA	Actin, aortic smooth muscle	2.0E-03	1.7E-03
P60709	ACTB	Actin, cytoplasmic 1	2.2E-03	5.0E-04
P59998	ARPC4	Actin-related protein 2/3 complex subunit 4	0.19	0.04
P12814	ACTN1	Alpha-actinin-1	0.04	0.01
Q9BYX7	POTEKP	Beta-actin-like protein 3	1.3E-03	2.8E-03
P23528	CFL1	Cofilin-1	6.1E-04	3.1E-05
Q7L576	CYFIP1	Cytoplasmic FMR1-interacting protein 1	0.19	0.02
P15311	EZR	Ezrin	2.6E-04	0.06
P52907	CAPZA1	F-actin-capping protein subunit alpha-1	0.01	0.01
P47756	CAPZB	F-actin-capping protein subunit beta	0.19	8.8E-05
Q86UX7	FERMT3	Fermitin family homolog 3	0.18	0.03
P21333	FLNA	Filamin-A	0.11	0.01
P26038	MSN	Moesin	2.1E-03	0.07
P60660	MYL6	Myosin light polypeptide 6	0.18	0.05
P35580	MYH10	Myosin-10	0.18	0.05
P35579	MYH9	Myosin-9	0.09	2.6E-03
O94832	MYO1D	Myosin-Id	No data	0.04
B011T2	MYO1G	Myosin-Ig	0.09	3.8E-03
O14745	SLC9A3R1	Na(+)/H(+) exchange regulatory cofactor NHE-RF1	4.0E-04	2.3E-03
P55160	NCKAP1L	Nck-associated protein 1-like	No data	0.04
P07737	PFN1	Profilin-1	1.4E-03	3.0E-03
O00560	SDCBP	Syntenin-1	2.5E-03	6.2E-04
P09493	TPM1	Tropomyosin alpha-1 chain	0.04	0.04
P08670	VIM	Vimentin	0.08	0.04
Q9Y6W5	WASF2	Wiskott-Aldrich syndrome protein family member 2	0.19	0.04
<b>Señalización celular</b>				
P31946	YWHAB	14-3-3 protein beta/alpha	0.19	0.03
P62258	YWHAE	14-3-3 protein epsilon	No data	0.02
P61981	YWHAG	14-3-3 protein gamma	No data	0.05
P63104	YWHAZ	14-3-3 protein zeta/delta	0.01	7.2E-04
Q99828	CIB1	Calcium and integrin-binding protein 1	1.9E-03	0.01
P63096	GNAI1	Guanine nucleotide-binding protein G(i), alpha-1 subunit	7.3E-04	1.3E-04
P04899	GNAI2	Guanine nucleotide-binding protein G(i), alpha-2 subunit	3.9E-03	3.4E-03
P62873	GNB1	Guanine nucleotide-binding protein G(l)/G(s)/G(t) subunit	9.8E-04	1.1E-04
P08754	GNAI3	Guanine nucleotide-binding protein G(k) subunit alpha	0.02	0.01
P63244	GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1	No data	0.04
Q9P2D0	IBTK	Inhibitor of Bruton tyrosine kinase	0.19	0.04
Q14764	MVP	Major vault protein	0.19	0.03
O95819	MAP4K4	Mitogen-activated protein kinase kinase kinase 4	No data	0.01
Q5TEA3		Myotubularin-related protein 14 OS=Homo sapiens GN=MTMR1	0.21	0.04
Q9BYT8	NLN	Neurolysin, mitochondrial	No data	0.02
Q06830	PRDX1	Peroxiredoxin-1	0.10	0.05
P42356	PI4KA	Phosphatidylinositol 4-kinase alpha	No data	0.04
Q8WUM4	PDCD6IP	Programmed cell death 6-interacting protein	0.05	0.02
O75340	PDCD6	Programmed cell death protein 6	0.02	No data
Q96T49	PPP1R16B	Protein phosphatase 1 regulatory inhibitor subunit 16B	0.19	4.2E-03
P05109	S100A8	Protein S100-A8	0.12	0.03
P06241	FYN	Proto-oncogene tyrosine-protein kinase Fyn	0.19	0.02
P06239	LCK	Proto-oncogene tyrosine-protein kinase LCK	0.02	1.7E-03

# ANEXOS I. TABLAS SUPLEMENTARIAS

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	CD81 <sup>c</sup>	EWI-2 <sup>c</sup>
Q96Q04	LMTK3	Serine/threonine-protein kinase LMTK3	0.04	0.01
Q96Q15	SMG1	Serine/threonine-protein kinase SMG1	0.04	2.5E-03
P62136	PPP1CA	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	0.04	2.5E-03
P42224	STAT1	Signal transducer and activator of transcription 1-alpha	No data	0.01
P10599	TXN	Thioredoxin	5.9E-05	0.02
Q13077	TRAF1	TNF receptor-associated factor 1	No data	0.01
Q15025	TNIP1	TNFAIP3-interacting protein 1	No data	0.05
Q96S44		TP53-regulating kinase	0.04	No data
P37802	TAGLN2	Transgelin-2	No data	0.04
P21580	TNFAIP3	Tumor necrosis factor, alpha-induced protein 3	No data	0.02
P62988		Ubiquitin	2.3E-03	0.01
<b>Chaperonas</b>				
P34931	HS71L	Heat shock 70 kDa protein 1L	3.1E-05	0.04
P11142	HSPA8	Heat shock cognate 71 kDa protein	0.01	0.01
P08238	HSP90AB1	Heat shock protein HSP 90-beta	0.19	0.04
Q58FF7	HSP90AB2P	Putative heat shock protein HSP 90-beta-3	No data	0.04
<b>Histonas</b>				
Q02539	HIST1H1A	Histone H1.1	0.04	0.04
P16403	HIST1H1C	Histone H1.2	0.20	0.05
Q96QV6	HIST1H2AA	Histone H2A type 1-A	2.4E-03	4.8E-03
P04908		Histone H2A type 1-B/E	0.10	0.05
Q71UI9		Histone H2A.V	No data	0.05
P33778	HIST1H2BB	Histone H2B type 1-B	0.01	0.01
P62807		Histone H2B type 1-C/E/F/G/I	0.04	1.3E-03
Q16695	HIST3H3	Histone H3.1t	2.4E-03	3.7E-03
P62805		Histone H4	1.7E-03	2.6E-03
<b>Proteínas de membrana y receptores</b>				
P11836	MS4A1	B-lymphocyte antigen CD20	0.04	No data
P16070	CD44	CD44 antigen	0.01	0.09
P48960	CD97	CD97 antigen	2.4E-04	0.09
P61073	CXCR4	C-X-C chemokine receptor type 4	0.05	No data
O14672	ADAM10	Disintegrin and metalloproteinase domain-containing protein 10	0.02	0.19
P30443	HLA-A	HLA class I histocompatibility antigen, A-1 alpha chain	3.1E-03	0.01
P01892	HLA-A	HLA class I histocompatibility antigen, A-2 alpha chain	4.0E-03	0.04
P18462	HLA-A	HLA class I histocompatibility antigen, A-25 alpha chain	0.04	0.05
P01889	HLA-B	HLA class I histocompatibility antigen, B-7 alpha chain	0.02	0.03
P04233	CD74	HLA class II histocompatibility antigen gamma chain	0.04	1.00
P01903	HLA-DRA	HLA class II histocompatibility antigen, DR alpha chain	0.04	No data
P04229	HLA-DRB1	HLA class II histocompatibility antigen, DRB1-1 beta chain	0.04	0.19
P26010	ITGB7	Integrin beta-7	0.01	No data
P13164	IFITM1	Interferon-induced transmembrane protein 1	0.09	0.04
P08575	PTPRC	Leukocyte common antigen	4.8E-04	0.02
P16150	SPN	Leukosialin	1.1E-03	0.19
Q14761	PTPRCAP	Protein tyrosine phosphatase receptor type C-associated protein 1	0.04	0.19
Q14242	SELPLG	P-selectin glycoprotein ligand 1	0.01	0.12
Q9Y3P8	SIT1	Signaling threshold-regulating transmembrane adapter 1	4.1E-03	2.5E-03
P01850		T-cell receptor beta chain C region	0.02	No data
P06729	CD2	T-cell surface antigen CD2	0.01	No data
P04234	CD3D	T-cell surface glycoprotein CD3 delta chain	2.8E-03	No data
P07766	CD3E	T-cell surface glycoprotein CD3 epsilon chain	1.4E-03	No data
P09693	CD3G	T-cell surface glycoprotein CD3 gamma chain	0.03	No data



# ANEXOS I. TABLAS SUPLEMENTARIAS

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	CD81 <sup>c</sup>	EWI-2 <sup>c</sup>
P20963	CD3Z	T-cell surface glycoprotein CD3 zeta chain	1.9E-03	No data
P01730	CD4	T-cell surface glycoprotein CD4	0.04	0.09
P06127	CD5	T-cell surface glycoprotein CD5	0.02	No data
P01732	CD8A	T-cell surface glycoprotein CD8 alpha chain	4.6E-03	0.19
<b>Metabolismo</b>				
P06733	ENO1	Alpha-enolase	2.7E-04	0.04
P01040	CSTA	Cystatin-A	0.04	0.22
P04080	CSTB	Cystatin-B	0.04	0.09
P04406	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	4.4E-04	3.6E-06
O95573	ACSL3	Long-chain-fatty-acid--CoA ligase 3	0.19	0.04
P62937	PPIA	Peptidyl-prolyl cis-trans isomerase A	0.19	0.04
P18669	PGAM1	Phosphoglycerate mutase 1	0.19	0.04
P14618	PKM2	Pyruvate kinase isozymes M1/M2	2.8E-03	9.7E-04
<b>Proteínas relacionadas con ácidos nucleicos</b>				
P13010	XRCC5	ATP-dependent DNA helicase 2 subunit 2	0.18	0.04
Q9HC52		Chromobox protein homolog 8	0.10	0.05
P42695	NCAPD3	Condensin-2 complex subunit D3	No data	0.04
P68104	EEF1A1	Elongation factor 1-alpha 1	0.01	2.4E-04
P13639	EEF2	Elongation factor 2	No data	0.04
Q9HAJ7	SAP30L	Histone deacetylase complex subunit SAP30L	0.15	0.05
P53990	KIAA0174	IST1 homolog	0.19	0.04
Q13310		Polyadenylate-binding protein 4	0.04	0.20
<b>Proteínas relacionadas con la vía ras</b>				
P60953	CDC42	Cell division control protein 42 homolog	0.01	0.01
Q96BY6	DOCK10	Dedicator of cytokinesis protein 10	0.05	3.1E-03
Q9H4E7	DEF6	Differentially expressed in FDCP 6 homolog	0.04	0.01
P62826	RAN	GTP-binding nuclear protein Ran	0.19	0.05
P46940	IQGAP1	Ras GTPase-activating-like protein IQGAP1	0.01	0.01
Q13576	IQGAP2	Ras GTPase-activating-like protein IQGAP2	0.37	0.01
P63000	RAC1	Ras-related C3 botulinum toxin substrate 1	1.6E-03	4.8E-03
P15153	RAC2	Ras-related C3 botulinum toxin substrate 2	0.01	0.03
P61006		Ras-related protein Rab-8A	3.1E-03	0.09
Q92930	RAB8B	Ras-related protein Rab-8B	0.02	0.01
P11233	RALA	Ras-related protein Ral-A	9.3E-05	2.0E-04
P62834	RAP1A	Ras-related protein Rap-1A	0.04	0.19
P10114		Ras-related protein Rap-2a	0.10	0.01
P61225	RAP2B	Ras-related protein Rap-2b	0.04	0.19
P52566	ARHGD1B	Rho GDP-dissociation inhibitor 2	No data	0.01
Q9HBH0	RHOF	Rho-related GTP-binding protein RhoF	0.19	0.04
Q9BXI6	TBC1D10A	TBC1 domain family member 10A	0.19	0.01
<b>Proteínas ribosomales</b>				
P62847	RPS24	40S ribosomal protein S24	No data	0.04
P62753	RPS6	40S ribosomal protein S6	No data	0.05
P62081	RPS7	40S ribosomal protein S7	No data	0.04
P62241	RPS8	40S ribosomal protein S8	0.19	0.03
P46781	RPS9	40S ribosomal protein S9	No data	0.05
P27635	RPL10	60S ribosomal protein L10	No data	0.04
P62913	RPL11	60S ribosomal protein L11	No data	0.04
P61353	RPL27	60S ribosomal protein L27	No data	0.05
P46776	RPL27A	60S ribosomal protein L27a	No data	0.01
P36578	RPL4	60S ribosomal protein L4	No data	0.04

# ANEXOS I. TABLAS SUPLEMENTARIAS

Nº Acceso	Símbolo IPA <sup>a</sup>	Categoría/Proteína <sup>b</sup>	CD81 <sup>c</sup>	EWI-2 <sup>c</sup>
Q02878	RPL6	60S ribosomal protein L6	No data	0.05
P62424	RPL7A	60S ribosomal protein L7a	No data	0.04
P62917	RPL8	60S ribosomal protein L8	No data	0.04
Q00839		Heterogeneous nuclear ribonucleoprotein U	No data	0.05
<b>Proteínas de suero</b>				
P12821	ACE	Angiotensin-converting enzyme	0.10	0.05
P61769	B2M	Beta-2-microglobulin	2.4E-04	0.04
P12259	F5	Coagulation factor V	0.04	0.09
P69905		Hemoglobin subunit alpha	1.5E-04	5.5E-04
P68871	HBB	Hemoglobin subunit beta	0.05	1.4E-03
P19823	ITIH2	Inter-alpha-trypsin inhibitor heavy chain	0.03	0.05
P00734	F2	Prothrombin	0.19	0.01
P22105	TNXB	Tenascin-X	0.04	8.8E-05
P04004	VTN	Vitronectin	0.01	3.6E-03
<b>Proteínas relacionadas con el tráfico celular</b>				
O00299	CLIC1	Chloride intracellular channel protein 1	0.07	0.01
Q9H4G4	GLIPR2	Golgi-associated plant pathogenesis-related protein 1	0.05	No data
P43007	SLC1A4	Neutral amino acid transporter A	1.8E-03	No data
Q15758	SLC1A5	Neutral amino acid transporter B(0)	0.01	No data
P05023	ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1	0.02	No data
Q8IU68	TMC8	Transmembrane channel-like protein 8	0.02	0.01
<b>Tubulinas</b>				
Q71U36	TUBA1A	Tubulin alpha-1A chain	0.02	2.9E-04
P68363	TUBA1B	Tubulin alpha-1B chain	0.19	0.01
P07437	TUBB	Tubulin beta chain	0.19	2.4E-04
Q9H4B7	TUBB1	Tubulin beta-1 chain	No data	7.0E-04
Q13885	TUBB2A	Tubulin beta-2A chain	4.8E-03	1.5E-03
P68371	TUBB2C	Tubulin beta-2C chain	0.09	1.3E-04
Q6ZT98		Tubulin polyglutamylase TTL7	0.06	0.05
<b>Desconocido</b>				
Q8N5I2	ARRDC1	Arrestin domain-containing protein 1	0.01	1.6E-03
Q9P265	DIP2B	Disco-interacting protein 2 homolog B	0.03	0.01
Q9H0Q0	FAM49A	Protein FAM49A	No data	0.04
Q6UWD8	C16orf54	Transmembrane protein C16orf54	0.01	1.4E-04
Q5TEA3		Uncharacterized protein C20orf194	0.04	0.04
<b>Transporte vesicular</b>				
P84077	ARF1	ADP-ribosylation factor 1	0.01	0.02
P62330	ARF6	ADP-ribosylation factor 6	3.7E-03	1.2E-03
P04083	ANXA1	Annexin A1	0.05	0.16
P50995	ANX11	Annexin A11	0.01	No data
P07355	ANXA2	Annexin A2	2.7E-03	0.01
P08133	ANXA6	Annexin A6	3.1E-03	0.09
Q00610	CLTC	Clathrin heavy chain 1	0.09	0.02
Q9H4M9	EHD1	EH domain-containing protein 1	0.01	1.1E-03
Q9H223	EHD4	EH domain-containing protein 4	0.01	0.02
Q14974	KPNB1	Importin subunit beta-1	0.09	0.02
Q8WV92	MITD1	MIT domain-containing protein 1	0.04	0.04
P14222		Perforin-1	0.19	0.04
Q9UNF0	PACSIN2	Protein kinase C and casein kinase substrate in neurons	0.04	No data
Q15833	STXBP2	Syntaxin-binding protein 2	0.09	0.04

<sup>a</sup> Referencia de las proteínas se tomó del software del IPA (<http://www.ingenuity.com>).

<sup>b</sup> Proteínas que interaccionan con CD81 y/o EWI-2.

<sup>c,d</sup> p-valor calculado por el test de Student. Los valores en rojo corresponden a las interacciones; la intensidad del color depende la significatividad con que la proteína interacciona. No data corresponde a la no identificación de dicha proteína para ese ligando.

# ANEXOS I. TABLAS SUPLEMENTARIAS

**Tabla suplementaria 4.** Cambios en la abundancia de proteínas detectados en exosomas de linfoblastos T de ratones CD81 +/- *Knock-out* respecto a los exosomas de animales sin deplecionar con  $|Zq| > 2.4$

Número <sup>a</sup>	Número de acceso	Nombre de la proteína <sup>b</sup>	$\log_2$ -ratio corregido <sup>c</sup>	Cambio de abundancia	$Z_q$ <sup>d</sup>	FDR <sub>q</sub> <sup>e</sup>	
1	Q64514	Tripeptidyl-peptidase 2	-1.05	2.07	UP	-9.35	0.00E+00
2	Q05CR0	LOC72520 protein (Fragment)	-0.61	1.53	UP	-4.22	8.82E-04
	P07724	Serum albumin	-0.51	1.43	UP	-2.99	0.07
	D4A157	RAT Putative uncharacterized protein Rpp14	-0.79	1.73	UP	-2.95	0.07
	O08614	utrophin	-0.37	1.29	UP	-2.79	0.10
	P62245	40S ribosomal protein S15a	-0.81	1.76	UP	-2.79	0.10
	Q8BPU7	Engulfment and cell motility protein 1	-0.80	1.74	UP	-2.75	0.11
	Q8CC06	Integrin alpha 6	-0.61	1.53	UP	-2.71	0.12
	Q99LU0	Charged multivesicular body protein 1b-1	-0.42	1.34	UP	-2.66	0.13
	P06332	T-cell surface glycoprotein CD4	-0.50	1.42	UP	-2.65	0.14
	Q60769	Tumor necrosis factor alpha-induced protein 3	-0.45	1.37	UP	-2.51	0.19
	Q9JHU9	Inositol-3-phosphate synthase 1	-0.59	1.51	UP	-2.46	0.21
	Q8BFZ3	Beta-actin-like protein 2	-0.53	1.44	UP	-2.44	0.22
	Q9DCH4	Eukaryotic translation initiation factor 3 subunit F	-0.65	1.57	UP	-2.42	0.21
	P50396	Rab GDP dissociation inhibitor alpha	0.43	1.34	DOWN	2.40	0.22
	P62137	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	0.27	1.21	DOWN	2.41	0.22
	P05213	Tubulin alpha-1B	0.35	1.27	DOWN	2.42	0.22
	Q99JI6	Ras-related protein Rap-1b	0.35	1.27	DOWN	2.43	0.21
	P10404	MLV-related proviral Env polyprotein	0.29	1.23	DOWN	2.46	0.21
	P62889	60S ribosomal protein L30	0.59	1.51	DOWN	2.61	0.14
	Q62178	Semaphorin-4A	0.39	1.31	DOWN	2.62	0.15
	P34884	Macrophage migration inhibitory factor	0.56	1.48	DOWN	2.80	0.10
	Q9D8N0	Elongation factor 1-gamma	0.36	1.28	DOWN	2.83	0.09
	Q9Z1Q5	Chloride intracellular channel protein 1	0.31	1.24	DOWN	2.86	0.09
	P47962	60S ribosomal protein L5	0.43	1.35	DOWN	2.88	0.08
	P84096	Rho-related GTP-binding protein RhoG	0.43	1.35	DOWN	2.91	0.08
	Q62422	Osteoclast-stimulating factor 1	0.80	1.75	DOWN	2.94	0.08
	Q61411	GTPase Hras	0.61	1.53	DOWN	3.03	0.06
3	P19437	B-lymphocyte antigen CD20	0.39	1.31	DOWN	3.17	0.04
4	Q9WTR1	Transient receptor potential cation channel subfamily V member 2	0.69	1.61	DOWN	3.23	0.04
5	P62814	V-type proton ATPase subunit B, brain isoform	0.69	1.62	DOWN	3.34	0.02
6	P14483	H-2 class II histocompatibility antigen, Abeta chain	0.40	1.32	DOWN	3.50	0.01
7	Q6MGB8	RT1 class I, A2	0.64	1.56	DOWN	3.65	0.01
8	Q05144	Ras-related C3 botulinum toxin substrate 2	0.51	1.42	DOWN	3.90	3.03E-03
9	Q9QUJ0	Transforming protein RhoA	0.55	1.47	DOWN	4.35	4.88E-04
10	P31254	Ubiquitin-like modifier-activating enzyme 1 Y	0.89	1.85	DOWN	4.66	1.17E-04
11	P01872	Ig mu chain C region secreted form	0.53	1.44	DOWN	4.70	1.10E-04
12	P01590	Interleukin-2 receptor subunit alpha	0.67	1.60	DOWN	4.74	9.18E-05
13	Q9QZ19	Serine incorporator 3	0.70	1.63	DOWN	4.92	3.70E-05
14	Q91V55	Ribosomal protein S5	0.89	1.85	DOWN	4.99	3.20E-05
15	P01898	H-2 class I histocompatibility antigen, Q10 alpha chain	0.74	1.67	DOWN	5.03	2.56E-05
16	Q64467	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	0.87	1.83	DOWN	5.08	1.99E-05
17	Q920N1	Eukaryotic translation initiation factor 2 subunit 3, X-linked	1.05	2.08	DOWN	5.10	2.09E-05
18	Q80UL9	Junctional adhesion molecule-like	0.96	1.94	DOWN	5.14	1.75E-05
19	P13597	Intercellular adhesion molecule 1	0.74	1.67	DOWN	5.19	1.81E-05
20	A8DUN2	Beta-globin	1.00	2.00	DOWN	5.74	8.12E-07
21	P14046	alpha-1-inhibitor 3	1.05	2.06	DOWN	5.96	2.13E-07
22	P16045	Galectin-1	1.20	2.30	DOWN	6.27	4.12E-08
23	P97351	40S ribosomal protein S3a	1.26	2.40	DOWN	6.51	8.48E-09
24	P62880	Guanine nucleotide-binding protein G(i)/G(s)/G(t) subunit beta-2	0.97	1.96	DOWN	6.81	2.28E-09
25	P60766	Cell division control protein 42 homolog	1.02	2.03	DOWN	7.08	3.41E-10
26	P01882	Ig delta chain C region membrane-bound form	1.79	3.47	DOWN	10.15	0
27	P35762	CD81 antigen	2.34	5.07	DOWN	12.95	0

<sup>a</sup> Las proteínas que muestran los cambios de abundancia estadísticamente significativas están numeradas como en la figura 25.

<sup>b</sup> Las proteínas señaladas en rojo corresponden a las proteínas ribosomales, las proteínas en verde representan las proteínas relacionadas con los dominios tetraspaninas, y las proteínas en azul representan las proteínas relacionadas con la vía ras. Estas proteínas son mostradas en la Figura 26.

<sup>c</sup> El  $\log_2$ -ratio se corrigió restando el valor de la gran media.

<sup>d</sup> Los valores normales estandarizados se obtienen dividiendo la proteína  $\log_2$ -ratios por su varianza. La magnitud del cambio de la variable estandarizada ( $Z_q$ ) se muestra de acuerdo a la escala de colores indicada en la parte superior de la tabla.

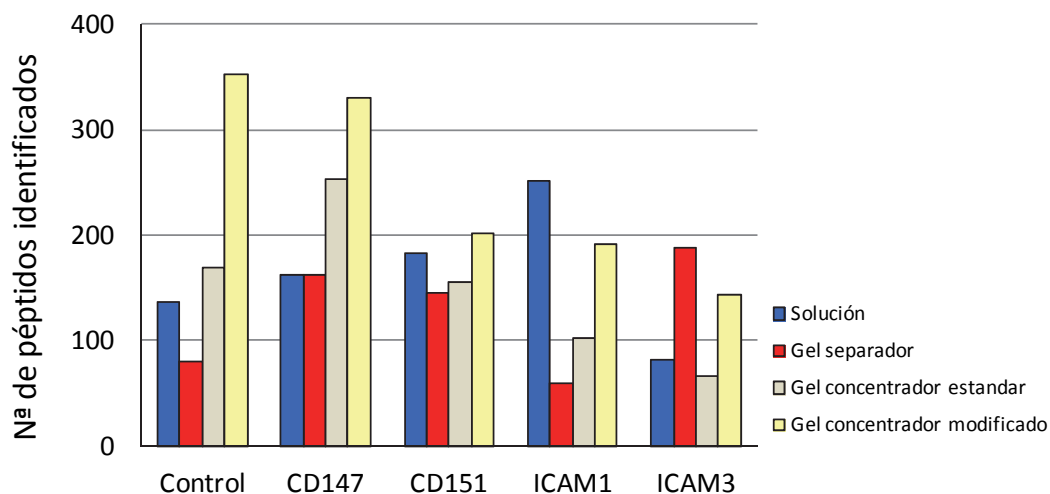
<sup>e</sup> Tasa de posibles falsos errores de cuantificación a nivel de proteínas (ver ref ()).



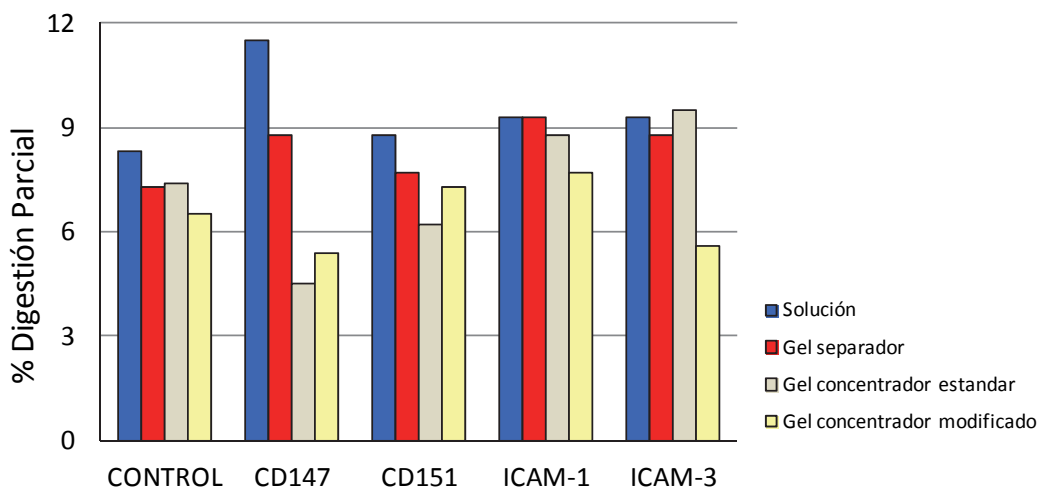
## **ANEXOS II. FIGURAS SUPLEMENTARIAS**



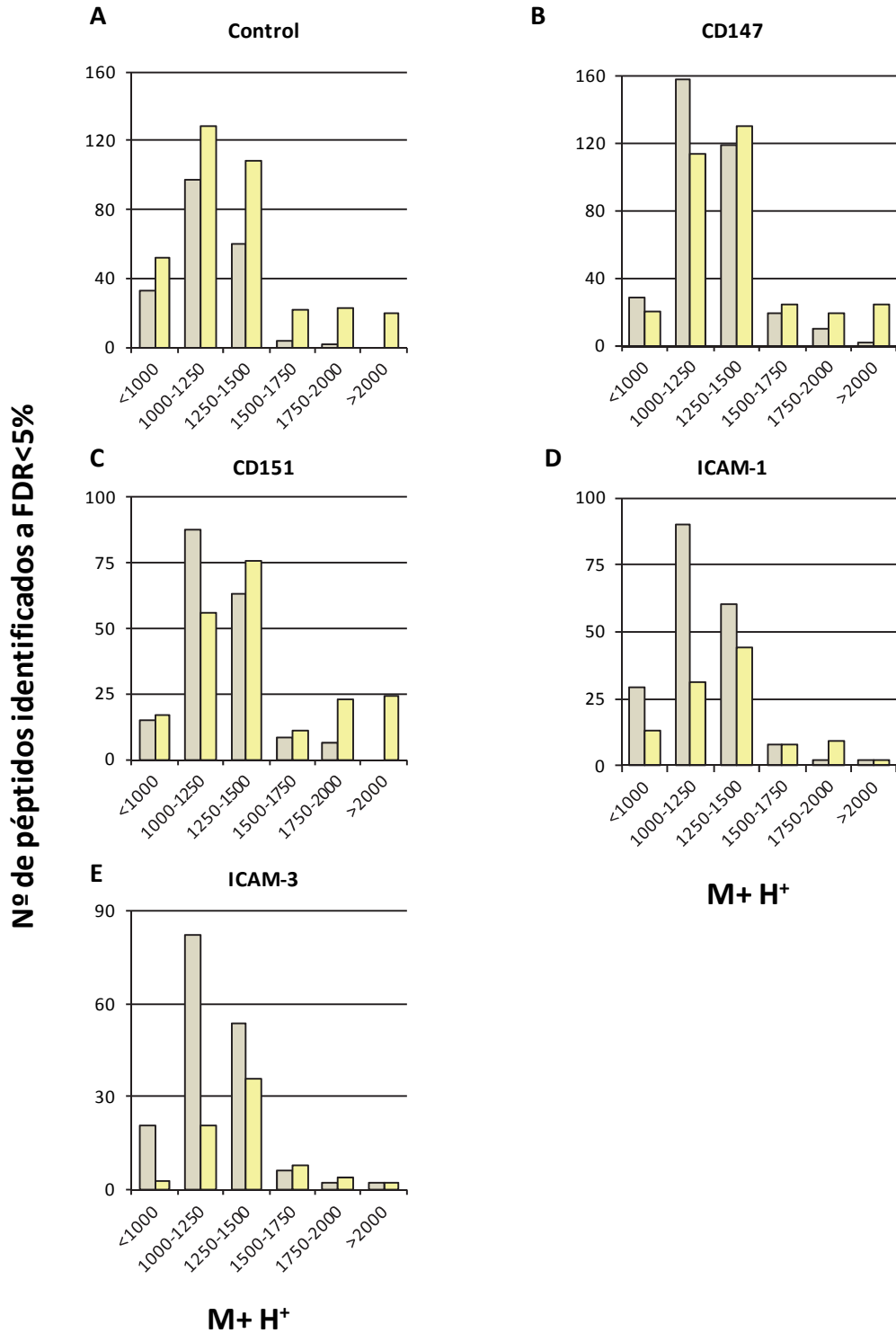
## ANEXOS II. FIGURAS SUPLEMENTARIAS



**Figura Suplementaria 1: Número de identificaciones peptídicas para cada uno de los ligandos.** Se digirieron las proteínas del mismo donante con las distintas aproximaciones metodológicas. El análisis por EM usado fue el mismo para todos los extractos (mismo gradiente RP-HPLC, misma exclusión dinámica, etc.), usando una tasa de error del 5% en la identificación. Se muestra que en 3 de los 5 anzuelos probados el número de péptidos identificados es mayor en la digestión en gel concentrador con el tampón modificado respecto a las otras técnicas. La digestión en gel separador y la digestión en solución son inferiores en número de péptidos.



**Figura Suplementaria 2: Porcentaje de digestión parcial de los péptidos identificados a una FDR del 5% en los distintos ligandos.** El cálculo se realizó, para cada ligando y en cada aproximación metodológica, dividiendo el número de péptidos únicos parcialmente digeridos por el número total de péptidos únicos identificados. La identificación de los subpéptidos se hizo con el programa bioinformático de análisis QuiXoT.



**Figura S3: Péptidos identificados en digestión en gel concentrador con el tampón de digestión estándar o modificado.** Se dividió el número de identificaciones por su M+H<sup>+</sup>, determinando que con el nuevo tampón se produce una recuperación de péptidos en la mayoría de rangos de masa con el tampón modificado. En el control (A), CD147 (B) y CD151 (C) este enriquecimiento es más significativo en los péptidos de mayor tamaño respecto a la digestión con tampón estándar.



## **ARTÍCULOS PUBLICADOS**

