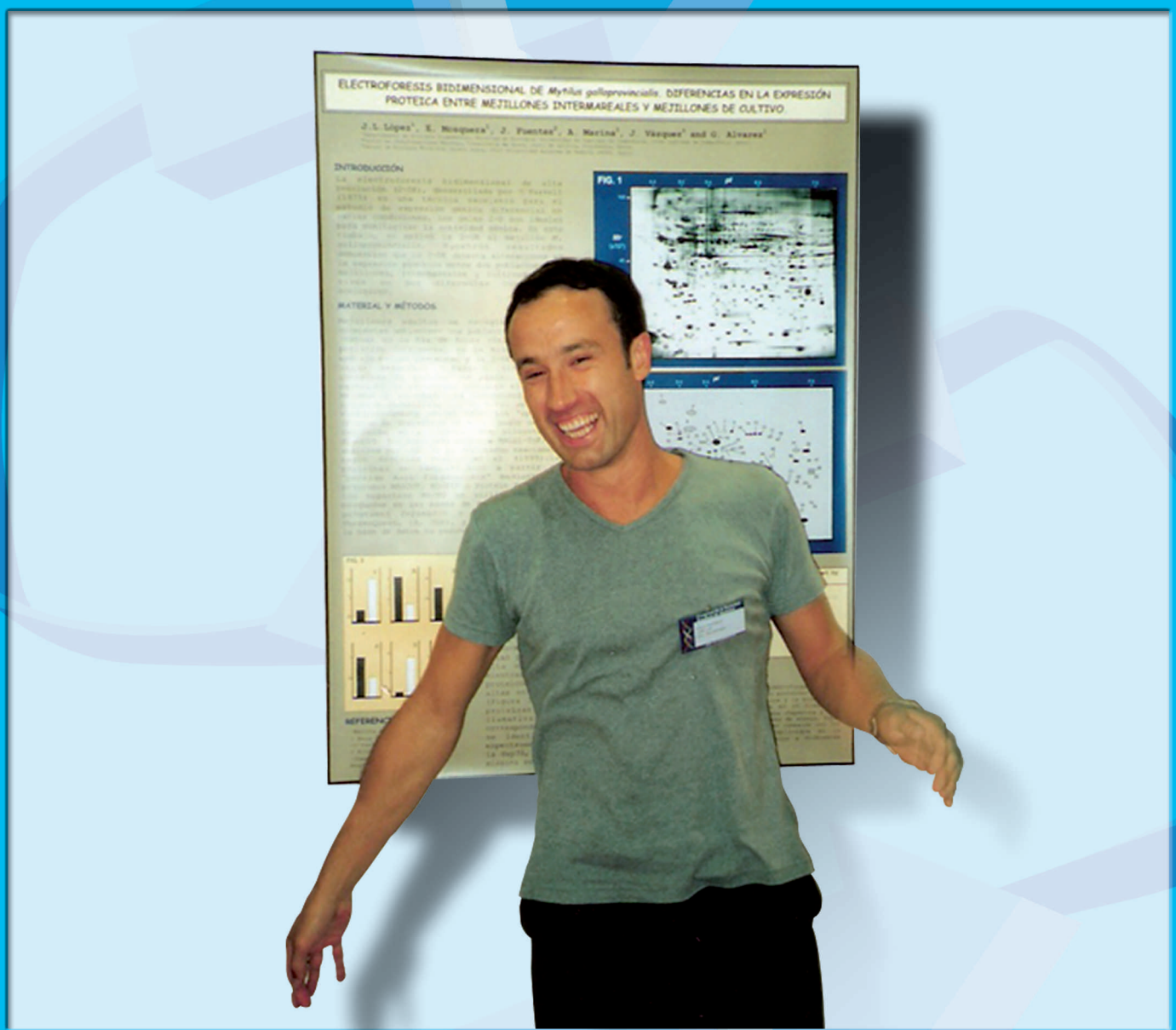
PROTEÓMICA

Revista de la Sociedad Española de Proteómica

http://www.cbm.uam.es/seprot/ Número 2 • Diciembre 2008



Servicio de Publicaciones
UNIVERSIDAD Ð CÓRDOBA

SEPROT
SOCIEDAD ESPAÑOLA DE PROTEÓMICA

Sede social: Instituto de Biomedicina de Valencia, C.S.I.C.
c/. Jaime Roig, 11. 46010 Valencia
Tel. 96 339 1778 • Fax 96 369 0800
C.I.F.: G-97465629.
Nº. Registro Nacional de Asociaciones: 584180.

http://www.cbm.uam.es/seprot

Junta Directiva:

Juan José Calvete.
Presidente
Instituto de Biomedicina de Valencia
C.S.I.C.
jcalvete@ibv.csic.es

Concha Gil.
Vicepresidenta
Departamento de Microbiología II.
Universidad Complutense de Madrid
conchagil@farm.ucm.es

Jesús V. Jorrín.
Secretario
Universidad de Córdoba
bf1jonoj@uco.es

David Andreu.
Tesorero
Universidad Pompeu Fabra. Barcelona
david.andreu@upf.edu

Juan Pablo Albar.
Vocal
Centro Nacional de Biotecnología
U.A.M.-C.S.I.C. Madrid
jpalbar@cnb.uam.es

Fernando J. Corrales.
Vocal
Universidad de Navarra
fjcorrales@unav.es

Ángela Moreno.
Vocal
Universidad de Córdoba-C.S.I.C.
ge1moloa@uco.es

Jesús Vázquez.
Vocal
Centro de Biología Molecular "Severo Ochoa". U.A.M.-C.S.I.C. Madrid
jvazquez@cbm.uam.es

PROTEÓMICA

Revista de la Sociedad Española de Proteómica

Número 2, Diciembre 2008

COMITÉ EDITORIAL:

Jesús V. Jorrín (Universidad de Córdoba)
Jesús Vázquez (CBMSO, Madrid)
Fernando Corrales (CIMA-Universidad de Pamplona)
Ángela Moreno (CSIC- Universidad de Córdoba)
Juan J. Calvete (IBV-CSIC, Valencia)
Ángel García (Universidad de Santiago de Compostela)
Antonio Martínez Ruiz (Hospital de la Princesa, Madrid)

CORRESPONDENCIA EDITORIAL:
Jesús V. Jorrín Novo. Dpto de Bioquímica y Biología Molecular, Universidad de Córdoba. Campus de Rabanales, Ed. Severo Ochoa (C6), 14071 Córdoba.
E-mail: bf1jonoj@uco.es.

EDITA:

Servicio de Publicaciones
UNIVERSIDAD Ð CÓRDOBA

PROTEÓMICA. Revista de la Sociedad Española de Proteómica

Editada por: Servicio de Publicaciones de la Universidad de Córdoba

Periodicidad: Semestral (2 números por año: enero-febrero y julio-septiembre).

Contenidos: se publicarán artículos originales, comunicaciones breves, artículos de revisión, artículos de difusión, opiniones, notas, comentarios sobre cualquier aspecto relacionado con la proteómica. Se priorizarán artículos originales sobre aspectos metodológicos o de aplicación al estudio de sistemas biológicos. Incluye información sobre nuestra Sociedad, personas, grupos e instituciones que la componen.

Idioma: será el castellano, aunque se admiten contribuciones en otras lenguas, preferentemente inglés.

Distribución: España y Latinoamérica, aunque pretendemos que tenga un carácter internacional mediante la distribución a otros países. Se enviarán, sin coste alguno, a los socios de la SEProt, Unidades y Servicios, así como a instituciones y organizaciones públicas o privadas miembros de la sociedad o con actividad relevante en el campo de la proteómica.

Publicación: habrá una versión impresa, editada por el Servicio de Publicaciones de la UCO, y una versión "on-line" que aparecerá en la página web de la SEProt y de la Universidad de Córdoba.

Editores:
Jesús V. Jorrín (Universidad de Córdoba)
Jesús Vázquez (CBMSO-CSIC, Madrid)
Fernando Corrales (CIMA-Universidad de Pamplona)
Ángela Moreno (CSIC- Universidad de Córdoba)
Juan J. Calvete (IBV-CSIC, Valencia)
Ángel García (Universidad de Santiago de Compostela)
Antonio Martínez Ruiz (Hospital de la Princesa, Madrid)

Correspondencia editorial:
Jesús V. Jorrín Novo. Dpto de Bioquímica y Biología Molecular, Universidad de Córdoba. Campus de Rabanales, Ed. Severo Ochoa (C6), 14071 Córdoba. E-mail: jornadas-proteomica@uco.es.

Instrucciones a los autores:
http://www.cbm.uam.es/seprot/

Envío de los manuscritos:
Mediante correo electrónico a:bf1jonoj@uco.es

I.S.S.N.: 1888-0096

Depósito Legal: CO-1005-07

Edita: Servicio de Publicaciones de la Universidad de Córdoba.
Campus de Rabanales. Ctra. Nacional IV, Km. 396. 14071 CÓRDOBA
Tlfnos.: 957 21 21 65. Fax: 957 21 81 96
http://www.uco.es/publicaciones
Correo electrónico: publicaciones@uco.es

Imprime: Argos Impresores S.L. Córdoba.

El Servicio de Publicaciones de la Universidad de Córdoba, a los efectos previstos en el artículo 32.1, párrafo 2º, del vigente TRLPI, se opone expresamente a que cualquiera de las páginas de *Proteómica*, o partes de ellas, sean utilizadas para la realización de resúmenes de prensa. Cualquier acto de explotación de sus contenidos precisará de la oportuna autorización, que será concedida por CEDRO mediante licencia.

ÍNDICE DE CONTENIDOS

Editorial
Jesús V. Jorrín Novo 6

La SEProt cumple 5 años
Juanjo Calvete 7

Noticias de la SEProt
Juanjo Calvete 9

El proceso de elección de la "nueva" Junta Directiva
Jesús V. Jorrín Novo 11

José Luis López: el amigo y el científico
Esther Mosquera, Anabel Marina, Jesús Vázquez, Jesús Jorrín 13

Introducción a la espectrometría de masas para la caracterización de péptidos y proteínas en proteómica
Joaquín Abián, Montserrat Carrascal, Marina Gay 16

La inevitable estadística… en el mundo de la proteómica
Ana María Rodríguez Piñeiro, María Páez de la Cadena, Francisco Javier Rodríguez Berrocal 36

El plasma ICP para la ionización elemental en Espectrometría de Masas: su aplicación en Proteómica Cuantitativa
Alfredo Sanz-Medel 47

Estratégias proteómicas para el análisis de grupos carbonilo como marcadores de daño oxidativo en proteínas
Verónica Irazusta, Armando Moreno-Cermeño, Joaquim Ros, Jordi Tamarit 51

Aplicación del iTRAQ para el análisis de proteínas de membrana implicadas en la glicosilación en levaduras
Raquel Martínez-López, Pilar Ximénez-Embún, Dolores Gutiérrez, Lucía Monteoliva, J. Arroyo, María Molina, Concha Gil 59

Analysis of thylakoid membrane protein complexes from maize by Blue-Native Gel Electrophoresis and Tandem Mass Spectrometry
A. Campos, A. Coelho, L. Arike, R. Santos, C. Franco , P. Carvajal-Vallejos, M.A. Santos, J.M. Torné, A.V. Coelho 60

Proteomics as a complementary tool for identifying unintended side effects occurring in transgenic maize seeds as a result of genetic modifications.
Maria G. Egidi, Anna Maria Timperio, Sara Rinalducci, Giovanni Monastra, Lello Zolla 62

Lista de términos y expresiones técnicas en inglés y su traducción recomendada al castellano
Joaquín Abián, Jesús Vázquez 63

Instrucciones a los Autores 65

Poetómica
Jesús Vázquez 68

EDITORIAL

Nuestra más sincera y merecida muestra de cariño a José Luis López.

El pasado 12 de octubre falleció José Luis López, uno de los pioneros de la proteómica en nuestro país. Desde esta modesta revista queremos rendirle un más que merecido homenaje. Su humanismo superó con creces su excelente trabajo como Profesor Universitario e Investigador, de lo que dan fe dos de las cartas publicadas en este número y de las que son autoras Esther Mosquera y Anabel Marina, quizás las personas que mejor le conocieron en la doble faceta, la personal y la profesional.

La vida sigue, y nuestra Sociedad continúa con sus múltiples actividades. Estamos en periodo de renovación de la Junta Directiva, y la elección del nuevo equipo Directivo tendrá lugar en el próximo Congreso de la SEProt de Pamplona. A pesar de su juventud, la sociedad la podemos considerar consolidada, tal y como refleja Juanjo Calvete en su carta de despedida como presidente. El número de socios en la actualidad ronda los 200, una cifra impensable hace un par de años. No obstante, y tal y como se indica en el artículo "Noticias de la SEProt" existen importantes retos y proyectos que abordar en este segundo mandato. Entre estos proyectos hay que incluir el de la consolidación de la revista, cuyos números ven la luz con cierto retraso y gracias al trabajo de sólo unos pocos. Desde aquí hacemos una llamada a todos los socios para que nos envíen contribuciones y a la vez agradecemos los trabajos y revisiones recibidas. Entre estas incluimos la excelente revisión sobre espectrometría de masas, firmada, como no podía ser de otra manera, por Joaquín Abián, y a la que han contribuido Montse Carrascal y Marina Gay. He de reconocer que el nombre de Joaquín Abián surge frecuentemente en nuestras reuniones de la Junta Directiva ya que le consideramos como una de las personas claves en el quehacer de la SEProt. Aunque no esté entre los miembros de la Junta Directiva, siempre ha atendido nuestras peticiones de colaboración. Muestra de ello es no sólo la revisión mencionada, los cursos que imparte, y últimamente, la elaboración de una lista de términos y expresiones técnicas en inglés y su traducción recomendada al castellano, preparada en colaboración con otro "insustituible", mi querido amigo Jesús Vázquez. Este documento será, sin duda alguna, de gran valor y utilidad, en especial a la hora de redactar tesis Doctorales.

La aplicación correcta de la estadística y métodos estadísticos es uno de los fallos más generales en un buen número de artículos de proteómica publicados hasta la fecha. Los evaluadores de nuestros trabajos y la política editorial de las revistas del área son más estrictas a este respecto, tal y como se recoge en los documentos MIAPE. La revisión firmada por Ana María Rodríguez Piñeiro como primera autora será de gran ayuda a la hora de diseñar y analizar correctamente los experimentos y datos, necesario para interpretar y validar correctamente, y desde un punto de vista biológico, nuestros datos y conclusiones.

En este número aparecen, por último, dos excelentes revisiones, la primera, firmada por Alfredo Sanz Medel, sobre ICP masas, un área muy novedosa, y la segunda por Jordi Tamarit y colaboradores, sobre el análisis de grupos carbonilo como marcadores de daño oxidativo en proteínas.

En la contraportada aparece, como es habitual, la sección poetómica de Jesús Vázquez. ¿Quién habla de la separación entre ciencias y letras, técnica y humanismo? Conociendo a personas como Jesús Vázquez y Juanjo Calvete, diríamos que no son apuestas alternativas, sino que van de la mano. Cuando sea mayor quiero ser como vosotros.

Finalmente, dar la bienvenida a los dos nuevos miembros del equipo editorial de la revista Ángel García (Universidad Santiago Compostela) y Antonio Martínez Ruiz (Hospital de la Princesa, Madrid). No sabéis donde os habéis metido; en cualquier caso, gracias.

Jesús V. Jorrín Novo

LA SEProt cumple 5 años

El próximo trece de abril (Acta Fundacional) o nueve de Diciembre (Registro de Asociaciones) nuestra Sociedad cumple 5 años. Parece que fue ayer pero hace ya cuatro años que la actual Junta Directiva, la primera elegida en Asamblea General durante el 1er Congreso de la SEProt celebrado en Córdoba (13-17 Febrero 2005), dirige la marcha de nuestra Sociedad. Atrás queda también el 2º Congreso (Valencia, 10-14 2007) que consolidó la decidida vocación internacionalista de la SEProt, tanto en el plano europeo -contribuyendo actualmente de manera destacada en los órganos de dirección y comités de la Federación Europea de Sociedades de Proteómica- como en el ámbito latinoamericano. El tercer Congreso que este 2009 celebramos en Pamplona tiene también un punto de mira en el fortalecimiento de las relaciones entre países separados por el Océano Atlántico pero culturalmente cercanos. Hubo intentos de acercamiento anteriores, incluyendo la posibilidad de creación de una red multinacional de temática proteómica que no prosperó por reticencias de alguna Sociedad de Proteómica a un supuestamente encubierto "neocolonialismo cultural". No caeré en la tentación de comentar este sinsentido, aunque tampoco cejaré en promover el fortalecimiento de lazos proteómicos entre ambos lados del charco... En nuestro país, en este primer lustro de existencia de nuestra Sociedad, se han consolidado algunos programas de educación a través de becas de asistencia a cursos nacionales e internacionales de formación en el ámbito de la Proteómica, merced a la saneada economía de la SEProt, consecuencia directa de la excelente gestión por parte de nuestro Tesorero David Andreu de los ingresos por cuotas individuales y corporativas.

Reconocer la excelencia investigadora premiando contribuciones de especial relevancia es una de las tareas más gratas de una Sociedad científica. Así lo han entendido también las empresas del ramo que han apoyado esta iniciativa patrocinando diversos galardones. Gabriel Padrón (Centro de Ingeniería Genética y Biotecnología, La Habana, Cuba) recibió en Valencia el Premio SEProt-Promega "por su contribución al desarrollo de la espectrometría de masas y su aplicación a la revolución proteómica". Applied Biosystems patrocinó los Premios SEProt-ABI para pósters presentados por grupos no españoles. Durante el 3er Congreso se otorgarán los Premios SEProt a un artículo publicado durante el bienio 2007-2008 y a un trabajo original presentado en forma de panel al Congreso. Esta tercera edición, al igual que las anteriores, está patrocinada por Bruker Biosciences Española S.A. El 3. Premio SEProt está dedicado a la memoria de José Luís López Rodríguez, entrañable y excelente proteómico de *Mytilus*, de la Universidad de Santiago de Compostela, que nos dejó tempranamente el pasado 12 de Octubre. Sus muchas virtudes han sido comentadas por quienes más le conocieron en el portal electrónico de la SEProt. Dando su nombre a nuestro galardón queremos también rendir un sentido y público homenaje institucional a nuestro compañero José Luís.

Durante el pasado congreso, nuestra Sociedad otorgó membresía de Honor a John B. Fenn, galardonado con el Premio Nobel de Química en 2002 por su trabajo pionero en el desarrollo de la ionización de macromoléculas mediante electronebulización (Proteómica 0:11-14, Julio 2007). Otra técnica esencial de la Proteómica es la ionización mediada por radiación láser y asistida por una matriz (MALDI). Este método de ionización fue desarrollado simultánea- e independientemente por Koichi Tanaka (Shimadzu, Japón) y por Franz Hillenkamp y Michael Karas (a la sazón en la Universidad de Münster, Alemania) durante los años 80 del pasado siglo. Los avatares del destino, quizás forzados por la presión de la poderosa compañía japonesa, hicieron que Tanaka compartiera con Fenn el Premio Nobel que la Academia Sueca negó a los alemanes. Nuestra Sociedad, nombrándoles Socios de Honor, ha querido sumarse a la lista de Instituciones que reconociendo esta injusticia histórica han rendido honores a Hillenkamp y Karas. Me alegra y enorgullece también que Peter Roepstorff, pionero de la espectrometría de masas biológica, testigo y actor privilegiado de la transición Química de Proteínas-Proteómica, y maestro de proteómicos, haya aceptado ser Miembro de Honor de la SEProt.

El Artículo 25c de nuestros Estatutos especifica que "el nombramiento de los Socios de Honor co-

rresponde a la Asamblea General, a propuesta de, al menos, diez Socios Ordinarios. El nombramiento se refrendará en votación secreta, debiendo alcanzar un número de votos favorables de, al menos, dos tercios de la totalidad de los miembros de la Junta Directiva. La decisión favorable deberá ser refrendada por la Asamblea General por mayoría simple". Quisiera desde esta tribuna proponer la candidatura a Socios de Honor de dos pioneros de las técnicas proteómicas, Jasna Peter-Katalinić (http://impb.klinikum.uni-muenster.de/research/peterkatalinic/index.html) y Pier-Giorgio Righetti (Electrophoresis 27 (8): 1435 - 1692 (No. 8 April 2006) Special Issue: Dedicated to Professor Pier Giorgio Righetti at his 65th Birthday). Jasna ha contribuido durante los últimos 25 años al desarrollo de la glicoproteómica; Pier-Giorgio, prolífico e imaginativo innovador de las técnicas de separación y su aplicación a la proteómica, ha publicado más de 450 artículos originales. A él debemos la invención de las tiras de gradiente de pH inmovilizado o la reciente técnica de "democratización" de proteomas complejos o ProteoMiner comercializado por BioRad.

Suenan tambores que anuncian la necesidad de renovación. Y no sólo porque así lo marquen nuestros Estatutos, sino porque el funcionamiento de toda Sociedad se rige por ciclos que garantizan el progreso de las ideas. Seguimos siendo una Sociedad joven, muy joven, que necesita el apoyo de todos sus socios. La Junta Directiva que surja de la Asamblea General en Pamplona va a necesitar el apoyo de *proteómicos* experimentados y aprendices de brujo para seguir contribuyendo al desarrollo de nuestra disciplina, en España y fuera de nuestras fronteras. Muchos han sido los logros conseguidos en este quinquenio, pero muchos siguen siendo los proyectos iniciados que aun necesitan el espaldarazo de la consolidación. En Julio del 2007 vio la luz el número 0 de la revista Proteómica y unos meses después (Febrero del 2008) se publicó el número 1 con ocasión de las Primeras Jornadas Bienales de Proteómica (Sitges, 21-22 Febrero 2008). Estas iniciativas, revista y Jornadas, han tenido como principales valedores a los *Jesuses*, Jorrín y Vázquez, respectivamente. Conociendo el entusiasmo que siempre han demostrado hacia la SEProt, seguro que no yerro en el convencimiento de que ambos seguirán brindando su colaboración en esos proyectos de capital importancia para la SEProt.

Aprovecho esta ocasión para despedirme como Presidente de la SEProt y desearle al nuevo Presidente, que será elegido por la Asamblea General durante el 3er Congreso, las mismas satisfacciones, sensación de compañerismo que yo he experimentado durante estos cuatro años al frente de la Junta Directiva. A Ángela, Concha, David, Jesús J., Jesús V., José Manuel, Juan Pablo, y Fernando, mi reconocimiento y gratitud por vuestra dedicación y apoyo. ¡Haber trabajado con vosotros por nuestra Sociedad compensa con creces el tiempo y esfuerzo robado a otras actividades! A la nueva Junta Directiva le esperan nuevos retos, como la tarea de profesionalizar la Secretaría Técnica de la Sociedad, incluyendo asesoría fiscal, y la edición de un libro en castellano sobre Proteómica. Todos los que impartimos Cursos de Doctorado, de Licenciatura, de especialización, Master, etc., seguramente coincidiremos en la necesidad de disponer de una monografía en nuestro idioma que recoja la historia, el desarrollo de la tecnología, así como los principios y aplicaciones de la Proteómica. Podemos esperar a que alguien lo haga o ponernos manos a la obra. Somos muchos los que pensamos que podemos y debemos hacerlo desde la SEProt. Como siempre, cualquier sugerencia y ayuda serán más que bienvenidas. ¿Seguimos hablando en Pamplona?

Juan J. Calvete

Noticias de la SEProt

Del día a día de la Sociedad recibís puntualmente información a través de correos electrónicos y de la página web (http://www.cbm.uam.es/seprot), incluyendo noticias sobre cursos, congresos, talleres y otras reuniones científicas, ofertas del trabajo, convocatorias de becas. Seguidamente, llamaremos vuestra atención sobre algunos de los aspectos que consideramos claves para la buena marcha y funcionamiento de la Sociedad. Dichos aspectos han sido tratados en detalle en las tres reuniones que la Junta Directiva ha tenido desde que salió a la luz el número 1 de la revista. La primera tuvo lugar en Sitges, el 20 de febrero, durante las I Jornadas Bienales; la segunda en Valencia, el 14 de Junio, y la tercera y última, en Madrid, el pasado 22 de diciembre.

Vaya por delante nuestra felicitación a las organizadoras de las I Jornadas Bienales, Monte Carrascal y Marina Gay, por el éxito de la reunión, y a todos los participantes por la calidad de sus contribuciones. Cualquier duda existente respecto al interés y conveniencia del proyecto ha quedado disipada. Seguro que hay detalles que mejorar, tales como el incrementar el tiempo dedicado a mesas redondas y sesiones de discusión, pero las bases de las futuras reuniones están sentadas y los próximos organizadores tienen ante sí un reto difícil de superar. Por añadidura, las Jornadas han tenido "superávit", lo que siempre viene bien a las arcas.

Siguiendo con el apartado de felicitaciones y agradecimientos, le toca el turno a nuestro compañero y amigo Fernando Corrales, así como a su grupo, por la excelente organización del "Tercer Congreso de la SEProt". El programa, contenidos, y demás detalles aparecen recogidos en la página web (http://www.unav.es/proteomicsmeeting/index.html). La participación de grupos latinoamericanos, la organización del curso precongreso, y la asistencia de científicos de gran relevancia han sido auténticos logros. Entre estos últimos cabe mencionar a los premios Nobel Franz Hillenkamp y Michael Karas, así como a Peter Roepstorf, quienes han sido propuestos como Socios de Honor de la Sociedad. Desde aquí nuestro agradecimiento a todas las instituciones públicas y privadas que van a colaborar en el evento.

Es hora de ir pensando dónde va a tener lugar el IV Congreso de la SEProt. Os animamos a que presentéis, a nivel individual o institucional, candidaturas. En Galicia, Cataluña, Madrid, y otros puntos de la geografía española hay excelentes grupos que serían candidatos idóneos.

En esta sección queremos reconocer la labor realizada por FUNDECOR-UCO, concretamente por María Teresa Montero, a cargo de la Secretaría Técnica. ¡Ojala podamos seguir contando con ellos en el futuro próximo!

El próximo mes de febrero, coincidiendo con nuestro Congreso, tendrá lugar la votación para la renovación de la Junta Directiva. Desde aquí animamos a todos los socios a que presenten candidatura. La nueva Junta Directiva tiene, ante sí, importantes retos, entre los que habría que citar, bajo nuestro punto de vista, y como más relevantes:

i) la consolidación de la revista Proteómica;

ii) la edición de una monografía en castellano sobre proteómica en la que se incluyan: aspectos históricos, fundamentos, metodologías, protocolos y aplicaciones;

iii) la celebración del próximo Congreso y las Jornadas Bienales;

iv) potenciar la relación con grupos latinoamericanos de proteómica.

El llevar a buen puerto y consolidar dichos proyectos depende en gran medida no tanto de quien asuma el reto de la dirección de la Sociedad, sino de la participación activa de los socios, algo que por ahora dista mucho de lo que sería deseable. Es por ello que desde estas líneas hacemos una llamada a todos los socios, sin cuya colaboración la Sociedad carece de sentido y significado.

Son muchos los ejemplos que se pueden citar en favor de la relevancia de la Sociedad y del buen hacer de la actual Junta Directiva, entre ellos, la participación de sus miembros en la Dirección de la EuPA y sus distintos Comités, en ProteoRed, en el Comité Editorial del Journal of Proteomics,

en la organización de un buen número de cursos y seminarios, así como de diferentes proyectos de investigación nacionales e internacionales. El reconocimiento de nuestra Sociedad queda reflejado, de nuevo, en nuestra presencia significativa en la renovados Comités de la EuPA. Esto no es ni más ni menos que un ejemplo anecdótico del excelente nivel que la proteómica está alcanzando en nuestro país, con cada día más grupos implicados, con un aumento en la membresía, en el número de publicaciones y su calidad. Lo podremos constatar durante el Congreso de Pamplona, y de ello da fe el número y calidad de artículos publicados en el último año que optan al "Premio de la SEProt José Luis López"

Una de las cosas que ha llamado negativamente nuestra atención es el poco eco que ha tenido, a pesar de su difusión, la convocatoria de becas SEProt. Tal es así que en, hasta la fecha, solo hemos recibido una solicitud. Esta Junta Directiva siempre ha tenido como prioridad ayudar a los socios, en especial a los más jóvenes, con medidas tales como la del programa de becas. Pero como dice el chiste, no te tocará la lotería si no compras un número. Entendemos que afortunadamente, hoy en día, existen muchas convocatorias de ayudas y quizás en mayor cuantía que la nuestra. Seguiremos con nuestro programa de becas y en la medida de nuestras posibilidades trataremos de aumentar la asignación.

Junta Directiva de la SEProt

El proceso de elección de la "nueva" Junta Directiva

Según nuestros estatutos[1], el manwdato de los miembros de la Junta Directiva tendrá una duración de cuatro años, por lo que la renovación de los cargos ha de tener lugar el próximo mes de febrero de 2009. Es por ello que, la actual Junta Directiva, en su reunión del 4 de julio de 2008 (IBV-CSIC, Valencia), acordó la convocatoria de dichas elecciones, lo que se comunicó, mediante correo electrónico, a todos los socios. Tendrán lugar en Pamplona durante el III Congreso de la SEProt.

Se ha establecido el siguiente calendario:

Fecha tope	*Fase de las elecciones*
10 de julio de 2008	Comunicado a los socios
10 enero 2009	Envío, por parte de los socios, de candidaturas a la Secretaría
20 de enero de 2009	Proclamación de candidaturas
21 de enero de 2009	Envío de candidaturas a los socios
11-13 de febrero de 2009	Elección, que tendrá lugar durante el 3º Congreso de la SEProt (Pamplona)

En la votación participarán todos los socios de la SEProt que estén al corriente en la cuota de pago. Se podrá votar mediante los siguientes sistemas:

1. En persona, y durante el Congreso. Se pondrá a disposición de los socios la papeleta de votación y se habilitará una urna sellada, que se ubicará en la mesa asignada a la SEProt. Para votar será necesario presentar el DNI y el carné de socio. El plazo máximo para votar será de 2 horas antes de que empiece la Asamblea General.

2. Mediante correo electrónico dirigido al Secretario (bf1jonoj@uco.es). Se anexará al correo copia del DNI y del carné de socio. En el asunto se pondrá: "Elecciones SEProt". En el texto se indicará una de las siguientes alternativas:

   i) doy mi voto a la candidatura presentada por, seguida del nombre del candidato;
   ii) me abstengo en la votación;
   iii) voto en blanco.

   El plazo máximo para votar por correo electrónico será el 7 de febrero de 2009.

3. Mediante correo ordinario. En un sobre cerrado, y firmado en la solapa, se introducirá una hoja en blanco en el que se indicará una de las siguientes alternativas:

   i) doy mi voto a la candidatura presentada por, seguida del nombre del candidato;
   ii) me abstengo en la votación;
   iii) voto en blanco.

[1] La **SEProt** será dirigida, representada y administrada por una Junta Directiva formada, al menos, por un Presidente, un Vicepresidente, un Secretario, un Tesorero y dos vocales. Todos los cargos que componen la Junta Directiva son gratuitos, serán designados y revocados por la Asamblea General Extraordinaria y sus mandatos tendrán una duración de 4 años.
Nadie podrá ser reelegido para el mismo cargo en dos cuatrienios consecutivos.
La Junta se renovará por mitades cada dos años, alternando el cese de Presidente, Tesorero y la mitad de los vocales con los de Vicepresidente, Secretario y la otra mitad de los vocales. Los candidatos podrán ser propuestos por cualquier Socio Ordinario siempre que su petición sea apoyada por la firma de otros cuatro Socios Ordinarios y se reciba en la Secretaría, al menos treinta días antes de la votación correspondiente. Asimismo la Junta Directiva podrá proponer candidatos para las vacantes existentes. La relación de candidatos que cumplan los requisitos anteriores será cursada por la Secretaría a todos los socios con voto, al menos veinte días antes de la Asamblea General. Durante ésta se elegirá entre los asistentes una mesa electoral constituida por un Presidente y dos Vocales, de los que el más joven actuará como secretario de la mesa. La votación será secreta, considerándose los votos recibidos por correo y serán proclamados los candidatos que hayan obtenido en cada caso el mayor número de votos. La Junta Directiva anunciará, al menos con dos meses de antelación, la convocatoria para renovar los cargos de la misma, así como los de la Comisión de Admisiones cuando corresponda. A dicha convocatoria se acompañará un calendario electoral. Para facilitar el planteamiento de su labor futura, los cargos de Presidente y Secretario se elegirán con dos años de antelación sobre el que corresponda a la renovación efectiva. El Presidente y Secretario Electos, durante el bienio comprendido entre su elección y su toma de posesión, podrán asistir a las reuniones de la Junta Directiva con voz, aunque sin voto.

Dicho sobre se introducirá en otro, acompañado de copia del DNI y del carné de socio. Se firmará en la solapa. Dirigir la carta a:

Jesús V. Jorrín Novo, Secretario de la SEProt. Dpto. Bioquímica y Biología Molecular, Universidad de Córdoba. Campus de Rabanales, Ed. Severo Ochoa, planta baja.14071 Córdoba

Se contabilizarán los votos que se reciban, como plazo máximo, el 7 de febrero de 2009.

El proceso de recuento de los votos y la proclamación de la candidatura más votada tendrá lugar durante la Asamblea General.

José Luis López: el amigo y el científico

Para muchos de nosotros el 12 de octubre de este año será un día difícil de olvidar, porque fue el día en que un gran amigo nos dejó. Empecé a trabajar con José Luís cuando la proteómica no era aún la proteómica, cuando mirar a los geles de electroforesis bidimensional era, tal como decía Hochstrasser, como mirar a las estrellas ("It's like looking at the sky, you see the stars but you don't know which ones they are"). Pero José Luís no dejó de mirar las estrellas, por el contrario, hizo que cada vez fuésemos más los que no podíamos dejar de mirarlas.

Jose tenía visión, tenía tesón y sabía contagiar como nadie la ilusión por las cosas. Nos contagió a muchos la pasión por la proteómica cuando estaba de moda y cuando todavía no lo estaba también. Y así fue enganchando a ese tren cada vez a más gente, y se fue encontrando por el camino con gente maravillosa, muchos de vosotros, que seguíais el mismo camino.

Os aseguro que los comienzos fueron complicados pero Jose también era muy constante, por eso no se rindió a pesar de que las cosas no le fueron fáciles. Ese era su espíritu "Coe", porque Jose fue un buen atleta del medio-fondo como deportista, fue un buen atleta de la proteómica como investigador y fue un gran atleta de la vida como persona. Ninguna adversidad, ni siquiera su enfermedad durante este año y medio pudo erosionar ni un ápice su ilusión, su nobleza, sus ganas de vivir. Ha sido tremendamente positivo y valiente hasta el final, os lo aseguro.

Desde aquellos inicios en la segunda mitad de los 90 hemos alcanzado logros, superado dificultades, nos hemos reído tanto, hemos compartido tantos momentos, tantas alegrías (también alguna que otra tristeza)… y no sólo a nivel profesional: os aseguro que he aprendido de Jose y con Jose mucho más que proteómica y sobre todas las cosas me ha enseñado cómo un compañero puede convertirse en un amigo y un amigo puede convertirse en un hermano.

Estoy convencida de que a todos aquellos que le habéis conocido en mayor o menor medida, no os resultará difícil reconocerle en mis palabras. Siempre he dicho que tenía una tendencia natural a engancharse a las ilusiones. Era una persona alegre y vital a la que le gustaba la gente, era el fan número uno de sus amigos y sobre todo adoraba la vida. Tenía además una nobleza absolutamente inusual. Todo ello hacía de José Luís una persona dotada de un magnetismo especial. Era, como muchos de vosotros me habéis dicho, muy fácil de querer, era casi imposible no cogerle cariño, a poco que se le conociese.

Y ¿sabéis qué os digo? Será muy difícil de olvidar el 12 de Octubre de este año pero también por fortuna, durante todo el tiempo que pudimos disfrutar de su energía y su vitalidad, de su pasión por la vida, por sus amigos, por la proteómica, por la investigación, por los Macintosh, por el atletismo, por tantas y tantas cosas, José Luís nos dejó la vida tan llena de buenos recuerdos que lo que sin duda no podremos hacer jamás es olvidarle a él. ¿A que no podéis pensar en Jose y no esbozar una sonrisa?

Esther Mosquera
(esther_mosquera@merck.com)

Me han pedido que escriba unas palabras en recuerdo a mi José (mi Josiño), pero es tan difícil… y no porque no tenga miles de cosas que decir sino porque nunca se me pasó, ni por lo más remoto de mi imaginación, hablar de él en su ausencia.

Yo soy una de las afortunadas que le conoció y una de las privilegiadas que puede decir que fui su amiga, de las que le quisieron, de las que fui querida. Hace 9 años ya que durante mi anterior baja maternal (es paradójico que ahora también esté en la misma situación) me escapase para colaborar en uno de los primeros cursos teórico-prácticos en Proteómica que impartíamos en nuestro Servicio, a las órdenes de mi querido jefe Jesús Vázquez. Por

entonces apenas reparé en los alumnos porque mi hermana pequeña se encontraba muy enferma en el hospital y mi cabeza andaba en otros Lares… aún así veo a José en clase consultando dudas y haciendo preguntas. Esto fue solo el principio de una fructuosa relación profesional y una maravillosa relación personal.

Todos sabéis de sus "manos" para confeccionar geles bidimensionales, habilidad que transmitió a sus pupilos y con los que seguimos trabajando. Pero esta era solo una habilidad laboral, sus habilidades y virtudes personales estaban muy por encima de la media. José era una gran persona en el más amplio sentido de la palabra, modesto, educado, respetuoso, detallista, elegante y un largo etc. de calificativos que formaban parte de él de modo innato, natural. Era un "señor". Era mi amigo, pero no un amigo cualquiera, fue el amigo que estuvo a mi lado, muy, muy cerca, más en mis malos que en mis buenos momentos, apoyándome siempre, alabándome, mimándome, jamás me abandonó un instante en estos 9 años tan especiales. Yo era su amiga la de "zumosol", así le gustaba llamarme cuando necesitaba de mis consejos proteómicos; le encantaba que le explicara los fundamentos de la ionización por electrospray, con las repulsiones coulómbicas que generan gotas cada vez más pequeñas, o la transferencia de protones de la matriz al péptido en la ionización MALDI. Recuerdo aquel verano en la playa, en Galicia, pinto rojeando folios que guardaba como un tesoro. Con él vi. la puesta de sol en Finisterre, conocí Galicia, su preciosa tierra, acompañado por Sara, su perra fiel, adorable, como todo lo que le rodeaba.

Han sido tantos buenos momentos, tantos recuerdos, tantas vivencias compartidas…recuerdas Esther?, qué bien lo pasamos en su oposición con aquel miembro del tribunal al que tuvimos que acompañar por todo Santiago a comprar un peluche para su hija!!!

¡Qué oposición tan dura le tenía preparada la vida! Esta también la superó con "nota". Nos dejó a todos perplejos, yo nunca olvidaré la lección de fuerza y valentía, de ilusión con la que se enfrentó a su enfermedad. Se le ha llevado pero estad seguros de que no le venció.

Bueno, creo que ya no puedo escribir más, esto es demasiado duro para todos los que le queremos, porque él siempre estará en nuestro presente ya que ocupaba demasiado espacio en nuestras vidas como para dejarnos sin más.

Anabel Marina

La noticia de la muerte de José Luis me ha impactado enormemente. José Luis era una buenísima persona, a quien, como ya han dicho sus buenos amigos, era difícil no apreciar desde el primer momento. Nos visitaba frecuentemente en nuestro laboratorio, donde analizábamos por espectrometría de masas las muestras que él nos traía, y recuerdo que aprovechaba la menor ocasión para coger su coche en Santiago y plantarse en Madrid. Con José Luis no cabían desánimos ni desasosiegos, era de un optimismo incorregible y contagioso y estaba siempre maquinando nuevos experimentos e invitándonos a participar en incipientes proyectos. Uno de nuestros pioneros de la electroforesis bidimensional, de José Luis y sus mejillones eran los mejores geles que recuerdo, plagados de *spots* resueltos a la perfección y de un enorme tamaño, que siempre nos llamaban la atención. Pero hoy en día, que están de moda los péptidos proteotípicos y el PeptideAtlas, pocos recuerdan que de José Luis fue la primera idea de usar un péptido de la tropomiosina, que acabábamos de secuenciar, para determinar de forma absolutamente específica, frente a cualquier otra especie marina o terrestre conocida, y usando una variante de lo que ahora llamaríamos MRM, la identidad de las tres especies comerciales de mejillones más importantes de Europa.

Pero ahora no acierto a acordarme de aquellos trabajos, donde a trancas y barrancas íbamos descubriendo, con sorpresas, el potencial de la tecnología que teníamos entre manos. Sólo me vienen a la cabeza los buenos momentos pasados en su compañía, en Madrid o en Santiago. Y se me hace imposible aceptar cómo es posible que su corta y prometedora carrera, su generosa amistad y la alegría que nos ofrecía se hayan podido truncar de esta manera.

Jesús Vázquez

Queridos amigos, nunca pensé en enviar a los socios de la SEProt una noticia tan dolorosa e impactante como es la del fallecimiento, el pasado 12 de octubre, de José Luis López Rodríguez, de la Universidad de Santiago. Muchos de vosotros le conocisteis, por lo que sobran las palabras. Para los que no tuvisteis la oportunidad de coincidir con él os diré que pertenecía a los que podíamos denominar grupo de "pioneros españoles de la proteómica". Participó activamente en todas las iniciativas que desde el año 2003 se pusieron en marcha y que desembocaron en la creación de la SEProt y otros proyectos ya consolidados. De su calidad científica, en especial en el campo de la electroforesis bidimensional, dan fe sus publicaciones, colaboradores y discípulos. Por encima de ello quiero resaltar su humanidad y bondad, definiéndolo, "en el mejor sentido de la palabra, bueno". Los que le conocimos y, en especial, los miembros de la Junta Directiva, nos gustaría rendirle un más que merecido y justo, aunque insuficiente, homenaje. Es por ello que en la página web del SEProt se va a activar un sistema de envío de mensajes que pretendemos hacer llegar a la familia. También pensamos rendirle público homenaje en Pamplona durante la celebración del próximo Congreso de la SEProt.

Jesús V. Jorrín Novo

Introducción a la Espectrometría de Masas para la caracterización de péptidos y proteínas en Proteómica

Joaquín Abián, Montserrat Carrascal, Marina Gay

Laboratori de Proteòmica CSIC/UAB, Instituto de Investigaciones Biomédicas de Barcelona IIBB-CSIC, IDIBAPS, Barcelona.

Resumen

La identificación de péptidos y proteínas en mezclas complejas se lleva a cabo actualmente mediante una serie de procedimientos que, si bien pueden utilizar una instrumentación muy variada, se asienta en dos estrategias básicas: el denominado mapeo de masas peptídicas mediante espectrometría de masas y la secuenciación de péptidos mediante espectrometría de masas en tándem. En el primer caso, la utilización de bases de datos de secuencias de genes o proteínas es obligada, mientras que no lo es necesariamente en el segundo. En este tutorial analizaremos las diferentes técnicas instrumentales utilizadas en estos procedimientos haciendo hincapié en las ventajas e inconvenientes de cada una de ellas.

1. Espectrometría de masas

La espectrometría de masas *(mass spectrometry*, MS) se introdujo en el ámbito del análisis de biomoléculas al final de la década de los años 70 del siglo pasado con la aparición de diversas técnicas de ionización suave, como la desorción por campo eléctrico (FD), la desorción por plasma (PD) y la ionización por bombardeo con átomos rápidos (FAB) (Roepstorff y Richter, 1992). Estas técnicas permitían la ionización de moléculas termolábiles de gran tamaño, como los péptidos y las proteínas, sin producir la extensiva degradación característica de las denominadas técnicas de ionización dura, cuyo representante de referencia es el impacto electrónico. En especial el FAB (Barber et al., 1981), desarrollado a partir de las técnicas de espectrometría de masas de iones secundarios (SIMS), constituyó la herramienta básica para el análisis y secuenciación de péptidos durante más de una década. Posteriormente, en el inicio de los años 90, otros dos métodos de ionización suave, el electrospray (ESI) y la desorción por láser asistida por matriz (MALDI), tomaron el relevo al FAB y se constituyeron en los pilares básicos de la espectrometría de masas en la proteómica contemporánea.

El desarrollo de analizadores de masa (Figura 1) de creciente resolución y velocidad de análisis ha sido primordial en este proceso. En contraste con los métodos de ionización por ESI y MALDI, con una tecnología que quedó prácticamente establecida en los primeros años después de su aparición, el incremento continuo de las prestaciones de los analizadores de iones ha ido marcando y ampliando los límites con los que el investigador ha podido abordar el estudio del proteoma por espectrometría de masas en los últimos años.

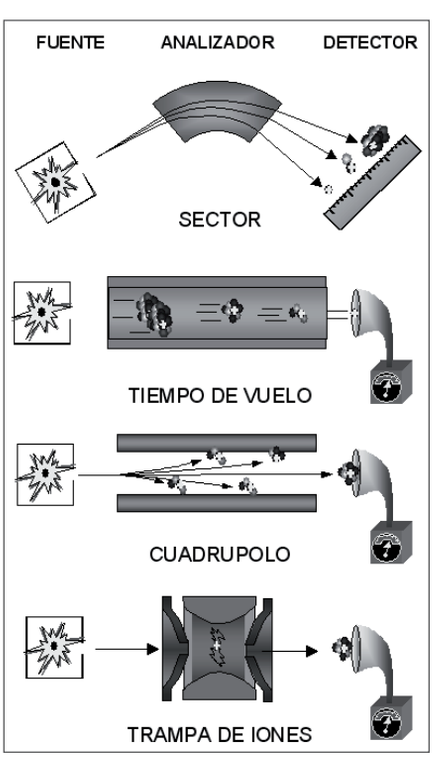


Figura 1.- *Analizadores de masa*

De los cuatro grandes grupos en que pueden dividirse los analizadores de masas, vgr. analizadores de sectores (eléctricos, E y/o magnéticos, B), de cuadrupolo (Q), de tiempo de vuelo (TOF) y de confinamiento de iones (Ion Traps, Orbitraps, y analizadores de resonancia ciclotrónica de iones, ICR), sólo los primeros no tienen una utilización amplia en proteómica (Figura 1). Las ventajas únicas que ofrecían en el pasado los sistemas de sectores en cuanto a su resolución y a su capacidad para llevar a cabo colisiones de alta energía para la secuenciación de péptidos, puede ser proporcionada ahora por sistemas relativamente menos complejos, voluminosos y costosos como los analizadores de tiempo de vuelo provistos de reflectrón (reTOF) o el *Orbitrap*. Por otro lado, los sistemas actuales de ICR mediante transformada de Fourier (FT-ICR), superan ampliamente las características de los mejores sistemas de sector, por lo que en definitiva estos últimos han dejado de ocupar un espacio visible en el área de las aplicaciones en proteómica.

2. Analizadores de masa

2.1. Cuadrupolo

Los cuadrupolos son un caso particular de sistemas multipolares que, como los hexapolos u octapolos, están presentes en casi todos los tipos de espectrómetros de masa modernos como filtros y guías de banda ancha de iones. El funcionamiento de este analizador fue descrito ya por Paul y Steinwegen en 1953 (Paul y Steinwegen, 1953). Un cuadrupolo consiste en cuatro barras paralelas de sección hiperbólica en la cara interna, generalmente de unos 15-20 cm de largo y 0.5 cm de radio, separadas entre sí unos 2 cm, a las que se aplica un potencial combinado de corriente continua y de radiofrecuencia que crean en su interior un campo denominado cuadrupolar. Los iones generados en la fuente deben atravesar longitudinalmente el recinto limitado por estas barras para incidir en el detector. Estos iones, que entran en el analizador a una energía de unos pocos electronvoltios, son sometidos al efecto del campo cuadrupolar que los hace oscilar y los desvía en función de su valor m/z (masa por unidad de carga) de forma que para una combinación de potenciales sólo los iones en un estrecho rango de valores m/z llega a incidir en el detector. Los analizadores de cuadrupolo actúan por tanto como filtros de iones y los espectros de masas en estos sistemas se obtienen mediante un barrido del potencial aplicado a las barras. En consecuencia, en cada instante sólo una pequeña fracción del total de iones es monitorizada mientras que el resto se desecha. Este modo de funcionamiento repercute negativamente en el límite de detección de estos instrumentos, especialmente cuando se requiere la obtención de espectros completos en rangos de masa amplios. Cuando a un cuadrupolo se le aplica únicamente el potencial de radiofrecuencia, este sistema actua como un filtro de banda ancha mucho más eficiente que los formados por lentes electrostáticas en analizadores de baja energía. Este modo de trabajo, conocido como *RF-only*, es el que utilizan las cámaras de colision en los instrumentos de triple cuadrupolo y en otros espectrómetros en tándem que emplean colisiones de baja energía. Además, en los sistemas de triple cuadrupolo, cuando se realizan barridos convencionales (MS) uno de los dos analizadores debe trabajar también en modo *RF-only.*

2.2. TOF

El analizador TOF (Wiley y McLaren, 1955) se basa en el hecho de que los iones acelerados por un campo eléctrico adquieren distintas velocidades según el valor de su relación m/z y por tanto tardan distinto tiempo en recorrer una determinada distancia. Considerando constantes la carga y la energía cinética de los iones formados en la fuente, la medida del tiempo de vuelo permite determinar de forma muy precisa la masa de cada uno de estos iones. A diferencia de los sistemas de cuadrupolo o de sectores, que como se indicó anteriormente filtran en cada instante grupos de iones dentro de un pequeño rango de valores m/z desechando el resto de la población de iones, el analizador TOF separa y detecta en una escala de tiempo (tiempo de vuelo) el paquete completo de iones procedente de la fuente. El sistema trabaja por este motivo en régimen discontinuo por lo que es un detector indicado para técnicas de ionización de carácter pulsante como el MALDI. El acoplamiento del TOF a fuentes de iones de producción continua como el ESI, se lleva a cabo mediante sistemas de confinamiento de iones intermedios que mediante campos eléctricos, acumulan e inyectan, perpendicularmente a la trayectoria inicial, paquetes de iones en el analizador en ciclos con frecuencia del orden del KHz (Boyle y Whitehouse, 1992). El TOF es un analizador relativamente simple, barato, con un ciclo de barrido rápido y con un rango de masas en teoría ilimitado. Debe indicarse que este instrumento fue el primero en utilizarse en sistemas acoplados para cromatografía de gases-MS (GC/MS) en los años 50 del siglo pasado (Gohlke, 1959) y el

primero también utilizado para el estudio de biomoléculas mediante su acoplamiento a la ionización por desorción por plasma (PD) en 1974 (Thorgerson et al., 1974). En general las técnicas de láser desarrolladas entonces (PD, LD) se han utilizado casi exclusivamente con analizadores TOF. Con algunas excepciones, lo mismo puede decirse actualmente de las fuentes MALDI.

2.3. Trampas de Iones

Este tipo de analizadores permiten el confinamiento de iones dentro de una cámara de pequeño tamaño utilizando campos eléctricos (*Paul trap* o *ión trap, Orbitrap*) o magnéticos (analizador de resonancia ciclotrónica de iones mediante transformada de Fourier, FT-ICR). Este tipo de analizadores permiten almacenar, seleccionar y analizar los iones formados en la misma trampa o en fuentes de ionización externas. Los iones pueden mantenerse en el interior de la trampa durante tiempos prolongados con objeto de favorecer la observación de descomposiciones metaestables o de fragmentos producidos por colisión con moléculas de gas. Los iones fragmentos generados en estos procesos pueden ser a su vez seleccionados y fragmentados de nuevo en la misma trampa de forma que el sistema equivale a un sistema de espectrometría de masas en tándem múltiple.

El denominado *ion trap* o trampa de iones cuadrupolar (QIT) fue inicialmente descrito por el premio Nobel de física Wolfgang Paul en la misma patente en la que describía el cuadrupolo (Paul y Steinwedel, 1960). Los principales avances en el desarrollo de este analizador fueron llevados a cabo por Finnigan MAT (Stafford et al., 1984) que en 1983 comercializó este analizador en sistemas para GC/MS. En un QIT, el espacio donde se atrapan los iones queda definido por las superficies hiperbólicas de tres electrodos: un anillo circular central y dos tapas laterales, estas últimas provistas en su centro de pequeños agujeros para la entrada de los iones a la fuente y su salida hacia el detector (March, 1997). Generalmente las tapas laterales están conectadas a tierra o tienen aplicado un campo de corriente continua mientras que al electrodo anular se le aplica un voltaje de radiofrecuencia que genera un campo de tipo cuadrupolar en el interior del analizador. Los iones que se introducen en la trampa quedan sometidos a este campo de radiofrecuencia que los mantiene oscilando en una órbita compleja formada por una componente axial y otra radial. Para generar el espectro de masas, los iones pueden extraerse de la trampa en función de su relación m/z utilizando diferentes procedimientos. El primer método utilizado en sistemas comerciales fue el modo de desestabilización selectiva de masas (*mass-selective instability*) que hace un barrido de la corriente de radiofrecuencia aplicada al electrodo anular. El aumento de la amplitud de esta corriente aumenta la energía de las oscilaciones de los iones y desestabiliza sus trayectorias produciendo la expulsión axial de la trampa de iones con valores de m/z creciente (Stafford et al., 1984). Posteriormente el procedimiento se mejoró mediante el modo de eyección resonante, donde se aplica un voltaje de radiofrecuencia adicional a las placas laterales para la desestabilización y expulsión de los iones.

Un rasgo adicional de la trampa es que ésta permite aislar iones individuales que luego pueden, mediante la aplicación de un voltaje de radiofrecuencia a las tapas laterales, excitarse para su fragmentación por CID con moléculas de Helio introducidas en la trampa. Para el aislamiento, primero se eliminan los iones de valor m/z inferior al del ión de interés, aumentando la amplitud de la radiofrecuencia aplicada al electrodo anular tal como se indicó anteriormente, para seguidamente aplicar a éste un voltaje negativo que produce la eliminación en el eje radial de los iones de m/z superior. Este proceso de aislamiento/activación puede realizarse repetidamente y si bien en la práctica del análisis proteómico raramente se requieren más de 3 de estas etapas, las trampas iónicas comerciales son capaces de realizar MS^n con n=15.

El primer instrumento basado en el QIT dirigido al análisis de biomoléculas (ESQUIRE) fue presentado por Bruker Instruments en 1994. Este sistema podía utilizar diversos métodos de ionización como el ESI, el MALDI o el FAB y presentaba un límite del rango de barrido de 6000 Th. En 1995, Finnigan MAT por su parte introdujo su LCQ, sistema especialmente diseñado para ionización por ESI e ionización química a presión atmosférica (APCI) con un límite de 2000 Th. Para entonces el QIT había superado muchos de sus inconvenientes iniciales (rango de masas bajo, limitación al análisis de iones positivos) y mostraba las ventajas únicas que los sistemas de confinamiento de iones aportaban, entre ellas, su capacidad de acumulación de iones y de realizar análisis en tándem múltiple. La aparición de las fuentes de micro (Emmet y Caprioli, 1994) y nanospray (Wilm y Mann, 1994) en 1995 hizo de estos sistemas uno de los analizadores más populares para el análisis de biomoléculas en las últimas dos décadas.

2.4. Trampa de iones lineal

El confinamiento de iones puede también realizarse en sistemas de cuadrupolo provistos en los extremos de sus barras de unos electrodos terminales con la misma geometría que éstas pero más cortos y aislados eléctricamente. A estos electrodos se les aplica un potencial que actúa como barrera para los iones atrapados en su interior. Este tipo de analizadores, denominados trampas iónicas lineales (LIT, *linear ion trap*) o cuadrupolos de confinamiento lineal (LTQ, *linear trap quadrupole*), mejoró muchas de las características de los QITs aumentando su capacidad de almacenamiento de iones, velocidad de barrido y sensibilidad (Schwartz et al., 2002). La trampa LIT puede utilizarse además como un cuadrupolo convencional por lo que, situada como el segundo analizador en sistemas de triple cuadrupolo, permite trabajar en cualquiera de los modos de espectrometría de masas en tándem posibles en este tipo de instrumentos (barridos de precursores, pérdidas neutras y de iones producto) además de los modos propios de la trampa. Este tipo de analizadores puede disponer de un detector situado axialmente, que es la configuracion convencional de los analizadores de cuadrupolo, o bien de dos detectores situados en posiciones opuestas en el eje perpendicular del analizador. Dado que el modo de barrido de las trampas implica la expulsión de iones en dos direcciones opuestas, esta segunda configuración permite aprovechar la señal producida por ambos haces de iones y por tanto mejora la sensibilidad de estos sistemas.

2.5. Orbitrap

El orbitrap fue desarrollado y patentado por Makarov a finales de los años 90 (Makarov, 2000). Este analizador deriva de las trampas de Kingdon y Knight (Knight, 1981) que utilizaban dos electrodos, uno tubular externo y un filamento interno coaxial, para atrapar iones en el campo electroestático formado entre ellos. En el orbitrap, el filamento interno se ha convertido en una barra en forma de huso y las paredes del cilindro que lo contienen han adquirido esta misma forma (Hu et al., 2005). Los iones introducidos de forma perpendicular en este sistema adquieren un movimiento radial alrededor del huso combinado con un movimiento axial periódico cuya frecuencia es una función de su valor m/z. Este movimiento podría visualizarse como si estuviéramos bobinando un hilo en un huso convencional. En el orbitrap, el electrodo externo esta dividido en su centro y perpendicularmente en dos mitades, de forma que el movimiento periódico de los paquetes de iones de un lado a otro de la cavidad induce entre estas dos mitades una señal imagen cuya frecuencia puede medirse para determinar el valor m/z correspondiente. En la práctica, la señal registrada es una señal compleja o interferograma formada por la combinación de las señales senoidales de distintas frecuencias (m/z) e intensidades (abundancia) de cada uno de los iones. La aplicación a esta señal de la trasformada de Fourier (*Fourier Transform*, FT), método matemático que permite separar los componentes de una señal periódica compleja, permite determinar la frecuencia derivada de cada uno de los iones y obtener el espectro de masas. El orbitrap es quizás el analizador de iones más joven en espectrometría de masas aunque, por sus características se ha introducido rápidamente en el área de la proteómica (Scigelova y Makarov, 2006). Está comercializado por una única compañía, ThermoFisher, acoplado a un instrumento de trampa lineal que se utiliza para la acumulación de iones y su inyección en el Orbitrap así como para llevar a cabo la fragmentación de iones en alguno de sus modos de espectrometría de masas en tándem.

2.6. FT-ICR

La espectrometría de masas de resonancia ciclotrónica de iones con transformada de Fourier fue introducida por Marshall y Comisarow en 1974 y ya en 1989 Hunt y cols. mostraron su potencial en el análisis de biomoléculas utilizando una fuente de ESI (Henry et al., 1989). Debido a la complejidad, tamaño y coste derivados en gran medida de los imanes superconductores utilizados en los sistemas FT-ICR, estos instrumentos quedaron restringidos durante muchos años a grandes laboratorios del área de la física y de la fisicoquímica o de la industria química. La continua disminución en los costes de los sistemas computacionales y electrónicos que gravan estos instrumentos, así como la solución de diversos problemas inherentes a las estrictas condiciones de funcionamiento requeridas por la celda de ICR, ha permitido en los últimos años su introducción en otras áreas de investigación y hoy constituye una de las herramientas más potentes en el ámbito de la proteómica.

El centro del analizador FT-ICR es una trampa de Penning, que consiste en un campo eléctrico cuadrupolar combinado con un campo magnético homogéneo que permite confinar los iones en su interior

(el campo cuadrupolar confina los iones axialmente mientras que el campo magnético lo hace radialmente). Esta celda ICR puede tomar diferentes formas (cúbicas, hiperbólicas, abiertas de segmentos cilíndricos…) con un número variable de electrodos.

La FT-ICR se basa en la utilización del proceso de resonancia ciclotrónica para la determinación de la masa de un ión. Una partícula cargada en movimiento perpendicular a un campo magnético adquiere una trayectoria circular de una frecuencia de giro (frecuencia ciclotrónica) que depende de su relación m/z y del valor del campo. La aplicación de una señal eléctrica de frecuencia igual a la frecuencia de giro de la partícula hace que esta absorba energía de forma resonante y aumente su radio de giro. En las celdas ICR originales, este aumento del radio de giro permitía a la partícula alcanzar un detector con el que colisionaba (Figura 2). Estas celdas eran por tanto destructivas. En los sistemas FT-ICR, en cambio, la frecuencia de giro puede medirse a partir de la señal imagen inducida capacitativamente por la carga durante su giro en unos electrodos y a partir de ella determinar la masa del ión. En la práctica, cuando se trata de una mezcla de iones, el conjunto se excita durante unos milisegundos con una señal de frecuencia variable lo que sitúa a todos los iones en órbitas cercanas a las placas detectoras que colectan el correspondiente interferograma. Tal como se describió en el caso del Orbitrap, esta señal puede entonces deconvolucionarse en sus componentes mediante su análisis de FT para generar el correspondiente espectro de masas.

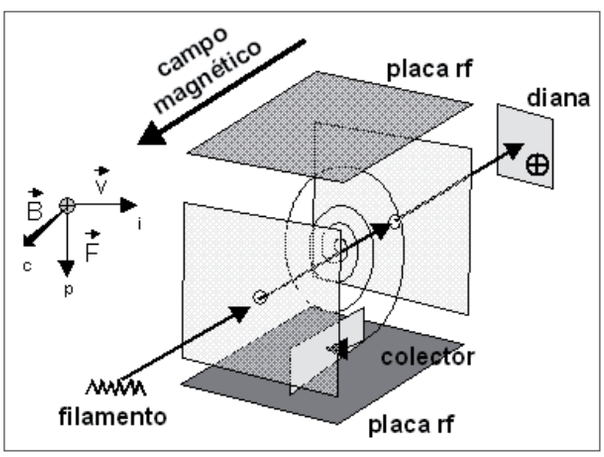


Figura 2.- *Celda ICR clásica. La ionización de las moléculas en el interior de la celda se lleva a cabo mediante ionización electrónica. Las placas frontal y trasera tienen aplicado un potencial para el confinamiento de iones en la dirección axial. Entre las placas superior e inferior se aplica una señal de radiofrecuencia resonante con la frecuencia ciclotrónica del ión, llevándolo a una órbita que intersecta con la placa detectora. El sistema de vectores incluye la clave mnemotécnica de la regla de mano derecha para la fuerza (F) que sufre una partícula cargada positivamente en movimiento (v) dentro de un campo magnético (B) (dedos: c, corazón; p, pulgar; i, índice)*

El FT-ICR utiliza imanes superconductores para generar campos de entre 2 y 14 Teslas y que determinan el grado de resolución alcanzable en la medida. Como en el caso del orbitrap, el FT-ICR se utiliza acoplado a un sistema de inyección de iones, generalmente una trampa LIT o QIT.

2.7. Características de los diferentes analizadores

Los analizadores de iones pueden clasificarse mediante diversas características entre las cuales podríamos destacar: su poder de resolución, la precisión de la medida, y el rango de masa accesible en el análisis. Otros factores a tener en cuenta son su rango dinámico, el ciclo de trabajo y su rendimiento de transmisión de iones (Tabla 1).

La resolución determina el grado hasta el cual un instrumento puede diferenciar entre dos iones adyacentes en el eje de masas. La resolución depende de la masa y por lo tanto es conveniente disponer de su valor a distintas masas para evaluar las características del analizador. Existen diversos términos en espectrometría de masas para expresar la resolución. La denominada resolución a 10% de valle se define como el cociente $M/\Delta M$ para dos picos iguales de masa M y $M+\Delta M$ que se encuentran separados de forma que la altura del collado entre ellos es igual al 10% de su altura máxima. La resolución expresada en términos de anchura de pico se define como el valor $M/\Delta M$ para un pico de masa M y que presenta un ancho ΔM a una altura determinada típicamente establecida a 0.5, 5 o 50% de su altura máxima. El valor de resolución a 50% de altura máxima es un valor estándar frecuentemente utilizado en la descripción de la forma de distintos tipos de señales incluida la espectrometría de masas y se indica con el término FWHM (*full widht half mass*) (Price, 1991). Como referencia, cuando se trata de dos picos iguales y simétricos, la resolución a un determinado valor de porcentaje de valle es la misma que a un porcentaje de altura máxima correspondiente a la mitad de ese valor. En el caso de los analizadores de cuadrupolo se ha utilizado frecuentemente el termino de "resolución unidad" que se refiere a que en el rango de masas de trabajo este analizador es capaz de resolver dos señales (generalmente a 50% de altura de valle) separadas una unidad de masa.

Tabla 1.- Resumen de características de algunos de los analizadores comerciales actualmente en uso en proteómica . En algunos casos los valores de la tabla se han recalculado con objeto de homogeneizar las unidades. NA= información no disponible o no encontrada en las fuentes consultadas.

INSTRUMENTO	*Rango de Barrido*	*Resolución*	*Precisión (ppm)[1]*	*MS/MS*	*Ciclo de Barrido[2]*	*Límite de Detección*	*Rango Dinámico*	*Etapas MS^n*
QIT, LIT								
QIT	50 – 2200 200 - 6000	FWHM (m/z 50-2200) > 3700 a 26000 u/s > 8800 a 800 u/s	Max 100 (barrido completo)	CID, ETD	800-27,000 u/s	250 fg reserpina - en columna	$>10^3$	5-11
LIT	15 – 200 50 – 2000 200 – 4000	Max 40000 (ZoomScan)	NA	CID, ETD	4 sc/s	250 fg reserpina - en inyector S/N>100	NA	15
QqLIT	5 - 2800 en Q1/Q3 50 - 2800 en LIT	>3000 (m/z 609, 250 u/s)	120 Mr 16951, TurboIonSpray, Q1	CID	<4,000 u/s LIT <2,400 u/s RF/DC	1 pmol Mioglobin Q1 barrido completo (400-1800 amu) S/N > 6 para pico base	$>10^6$	NA
Orbitrap								
20	50-2000 200-4000	7500 - 100000 at m/z 400 (60000 a 1Hz)	< 3 EC < 2 IC	CID	0.5 sc/s a $R=10^5$	< fmol	> 4 x10^3	10
FT								
30	100 - 10000	9E5 (m/z 400, 7.0 T) 1.5E6 (m/z 400, 9.4 T) Mayor a 12-15 T.	Sub-ppm	ECD, CID, IRMPD	2 sc/s	< fmol	10^4	5
31 (7T)	50 - 4000	100,000 (m/z 400, 1 sc/s) > 750,000 m/z 400, < 1 sc/s)	< 1.2 EC < 1 IC	ECD, CID, IRMPD	1 sc/s at $R=10^5$	amol	> 4 x10^3	NA
40	5-40000 (TOF) 5-6000 (QMS/MS)	<1 Da para selección de Precursor	<3	CID	< 20 sc/s	NA	10^4	2
41	25–20000 TOF	≤ 20,000	<1 MS 2 MS/MS	CID	20 sc/s (MS) 10 sc/s (MS/MS)	fmol-amol en columna	10^5	2
42	20-1000,000 (TOF) 20-8,000 (Q)	17500 (m/z 956)	2	CID	10 sc/s	amol	10^4	2
43	50 to 20,000	17,500 FWHM a m/z 922 a max sensibilidad	< 2 IC < 5 EC	CID or ISCD	20 sc/s	amol	10^5	2
TOF-TOF								
50 (Lineal)	NA	≥ 1000	≤ 100					2
50 (Reflector)	0 to 500000	> 25,000 para péptidos	≤ 3	CID	200 sc/s	250 amol with S/N > 100:1	10^3-10^4	2
51 (Lineal)	NA	> 800 a 11674 Da > 85 a 66431 Da	500 at 11674 10000 at 66431	CID	NA	Tioredoxina 1pmol 500 disparos, S/N >200:1 BSA 1pmol 1000 disparos, S/N >80:1	NA	2
51 (Reflector)	NA	>20.000 de 1000 a 3600 Da	2.5 IC 20 EC	CID	NA	Neurotensina 250 amol 400 disparos, S/N >100:1	NA	2
51 (MS/MS)	NA	**Precursor:** 1570.6774, **Shots:** 500 **Fragmento Resolucion** 175.1195 > 2000 1056.475 > 3500 1441.634 > 5000	<40 a 175-1441	CID	NA	**Precursor:** 1296.69 1.5 fmol **Disparos:** 500 **Fragmento, S/N Specific** 255.109 S/N > 65 784.411 S/N > 30 1183.601 S/N > 70	NA	2

1) EC: Calibración Externa, IC: Calibración Interna
2) u/s: unidades de masa/segundo, sc/s: barridos por segundo

La resolución de un cuadrupolo o de un analizador TOF es de aproximadamente 500-1000 a m/z 1000. La resolución del analizador TOF se mejoró notablemente mediante la adición del reflectrón y las técnicas de extracción retardada de iones (DE, *delayed extraction*). El sistema de reflectrón (reTOF)

utiliza un espejo de lentes electrostáticas para corregir la dispersión en la velocidad de los distintos iones derivada del proceso de ionización. En comparacion con el trabajo en modo convencional (denominado "lineal"), que ofrece un rango de barrido comunmente calificado como ilimitado, el uso del reflectrón limita este rango a valores de hasta 5000-6000 Th (Tabla 1). Esta limitación es sin embargo poco relevante en el ámbito de la identificación rutinaria de proteínas donde, tal como veremos más adelante, el material a analizar (digeridos peptídicos) tiene generalmente pesos moleculares menores de 5000 Da. El mismo objetivo de aumento de resolución se consigue mediante la DE que consiste en acumular los iones producidos en la fuente por el disparo láser para enviarlos al tubo de vuelo en un paquete compacto. Un analizador reTOF puede alcanzar resoluciones superiores a 20000 FWHM a este valor de m/z, mientras que un QIT convencional o una trampa lineal suelen ofrecer resoluciones cercanas a 5000 y 10000, respectivamente, en modo de barrido completo, y en el entorno de 10000 y 40000 respectivamente utilizando modos de barrido lento de rango corto (ZoomScan y UltraZoomScan). En comparación, un analizador orbitrap o un FT-ICR con campo de 7 Teslas pueden alcanzar resoluciones del orden de 100.000, mientras que un FT-ICR trabajando con tiempos de análisis mayores o utilizando imanes más potentes puede alcanzar valores superiores a 10^6.

La precisión de un analizador indica el error con el cual puede medirse la masa de un ión. La precisión está relacionada con la reproducibilidad de la medida y determina, junto con la resolución, la capacidad de un instrumento para la medida de masas exactas en mezclas complejas, que es el campo habitual del análisis proteómico. La determinación de masas exactas es una de las tareas más recurrentes en síntesis orgánica donde se persigue su caracterización con un error inferior a 5 ppm como requisito, un tanto arbitrario, para la confirmación de la composición elemental de los compuestos descritos en los trabajos de investigación. Esta medida se ha llevado a cabo clásicamente mediante espectrómetros de masas de sectores, frecuentemente mediante ionización electrónica o por FAB y puede llevarse a cabo a nivel inferiores a 1 ppm (Hass, 1999). Las medidas de masa exacta con estos instrumentos requiere sin embargo la utilización de métodos de calibración interna como el *peak matching* o el *bracketing* poco adecuados para su utilización en el análisis proteómico rutinario. Los analizadores TOF en los sistemas modernos MALDI-reTOF permiten la medida exacta de masa al nivel de los 5-10 ppm de forma rutinaria mientras que un orbitrap puede alcanzar los 2-5 ppm. El límite inferior se consigue de nuevo con los sistemas FT-ICR que pueden alcanzar errores inferiores a 1 ppm para péptidos de masa alrededor de 5 KDa.

La limitación en el rango de barrido de los analizadores actuales no suele tener relevancia en los procedimientos convencionales tipo "de abajo hacia arriba" (*bottom-up*) utilizados para la identificación de proteínas mediante mapeo peptídico o secuenciación de sus digeridos enzimáticos, ya que, en general, las señales a identificar producidas mediante MALDI o ESI se encuentran mayoritariamente por debajo de 4000 Th. En cualquier caso, la limitación de los cuadrupolos y trampas cuadrupolares y en menor grado de los sectores en cuanto al máximo valor de la masa analizable (max. 2000-8000 Da y 15000 Da, respectivamente) pasó a un segundo plano con la aparición del ESI, método de ionización que se caracteriza por la producción de abundantes iones multicargados (z>1). Estos iones presentan relaciones m/z que son fracciones del peso molecular del compuesto, por lo que moléculas con pesos moleculares muy por encima del rango de estos analizadores pueden igualmente ser analizadas y detectadas. Por otro lado, el analizador TOF, que en modo lineal se caracteriza por un rango de masas teóricamente ilimitado, puede también alcanzar límites superiores a 4-5 KDa en modo reflectrón. En general, el análisis de iones de masas superiores a 4-5 KDa implica una pérdida de resolución y exactitud de la medida. Aun así, los sistemas FT-ICR permiten alcanzar límites de m/z superiores a 20000 con alta resolución. Ya en 1995 se habían descrito, mediante MALDI-FTICR, resoluciones de 90000 para la insulina bovina (Mr 5733.6) y de más 10^6 para Arg^8-vasopresina (Hunter el al., 1994; Pastor et al., 1995) con errores inferiores a 1 ppm en el cálculo de masa.

3. Métodos de ionización

La mayoría de los métodos de análisis de iones en fase gaseosa descritos anteriormente se conocían y utilizaban hace ya decenios. El factor principal que permitió la introducción de la espectrometría de masas en el análisis de biomoléculas no fue sino el desarrollo en los años 80 de diversas técnicas de ionización suave que permitían observar iones moleculares de compuestos de masa alta que de otra forma sufrían degradaciones extensivas resultado del exceso de energía que las otras técnicas depositaban sobre ellos.

ESI y MALDI, las dos técnicas de ionización utilizadas de forma general en la actualidad en el análisis proteómico, fueron el núcleo de una revolución que cambió nuestra forma de abordar el estudio químico de los sistemas biológicos. No obstante, a partir de mediados de los 90 estas técnicas estaban ya establecidas (Figura 3) y, excepto por diversos avances en el campo de los sistemas basados en *chips*, su evolución técnica en los últimos años no ha tenido la relevancia de la derivada del desarrollo de los nuevos analizadores de iones. Por este motivo entraremos muy por encima en la descripción de estos dos métodos de ionización.

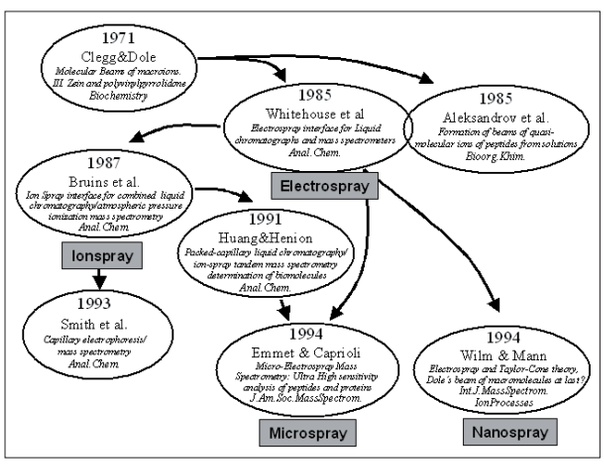


Figura 3. *Filogenia de la familia del Electrospray*

El ESI fue desarrollado simultáneamente por los grupos de Fenn y Aleksandrov durante los años 80 (Whitehouse et al., 1995; Alexandrov et al., 1995). El ESI es una técnica de ionización a presión atmosférica en la que, mediante la aplicación de un fuerte campo eléctrico, se produce la nebulización de una solución del analito a su salida de un tubo capilar. En este proceso los iones presentes en fase líquida pueden son desorbidos a fase gaseosa de una forma tan suave que se pueden conservar incluso complejos moleculares derivados de interacciones no covalentes existentes en la solución. La ionización por ESI de péptidos y proteínas genera iones multicargados siendo la extensión de la carga dependiente del pH de la solución y del número de grupos básicos (o ácidos) de la molécula.

En los trabajos pioneros de Fenn y Aleksandrov de 1985 se demostraba ya la capacidad del ESI para el análisis de péptidos, presentándose resultados obtenidos con la ciclosporina A (Mr 1203), la gramicidina S (Mr 1140), la bleomicina (Mr 1375), la substancia P (Mr 1348) y la bradiquinina (Mr 1060). Poco después se demostró la capacidad del ESI para el análisis de proteínas de gran tamaño como la alcohol deshidrogenasa (Mr 40000) (Meng et al., 1988), haciendo aparición los primeros sistemas comerciales. A partir de este momento la tecnología inicial se desarrolló rápidamente para dar lugar a interfases acopladas a la cromatografía líquida convencional (ionspray) y capilar (microelectrospray, Emmet y Caprioli, 1994) y a interfases para análisis *off-line* (nanospray, Wilm y Mann, 1994). La utilización del sonoro término "nano" actualmente engloba, probablemente de forma incorrecta, las interfases de ESI para cromatografía capilar en línea.

Los sistemas de ESI se acoplaron inicialmente a instrumentos de cuadrupolo simple y de triple cuadrupolo y en menor medida a analizadores de sectores. Su acoplamiento a los nuevos sistemas QIT a mediados de los años 90, que aparecieron en paralelo con las nuevas fuentes de nanospray comercializadas por Protana (actualmente MSD Proteomics), dio un fuerte empuje al uso de este tipo de espectrómetros. El uso de la fuente ESI en analizadores Q-TOF fue otra de las combinaciones más exitosas en proteómica ya que permitía obtener espectros de alta resolución. El ESI puede utilizarse actualmente acoplado a cualquiera de los tipos de analizadores de altas prestaciones descritos, especialmente a trampas lineales, orbitraps y FT-ICR.

El MALDI fue desarrollado por los grupos de Karas y Hillenkamp (Karas y Hillenkamp 1988) y del premio Nobel Koichi Tanaka (Tanaka et al., 1988) a partir de las técnicas de desorción por láser aplicadas a la determinación de la composición atómica de superficies. La principal diferencia entre el MALDI y estas técnicas es la utilización, como en el caso del FAB, de una matriz química para la deposición de la muestra. Esta matriz está constituida por un compuesto químico sólido con capacidad de absorber radiación a la frecuencia del láser utilizado (UV o IR) y que juega un papel tanto en la dispersión de las moléculas del analito con el que se co-cristaliza (el primer paso hacia la fase gaseosa) como en la absorción de la energía láser para producir, a través de reacciones fotoquímicas, iones del analito.

La formación de iones en MALDI es pulsante por lo que se ha utilizado preferentemente acoplado a analizadores TOF. Estos sistemas, como comentaremos seguidamente, han sido el paradigma de la instrumentación para identificación de proteínas en base de datos en proteómica. El MALDI se ha acoplado también a sistemas de QIT utilizando modalidades de ionización tanto a alta como a baja presión, así como a sistemas Q-TOF, orbitrap y FT-ICR.

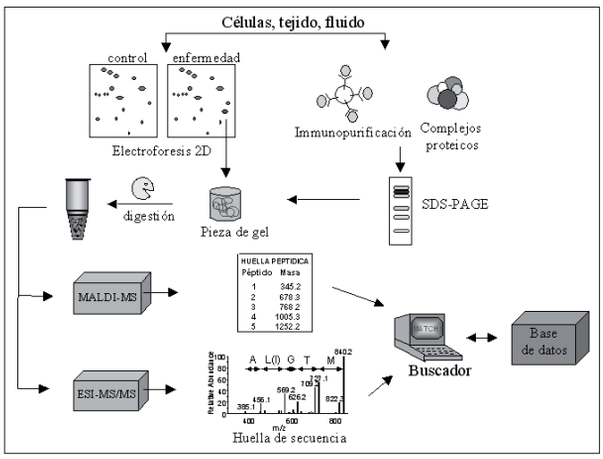


Figura 4.- *Esquema del procedimiento básico para la identificación de proteínas en bases de datos (J.Abián, Tratado de Medicina Interna Rodes-Guardia, Ed Massons (2004))*

4. Identificación de proteínas mediante mapeo peptídico

Mediante la estrategia denominada mapeo de masas peptídicas o huella de masas peptídicas, una proteína aislada es convertida en péptidos de tamaño medio mediante la acción de enzimas específicas, típicamente la tripsina, pero también V8, endolisina u otras. El conjunto de péptidos derivados de una proteína es una característica propia de ésta y puede utilizarse para su identificación tal como mostraron ya en 1977 Cleveland et cols (Cleveland et al, 1977). Así, la determinación de la masa de los péptidos de digestión de una proteína permite hacer comparaciones con las masas teóricas del conjunto de péptidos esperados para las secuencias almacenadas en una base de datos y sugerir, en último término, aquellas que se ajustan más a los datos experimentales (figura 4). La capacidad de este procedimiento para identificar proteínas en bases de datos fue demostrada por Henzel et al en el congreso de la *Protein Society* de 1989 utilizando un espectrómetro de masas de sector con una fuente FAB (Henzel et al., 1993). En 1992 Rosenfeld describió el método para digerir proteínas separadas en geles de electroforesis in-situ que sustituyó a los procedimientos clásicos que requerían una electrotransferencia y la digestión en la membrana (Rosenfeld et al., 1992). En 1993 los grupos de Henzel, Roepstorf, James, Yates y Pappin describieron en paralelo la utilización del mapeo peptídico para la identificación de proteínas en geles (Henzel et al, 2003) y el último de estos autores fue el primero en acuñar el término *Peptide Mass Fingerprinting* (PMF) (Pappin et al, 1993).

La identidad entre el repertorio de péptidos experimental y el calculado *in silico* se valora en función del número de señales de masa identificadas y de la diferencia entre la teórica y experimental utilizando diferentes algoritmos matemáticos (Tabla 2).

Tabla 2. Herramientas para la identificación de proteínas mediante mapeo peptídico o secuenciación de péptidos en bases de datos o *de novo*.

Mapeo Peptídico	
Aldente	www.expasy.org/tools/aldente
Mascot*	www.matrixscience.com
Mowse	www.hgmp.mrc.ac.uk/Bioinformatics/Webapp/mowse
MS-Fit	www.prospector.ucsf.edu
PepMAPPER	wolf.bms.umist.ac.uk/mapper/
Pepsea*	www.pepsea.protana.com
ProFound	www.proteometrics.com
Secuenciación de péptidos	
GPM X! Hunter	h201.thegpm
GutenTag	fields.scripps.edu/GutenTag/index.html
EigenMS	[Bern y Goldberg, 2006]
InsPectT	peptide.ucsd.edu/
Mascot	www.matrixscience.com
MCPS	ras-0.comp.nus.edu.sg/~msms/
MS-Tag	www.prospector.ucsf.edu
OMSSA	pubchem.ncbi.nlm.nih.gov/omssa/
Pepsea	www.pepsea.protana.com
PeptideSearch	www.narrador.embl-heidelberg.de/GroupPages/Homepage.html
Phenyx	phenyx.vital-it.ch/pwi/
Popitam	www.expasy.org/tools/popitam/
SEQUEST	www.fields.scripps.edu/sequest
SHERPA	www.hairyfatguy.com/Sherpa/
Sonar	www.proteometrics.com
SPIDER	bif.csd.uwo.ca/spider/
X!TANDEM	www.thegpm.org/TANDEM/
De Novo	
DeNovoID	prometheus.brc.mcw.edu/denovoid/
DenovoX	www.thermo.com/
Lutefisk	www.hairyfatguy.com/lutefisk/
MSNovo	msms.usc.edu/supplementary/msnovo
NovoHMM	people.inf.ethz.ch/befische/proteomics/
OpenSea	medir.ohsu.edu/ ~geneview.
Peaks	www.bioinformaticssolutions.com/products/peaks/
PepNovo	peptide.ucsd.edu/
SHERENGA	[Dancik et al., 1999]

En un análisis por mapeo peptídico es por tanto necesario tener en cuenta diversos factores que determinaran el éxito del análisis, entre ellos, el grado de purificación de la proteína, el número de péptidos

de ésta obtenido en el proceso de digestión-análisis (cobertura o *coverage*), la sensibilidad y precisión de la medida y la calidad de la base de datos utilizada.

El mapeo peptídico se lleva a cabo actualmente casi exclusivamente mediante técnicas de MALDI-reTOF. La rapidez y capacidad de automatización de estos instrumentos permite analizar decenas de muestras por hora de forma automatizada. Esta estrategia provee por tanto de un método sensible y rápido que es actualmente la herramienta de primera elección para identificar una proteína aislada en solución o separada por electroforesis en gel. En un trabajo reciente se mostraba que en un muestreo que cubría los años 2003-2007 en el 68% de las publicaciones se utilizaba el mapeo peptídico para la identificación, si bien es también destacable que sólo en el 35% del total de los casos esta técnica se utilizaba sola frente al 32% del uso único de la aproximación basada en la secuenciación por MS/MS (Damodaran et al, 2007).

Como ya se indicó, en PMF la identificación de una proteína en una base de datos a partir de la huella dactilar de masas peptídicas se lleva a cabo mediante programas informáticos. Existen diversas herramientas de este tipo, tanto comerciales como libres (Tabla II) que utilizan diferentes sistemas de puntuación para evaluar la bondad de la asignación. Los sistemas de puntuación más simples, utilizados inicialmente, se basaban en determinar el número de péptidos que se ajustaban a las masas esperadas con un determinado error. Este es el caso del programa Fragfit descrito en 1989 por Henzel y col (Henzel et al., 1993) que, teniendo en cuenta el tamaño de las bases de datos disponibles entonces, era capaz, utilizando un computador VAX de buenas prestaciones, de completar una búsqueda… ¡en menos de 5 minutos!. Esta búsqueda simple fue mejorada con la aparición del programa MOWSE ("*Molecular Weight Search*"), en el que se basan los puntuadores actuales de MASCOT y Protein Prospector, y que tenía en cuenta el número de péptidos para cada valor de peso molecular en la base de datos utilizada (Pappin et al. 1994). Profound (Zhang y Chait, 2000) utiliza una aproximación diferente al problema aplicando un algoritmo de probabilidad bayesiano.

Independientemente del algoritmo utilizado, la exactitud en la medida de masa es un factor determinante en una identificación por PMF. Los sistemas MALDI-reTOF menos sofisticados utilizados en proteómica permiten la medida de masas peptídicas con errores inferiores a los 20-50 ppm mientras que en sistemas de última generación este error puede bajar de los 3 ppm. La probabilidad de una identificación correcta y unívoca del espectro con una secuencia en la base de datos aumenta con el numero de señales de la proteína detectables en el espectro. En general una identificación tentativa fiable requiere al menos 5 péptidos medidos con errores de 10 a 30 ppm (Clauser et al., 1999).

Por otro lado, dado que la masa de un péptido viene determinada por la combinación de las masas de cada uno de sus aminoácidos y que el número de estos es finito, la masa del péptido no puede a su vez alcanzar cualquier valor. En consecuencia la determinación precisa del valor m/z de una señal permite no sólo determinar si esta señal corresponde a un péptido sino también su composición aminoacídica. A un valor de m/z entorno a 1005.4433 se calcula existen cerca de 4000 combinaciones posibles de aminoácidos que se ajustarían a esta masa con un error de 10 ppm (Figura 5). Aún aceptando un error máximo de 0.5 ppm es posible encontrar todavía 20 posibles combinaciones (Spengler, 2004). Sin embargo, sólo un reducido número de las posibles permutaciones del conjunto de 10-20 aminoácidos que forman los péptidos generalmente analizados en los digeridos enzimáticos están presentes en un proteoma. Esta circunstancia hace probable la identificación de una proteína en una base de datos a partir de la medida de la masa exacta de sólo uno de sus péptidos de digestión (en el supuesto de que se tratara de un péptido convencional, no modificado). La precisión requerida para esta determinación unívoca de la secuencia del péptido se calcula se encuentra en el orden de 0.1 ppm, lo que pone esta estrategia al alcance de los sistemas FT-ICR de alto campo.

En la práctica, como se ha indicado anteriormente, el procedimiento de PMF implica la medida de la masa de varias señales de la misma proteína con la mayor exactitud posible. En los sistemas MALDI-reTOF comúnmente empleados esto se consigue mediante la calibración del eje de masas con patrones internos. Generalmente se lleva a cabo una calibración externa seguida de una calibración interna utilizando señales conocidas normalmente presentes en todas las muestras, como las derivadas de los péptidos de autodigestión de la tripsina. La calibración externa se realiza mediante muestras de calibrante, situada muy cercanas a un grupo de muestras para evitar errores derivados de la no planaridad de la placa de MALDI, y de las que se adquieren espectros de referencia alternativamente al espectro de cada muestra. Esta calibración inicial sitúa el error de la medida alrededor de 50-100 ppm en un sistema convencional (Voyager DEPro de Applied Biosystems).

La calibración interna, finalmente, permite bajar el error a 10-50 ppm en estos sistemas.

Cuando se utilizan algoritmos como el MOWSE, para una determinada base de datos, la probabilidad de una identificación correcta aumenta con el número y el peso molecular de los péptidos identificados y disminuye con la tolerancia de masa aceptada y la necesidad de asumir factores extras capaces de originar péptidos de masas diferentes a las esperadas. Estos factores extras pueden derivar del proceso de digestión, como por ejemplo la degradación proteolítica incompleta o las roturas no específicas, o del tratamiento previo de la muestra como es el caso de diversos artefactos químicos que incluyen la oxidación de metioninas, deamidación y propionamidación de la cisteina por la acrilamida del gel. Estos factores, aunque de extensión controlable, están siempre presentes en mayor o menor grado y no es posible a priori determinar si las identificaciones de masas asignadas se deben realmente a estos fenómenos o corresponden a la presencia de señales cercanas de otras proteínas contaminantes.

Es posible realizar calibrados internos de la muestra de forma iterativa utilizando los propios iones de ésta que pueden ayudar a distinguir entre señales propias de la proteína y posibles interferencias o asignaciones erróneas. Según este procedimiento, una vez obtenida una identificación probable en la base de datos, el espectro se vuelve a calibrar utilizando los péptidos identificados como referencia y asumiendo que su masa correcta es la derivada de la base de datos. En el caso de que las señales de referencia correspondan realmente a la proteína en cuestión se obtiene un calibrado excelente que permite discriminar con más precisión entre las señales propias y de otras proteínas. Este procedimiento fué incluido en sistemas comerciales (Intellical de Applied Biosystems).

Determinar cuando una identificación es correcta, entendiendo por este término que la mayor parte de las señales medidas corresponde a péptidos cuyas secuencias están contenidas en la de la proteína identificada, es raramente fácil de llevar a cabo únicamente en base al análisis por PMF. El uso de las puntuaciones (*scores*) y de cierto sentido común para limitar la población de posibles candidatos es una práctica generalizada aunque poco estandarizada. Se han sugerido, por ejemplo, valores mínimos de la puntuación Mascot que oscilan entre 69 y 95 (Damodaran et al, 2007) para aceptar como válida una identificación, así como métodos que utilizan valoraciones combinadas entre dos algoritmos de búsqueda como

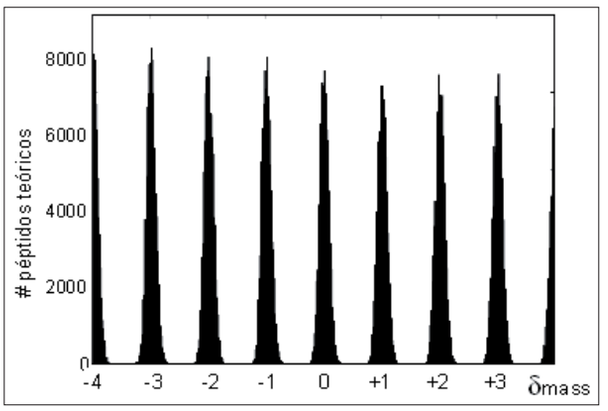


Figura 5.- *Distribución del número de péptidos teóricos posibles en un rango de cuatro unidades de masa alrededor de 1000.42 Da. Se han tenido en cuenta la presencia de los aminoacidos tirosina, serina y treonina fosforilados asi como de metionina oxidada y cisteina carbamidometilada. La figura es un histograma con un total de 500 divisiones generada mediante un script Python.*

Mascot y Profound que, al utilizar métodos de calculo diferentes, aumentan la confianza en la valoración final (Naranjo et al., 2007). En cualquier caso, ante una identificación ambigua o con baja puntuación la estrategia común es llevar a cabo un análisis de secuenciación mediante espectrometría de masas en tándem tal como se describirá más adelante.

En términos de sensibilidad, una regla práctica que suele utilizarse y que es válida en la mayoría de los casos, es que es posible identificar cualquier proteína cuya mancha sea visible en un gel teñido con nitrato de plata (sin especial esfuerzo ni herramientas de tratamiento de imagen, naturalmente). La sensibilidad de los espectrómetros MALDI-TOF actuales es muy superior a las cantidades que esto supone (se requieren entre 10 y 50 fmol de proteína en un gel para generar una banda observable), situándose su límite de detección en unas decenas de atomoles de material situados en la placa. En la práctica, sin embargo, el rendimiento de recuperación de los péptidos de digestión puede ser relativamente bajo, debido a las adsorciones en superficies, al rendimiento de extracción del gel, a la dispersión de la muestra en la placa de MALDI, a las características de la cristalización, etc.

La sensibilidad propia del análisis MALDI-reTOF ha sido mejorada mediante sistemas como el *AnchorChip* (Bruker Daltonik, Alemania) (Schuerenberg et al., 2000) o las placas con matriz predepositada (Qiagen, Alemania). Aunque se ha demostrado la utilidad de diferentes matrices y aditivos para mejorar la respuesta del MALDI-TOF en casos concretos (análisis de péptidos fosforilados, pépti-

dos ácidos), el abanico de matrices utilizadas en la identificación genérica de proteínas mediante PMF no ha sufrido modificaciones en muchos años y el ácido hidroxicinámico sigue siendo la matriz preferida para este tipo de análisis. Pan and col. han revisado los avances más recientes en este sentido así como en los métodos de trabajo y tratamiento de muestra utilizados en el análisis MALDI-TOF de muestras biológicas (Pan et al., 2007).

Uno de los principales factores de éxito en relación a la matriz es la adecuación de la cantidad utilizada en función del tipo de muestra y el procedimiento utilizado para su cristalización. Es además imprescindible el uso de material de alta pureza. El ácido hidroxicinámico se degrada con el tiempo con un efecto negativo sobre la sensibilidad, por lo que es necesario prepararlo fresco. Hace pocos años era común, además, recristalizar la matriz en el laboratorio para mejorar su pureza; sin embargo, en la actualidad este material puede ya adquirirse recristalizado y alicuotado en pequeñas cantidades de forma que se hace innecesario este procedimiento.

La ionización por MALDI es menos sensible a la presencia de sales y otros contaminantes en el medio que el ESI y, en muchos casos, es posible llevar a cabo el análisis por PMF sobre el propio digerido enzimático. Sin embargo, excesos de sales o ciertos aditivos comunes presentes en las muestras como el Tritón o el CHAPS pueden afectar la ionización disminuyendo la sensibilidad del análisis (Zhang y Li, 2004). Por este motivo, en análisis que requieren alta sensibilidad o donde se haya observado una mala cristalización, es aconsejable el desalado y concentración de la muestra. Para ello se utilizan generalmente minicolumnas de fase reversa comerciales como los *Ziptip* de Millipore o las de las empresas New Objective y Proxeon. Estas minicolumnas, similares a una punta desechable de pipeta, contienen en su extremo un pequeño volumen de absorbente cromatográfico (0,05-1 µL) que permite desalar los digeridos (aprox 20-50 µL) y concentrarlos hasta 0.5-1 µL. Este procedimiento puede además automatizarse utilizando robots de digestión y/o deposición de muestra en placas de MALDI (*spotters*).

Uno de los pasos limitantes en el análisis mediante PMF es el de la digestión enzimática de las proteínas en el gel. Diversos autores han estudiado los factores que determinan esta velocidad y se han descrito condiciones que permiten reducir el tiempo de digestión hasta los 30 minutos (Havlis et al., 2003). Se han utilizado además diversas estrategias como la utilización de microondas (Pramanik et al., 2002) y ultrasonidos (López-Ferrer et al, 2005) que favorecen la difusión de las moléculas en el gel y su extracción y que permiten realizar la digestion en menos de 1 minuto.

Por otro lado, se dispone de minicolumnas comerciales con tripsina inmobilizada que permiten disminuir enormemente los tiempos de digestión a unos pocos minutos. Estas columnas sólo son aplicables a proteínas en solución, situación que no es habitual en la identificación por PMF rutinaria y que mayoritariamente se aplica a proteínas separadas en gel. Sin embargo, los inconvenientes derivados de la utilización de geles bidimensionales, y la popularización de las técnicas de análisis del tipo "de arriba hacia abajo" (*top-down)* podrían favorecer en el futuro un mayor desarrollo de las técnicas basadas en el aislamiento de proteínas en fase líquida. Los avances relacionados con la utilización de soportes miniaturizados o microchips que pueden aumentar la velocidad y eficiencia del mapeo peptídico podrían también contribuir a ello. Un ejemplo de estas nuevas tecnologías son los sistemas µTAS de la empresa Gyros que se han utilizado para la preparación de muestra y como placas de MALDI (Palm et al., 2001). Asimismo Marko-Varga y col. han desarrollado un sistema de identificación de proteínas que incluye un digestor, un depositador piezoeléctrico de muestra y placas de MALDI con microceldas, todo realizado en obleas de silicio (Ekstromm et al, 2000), y que permite realizar el analisis PMF completo de forma automatizada a una velocidad de 30 muestras/hora.

Los sistemas MALDI-TOF/TOF (Figura 6) permiten llevar a cabo análisis PMF combinados con análisis de secuenciación automática de uno o más de los péptidos en la muestra. La disponibilidad en muchos laboratorios de este tipo de espectrómetros así como la reciente introducción de los sistemas MALDI-Orbitrap y MALDI-FTICR está generando un cambio drástico en la aplicación del PMF a la identificación de proteínas. Esta aproximación ofrece una gran fiabilidad en las identificaciones obtenidas, permite el análisis de mezclas de proteínas y elimina en muchos casos la necesidad de recurrir a un análisis de secuenciación posterior de las muestras no identificadas. Este análisis doble puede llevarse a cabo también en sistemas MALDI-QIT que, si bien no pueden competir en cuanto a la precisión obtenida en la determinación de masas por MALDI-TOF, compensan esta limitación con una mayor resolución en el aislamiento de los iones precursores y una alta eficiencia en la fragmentación. El MALDI-QIT, sin embargo, tiene que lidiar con la fragmentación de iones monocargados que, a

baja energía de colisión, suelen dar fragmentaciones menos informativas que los iones dicargados y tricargados habitualmente encontrados en las técnicas basadas en ESI (figura 7). Dado que la fragmentación en un sistema TOF/TOF se lleva a cabo a alta energía mientras que en un QIT las colisiones son de baja energía, ambos tipos de espectros muestran diferentes tipos de fragmentos que, pueden ser útiles para la elucidación de la secuencia en casos dudosos utilizando la misma preparación de muestra.

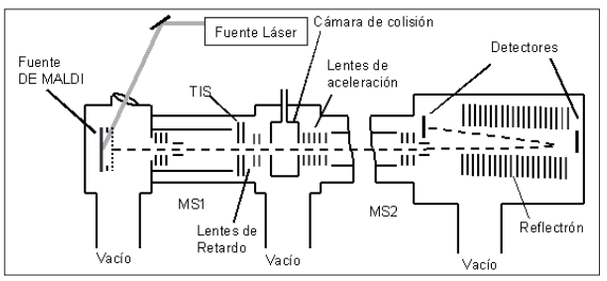


Figura 6.- *Sistema MALDI-TOF/TOF (Modificado de Applied Biosystems*

5. Identificación de proteínas mediante espectrometría de masas en tándem

En 1995 el grupo de John Yates III describía una estrategia para identificar proteínas a gran escala mediante los espectros de masas en tándem obtenidos a partir del ión correspondiente a uno de los péptidos de digestión de la proteína y su comparación con los espectros teóricos derivados de las secuencias anotadas en una base de datos (Yates et al., 1995). El procedimiento puede aplicarse tanto a la identificación de una única proteína aislada mediante cromatografía o electroforesis en gel, de forma similar a como se describió para el PMF, o a la de una mezcla compleja de proteínas mediante el análisis exhaustivo de la colección de péptidos obtenidos después de su digestión (figura 4). Esta última estrategia es la base de las técnicas de análisis de proteomas denominadas *shotgun proteomics* (Figura 7).

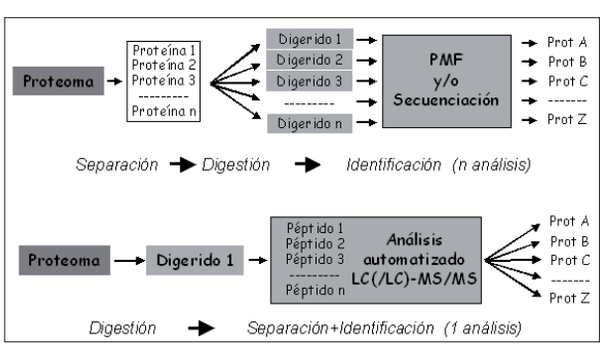


Figura 7.- *Análisis convencional comparado con un análisis shotgun de un proteoma.*

Los analizadores de masas más frecuentemente utilizados para la identificación de proteínas mediante secuenciación han sido las trampas de iones y los sistemas híbridos Q-TOF y Q-LIT, si bien cabe destacar que los trabajos pioneros del grupo de Hunt a principios de los años 90 se llevaron a cabo mediante sistemas de triple cuadrupolo (Hunt et al., 1992). Con este tipo de analizador y utilizando una fuente de microelectrospray este grupo fue capaz de secuenciar diversos péptidos de tamaños entre 8 y 11 aminoácidos en cantidades del orden de los 50 fmols. El "truco" utilizado para conseguir la mayor sensibilidad posible del analizador era la reducción de la resolución de los dos cuadrupolos, especialmente del primero, de forma que se aprovechaba el mayor número posible de iones correspondiente al precursor a pesar de que su señal podía alcanzar alrededor 2 Th de ancho, lo que equivale a una resolución inferior a… ¡300 FWHM!

Las trampas de iones han sido uno de los sistemas más utilizados para la identificación de proteínas tanto en formato de acoplamiento LC/MS como mediante nanospray. Muchos de los trabajos pioneros de *shotgun proteomics* fueron realizados mediante este tipo de sistemas (Wolters et al., 2001). Las trampas son analizadores estables, con una resolución moderada pero que ofrecen un alto rendimiento de fragmentación. En los procedimientos de secuenciación basados en el nanospray, comúnmente utilizados a principios de la década, la capacidad para acumular iones de muestras con señales muy débiles hasta conseguir acumular suficiente señal para obtener un espectro utilizable, hacía de este detector un sistema único.

Los sistemas Q-TOF permitían por su parte la obtención de espectros de fragmentación a resolución alta, con errores entre 2 y 5 ppm. Como en otros sistemas que utilizan un cuadrupolo o trampa como primer analizador, el ión precursor puede asimismo seleccionarse en una ventana relativamente estrecha, en algunos casos inferior a 1 Da. Ambos factores determinan la obtención de espectros adecuados para la búsqueda posterior en bases de datos.

El método de ionización casi exclusivo para llevar a cabo análisis de secuenciación ha sido durante muchos años el ESI en sus diferentes formatos. Mas recientemente, la disponibilidad de los analizadores TOF/TOF (Figura 7) y de los nuevos orbitraps y FT-ICR ha favorecido la reintroducción del MALDI como método de ionización en análisis de secuenciación.

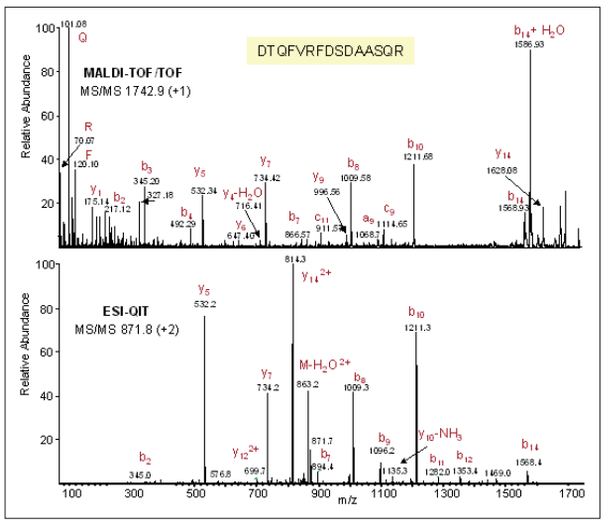


Figura 8.- *Espectros de secuenciación mediante MALDI-TOF/TOF y ESI-QIT del péptido DTQVRFDSDAASQR (M.Carrascal, Tesis de licenciatura, UB 2005).*

El método de fragmentación por excelencia en espectrometría de masas es la disociación inducida por colisión (CID), en el que se producen fragmentos de las moléculas a analizar mediante colisión con átomos o moléculas de gases nobles como el Argón o Xenón (triples cuadrupolos) o Helio (trampas iónicas). Remitimos al lector a otros trabajos para una descripción mas detallada de la fragmentación peptídica mediante CID (Burlingame et al, 1994). La utilización en los sistemas MALDI-reTOF de la fragmentación denominada PSD (*Post Source Decay*) no ha tenido relevancia en este sentido debido probablemente tanto a la complejidad del proceso de adquisición, donde el espectro PSD se obtenía por combinación de varios espectros optimizados cada uno de ellos en un rango de masas particular, como a la baja eficiencia de fragmentación. El primer problema se solucionó en parte mediante reflectores con una distribución hiperbólica de las lentes electrostáticas que permitían un barrido del espectro en una sola etapa y la automatización del análisis. Un metodo para aumentar la eficiencia de fragmentación consiste en la derivatización de los péptidos con el reactivo CAF (*Chemically Activated Fragmentation*, General Electric). Este reactivo une al grupo amino terminal un grupo cargado negativamente que induce una fragmentación PSD más extensa y simplifica además el perfil de fragmentación. Probablemente el inconveniente de añadir una etapa de derivatización que requiere un cuidado especial en las condiciones de reacción, el coste y corta vida media del reactivo asi como la aparición de los sistemas TOF/TOF han sido responsables de la poca difusión de esta herramienta.

Recientemente se han introducido métodos de fragmentación basados en la irradiación de la muestra con luz láser, como en el caso de la disociación por haces de fotones infrarrojos (*infrared multiphoton dissociation*, IRMPD), o con electrones libres o radicales químicos, como en las denominadas disociación por captura electrónica (ECD) y disociación por transferencia electrónica (ETD), respectivamente, que ofrecen una información distinta a la obtenida por CID. En el primer procedimiento, la radiación láser se utiliza para transferir energía y fragmentar los iones del péptido. La ECD por su parte utiliza haces de electrones de baja energía que pueden ser absorbidos de forma resonante por los iones del péptido. Este proceso produce principalmente fragmentaciones en la cadena lineal del péptido, respetando en gran medida las modificaciones de las cadenas laterales por lo que es un procedimiento de interés en el análisis de PTMs y es especialmente eficiente en el caso de polipéptidos con alto número de carga y de tamaño grande. La ECD implica un proceso de fragmentación muy rápido en el que no tiene lugar una redistribución de la energía absorbida en forma de energía vibracional en la molécula. En consecuencia, y a pesar de una eficiencia de fragmentación relativamente baja, la ECD produce un perfil de fragmentación más homogéneo e independiente de la secuencia aminoacídica, más rico en información de secuencia que los obtenidos por CID. La necesidad de una nube de electrones térmicos para la captura electrónica no permite la utilización de esta técnica con trampas del tipo QIT o LIT debido a que los electrones adquieren rápidamente energía dentro del campo de radiofrecuencia de estos analizadores. Para resolver este problema, el grupo de Hunt, utilizando una vez más su experiencia en los procesos de ionización química en fase gaseosa, desarrolló el método ETD en el que se utilizan radicales aniónicos capaces de interaccionar con los iones con carga positiva del péptido y transferirles un electrón produciendo un tipo de fragmentación parecido al de ECD, rico en iones c y z (Syka et al., 2004; Chi et al., 2007).

Tanto el ECD como el IRMPD se encuentran implementados principalmente en sistemas FT-ICR ya que su uso no afecta al mantenimiento del alto vacío necesario en estos analizadores. Otros procedimientos como el CID o ETD deben realizarse fuera de la celda ICR o bien utilizar técnicas como la CID mediante irradiación sostenida no resonante (SORI-CID, *sustained off-resonance irradiatión collision-induced dissociation*) que implican el llenado de la

celda con gas de colisión y su posterior eliminación previa al análisis de los fragmentos. La disociación inducida por superficie (SID, Surface Induced Dissociation), es otro procedimento de activacion que se adapta bien a los FT-ICR especialmente combinado con el MALDI y que se caracteriza por proveer de una mayor información estructural que el CID. Aunque el SID ha sido objeto de atención en los últimos años (Laskin y Futrell, 2004; Fernandez et al, 2006) su uso es todavía muy reducido.

La búsqueda en bases de datos de las proteínas en base a su espectro de fragmentación se lleva a cabo, como en el caso del PMF, mediante programas informáticos (Tabla 2). En los trabajos pioneros de Yates y Eng estos autores desarrollaron un algoritmo de correlación que fue el origen del conocido programa SEQUEST. Actualmente existen diversos programas para realizar esta tarea (Ver Tabla 1). A pesar de este abanico de herramientas, la valoración manual por un experto de las secuencias asignadas a un espectro sigue siendo la validación más robusta de esta asignación. La complejidad del proceso de identificación difiere de forma importante entre un procedimiento convencional donde se identifican una o unas pocas proteínas aisladas en cada análisis y donde frecuentemente se ha realizado un análisis PMF previo y un análisis shotgun en el que deben analizarse de forma exhaustiva grandes colecciones de péptidos con objeto de describir los componentes de la muestra. En el primer caso es frecuente disponer de información sobre varios espectros que apuntan a zonas distintas de la misma proteína validando por tanto la identificación.

En los casos en los que la cantidad de material es muy baja, es posible, sin embargo, que la identificación dependa de un único espectro. Cuando se dispone de un análisis PMF que apunta hacia una proteína con alguna secuencia coherente con el espectro de fragmentación este conjunto de datos puede permitir validar definitivamente la asignación incluso cuando el espectro de fragmentación por si solo no permite inferir un subsecuencia (*tag*) utilizable. El proceso inverso también es posible, es decir, que una identificación obtenida mediante secuenciación pueda confirmarse mediante la localización de algunos de los péptidos esperados para esa proteína en un espectro de PMF que por algún motivo (interferencias, error de medida) no resultó previamente en una asignación clara. En un análisis de tipo shotgun a gran escala no es posible utilizar estos procedimientos de valoración ya que las proteínas han sido digeridas previamente al análisis y el concepto de una familia de péptidos agrupada en una determinada unidad de análisis (una mancha de gel o una fracción cromatográfica) no es aplicable. Dado que los algoritmos de búsqueda pueden producir identificaciones de alta puntuación incluso con espectros de ruido y que la validación manual es sólo practicable con un número reducido de espectros, resulta dificil separar de forma absolutamente fiable las asignaciones correctas de las incorrectas, siendo este uno de los problemas más graves en este tipo de tareas. La utilización de filtros de calidad de los espectros previa a la búsqueda es un factor vital para evitar la producción de asignaciones casuales y el uso de algoritmos diferentes de búsqueda permite aumentar la confianza en las asignaciones realizadas. La proporcion estimada de falsos positivos en una poblacion de peptidos identificados con una determinada confianza viene expresada por la tasa de asignaciones incorrectas o FDR (*false discovery rate*). La FDR puede determinarse de forma eficaz determinando el número de asignaciones de la misma confianza que se obtendría a partir de la misma colección de espectros en una busqueda contra una base de datos señuelo conteniendo secuencias aleatorizadas o invertidas (“mock” o “decoy databases” (Elias y Gygi, 2007) (Figura 8). Comúnmente la colecciones descritas en la literatura se selecionan con valores de FDR inferiores al 2%.

La identificación de proteínas o péptidos en el caso de especies con genomas poco o nada secuenciados requiere el uso de métodos de secuenciación junto con programas de interpretación *de novo* de espectros. En el análisis de novo las secuencias o trozos de secuencia peptídica deben caracterizarse mediante métodos de elucidación manual o algoritmos matemáticos. La proteína o el posible gen de origen deben identificarse en un segundo paso mediante análisis de similitud de secuencias mediante BLAST o técnicas similares. Cuando se utiliza BLAST debe seleccionarse una matriz adecuada para el análisis que tenga en cuenta el tamaño pequeño de las secuencias problema. Asimismo debe tenerse en cuenta que la puntuación de esta herramienta valora la semejanza funcional de las secuencias y no su semejanza química.

En general, los algoritmos disponibles actualmente para el análisis *de novo* producen resultados cuya bondad resulta a veces difícil de evaluar. Algunos de los problemas asociados a este tipo de búsqueda ya han sido comentados más arriba en relación a los algoritmos de búsqueda en bases de datos. Todos los programas de búsqueda que hemos podido utilizar tienden a producir una gran cantidad de secuencias no correctas y son frecuentes las secuencias puntuadas con alta fiabilidad a partir de

espectros de ruido que no corresponden a péptidos. Los espectros de fragmentación por CID no suelen contener una información completa de la secuencia lo que produce zonas ambiguas donde más de una combinación de aminoácidos o un distinto orden de estos es posible. Algunas herramientas de búsqueda como Lutefisk marcan estas zonas de secuencia con el valor de su peso molecular mientras que otras, más voluntariosas, rellenan estos huecos siempre. La problemática de esta estrategia se evidencia en alguno de estos buscadores por secuencias asignadas de alta probabilidad donde estos huecos se rellenan con secuencias conteniendo un alto numero de W, C y Y. Aunque el triptófano es un aminoácido poco frecuente, su alto peso molecular permite rellenar los espacios vacíos de forma más efectiva, objetivo final del algoritmo de cálculo.

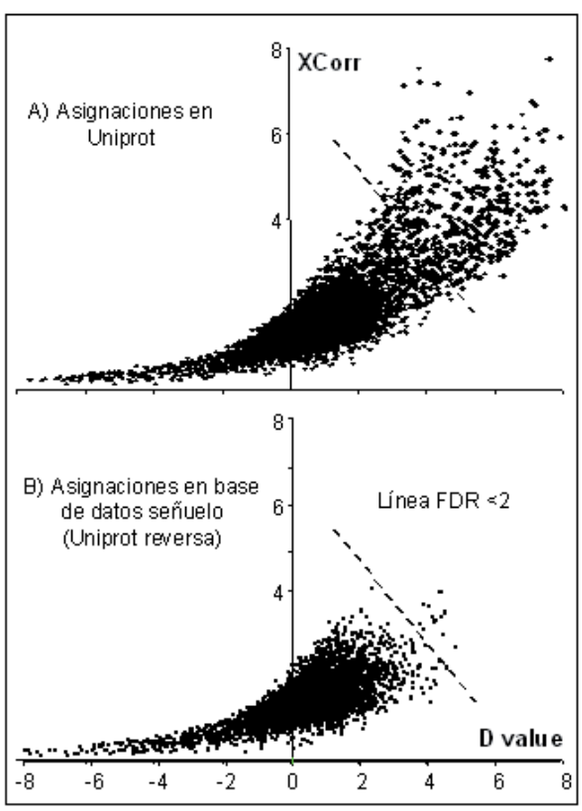


Figura 9.- *Utilización de los parámetros Xcorr (SEQUEST) y D (Peptide Prophet) para la determinación de los valores de corte (líneas discontinuas) que producen resultados con un bajo porcentaje de falsos positivos (FDR < 2) en una muestra enriquecida en fosfopéptidos. La base de datos señuelo se ha generado a partir de la base de datos Uniprot. Imagen superior: espectros que apuntan a secuencias "reales" en la base de datos Uniprot. Imagen inferior: espectros que apuntan a secuencias invertidas. (Carrascal et al. Journal Proteomics Research 7, 5167-5176, 2008).*

El uso de instrumentos de alta precisión en la medida de la masa, tanto de los iones precursores como de los productos, supone una importante ventaja para la asignación de las secuencias correctas (Frank et al., 2007). Tal como se describió previamente, la principal ventaja de los sistemas tipo MALDI acoplados a analizadores TOF/TOF o TOF-LIMS es la capacidad para llevar a cabo para el análisis combinado por PMF y de fragmentación en ambos casos con errores en el cálculo de masa del orden de los 5 ppm. Este mismo tipo de estrategia puede llevarse a cabo en sistemas orbitrap y FT-ICR con precisiones aun mayores.

La eficiencia del análisis de secuencia puede aumentarse utilizando datos espectrométricos complementarios como la composición de aminoácidos obtenida a partir del análisis de los iones imonio o de medidas de alta precisión, o el uso de marcaje isotópico. Uno de los métodos más extendidos y prácticos es la combinación de datos MS^2 y MS^3 del péptido problema que pueden obtenerse en sistemas QIT o LIT (Olsen y Mann, 2004). ECD y ETD producen fragmentos con información de secuencia que es complementaria a la obtenida mediante CID (Creese y Cooper, 2007). McLafferty y cols desarrollaron un programa para el análisis conjunto de la información obtenida mediante CID y ECD (Horn et al., 2000). Más recientemente esta estrategia se ha aplicado en sistemas FT-ICR, produciendo espectros CID y ECD alternativos para cada una de los iones de interés (Savitski et al., 2005).

6. Identificacion "*top-down*" de proteinas

La nueva instrumentación para espectrometría de masas exactas o de alta resolución introducida en el campo de la proteómica (FT-ICR y Orbitrap) permite la identificación de proteínas sin necesidad de la digestión previa de las muestras para producir péptidos de pequeño tamaño susceptibles de ser fragmentados o analizados con las técnicas convencionales. En estos instrumentos es posible fragmentar el ión molecular de proteínas completas con suficiente eficiencia como para, a partir de la medida exacta de la masa del ión molecular y de los fragmentos producidos, identificar la secuencia de procedencia. Esta estrategia, denominada de análisis de arriba hacia abajo (*top-down proteomics*), fue descrita ya en 1996 por Mortz et al (Mortz et al., 1996). Previamente, los grupos de Smith y Feng habian demostrado la posibilidad de romper iones multicargados de hasta 150 Kda mediante CID en analizadores de triple quadrupolo (Fenn y Konishi, 1993) o por colision en la fuente de electrospray

(Loo et al., 1990) y McLafferty había aplicado con el mismo objetivo la fragmentacion por SORI-CID en un FTICR (Senko et al., 1994).

Una de las ventajas de este tipo de aproximación es que, al no fragmentar la muestra mediante digestión enzimática y trabajar sobre la molécula proteica completa, se disminuye enormemente la complejidad de la muestra y se eliminan las ambiguedades derivadas de la presencia de proteínas con secuencias homólogas característica del análisis de abajo a arriba aumentando por tanto la especificidad de la identificación.

Actualmente, las técnicas *top-down* utilizan ventajosamente métodos de fragmentación como el ETD o ECD que se han descrito anteriormente y que se adaptan bien a los analizadores FT-ICR y Orbitrap. Si bien la espectrometría de masas de alta resolución no está actualmente al alcance de todos los laboratorios, la rápida introducción de estos instrumentos va a tener una notable influencia en la forma en que los laboratorios de proteómica abordaran la identificación de proteínas en los próximos años.

Agradecimientos

Parte de los datos experimentales de este trabajo han sido obtenidos dentro de los proyectos BIO2004-01788 y BIO2008-03365 del Ministerio de Ciencia y Tecnología. LP-CSIC/UAB es miembro del Instituto Nacional de Proteómica (ProteoRed, http://www.proteored.org/), financiado por Genoma España.

Referencias

Aleksandrov M L, Barama G I, Gall L M, Krasnov N V. et al. 1985. *Formation of beams of quasi-molecular ions of peptides from solutions.* Bioorganicheskaya Khimiya 11:700-704.

Barber M, Bordoli R S, Sedgwick R D, Tyler A N. 1981. *Fast atom bombardment of solids (FAB): a new ion source for mass spectrometry.* Journal of the Chemical Society, Chemical Communications 7:325-327.

Bern M W, Goldberg D. 2006. *De novo analysis of peptide tandem mass spectra by spectral graph partitioning.* Journal of Computational Biology 13: 364-78.

Boyle J G, Whitehouse C M. 1992. *Time-of-flight mass spectrometry with an electrospray ion beam.* Analytical Chemistry 64:2084-2089.

Burlingame A L, Boyd R K, Gaskell S J. 1994. *Mass Spectrometry*. Analytical Chemistry 66:634R-683R.

Clauser K R, Baker P, Burlingame A L. 1999. *Role of Accurate Mass Measurement (10 ppm) in Protein Identification Strategies Employing MS or MS/MS and Database Searching.* Analytical Chemistry 71:2871-2882.

Cleveland D W, Fischer S G, Kirschner M W, Laemmli U K. 1977. *Peptide Mapping by Limited Proteolysis in Sodium Dodecyl Sulfate and Analysis by Gel Electrophoresis.* Journal of Biological Chemistry 252:1102-1106.

Cooper H J, Håkansson K, Marshall A G. 2005. *The role of electron capture dissociation in biomolecular analysis.* Mass Spectrometry Reviews 24:201–222.

Creese A J, Cooper H J. 2007. *Liquid Chromatography Electron Capture Dissociation Tandem Mass Spectrometry (LC-ECD-MS/MS) versus Liquid Chromatography Collision-induced Dissociation Tandem Mass Spectrometry (LC-CID-MS/MS) for the Identification of Proteins.* Journal of the American Society for Mass Spectrometry 18:891–897.

Chamrad D C, Körting G, Stühler K, Meyer H E et al. 2004. *Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data.* Proteomics 4:619-628.

Chi A, Huttenhower C, Geer L Y, Coon J J. et al. 2007. *Analysis of phophorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America 104:2193-2198.

Dancík V, Addona T A, Clauser K R, Vath J E. et al. 1999. *De novo peptide sequencing via tandem mass spectrometry.* Journal of Computational Biology 6:327-42.

Damodaran S, Wood T D, Nagarajan P, Rabin R A. 2007. *Evaluating Peptide Mass Fingerprinting-based Protein Identification.* Genomics, Proteomics & Bioinformatics 5:152-157.

Elias J E, Gygi S P. 2007. *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry* Nature Methods 4:207-214.

Emmett M R, Caprioli R M. 1994. *Micro-electrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins.* Journal of

the American Society for Mass Spectrometry 5:605-613.

Eng J K, McCormack A L, Yates III J R. 1994. *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* Journal of the American Society for Mass Spectrometry 5:976-989.

Ekström S, Onnerfjord P, Nilsson J, Bengtsson M. et al. 2000. *Integrated microanalytical technology enabling rapid and automated protein identification.* Analytical Chemistry 72:286-293.

Fernandez F M, Wysocki V H, Futrell J H, Laskin J. 2006. *Protein Identification via Surface-Induced Dissociation in an FT-ICR Mass Spectrometer and a Patchwork Sequencing Approach*. Journal of the American Society for Mass Spectrometry 17:700-709.

Feng R, Konishi Y. 1993. *Collisionally-Activated Dissociation of Multiply Charged 150-kDa Antibody Ions*. Analytical Chemistry 65:645-649.

Frank AM, Savitski MM, Nielsen ML, Zubarev RA. et al. 2007. *De Novo Peptide Sequencing and Identification with Precision Mass Spectrometry.* Journal of Proteome Research 6:114-123.

Gohlke RS. 1959. *Time-of-flight mass spectrometry and gas-liquid partition chromatography.* Analytical Chemistry 31:535-541.

Hass MJ. 1999. *Fully Automated Exact Mass Measurements by High-resolution Electrospray Ionization on a Sector Instrument.* Rapid Communications in Mass Spectrometry 13:381–383.

Havlis J, Thomas H, Sÿebela M, Shevchenko A. 2003. *Fast-Response Proteomics by Accelerated In-Gel Digestion of Proteins*. Analytical Chemistry 75:1300-1306.

Henzel J, Watanabe C, Stults J T. 2003. *Protein Identification: The Origins of Peptide Mass Fingerprinting*. Journal of the American Society for Mass Spectrometry 14:931–942.

Henry K D, Williams E R, Wang B -H, McLafferty F W. et al. 1989. *Fourier-transform mass spectrometry of large molecules by electrospray ionization.* Proceedings of the National Academy of Sciences of the United States of America 86:9075–9078.

Henzel W J, Billeci T M, Stults J T, Wong S C. et al. 1993. *Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases (mass spectrometry/in situ digestion/two-dimensional electrophoresis).* Proceedings of the National Academy of Sciences of the United States of America 90:5011-5015.

Horn, D M, Zubarev, R A., McLafferty, F W. 2000. *Automated De Novo Sequencing of Proteins by Tandem High-resolution Mass Spectrometry*. Proceedings of the National Academy of Sciences of the United States of America 97, 10313–10317.

Hu Q, Noll R J, Li H, Makarov A. et al. 2005. *The Orbitrap: a new mass spectrometer*. Journal of Mass Spectrometry 40:430–443.

Hunt D F, Henderson R A, Shabanowitz J, Sakaguchi K. et al. 1992. *Characterization of peptides bound to Class I MHC Molecule HLA-A2 1 by mass spectrometry*. Science 255:1261-1263.

Hunt D F, Michel H, Dickinson T A, Shabanowitz J. et al. 1992. *Peptides presented to the immune system by the murine class II mayor histocompatibility complex molecule I-A.* Science 256:1817-1820.

Hunter R L, McIver R T, Li Y, Hunter R L. 1994. *Ultra-high resolution laser desorption mass spectrometry of peptides and small proteins*. Proceedings of the 42nd ASMs Conference on Mass Spectrometry and Allied Topics, Chicago, Illinois, May 29-June 3, pg245.

Karas M, Hillenkamp F. 1988. *Laser desorption ionization of proteins with molecular masses exceeding 10000 Da.* Analytical Chemistry 60:2299-2301.

Knight RD. 1981. *Storage of ions from laser induced plasmas.* Applied Physics Letters 38:221-223.

Laskin J, Beck K M, Hache J J, Futrell J H. 2004. *Surface-Induced Dissociation of Ions Produced by Matrix-Assisted Laser Desorption/Ionization in a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer.* Analytical Chemistry 76:351-356.

Laskin J, Futrell J H. 2005. Activation of large ions in FT-ICR mass spectrometry. Mass Spectrometry Reviews 24:135-67.

Little D P, Speir J P, Senko M W, O'Connor P B. et al. 1994. Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. Analytical Chemistry 66:2809-2815.

Loo J A, Edmonds C G, and Smith R D. 1990. Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. Science 248:201-204.

Lopez-Ferrer D, Capelo J L, Vazquez J. 2005. *Ultra Fast Trypsin Digestion of Proteins by High Intensity Focused Ultrasound*. Journal of Proteome Research 4:1569-1574.

Makarov A. 2000. *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis.* Analytical Chemistry 72:1156.

March R E. 1997. *An Introduction to Quadrupole Ion Trap Mass Spectrometry.* Journal of Mass Spectrometry 32:351-369.

Medzihradszky K F, Campbell J M, Baldwin M A, Falick A M. et al. 2000. *The Characteristics of Peptide Collision-Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer.* Analytical Chemistry 72:552-558.

Meng C K, Mann M, Fenn J B. 1988. Of Protons or Proteins - A Beam‘s a Beam For a‘That (Burns,O.S.). Zeitschrift Fur Physik D-Atoms Molecules And Clusters 10:361-368.

Mikesh L M, Ueberheide B, Chi A, Coon J J. et al. 2006. *The utility of ETD mass spectrometry in proteomic analysis.* Biochimica et Biophysica Acta 1764:1811-1822.

Mortz E, O’connor P B, Roepstorff P, Kelleher N L. et al. 1996. *Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases.* Proceedings of the National Academy of Sciences of the United States of America 93:8264-8267.

Murray K K, Boyd R K, Eberlin M N, Langley G J, et al. International Union of Pure and Applied Chemistry Analytical Chemistry Division *Standard definitions of terms relating to mass Spectrometry* (IUPAC Recommendations 2006)

Naranjo V , Villar M , Martín-Hernando M P , Vidal D , et al. 2007. *Proteomic and transcriptomic analyses of differential stress/inflammatory responses in mandibular lymph nodes and oropharyngeal tonsils of European wild boars naturally infected with Mycobacterium bovis.* Proteomics 7:220-231.

Olsen J V and Mann M. 2004. *Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation.* Proceedings of the National Academy of Sciences of the United States of America 101:13417-13422.

Palm A , Wallenborg S R , Gustafsson M , Hedstrom A , et al. 2001. *En Micro Total Analysis Systems* (Ramsey, M , van den Berg, A , Eds) Kluwer Academic Publishers: Dordrecht, The Netherlands, 216-218.

Pan Ch, Xu S, Zhou H, Fu Y. et al. 2007. *Recents developments in methods and technology for analysis of biological samples by MALDI-TOF-MS*. Analytical and Bioanalytical Chemistry 387:193-204.

Pappin D J C, Hojrup P, Bleasby A J. 1993. *Rapid Identification of Proteins by Peptide Mass Fingerprinting.* Current Biology 3:327–332.

Pappin D J C, Rahman D, Hansen H F, Bartlet-Jones M. et al. 1994. *Further development of a peptide-mass database for the rapid identification of proteins by mass spectrometric techniques* Proceedings of the 42nd ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, Illinois, May 29-June 3, pg 312.

Pastor S J, Castoro J A., Wilkins Ch L. 1995. *High Mass analysis using quadrupolar excitation/ion cooling in a Fourier transform Mass Spectrometer*. Anal.Chem. 67, 379-384.

Paul W and Steinwedel H S. 1953. *Mass spectrometer without magnetic field* Z Naturforsch , 8a, 448-450.

Paul W , Steinwedel H, U S Patent 1960, 2939952.

Pramanik B N , Mirza U A , Ing Y H , Liu Y H , et al. 2002. *Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes.* Protein Sciences , 11:2676-2687.

Price P. 1991. *Standard Definitions of Terms Relating to Mass Spectrometry.* Journal of the American Society for Mass Spectrometry 2, 336-348 [Ver también Murray et al].

Roepstorff P, Richter W J. 1992. *Status of, and developments in, mass spectrometry of peptides and proteins.* International Journal of Mass Spectrometry and Ion Processes 118/119:789-809.

Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P. 1992. *In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis* Analytical Biochemistry 203:173-179.

Savitski M M, Kjeldsen F, Nielsen M L, Zubarev R A. 2006. *Complementary sequence preferences of electron-capture dissociation and vibrational excitation in fragmentation of polypeptide polycations*. Angewandte Chemie International Edition English 45:5301-3.

Savitski M M, Nielsen M L, Kjeldsen F, Zubarev R A. 2005. *Proteomics-grade De Novo Sequencing Approach*. Journal Proteome Research 4:2348–2354.

Scigelova M and Makarov A. 2006. Orbitrap Mass Analyzer – *Overview and Applications in Proteomics* Practical Proteomics 1-2:16-21

Schuerenberg M , Luebbert C , Eickhoff H , Kalkum M , et al. , 2000. *Prestructured MALDI-MS sample supports*. Analytical Chemistry 72:3436-3442.

Schwartz J C, Senko M W, Syka J E P. 2002. *A Two-Dimensional Quadrupole Ion Trap Mass Spectrometer* . Journal of the American Society for Mass Spectrometry 13:659–669.

Shadforth I, Crowther D, Bessant C. 2005. *Protein and peptide identification algorithms using MS for use in high-throughput,automated pipelines*. Proteomics 5:4082-4095.

Senko M W, Speir J P, McLafferty F W. 1994. *Collisional Activation of Large Multiply Charged Ions Using Fourier Transform Mass Spectrometry*. Analytical Chemistry 66:2801-2808.

Spengler B. 2004. *De Novo Sequencing, Peptide Composition Analysis, and Composition-Based Sequencing: A New Strategy Employing Accurate Mass Determination by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. Journal American Society Mass Spectrometry 15, 703–714.

Stafford G C, Kelley P E, Syka J E P, Reynolds WE, Todd J F J. 1984. *Recent improvements in and analytical applications of advanced ion trap technology*. International Journal of Mass Spectrometry and Ion Processes 60:185-98.

Syka J E, Coon J J, Schroeder M J, Shabanowitz J. et al. 2004. *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry* Proceedings of the National Academy of Sciences of the United States of America 101:9528-33.

Syrstad E A, Turecek F. 2005. *Toward a general mechanism of electron capture dissociation.* Journal of the American Society for Mass Spectrometry 16:208-24.

Tanaka K, Waki H, Ido Y, Akita S, Yoshida I, Yoshida T. 1988. *Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry*. Rapid Communications in Mass Spectrometry 2:151-153.

Thorgerson D F, Mcfarlane R D and Skowronski R P. 1974. *New approach to the mass spectroscopy of non volatile compounds.* Biochemical Biophysical Research Communications 60:616-621.

Wiley W C, McLaren J B. 1955.*Time-of-flight mass spectrometer with improved resolution.* Review of Scientific Instruments 26:1150-1157.

Wilm M S, Mann M. 1994. *Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?*. International Journal of Mass Spectrometry and Ion Processes 136:167-180.

Whitehouse C R, Dreyer R N, Yamashita M, Fenn J B. 1985. *Electrospray interface for liquid chromatographs and mass spectrometers*. Analytical Chemistry 57:675-679.

Wöhlbrand L, Kallerhoff B, Lange D, Hufnagel P. et al. 2007. *Functional proteomic view of metabolic regulation in "Aromatoleum aromaticum "strain EbN1*. Proteomics 7:2222-2239.

Wolters D A, Washburn M P, Yates J R 3rd. 2001. *An automated multidimensional protein identification technology for shotgun proteomics.* Anal Chem 73:5683-90.

Yates III J R, Eng J K, McCormack A L, Schieltz D. 1995. *Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database.* Analytical Chemistry 67:1426-1436.

Zhang W, Chait B T. 2000. *ProFound: An Expert System for Protein Identification Using Mass Spectrometric Peptide Mapping Information* Analytical Chemistry 72:2482-2489.

Zhang N, Li L. 2004. *Effects of common wsurfactants on protein digestion and matrix assisted laser desorption ionization mass spectrometric analysis of the digested peptides using two-layer sample preparation*. Rapid Communications in Mass Spectrometry 18:889-896.

Zubarev R A, Kelleher N L, McLafferty F W. 1998. *Electron Capture Dissociation of Multiply Charged Protein Cations - a Nonergodic Process* . Journal of the American Chemical Society 120:3265-3266.

La inevitable estadística… en el mundo de la proteómica

Ana María Rodríguez Piñeiro[1]*, María Páez de la Cadena, Francisco Javier Rodríguez Berrocal*

Departamento de Bioquímica, Genética e Inmunología, Facultad de Biología, Campus As Lagoas-Marcosende, Universidad de Vigo, 36310 Vigo (Pontevedra)

Resumen

En los estudios de proteómica, especialmente en los orientados a la búsqueda de biomarcadores, resulta imprescindible el uso de métodos estadísticos. La proteómica conlleva un tipo de experimentación que implica en la mayoría de los casos el manejo de un gran número de variables (los *spots* en los geles bidimensionales o los picos en los espectros de masas) pero, por el contrario, un escaso número de muestras. Los conjuntos de datos obtenidos en proteómica pueden tratarse mediante métodos univariantes, pero este tipo de técnicas incrementa la posibilidad de que aparezcan falsos positivos y no permite la detección de tendencias. Por el contrario, su análisis mediante métodos estadísticos multivariantes resulta mucho más eficiente. Estos métodos permiten la reducción de la dimensionalidad del conjunto de datos, o la utilización de estrategias de validación estadística como la validación cruzada o las pruebas de permutación siendo en general métodos más robustos y menos susceptibles a factores que induzcan confusión.

Por otro lado, un paso imprescindible en el establecimiento de marcadores es su validación biológica. Nuevamente, aquí, es necesario el uso de métodos estadísticos que permitan conocer el valor o la utilidad que tiene una determinada proteína, o varias, a la hora de diagnosticar la enfermedad, utilizarse como indicador pronóstico o para seguir la evolución de un paciente. Basándonos en nuestra propia experiencia, en este trabajo se revisan los métodos estadísticos, tanto univariantes como multivariantes, que se utilizan más frecuentemente en la búsqueda de biomarcadores mediante 2-DE y en su validación.

Palabras clave:

Estadística, univariante, multivariante, electroforesis bidimensional, proteómica, biomarcadores.

1. Necesidad de la utilización de pruebas estadísticas en proteómica: consideraciones previas

La proteómica se considera, al cabo de más de 10 años de su definición por Wasinger y Wilkins (Wasinger *et al.*, 1995; Wilkins *et al.*, 1996), como el estudio de proteínas a gran escala, incluyendo su identificación, cuantificación, localización, estructura, modificaciones, interacciones, actividades y función. Aunque estos objetivos abarcan una gran variedad de técnicas, los métodos proteómicos por excelencia son la 2-DE[2] y la espectrometría de masas. Ambas técnicas son ampliamente utilizadas debido a su gran resolución, ya que permiten la detección de cientos a miles de proteínas en un solo análisis. Sin embargo, esto se traduce en complicaciones a la hora de evaluar conjuntamente la gran cantidad de datos obtenidos y de ahí la importancia de la aplicación de las técnicas estadísticas correctas, que permitan analizar esa ingente cantidad de datos y obtener conclusiones significativas.

En particular, en los estudios de proteómica orientados a la búsqueda de biomarcadores, es decir, proteínas concretas diferenciadoras de un estado biológico (por ejemplo, una enfermedad), resulta imprescindible el uso de métodos estadísticos. Como ya han argumentado otros autores, aunque los conjuntos de datos obtenidos en proteómica pueden tratarse mediante métodos univariantes, este tipo de técnicas incrementa la posibilidad de que aparezcan

[1] Autor para la correspondencia

[2] 2-DE: electroforesis bidimensional.

falsos positivos y no permite la detección de tendencias (Dudoit *et al.*, 2003; Karp *et al.*, 2005; Wilkins *et al.*, 2006). Por el contrario, su análisis mediante métodos estadísticos multivariantes resulta mucho más eficiente (Karp *et al.*, 2005).

El desarrollo de este tipo de experimentación implica en la mayoría de los casos dos problemas estadísticos que se han dado en llamar maldiciones (de Noo *et al*., 2006). La primera es la "maldición de la dimensionalidad" (*curse of dimensionality*), provocada porque el número de variables (los *spots* en los geles bidimensionales o los picos en los espectros de masas) está en el orden de millares a decenas de millares. La segunda es la "maldición de la dispersión del conjunto de datos" (*curse of dataset sparsity*), debida a un escaso número de muestras. Estos dos problemas afectan especialmente a los métodos univariantes, y se traducen en ocasiones en el descubrimiento de proteínas que discriminan las poblaciones a estudio, aun cuando estas no son en realidad diferentes. Este fenómeno se denomina "sobreajuste" (*overfitting*), y da lugar a que el modelo de discriminación encontrado no funcione adecuadamente cuando se analizan nuevas muestras. Para evitar este inconveniente existen varias soluciones, siendo una de ellas la reducción de la dimensionalidad del conjunto de datos, o la utilización de estrategias de validación estadística como la validación cruzada o las pruebas de permutación (Smit *et al.*, 2007). Aunque las aproximaciones multivariantes también se ven afectadas por estas dos maldiciones, estos métodos son en general más robustos y menos susceptibles a factores que induzcan a confusión.

Por otro lado, un paso imprescindible en el establecimiento de marcadores es su validación biológica. Estos estudios tienen, si cabe, una importancia mayor cuando se trata de encontrar buenos marcadores en relación con una determinada enfermedad. Nuevamente aquí, es imprescindible el uso de métodos estadísticos que permitan conocer el valor o la utilidad que tiene una determinada proteína, o varias, a la hora de diagnosticar la enfermedad, utilizarse como indicador pronóstico, o para seguir la evolución de un paciente. Estas capacidades se reflejan, en general, en la certeza con la que las proteínas seleccionadas clasifican correctamente las muestras en su verdadero grupo de origen.

Basándonos en nuestra propia experiencia (Ayude *et al.*, 2000; Rodríguez-Piñeiro *et al*., 2005; Lemos-González *et al.*, 2007; Rodríguez-Piñeiro *et al*., 2007) en este trabajo se revisan los métodos estadísticos, tanto univariantes como multivariantes, que se utilizan más frecuentemente en la búsqueda de biomarcadores mediante 2-DE. En general, la estadística está presente en casi todas las fases del análisis de datos, incluso cuando el investigador no es consciente de ello (por ejemplo si está empleando programas especializados que "esconden" los procesos matemáticos del análisis). Sin embargo, en esta revisión nos centraremos en aquellos métodos que el investigador emplea de forma "consciente". Por último, dedicaremos un apartado a describir las pruebas estadísticas que permiten determinar la sensibilidad y especificidad de un marcador o su capacidad de pronóstico.

2. Análisis estadístico de mapas bidimensionales de proteínas

Una de las ventajas de la utilización de la 2-DE es que el proteoma puede visualizarse mediante mapas en donde las proteínas aparecen como manchas o *spots*, separadas en función de su punto isoeléctrico y de su masa molecular. Un proceso fundamental antes de su obtención es un cuidadoso diseño experimental, lo que ayuda a reducir los errores sistemáticos, mejorando la precisión de los análisis estadísticos posteriores y disminuyendo el número de falsos positivos. A este respecto, Chich *et al*. (2007) han propuesto recientemente la utilización de un diseño por bloques de muestras que sean robustos y la aleatorización de las muestras en cada bloque.

Dada la complejidad de estos mapas, su análisis ha de realizarse a través de un programa específico para imágenes 2D[3]. En general, estos programas se basan en asignar, a cada uno de los puntos (píxeles) detectados en un gel, unas coordenadas *X* e *Y*, que corresponden a su posición horizontal y vertical en la imagen, y un valor *Z* que determina la densidad óptica. De esta forma, las coordenadas *X*, *Y* y *Z* de cada mancha proteica vienen determinadas por las correspondientes a los puntos que la forman. Esto proporciona el área de cada *spot*, así como su valor de intensidad óptica, y por lo tanto una medida cuantitativa (volumen) de cada proteína.

Cuando las manchas proteicas son detectadas, la imagen original del gel es filtrada y el volumen de los *spots* se ajusta a un modelo gaussiano, sobre el que se puede realizar la comparación. En caso

[3] 2D: bidimensional/es.

necesario, existen varios métodos con los que el investigador puede estimar los datos perdidos, como reemplazar los valores por 0, por un valor de densidad mínimo detectable en el gel, por la media de las medidas del mismo *spot* en otras muestras, etc. (Chang *et al.*, 2004; Chich *et al*., 2007). A continuación se suelen normalizar los valores de los *spots* en todos los geles analizados, en la mayoría de los casos obteniendo su volumen relativo, que es comparable entre geles y que permite la aplicación de los métodos estadísticos. En relación a este último paso, cabe destacar el estudio de Randic *et al*. (2006), quienes recientemente han demostrado que existe un número máximo de *spots*, que ellos proponen ser los 300 *spots* más intensos, a partir del cual no es posible obtener más información estadística, es decir, que los datos derivados de las restantes proteínas no aportan mayor información sobre el mapa proteómico que se quiere caracterizar. Esta afirmación no es aplicable en el caso de la búsqueda de variaciones en proteínas concretas, especialmente en la búsqueda de biomarcadores, que suelen ser proteínas de baja abundancia. Sin embargo, sí debe ser tenido en cuenta cuando se pretende utilizar el propio mapa proteico para caracterizar a un grupo de población.

Por otra parte, antes de realizar la comparación entre mapas 2D de diversas muestras poblacionales, el investigador debe comprobar la reproducibilidad de la técnica, proceso en el que también está presente la estadística ya que, por ejemplo, es necesario emplear métodos de regresión para estimar el coeficiente de variación entre réplicas. Desde el punto de vista estadístico cabe destacar que, a pesar de la conocida variabilidad de la 2-DE, las réplicas biológicas tienen más valor que las réplicas técnicas (Chich *et al*., 2007), y por ello son más recomendables, especialmente en el caso de muestras complejas como son las humanas. Con el fin de analizar la utilidad de un diseño experimental, Hunt *et al*. (2005) proponen, antes de comenzar un estudio, llevar a cabo estudios piloto con 3 muestras por grupo y 3 geles por muestra, y una serie de herramientas estadísticas para evaluar la potencia del diseño y determinar el número mínimo de muestras y geles necesarios para el experimento definitivo.

Por último, hay que recordar que los procesos de escaneado de la imagen, el pre-procesado (normalización, filtrado, etc.) y la propia comparación (alineamiento de imágenes y *spots*) también conllevan una serie de errores que, en gran medida, dependen del tipo de densitómetro empleado y del programa de análisis 2D, y que afectarán a los resultados de los análisis estadísticos (Meleth *et al*., 2005). Los distintos programas disponibles para 2-DE han sido comparados por varios autores, y así por ejemplo Garrels *et al*. (1989) estimaron la reproducibilidad del sistema QUEST en un 97,6%, y Rosengren *et al*. (2003) determinaron que el PD-Quest derivado de este sistema es más fiable que otros programas en el alineamiento de *spots*, con una menor tasa de falsos positivos.

3. Tratamientos estadísticos para comparar la expresión proteica

3.1. Métodos univariantes: comparación de la expresión de proteínas individuales en mapas 2D

De modo general, en la búsqueda proteómica de biomarcadores se consideran proteínas candidatas a aquellas cuya expresión (bien en términos de cantidad absoluta o de volumen relativo) varía en los mapas problema en relación a los controles. Las pruebas estadísticas univariantes que se emplean en estos análisis de mapas 2D se pueden clasificar como métodos de filtrado, que se aplican sobre los datos pre-procesados (tratados como se ha descrito en el apartado anterior), y que posteriormente se pueden combinar con métodos de reducción de la dimensionalidad o de clasificación, como se comentará más adelante. Las pruebas univariantes más comunes son las de significación, generalmente con el fin de comparar medias/medianas de las variables. La prueba *t* de Student compara diferencias entre las medias del grupo de casos y del grupo control, asumiendo la normalidad de los datos (prueba paramétrica). Aunque este test se ha aplicado en numerosos estudios proteómicos, los datos obtenidos a partir de mapas 2D no suelen cumplir los requisitos paramétricos, es decir, que las variables medidas sean continuas y presenten una distribución normal, con varianzas homogéneas en las dos muestras poblacionales, y que los datos sean independientes. Por ello es más correcto emplear el test *U* de Mann-Whitney, que no asume estos requisitos (prueba no paramétrica). Para ambos tipos de test, se considera que una variable difiere en los grupos significativamente cuando el valor de probabilidad *P* asociado al estadístico es menor que un valor α (generalmente 0,05 para un 95% de confianza en el resultado, o 0,01 para un 99% de confianza). Estas dos pruebas son las más empleadas en 2-DE, ya que normalmente las muestras a comparar no están relacionadas. Sin embargo, en algunos casos se trabaja con muestras

relacionadas (como es el caso del análisis de tejidos sanos y enfermos de un mismo paciente) y para ellos existen pruebas específicas: la prueba *t* para muestras relacionadas (prueba *t* dependiente) en el caso paramétrico, y la prueba *T* de Wilcoxon en el caso no paramétrico. Por otra parte, existen otros tipos de análisis univariantes aplicables a los datos de 2-DE, como por ejemplo el ANOVA[4] cuando lo que se desea analizar es la varianza muestral.

Como se indicó anteriormente, aplicar un valor de α para un conjunto de muchas variables, las cuales son analizadas de manera univariada, conduce a un incremento notable de la probabilidad de obtener falsos positivos. Por ejemplo, analizando 1000 *spots* con una prueba como la *t* de Student, y con un nivel de confianza del 95%, se pueden seleccionar hasta 50 marcadores por azar. Por ello es necesario aplicar *a posteriori* una corrección (Horgan, 2007). Uno de los métodos de corrección más empleados es el de Bonferroni, que fija un valor de α para el conjunto de datos, y compara el estadístico obtenido para cada variable con el valor [α/número de variables], disminuyendo así la tasa de falsos positivos (Smit *et al*., 2007). Sin embargo, es necesario tener en cuenta que la corrección de Bonferroni, considerada como una "prueba familiar" (*familywise*) ya que fija un valor de α idéntico para todas las variables, puede ser contraproducente en el caso de grandes conjuntos de datos, ya que puede llegar a anular la significación estadística de variables que realmente son significativas (Perco *et al*., 2006). Para estos casos existen otras pruebas que consideran distintos valores de α para cada variable, y que pueden ser más apropiadas para el cálculo de la tasa de falsos positivos (Smit *et al*., 2007).

Una consideración necesaria en este tipo de comparaciones es que, desde un punto de vista estadístico, el análisis de un número limitado de muestras de pacientes conlleva el riesgo de no detectar algunas proteínas alteradas (lo que se denomina "error de tipo II"), pero eso no significa que, si se incluyeran más muestras en el estudio, proteínas que se hubieran detectado mostrando una variación significativa en su expresión, no fueran nuevamente detectadas (lo que sería un "error de tipo I") (Byrjalsen *et al.*, 1999).

En nuestro grupo, hemos aplicado diversas técnicas estadísticas univariantes para el análisis de mapas 2D. Por ejemplo, para la comparación de proteínas séricas de individuos sanos y de pacientes con CCR[5] hemos empleado la prueba de Mann-Whitney (Rodríguez-Piñeiro *et al.*, 2004), mientras que para la comparación de tejido sano y tumoral de pacientes se empleó la prueba de Wilcoxon (Álvarez-Chaver *et al.*, 2007).

3.2. Métodos multivariantes: comparación de los niveles de expresión de conjuntos de proteínas en mapas 2D

Una estrategia diferente consiste en buscar no una, sino un patrón o conjunto de proteínas que varíen o que contribuyan a la diferenciación entre el grupo de casos y el grupo de controles. Fundamentalmente, en el análisis de mapas 2D se utilizan dos tipos de métodos multivariantes: los que permiten la reducción de la dimensionalidad del conjunto de datos (apartado 3.2.1), y aquellos que permiten la clasificación de las muestras en diferentes grupos (apartado 3.2.2). Por otra parte, los métodos multivariantes pueden ser empleados con dos finalidades: en primer lugar, permiten seleccionar un conjunto de proteínas cuya variación combinada discrimina los grupos analizados (apartado 4), proteínas que más tarde se valoran en la práctica clínica mediante otras técnicas distintas de la 2-DE; en segundo lugar, las pruebas multivariantes pueden en sí mismas emplearse para extraer toda la información que contiene el subproteoma visualizado en el mapa 2D, de forma que el propio mapa es la herramienta discriminatoria de los grupos (apartado 5).

3.2.1. Métodos de reducción de datos

3.2.1.1. Análisis de componentes principales (ACP)

El ACP[6] es una clase de análisis factorial que intenta identificar una serie de factores o variables subyacentes a los datos obtenidos (variables predictoras), de forma que expliquen la mayor parte de la varianza observada para el mayor número posible de variables medidas. Cada uno de estos nuevos factores, denominados componentes principales (CP[7]), es una combinación lineal de las variables originales, cumpliendo además la condición de ser ortonor-

[4] ANOVA: análisis de varianza.

[5] CCR: cáncer colorrectal.

[6] ACP: análisis de componentes principales.

[7] CP: componentes principales.

males entre sí, de modo que no existe correlación entre las CP, es decir, recogen distinta "parte" de la información contenida en las variables originales. El ACP genera un número de CP igual al número de variables originales, y cada componente expresa en orden decreciente un porcentaje de la varianza total de los datos, de modo que a menudo bastan las tres primeras CP para explicar una proporción suficiente de la varianza, reduciendo así de forma efectiva el número de variables que es necesario manejar. Por lo tanto, el ACP permite reducir un número elevado de variables (por ejemplo, el volumen relativo de todos los *spots* comunes a controles y casos) a unos pocos factores que permiten establecer diferencias entre los dos grupos de población. Esta prueba requiere que los datos sean variables cuantitativas (no categóricas), con una distribución normal bivariada para cada pareja de variables, y que las observaciones sean independientes. El modelo de análisis factorial especifica que las variables vienen determinadas por los factores comunes (los factores estimados por el modelo) y por factores únicos (los cuales no se superponen entre las distintas variables observadas); las estimaciones calculadas se basan en el supuesto de que ningún factor único está correlacionado con los demás, ni con los factores comunes.

El ACP es uno de los análisis multivariantes más populares en 2-DE, y ha sido aplicado en numerosos estudios, tanto en nuestro grupo (Rodríguez-Piñeiro *et al*., 2007) como en otros: Kovarova *et al*. (2000) lo aplicaron para caracterizar cambios en líneas celulares de leucemia en respuesta a tratamiento; Roblick *et al*. (2004) estudiaron las diferencias relativas a la progresión del CCR analizando el proteoma de tejido normal, pólipos, adenomas, tejidos tumorales y metastáticos; recientemente, Verhoeckx *et al*. (2005) combinaron DIGE con ACP para corroborar las diferencias encontradas en mapas analizados mediante técnicas univariantes. Por otra parte, Marengo *et al*. (2003) describieron un nuevo método denominado ACP de tres vías, que incluye información tanto sobre la intensidad como sobre la posición de los *spots*.

3.2.1.2. Mínimos cuadrados parciales (MCP)

La técnica de MCP[8], más conocida como PLS por sus siglas en inglés (*partial least squares*) es similar al ACP, distinguiéndose de éste en que también tiene en cuenta la covarianza de los datos a la hora de buscar las relaciones con los grupos. De esta forma, se pretende encontrar una variable predictora Y en términos de un conjunto de variables X, es decir, las variables no se consideran de forma independiente, sino que se dividen en variables dependientes e independientes, y por ello es que se incluye el estudio de covarianzas.

3.2.2. Métodos de clasificación

3.2.2.1. Análisis discriminante (AD)

El AD[9] permite clasificar las muestras en una serie de grupos previamente establecidos. Así se obtiene un modelo predictivo para pronosticar el grupo de pertenencia de una muestra particular, a partir de las características observadas. Existen diversos tipos de AD: lineal, cuadrático, regularizado y diagonal. En el modelo lineal para dos grupos, usualmente empleado en proteómica, se genera una función discriminante basada en la combinación lineal de las variables predictoras que proporcionan la mejor discriminación entre los dos grupos. La función se genera a partir de una muestra de casos independientes y mutuamente exclusivos, cuyo grupo de pertenencia (único) es conocido, y posteriormente se puede aplicar a nuevos casos no clasificados (siempre y cuando se disponga de las medidas de las variables predictoras). Concretamente, cuando se trata de mapas 2D se suele aplicar el AD lineal a los valores de cantidad o de volumen relativo de los *spots* comunes a todos los mapas de los grupos analizados.

Una variante del AD que se está haciendo cada vez más popular en proteómica se denomina **AD de componentes principales** (ADCP, más conocido como PC-DA por sus siglas en inglés), y consiste en la reducción de datos mediante ACP, y la realización de un AD posterior sobre las puntuaciones así obtenidas. Esta combinación de métodos ha sido empleada en proteómica bajo diversos nombres (revisado en Smit *et al*., 2007).

Otro método de reducción de datos combinable con el AD es el método de MCP, que en ese caso de denomina **AD-MCP** (PLS-DA en inglés). Recientemente, Karp *et al*. (2005) demostraron su utilidad y complementariedad con métodos univariantes aplicados a la proteómica de expresión en bacterias, aunque también existen numerosos ejemplos de su

[8] MCP: mínimos cuadrados parciales.

[9] AD: análisis discriminante.

utilización en estudios de metabolómica (revisado en Smit *et al*. 2007). En relación a la búsqueda de biomarcadores, Verhoeckx *et al*. (2005) emplearon este método en combinación con técnicas univariantes para estudiar efectos antiinflamatorios.

3.2.2.2. Otros métodos de clasificación

El método denominado **máquinas de vectores de soporte (SVM en inglés)** realiza la separación de los grupos construyendo un hiperplano que separa los casos de los controles, asignando el grupo de pertenencia de cada muestra según en qué lado del hiperplano se encuentre. También es combinable con el ACP y el MCP. Esta técnica es bastante popular en proteómica clínica, y especialmente se ha empleado para discriminar varios tipos de cáncer (ver ejemplos en Smit *et al*., 2007).

La **regresión logística** es un método de clasificación basado en aplicar una regresión lineal para ajustar los datos al logaritmo natural de la probabilidad de que una muestra de una clase esté fuera de esa clase. Intuitivamente, este análisis realiza un proceso similar al AD, y por ello también puede combinarse con métodos previos de reducción de datos. La regresión logística ha sido recientemente empleada en proteómica del cáncer (Smit *et al*., 2007).

Las **redes neuronales artificiales** son métodos no lineales, en los que se construye un modelo con una capa neuronal interna formada por los datos, una o más capas neuronales intermedias y ocultas, que consisten en las transformaciones de los datos de origen, y una capa neuronal externa que constituye la respuesta predicha. Este método también está adquiriendo popularidad en proteómica clínica (ver Smit *et al*., 2007).

El método de los **centroides reducidos más cercanos** (*nearest shrunken centroids*) asigna las muestras a aquella clase cuya mediana de clase es más cercana al valor de la muestra. Aunque su fundamento parece *a priori* demasiado simple para producir resultados significativos, esta prueba ha sido empleada con éxito por Kemperman *et al*. (2007) para distinguir pacientes con y sin proteinuria.

Los **árboles de clasificación** se basan en el uso de un algoritmo que sucesivamente va dividiendo los datos de cada "nódulo paterno" en subconjuntos que constituyen los "nódulos hijos", maximizando la homogeneidad en los hijos mediante la elección de la/las proteína/s que más disminuyen la heterogeneidad. En proteómica clínica, se pueden encontrar ejemplos como la detección de pacientes con cáncer de páncreas (Bhattacharyya *et al.*, 2004), o del efecto del tratamiento de pacientes con leucemia (Albitar *et al.*, 2006).

Los **ensambles o conjuntos de clasificadores** (*ensemble classifiers*) combinan varios tipos de reglas de clasificación (clasificadores básicos) para construir un clasificador con mayor resolución. Cada muestra es clasificada primero por las reglas básicas, y la predicción se realiza en forma de "voto por mayoría". Un ejemplo es el bosque de clasificación, donde se construyen múltiples árboles de clasificación cuyos resultados se ensamblan posteriormente (Tong *et al*., 2003).

4. Tratamiento estadístico para la selección de proteínas candidatas como biomarcadores

Como se ha comentado anteriormente, una utilidad importante de los métodos multivariantes en el análisis de mapas 2D es poder seleccionar proteínas con potencial valor como biomarcadores, antes de realizar los estudios de validación biológica. Por ejemplo, es muy frecuente que al comparar el proteoma de individuos afectados (por enfermedad, tratamiento, etc.) con el de individuos control se detecte un número elevado de proteínas cuya expresión individual está alterada. Estas alteraciones, aun teniendo relación con la enfermedad o situación, pueden no tener mucho valor como marcadores para distinguir los dos grupos considerados, o incluso pueden deberse al azar (error tipo I). En estos casos, la aplicación de métodos multivariantes sobre el conjunto de *spots* cuya alteración se ha detectado mediante métodos univariantes permite agrupar aquellas proteínas con mayor influencia en la distinción de grupos, y seleccionar las que teóricamente serían más útiles en la práctica clínica.

Nuestro grupo se ha beneficiado de esta aplicación de las técnicas multivariantes para la selección de potenciales marcadores (Rodríguez-Piñeiro *et al*., 2007). En una comparación de mapas 2D de glicoproteínas séricas de individuos sanos y pacientes con CCR, se aplicó primero la prueba *U* de Mann-Whitney para seleccionar aquellas proteínas alteradas por la patología, encontrando 28 proteínas significativamente diferentes entre los grupos comparados. Sobre ellas se aplicaron métodos de reducción de datos (ACP) y clasificación (AD), re-

duciendo el grupo de proteínas a los mejores candidatos a biomarcadores en base a su capacidad de discriminación de los grupos. Así, finalmente se pudo concluir que 3 de ellas (clusterina, factor I del complemento y β-2-glicoproteína I) constituían un potencial panel de biomarcadores para el CCR.

5. Utilización de los mapas proteicos como herramientas diagnósticas

Las pruebas multivariantes también sirven para determinar si los mapas 2D tienen en sí mismos utilidad diagnóstica, es decir, si el proteoma de un individuo enfermo contiene la información necesaria para su diagnóstico, del mismo modo que los patrones de expresión génica observados mediante microarrays. En este caso, las pruebas multivariantes se aplican a todo el conjunto de proteínas de un mapa 2D para determinar si la información global que contienen permite discriminar los grupos de población estudiados (si el fin es diagnóstico, individuos sanos y pacientes).

Las diferencias entre los mapas de casos y controles se pueden determinar basándose en un rasgo cuantitativo como es la cantidad de proteína (volumen relativo), pero también puede hacerse utilizando un criterio cualitativo como es la posición de las proteínas en los mapas. Dentro de los estudios realizados por nuestro grupo, hemos comprobado la utilidad diagnóstica de los mapas proteicos utilizando ambos criterios, cuantitativo (Rodríguez-Piñeiro *et al*., 2007) y cualitativo (Rodríguez-Piñeiro *et al*., 2005). Resumimos, a continuación, los resultados de estos trabajos.

5.1. Comparando cantidad: ACP y AD

Cuando las pruebas multivariantes se aplican a los valores de todas las proteínas comunes a los mapas de casos y controles, sin tener en cuenta los resultados en los análisis univariantes, se mejora la detección de estados patológicos (en general, de los casos) ya que puede ocurrir que proteínas que no presentan por sí mismas una alteración significativa contribuyan al estado patológico en combinación con otras proteínas (LaBaer, 2005).

En nuestro grupo, se han empleado el ACP y el AD para determinar si el glicoproteoma sérico completo permitía distinguir a personas sanas de pacientes con CCR (Rodríguez-Piñeiro *et al*., 2007). En primer lugar, se aplicó ACP sobre todos los datos de volumen relativo (cantidad) de 363 *spots* comunes a controles y pacientes. Esta prueba permitió reducir el 100% de la varianza entre los grupos, dada inicialmente por las 363 variables, a 9 CP. Sobre ellos se aplicó una prueba de homogeneidad de varianza (mediante el estadístico de Levene), comprobándose que no existían diferencias significativas entre las varianzas del grupo de controles y del grupo de pacientes. Por ello, se aplicó una prueba de análisis de varianza de tipo ANOVA, encontrando que el segundo CP daba cuenta de las diferencias entre los grupos. Esto se pudo corroborar gráficamente.

Por otra parte, la aplicación del AD a los 363 *spots* comunes permitió encontrar una función discriminante (de clasificación o diagnóstica en este caso) para los grupos control y paciente, que explicaba el 100% de la varianza muestral, obteniéndose una puntuación para cada individuo que permite clasificarlo en uno u otro grupo con un nivel de confianza del 99,2%.

Estos resultados demuestran que, al igual que se había observado mediante ACP, es posible determinar la condición de un individuo en función de su patrón 2D de expresión de glicoproteínas séricas. Sin embargo, cabe reconocer que los valores de confianza obtenidos son mucho mejores cuando el análisis se restringe a aquellas proteínas detectadas previamente como diferenciales mediante técnicas univariantes.

5.2. Comparando posición: análisis de deformaciones relativas

Aunque clásicamente los análisis de mapas 2D se centran en la comparación de intensidad, volumen o cantidad de proteína, o en la presencia o ausencia de determinados *spots*, también es posible incluir variables de posición (datos de masa y punto isoeléctrico), como en el ACP de tres vías propuesto por Marengo *et al*. (2003) y comentado anteriormente. Dentro del estudio de la utilidad de los mapas como herramienta diagnóstica, nuestro grupo ha utilizado otro tipo de análisis multivariante que no tiene en cuenta la cantidad de proteína, sino sólo sus coordenadas de posición en los ejes *X* e *Y* de los mapas 2D. Esta técnica podría considerarse similar al ACP, y se denomina análisis de deformaciones relativas (RWA[10] por sus siglas en inglés) (Rodríguez-Piñeiro *et al*., 2005). Este método es de

[10] RWA: análisis de deformaciones relativas.

uso generalizado y ha sido ampliamente comprobado en morfometría aplicada a sistemas biológicos (Bookstein, 1991; Rohlf, 1993).

En el estudio que realizamos comparando mapas de suero de donantes y pacientes con CCR, la aplicación de la morfometría geométrica nos permitió la detección de una variación local significativa en punto isoeléctrico y masa molecular, que claramente separaba el grupo control del grupo de pacientes con CCR (Rodríguez-Piñeiro *et al*., 2005). Esta utilidad podría ser una futura aplicación diagnóstica del método, de forma análoga al ACP y AD, ya que se podrían almacenar los patrones de posición de *spots* de un grupo numeroso de muestras de individuos sanos y de casos, y posteriormente comprobar estadísticamente si una muestra desconocida, con un patrón proteico concreto, corresponde al fenotipo de una persona sana o afectada por una patología o condición fisiológica.

Aún más, el método permite la detección gráfica de aquellas manchas proteicas que contribuyen en mayor medida a los cambios detectados entre los grupos, teóricamente relacionados con la condición o enfermedad estudiada. Existen varios fenómenos que explican las variaciones en posición: modificaciones post-traduccionales tales como fosforilación, glicosilación, acetilación, y otras; cambios en la secuencia aminoacídica de una proteína (polimorfismos de aminoácidos, por ejemplo); la existencia de formas proteicas truncadas; etc. Tanto la magnitud como la dirección de estos cambios se pueden mostrar en forma de vectores en los mapas 2D, y así en nuestro estudio se pudieron identificar aquellas proteínas que mostraban grandes desviaciones (alteraciones), junto con proteínas que mostraban cambios menores o incluso proteínas que no variaban en relación a la condición estudiada.

En comparación con la estrategia de Marengo *et al*. (2003), quienes aplicaron el ACP de tres vías para diferenciar grupos en dos conjuntos de muestras de controles y casos, el RWA separa la variación local de los efectos globales (Bookstein, 1991; Rohlf, 1993; Zelditch *et al*., 2004). Aplicado a mapas 2D, corrige así el sesgo debido a variaciones que afecten a todas las manchas proteicas simultáneamente en cualquiera de las dos dimensiones de separación. Esto tiene una especial importancia en el caso de la 2-DE, ya que los errores técnicos y otras variaciones aleatorias tendrían probablemente un efecto en la localización de todos los *spots* resueltos (al menos en una de las dos dimensiones).

6. Métodos estadísticos empleados para la validación biológica de los marcadores.

6.1. Tratamiento estadístico para la validación de los biomarcadores

Otro campo dentro de la búsqueda de biomarcadores que requiere la utilización de métodos estadísticos es la validación de los marcadores. Un marcador para una enfermedad puede ser útil en el diagnóstico, para determinar el pronóstico del enfermo, o en el seguimiento de la evolución del paciente. Para evaluar la utilidad del marcador en cada uno de estos aspectos se requiere cuantificar la proteína (comúnmente con métodos tipo ELISA[11]) en un número amplio de pacientes y controles, o de pacientes en diferentes situaciones (por ejemplo, con tratamiento y sin tratamiento) y analizar los datos con diferentes pruebas estadísticas. Sin embargo, antes de realizar la comparación debe observarse la distribución de los datos, por ejemplo aplicando la prueba de Kolmogorov-Smirnov para cada muestra poblacional, lo que permite determinar si la población se ajusta o no a una distribución normal. A continuación se aplica la prueba de Levene para determinar si existe homogeneidad de varianza. Como se mencionó anteriormente, estas dos condiciones permiten aplicar pruebas paramétricas, como la *t* de Student de contraste de hipótesis para dos muestras independientes.

En general, puede decirse que existen dos aproximaciones estadísticas para evaluar biomarcadores. La primera modeliza el riesgo de enfermedad (o el resultado de la enfermedad) con técnicas como la regresión logística; así, un marcador es considerado útil si tiene un efecto importante en el riesgo. La segunda evalúa la efectividad de la clasificación dada por el marcador, utilizando parámetros tales como la sensibilidad, especificidad, valores predictivos, y las curvas ROC[12] (Pepe *et al*., 2007).

6.2. Estimación de los parámetros diagnósticos

La validez de la determinación de los niveles de una proteína sérica como prueba diagnóstica para una patología se estudia mediante la elaboración de una curva ROC, siguiendo el procedimiento descrito por Beck y Shultz (1986). En dicha curva se representa, para cada uno de los posibles puntos de corte tomados en función de los niveles del marcador en

[11] ELISA: ensayo inmunoabsorbente ligado a enzima.

[12] ROC: características operativas relativas o para el receptor.

la muestra poblacional de pacientes, el porcentaje de verdaderos positivos (sensibilidad) en el eje de ordenadas, frente al porcentaje de falsos positivos (100-especificidad) en el eje de abscisas (Figura 1). El porcentaje de verdaderos positivos se calcula en función del número de individuos correctamente clasificados como enfermos por la prueba, con respecto al total de individuos de la muestra. El porcentaje de falsos positivos se calcula en función del número de individuos que, estando sanos, son erróneamente clasificados como enfermos por la prueba, con respecto al total de individuos de la muestra.

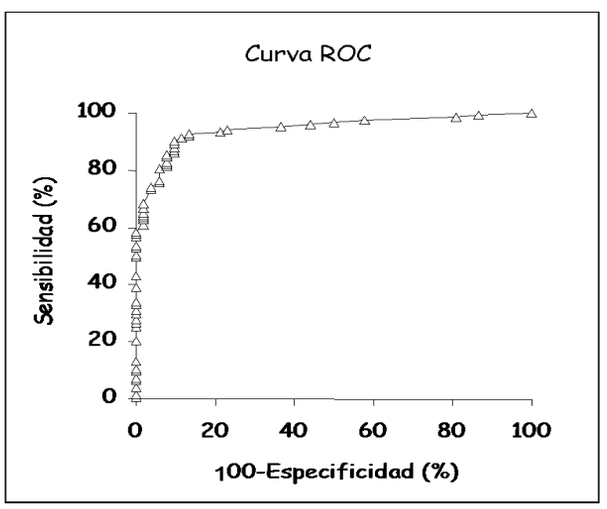


Figura 1. *Ejemplo de una curva ROC obtenida para un marcador tumoral sérico. La gráfica muestra en ordenadas el porcentaje de verdaderos positivos (sensibilidad), y en abscisas el porcentaje de falsos positivos (100-especificidad). El área bajo la curva indica la eficacia diagnóstica del marcador.*

Un test diagnóstico perfecto tendría un 100% de especificidad y sensibilidad. Sin embargo, esto es una utopía. Lo que ocurre en la realidad es que si, utilizando la curva ROC, cambiamos el punto de corte para aumentar la sensibilidad, disminuiremos la especificidad y viceversa. La decisión depende de lo que se quiera conseguir. Por ejemplo, si queremos confirmar la presencia de la enfermedad, el test debería tener una alta especificidad (es decir, pocos falsos positivos); por el contrario, si quisiéramos descartar la enfermedad, entonces el test debería ser muy sensible (es decir, presentar pocos falsos negativos).

Tanto la sensibilidad como la especificidad proporcionan información acerca de la probabilidad de obtener un resultado concreto (positivo o negativo) en función de la verdadera condición del individuo con respecto a la enfermedad. Sin embargo, el clínico normalmente desconoce si el individuo está enfermo o no, y el test debe ayudarle a discernirlo. En este sentido son útiles los parámetros "valor predictivo positivo" y "valor predictivo negativo". El VPP[13] es la probabilidad de padecer la enfermedad si se obtiene un resultado positivo en el test, mientras que el VPN[14] es la probabilidad de que un sujeto con un resultado negativo en la prueba esté realmente sano. No obstante, a pesar de su utilidad clínica hay que tener en cuenta que estos parámetros dependen en gran medida de la prevalencia de la enfermedad (Peng, 2006).

Finalmente, a la hora de validar marcadores debe aprovecharse la circunstancia que ofrece la proteómica al detectar simultáneamente varias proteínas alteradas y por lo tanto candidatas. La determinación de parámetros diagnósticos puede hacerse combinando los valores de varios marcadores y aplicando pruebas multivariantes. A modo de ejemplo, en trabajos realizados en nuestro laboratorio hemos observado que combinando los valores determinados en suero para el receptor del factor de crecimiento epidérmico (sEGFR[15]) y tres de sus ligandos (EGF, TGF-α y anfirregulina) se mejora enormemente la sensibilidad y la especificidad obtenidas para cada marcador de forma independiente (Lemos-González *et al*., 2007). Por otro lado, mediante la aplicación del AD, y con la combinación de los niveles séricos del sEGFR y de sus ligandos EGF y TGF-α, se clasificaron correctamente un 90,2% de individuos sanos y un 100% de pacientes de cáncer no microcítico de pulmón, mientras que la aplicación del modelo de regresión logística, con la misma combinación de sEGFR, EGF y TGF-α, permitió clasificar correctamente un 100% de donantes sanos y un 100% de pacientes de cáncer de cabeza y cuello.

7. Análisis estadístico en espectrometría de masas

Aunque no haremos una revisión detallada de la estadística en espectrometría de masas, queremos recordar que los métodos estadísticos son también relevantes en este campo de la proteómica. En general, las pruebas descritas para 2-DE son también aplicables sobre los espectros de masas, donde cada pico de masas es una variable similar a los *spots* de los mapas 2D.

---

[13] VPP: valor predictivo positivo.

[14] VPN: valor predictivo negativo.

[15] sEGFR: receptor del factor de crecimiento epidérmico.

Un proceso fundamental en el que la estadística está presente es la identificación de proteínas a partir de los espectros obtenidos. Así, motores de búsqueda como ProFound están basados en métodos Bayesianos y utilizan filtros "finos" para establecer las asociaciones entre los datos empíricos y las bases de datos. Otros motores, como SEQUEST, utilizan filtros más "burdos" pero que requieren menos cálculos y son por ello más rápidos aunque sin una pérdida de fiabilidad. PeptideProphet es un programa que emplea pruebas como el AD lineal y la regla de Bayes para convertir los resultados de SEQUEST en una probabilidad (Alterovitz *et al*., 2007).

En ocasiones los datos de espectrometría de masas no se emplean para la identificación de proteínas, sino para la búsqueda de patrones diferenciales, como en los experimentos de SELDI-TOF[16]. En estos casos, se suelen emplear métodos multivariantes de reducción de datos y clasificación como los anteriormente descritos, o análisis de correlaciones y distancias (Cox *et al*., 2005). Un ejemplo de la aplicación de métodos de reducción de datos se puede encontrar en el análisis de la localización subcelular de proteínas de membrana presentado por Sadowski *et al*. (2006). Por otra parte, las técnicas de clasificación han sido recientemente empleadas por Meuwis *et al*. (2007), quienes han utilizado SELDI-TOF para la detección de biomarcadores séricos de enfermedad inflamatoria intestinal.

8. Conclusiones

La combinación de técnicas estadísticas univariantes y multivariantes es de gran utilidad en la búsqueda de biomarcadores mediante técnicas proteómicas. No sólo maximiza la información que puede obtenerse a partir de mapas 2D, y mejora la detección de cambios proteómicos verdaderos y relevantes como marcadores, sino que también ayuda a su validación biológica.

Agradecimientos

Los trabajos mencionados han sido parcialmente financiados por la Xunta de Galicia (PGIDIT05PXIB31002PR). A.M. Rodríguez Piñeiro disfruta de un contrato Ángeles Alvariño (Xunta de Galicia).

[16] SELDI-TOF: ionización/desorción por láser potenciada por superficie, acoplada a tiempo de vuelo.

9. Referencias

Albitar M, Potts S J, Giles F J, O'Brien S *et al*. 2006. Proteomic-based prediction of clinical behavior in adult acute lymphoblastic leukemia. Cancer 106: 1587-1594.

Alterovitz G, Liu J, Afkhami E y Ramoni M F. 2007. Bayesian methods for proteomics. Proteomics 7: 2843-2855.

Álvarez-Chaver P, Rodríguez-Piñeiro A M, Rodríguez-Berrocal F J, Martínez-Zorzano V S *et al*. 2007. Identification of hydrophobic proteins as biomarker candidates for colorectal cancer. International Journal of Biochemistry and Cell Biology 39: 529-540.

Ayude D, Fernández-Rodríguez J, Rodríguez-Berrocal F J, Martínez-Zorzano V S *et al*. 2000. Value of the serum alpha-L-fucosidase activity in the diagnosis of colorectal cancer. Oncology 59: 310-316.

Beck J R y Schultz E K. 1986. The use of relative operating characteristic (ROC) curves in test performance evaluation. Archives of Pathology and Laboratory Medicine 110: 13-20.

Bhattacharyya S, Siegel E R, Petersen G M, Chari S T *et al*. 2004. Diagnosis of pancreatic cancer using serum proteomic profiling. Neoplasia 6: 674-686.

Bookstein F L. 1991. Morphometric tools for landmark data. Cambridge University Press, Nueva York, Estados Unidos.

Byrjalsen I, Mose Larsen P, Fey S J, Nilas L *et al*. 1999. Two-dimensional gel analysis of human endometrial proteins: characterization of proteins with increased expression in hyperplasia and adenocarcinoma. Molecular Human Reproduction 5: 748-756.

Chang J, van Remmen H, Ward W F, Regnier F E *et al*. 2004. Processing of data generated by 2-dimensional gel electrophoresis for statistical analysis: missing data, normalization, and statistics. Journal of Proteome Research 3: 1210-1218.

Chich J-F, David O, Villers F, Schaeffer B *et al*. 2007. Statistics for proteomics: Experimental design and 2-DE differential analysis. Journal of Chromatography B 849: 261-272.

Cox B, Kislinger T y Emili A. 2005. Integrating gene and protein expression data: pattern analysis and profile mining. Methods 35: 303-314.

De Noo M E, Tollenaar R A E M, Deelder A M y Bouwman L H. 2006. Current status and prospects

of clinical proteomics studies on detection of colorectal cancer: Hopes and fears. World Journal of Gastroenterology 12: 6594-6601.

Dudoit S, Shaffer J P y Boldrick J C. 2003. Multiple hypothesis testing in microarray experiments. Statistical Science 18: 71-103.

Garrels J I. 1989. The QUEST system for quantitative analysis of two-dimensional gels. Journal of Biological Chemistry 264: 5269-5282.

Horgan G W. 2007. Sample size and replication in 2D gel electrophoresis studies. Journal of Proteome Research 6: 2884-2887.

Hunt S M N, Thomas M R, Sebastian L T, Pedersen S K *et al*. 2005. Optimal replication and the importance of experimental design for gel-based quantitative proteomics. Journal of Proteome Research 4: 809-819.

Karp N A, Griffin J L y Lilley K S. 2005. Application of partial least squares discriminant analysis to two-dimensional difference gel studies in expression proteomics. Proteomics 5: 81-90.

Kemperman R F, Horvatovich P L, Hoekman B, Reijmers T H *et al*. 2007. Comparative urine analysis by liquid chromatography-mass spectrometry and multivariate statistics: method development, evaluation, and application to proteinuria. Journal of Proteome Research 6: 194-206.

Kovárová H, Hajdúch M, Korínková G, Halada P *et al*. 2000. Proteomics approach in classifying the biochemical basis of the anticancer activity of the new olomoucine-derived synthetic cyclin-dependent kinase inhibitor, bohemine. Electrophoresis 21: 3757-3764.

LaBaer J. 2005. So, you want to look for biomarkers (Introduction to the special biomarkers issue). Journal of Proteome Research 4: 1053-1059.

Lemos-González Y, Rodríguez-Berrocal F J, Cordero O J, Gómez C *et al.* 2007. Alteration of the serum levels of the epidermal growth factor and its ligands in patients with non-small cell lung cancer and head and neck carcinoma. British Journal of Cancer 96: 1569-1578.

Marengo E, Leardi R, Robotti E, Righetti P G *et al*. 2003. Application of three-way principal component analysis to the evaluation of two-dimensional maps in proteomics. Journal of Proteome Research 2: 351-360.

Meleth S, Deshane J y Kim H. 2005. The case for well-conducted experiments to validate statistical protocols for 2D gels: different pre-processing = different lists of significant proteins. BMC Biotechnology 5: 7-21.

Meuwis M A, Fillet M, Geurts P, de Seny D *et al*. 2007. Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling. Biochemical Pharmacology 73: 1422-1433.

Peng X. 2006. Developing and evaluating genomics- or proteomics-based diagnostic tests: statistical perspectives. Methods in Molecular Medicine 129: 27-39.

Pepe M S, Feng Z, Huang Y, Longton G *et al*. 2007. Integrating the predictiveness of a marker with its performance as a classifier. American Journal of Epidemiology (en prensa, DOI 10.1093/aje/kwm305).

Perco P, Rapberger R, Siehs C, Lukas A *et al*. 2006. Transforming omics data into context: Bioinformatics on genomics and proteomics raw data. Electrophoresis 27: 2659-2675.

Randic M. 2006. Quantitative characterizations of proteome: Dependence on the number of proteins considered. Journal of Proteome Research 5: 1575-1579.

Roblick U J, Hirschberg D, Habermann J K, Palmberg C *et al*. 2004. Sequential proteome alterations during genesis and progression of colon cancer. Cellular and Molecular Life Sciences 61: 1246-1255.

Rodríguez-Piñeiro A M, Ayude D, Rodríguez-Berrocal F J y Páez de la Cadena M. 2004. Concanavalin A chromatography coupled to two-dimensional gel electrophoresis improves protein expression studies of the serum proteome. Journal of Chromatography B 803: 337-343.

Rodríguez-Piñeiro A M, Carvajal-Rodríguez A, Rolán-Alvarez E, Rodríguez-Berrocal F J *et al*. 2005. Application of relative warp analysis to the evaluation of two-dimensional gels in proteomics: Studying isoelectric point and relative molecular mass variation. Journal of Proteome Research 4: 1318-1323.

Rodríguez-Piñeiro A M, Rodríguez-Berrocal F J y Páez de la Cadena M. 2007. Improvements in the search for potential biomarkers by proteomics: Application of principal component and discriminant analyses for two-dimensional maps evaluation. Journal of Chromatography B 849: 251-260.

Rohlf, F. J. 1993. Contributions to Morphometrics. Museo Nacional de Ciencias Naturales, Madrid, España.

Rosengren A T, Salmi J M, Aittokallio T, Westerholm J *et al*. 2003. Comparison of PDQuest and Progenesis software packages in the analysis of two-dimensional electrophoresis gels. Proteomics 3: 1936-1946.

Sadowski P, Dunkley T P J, Shadforth I P, Dupree P *et al*. 2006. Quantitative proteomic approach to study subcellular localization of membrane proteins. Nature Protocols 1: 1778-1789.

Smit S, Hoefsloot H C J y Smilde A K. 2007. Statistical data processing in clinical proteomics. Journal of Chromatography B (en prensa, DOI 10.1016/j.jchromb.2007.10.042).

Tong W, Hong H, Fang H, Xie Q *et al*. 2003. Decision forest: combining the predictions of multiple independent decision tree models. Journal of Chemical Information and Computer Sciences 43: 525-531.

Verhoeckx K C, Gaspari M, Bijlsma S, van der Greef J *et al*. 2005. In search of secreted protein biomarkers for the anti-inflammatory effect of beta2-adrenergic receptor agonists: application of DIGE technology in combination with multivariate and univariate data analysis tools. Journal of Proteome Research 4: 2015-2023.

Wasinger V, Cordwell S, Cerpa-Poljak A, Yan J *et al*. 1995. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. Electrophoresis 16: 1090-1094.

Wilkins M R, Pasquali C, Appel R D, Ou K *et al*. 1996. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. BioTechnology 14: 61-65.

Wilkins M R, Appel R D, Van Eyk J E, Chung M C M *et al*. 2006. Guidelines for the next 10 years of proteomics. Proteomics 6: 4-8.

Zelditch M L; Swiderski D L; Sheets H D; Fink W L. 2004. Geometric Morphometrics for Biologists: A Primer. Academic Press, Nueva York, Estados Unidos.

El plasma ICP para la ionización elemental en Espectrometría de Masas: su aplicación en Proteómica Cuantitativa”

Alfredo Sanz-Medel

Departamento de Química Física y Analítica, Universidad de Oviedo, C/Julián Clavería 8, 33006 Oviedo (Asturias). E-mail: asm@uniovi.es

Introducción

A estas alturas ya nadie discute que la Espectrometría de Masas (EM) se ha convertido en una de las técnicas clave, probablemente la más poderosa por su sensibilidad y aplicabilidad a problemas reales, para la caracterización de proteínas. No es de extrañar, pues, que las técnicas de EM ya clásicas (e.g. MALDI-MS y ESI-$(MS)^n$) sean hoy imprescindibles en estudios de proteómica estructural y funcional. Asimismo, la importancia y aplicaciones de la EM analítica crece enormemente en los últimos años también en el área emergente de la biología de sistemas.

En cualquier caso, tales técnicas “convencionales” de EM proporcionan una información química molecular, es decir, de las moléculas, de los aminoácidos, péptidos y proteínas buscados. Mucho menos conocida en el mundo de la Proteómica (mi asistencia a congresos de Espectrometría de Masas en los simposios dedicados a su aplicación en proteínas así lo corrobora de forma irrebatible) es la Espectrometría de Masas Elemental, que utiliza como fuente de ionización un plasma inducido de radiofrecuencias (el ICP-MS). En este caso, la información directa obtenida es atómica, es decir, de los elementos constitutivos de la biomolécula. Por eso, para conocer la naturaleza de la especie química concreta (biomolécula) donde se halla el elemento medido por ICP-MS es preciso acoplar ese detector a un sistema de separación previa y potente de la biomolécula estudiada (p.e. HPLC, CE o electroforesis 2D) en la

cual se halla dicho elemento. Desde luego, "el tiempo de retención" (p.e. en HPLC) da idea de la naturaleza química del compuesto que origina un pico cromatográfico dado y, por tanto, realizar un barrido de la presencia del elemento en estudio en los picos cromatográficos obtenidos (biocompuestos) con esta detección por ICP-MS es rapidísimo. Por ello, la técnica "híbrida" o acoplada HPLC-ICP-MS (ver esquema típico de la instrumentación y del tipo de cromatograma en la Figura 1) se ha desarrollado con gran rapidez en el campo del análisis de trazas metálicas en materiales biológicos. Dicha estrategia ha demostrado recientemente poseer también un extraordinario potencial para Proteómica, como una herramienta complementaria a las otras técnicas de EM moleculares. La capacidad excepcional del ICP-MS de monitorizar y determinar cualquier "heteroelemento" (significando así cualquier elemento en la biomolécula excepto los mayoritarios: C, H, O y N) de forma casi específica, con pocas interferencias y a niveles de ppt ($ng.L^{-1}$) en muestras reales está hoy bien documentada.

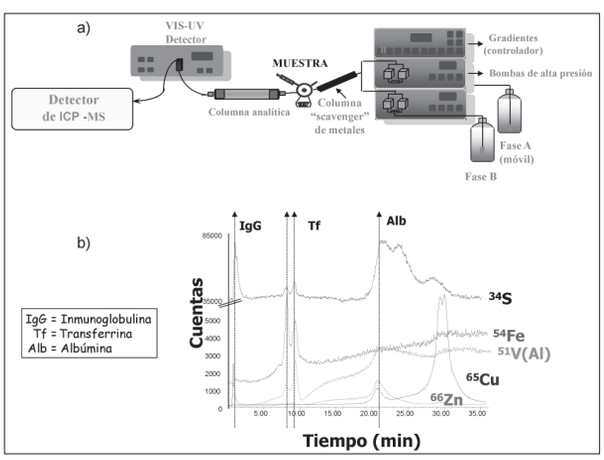


Figura 1. *a) Esquema del dispositivo instrumental típico de acoplamiento de una separación por HPLC seguida de una detección molecular (VIS-UV) y elemental (ICP-MS); b) Un cromatograma típico con detección "elemental" (por ICP-MS), siguiendo la pista simultáneamente a S, Fe, V, Al, Cu y Zn*

Gracias al empleo de un plasma tan robusto como el ICP para ionizar (ver un esquema de un ICP-MS completo en la Figura 2) en la zona más caliente de dicho plasma llegan a alcanzarse los 10.000º K, de modo que es posible conseguir el "screening" rápido y fiable de la presencia de heteroátomos(con la relación masa/carga ,m/z, deseada) de gran importancia biológica (e.g. metales, semimetales, P, S, etc) y que son parte constitutiva de las proteínas de interés.

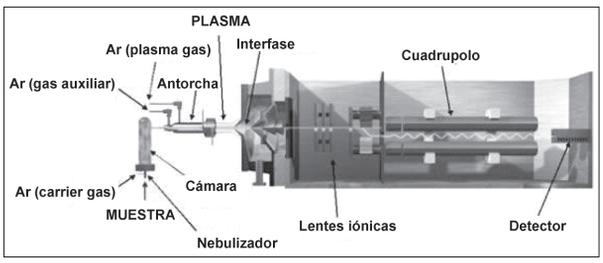


Figura 2. *Un esquema de componentes de un ICP-MS clásico (de cuadrupolo)*

Concepto de Proteómica guiada por un heteroátomo

Tras una separación adecuada por HPLC y detección final por ICP-MS, mezclas muy complejas de proteínas u otras biomoléculas (e.g. metaloproteínas o selenoproteínas en un tejido) se pueden estudiar y determinar de modo rápido y con gran fiabilidad ya que el ICP-MS solo "ve" el heteroátomo (o hetero-isótopo, m/z) deseado en dicha biomolécula. Esta estrategia de simplificar el problema inicial ha dado lugar muy recientemente a un concepto nuevo denominado "heteroatom(isotope)-tagged proteomics" (Sanz-Medel, 2008). Dicha estrategia consiste en la simplificación inicial de las complicadas mezclas de proteínas, resultantes de un material biológico real, gracias a emplear un detector elemental, el ICP-MS, que solo sigue la pista en tales mezclas a un heteroátomo(isótopo) dado. Como se ve en la Figura 3, que resume tal estrategia de trabajo, la detección inicial por ICP-MS (detector mucho más sensible y robusto que los MS más convencionales) permite hacer el primer "screening" de la presencia del heteroátomo deseado en cualquier biocompuesto separado de la mezcla (1ª etapa en la Figura 3). Si existe en el pico (compuesto) dicho elemento, se pasa a la 2ª etapa, la de aislar tal especie y proceder al escrutinio "molecular" de cada compuesto por MALDI- y/o ESI-$(MS)^n$.

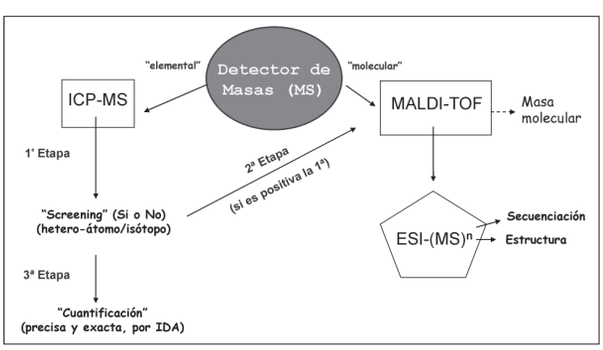


Figura 3. *Una estrategia general de trabajo en "heteroatom-tagged proteomics"*

Finalmente, una vez identificada/caracterizada la(s) biomolécula(s) de interés conteniendo el heteroátomo por tales técnicas, es posible proceder a la 3ª etapa, es decir, la cuantificación segura, fiable y exacta de la masa absoluta de la proteína seleccionada, sea por referencia a una curva de calibrado (con estándares) o bien recurriendo a la técnica de Disolución Isotópica.

La importancia del concepto de "heteroatom (isotope)-tagged proteomics" (es decir, "proteómica inicial guiada por la medida de un heteroátomo concreto") se amplifica por la capacidad del ICP-MS para medir con precisión abundancias y relaciones isotópicas. Esta facultad abre nuevos horizontes en estudios cuantitativos (vía análisis por dilución isotópica) así como en estudios de metabolismo con isótopos "marcados" (que, contrariamente a lo que ocurre con los radioisótopos, son estables, no emiten radiación y pueden investigarse tras su administración en humanos).

Desde luego, el uso del ICP-MS en dichas estrategias no excluye el empleo de los métodos de EM ya clásicos (o "moleculares"). Por el contrario, su papel como "detector elemental e isotópico" es complementar a dichos métodos y técnicas "moleculares", que siguen siendo indispensables para la identificación y caracterización de las biomoléculas que contenían el heteroátomo (moléculas que, tal vez, pudieron ser detectadas por primera vez utilizando el ICP-MS, debido a su gran sensibilidad).

La abrumadora importancia actual de las fuentes MALDI- o ESI- (MS) en Proteómica no se eclipsa por la aceptación de una nueva fuente de iónes como el ICP que viene a complementar a las que ya recibieron la mitad del Premio Nobel de Química en 2002 (J. Fenn, por su introducción del Electrospray [ESI] y K. Tanaka por su trabajo con la "matrix-assisted laser-desorption ionization" [MALDI]).

Algunas Aplicaciones ilustrativas

La importancia actual de esta "Proteómica guiada" por el ICP-MS (a través de un elemento o heteroátomo de importancia biológica) es cada vez más clara y trataré de ilustrarla brevemente, a tres niveles, con experimentos escogidos de nuestro propio laboratorio:

a) "Screening" de metales en proteínas y en sus mezclas complejas (e.g. leche humana).

b) Proteómica cuantitativa en modificaciones post-traduccionales de proteínas (PTMs) "naturales", p.e. en metaloproteómica y fosfoproteómica.

c) Finalmente, el potencial del ICP-MS en PTMs realizadas "artificialmente".

Una de las aplicaciones más interesantes del empleo de la estrategia HPLC-ICP-MS es, como hemos adelantado ya, la posibilidad de llevar a cabo un "screening" rápido y fiable de la presencia (o ausencia) de metales de interés (e.g. toxicológico o nutricional) en proteínas. Ejemplos ilustrativos de dicha aplicación se pueden ver en nuestros trabajos sobre la distribución y especiación de nutrientes esenciales en leche humana y sus proteínas (p.e. Fe,Cu,Zn,Se, I) y su comparación semicuantitativa con leches fórmula típicas utilizadas en los hospitales en alimentación de bebés prematuros y a término (Michalke *et al.,* 2007). Dicho trabajo puso de manifiesto la gran diferencia de composición molecular ("especiación química" de los elementos traza esenciales) entre la leche humana materna y las distintas leches "fórmula" utilizadas en alimentación de recién nacidos prematuros y a término.

El empleo de la estrategia HPLC-ICP-MS en la investigación cuantitativa de PTMs "naturales" en proteínas es otro campo de extraordinaria importancia y actualidad. Utilizando estrategias típicas de "especiación de Fe" en suero humano (ver Figura 4), hemos desarrollado importantes metodologías para la determinación exacta de las "glicoformas" de la transferrina. Tras un análisis por dilución isotópica (IDA) en línea, que utiliza ^{57}Fe enriquecido, es posible realizar determinaciones absolutas de la cantidad de cada una de las isoformas Fe-Tf del suero humano por HPLC-(IDA)ICP-MS. En dichos estudios se ha demostrado (3) que la cuantificación de tales isoformas individuales puede ser de gran utilidad como "biomarcadores" de enfermedades, p.e. en deficiencias genéticas en la glicosilación de proteínas (carbohydrate deficient transferrin, CDT) y en el control analítico de alcoholismo (ver Figura 4 y del Castillo *et al.,* 2005).

Otra de las PTMs de extraordinaria relevancia en Proteómica es la fosforilación, una modificación que regula, entre otros, los importantes procesos de señalización en las células y la eventual aparición de enfermedades relacionadas con disfunciones en dicha señalización (e.g. cáncer). Recientemente hemos demostrado los espectaculares resultados que se pueden conseguir utilizando HPLC capilar acoplada a un ICP-MS (que detecta el ^{31}P) para la determinación de fosfopéptidos, resultantes de la digestión

tríptica de mezclas complejas de proteínas fosforiladas (Pereira Navaza et al., 2007a). El impacto de estas técnicas para llevar a cabo investigaciones cuantitativas de pequeños cambios de la fosforilación (Pereira Navaza et al., 2008) y, sobre todo, en estudios de la cinética de dicha fosforilación de las proteínas en la célula ha sido reconocido muy recientemente en el mundo de la Biología Química (Pereira Navaza et al., 2007b).

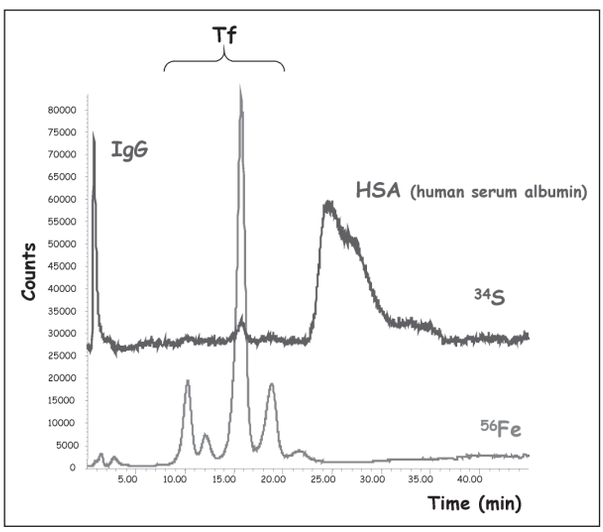


Figura 4. *Un cromatograma típico HPLC-ICP-MS, con detección de Fe y S simultáneas, de un suero humano (cada pico cromatográfico observado corresponde a una glicoforma diferente de transferrina)*

Para terminar esta breve exposición, solo una mirada al futuro: la estrategia comentada de "heteroatom(isotope)-tagged proteomics" solo puede usarse de forma directa en proteínas naturales que contienen los heteroátomos (e.g metales, Se, P, S, etc.). Sin embargo, hoy existen amplios y variados conocimientos sobre "bioconjugación" que permiten introducir artificialmente una "etiqueta" o marcador elemental (i.e. un heteroátomo) en las proteínas deseadas. Esta introducción, mediante un reactivo adecuado, haría que las proteínas sin heteroátomos pudieran también hacerse "visibles" y "medibles" para el ICP-MS. Hoy es ésta una posibilidad de futuro que podría ampliar el campo de utilización de estas estrategias de Proteómica "elemental", aquí delineadas, a prácticamente todas las proteínas conocidas.

En resumen, los ejemplos ilustrativos expuestos han sido seleccionados de los trabajos de mi grupo en Oviedo, pero, desde luego, existen ya otros autores y trabajos muy importantes en esta línea (Lobinski *et al.,* 2006) y todos ellos demuestran que el concepto y estrategias denominadas"heteroatom(isotope)-tagged proteomics" ofrecen un interés y un potencial extraordinarios tanto desde el punto de vista analítico como bioquímico. En todo caso, si como parece cercano, en el futuro es posible extender su campo de acción a todas las proteínas (vía "etiquetado" o "marcado" con heteroátomos) este potencial se muestra ya tan atractivo como para animar a los "viejos" espectroscopistas atómicos a adentrarnos en un bosque nuevo tan atractivo, frondoso y lejano a nosotros como el de la Proteómica, particularmente el de la Proteómica cuantitativa (del Castillo *et al.,* 2005; Pereira Navaza *et al.,* 2007a). A su vez, solo si nuestros colegas bioquímicos, biólogos, médicos clínicos, etc. son conscientes y conocedores de estas técnicas y estrategias, y sobre todo de sus aplicaciones en Biociencia, dicho potencial de futuro podrá finalmente desarrollarse en toda su plenitud. En todo caso, multidisciplinaridad y cooperación interdisciplinar es ahora la clave para hacer posible y eficaz ese esperable y prometedor desarrollo.

Nota final

El objetivo de este resumen es solo dar a conocer en el mundo de la Proteómica de habla hispana el enorme potencial de estas metodologías de MS Elemental, derivadas del empleo de un moderno ICP-MS. El lector interesado en conocer más a fondo las aplicaciones y el sentido exacto de las Figuras (particularmente las Figuras 1, 4 y 5, que intentan abrir una ventana a las aplicaciones mas ilustrativas/ interesantes) puede consultar los detalles en las referencias bibliográficas escogidas.

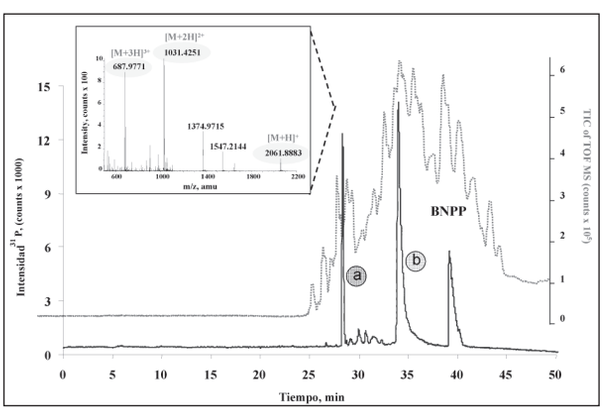


Figura 5. *Un cromatograma típico HPLC-ICP-MS, con detección de 31P (traza azul) y su comparación con el cromatograma más convencional HPLC-Electrospray (traza roja). La expansión del péptido "a" no es otra cosa que el espectro MALDI-TOF de dicha fracción "a" aislada (para confirmar su identidad)*

Referencias

del Castillo Busto ME, Montes-Bayón M, Blanco-González E, Meija J, Sanz-Medel A.2005. Analytical Chemistry 77, 5615-5621.

Lobinski R, Schaumlöffel D, Szpunar J. 2006. Mass Spectrometry In Bioinorganic Analytical Chemistry. Mass Spectrometry Reviews 25, 255-289.

Michalke B, Fernández-Sánchez ML, Sanz-Medel A. 2007. Elemental speciation in human milk and substitute food for newborns. The Determination of Chemical Elements in Food (Applications for Atomic Mass Spectrometry). Wiley & sons 3, 535-567.

Pereira Navaza A, Ruiz Encinar J, Sanz-Medel A. 2007a. Angewandte Chemie International Edition 46, 569-571.

Pereira Navaza A, Ruiz Encinar J, Carrascal M, Abian J, Sanz-Medel A. 2008. Analytical Chemistry 80, 1777-1787.

Pereira Navaza A, Ruiz Encinar J, Sanz-Medel A. 2007b. Chemical Biology (Instant Insight) 2, B79.

Sanz-Medel A. 2008. Heteroatom(isotope)-tagged proteomics via ICP-MS: screening and quantification of proteins and their post-translational modifications., Analytical & Bioanalytical Chemistry (Trends) 391, 885-894.

Estrategias proteómicas para el análisis de grupos carbonilo como marcadores de daño oxidativo en proteínas

Verónica Irazusta, Armando Moreno-Cermeño, Joaquim Ros, Jordi Tamarit

Grup de Bioquímica de l'Estrés Oxidatiu, Departament de Ciències Mèdiques Bàsiques, Facultat de Medicina, Universitat de Lleida, Lleida

Resumen

El estrés oxidativo se produce por un desequilibrio entre la formación y la destrucción de las especies reactivas del oxígeno. Una forma indirecta de determinar la existencia de estrés oxidativo es la medida de los productos derivados de la reacción de las especies reactivas del oxígeno con biomoléculas. En este sentido, la formación de grupos carbonilo en las proteínas ha sido uno de los marcadores de estrés oxidativo más estudiados debido a su estabilidad y fácil detección. Diversas herramientas proteómicas ofrecen un gran potencial para descubrir nuevas proteínas susceptibles al estrés oxidativo, cambios cuantitativos en el perfil de estas modificaciones bajo distintas condiciones biológicas, y para caracterizar el sitio exacto y tipo de modificación sufrida por una determinada proteína. En el presente trabajo se revisan las diferentes aproximaciones utilizadas para la detección de proteínas carboniladas y las herramientas proteómicas que permiten su identificación.

Palabras clave:

Estrés oxidativo, oxidación de proteínas, grupos carbonilo, DNPH, proteómica redox.

Introducción

El estrés oxidativo se describe como una situación en la cual las defensas antioxidantes son insuficientes para inactivar completamente a las especies reactivas del oxígeno (ROS) [1]. Este desequilibrio entre la producción y la destrucción de ROS puede afectar a gran parte de los componentes celulares, incluyendo lípidos, proteínas, carbohidratos y ácidos nucleicos [2-4]. Muchos de los cambios que ocurren durante el envejecimiento y durante la progresión de ciertas enfermedades son una consecuencia del estrés oxidativo [5-7]. Las proteínas se modifican oxidativamente al ser uno de los blancos principales de las ROS. Ello conlleva modificaciones estructurales

que, en general, suponen la pérdida total o parcial de su función [4, 8, 9]. En este sentido, la oxidación de proteínas juega un papel muy importante en el desarrollo del proceso de envejecimiento y de una lista muy extensa de enfermedades entre las que podríamos destacar la ateroesclerosis, el cáncer, y diversas enfermedades neurodegenerativas como el Alzheimer o el Parkinson [10]. Muchos residuos de aminoácidos de las proteínas son susceptibles a la oxidación por varias especies reactivas del oxígeno. El producto de la oxidación de proteínas más ampliamente estudiado en los últimos años ha sido la formación de grupos carbonilo debido a su estabilidad y fácil detección. Por tanto, su cuantificación se ha convertido en el método más aceptado para medir el daño oxidativo en proteínas en situaciones donde el estrés oxidativo juega un papel relevante.

Formación de grupos carbonilo

El estrés oxidativo puede causar modificaciones reversibles o irreversibles en las proteínas. Las modificaciones reversibles, generalmente en los residuos de cisteína, pueden ser reparadas mediante enzimas específicos como las glutaredoxinas o tioredoxinas [11, 12]. De la misma manera, la metionina se oxida fácilmente a metionina sulfóxido pudiendo ser reducida por la enzima metionina sulfóxido reductasa [13]. Habitualmente este tipo de modificaciones tienen la función de proteger ante el daño irreversible o de modular la función de las proteínas (regulación redox). Las modificaciones irreversibles no pueden ser reparadas, de forma que las proteínas inactivas se acumulan o son degradadas. Un ejemplo de modificaciones irreversibles es el caso de la carbonilación de las proteínas [14, 15].

Los grupos carbonilo pueden ser introducidos en las proteínas en diferentes sitios y por medio de diferentes mecanismos. Por un lado, los carbonilos son producidos en las cadenas laterales de los aminoácidos, especialmente en los residuos de prolina, arginina, lisina y treonina los cuales son oxidados a derivados aldehidos o cetonas [2, 16]. Este tipo de modificaciones se producen habitualmente por el ataque del radical hidroxilo (•OH), que puede ser originado por radiación ionizante o por la reacción de cationes metálicos con el peróxido de hidrógeno (H_2O_2, reacción de Fenton). En este sentido merece la pena comentar que la mayoría de los carbonilos formados en las proteínas mediante la oxidación catalizada por cationes metálicos son el semialdehido glutámico derivado de la arginina y el semialdehido aminoadípico derivado de la lisina (Fig. 1A) [17]. Por otro lado, los grupos carbonilo pueden también ser generados por proteólisis de las proteínas tras el ataque de radicales •OH, vía alfa-amidación o tras la oxidación de las cadenas laterales de residuos glutamilo, dando lugar a la formación de un péptido en el cual el aminoácido N-terminal está bloqueado por un derivado alfa-cetoacilo [16]. Existen también otros mecanismos que pueden dar lugar a la carbonilación proteica como la reacción con productos generados durante la peroxidación lipídica como 4-hidroxinonenal (4-HNE[1]) (Fig. 1B), 2-propenal (acroleina) y malondialdehido [18]. Este tipo de reacción denominada "Adición de Michael", conlleva la adición de grupos aldehídos reactivos a las cadenas laterales de los residuos de cisteína, histidina o lisina. Finalmente, grupos carbonilo reactivos pueden ser también generados por reacciones del grupo amino de los residuos de lisina con azúcares reducidos o por la oxidación de sus productos de reacción (glicación/glicoxidación), dando lugar a, por ejemplo, carboximetil lisina [2, 3].

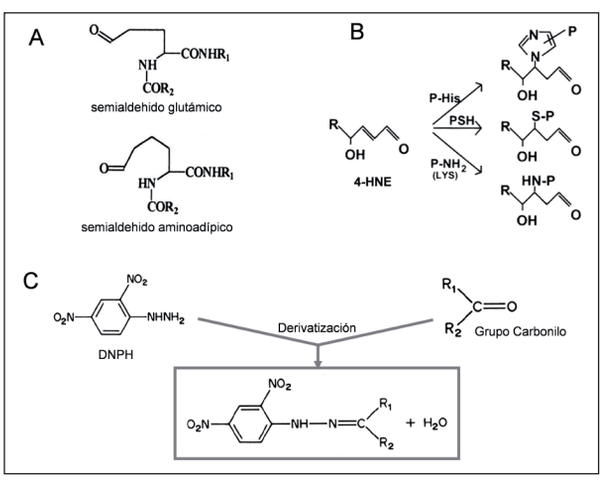


Figura 1. ***Formación de grupos carbonilo y derivatización mediante DNPH. A:*** *productos generados por el ataque del radical •OH a las cadenas laterales de arginina (semialdehido glutámico) y lisina (semialdehido aminoadípico) [17];* ***B:*** *la reacción de 4-HNE con las cadenas laterales de los residuos de cisteína, histidina y lisina conlleva la adición de grupos aldehídos reactivos a las proteínas [16];* ***C:*** *formación de 2,4-dinitrofenilhidrazona tras la reacción del DNPH con grupos carbonilo.*

Medidas de daño oxidativo a proteínas: carbonilos

Debido a que las ROS son generalmente muy reactivas y tienen una vida media muy corta, una

[1] 4-HNE: 4-hidroxinonenal

forma indirecta de determinar la existencia de estrés oxidativo es la medida de los productos formados por la reacción de las ROS con biomoléculas. Dado que en el caso de las proteínas uno de los productos generado son los grupos carbonilo su cuantificación se ha convertido en un parámetro ampliamente aceptado de daño oxidativo en condiciones de estrés. Los grupos carbonilo han sido cuantificados de diversas formas. El método más comúnmente utilizado es el basado en su reacción con 2,4-dinitrofenilhidrazina (DNPH[2]) para formar 2,4-dinitrofenilhidrazona. Este método se basa en que el DNPH es capaz de unirse covalentemente a los grupos carbonilo -formando una base de Schiff- para dar lugar a hidrazonas estables (Fig.1C). El grupo dinitrofenilo (DNP[3]) puede ser detectado y cuantificado espectrofotométricamente debido a que exhibe un espectro de absorción característico con una absorbancia máxima a 365-375 nm [8, 19]. Asimismo, debido a que existen excelentes anticuerpos comerciales contra el grupo dinitrofenilo [20, 21], la detección de grupos carbonilo puede realizarse mediante western blot, dot blot, inmunohistoquímica o ELISA. Se han desarrollado diversos métodos para la derivatización y la preparación de las muestras. La detección mediante western blot ha sido comúnmente denominada Oxi-blot. [10, 22-24].

Sin embargo otros métodos están emergiendo como alternativas al DNPH debido a que algunos autores cuestionan la estabilidad de la 2,4-dinitrofenilhidrazona en condiciones alcalinas o durante el almacenamiento prolongado. Entre los métodos alternativos más utilizados se encuentran los que utilizan la química de la hidrazida. Las hidrazidas reaccionan específicamente con los carbonilos de proteínas (principalmente en su forma aldehido) mediante la formación de una base de Schiff, la cual puede ser estabilizada por reducción con cianoborohidruro [25-27]. Las ventajas de esta aproximación no son sólo una mayor estabilidad del producto final de reacción sino también la presencia de un amplio catálogo de grupos funcionales acoplados a la hidrazida que permiten el diseño de diversas estrategias para la captura o detección de las proteínas carboniladas [28]. Entre estos grupos funcionales encontramos biotina, digoxigenina y diversos compuestos fluorescentes. De esta forma los grupos carbonilo pueden ser derivatizados con digoxigenina-hidrazida, e identificados utilizando anticuerpos contra digoxigenina [29]. Una técnica similar puede ser utilizada derivatizando los grupos carbonilo con biotina–hidrazida, identificando las proteínas derivatizadas utilizando avidina marcada con fluoresceína [30]. Otro compuesto utilizado para la detección fluorescente de grupos carbonilo es la fluoresceína-5-tiosemicarbazida. El grupo tiosemicarbazida presenta una reactividad frente a los grupos carbonilo similar a la hidrazida [31].

Finalmente, otra aproximación para medir el contenido de carbonilos totales en una muestra biológica es la determinación de productos específicos como el semialdehido glutámico y el semialdehido aminoadípico mediante cromatografía de gases y espectrometría de masas con dilución isotópica. Esta aproximación fue desarrollada por Requena y colaboradores y posteriormente utilizada para caracterizar la naturaleza de los distintos productos carbonilo presentes en muestras de enfermos de Alzheimer [17, 32].

Identificación de proteínas oxidadas mediante herramientas proteómicas

Podemos definir la "proteómica redox" como el conjunto de aproximaciones proteómicas que permiten la identificación de proteínas modificadas oxidativamente. La proteómica redox ofrece un gran potencial para descubrir nuevas proteínas susceptibles al estrés oxidativo, correlacionar cambios cuantitativos en el perfil de estas modificaciones con el estado de una enfermedad o con el envejecimiento, y para caracterizar el sitio exacto y tipo de modificación vía análisis MS/MS [3, 33]. Este tipo de aproximaciones pueden dar lugar a la identificación de grupos de proteínas selectivamente oxidadas en determinadas patologías y por tanto contribuir a aclarar los mecanismos moleculares involucrados en ellas. En esta revisión nos centraremos en la identificación de grupos carbonilo, uno de los productos de la oxidación de proteínas. Hasta la fecha diversos estudios proteómicos han demostrado que la carbonilación es altamente selectiva, debido a que en diferentes modelos de estudio se han identificado grupos de proteínas particularmente sensibles al estrés oxidativo. A continuación se revisan algunas de las estrategias utilizadas para la identificación de proteínas carboniladas. Un resumen de los estudios referidos puede encontrase en la Tabla 1, mientras que en la Figura 2 se presenta una visión esquemática de las principales estrategias utilizadas.

[2] DNPH: 2,4-dinitrofenilhidrazina

[3] DNP: 2,4-dinitrofenilo

Tabla 1. Estudios en que se identifican proteínas carboniladas

Referencia	*Organismo*	*Modelo*	*Compuesto*	*Estrategia*	*Proteínas identificadas*
			Sistemas basados en electroforesis bidimensional		
[40]	Levadura	Deficiencia en frataxina	DNPH	2DE Oxi-blot	14
[39]	Cerebro humano	Enfermedad de Huntington	DNPH	2DE Oxi-blot	4
[38]	Levadura	Envejecimiento	DNPH	2DE Oxi-blot	16
[37]	Cerebro humano	Alzheimer	DNPH	2DE Oxi-blot	4
[31]	Ratón	Envejecimiento	Fluorescein-5-Tiosemicarbazida	2DE/detección fluorescente	12
[30]	Levadura	H_2O_2	Biotina hidrazida	2DE/avidina-fluoresceína	20
[25]	Rata	Diabetes	Biotina hidrazida	2DE avidin blot	7
[48]	*Arabidopsis thaliana*	Ciclo de vida	DNPH	2DE Oxi-blot	5
			Sistemas libres de gel		
[28]	Levadura	H_2O_2	Biotina hidrazida	Afinidad/LC/proteólisis/MS-MS	99
[26]	Mitocondria de rata	Basal	Biotina hidrazida + iTRAQ	Afinidad/proteólisis/iTRAQ/LC-MS-MS	200
[27]	Cerebro de ratón	Envejecimiento	Biotina hidrazida	Afinidad/LC/proteólisis/LC-MS-MS	100
[43]	Células tumorales	Tratamiento fotodinámico	Biotina hidrazida	Afinidad/1DE/proteólisis/ LC-MS-MS	81
[47]	Levadura	H_2O_2	Reactivo de Girard	LC-MS/MALDI-MS-MS	36
[44]	Tejido adiposo	Obesidad	Biotina hidrazida	Afinidad/proteólisis/ LC-MS-MS	37

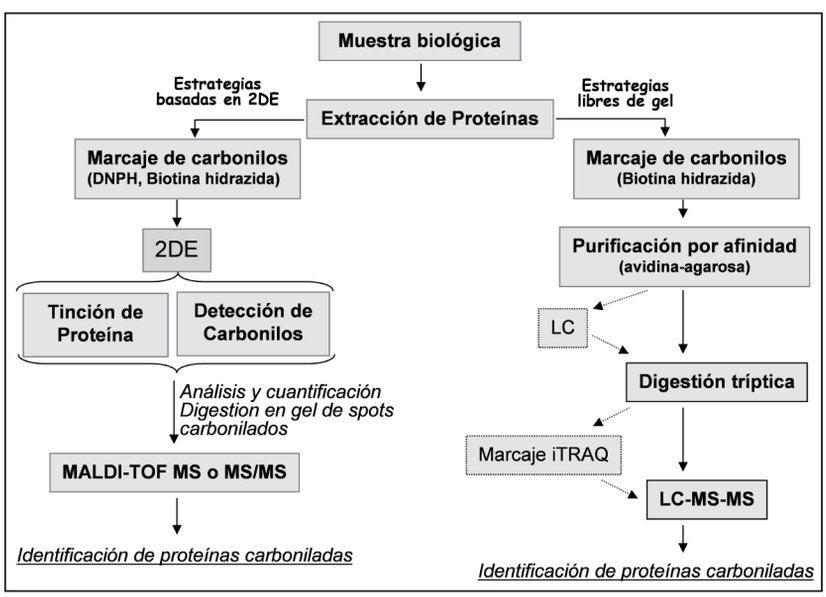


Figura 2. *Visión esquemática de diversas estrategias analíticas utilizadas para la identificación y cuantificación de proteínas carboniladas.*

Sistemas basados en electroforesis bidimensional

La electroforesis bidimensional (2DE) ha sido ampliamente utilizada para la identificación de proteínas carboniladas a partir de muestras complejas. La detección de esta modificación oxidativa se puede realizar mediante western blot con anticuerpos que reconocen directamente la modificación carbonilo en la proteína (por ejemplo: anti –4-HNE)[34], o bien con anticuerpos anti-DNP si la muestra ha sido previamente derivatizada con este compuesto [35-37]. Un ejemplo de este tipo de aproximación se muestra en la Figura 3. En otros casos, los carbonilos reactivos son marcados covalentemente con biotina-hidrazida y detectados en el gel mediante avidina-fluoresceína [30] o avidina-peroxidasa [25]. El uso de geles bidimensionales ha demostrado ser muy útil para la identificación de proteínas oxidadas, especialmente cuando se comparan dos condiciones fisiológicas [38-40]. Presenta la ventaja de poder contar por un lado con las imágenes obtenidas tras la detección del grupo funcional unido a los carbonilos y por otro con aquellas obtenidas tras la tinción de los geles mediante colorantes de proteínas (tinción con plata, Sypro Ruby, Coomassie Brilliant blue, etc.). Esto permite la normalización de la señal de carbonilo respecto a la cantidad de proteína. Hay que tener en cuenta que como en cualquier análisis proteómico, el estudio de las proteínas carboniladas está sesgado a favor de aquellas proteínas más abundantes. Por tanto la determinación del nivel de carbonilación relativo de una determinada proteína es más informativo que el contenido total de proteína modificada. De lo contrario, aquellas proteínas abundantes pero con bajo contenido de carbonilos específicos pueden aparecer como predominantes en oxi-blots, mientras que

proteínas poco abundantes pero con niveles altos de carbonilación específica podrían pasar desapercibidas [10]. Asimismo un cambio en el nivel de expresión de una determinada proteína entre dos condiciones de estudio podría identificarse erróneamente como un cambio en el grado de daño de dicha proteína. En este sentido conviene destacar que diversos estudios basados en geles bidimensionales han permitido identificar proteínas poco abundantes como altamente sensibles al estrés oxidativo. Un ejemplo serian las chaperonas mitocondriales [40-42].

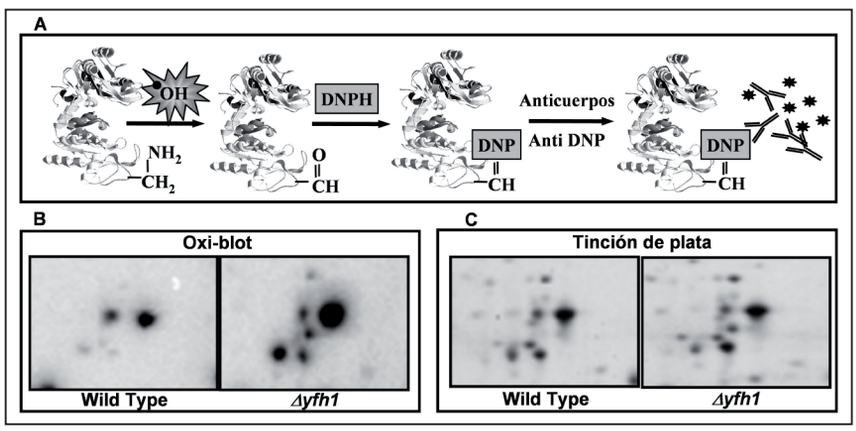


Figura 3. ***Formación de grupos carbonilo y detección mediante anticuerpos anti-DNP (Oxi-Blot). A:*** *El ataque de un radical •OH a la cadena lateral de un aminoácido genera un grupo carbonilo. El DNPH reacciona específicamente con el grupo carbonilo formándose una hidrazona estable que puede ser detectada inmunológicamente por medio de anticuerpos específicos anti-DNP.* ***B:*** *Detección de proteínas carboniladas en extractos celulares de dos cepas de levadura mediante 2DE y Oxi-Blot.* ***C:*** *La tinción de proteínas mediante plata de la misma región del gel bidimensional permite la normalización de los niveles de carbonilación. Puede observarse que diversas proteínas presentan un mayor grado de carbonilación en la cepa Δyfh1 (para detalle ver ref. 40).*

A pesar de la gran información obtenida mediante el análisis proteómico basado en geles, existen limitaciones de este tipo de aproximación para el caso de análisis a gran escala así como para la identificación de proteínas de membrana o proteínas poco abundantes. Por otro lado, el uso de geles bidimensionales presenta como principal problema la validación de los resultados obtenidos. La comparación de las imágenes obtenidas mediante tinción específica (anti-DNP, por ejemplo) y proteína total no siempre es fácil, y el criterio a seguir en cada trabajo queda en manos del investigador. La doble tinción fluorescente permite, en teoría, un mejor emparejamiento entre las distintas imágenes obtenidas. La confirmación mediante MS de la presencia de péptidos modificados podría ser una forma de confirmar el daño a una proteína determinada. Sin embargo, esta estrategia no se acostumbra a utilizar ya que para resultar concluyente requeriría del desarrollo de aproximaciones específicamente encaminadas a ese fin. Una forma indirecta de validación de los resultados es la medida de las actividades enzimáticas de aquellas proteínas identificadas como oxidadas. Esta aproximación ha sido utilizada con éxito en diversos estudios [9, 38, 40], aunque su principal limitación se encuentra en el hecho de que solo puede aplicarse a enzimas para los cuales exista un método eficaz para medir su actividad enzimática.

Sistemas "libres de gel"

Diversos autores han explorado métodos libres de gel, basados en el enriquecimiento mediante afinidad de proteínas que presentan una modificación tipo carbonilo y su posterior detección mediante espectrometría de masas. Estas aproximaciones permitirían en teoría obviar algunas de las limitaciones de los sistemas basados en 2DE descritas en el apartado anterior. Sin embargo la desventaja de estas aproximaciones es que no permiten un análisis cuantitativo del grado de carbonilación de las distintas proteínas identificadas en diferentes muestras, lo que limita su utilidad como herramienta para el estudio de cambios en el patrón de oxidación de proteínas en distintas condiciones biológicas. El sistema más utilizado para el enriquecimiento de las proteínas carboniladas es la cromatografía de afinidad de avidina, para carbonilos previamente derivatizados con biotina-hidrazida. Posteriormente las proteínas enriquecidas son digeridas y los péptidos obtenidos son identificados mediante LC-MS/MS [26-28, 43, 44]. Utilizando esta tecnología Soreghan y colaboradores identificaron 100 proteínas carboniladas en un único experimento a partir de homogenados de cerebro de ratones de diferentes edades. Entre ellas destacan receptores poco abundantes y proteínas mitocondriales implicadas en el metabolismo energético y en la respuesta a estrés, así como receptores con residuos de tirosina fosforilables. Además, fueron capaces de observar carbonilos en proteínas citoplasmáticas abundantes previamente identificadas mediante electroforesis bidimensional [27]. Una extensión de esta tecnología utilizando iTRAQ ha permitido distinguir las proteínas carboniladas de aquellas con afinidad a la avidina. De esta

forma se han logrado identificar 200 proteínas carboniladas en extractos mitocondriales de músculo de rata, lo cual muestra la gran susceptibilidad de este tipo de proteínas ante el estrés oxidativo [26]. Otra alternativa consiste en el fraccionamiento mediante LC de las proteínas derivatizadas con biotina-hidrazida tras su purificación mediante cromatografía de afinidad de avidina y previamente a su digestión. Utilizando esta aproximación "top-down" Mirzaei y Regnier identificaron 99 proteínas carboniladas en levaduras tratadas con H_2O_2 [28].

Identificación del sitio de carbonilación

En los últimos años se han intentado desarrollar metodologías para identificar el aminoácido modificado oxidativamente. Esta información es importante para entender en profundidad los mecanismos oxidativos por los cuales se están produciendo los grupos carbonilo, así como para conocer los efectos funcionales que una modificación a un determinado aminoácido puede causar en la proteína. Existen diferentes estrategias analíticas para la detección del aminoácido modificado. La mayoría de estas metodologías están basadas en la observación de la variación de la masa molecular específica de un aminoácido modificado covalentemente, que es fácilmente detectable mediante mediciones de espectrometría de masas. En algunos de los estudios en que se ha usado la biotina-hidrazida para el marcaje y purificación de proteínas carboniladas, la identificación de péptidos biotinilados ha permitido confirmar la modificación oxidativa y asimismo conocer su localización [28, 43]. Por otro lado, la reacción de 4-HNE con los residuos de lisina, histidina y cisteína puede ser monitorizada en proteínas intactas, detectando los productos correspondientes con un incremento de 156 Da. Asimismo, puede ser identificado el péptido específico que lleva esta diferencia de masa [3, 34, 45]. Existe otra aproximación que permite la identificación mediante MALDI-TOF de péptidos modificados oxidativamente mediante la adición de DNP por reacción con 4-HNE. Los péptidos carbonilados pueden ser identificados por la diferencia de masa de 180 Da (correspondiente al DNP) existente entre muestras derivatizadas y sin derivatizar. Dada la baja estabilidad de la hidrazona en las condiciones habitualmente utilizadas en MALDI-TOF, esta estrategia requiere del uso de DNPH como matriz de ionización [46]. Finalmente, Mirzaei y Regnier utilizaron dos formas isotópicas del reactivo de Girard (que contiene un grupo hidrazida) para caracterizar el lugar de carbonilación de proteínas provenientes de extractos de levaduras tratadas con H_2O_2 [47].

Queda claro que las técnicas proteómicas están ayudando a profundizar en la detección, identificación y los mecanismos subyacentes en las modificaciones oxidativas de las proteínas. Con ello se contribuye a una mejor comprensión de la trascendencia de estas modificaciones a nivel celular. El impacto de ellas sobre la fisiología celular será, sin duda, el próximo reto, es decir, cómo se afecta el metaboloma de la célula como consecuencia de la alteración en la función de determinadas proteínas.

Referencias

[1] Cadenas, E.; Wefers, H.; Muller, A.; Brigelius, R.; Sies, H. Active oxygen metabolites and their action in the hepatocyte. Studies on chemiluminescence responses and alkane production. Agents Actions Suppl 11:203-216; 1982.

[2] Stadtman, E. R.; Berlett, B. S. Reactive oxygen-mediated protein oxidation in aging and disease. Drug Metab Rev 30:225-243; 1998.

[3] Dalle-Donne, I.; Scaloni, A.; Giustarini, D.; Cavarra, E.; Tell, G.; Lungarella, G.; Colombo, R.; Rossi, R.; Milzani, A. Proteins as biomarkers of oxidative/nitrosative stress in diseases: the contribution of redox proteomics. Mass Spectrom Rev 24:55-99; 2005.

[4] Dalle-Donne, I.; Aldini, G.; Carini, M.; Colombo, R.; Rossi, R.; Milzani, A. Protein carbonylation, cellular dysfunction, and disease progression. J Cell Mol Med 10:389-406; 2006.

[5] Harman, D. Aging: a theory based on free radical and radiation chemistry. J Gerontol 11:298-300; 1956.

[6] Balaban, R. S.; Nemoto, S.; Finkel, T. Mitochondria, oxidants, and aging. Cell 120:483-495; 2005.

[7] Giorgio, M.; Trinei, M.; Migliaccio, E.; Pelicci, P. G. Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? Nat Rev Mol Cell Biol 8:722-728; 2007.

[8] Levine, R. L.; Williams, J. A.; Stadtman, E. R.; Shacter, E. Carbonyl assays for determination of oxidatively modified proteins. Methods Enzymol 233:346-357; 1994.

[9] Tamarit, J.; Cabiscol, E.; Ros, J. Identification of the major oxidatively damaged proteins in Es-

cherichia coli cells exposed to oxidative stress. J Biol Chem 273:3027-3032; 1998.

[10] Levine, R. L., Stadtman, E. R. *Carbonylated proteins and their implication in physiology and pathology*. In: Dalle-Donne I., Scaloni, A., Butterfield, D. A., eds. Redox Proteomics: From protein modifications to cellular dysfunction and diseases. United States of America: Wiley-Interscience; 2006: 123-168.

[11] Lillig, C. H.; Berndt, C.; Holmgren, A. Glutaredoxin systems. Biochim Biophys Acta 1780:1304-1317; 2008.

[12] Oktyabrsky, O. N.; Smirnova, G. V. Redox regulation of cellular functions. Biochemistry (Mosc) 72:132-145; 2007.

[13] Vogt, W. Oxidation of methionyl residues in proteins: tools, targets, and reversal. Free Radic Biol Med 18:93-105; 1995.

[14] Stadtman, E. R.; Levine, R. L. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids 25:207-218; 2003.

[15] Dean, R. T.; Fu, S.; Stocker, R.; Davies, M. J. Biochemistry and pathology of radical-mediated protein oxidation. Biochem J 324 (Pt 1):1-18; 1997.

[16] Berlett, B. S.; Stadtman, E. R. Protein oxidation in aging, disease, and oxidative stress. J Biol Chem 272:20313-20316; 1997.

[17] Requena, J. R.; Chao, C. C.; Levine, R. L.; Stadtman, E. R. Glutamic and aminoadipic semialdehydes are the main carbonyl products of metal-catalyzed oxidation of proteins. Proc Natl Acad Sci U S A 98:69-74; 2001.

[18] Sayre, L. M.; Lin, D.; Yuan, Q.; Zhu, X.; Tang, X. Protein adducts generated from products of lipid oxidation: focus on HNE and one. Drug Metab Rev 38:651-675; 2006.

[19] Levine, R. L.; Garland, D.; Oliver, C. N.; Amici, A.; Climent, I.; Lenz, A. G.; Ahn, B. W.; Shaltiel, S.; Stadtman, E. R. Determination of carbonyl content in oxidatively modified proteins. Methods Enzymol 186:464-478; 1990.

[20] Shacter, E.; Williams, J. A.; Lim, M.; Levine, R. L. Differential susceptibility of plasma proteins to oxidative modification: examination by western blot immunoassay. Free Radic Biol Med 17:429-437; 1994.

[21] Keller, R. J.; Halmes, N. C.; Hinson, J. A.; Pumford, N. R. Immunochemical detection of oxidized proteins. Chem Res Toxicol 6:430-433; 1993.

[22] Robinson, C. E.; Keshavarzian, A.; Pasco, D. S.; Frommel, T. O.; Winship, D. H.; Holmes, E. W. Determination of protein carbonyl groups by immunoblotting. Anal Biochem 266:48-57; 1999.

[23] Davies, S. M.; Poljak, A.; Duncan, M. W.; Smythe, G. A.; Murphy, M. P. Measurements of protein carbonyls, ortho- and meta-tyrosine and oxidative phosphorylation complex activity in mitochondria from young and old rats. Free Radic Biol Med 31:181-190; 2001.

[24] Buss, H.; Chan, T. P.; Sluis, K. B.; Domigan, N. M.; Winterbourn, C. C. Protein carbonyl measurement by a sensitive ELISA method. Free Radic Biol Med 23:361-366; 1997.

[25] Oh-Ishi, M.; Ueno, T.; Maeda, T. Proteomic method detects oxidatively induced protein carbonyls in muscles of a diabetes model Otsuka Long-Evans Tokushima Fatty (OLETF) rat. Free Radic Biol Med 34:11-22; 2003.

[26] Meany, D. L.; Xie, H.; Thompson, L. V.; Arriaga, E. A.; Griffin, T. J. Identification of carbonylated proteins from enriched rat skeletal muscle mitochondria using affinity chromatography-stable isotope labeling and tandem mass spectrometry. Proteomics 7:1150-1163; 2007.

[27] Soreghan, B. A.; Yang, F.; Thomas, S. N.; Hsu, J.; Yang, A. J. High-throughput proteomic-based identification of oxidatively induced protein carbonylation in mouse brain. Pharm Res 20:1713-1720; 2003.

[28] Mirzaei, H.; Regnier, F. Identification of yeast oxidized proteins: chromatographic top-down approach for identification of carbonylated, fragmented and cross-linked proteins in yeast. J Chromatogr A 1141:22-31; 2007.

[29] Bautista, J.; Mateos-Nevado, M. D. Immunological detection and quantification of oxidized proteins by labelling with digoxigenin. Biosci Biotechnol Biochem 62:419-423; 1998.

[30] Yoo, B. S.; Regnier, F. E. Proteomic analysis of carbonylated proteins in two-dimensional gel electrophoresis using avidin-fluorescein affinity staining. Electrophoresis 25:1334-1341; 2004.

[31] Chaudhuri, A. R.; de Waal, E. M.; Pierce, A.; Van Remmen, H.; Ward, W. F.; Richardson, A. Detection of protein carbonyls in aging liver tissue: A fluorescence-based proteomic approach. Mech Ageing Dev 127:849-861; 2006.

[32] Pamplona, R.; Dalfo, E.; Ayala, V.; Bellmunt, M. J.; Prat, J.; Ferrer, I.; Portero-Otin, M. Proteins in human brain cortex are modified by oxidation, glycoxidation, and lipoxidation. Effects of Alzheimer disease and identification of lipoxidation targets. J Biol Chem 280:21522-21530; 2005.

[33] Butterfield, D. A.; Sultana, R. Redox proteomics: understanding oxidative stress in the progression of age-related neurodegenerative disorders. Expert Rev Proteomics 5:157-160; 2008.

[34] Bennaars-Eiden, A.; Higgins, L.; Hertzel, A. V.; Kapphahn, R. J.; Ferrington, D. A.; Bernlohr, D. A. Covalent modification of epithelial fatty acid-binding protein by 4-hydroxynonenal in vitro and in vivo. Evidence for a role in antioxidant biology. J Biol Chem 277:50693-50702; 2002.

[35] Costa, V. M.; Amorim, M. A.; Quintanilha, A.; Moradas-Ferreira, P. Hydrogen peroxide-induced carbonylation of key metabolic enzymes in Saccharomyces cerevisiae: the involvement of the oxidative stress response regulators Yap1 and Skn7. Free Radic Biol Med 33:1507-1515; 2002.

[36] Korolainen, M. A.; Goldsteins, G.; Alafuzoff, I.; Koistinaho, J.; Pirttila, T. Proteomic analysis of protein oxidation in Alzheimer's disease brain. Electrophoresis 23:3428-3433; 2002.

[37] Castegna, A.; Aksenov, M.; Thongboonkerd, V.; Klein, J. B.; Pierce, W. M.; Booze, R.; Markesbery, W. R.; Butterfield, D. A. Proteomic identification of oxidatively modified proteins in Alzheimer's disease brain. Part II: dihydropyrimidinase-related protein 2, alpha-enolase and heat shock cognate 71. J Neurochem 82:1524-1532; 2002.

[38] Reverter-Branchat, G.; Cabiscol, E.; Tamarit, J.; Ros, J. Oxidative damage to specific proteins in replicative and chronological-aged Saccharomyces cerevisiae: common targets and prevention by calorie restriction. J Biol Chem 279:31983-31989; 2004.

[39] Sorolla, M. A.; Reverter-Branchat, G.; Tamarit, J.; Ferrer, I.; Ros, J.; Cabiscol, E. Proteomic and oxidative stress analysis in human brain samples of Huntington disease. Free Radic Biol Med 45:667-678; 2008.

[40] Irazusta, V.; Moreno-Cermeno, A.; Cabiscol, E.; Ros, J.; Tamarit, J. Major targets of iron-induced protein oxidative damage in frataxin-deficient yeasts are magnesium-binding proteins. Free Radic Biol Med 44:1712-1723; 2008.

[41] Magi, B.; Ettorre, A.; Liberatori, S.; Bini, L.; Andreassi, M.; Frosali, S.; Neri, P.; Pallini, V.; Di Stefano, A. Selectivity of protein carbonylation in the apoptotic response to oxidative stress associated with photodynamic therapy: a cell biochemical and proteomic investigation. Cell Death Differ 11:842-852; 2004.

[42] Cabiscol, E.; Belli, G.; Tamarit, J.; Echave, P.; Herrero, E.; Ros, J. Mitochondrial Hsp60, resistance to oxidative stress, and the labile iron pool are closely connected in Saccharomyces cerevisiae. J Biol Chem 277:44531-44538; 2002.

[43] Tsaytler, P. A.; M, C. O. F.; Sakharov, D. V.; Krijgsveld, J.; Egmond, M. R. Immediate protein targets of photodynamic treatment in carcinoma cells. J Proteome Res 7:3868-3878; 2008.

[44] Grimsrud, P. A.; Picklo, M. J., Sr.; Griffin, T. J.; Bernlohr, D. A. Carbonylation of adipose proteins in obesity and insulin resistance: identification of adipocyte fatty acid-binding protein as a cellular target of 4-hydroxynonenal. Mol Cell Proteomics 6:624-637; 2007.

[45] Grimsrud, P. A.; Xie, H.; Griffin, T. J.; Bernlohr, D. A. Oxidative stress and covalent modification of protein with bioactive aldehydes. J Biol Chem 283:21837-21841; 2008.

[46] Fenaille, F.; Tabet, J. C.; Guy, P. A. Identification of 4-hydroxy-2-nonenal-modified peptides within unfractionated digests using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal Chem 76:867-873; 2004.

[47] Mirzaei, H.; Regnier, F. Enrichment of carbonylated peptides using Girard P reagent and strong cation exchange chromatography. Anal Chem 78:770-778; 2006.

[48] Johansson, E.; Olsson, O.; Nystrom, T. Progression and specificity of protein oxidation in the life cycle of Arabidopsis thaliana. J Biol Chem 279:22204-22208; 2004.

Aplicación del iTRAQ para el análisis de proteínas de membrana implicadas en la glicosilación en levaduras[1]

Raquel Martínez-Lópezl[1], Pilar Ximénez-Embún[2], Dolores Gutiérrez[2], Lucía Monteoliva[1], J. Arroyo[1], María Molina[1], Concha Gil[1]

[1] Departamento de Microbiología II. Facultad de Farmacia. Universidad Complutense de Madrid. [2] Unidad de Proteómica. Universidad Complutense de Madrid-Parque Científico de Madrid (UCM-PCM).

Introducción

La glicosilación de proteínas constituye una de las modificaciones postraduccionales más relevantes y mayoritaria. Como modelo eucariota en el estudio de la glicosilación se ha utilizado la levadura *Saccharomyces cerevisiae*. La mayoría de las proteínas presentes tanto en la pared celular como en las membranas fúngicas están glicosiladas. En este trabajo se ha estudiado la expresión diferencial de proteínas presentes en la fracción de membrana, enriquecida en proteínas glicosiladas, de una cepa parental SEY6210 respecto a un mutante defectuoso en el proceso de O-glicosilación (*pmt4*Δ). Debido a la naturaleza de estas proteínas se seleccionó el método de marcaje isobárico "iTRAQ", no basado en gel, para su estudio.

Materiales y métodos

Para la extracción de proteínas de membrana se llevó a cabo la eliminación de la pared celular de la levadura mediante la formación de protoplastos. Tras la rotura mecánica de los mismos y ultracentrifugación de la muestra en un gradiente de sacarosa se obtuvo la fracción enriquecida en membranas. El posterior tratamiento de dicha fracción con carbonato sódico (abre vesículas) y el detergente ASB, condujo a la obtención final de la muestra enriquecida en proteínas de membrana.

Puesto que la metodología iTRAQ permite comparar hasta 4 muestras simultáneamente, cada cepa se marcó por duplicado utilizando los reporters 114 y 116 para dos alícuotas de la cepa parental; y 115 y 117 para dos alícuotas de la cepa mutante. Así pudimos estimar la reproducibilidad del marcaje sobre una misma réplica biológica, exactamente en las mismas condiciones experimentales (Gan *et al*, 2007). La muestra resultante se separó mediante 2D-LC *off-line*. El eluído, aplicado de forma automática sobre una placa, se analizó mediante MALDI-TOF/TOF. Los espectros se enfrentaron contra la base de datos SwissProt usando MASCOT como motor de búsqueda. Los ratios de interés se calcularon utilizando el área bajo la curva de cada ion marcador y se normalizaron respecto a la mediana.

Resultados

De 458 proteínas identificadas (CI%>95%), sólo se consideraron aquellas con 2 ó más péptidos: 240 (CI%>99.99%). Al enfrentar la muestra mutante frente a la control, y teniendo en cuenta la reproducibilidad del marcaje, se observa que un porcentaje importante de las proteínas del mutante han disminuido su expresión. Cabe destacar entre estas proteínas subexpresadas en el mutante, proteínas implicadas o relacionadas con el proceso de glicosilación como Ktr3p, Anp1p, Dpm1p, Stt3p y Mnn1p.

En la actualidad se están analizando nuevas réplicas biológicas.

Referencias

Gan, C. S.; Chong, P. K.; Pham, T. K.; Wright, P. C. 2007. Technical, Experimental, and Biological Variations in Isobaric Tags for Relative and Absolute Quantitation (iTRAQ). Journal of Proteome Research 6: 821-827.

[1] Este trabajo se presentó en las "I Jornadas Bienales de Proteómica" (SEProt, Sitges, 21-22 febrero 2008)

Analysis of thylakoid membrane protein complexes from maize by Blue-Native Gel Electrophoresis and Tandem Mass Spectrometry[1]

A. Campos.[1,2], A. Coelho.[3], L. Arike [4], R. Santos R, C. Franco, P. Carvajal-Vallejos, M.A. Santos[2], J.M. Torné.[2], A.V. Coelho[1,5]

[1)] Instituto de Tecnologia Química e Biológica, Oeiras, Portugal; [2)] Institut de Biologia Molecular de Barcelona , Consorci CSIC-IRTA, Barcelona, Spain; [3)] Institut Nacional des Sciences Apliqués, Toulouse, France; [4)] Tallinn University of Technology, Tallinn, Estonia; [5)] Universidade de Évora, Évora, Portugal. Contact e-mail: acampos@itqb.unl.pt

Introduction

Blue-Native Polyacrylamide Gel Electrophoresis (BN-PAGE) is a powerful technique to separate and analyse protein complexes from biological membranes. The technique was developped to separate protein complexes from mitochondria (Schägger and Von Jagow, 1991) but has been extended to the analysis of proteins of different membrane systems (Heinemeyer et al., 2007). Regarding a research project aimed in the study of the function of maize chloroplastidial transglutaminase (chlTGase) and the characterisation of protein post-translational modifications (PTM's) mediated by this enzyme (Villalobos et al., 2004; Carvajal-Vallejos et al., 2007; Santos et al., 2007) we undertaken a work to optimize the solubilisation of thylakoid protein complexes from maize and characterize the complexes by BN-PAGE and mass spectrometry (MS). Using the non-ionic detergent n-dodecyl-β-D-maltoside (DDM), protein complexes could be efficiently solubilised from thylakoid extracts and separated by BN-PAGE. Protein bands were subjected to trypsin digestion and analysed by micro LC-MS/MS. Thylakoid proteins could be identified and the different complexes separated by BN-PAGE inferred. We regard this approach of relevant analytical power to study the maize thylakoid membrane complexes and modifications occurring in the subunits of each complex upon environmental stimulus. To the best of our knowledge this work represents the first characterization of maize thylakoid protein complexes by BN-PAGE and MS.

Materials and methods

Leaves from 20 day greenhouse growth maize seedlings were collected and thylakoid enriched fractions were obtained (Berthold et al., 1981). Protein concentration was determined by Bradford method. A first experiment was performed in which thylakoid protein complexes were solubilized with non-ionic detergent n-dodecyl-β-D-maltoside (DDM) at 1.3% (w/v) (Heinemeyer et al., 2007), to protein concentrations of 2, 3 and 6 mg /ml. Efficiency of protein solubilization was assessed by BN-PAGE (Heinemeyer et al., 2007). Results are presented in terms of the ratio between detergent and protein weight in the sample. A second experiment was made to verify the effects of detergent concentration (0.95%-1.8%) on BN-PAGE protein patterns, taking as reference the most suitable solubilization condition considered in the first experiment. Protein bands from BN-PAGE gels were isolated, digested according to Pandey et al. (2000) and analyzed by micro LC-MS/MS (LTQ mass spectrometer, Thermo Finnigan). Peptides were concentrated in a Peptrap (ThermoElectron) and separated in a reversed-phase column (ThermoElectron) with a linear gradient of 5% to 80% acetonitrile with 0.1% formic acid at flow rate of 2μL/min. All product ion spectra were searched using TurboSEQUEST (Bioworks 3.3.1, Thermo Finnigan, San Jose, CA, USA) against Uniprot database (downloaded from NCBI website, 5128948 protein entries). Peptide recognition take in account a peptide probability superior to 0,01 and Xcorr correlation with charge state of 1.5, 2.0, 2.5 and 3.0 for 1+, 2+, 3+ and 4+ charged peptides.

Results

Proteins from thylakoid membranes were solubilized with different ratios of detergent to protein

[1] This work was presented at the "I Jornadas Bienales de Proteómica" (SEProt, Sitges, 21-22 febrero 2008)

and separated by BN-PAGE (Fig. 1). The lowest ratio tested that allowed obtaining a clear BN-PAGE pattern was 5:1. This BN-PAGE pattern is consistent with the BN-PAGE patterns reported for other species thylakoid membranes, which lead us to consider this detergent:protein ratio appropriated for the solubilization of thylakoid membrane complexes from maize. Further increase in detergent to protein ratio lead to changes in BN-PAGE pattern resulting in the increase of the complex with approximately 242 kDa and the decrease of the high MW complexes (Fig.1). In order to identify the complexes being separated by BN-PAGE, protein bands were isolated, subjected to in-gel trypsin digestion and the resulting peptide mixtures analyzed by LC-MS/MS.

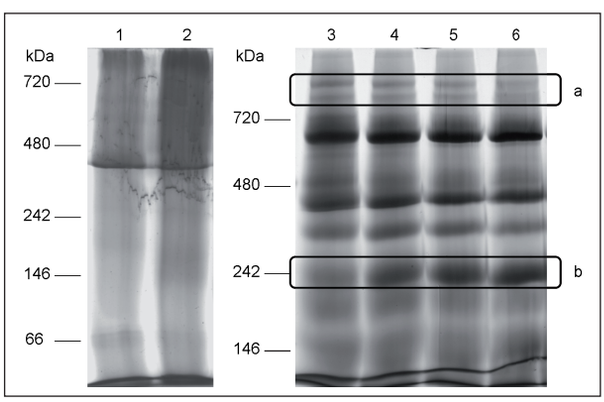
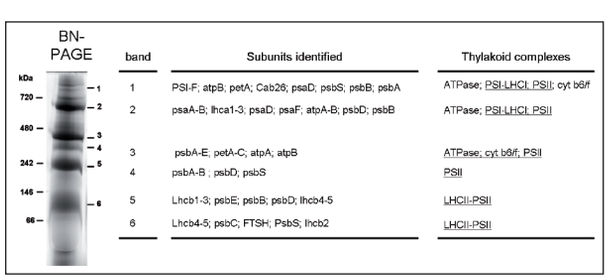


Figure 1. *Solubilization of thylakoid protein complexes and separation by BN-PAGE. Solubilization performed at a DDM to protein weight ratio of 1:1 (1), 2:1 (2), 5:1 (3), 6.4:1 (4), 8:1 (5) and 10:1 (6). Decrease (a) and increase (b) in abundance of thylakoid complexes.*

Distinct proteins that are part of the composition of the major thylakoid protein complexes were identified (Fig. 2). With this information we could infer the presence of the several complexes separated in BN-PAGE and respective level of organization (Fig.2).



band	Subunits identified	Thylakoid complexes
1	PSI-F; atpB; petA; Cab26; psaD; psbS; psbB; psbA	ATPase; PSI-LHCI; PSII; cyt b6/f
2	psaA-B; lhca1-3; psaD; psaF; atpA-B; psbD; psbB	ATPase; PSI-LHCI; PSII
3	psbA-E; petA-C; atpA; atpB	ATPase; cyt b6/f; PSII
4	psbA-B ; psbD; psbS	PSII
5	Lhcb1-3; psbE; psbB; psbD; lhcb4-5	LHCII-PSII
6	Lhcb4-5; psbC; FTSH; PsbS; lhcb2	LHCII-PSII

Figure 2. *Protein identification by LC-MS/MS and assessment of thylakoid complexes separated by BN-PAGE. Underlined are the complexes that were reported also by other authors.*

A general BN-PAGE pattern for maize can be described that is in accordance with previous works of BN-PAGE from thylakoid membranes of higher plants. High similarity is verified with barley thylakoid protein complexes separated by BN-PAGE (Ciambella et al., 2005) consisting in the existence of LHCII-PSII and LHCI-PSI supercomplexes (> 600 kDa), LHCI-PSI complex and PSII dimeric core (band 2), Cyt b6/f dimer, ATPase and PSII monomer core complexes (bands 3 and 4) and the trimer and monomer LHCII (bands 5 and 6, respectively).

Conclusions

In this work we were able to establish a procedure to solubilize maize thylakoid protein complexes. Blue-native PAGE allowed the separation of complexes and further analysis by MS. Protein identification was achieved and sequence information of the several complex subunits obtained. This approach allows a level of fractionation of thylakoid membrane proteome suitable for a deeper characterization of the subunits of the complexes including LHCII proteins. Ongoing work consists in establish the second-dimension from BN-PAGE in denaturing conditions, the analysis of thylakoid proteins by MS in conditions of chloroplastidial transglutaminase activity and identification of post-translational modifications mediated by this enzyme.

References

- Berthold, D.A., Babcock, G.T., Yocum, C.F., A highly resolved, oxygen-evolving photosystem II preparation from spinach thylakoid membranes : EPR and electron-transport properties. *FEBS Lett* 1981, *134*, 231-234.

- Carvajal-Vallejos, P.K., Campos, A., Fuentes-Prior, P., Villalobos, E., Almeida, A.M., *et al*., Purification and in vitro refolding of maize chloroplast transglutaminase over-expressed in Escherichia coli. *Biotechnol Lett* 2007, *29*, 1255-1262.

- Ciambella, C., Roepstorff, P., Aro, e.M., Zolla, L., A proteomic approach for investigation of photosynthetic apparatus in plants. *Proteomics* 2005, *5*, 746-757.

- Heinemeyer, J., Lewejohann, D., Braun, PT., Blue-native gel electrophoresis for the characterization of protein complexes in plants, in: Thiellement, H., Zivy, M., Damerval, C., Méchin, V., (Eds.), *Methods in Molecular Biology,* Humana Press, Totowa, NJ 2007, *335,* pp. 343-353.

- Pandey, A., Andersen, J.S., Mann, M., Use of mass spectrometry to study signaling pathways. *Science's Stke* 2000, *37*, 1-12.

- Santos, M., Villalobos, E., Carvajal-Vallejos, P., Barberá, E., Campos, A., *et al*., Immunolocalization of maize transglutaminase and its substrates in plant cells and in Escherichia coli transformed cells, in: Mendez Vilas, A., Díaz Alvarez, J., (Eds.), *Modern research and Educational Topics in Microscopy* 2007, *1*, 212-213.

- Schägger, H., Von Jagow, G., Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* 1991, *199*, 223-231.

- Villalobos, E., Santos, M., Talavera, D., Rodríguez-Falcón, M., Torné, J.M., Molecular cloning and characterization of a maize transglutaminase complementary DNA. *Gene* 2004, *336*, 93-104.

Proteomics as a complementary tool for identifying unintended side effects occurring in transgenic maize seeds as a result of genetic modifications[1]

Maria G. Egidi, Anna Maria Timperio, Sara Rinalducci, Giovanni Monastra, Lello Zolla*

Lab. of Proteomics and Genomics, Department of Environmental Sciences, University of Tuscia, 01100, Viterbo, Italy. * Istituto Nazionale di Ricerca per gli Alimenti e la Nutrizione (INRAN), Roma, Italy

In order to improve the probability of detecting unintended side effects during maize gene manipulations by bombardment, proteomics was used as an analytical tool complementary to the existing safety assessment techniques. Since seed proteome is highly dynamic, depending on the species variability and environmental influence we analyzed the proteomic profiles of one transgenic maize variety (event MON 810) in T5 and T6 generation with their respective isogenic controls (F5 and F6). Thus, by comparing the proteomic profiles of F5 with F6 we could determine the environmental effects, while the comparison between F6 and T6 seeds from plants grown under controlled conditions allowed us to investigate the effects of DNA manipulation. Finally, by comparing T5 with T6 seed proteomes it was possible to get some indications about similarities and differences between the adaptations of transgenic and isogenic plants to the same strictly controlled growth environment. Approximately 100 total proteins resulted differentially modulated in the expression level as a consequence of the environmental influence (F6 *vs.* F5), whereas 43 proteins resulted up- or down-regulated in transgenic seeds with respect to their controls (T6 *vs.* F6), which could be specifically related to the insertion of a single gene into a maize genome by particle bombardment. Transgenic seeds responded differentially to the same environment as compared to their respective isogenic controls, as result of the genome rearrangement derived from gene insertion. To conclude, an exhaustive differential proteomic analysis allows to determine similarities and differences between traditional food and new products (substantial equivalence), and a case-by-case assessment of the new food should be carried out in order to have a wide knowledge of its features.

[1] This contribution was presented at the II Meeting (WG 2) of the Plant Proteomics in Europe- COST Action FA 0603

Lista de términos y expresiones técnicas en inglés y su traducción recomendada al castellano

Joaquín Abián, Jesús Vázquez

Esta lista será actualizada en los siguientes números de la revista, de acuerdo a las sugerencias recibidas por los socios. Esperamos vuestras contribuciones y comentarios, que podéis enviar a Joaquín Abian (Joaquim.Abian.csic@uab.cat) o Jesús Vázquez (jvazquez@cbm.uam.es).

Se recomienda usar las abreviaturas inglesas en todos los casos. También se recomienda indicar la traducción al castellano la primera vez que se usa el término o la expresión cuando se prefiera utilizar la notación inglesa original.

Término en Inglés	***Traducción recomendada al Castellano***
hard/soft ionization	ionización dura/suave
field desorption (FD)	desorción por campo eléctrico
plasma desorption (PD)	desorción por plasma
fast atom bombardment (FAB)	bombardeo con átomos rápidos
electron ionization (EI)	ionización electrónica o ionización por electrones
secondary ion mass spectrometry (SIMS)	espectrometría de masas de iones secundarios
electrospray ionization (ESI)	ionización por *electrospray*
Microelectrospray	*microelectrospray*
Nanospray	*nanospray*
Ionspray	*ionspray*
matrix-assisted laser desorption ionization (MALDI)	ionización mediante desorción por láser asistida por matriz
surface enhanced laser desorption ionization (SELDI)	ionización por láser inducida en superficie
time of flight (TOF)	tiempo de vuelo
ion trapping	confinamiento de iones
ion trap (IT)	trampa iónica o trampa de iones
quadrupole ion trap (QIT)	trampa iónica cuadrupolar
3D-ion trap	trampa iónica tridimensional
lnear ion trap (LIT)	trampa iónica lineal
linear trap quadrupole (LTQ)	cuadrupolo de confinamiento lineal
reflector (for TOF analyzers)	reflectrón
ion cyclotron resonance (ICR)	resonancia ciclotrónica de iones
Fourier transform (FT)	transformada de Fourier
Fast Fourier transform (FFT)	transformada rápida de Fourier
Orbitrap	*Orbitrap*
ion monitoring	monitorización de iones
full width half mass (FWHM)	anchura de pico a media altura
delayed extraction (DE)	extracción retardada

Término en Inglés	*Traducción recomendada al Castellano*
bottom-up approaches	métodos *bottom-up* o de abajo hacia arriba
top-down approaches	métodos *top-down* o de arriba hacia abajo
on-line (with)	*on-line*, en línea ó acoplado (a)
off-line	*off-line*, fuera de línea o desacoplado
mass selective instability	desestabilización selectiva de masas
(protein) band/spot	banda/mancha de proteína
peptide mapping	mapeo peptídico
peptide mass fingerprint	huella de masas peptídicas
(protein identification by) peptide mass fingerprinting (PMF)	identificación de proteínas a partir de la huella de masas peptídicas
(protein) sequence coverage	cobertura de secuencia (de una proteína)
collision-induced dissociation (CID)	disociación inducida por colisión
surface-induced dissociation (SID)	disociación inducida por superficie
infrared multiphoton dissociation (IRMPD)	disociación por (haces de) fotones infrarrojos
electron-capture dissociation (ECD)	disociación por captura de electrones
electron-transfer dissociation (ETD)	disociación por transferencia de electrones
sustained off-resonance irradiation collision-induced dissociation (SORI-CID)	disociación inducida mediante colisión por irradiación sostenida no resonante
tandem mass spectrometry (MS/MS)	espectrometría de masas en tándem
sequence tag	marca ó etiqueta de secuencia, subsecuencia
database search engine	motor de búsqueda en base de datos
mock or decoy database	base de datos señuelo ó aleatoria
post-translational modification (PTM)	modificación post-traduccional
shotgun analysis	???*

***** Animamos a los socios a presentar sugerencias para la traducción al castellano de esta expresión. Las sugerencias enviadas se discutirán en el próximo número de la revista.

Instrucciones a los autores

Proteómica es la revista oficial de la Sociedad Española de Proteómica. Su periodicidad será semestral y será publicada por el Servicio de Publicaciones de la Universidad de Córdoba.

La revista publicará artículos originales y comunicaciones breves de contenido científico o técnico, así como artículos de revisión, tutoriales y opiniones, notas o comentarios sobre cualquier aspecto relacionado con la proteómica. Incluirá información sobre nuestra Sociedad y sobre los socios, grupos e instituciones que la componen. El idioma será el castellano, aunque se admitirán contribuciones en otras lenguas, preferentemente el inglés.

Aunque *Proteómica* se distribuirá preferentemente en España y Latinoamérica, pretende tener un carácter internacional, extendiendo su distribución a otros países. Esta revista se envía, sin coste alguno, a los socios de la SEProt, y a las instituciones y organismos públicos o privados vinculados a la SEProt o al campo de la proteómica. Existirá una versión impresa, editada por el Servicio de Publicaciones de la UCO, y una versión "on-line" que aparecerá en la página web de la SEProt y de la Universidad de Córdoba.

Todas las contribuciones serán revisadas por el comité editorial, y su formato deberá ajustarse a las instrucciones que se adjuntan. Los artículos originales, las comunicaciones breves, las revisiones y los tutoriales serán evaluados científicamente por uno o dos revisores elegidos por el comité editorial. Todas las contribuciones reflejan la opinión de sus autores y no necesariamente la opinión del Comité Editorial ni de la Junta Directiva de la SEProt.

Los manuscritos se enviarán por correo electrónico, a la dirección bf1jonoj@uco.es.

Los artículos originales, las revisiones y los tutoriales deberán ir acompañados de una carta de presentación del trabajo (cover letter) y tendrán una extensión máxima de 15 páginas A4; las comunicaciones breves también deberán incluir una carta de presentación y tendrán una extensión máxima de 8 páginas A4; el resto de las contribuciones tendrá una extensión máxima de 4 páginas. Para el cálculo de la extensión se tendrán en cuenta, además del texto, las figuras, las tablas, las ilustraciones y las referencias. No se considera la publicación en color, por lo que las ilustraciones, fotografías y gráficos aparecerán en blanco y negro. Deberán enviarse, en archivos separados, por una parte el texto y las tablas (preferentemente en Word) y por otra las figuras (en formato pdf, a 250-300 dpi de resolución y al tamaño de impresión final). Los autores deberán verificar que los detalles de las figuras se imprimen a partir de los ficheros con la calidad requerida y que el texto que incluyan las figuras sea legible al tamaño de reproducción final. Las tablas y leyendas de figuras deben ir al final del texto. El texto debe ajustarse al siguiente formato: letra Times New Roman, tamaño 12, doble espacio, páginas y líneas numeradas, justificación total.

Los artículos originales deben tener las siguientes secciones:

Portada. Incluirá el título, la lista de autores (nombre y apellido(s)) y su afiliación, y los datos completos del autor con el que se mantendrá la correspondencia. La afiliación de los autores, en el caso de que exista más de una, se indicará mediante un símbolo en formato superíndice detrás del nombre de cada autor.

Resumen (máximo de 250 palabras) y palabras clave (máximo de 6, separadas por comas).

Introducción. Deberá evitarse una revisión demasiado extensa del tema y deberá justificar adecuadamente el contenido del trabajo.

Materiales y métodos. Esta sección deberá ser breve, limitándose a describir los procedimientos que sean novedosos y referenciando aquellos ya descritos. Las descripciones tendrán el suficiente detalle para permitir la repetición de los experimentos.

Resultados.

Discusión. Los Resultados y la Discusión podrán agruparse en una única sección.

Referencias

Agradecimientos

Las abreviaturas se incluirán como nota al pié de página la primera vez que aparezcan en el texto.

Como norma general, no deben incluírse abreviaturas ni en el título ni en el resumen.

El sistema de citas bibliográficas y referencias, a partir del número 3 de la revista, será un formato basado en la revista *Journal of Proteomics*. Las referencias se citarán en el texto mediante números entre corchetes cuadrados de acuerdo a su orden de citación: cita simple[1], cita de varias referencias[2-4], cita de referencias alternadas[1, 3-5]. Todas las referencias deben ser numeradas de forma consecutiva. Las referencias a trabajos que no hayan sido publicados no se incluirán en la sección de Referencias; dichos trabajos deberán citarse, en el texto y entre paréntesis de acuerdo al siguiente formato: (Corrales F, Jorrín J, resultados no publicados), (Corrales F, Jorrín J, manuscrito enviado), (Jorrín J, Corrales F, comunicación personal).

Las publicaciones serán citadas en la sección de Referencias según su orden de aparición y de acuerdo a los siguientes ejemplos:

[1] Storey JD y Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440-5.

[2] Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 2002;1:376-86.

[3] Stults JT, Henzel WJ, Wong SC y Watanabe C, *Identification of Electroblotted Proteins by Peptide Mass Searching of a Sequence Database* en Mass Spectrometry in the Biological Sciences (editores: Burlingame A.L. y Carr S.A.), Humana Press, Totowa, New Jersey 1996, pp. 151-70.

[4] Kyte J. Mechanism in Protein Chemistry, Garland Publishing, Inc., 1995.

[5] Castillejo MA. Análisis de los cambios de expresión producidos por un inhibidor de quinasas en células endoteliales. Tesis Doctoral, Universidad de Córdoba, 2005.

Con objeto de facilitar la edición de los manuscritos se ha preparado un fichero de estilo para el programa EndNote denominado *Proteomica.ens*, accesible desde nuestra página web. Para poder usar este fichero hay que colocarlo en el subdirectorio Styles del EndNote.

Las tablas y figuras llevarán numeración arábica y se citarán en el texto en el siguiente formato: tabla 3, figura 2. Las tablas deberán llevar un título y podrán incluir notas a pie de tabla, que se referirán al contenido de la tabla usando símbolos en formato superíndice. Todas las figuras deberán llevar una leyenda conteniendo un titulo lo más representativo posible del contenido de la figura, en negrita, y un texto explicativo lo más conciso posible; la descripción detallada o la interpretación de la figura, en su caso, deberá hacerse en el texto del manuscrito.

Las comunicaciones breves deben ajustarse al mismo formato que los artículos originales, y contendrán un resumen (máximo 250 palabras), una lista de palabras clave (máximo de 6, separadas por comas), una única sección conteniendo el texto principal, y las referencias y agradecimientos, además de las tablas y figuras. El texto principal deberá introducir adecuadamente el tema, explicar la metodología utilizada, y comentar y discutir los resultados.

Los artículos de revisión y los tutoriales tendrán un formato libre pero deberán incluír un Resumen (máximo 250 palabras) y una lista de palabras clave (máximo de 6, separadas por comas). El resto de las contribuciones tendrá un formato libre. Todas las contribuciones deberán respetar el formato de los artículos originales en cuanto a tablas, figuras, abreviaturas y referencias.

POETÓMICA
"Review del ruido"

El ruido campa
Por edredones
De madrugada.
El ruido traza
De sombra y grana
Largos girones
En la mañana.

El ruido alegre
Fragmenta en serio.
El ruido pierde
El agua duende
Y amonio entero.
Pero no tiene
Secuencia, anhelo
Mi ruido verde.

La nube es ruido
De azul prendido
Que al sol venera.
La nube llena
De azar fluído
Es un latido
De altura etérea.

Los cambios claros
Se ven a ojo
Mas los dudosos
No son tan raros.
Y el ruido, ¿es cojo?
Va retozando,
Cambios creando
De puro antojo.

Umbral de gloria
Rozan espectros
Frente a los péptidos
Vueltos de cola.
Y aunque se asoman
Entre los ciertos
Ruido son, vértigo
De carambola.

Del ruido brota
La maravilla.
El ruido borda
De pliegue y roca
La cumbre alpina.
De ruido herida
Y arista rota,
La nieve explota
De luz, arriba.

La ciencia inquiere
Si el ruido puede
Copiar su esquema.
¿Tendrá la idea
Cimiento fuerte
Si unidos crean
Azar y suerte
Indicio leve
De coincidencia?

El ruido es vida
Si se elimina
Su torbellino.
La vida es ruido
Con lente fina.
Mas con precisa
tecnología
Salen del ruido
-¡Fourier, amigo!-
Las proteínas.

SEProt, pequeña,
Tú ya lo sabes,
El euro, ingrave,
Nos escasea.
Mas, eupera,
Tu joven gesta
Saldrá adelante
Seguro, a base
De ruido y clase.

Llevan al ruido
Todas las rutas,
Donde se busca
Siempre escondido.
¡Pero mi ruido!
Ruido dormido,
Ruido que arrullas
Hasta en mi oído.

El ruido corre
Lleno de arrojos
Ante mis ojos.
Y sin sonrojo
La tarde escoge
Para su pose
Fractales rojos.

Jesús Vázquez