

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.

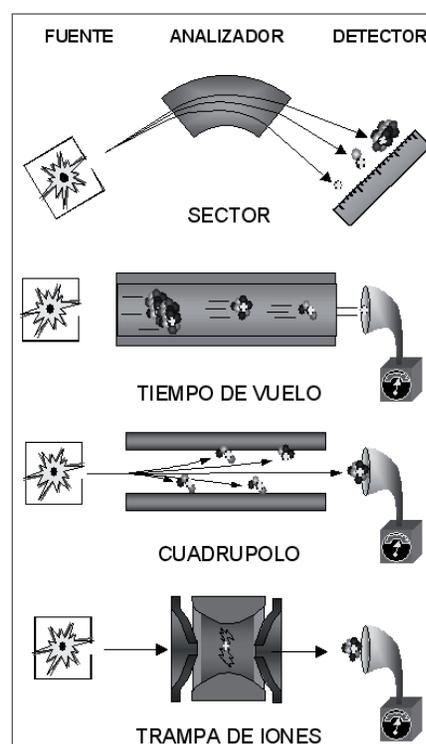


Figura 1.- Analizadores de masa

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 Steinweger en 1953 (Paul y Steinweger, 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 actúa 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 colisión 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 configuración 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 electrostático 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 senoideas de distintas frecuencias (m/z) e intensidades (abundancia) de cada uno de los iones. La aplicación a esta señal de la transformada 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 capacitivamente 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 co-

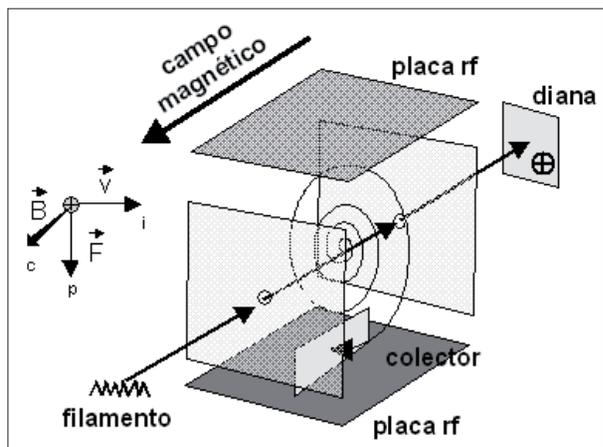


Figura 2.- Celda ICR clásica. La ionización de las moléculas en el interior de la celda se lleva a cabo mediante ionización electrónica. Las placas frontal y trasera tienen aplicado un potencial para el confinamiento de iones en la dirección axial. Entre las placas superior e inferior se aplica una señal 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 , pulgón; i , índice)

lectan 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.

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 width 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 término 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 Resolución 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 comparación con el trabajo en modo convencional (denominado “lineal”), que ofrece un rango de barrido comúnmente 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 Arg⁸-vasopresina (Hunter et 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.

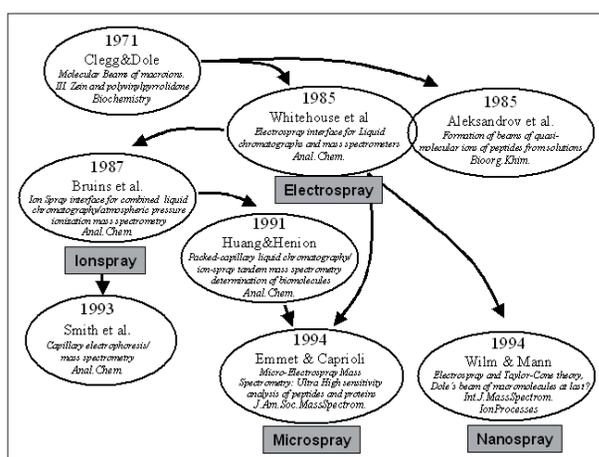


Figura 3. Filogenia de la familia del Electrospray

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 ser 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 sustancia 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.

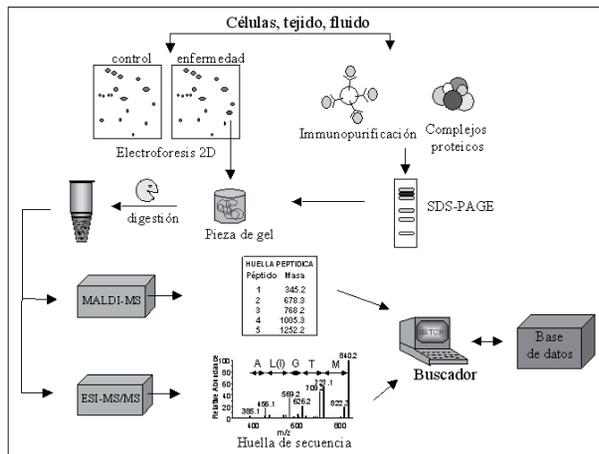


Figura 4.- Esquema del procedimiento básico para la identificación de proteínas en bases de datos (J. Abián, *Tratado de Medicina Interna Rodes-Guardia, Ed Massons (2004)*)

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, Roepstorff, 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/denovo/
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 determinarían 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 número 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 cisteína 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

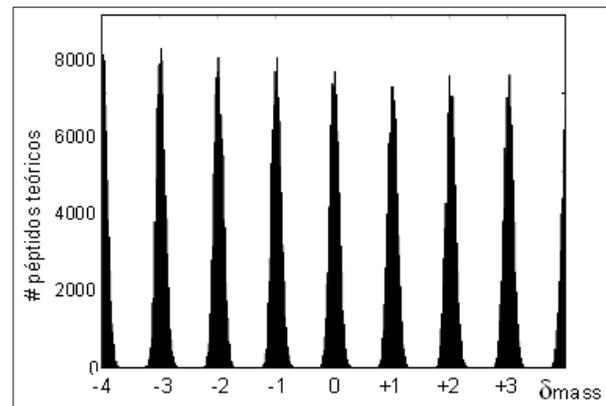


Figura 5.- Distribución del número de péptidos teóricos posibles en un rango de cuatro unidades de masa alrededor de 1000.42 Da. Se han tenido en cuenta la presencia de los aminoácidos tirosina, serina y treonina fosforilados así como de metionina oxidada y cisteína carbamidometilada. La figura es un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

Mascot y Profound que, al utilizar métodos de cálculo 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 predeposada (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 μ L) que permite desalar los digeridos (aprox 20-50 μ L) y concentrarlos hasta 0.5-1 μ L. 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 digestión en menos de 1 minuto.

Por otro lado, se dispone de minicolumnas comerciales con tripsina inmovilizada 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 μ TAS 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 (Ekstrom et al, 2000), y que permite realizar el análisis 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 tri-cargados 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.

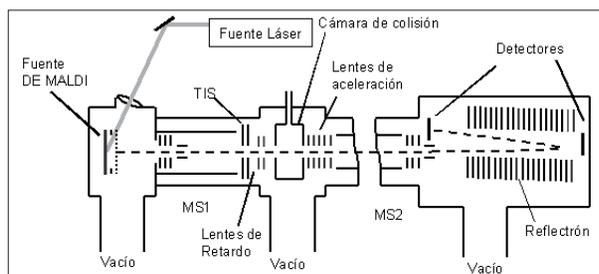


Figura 6.- Sistema MALDI-TOF/TOF (Modificado de Applied Biosystems)

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).

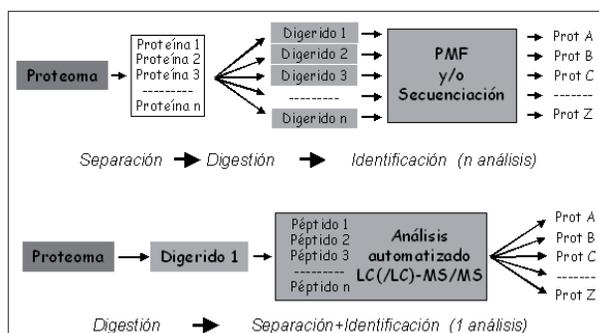


Figura 7.- Análisis convencional comparado con un análisis *shotgun* de un proteoma.

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.

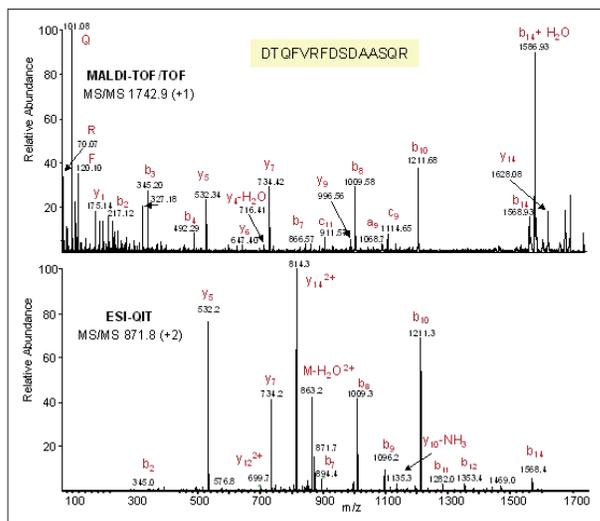


Figura 8.- Espectros de secuenciación mediante MALDI-TOF/TOF y ESI-QIT del péptido DTQVRFDSDAASQR (M.Carrascal, Tesis de licenciatura, UB 2005).

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 interactuar 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 (SORICID, *sustained off-resonance irradiation 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 procedimiento de activación 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 sí 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 difícil 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 proporción estimada de falsos positivos en una población de péptidos 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 búsqueda 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 seleccionan 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 Lutfisk 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 rellenan los espacios vacíos de forma más efectiva, objetivo final del algoritmo de cálculo.

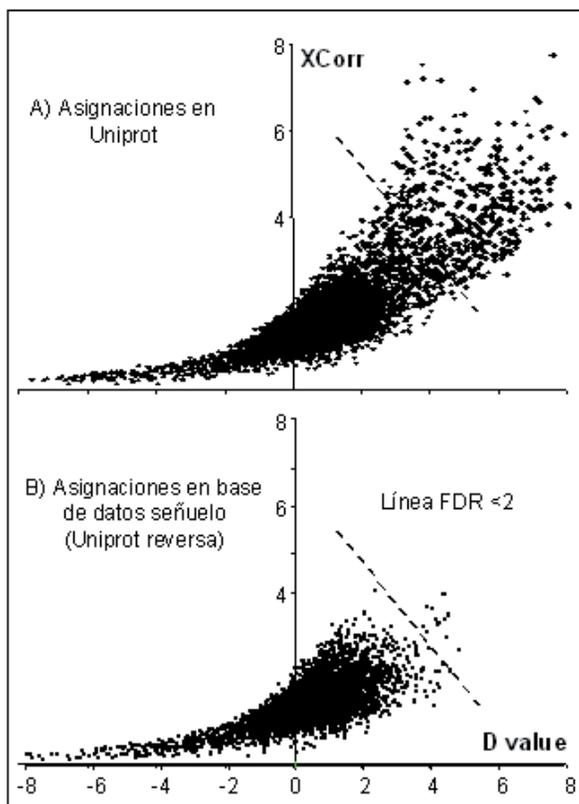


Figura 9.- Utilización de los parámetros Xcorr (SEQUEST) y D (Peptide Prophet) para la determinación de los valores de corte (líneas discontinuas) que producen resultados con un bajo porcentaje de falsos positivos (FDR < 2) en una muestra enriquecida en fosfopéptidos. La base de datos señuelo se ha generado 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 precursor-

res 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 MS² y MS³ 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. Identificación “top-down” de proteínas

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 cuadrupolo (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 fragmentación 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 ambigüedades 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 abordan 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 (Proteo-Red, <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, Hunttenhower C, Geer L Y, Coon J J. et al. 2007. *Analysis of phosphorylation 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, Syebela 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 major 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.
- Medzihradsky 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. *Recent 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, Bartlett-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 surfactants 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.