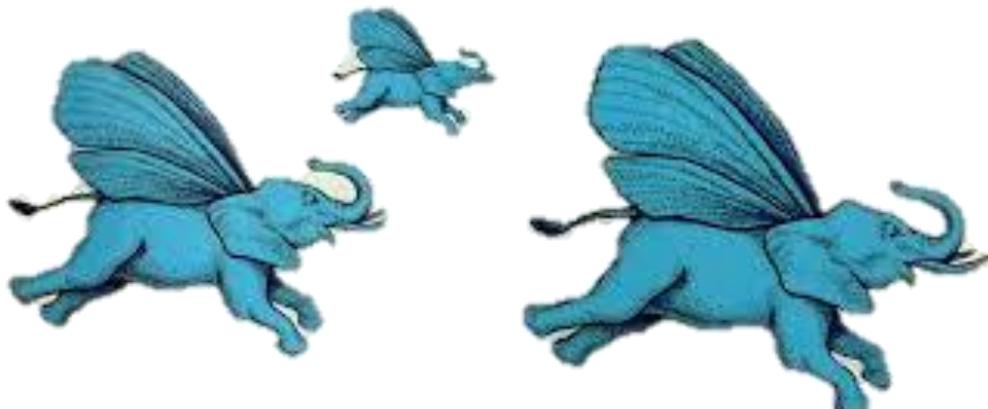


UNIVERSIDAD DE SEVILLA



EL VUELO DE LOS ELEFANTES

El nacimiento de una nueva era



Enedina Muñoz Muñoz

Facultad de Farmacia, Universidad de Sevilla

Curso 2017/18



UNIVERSIDAD DE SEVILLA
FACULTAD DE FARMACIA
GRADO EN FARMACIA



TRABAJO FIN DE GRADO

Revisión bibliográfica

EL VUELO DE LOS ELEFANTES

El nacimiento de una nueva era

Presentado por
Enedina Muñoz Muñoz

Tutor: Jose Luis Espartero Sánchez
Dpto. Química Orgánica y Farmacéutica

Sevilla, julio de 2018

RESUMEN

La espectrometría de masas (EM) es una técnica analítica usada actualmente para el análisis de macromoléculas biológicas que nos permite determinar el peso molecular y la estructura química de éstas. Un espectrómetro de masas tiene tres componentes fundamentales que son la fuente de ionización, el analizador de masas y el detector. En un espectrómetro de masas se convierte a las moléculas en iones gaseosos que luego son diferenciados en el analizador en función de su relación masa/carga (m/z) y, por último, estos iones pasan al detector. Posteriormente, mediante un sistema informático obtenemos la representación gráfica de estos iones frente a su abundancia relativa, que es lo que denominamos el espectro de masas. Las dos técnicas de ionización más utilizadas en la actualidad para el estudio de macromoléculas son la ionización por electrospray (ESI) y la desorción/ionización por láser asistida por una matriz (MALDI), obteniéndose iones multicargados mediante la técnica ESI e iones monocargados mediante la técnica MALDI. Los principales analizadores más utilizados en dichos análisis son cuadrupolo (Q), Orbitrap, tiempo de vuelo (TOF) y la combinación Q-TOF. La técnica de ionización MALDI suele usarse asociado a un analizador TOF.

Antes de que existieran estos métodos de ionización extremadamente suaves en EM, las proteínas se analizaban mediante métodos químicos de degradación como, por ejemplo, la degradación de Edman. Fue a partir de los años 90 cuando se empieza a utilizar la EM en el estudio de las proteínas y, a partir de ahí, surge la proteómica que se ocupa del estudio de las proteínas. La EM presenta actualmente múltiples aplicaciones, siendo una de las más novedosas la del Proyecto del Proteoma Humano.

Palabras clave: espectrometría de masas, proteómica, ESI, MALDI.

ÍNDICE

GLOSARIO DE ABREVIATURAS	5
1. INTRODUCCIÓN.....	7
2. OBJETIVOS DE LA REVISIÓN	11
3. METODOLOGÍA	11
4. RESULTADOS Y DISCUSIÓN	13
4.1. ESPECTROMETRÍA DE MASAS	13
4.1.1. Fuente de ionización	15
4.1.1.1. ESI.....	15
4.1.1.2. MALDI	16
4.1.2. Analizador	18
4.1.2.1. CUADRUPOLO	18
4.1.2.2. ORBITRAP	19
4.1.2.3. TOF.....	19
4.1.2.4. Q-TOF	21
4.1.3. Detector	21
4.2. MÉTODOS QUÍMICOS PARA EL ANÁLISIS DE PROTEINAS.....	21
4.3. ESPECTROMETRÍA DE MASAS EN EL ANÁLISIS DE PROTEÍNAS. PROTEÓMICA ..	23
4.3.1. Bottom-up.....	26
4.3.2. Top-down.....	26
4.4. APLICACIONES.....	27
5. CONCLUSIONES.....	31
6. BIBLIOGRAFÍA.....	33

GLOSARIO DE ABREVIATURAS

CI: ionización química

EI: impacto electrónico

EM: espectrometría de masas

ESI: ionización por electrospray

FAB: bombardeo con átomos rápidos

HPLC: cromatografía líquida de alta resolución

IT: analizador de trampa de iones

MALDI: desorción/ionización por láser asistida por una matriz

m/z: relación masa/carga

PITC: fenilisotiocianato

PTH: fenilhidantoína

PTM: modificaciones post-traduccionales

Q: analizador de cuadrupolo

Q-TOF: analizador híbrido cuadrupolo-tiempo de vuelo

RMN: Resonancia Magnética Nuclear

SDS-PAGE: electroforesis unidimensional

TFG: trabajo fin de grado

TOF: analizador de tiempo de vuelo

UFMP: ultra fine metal powder

2D-PAGE: electroforesis bidimensional

1. INTRODUCCIÓN

“El descubrimiento de la estructura del ADN por Watson y Crick (Watson and Crick, 1953) abrió una etapa revolucionaria en la biología permitiendo el conocimiento de los mecanismos que regulan la síntesis de proteínas, la división y diferenciación celular, las bases moleculares de múltiples patologías y un largo etc. El desarrollo de la genómica alcanza su culmen con la descripción del genoma humano y de otras especies animales y vegetales. La genómica, sin embargo, nos ha mostrado también algunas de sus limitaciones. Los estudios de identificación y regulación de genes no aclaran que sucede *a posteriori* con las proteínas, siendo el estudio de sus modificaciones post-traduccionales un campo nuevo con infinitas posibilidades. Nos permitirá conocer como los diferentes elementos intracelulares se relacionan entre sí y como unas células interaccionan con el resto” (Aguilar, 2007).

Las proteínas son macromoléculas orgánicas que están formadas por la unión de varios aminoácidos formando enlaces peptídicos: enlace entre el grupo amino (NH_2) de un aminoácido y el grupo carboxilo (COOH) del siguiente. Las proteínas están presentes en todos los organismos, siendo las macromoléculas biológicas más abundantes. Desempeñan una amplia variedad de funciones, siendo esenciales para el funcionamiento celular y, por tanto, para la vida. Algunas de sus funciones son catalizar las reacciones químicas de nuestro metabolismo, función estructural, también participan en la respuesta inmune, etc. Las proteínas son fabricadas en la célula a partir de la información genética codificada por el ADN, con la particularidad de que cada proteína tiene una secuencia de aminoácidos distinta (Pando and Ferreira, 2007).

En este Trabajo Fin de Grado (TFG) vamos a tratar sobre la Espectrometría de Masas (EM) y el análisis de proteínas, así como de la relación entre ambos. La EM es una técnica muy empleada actualmente en el análisis de proteínas, que nos sirve para determinar el peso molecular y la estructura química de las macromoléculas.

Las aplicaciones de la EM al principio se vieron limitadas debido a que solo podía usarse con analitos termorresistentes, de bajo peso molecular (700-800 Da) y muy volátiles. En la década de los 70 se consiguió ionizar moléculas termolábiles pero el análisis de

grandes macromoléculas seguía siendo un problema ya que éstas se descomponían al intentar ionizarlas (Erra-Balsells, 2004).

En los primeros años de la década de 1980, el ingeniero japonés Koichi Tanaka se encontraba realizando estudios en los que empleaba como analitos polímeros (polietilenglicoles) y proteínas que mezclaba con polvo metálico (ultra fine metal powder: UFMP); a esta mezcla le añadía acetona que al ser introducida en la cámara de ionización se evaporaba quedando solamente el analito embebido en el polvo metálico. Un día, por error, en lugar de usar acetona usó glicerina, pero en vez de tirar la preparación la introdujo en la cámara de ionización y realizó el experimento antes de que la glicerina se evaporara. Esta preparación al ser bombardeada con un láser ultravioleta de nitrógeno produjo el resultado que Tanaka estaba buscando: la ionización de macromoléculas (proteínas y polímeros) (Tanaka, 2002). Estos resultados fueron publicados en 1988 (Tanaka et al., 1988).

Coincidiendo en el tiempo, el Prof. Fenn y sus colaboradores estaban trabajando con un método diferente de ionización mediante electrospray. Cuando lo aplicaron a moléculas de proteínas, encontraron que podían obtener el peso molecular de las mismas a partir de los iones formados a partir de ellas. Los resultados de su trabajo pionero fueron publicados en 1989 (Fenn et al., 1989) en la revista Science.

Por el desarrollo de ambas técnicas, en el año 2002 se les concedió el Premio Nobel de Química a ambos investigadores (Griffiths, 2008), premio compartido con Kurt Wüthrich, por el desarrollo de la técnica TROSY en espectroscopia de Resonancia Magnética Nuclear (RMN) (Erra-Balsells, 2004).

The Nobel Prize in Chemistry 2002



John B. Fenn
Prize share: 1/4



Koichi Tanaka
Prize share: 1/4



Kurt Wüthrich
Prize share: 1/2

Figura 1
Galardonados con el Premio Nobel de Química 2002

El prof. Fenn tituló su discurso de aceptación del Premio Nobel con el sugerente título de *“Electrospray Wings for Molecular Elephants”* (Fenn, 2002), que nos ha inspirado el título *“El vuelo de los elefantes”* para este TFG: las proteínas son macromoléculas que, con su elevado peso molecular, se asemejan a elefantes moleculares.

Ambas técnicas, ESI y MALDI, siguen usándose en la actualidad y serán explicadas más adelante.

2. OBJETIVOS DE LA REVISIÓN

El objetivo general de este trabajo es la puesta al día del tema relacionado con la EM y su aplicación en proteómica.

3. METODOLOGÍA

Las principales palabras clave que se han utilizado (en inglés) para llevar a cabo la revisión bibliográfica del tema han sido: mass spectrometry, proteomics, review, electrospray, MALDI.

El proceso de documentación de este TFG se ha realizado en varias etapas. En primer lugar, realizamos una búsqueda que nos permitió una primera aproximación al tema objeto de nuestro estudio centrándonos en los trabajos de revisión anteriormente publicados y que nos permitió definir las palabras clave para una segunda búsqueda más específica. Por último, se realizó una tercera búsqueda más exhaustiva, para localizar o intentar localizar la bibliografía citada en los artículos seleccionados que consideramos de especial interés.

Las búsquedas se realizaron en las bases de datos PubMed, SciFinder, así como en las páginas web de la biblioteca de la Universidad de Sevilla (FAMA), del CSIC y DIALNET donde se ha podido encontrar documentación en español.

Una vez finalizada la búsqueda bibliográfica e identificados los artículos más relevantes para el tema de la revisión, se procedió a una selección de los mismos para incluirlos en la bibliografía de este TFG. Se han seleccionado aquellos artículos cuyo resumen se adecuaba más a los objetivos de la revisión.

Antes de comenzar el trabajo he asistido a un curso para orientarme y saber cómo utilizar de forma correcta todos los recursos que hay disponibles para la búsqueda de información y, además, para aprender a usar correctamente la herramienta Mendeley (gestor de referencias bibliográficas).

4. RESULTADOS Y DISCUSIÓN

4.1. ESPECTROMETRÍA DE MASAS

La EM es una técnica analítica que permite estudiar sustancias muy diversas, tanto orgánicas como inorgánicas. Está basada en el estudio detallado de los iones que se forman al suministrar energía a una muestra y así obtener información tanto cualitativa como cuantitativa. Entre sus grandes ventajas se encuentran las de ser una técnica muy rápida y de alta sensibilidad. La historia de la EM comienza con el profesor Joseph John Thomson en los laboratorios Cavendish de la Universidad de Cambridge. Sus experimentos con el flujo de partículas (electrones) que componen los rayos catódicos le llevó al descubrimiento del electrón en 1897, por el cual le fue otorgado el Premio Nobel en Física en 1906 (Griffiths, 2008). En el año 1912, Thomson construyó el primer espectrómetro de masas, en el cual los iones eran separados por sus distintas trayectorias parabólicas en campos electromagnéticos y la detección era registrada por los choques de los iones en una pantalla fluorescente o en una placa fotográfica (Ayala and De Regil, 2004).

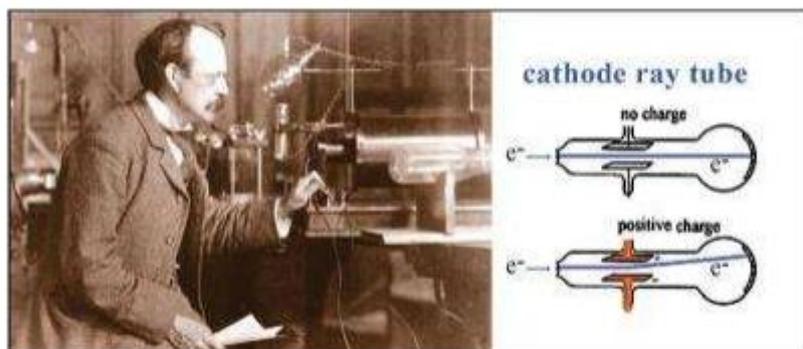


Figura 2

J.J. Thomson y el tubo de rayos catódicos usado para realizar las primeras mediciones de la relación m/z . La desviación del electrón se observaba una vez que se encendía el campo eléctrico (Ayala and De Regil, 2004)

Unos años más tarde, Francis W. Aston y A. J. Dempster construyeron otros espectrómetros de masas que mejoraron la resolución y permitieron separar y probar la existencia de isótopos (Griffiths, 2008) . Por tanto, Thomson, Aston y Dempster contribuyeron al desarrollo de la teoría de la EM.

Esta técnica a partir de que se hizo comercial empezó a ser muy utilizada para la determinación y caracterización estructural de pequeñas moléculas. Hoy en día, los espectrómetros de masas son utilizados para el estudio de muchas sustancias con alta precisión y sensibilidad (Plascencia, 2003).

La EM es una técnica que puede usarse actualmente también para el análisis de macromoléculas biológicas, en especial para las proteínas (Erra-Balsells, 2004). Esta técnica informa sobre el peso y fórmula molecular, tanto de la molécula en sí como de las distintas unidades estructurales que la forman.

Un espectrómetro de masa tiene tres componentes fundamentales: la fuente de ionización, el analizador de masas y el detector. Un cuarto elemento serían las bombas de vacío, que regulan el gradiente de vacío entre las tres partes mencionadas (Erra-Balsells, 2004).

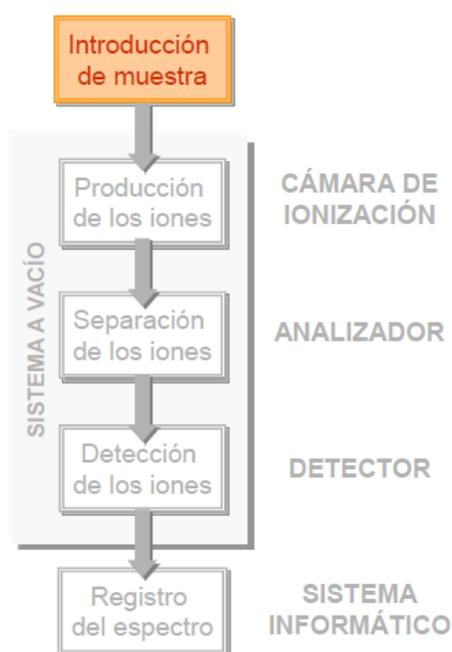


Figura 3
Esquema general de un espectrómetro de masas

“Básicamente, en un espectrómetro de masa se convierte a las moléculas (analito) en iones gaseosos y éstos son diferenciados por el analizador según su relación m/z obteniéndose así información de su peso molecular” (Erra-Balsells, 2004).

4.1.1. Fuente de ionización

En la fuente de ionización la muestra (analito) es ionizada generándose iones gaseosos del analito. Las técnicas de ionización clásicas incluyen la ionización mediante impacto electrónico (EI), la ionización química (CI) y el bombardeo con átomos rápidos (FAB). No obstante, las técnicas de ionización más utilizadas para el análisis de macromoléculas son ESI y MALDI. Estas últimas se consideran técnicas de ionización suave.

4.1.1.1. ESI

ESI, es una técnica popular en EM para la ionización de muestras. Para ello, la muestra de analito se mezcla con un solvente y se hace pasar a través de un capilar de acero inoxidable que termina en una punta muy fina. Este capilar está sometido a alto voltaje de manera que carga las moléculas en el solvente. A la salida del capilar la solución se dispersa en gotículas altamente cargadas con la misma carga. Estas moléculas, una vez que son empujadas a través de la boquilla hacia la cámara de evaporación, se repelen entre sí casi violentamente debido a que las gotitas contienen exceso de cargas en su superficie (cationes o aniones dependiendo de la polaridad del campo aplicado). Cuando el líquido cargado sale por la punta cónica, forma una especie de menisco cónico conocido como *cono de Taylor*, antes de que las gotas se separen unas de otras en un fino spray (Fenn, 2002). Una vez que salen a través de la boquilla, las gotas en el aerosol pasan por una serie de divisiones. Esto ocurre porque el solvente dentro de estas gotitas se evapora gradualmente (con la ayuda de gas nitrógeno bombeado hacia la cámara), reduciendo el tamaño de la gota y agrupando las cargas superficiales. Cuando estos iones son empujados, estando más juntos entre ellos, se terminan repeliendo unos a otros (este comportamiento es conocido como repulsión de Coulomb), lo que hace que las gotas se dividan en gotas más pequeñas (The National High Magnetic Field Laboratory, 2014). Además, *“la repulsión de Coulomb resulta en una divergencia de sus trayectorias para formar una “electropulverización” de gotas cargadas”* (Fenn, 2002). En la figura 4 podemos ver un esquema técnico de la ionización ESI.

El análisis por electrospray se puede hacer provocando una ionización positiva o negativa, depende del voltaje que le apliquemos al capilar. Con esta técnica de ionización obtenemos iones multicargados.

Una de las ventajas de este método de ionización es que las moléculas se mantienen intactas y no se rompen. Este es el motivo por el que este tipo de ionización se llama ionización suave (The National High Magnetic Field Laboratory, 2014).

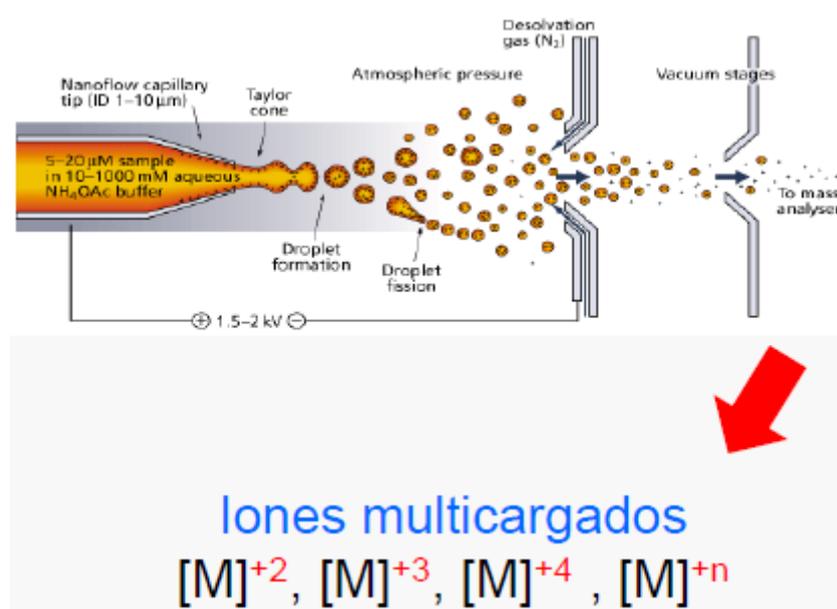


Figura 4
Esquema técnico de ionización ESI

El ESI puede utilizarse acoplado a varios tipos de analizadores como son cuadrupolo (Q), trampa de iones (IT), Orbitrap, tiempo de vuelo (TOF), etc. así como combinaciones de ambos Q-TOF, Q-TRAP, TOF-TOF, siendo el más exitoso el Q-TOF (Abián et al., 2008). Algunos de ellos serán explicados más adelante.

4.1.1.2. MALDI

MALDI, es un método de ionización que ha sido muy utilizado desde su introducción a finales de la década de 1980.

La muestra se tiene que disolver y luego la disolución de la muestra se mezcla con una sustancia matriz que se encuentra en mayor proporción que la muestra (10.000:1). La mezcla se deposita en un dispositivo portamuestras diseñado especialmente para este tipo de ionización. Las dos características clave que debe tener la matriz es que sea de bajo peso molecular y que absorba fácilmente la luz ultravioleta. A continuación, hay que evaporar el disolvente, esto lo conseguimos con una cámara de vacío, al eliminar el aire de la cámara el disolvente se evapora y nos quedan los cristales de la mezcla

muestra-matriz. Tras la evaporación del disolvente, los cristales de la mezcla muestra-matriz se irradian con pulsos cortos de luz láser UV que desorben e ionizan las moléculas de la muestra y la matriz, manteniéndolas en fase gaseosa (Martín and Ballesteros, 2010). Esto se puede ver esquematizado en la figura 5.

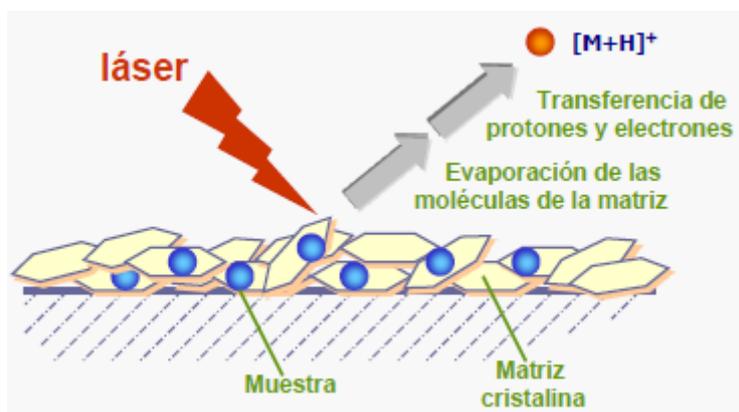


Figura 5
Esquema técnico de ionización MALDI

La longitud de onda del láser utilizado debe coincidir con el máximo de absorción de la matriz, de manera que la matriz es capaz de absorber gran cantidad de energía a la longitud de onda del láser y de transferirla a la muestra y finalmente es la muestra la que se ioniza dando lugar a iones monocargados.

Existen diferentes tipos de matrices en función de la naturaleza del analito y de las aplicaciones como son matrices orgánicas sólidas y líquidas, líquidos iónicos y materiales inorgánicos. La matriz desempeña un papel muy importante en este tipo de técnica y tiene dos funciones, la primera es absorber la energía procedente del láser y la segunda es que actúa como un dispersante del analito reduciendo las fuerzas intermoleculares y disminuyendo la agregación de las moléculas de analito (Martín and Ballesteros, 2010). Esta matriz sirve como una especie de amortiguador entre la muestra y el láser, ya que, si no hubiera una matriz, es decir, si la muestra hubiera sido expuesta directamente al láser, las moléculas de la muestra se habrían roto.

Esta técnica de ionización se usa preferentemente asociada a un analizador TOF y a este equipo se le denomina MALDI-TOF.

4.1.2. Analizador

A continuación, tenemos el analizador de masas donde los iones son separados en función de su relación m/z . El analizador clásicamente utilizaba un campo eléctrico o magnético, o una combinación de ambos, para afectar la trayectoria o la velocidad de las partículas cargadas. Actualmente, además, se están utilizando otros tipos de analizadores, como son cuadrupolo, trampa de iones (IT), Orbitrap, TOF, etc. así como combinaciones de ellos: Q-TOF, Q-TRAP, TOF-TOF. A continuación se explican algunos de ellos.

4.1.2.1. Cuadrupolo

El funcionamiento de este analizador fue descrito por Paul y Steinweger en 1953. Un cuadrupolo consiste en cuatro barras paralelas de sección hiperbólica en la cara interna a través de las cuales se hacen pasar los iones que queremos separar (figura 6). Generalmente estas barras son 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. Estos iones procedentes de la fuente de ionización, 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 relación m/z , de forma que dependiendo del campo eléctrico producido, solamente iones con cierta relación m/z serán dirigidos hacia el detector, todos los demás iones serán desviados hacia los polos. Por lo que, sólo una pequeña fracción del total de iones llega al detector y es monitorizada mientras que el resto se desecha (Abián et al., 2008).

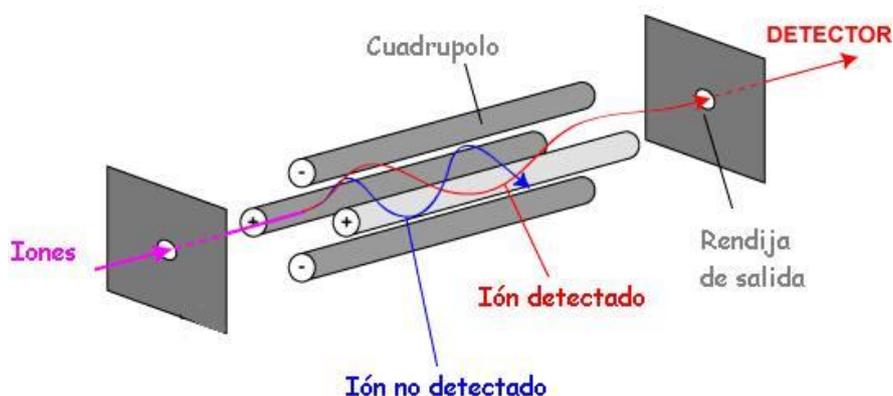


Figura 6
Analizador de masa tipo cuadrupolo

Este analizador permite el uso combinado con técnicas cromatográficas. La principal desventaja de este tipo de analizador es que tiene un poder de resolución muy pequeño, del orden de 500 a 1000, y su limitado rango de utilización, ya que sólo permite la separación de iones con relaciones m/z de hasta 1000 (Abián et al., 2008).

4.1.2.2. Orbitrap

El Orbitrap fue inventado por Alexander Makarov a finales de los años 90. Este analizador consiste en dos electrodos, uno externo en forma de tubo y otro interno que es una barra en forma de huso, las paredes del cilindro que lo contienen han adquirido esta misma forma (figura 7). Los iones son introducidos en el analizador de forma perpendicular y dan vueltas alrededor del huso pero, además, también se mueven hacia delante y hacia detrás a través del eje del electrodo central. Este movimiento podría visualizarse como si estuviéramos bobinando un hilo en un huso convencional. En el Orbitrap el electrodo externo está dividido en el 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 la relación m/z correspondiente.

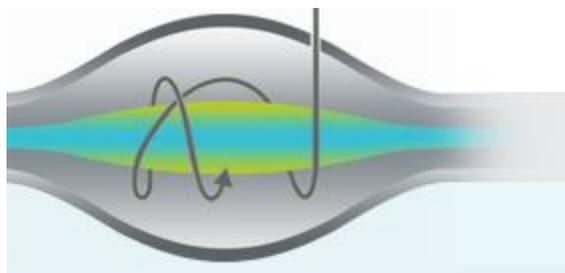


Figura 7
Analizador de masa Orbitrap

El Orbitrap es el analizador de masas más reciente en EM, aunque por sus características se ha introducido rápidamente en la proteómica. Lo interesante de este tipo de analizador es que tiene un alto poder de resolución del orden de 7.500 a 100.000 (a m/z de 400) (Abián et al., 2008).

4.1.2.3. TOF

El analizador de tiempo de vuelo es el más simple de todos. Su funcionamiento se basa en la separación de iones según su relación m/z midiendo sus respectivos tiempos de

vuelo, es decir, el tiempo que tarda cada ión en llegar desde la fuente hasta el detector (los iones generados en la fuente son acelerados por medio de un pulso de potencial eléctrico). El ión más ligero viaja a una velocidad más alta, llegando primero al detector (Abián et al., 2008).

Las principales ventajas que presenta este tipo de analizador son un tiempo de análisis bastante corto con respecto a otros analizadores y, además, tiene un rango de masas ilimitado, es decir, no existe limitación en la masa de los iones que pueden ser separados (Fleurence et al., 1999).

También presenta algunas limitaciones como son baja reproducibilidad y baja resolución. Si queremos aumentar la resolución podemos usar dispositivos de enfoque secundario llamados “reflectrones”. En teoría cuando los iones son acelerados por medio de un pulso de potencial eléctrico todos salen de la fuente con la misma energía cinética ya que se les ha aplicado la misma diferencia de potencial, pero en realidad esto no ocurre así, para corregir esto es para lo que se usan los reflectrones (Tanaka, 2002).

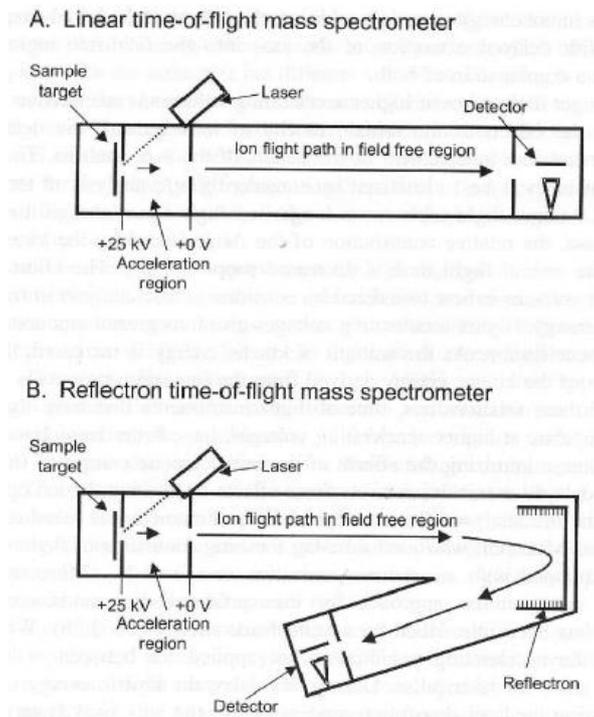


Figura 8

Espectrómetro de masas con tiempo de vuelo. (A) Un sistema lineal de tiempo de vuelo en el cual los iones generados por desorción láser se aceleran desde la fuente de iones hacia el tubo de campo libre que va al detector. (B) Un sistema de tiempo de vuelo con un reflectrón de iones que invierte la trayectoria de vuelos de los iones de manera que corrige las diferencias en la energía cinética de los iones con la misma m/z (Ayala and De Regil, 2004)

4.1.2.4. Q-TOF

En este tipo de analizador híbrido se combinan analizadores TOF y cuadrupolo. Usa un cuadrupolo (cuatro barras paralelas dispuestas en forma cuadrada), una celda de colisión y una unidad de tiempo de vuelo para producir los espectros. Los iones más ligeros avanzan más rápido por el tubo de vuelo y, por tanto, van a llegar antes al detector, determinando así la relación m/z de los iones (Labcompare, 2009).

Los sistemas Q-TOF permiten la obtención de espectros de masa de alta resolución, con errores entre 2 y 5 ppm (Abián et al. 2008).

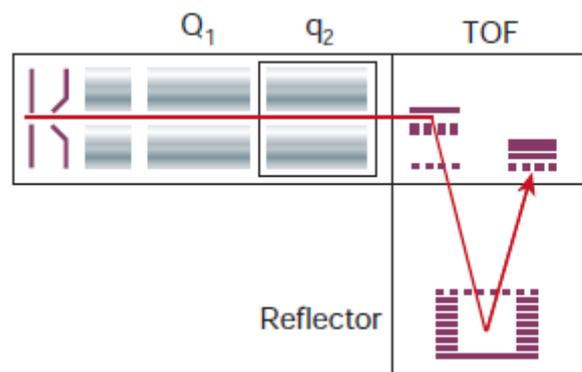


Figura 9

Analizador de masa Q-TOF (Aebersold and Mann, 2003)

4.1.3. Detector

Y, por último, nos encontramos con el detector el cual recibe a los iones diferenciados por el analizador.

4.2. MÉTODOS QUÍMICOS PARA EL ANÁLISIS DE PROTEÍNAS

Inicialmente se analizaban las proteínas mediante métodos químicos de degradación o marcado como, por ejemplo, la degradación de Edman (1950), en la cual el residuo amino terminal es marcado y separado del péptido o proteína sin afectar a los enlaces peptídicos entre los otros residuos. En esta misma década, Frederick Sanger recibió su primer Premio Nobel por determinar la secuencia de la insulina en el año 1958 (Mann, 2016).

El método de la degradación de Edman consiste en la eliminación repetida uno a uno de todos los aminoácidos del extremo N-terminal de proteínas o péptidos y su posterior

determinación (Hernandez et al., 2006). Para ello se hace reaccionar el extremo N-terminal del péptido en cuestión con fenilisotiocianato (PITC) y, de esta manera, se elimina ese extremo en forma de derivado de fenilhidantoína (PTH). A continuación, ese residuo N-terminal es identificado por HPLC. Este procedimiento se va repitiendo “n” veces hasta que se ha determinado toda la secuencia (Chicano, 2011). En la figura 10 se puede ver de forma esquemática todo el procedimiento.

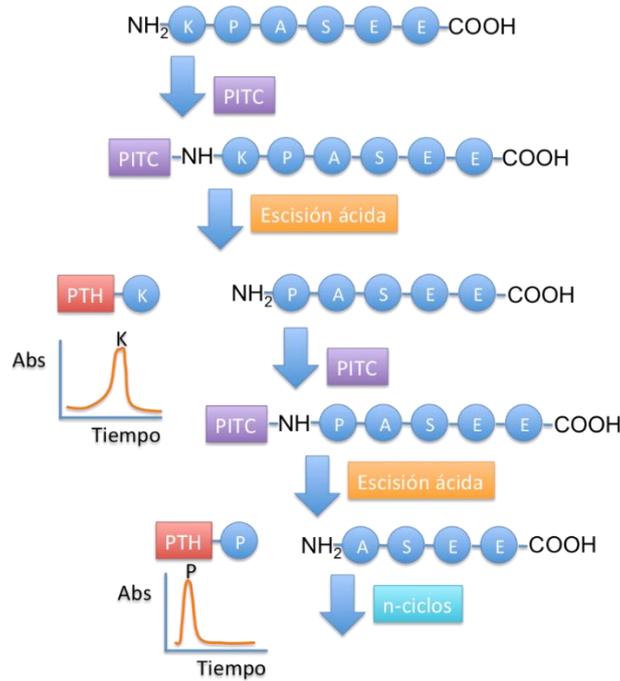


Figura 10
Esquema de la degradación de Edman

Otro método para la secuenciación de péptidos y determinación de la estructura primaria de proteínas combinaba la cromatografía de gases con la EM clásica. Para ello había que reducir los péptidos a poliaminoalcoholes debido a su no volatilidad por su carácter zwitteriónico. Esto se lograba convirtiendo el carboxilato del extremo C-terminal en un éster metílico y el ión amonio del extremo N-terminal en un grupo amido. Posteriormente se reducían todos los grupos carbonílicos presentes con LiAlD₄, para marcar sus posiciones respectivas (figura 11). De esta manera se eliminaba el zwitterión dando lugar a un compuesto más volátil (aminoalcohol). Esto permitía su separación por cromatografía de gases y, posterior introducción de cada fracción en el espectrómetro de masas para producir el espectro (Biemann, 2007).

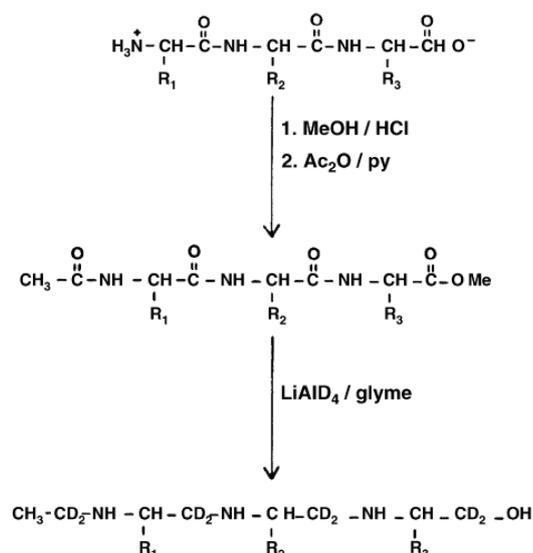


Figura 11

Esquema de reacción para la reducción de péptidos a poliaminoalcoholes (Biemann, 2007)

Los métodos descritos anteriormente se han convertido en la base de la proteómica de hoy en día, siendo actualmente las dos técnicas de ionización más utilizadas para el análisis de macromoléculas el ESI y el MALDI.

4.3. ESPECTROMETRÍA DE MASAS EN EL ANÁLISIS DE PROTEÍNAS. PROTEÓMICA

El término “proteoma” se utilizó por primera vez en 1996 para describir las proteínas codificadas por un genoma (Wilkins et al., 1996a). De forma involuntaria el proteoma se ha ido transformando en “proteómica”, un conjunto de técnicas al que se ha elevado al rango de disciplina científica. ¿Qué es la proteómica? En esencia es el estudio, a gran escala, de la estructura y funciones de las proteínas que integran un sistema biológico.

El peso molecular y el punto isoeléctrico de una proteína, información que proporciona la electroforesis bidimensional, es insuficiente para su identificación y realmente el punto clave para aprovechar la información obtenida en los geles reside en la capacidad para identificar (y caracterizar) las proteínas de interés. Dos hechos han dado un vuelco a esta situación en los últimos años: el desarrollo de técnicas de EM de alta resolución junto con el crecimiento de los bancos de datos de secuencias.

Por tanto, y resumiendo, podemos decir que la proteómica se basa fundamentalmente en dos pilares. En primer lugar, la separación de las mezclas complejas de proteínas, proceso que habitualmente se lleva a cabo mediante electroforesis bidimensional o cromatografía líquida y, en segundo lugar, la identificación de las proteínas separadas para lo que actualmente se recurre a la EM (Fernández-Santarén, 2004).

Las primeras revisiones bibliográficas sobre este tema aparecen en los años 1996-1997 (Wilkins et al., 1996b; Roepstorff, 1997), mientras que el número de artículos publicados desde entonces ha crecido exponencialmente, como se observa en la siguiente figura.

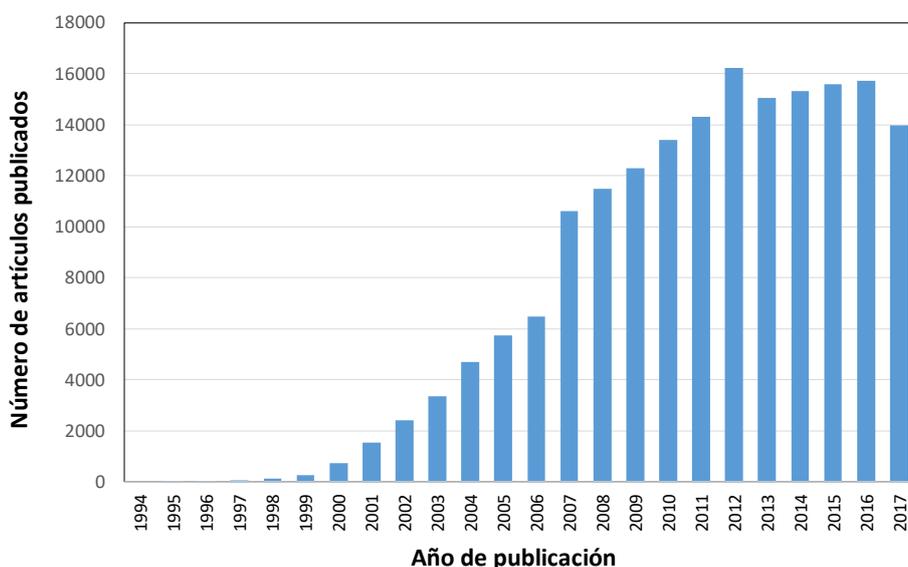


Figura 12

Artículos publicados relacionados con proteómica en la base de datos Scifinder (elaboración propia, 2018)

Inicialmente se empezó a utilizar la EM en el análisis de proteínas por la aplicación del método de ionización FAB (1981). Este método era fácil de utilizar, sólo había que ajustar una pistola atomizadora de argón a la fuente de iones del espectrómetro de masas, esto producía iones muy estables que eran difíciles de fragmentar (Biemann, 2007). No obstante, el método FAB tenía una aplicación muy limitada al estudio de proteínas.

A principios de los años 90, de forma casi simultánea, se desarrollan dos nuevas técnicas de ionización suave, ESI y MALDI, como se ha explicado en la introducción. Estas son las dos técnicas de ionización utilizadas hoy en día para el análisis de macromoléculas, como son las proteínas. Gracias al desarrollo de estas dos técnicas se pudo empezar a

investigar moléculas grandes. De hecho, no es hasta la invención de estas dos técnicas que no se empieza a aplicar masivamente la EM al análisis de proteínas.

La EM permite la identificación de las proteínas, así como la determinación de su estructura primaria, es decir, la secuencia de aminoácidos que componen las proteínas.

Una metodología general para el análisis proteómico es la siguiente:

1. Separación de la muestra: la muestra de proteínas se separa por técnicas cromatográficas o electroforéticas. Las más utilizadas son: electroforesis unidimensional (SDS-PAGE) y bidimensional (2-D PAGE), electroforesis capilar, cromatografía líquida de alta resolución (HPLC), cromatografía de afinidad, cromatografía de exclusión molecular y cromatografía de intercambio iónico. Actualmente la técnica más utilizada es la cromatografía líquida.
2. Procesamiento de la muestra: consiste en la reducción de los puentes disulfuro y la alquilación de las cisteínas de la proteína de interés. Esto tiene como objetivo la desnaturalización de las proteínas para que así puedan ser digeridas a péptidos fácilmente, generalmente se digieren con tripsina y, posteriormente se identifican los fragmentos (Pando and Ferreira, 2007).
3. Espectrometría de masas (explicado en el apartado 4.1).
4. Análisis de datos: actualmente el método más utilizado para la identificación de proteínas es el mapeo peptídico (Abián et al. 2008). Consiste en correlacionar los espectros obtenidos con péptidos teóricos de bases de datos. Entre estas están:

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

Figura 13
Bases de datos proteómica (Abián et al., 2008)

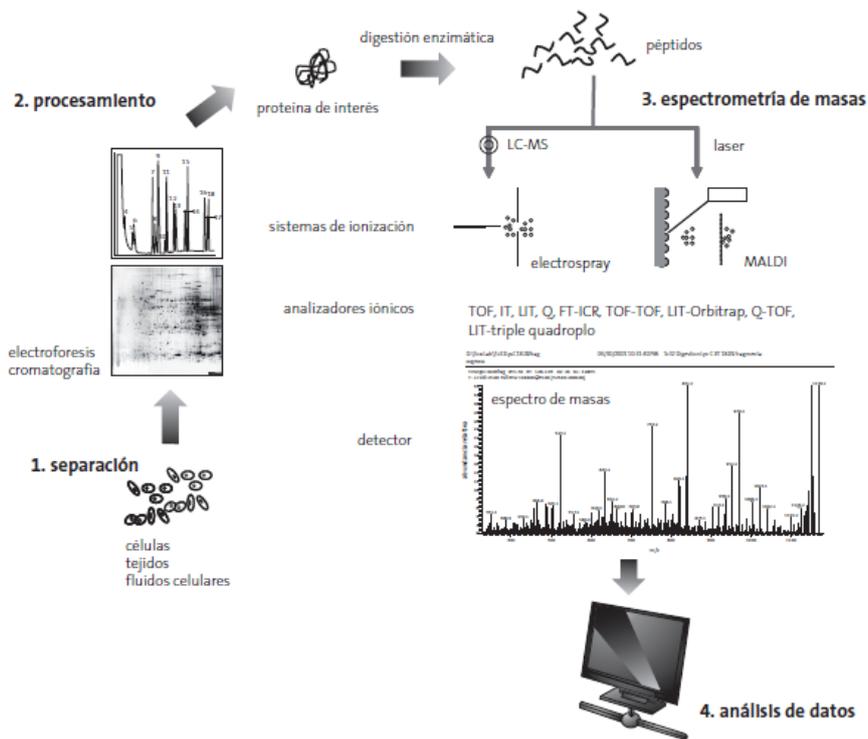


Figura 14
Esquema del análisis de proteínas (Pando and Ferreira, 2007)

Existen varias técnicas para el análisis de proteínas, siendo las dos más importantes las denominadas: Bottom-up y Top-down.

4.3.1. Bottom-up

Consiste en la digestión de la muestra de proteínas generalmente mediante tripsina o a veces digestión química, obteniéndose pequeños péptidos. Estos péptidos son posteriormente separados y analizados por EM en tándem EM-EM. Para separar mezclas complejas se acopla la cromatografía líquida con la EM o EM en tándem (Martín and Ballesteros, 2010). Primero obtenemos un espectro por EM de todo el conjunto de péptidos y, luego, seleccionamos un péptido en concreto y obtenemos un segundo espectro con todos los fragmentos correspondientes a los distintos aminoácidos que están constituyendo ese péptido.

4.3.2. Top-down

En este caso el análisis se hace sobre el conjunto de proteínas nativas intacta, sin digerir, produciéndose la fragmentación dentro del espectrómetro de masas (Martín and Ballesteros, 2010).

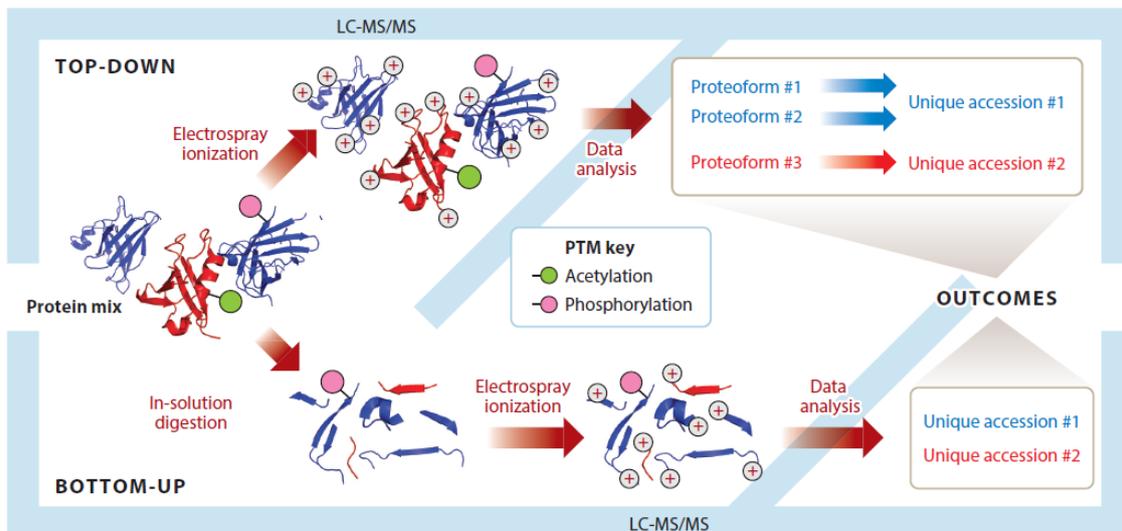


Figura 15

Diagramas de flujo divergentes en proteómica top-down y bottom-up. En el caso de top-down se elimina la etapa de digestión usada en bottom-up para generar los péptidos que se analizan. Abreviaturas: LC-MS/MS, cromatografía líquida acoplada a espectrometría de masas en tándem; PTM, modificación post-traducional (Toby et al., 2016).

Este método supone una evidente ventaja frente al anterior, bottom-up, donde se pierde una cantidad importante de información. Mediante la metodología top-down se obtiene información valiosa sobre las distintas **proteoformas** existentes. Estas proteoformas han sido definidas recientemente (Smith and Kelleher, 2013) como las “*diferentes formas moleculares en las que la proteína producto de un simple gen puede encontrarse, incluyendo cambios debidos a variaciones genéticas o modificaciones post-traduccionales (PTM)*”.

Por último, indicar que en los últimos años un conjunto de investigadores que utilizan esta metodología han creado el “Consortio para la Proteómica Top-Down” (<http://topdownproteomics.org/>).

4.4. APLICACIONES

La proteómica basada en EM puede aplicarse en distintas áreas como, por ejemplo, en el área de la microbiología para la identificación de microorganismos como bacterias u hongos. También se utiliza para el diagnóstico de enfermedades especiales ya que en estos casos se produce una alteración en la estructura de las proteínas. En concreto para

estas aplicaciones suele usarse MALDI-TOF, ya explicado anteriormente (Hosseini and Martinez-chapa, 2017).

En el caso de los recién nacidos, la EM en tándem (EM/EM) permite la identificación de enfermedades mediante el análisis de una pequeña cantidad de sangre del talón (Torres-Sepulveda et al., 2008).

La EM también está siendo utilizada en el **Proyecto del Proteoma Humano**, una iniciativa internacional tremendamente ambiciosa que trata de descifrar todas las proteínas que se expresan en el cuerpo humano y que están codificadas por los 20.300 genes identificados previamente gracias al **Proyecto del Genoma Humano**. Este proyecto está gestionado por la Organización del Proteoma Humano (HUPO) y permitirá una mejor comprensión de la función de las proteínas y, por tanto, avanzar en el diagnóstico y tratamiento de las enfermedades, para después poder utilizarlo en el desarrollo de nuevos fármacos (<https://hupo.org/human-proteome-project>).

El estudio del proteoma es mucho más complejo que el estudio del genoma por diversas razones. El genoma es estático, mientras que el proteoma es dinámico. El proteoma de un mismo individuo es diferente en distintas células de su cuerpo, y aunque el material genético está presente en las mismas cantidades, el nivel de expresión de la proteína que codifica puede variar en un rango de concentración de uno a un millón.

Para el correcto desarrollo de cualquier proceso biológico se necesitan proteínas que sean funcionales y que estén presentes a una determinada concentración en el organismo. Variaciones, aunque sean mínimas, en los niveles de expresión de algunas proteínas pueden desencadenar un proceso patológico (Calamia, 2018).

La correlación entre las proteínas, sus complejos y su estatus en una enfermedad dada acelerará la identificación de nuevas dianas de medicamentos que pueden usarse para diagnosticar y tratar enfermedades.

A continuación, se muestra una imagen (figura 16) del cuerpo humano (mitad hombre y mitad mujer) donde se señalan en cada órgano las proteínas que ya se han determinado. De las 20.055 proteínas humanas asignadas hasta ahora (neXtProt, febrero de 2016), 16.518 están rigurosamente confirmadas, otras 2.949 son actualmente proteínas PE2-

4 (que faltan por confirmar), mientras que las 588 restantes son proteínas PE5, y se consideran solo hipotéticas (Baker et al., 2017).

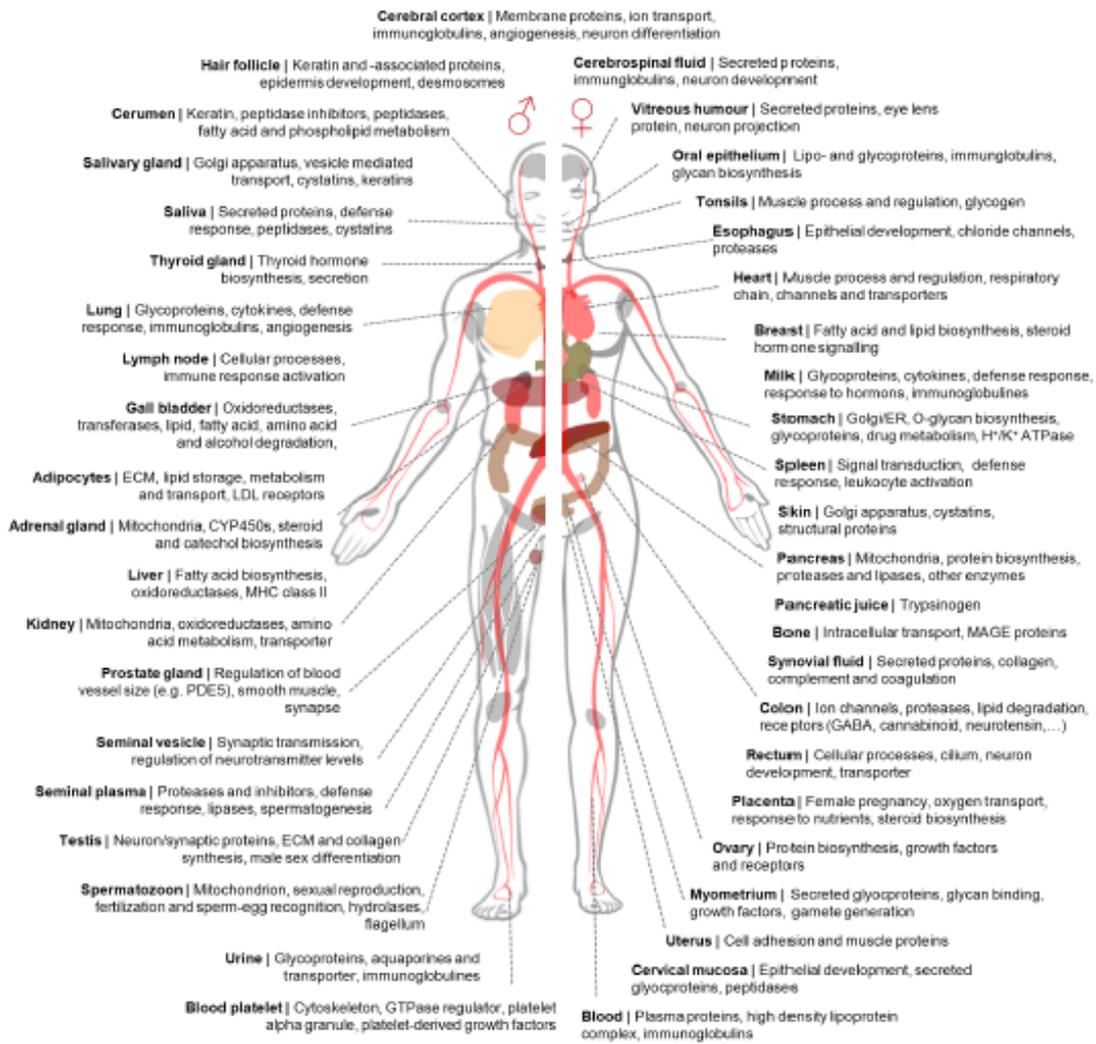


Figura 16
Mapa de expresión proteica (Wilhelm et al., 2014)

5. CONCLUSIONES

A pesar de la máxima délfica "*conócete a ti mismo*" inscrita en la explanada del Templo de Apolo en la antigua Grecia durante el siglo VI aC, todavía no tenemos una descripción exhaustiva de lo que significa ser humano en términos estrictamente moleculares, es decir: *genoma + transcriptoma + proteoma + metaboloma* (Baker et al., 2017).

La proteómica basada en EM es una herramienta de primera elección en el trabajo de multitud de equipos de investigadores que trabajan en el estudio de los sistemas biológicos y sus interrelaciones, gracias fundamentalmente al desarrollo de las técnicas de ionización suave conocidas como MALDI y ESI.

La capacidad de analizar las proteínas en su totalidad, en su entorno natural, a gran escala y de forma sistemática abrirá (está ya abriendo) el camino al estudio generalizado de los mecanismos moleculares de cualquier proceso fisiológico o patológico, y a la detección de biomarcadores de diagnóstico o pronóstico de cualquier enfermedad. Y el ritmo actual con el que están evolucionando los espectrómetros de masas nos hace pensar que en un futuro más próximo de lo que parece será posible cuantificar y caracterizar absolutamente todos los componentes del proteoma humano, en cuestión de minutos (Vázquez, 2013).

6. BIBLIOGRAFÍA

1. Abián J, Carrascal M, Gay M. Introducción a la Espectrometría de Masas para la caracterización de péptidos y proteínas en Proteómica. *Proteómica*. 2008; 2: 17–31.
2. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422(6928): 198–207.
3. Aguilar E. La proteómica: una apuesta fundamental de la Universidad de Córdoba. *Proteómica*. 2007.
4. Ayala C, de Regil R. Secuenciación de proteínas por espectrometría de masas. Instituto de Biotecnología, UNAM, México. 2004; 1–69. [Consultado en Mayo 2018]. Disponible en: http://www.ibt.unam.mx/computo/pdfs/met/secuenciacion_proteinas.pdf
5. Baker MS, Beom S, Mohamedali A, Islam MT, Cantor D, Verhaert PD et al. Accelerating the search for the missing proteins in the human proteome. *Nat. Commun*. 2017; 8(14271): 1-13.
6. Biemann K. Laying the groundwork for proteomics Mass spectrometry from 1958 to 1988. *Int. J. Mass Spectrom*. 2007; 259: 1–7.
7. Calamia V. *Proteómica*. 2018. [Consultado en Mayo 2018]. Disponible en: <http://www.ciber-bbn.es/>
8. Chicano E. Secuenciación de Edman | ProteomePlus [Internet]. 2011. [Consultado en Mayo 2018]. Disponible en: <https://proteomeplus.wordpress.com/2011/10/12/secuenciacion-de-edman/>.
9. Erra-Balsells R. “Del vuelo de las proteínas y de cómo lograrlo” (Espectrometría de masa UV-MALDI). *Química Viva*. 2004; 3(3): 1–24.
10. Fenn JB. Electrospray Wings for Molecular Elephants. Nobel Lecture. 2002; 154-184. [Consultado en Abril 2018]. Disponible en: https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/fenn-lecture.pdf

11. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for Mass Spectrometry of Large Biomolecules. *Science*. 1989; 246(4926): 64-71.
12. Fernández-Santarén J. De la tesis doctoral de Tiselius a la proteómica: setenta y cinco años de electroforesis de proteínas. *Arbor*. 2004; 698: 259-284.
13. Fleurence J, Fayaz M, Namitha KK, Murthy KN, Pérez-Pérez et al. Re. F. Farm. Espectrometría de Masas. *Trends Food Sci Tech*. 1999; 10: 25–28.
14. Griffiths J. A Brief History of Mass Spectrometry. *Anal. Chem*. 2008; 80 (15): 5678-5683.
15. Hernandez P, Müller M, Appel RD. Automated protein identification by tandem mass spectrometry: Issues and strategies. *Mass Spectrom. Rev*. 2006; 25(2): 235-254.
16. Hosseini S, Martinez-Chapa SO. *Fundamentals of MALDI-TOF-MS Analysis*. México: Springer; 2017.
17. Labcompare. Quadrupole Time of Flight Mass Spectrometer (QTOF MS). Labcompare.com [Internet]. 2009. [Consultado en Abril 2018]. Disponible en: <https://www.labcompare.com/Mass-Spectrometry/130-Quadrupole-Time-of-Flight-Mass-Spectrometer-QTOF-MS/Compare/?compare=9077305,11129986&catid=130>
18. Mann M. The rise of mass spectrometry and the fall of Edman degradation. *Clin. Chem*. 2016; 62(1): 293-294.
19. Martín C, Ballesteros M. Espectrometría de masas y análisis de biomarcadores. *Monogr. la Real*. 2010; 113-168.
20. Pando V, Ferreira C. Proteómica: hacia el entendimiento del lenguaje de las proteínas. *Biotecnología*. 2007; 14: 97–108.
21. Plascencia G. Espectrometría De Masas. *Inst. Biotecnol*. 2003; 1–40.
22. Roepstorff P. Mass spectrometry in protein studies from genome to function. *Current Op. Biotechnol*. 1997; 8: 6-13.

23. Smith LM, Kelleher NL. Consortium for Top-Down Proteomics. Proteoform: a single term describing protein complexity. *Nature Methods*. 2013; 10: 186-187.
24. Tanaka KH, Wake H, Ido Y, Akita S, Yoshida Y, Yoshida I. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom*. 1988; 2: 151-153.
25. Tanaka K. The Origin of Macromolecule Ionization by Laser Irradiation. Nobel Lecture. 2002; 197–217. [Consultado en Abril 2018]. Disponible en: https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/tanaka-lecture.pdf
26. The National High Magnetic Field Laboratory. Electrospray ionization (ESI) - MagLab [Internet]. 2014. [Consultado en Abril 2018]. Disponible en: <https://nationalmaglab.org/user-facilities/icr/techniques/esi>
27. Toby TK, Fornelli L, Kelleher NL. Progress in top-down proteomics and the analysis of proteoforms. *Annu. Rev. Anal. Chem*. 2016; 9: 499-519.
28. Torres-Sepulveda MD, Martinez-de Villarreal LE, Esmer C, Gonzalez-Alanis R, Ruiz-Herrera C, Sanchez-Pena A, et al. Expanded newborn screening using tandem mass spectrometry: Two years' experience in Nuevo Leon, Mexico. *Salud Publica Mex*. 2008; 50(3): 200-206.
29. Vázquez J. ¿Quo Vadis, Proteómica? http://dx.doi.org/10.18567/sebbmdiv_ANC.2013.01.1. SEBBM divulgación. 2013.
30. Watson JD, Crick FHC. Molecular Structure of Nucleic Acids. A Structure for Deoxyribose Nucleic Acid. *Nature*. 1953; 4356: 727.
31. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, et al. Mass-spectrometry-based draft of the human proteome. *Nature*. 2014; 509(7502): 582–587.
32. Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez JC et al. From Proteins to Proteomes: Large Scale Protein Identification by Two-Dimensional Electrophoresis and Amino Acid Analysis. *Nature Biotechnol*. 1996a; 14: 61-65.

33. Wilkins MR, Sanchez JC, Williams KL, Hochstrasser D. Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. *Electrophoresis*. 1996b; 17: 830-838.