

REVISIÓN

Importancia de las aplicaciones clínicas de la Proteómica

Ángel San Miguel Hernández*, Rafael San Miguel y Francisco Javier Martín-Gil

Comisión de Investigación y Servicio de Análisis Clínicos, Hospital Universitario Río Hortega, Valladolid, España

Recibido el 30 de octubre de 2008; aceptado el 26 de junio de 2009

Disponible en Internet el 25 de septiembre de 2009

PALABRAS CLAVE

Proteómica;
Aplicaciones clínicas;
Bioinformática

Resumen

La utilidad práctica de los resultados obtenidos con la Proteómica en relación con la salud está cobrando gran importancia. El descubrimiento de marcadores proteicos de enfermedades como las cardiovasculares, las neurológicas, las oncológicas, las metabólicas, entre otras, tiene una aplicación clínica en un futuro próximo en el diagnóstico, el seguimiento y el tratamiento de estas enfermedades.

La HUPO (Human Proteome Organization) (disponible en: www.hupo.org) se creó en el año 2001 para impulsar un mayor conocimiento de la importancia de la Proteómica y las oportunidades que ofrece en el diagnóstico, el pronóstico y el tratamiento de las enfermedades. Se han constituido posteriormente varios grupos: HPPP (Human Plasma Proteome Project), HLPP (Human Liver Proteome Project), PSI (Proteome Standards Initiative), HBPP (Human Brain Proteome Project) y MRPP (Mouse and Rat Proteome Project). Uno de los principales objetivos de la Proteómica es la identificación de marcadores de enfermedad. Un planteamiento ha sido comparar la expresión proteica de los tejidos normales y enfermos para identificar proteínas que se expresen de forma aberrante y que puedan representar nuevos marcadores. Otra estrategia es el análisis de las proteínas segregadas en líneas celulares y cultivos primarios y, finalmente, la obtención de perfiles proteicos en suero.

Los estudios proteómicos están adquiriendo en los últimos años una gran relevancia, fundamentalmente en lo que hace referencia a su aplicación a la enfermedad humana. Con este fin se está realizando un gran número de estudios en el plasma humano, los tejidos y diversos líquidos biológicos.

© 2008 AEBM, AEFA y SEQC. Publicado por Elsevier España, S.L. Todos los derechos reservados.

KEYWORDS

Proteomics;
Clinical applications;
Bioinformatics

Importance of the clinical applications of proteomics

Abstract

The practical usefulness of the results of Proteomics in relation to health is very important. The discovery of protein markers of diseases such as cardiovascular, neurological,

*Autor para correspondencia.

Correo electrónico: asanmiguel@hurh.sacyl.es (A.S. San Miguel Hernández).

oncologic, metabolic, among others, has an immediate clinical application in the diagnosis, monitoring and treatment of these diseases.

The Human Proteome Organization (HUPO, www.hupo.org) was established in 2001 to promote greater awareness of the importance of proteomics and opportunities in the diagnosis, prognosis and treatment of diseases. It has subsequently formed several groups: HPPP (Human Plasma Proteome Project), HLPP (Human Liver Proteome Project), PSI (Proteome Standards Initiative), HBPP (Human Brain Proteome Project) and MRPP (Mouse and Rat Proteome Project). One of the main goals of proteomics is the identification of markers of disease. One approach has been to compare the expression of proteins in normal and diseased tissues to identify proteins that are expressed abnormally and may represent new markers. Another strategy is the analysis of proteins in segregated cell lines and primary cultures and finally, obtaining serum protein profiles.

Proteomic studies are gaining a higher profile in recent years, mainly in its application to human pathology. To this end a large number of studies are being conducted in human plasma, tissues and various body fluids.

© 2008 AEBM, AEFA y SEQC. Published by Elsevier España, S.L. All rights reserved.

Introducción

La Proteómica es, con la Genómica, una de las nuevas disciplinas que más interés está adquiriendo en los últimos años. El auge de la Proteómica se debe a 2 razones fundamentales: en primer lugar, a que los grandes proyectos de secuenciación de genomas han generado una enorme cantidad de información y, en segundo lugar, a que la necesidad de descifrar esta ingente cantidad de información genómica ha estimulado considerablemente el estudio directo y a gran escala de los productos codificados por los genes, es decir, las proteínas. Simultáneamente, se ha estimulado el desarrollo de nuevas herramientas bioinformáticas para analizar y procesar esta información¹.

Las herramientas disponibles (principalmente el análisis de aminoácidos y la secuenciación de Edman) sólo permitían el análisis de proteínas de una forma poco sensible y excesivamente lenta y costosa. El reciente desarrollo de la Proteómica ha ido parejo al de las técnicas de ionización suave y nuevos analizadores o mejoras, como TOF (*time of flight* 'tiempo de vuelo'), cuadrupolo (Q) y trampa iónica (ION TRAP), que permiten la identificación y la secuenciación de proteínas con unos niveles de rapidez, sensibilidad y versatilidad sin precedentes¹.

El término proteoma se utiliza para describir el conjunto de proteínas que se expresan a partir de un genoma. El proteoma es un elemento altamente dinámico cuyos componentes varían en un organismo, un tejido, una célula o un compartimento subcelular dados como consecuencia de cambios en su entorno, situaciones de estrés, administración de fármacos, efectores o señales bioquímicas o su estado fisiológico o patológico. Estos factores incrementan de forma considerable la complejidad del proteoma como consecuencia de la activación o la supresión de la expresión de genes, las alteraciones en la intrincada pauta de interacciones intracelulares entre las proteínas o los cambios en sus modificaciones postraduccionales^{1,2}.

Si bien el término proteoma se aplica indiscriminadamente a cualquier conjunto de proteínas, su acepción más precisa es la que se refiere al conjunto de proteínas que componen una célula o un tipo celular dado en unas

condiciones determinadas. El proteoma de una célula es la expresión de su fenotipo característico, y las diferencias entre un tipo celular y otro o entre un mismo tipo celular en diferentes situaciones se pueden, con cierta propiedad, asignar a los cambios que tienen lugar en la expresión y la funcionalidad de sus proteínas¹⁻³.

La Proteómica, entendida como estudio del proteoma, supone una aproximación metodológica a las biociencias, que se perfila como una potente herramienta, tanto desde el punto de vista cualitativo como cuantitativo, para el estudio de aspectos fisiopatológicos en los seres humanos. Esto, unido al continuo y progresivo avance biotecnológico, ha contribuido a la aparición de abundante bibliografía sobre la utilización de técnicas proteómicas en el abordaje de algunas enfermedades¹⁻⁴ y su presencia en el laboratorio clínico^{5,6}.

Así, se pretende explicar el considerable volumen de proyectos de investigación generados en Proteómica y la necesidad de caracterizar esta investigación en términos de cuáles son las técnicas proteómicas utilizadas, las enfermedades estudiadas y el tipo de uso que se propone. Asimismo, se puede observar cuál es la tendencia y si existen en la actualidad aplicaciones de la Proteómica en la asistencia sanitaria hospitalaria e identificar posibles aplicaciones a corto o a medio plazo.

Según la tendencia observada, es previsible que el interés por la Proteómica en Biomedicina continúe durante muchos años y que siga siendo objeto de abundante investigación. El conocimiento aportado por esta investigación puede llevar a la identificación de nuevas moléculas de naturaleza amino-proteica de interés clínico⁷. La validación y el estudio de la utilidad de estas moléculas como biomarcadores para el uso asistencial⁸ pueden generar una futura demanda de actividad en el medio hospitalario.

Esto podría estar relacionado con la efectividad y la relación coste-efectividad del uso de técnicas proteómicas o bien de técnicas convencionales para la determinación de proteínas (como el enzimo-inmunoanálisis o la cromatografía líquida [LC]), todo ello en indicaciones concretas, como el diagnóstico, la predicción del pronóstico o las indicaciones de fármacos⁹.

Proteómica, Genómica y Metabolómica

Mientras el genoma de un organismo es constante en la totalidad de sus células, excepto mosaicismos y mutaciones selectivas, el proteoma supone un concepto mucho más variable¹⁰. Este dinamismo no sólo se observa en diferentes sistemas, tejidos e incluso células, sino en estos mismos cuando son receptores de distintas condiciones internas o externas. Todo ello se traduce en un proteoma cambiante, así como una gran cantidad de posibles interacciones entre las proteínas que lo componen. Si a esto se añade la ausencia de una metodología de amplificación proteica, puede intuirse que, a diferencia del genoma, el análisis completo de toda la realidad de proteomas es una tarea difícil^{11,12}.

Esta complejidad hace que el estudio del proteoma plantee retos importantes. Para dar respuesta a estos retos, se precisan continuos desarrollos y mejoras en las técnicas de análisis de proteínas, como se comentará en el siguiente apartado.

Las proteínas actúan dentro de otra red bioquímica mucho más amplia y compleja, que es el metabolismo de las células. El estudio general de las redes metabólicas se comienza a conocer como Metabolómica. Genoma, proteoma y metaboloma forman un entramado, cuyo estudio supone un salto cualitativo determinante en las ciencias biomédicas. La combinación de Proteómica con Genética, Bioquímica o Biofísica ha supuesto el descubrimiento de muchas proteínas y de la función de éstas¹³.

El gran número de proyectos orientados a estudiar proteomas de forma sistemática ha conducido a describir la Proteómica como la ciencia o el conjunto de tecnologías que estudian el proteoma. Algunos autores prefieren definir la Proteómica como la parte de la Genómica Funcional que centra su atención en el estudio directo de las proteínas.

Tipos de estudios en Proteómica

La Proteómica se ha aplicado fundamentalmente a 3 tipos de estudios¹:

- La caracterización sistemática de proteínas, con el objetivo de abordar la identificación a gran escala de los componentes de un proteoma.
La identificación a gran escala de los componentes de un proteoma permite la catalogación de éstos y la construcción de bases de datos, que suelen organizarse sobre la base de la imagen del proteoma que resulta de su separación mediante electroforesis bidimensional. Estas bases de datos tienen una enorme utilidad para los proyectos de investigación que necesiten un conocimiento detallado de esos proteomas.
- La identificación de los componentes del proteoma que presentan cambios en sus niveles de expresión a consecuencia de alteraciones fisiopatológicas o inducidas por agentes externos (Proteómica de Expresión Diferencial). Estos proyectos abordan la identificación de proteínas que presentan alteraciones en sus niveles de expresión a consecuencia de cambios en su entorno, situaciones de estrés, administración de fármacos, efectores o señales bioquímicas o su estado fisiológico o

patológico, lo que permite así determinar cuáles son las proteínas que intervienen en esos procesos o identificar proteínas con cambios de expresión que resulten característicos del proceso (identificación de marcadores diagnósticos). Este tipo de estudios puede también llevarse a cabo mediante la tecnología de los *chips* de ADN, aunque, en la práctica, se trata de 2 aproximaciones complementarias. Los niveles de expresión del ARN mensajero no siempre reflejan adecuadamente el nivel de expresión real de las proteínas correspondientes, por lo que la aproximación proteómica permite obtener resultados más concluyentes desde un punto de vista tanto fisiológico como farmacológico. Sin embargo, la tecnología de los *chips* de ADN permite concentrar el análisis en el nivel de expresión de determinados genes con una sensibilidad superior.

- La caracterización de las interacciones subcelulares existentes entre las proteínas y la determinación de los componentes de complejos macromoleculares (Proteómica de Mapa Celular).

Lo que se pretende estudiar es la función de las proteínas y caracterizar las interacciones que presentan en el interior de la célula. Estos proyectos responden a la noción cada vez más extendida de que las proteínas no actúan de forma aislada, sino a través de grandes complejos macromoleculares, que agrupan conjuntos de factores que realizan las diferentes etapas de un mismo proceso celular, de forma que su proximidad aumenta la eficiencia de cada etapa. Muchas veces, la identificación de los componentes de un complejo macromolecular o de los factores que interactúan con una proteína dada resulta el camino más rápido para determinar su funcionalidad. Mediante este tipo de proyectos, aplicados de forma sistemática, los investigadores pretenden construir un mapa físico de las interacciones existentes entre las proteínas celulares.

Tecnología en Proteómica

La Proteómica se basa en el uso sinérgico de técnicas y procedimientos quimicofísicos para el análisis global y la separación de los componentes del proteoma, matrices de proteínas para el estudio de las cualidades funcionales de las proteínas y sus interacciones, y herramientas informáticas para el procesado y la interpretación de los datos.

Procedimientos quimicofísicos y tratamiento de resultados

Análisis preespectrométricos

Incluyen sistemas electroforéticos y cromatográficos: la electroforesis en geles de poliacrilamida con dodecil sulfato sódico (SDS-PAGE) de 1 y 2 dimensiones, la electroforesis capilar con acoplamiento a espectrometría de masas (EM o MS) y los procedimientos cromatográficos para purificación de proteínas, incluyendo la cromatografía líquida de alta resolución (HPLC) de fase normal, fase reversa, exclusión molecular, intercambio iónico y basada en bioafinidad.

La metodología más empleada es la SDS-PAGE, que permite la separación de las proteínas de acuerdo con su peso molecular. Esta metodología se suele combinar con la

separación previa de las proteínas sobre la base de su punto isoelectrico (isoelectroenfoque). De este modo, al separar primero en función de sus propiedades eléctricas y después por su tamaño, es posible aislar prácticamente todas las proteínas de un proteoma en una matriz bidimensional (sistema de electroforesis bidimensional o 2-DE [fig. 1]). El resultado es una serie de manchas o *spots*, cada una de las cuales corresponde a una proteína. Con esta técnica se pueden separar varios miles de proteínas en un solo gel.

Teóricamente, esta técnica serviría para estimar el número de proteínas expresadas en la muestra en un momento dado, pero realmente el número de *spots* obtenidos sólo refleja un porcentaje del total de proteínas expresadas. La extracción proteica a partir del material de partida no es total¹⁵.

A pesar de haberse introducido numerosas mejoras en la técnica de 2-DE, ésta aún presenta algunas limitaciones: solapamiento de proteínas cuando sus cargas/pesos moleculares son muy próximos, sensibilidad limitada, resolución deficiente para proteínas muy grandes o muy pequeñas y alto consumo de tiempo. Para procedimientos automatizados, la reproducibilidad es alta. Sin embargo, para técnicas manuales la reproducibilidad es media y en ocasiones no proporciona el objetivo deseado debido a las dificultades en la preparación de la muestra desde el material de partida hasta que puede cargarse en el gel^{16,17}.

Cuando interesa el análisis cuantitativo de perfiles de expresión de proteínas, tanto el marcaje por fluorescencia (sistema DIGE [differential in gel electrophoresis]) como la

automatización resultan obligados: se utilizan escáneres para la obtención de imágenes de fluorescencia de geles 2-DE, un robot para el cortado de manchas seleccionadas de geles 2-DE y un *software* especializado para el análisis cuantitativo de diferencias.

Otra técnica basada en la separación de las proteínas, la LC, está cobrando cada vez más importancia por su utilidad cuando se quieren separar proteínas pequeñas, y se utiliza fundamentalmente para el aislamiento de péptidos. La LC es un método físico de separación basado en la distribución de los componentes de una mezcla entre 2 fases inmiscibles: una fija o estacionaria y otra móvil. La fase móvil es un líquido que fluye a través de una columna que contiene la fase estacionaria. La LC clásica se lleva a cabo en una columna generalmente de vidrio, que está rellena con la fase estacionaria. La fase estacionaria puede ser un sólido de diferentes propiedades químicas; en función de éstas, resultan diferentes tipos de cromatografía (véase más adelante). La fase móvil puede ser un solvente puro o una mezcla de solventes. Tras depositar la muestra en la parte superior, se hace fluir la fase móvil a través de la columna por efecto de la gravedad. Con el objeto de aumentar la eficiencia en las separaciones, el tamaño de las partículas de la fase estacionaria se fue disminuyendo hasta el tamaño de los micrones, lo que generó la necesidad de utilizar altas presiones para lograr que fluyera la fase móvil. De esta manera, nació la técnica de HPLC, que, basándose en diferentes tipos de interacciones químicas entre las sustancias analizadas y la columna cromatográfica, permite separar los componentes de una mezcla¹⁴. Los tipos de HPLC normalmente utilizados son los de fase normal, fase reversa, de exclusión molecular, de intercambio iónico y basada en bioafinidad. Las HPLC de fase normal y reversa separan por polaridad; la de exclusión molecular, por tamaño; la de intercambio iónico, por polaridad, y la de bioafinidad, por la diferente capacidad de las sustancias biológicamente activas de formar complejos estables, específicos y reversibles. Un desarrollo de reciente aparición, altamente ventajoso, es la combinación de la HPLC de intercambio con la HPLC de fase reversa.

Otro desarrollo también reciente en la escala de trabajo de la HPLC es la operatividad de sistemas de micro-HPLC y HPLC con nanoflujo (nano-HPLC), que exhiben sensibilidades nunca observadas. La utilización de nano-HPLC permite la separación de péptidos antes de introducirlos hacia la fuente de ionización así como la eliminación de pequeñas cantidades de contaminantes (sales, detergentes, etc.) que interfieren con el análisis.

Análisis espectrométricos

Incluyen los sistemas de EM o MS, que hacen uso de procedimientos de ionización, como ESI (*electrospray ionization*), MALDI (*matrix asisted laser desorption/ionization*) o SELDI (*surface-enhanced laser desorption/ionization*), y detectores/analizadores de masa tipo Q lineal, Q en versión ION TRAP, TOF o híbridos.

Un sistema de MS realiza básicamente 3 funciones:

- ionización de la muestra;
- separación de fragmentos o iones, y
- detección/análisis de los fragmentos o iones.

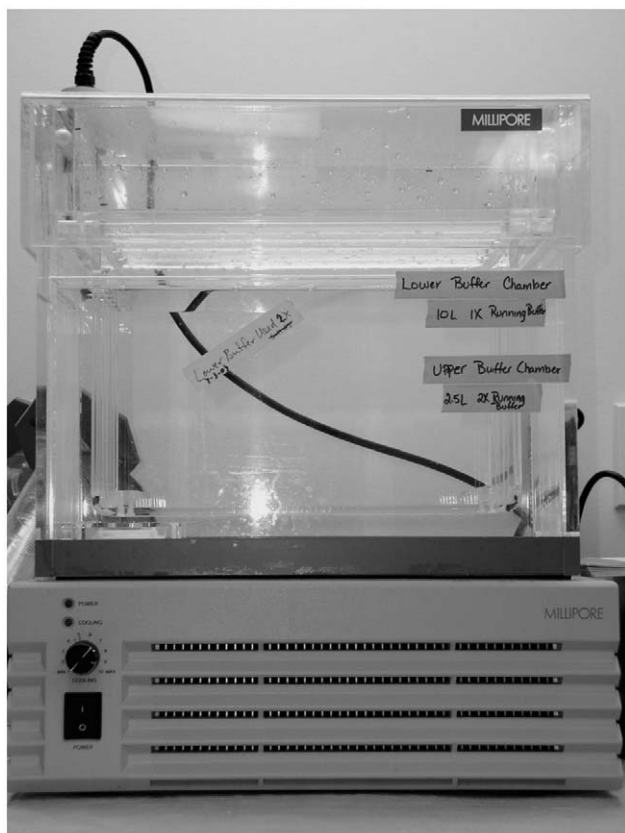


Figura 1 Sistema de electroforesis bidimensional.

Para realizar estas funciones, el espectrómetro de masas cuenta con las siguientes partes: una fuente de ionización, un analizador de masa/carga y un detector¹⁸.

Producir iones en fase gaseosa es fácil para compuestos volátiles de bajo peso molecular; sin embargo, la aplicación de la MS en el campo de las proteínas es relativamente reciente. Esto se debe a la dificultad de obtener iones de macromoléculas en fase gaseosa, dado que los péptidos y las proteínas son compuestos de baja volatilidad y alto peso molecular. Por esto, es necesario el uso de técnicas especiales de ionización. Entre las distintas variantes de estas fuentes de ionización, destacan por su aplicabilidad a los péptidos y a las proteínas la ionización por ESI y la desorción/ionización mediante láser según variantes MALDI o SELDI. Cabe reseñar que el original fue MALDI (hoy empleado para Proteómica de los tejidos con el nombre de MALDI Imaging)^{23,24} y SELDI es una adaptación para la Proteómica clínica de fluidos.

Las proteínas separadas por los métodos del apartado “Análisis preespectrométricos” pueden analizarse mediante MS según los tratamientos siguientes:

En el más común, se *recortan* las proteínas del gel bidimensional donde se han aislado, se digieren con una proteasa para producir un conjunto de péptidos y se analizan subsiguientemente estos péptidos mediante MS tipo MALDI-TOF (fig. 2). Esta técnica es particularmente útil para obtener el espectro de masas del conjunto de péptidos, también denominada “huella peptídica” (*peptide fingerprinting*).

Alternativamente, los péptidos producidos tras un proceso de eliminación de contaminantes se analizan mediante MS tipo ESI en combinación con un triple Q o una ION TRAP. Estas técnicas permiten el análisis en tiempo real de péptidos individuales presentes en la mezcla, sin necesidad de separarlos del resto, que consiste en la fragmentación del péptido y la generación de un espectro de masas, también llamado espectro de fragmentación o espectrofotometría de masas en tandem (MS/MS).

Las huellas peptídicas son características de cada una de las proteínas, y permiten identificarlas una a una en la base de datos mediante la utilización de técnicas bioinformáticas, que constituyen las herramientas de la tercera categoría. De la misma manera, los espectros MS/MS son

también característicos de cada uno de los péptidos, y permiten su identificación *in silico* en la base de datos. Ninguna de las 2 técnicas (huella peptídica o fragmentación) es de aplicabilidad universal, y cada una de ellas tiene sus ventajas y sus inconvenientes. Sin embargo, la disponibilidad sinérgica de ambas técnicas hace de la MS una formidable herramienta para la identificación sistemática de proteínas.

En la MS, se están comenzando a usar nuevos aparatos, que se basan principalmente en la combinación *híbrida* de técnicas ya conocidas. Como por ejemplo, se ha indicado en párrafos anteriores, los aparatos basados en la combinación de un doble Q y un TOF (qQ-TOF) mejoran sensiblemente la resolución de los espectros de fragmentación, y la combinación de la ionización MALDI con el analizador qQ-TOF o la ION TRAP permite fragmentar péptidos directamente una vez obtenida la huella peptídica.

La MS denominada MALDI-TOF/TOF conjuga todas las ventajas de la espectrometría MALDI-TOF convencional con la capacidad de producir espectros de fragmentación de péptidos de forma rápida y consistente. Esta técnica está comenzando a utilizarse para la identificación a gran escala de proteínas, pero su alto coste en el mercado ha limitado su disponibilidad a pocos laboratorios.

Otras aproximaciones de la MS se basan en el acoplamiento de un sistema de LC bidimensional con un analizador de ION TRAP o del tipo qQ-TOF. Este procedimiento, capaz de identificar millares de péptidos en un solo experimento, permite analizar de forma automática los componentes de mezclas muy complejas de péptidos, tales como las producidas por la digestión directa de un proteoma. Esta técnica (también denominada *shotgun proteomics*) se ha aplicado a la identificación directa de componentes de complejos macromoleculares e incluso al análisis de proteomas enteros sin necesidad de separar sus componentes mediante electroforesis.

Otra variante de esta metodología permite realizar estudios de expresión diferencial de proteínas en el proteoma sin necesidad de realizar 2-DE. Esta técnica, conocida como ICAT (*isotope-coded affinity tag* ‘marcaje isotópico diferencial’), se basa en el uso de reactivos de proteínas en forma de isótopos estables, que permiten determinar la proporción relativa de los péptidos derivados de los proteomas por comparar. Al utilizar programas informáticos desarrollados para este fin, este procedimiento permite la identificación automática y simultánea de cambios relativos en las concentraciones de proteínas y de la naturaleza de las proteínas que experimentan expresión diferencial.

Finalmente, estas metodologías, en combinación con procedimientos de purificación selectiva por afinidad, han permitido el análisis de péptidos modificados, tales como fosfopéptidos, a partir de proteomas completos, lo que ha dado lugar al concepto de *fosfoproteoma*.

Posibilidades múltiples de trabajo según procedimientos analíticos combinados

Una proteína purificada puede llevarse a MALDI-TOF o a ESI-MS y hallar su masa molecular, se puede digerir y analizar su huella peptídica o se puede analizar uno sólo o varios péptidos por fragmentación hasta identificar positivamente

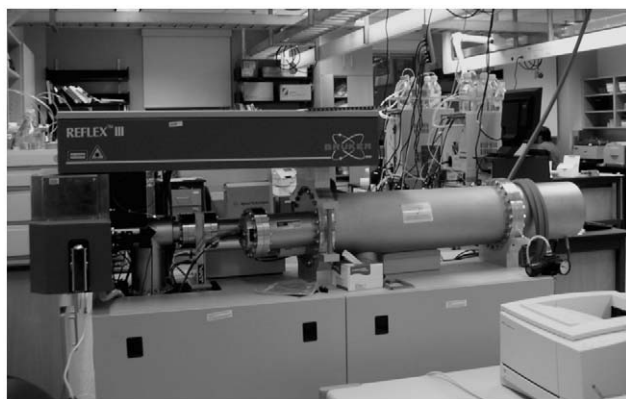


Figura 2 Sistema Matrix Asisted Laser Desorption Ionization-time of flight/espectrometría de masas (MALDI-TOF/MS) de Bruker®.

e incluso elucidar sus modificaciones postraslacionales. A efectos de cuantificación, es de especial interés el sistema AQUA[®], con el concurso de péptidos sintéticos marcados con isótopos estables.

Una mezcla de proteínas de cultivo, homogeneizado de tejido o cualquier fluido pueden llevarse a un gel-2D y separar sus proteínas. Ya puede actuarse sobre bases de datos de geles para la identificación o sobre los *spots* o las relaciones de los fluorocromos en el sistema DIGE para la cuantificación. Asimismo —como se ha dicho anteriormente— se pueden recortar *spots*, digerirlos y llevarlos a MALDI-TOF para la huella peptídica o a ESI-MS/MS para la secuenciación, por lo que ambas técnicas pueden llegar a la identificación positiva.

Para procesar mezclas extraordinariamente complejas de péptidos, se utilizan sistemas ICAT, con reactivos ligeros y pesados, de las unidades cisteína para subsiguientemente, tras digestión y separación por afinidad, proceder a su cuantificación por MS. También existe la posibilidad de que estas muestras complejas se puedan digerir sin separar y luego intentar resolverlas por cromatografía en una o 2 dimensiones (intercambio y fase reversa), y aislar péptidos para secuenciarlos en micro-ESI-MS/MS y, si es posible, con híbridos que permiten resolver péptidos múltiplemente cargados.

Tratamiento de espectros y bases de datos disponibles

De hecho, las estrategias descritas en apartados anteriores son asistidas de *softwares* que enfrentan los espectros obtenidos a bases de datos de gel, digestión, secuencia, posmodificaciones, etc. con motores de búsqueda muy potentes. Otras combinan la información obtenida del paso de ADN a posibles proteínas con digestiones simuladas *in silico* susceptibles de confrontación con los espectros obtenidos.

Entre los *softwares* de proteómica clínica más utilizados se encuentran los de *ClinProt Tools* para el estudio de perfiles y el *DeCyder* para análisis cuantitativo de diferencias.

Matrices de proteínas

Las matrices, *arrays* o *biochips* de proteínas (fig. 3) permiten la detección, la caracterización y la cuantificación proteica, así como el estudio de las cualidades funcionales de las proteínas y sus interacciones, tanto entre ellas como con moléculas de ADN o lípidos. Un ejemplo son los *arrays* de anticuerpos, que permiten estudiar la expresión de un elevado número de proteínas en una sola matriz. Esta técnica consiste en fijar anticuerpos a una superficie sólida, de manera ordenada, y localizada a modo de puntos (*spots*) distribuidos en 2 ejes. A continuación se pone en contacto la muestra objeto de estudio (p. ej., un lisado celular) con la matriz, con lo que se consigue el reconocimiento y la unión específica proteína-anticuerpo. De modo análogo, se puede unir al *array* otro tipo de moléculas^{19,20}.

La complejidad y la diversidad estructural proteicas han hecho que el desarrollo de los *microarrays* de proteínas haya sido técnicamente complicado.

Al contrario que los ácidos nucleicos, las proteínas no tienen una estructura homogénea ni un patrón de unión específico, sino que cada proteína posee unas características

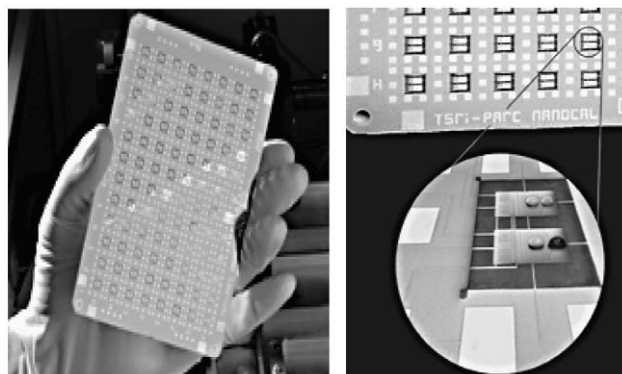


Figura 3 Sistema de matrices de proteínas. Cada celda detectora en el *array* consta de 2 regiones detectoras adyacentes para muestra y referencia.

bioquímicas particulares. En cuanto a los procesos de amplificación de la muestra, no existe una técnica equivalente a la reacción en cadena de la polimerasa en Genómica que sea capaz de amplificar o multiplicar la cantidad de proteína existente en una muestra.

Por otra parte, el ADN posee una carga negativa que puede utilizarse para inmovilizar la molécula sobre la superficie del *array* mediante fuerzas electrostáticas. Por el contrario, la carga de las proteínas es muy variable y, por este motivo, se han encontrado dificultades en la estandarización de materiales de soporte que sean adecuados para cada tipo de *microarray* de proteínas.

Hoy día, se cuenta con una nueva generación química de superficies de membrana y de vidrio, que no requieren del empleo de agentes bloqueantes para eliminar el ruido de fondo y que a la vez previenen el contacto directo de la proteína con la superficie. Con todo, la tecnología de *microarrays* de proteínas aún se encuentra ante dificultades técnicas en cuanto a la adquisición y a la unión estable de proteínas a superficies donde puedan interactuar con otras proteínas o ligandos y detectarse tal interacción^{21,22}.

Aplicaciones de la Bioinformática

Otra de las ciencias aplicadas al estudio de la Proteómica es la Bioinformática, que se podría definir como el conjunto de herramientas para el procesamiento y la interpretación de los datos obtenidos con cualquiera de las técnicas anteriormente expuestas u otras. Así, las herramientas bioinformáticas permiten identificar una proteína en una base de datos a partir de su huella peptídica (que es específica para cada proteína) o de su espectro MS/MS (que también es específico de cada uno de los péptidos que forman parte de una proteína)²³.

Pese a los avances en proteómica, muchas propiedades de las proteínas, sobre todo sus interacciones y sus modificaciones postraduccionales, no pueden, en general, predecirse a partir de la secuencia de ADN. En un proyecto reciente de particular interés, se ha efectuado un estudio a gran escala de las interacciones entre proteínas en levadura, que ha dado lugar a la construcción de un mapa de interacciones¹. La evidente complejidad de este mapa indica que la

funcionalidad de las proteínas no debe considerarse de forma aislada, sino dentro de una red global de componentes que interactúan mutuamente.

Se están desarrollando técnicas bioinformáticas capaces de interpretar el significado de esta compleja red de interacciones y, por otra parte, esta metodología no puede todavía aplicarse a células eucariotas, es decir, células de los organismos superiores. Sin embargo, este trabajo pionero indica que en un futuro próximo la Proteómica nos permitirá analizar las propiedades de sistemas complejos de proteínas desde una perspectiva global e integradora. Resultarán apasionantes los resultados que se deriven de aplicar esta estrategia experimental al estudio de procesos celulares de alta complejidad, tales como las cascadas de transducción de señales, la diferenciación celular, la degeneración neuronal, la apoptosis o el cáncer.

Una de las asignaturas pendientes de la Proteómica es que, en su concepción actual, no puede aplicarse de forma efectiva al estudio de componentes de microorganismos o especies de interés en Biotecnología o Tecnología de Alimentos, cuyas proteínas no están suficientemente representadas en las bases de datos. El desarrollo de nuevos algoritmos informáticos capaces de interpretar de forma inteligente la información generada por la MS permitirá en un futuro próximo aplicar la Proteómica al estudio de estas especies.

Estos nuevos desarrollos bioinformáticos podrían también aplicarse al estudio sistemático de modificaciones postraduccionales. Estas técnicas, en combinación con los métodos de LC bidimensional, podrían abrir el camino al estudio de modificaciones postraduccionales a gran escala y de forma dinámica, y a la manera como éstas modulan, desde una perspectiva global, el control de la inmensa mayoría de los procesos celulares.

Existe un gran número de proteínas y de nuevos genes descubiertos, de los que no se dispone aún de información acerca de su función. Los métodos de predicción de las funciones proteicas constituyen un apoyo teórico importante en el estudio de estas macromoléculas, tanto por la información que proporcionan como por la utilidad que pueden tener en el diseño de péptidos específicos y la prospección de nuevas funciones. Estos métodos de predicción de funciones proteicas son probabilísticos, por lo que dependen de la cantidad de datos o la información disponible en un principio; en otras palabras, del tamaño de la muestra por analizar. Por otra parte, el procedimiento que suele utilizarse para realizar predicciones es la comparación con datos de naturaleza semejante. Aquí también es, por tanto, importante el tamaño de la población de referencia, que se trata del volumen de datos con los que se va a comparar la muestra. A este respecto, las poblaciones de referencia son las actuales bases de datos disponibles sobre secuencias moleculares, estructuras secundarias y terciarias, homología total o parcial de genes y proteínas y un amplio etcétera²⁴.

La mínima información que ha de tenerse de una proteína para poder predecir su función es un fragmento representativo de su secuencia aminoacídica. De esta forma, al menos, se puede llevar a cabo un análisis a nivel de *secuencia*. Si además se dispone de la información estructural, del conocimiento de su naturaleza iónica, de los efectos mutacionales sobre su función, de la localización de la

proteína en la célula o en el organismo o de las interacciones con otras proteínas (es decir, de su *contexto*), podrán abordarse también otros aspectos analíticos que faciliten el objetivo de la predicción, que es la determinación precisa de su función²⁵.

Proteómica en estudios de biomedicina

La perspectiva ofrecida por la Proteómica se ha utilizado en estudios de investigación de diferentes áreas de la Medicina, incluida la Biomedicina. Estos estudios podrían clasificarse de distintas formas: en función del tipo de muestra empleada, de la enfermedad o el tipo de enfermedades que abordan, de la técnica o las técnicas utilizadas, del uso o la aplicación de éstas, etc.

Esto se orientará en torno al eje de clasificación según el tipo de muestra empleado. Sobre la base de ella, cabe destacar la investigación con líneas celulares, con tejidos sólidos y la Proteómica de fluidos.

Las líneas celulares son variantes derivadas de las células cancerosas o las células normales inmortalizadas mediante la inhibición de los mecanismos celulares causantes de la interrupción del crecimiento. A diferencia de la mayoría de las células aisladas directamente de los organismos, estas células son capaces de dividirse indefinidamente cuando se mantienen cultivadas en el laboratorio.

En la investigación con líneas celulares, las técnicas proteómicas se han utilizado para caracterizar los perfiles de expresión proteica en diferentes tipos de estas líneas como base para futuros experimentos comparativos^{26,27}.

Estos estudios han mostrado que el número de proteínas identificadas en cada proteoma varía de forma importante, lo que probablemente sea el reflejo de la existencia de diferencias en el rango de expresión de las distintas proteínas en cada línea celular²⁸.

Las líneas celulares tumorales humanas pueden constituir modelos válidos para el estudio del cáncer. De este modo, los estudios que comparan los proteomas de diversas líneas tumorales humanas (o células obtenidas directamente de un tumor) con sus correspondientes líneas celulares normales (o tejido sano) comienzan a ser habituales para identificar marcadores de enfermedad o biomarcadores que permitan la detección precoz, la clasificación y la predicción del pronóstico de los tumores así como para proponer nuevas dianas potenciales o dianas terapéuticas para mejorar su tratamiento.

Por consiguiente, se pueden realizar estudios proteómicos dirigidos a evaluar el potencial maligno de un tumor (potencial metastásico según exprese una proteína específica o no), la quimiosensibilidad o la quimiorresistencia de éste, etc²⁹. El uso de la Proteómica para el diseño, la indicación y la predicción de respuesta a un fármaco o a fármacos sería también extrapolable a otras enfermedades; es lo que conocemos como farmacoproteómica^{30,31}.

Por otra parte, las líneas celulares se han utilizado clásicamente para conocer los mecanismos adaptativos a distintos tipos de estrés. El desarrollo de la Proteómica ha encontrado una clara aplicación en estos estudios, ya que estas líneas celulares permiten la identificación de nuevas proteínas no relacionadas hasta el momento con el establecimiento de un fenotipo resistente a las condiciones que provocan el estrés celular³².

Además de los estudios en líneas celulares humanas descritos, las técnicas proteómicas se han aplicado recientemente en el análisis de tejidos sólidos y de fluidos para el estudio de situaciones fisiológicas (durante el desarrollo, en distintos estados metabólicos, frente a diversas condiciones ambientales) o patológicas (cáncer, autoinmunidad, infecciones, etc.)^{3,4}. Estos estudios con tejidos o fluidos se han empleado con objetivos similares a los estudios con líneas celulares (identificación de biomarcadores, búsqueda de dianas terapéuticas, farmacoproteómica, etc.).

La utilización de tejidos sólidos en estudios de Proteómica presenta, lógicamente, el inconveniente de la mayor dificultad de acceso a las muestras.

No obstante, la información obtenida en éstos puede ser determinante y, a veces, no existe mejor alternativa a la utilización de muestras tisulares sólidas³³. Tal puede ser el caso del estudio de mutaciones selectivas que alteren el proteoma de un tejido a nivel local, que no son detectables en otros compartimentos del organismo.

Además, los estudios realizados con técnicas de Proteómica que emplean muestras tisulares pueden completarse con la utilización de otros métodos. Tal sería el caso del empleo de *western blot* para estudiar la presencia de autoanticuerpos en suero a partir del proteoma obtenido de células tumorales. La caracterización de un conjunto determinado de autoanticuerpos, característico de cada tumor, puede ser de gran valor, ya que podría ayudar al diagnóstico precoz en pacientes de riesgo^{34,35}.

En cuanto a la Proteómica de fluidos, cabe reseñar que ha supuesto un hito considerable en el ámbito de la investigación clínica por la mayor facilidad de obtención y procesamiento de las muestras. De entre ellas, cabe destacar el plasma y el suero por su accesibilidad. La aplicación de técnicas proteómicas a estas muestras se ha utilizado para el estudio de enfermedades muy diversas (neoplásicas, reumáticas, endocrinológicas, toxicológicas, etc.)^{3,4,36-48}.

Otro ejemplo de Proteómica de fluidos, aunque la muestra es de más difícil obtención, es el análisis del proteoma del líquido cefalorraquídeo. Éste se ha empleado recientemente en el estudio de las bases etiopatogénicas y la identificación de biomarcadores de enfermedades neurológicas, como trastornos neuropsiquiátricos, tumores cerebrales y síndromes dolorosos lumbares^{41,42}.

Conclusión

El objetivo de la Proteómica en clínica es *comparar, a nivel molecular, el estado proteico celular en situaciones normales y patológicas*. Esta comparación se concreta, en unos y otros casos, en identificar las proteínas que hay presentes, determinar su proporción, dilucidar en qué procesos biológicos participan y averiguar cómo interactúan entre sí. Las aplicaciones biomédicas resultan inminentes en: el diseño de fármacos que actúen directa y selectivamente sobre las proteínas implicadas en una enfermedad.

Bibliografía

- Vázquez-Cobos J. Presente y futuro de la Proteómica. Disponible en: <http://www2.cbm.uam.es/jvazquez/PDFs/Proteomica.pdf>
- Zapico Muniz E, Mora Bruges J, Blanco Vaca F. Diagnóstico precoz del cáncer mediante análisis proteómicos del suero: ¿ficción o realidad? *Med Clin (Barc)*. 2005;124:181-5.
- González-Buitrago JM. Multiplexed testing in autoimmunity laboratory. *Clin Chem Lab Med*. 2006;44:1169-74.
- González-Buitrago JM, Ferreira L, Lorenzo I. Urinary proteomics. *Clin Chim Acta*. 2007;375:49-56.
- Righetti PG, Castagna A, Antonucci F, Piubelli C, Cecconi D, Campostri N, et al. Proteome analysis in the clinical chemistry laboratory: Myth or reality?. *Clin Chim Acta*. 2005;357:123-39.
- Master SR. Diagnostic proteomics: Back to basics?. *Clin Chem*. 2005;51:1333-4.
- Dhingra V, Gupta M, Andacht T, Fu ZF. New frontiers in proteomics research: A perspective. *Int J Pharm*. 2005;299:1-18.
- Lindsay MA. Target discovery. *Nat Rev Drug Discov*. 2003;2:831-8.
- Posadas EM, Simpkins F, Liotta LA, MacDonald C, Kohn EC. Proteomic analysis for the early detection and rational treatment of cancer: Realistic hope?. *Ann Oncol*. 2005;16:16-22.
- Yan JX, Tonella L, Sánchez JC, Wilkins MR, Packer NH, Gooley AA, et al. The dictyostelium discoideum proteome-the SWISS-2DPAGE database of the multicellular aggregate (slug). *Electrophoresis*. 1997;18:491-7.
- Kahn P. From genome to proteome: Looking at a cell's proteins. *Science*. 1995;270:369-70.
- Nielsen J, Oliver S. The next wave in metabolome analysis. *Trends Biotechnol*. 2005;23:544-6.
- Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, et al. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*. 1995;16:1090-4.
- McDonald WH, Yates JR. Shotgun proteomics and biomarker discovery. *Dis Markers*. 2002;18:99-105.
- Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: An updated protocol and implications for a functional analysis of the genome. *Electrophoresis*. 1995;16:1034-59.
- Fuchs D, Winkelman I, Johnson IT, Mariman E, Wenzel U, Daniel H. Proteomics in nutrition research: Principles, technologies and applications. *Br J Nutr*. 2005;94:302-14.
- Gorg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics*. 2004;4:3665-85.
- Aebersold A, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003;422:198-207.
- Khandurina J, Guttman A. Microchip-based high-throughput screening analysis of combinatorial libraries. *Curr Opin Chem Biol*. 2002;6:359-66.
- Templin MF, Stoll D, Schwenk JM, Potz O, Kramer S, Joos TO. Protein microarrays: Promising tools for proteomic research. *Proteomics*. 2003;3:2155-66.
- López M, Mallorquín P, Vega M. Aplicaciones de los microarrays y biochips en salud humana. Madrid: Fundación Española para el Desarrollo de la Investigación en Genómica y Proteómica/ Fundación General de la Universidad Autónoma de Madrid; 2006. p 24-37. [consultado 2008] Disponible en: http://www.madrimas.org/biotecnologia/Seleccion/Downloads_GetFile.aspx?id=5000
- MacBeath G. Protein microarrays and proteomics. *Nat Genet*. 2002;32:526-32.
- Fiehn O, Weckwerth W. Deciphering metabolic networks. *Eur J Biochem*. 2003;270:579-88.
- Lilien RH, Farid H, Donald BR. Probabilistic disease classification of expression-dependent proteomic data from mass spectrometry of human serum. *J Comput Biol*. 2003;10:925-1046.

25. Droit A, Poirier GG, Hunter JM. Experimental and bioinformatics approaches for interrogating protein-protein interactions to determine protein function. *J Mol Endocrinol*. 2005;34:263–80.
26. Issaq HJ, Conrads TP, Prieto DA, Tirumalai R, Veenstra TD. SELDITOF MS for diagnostic proteomics. *Anal Chem*. 2003;75:148–55.
27. Chaurand P, Sanders ME, Jensen RA, Caprioli RM. Proteomics in diagnostic pathology: Profiling and imaging proteins directly in tissue sections. *Am J Pathol*. 2004;165:1057–68.
28. Moore GE, Minowada J. Historical progress and the future of human cell culture research. *Hum Cell*. 1992;5:313–33.
29. Le Naour F. Contribution of proteomics to tumor immunology. *Proteomics*. 2001;1:1295–302.
30. Lage H. Proteomics in cancer cell research: An analysis of therapy resistance. *Pathol Res Pract*. 2004;200:105–17.
31. Jain KK. Role of pharmacoproteomics in the development of personalized medicine. *Pharmacogenomics*. 2004;5:331–6.
32. Yan Y, Weaver VM, Blair IA. Analysis of protein expression during oxidative stress in breast epithelial cells using a stable isotope labelled proteome internal standard. *J Proteome Res*. 2005;4:2007–14.
33. Bhattacharya SH, Gal AA, Murray KK. Laser capture microdissection MALDI for direct analysis of archival tissue. *J Proteome Res*. 2003;2:95–8.
34. Caldwell RL, Caprioli RM. Tissue profiling by mass spectrometry. *Mol Cell Proteomics*. 2005;4:394–401.
35. Reyzer ML, Caprioli RM. MALDI mass spectrometry for direct tissue analysis: A new tool for biomarker discovery. *J Proteome Res*. 2005;4:1138–42.
36. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*. 2002;359:572–7.
37. Liotta LA, Petricoin EF. Serum peptidome for cancer detection: Spinning biologic trash into diagnostic gold. *J Clin Invest*. 2006;116:26–30.
38. Tampona M, Brescia V, Fontana A, Maggiolini P, Zucano A, Pansini N. Proteomic: New advances in the diagnosis of rheumatoid arthritis. *Clin Chim Acta*. 2005;357:219–25.
39. Petricoin EF, Rajapaske V, Herman EH, Arekani AM, Ross S, Johann D, et al. Toxicoproteomics: Serum proteomic pattern diagnostics for early detection of drug induced cardiac toxicities and cardioprotection. *Toxicol Pathol*. 2004;32:122–30.
40. Clynen E, De Loof A, Schoofs L. The use of peptidomics in endocrine research. *Gen Comp Endocrinol*. 2003;132:1–9.
41. Romeo MJ, Espina V, Lowenthal M, Espina B H, Petricoin EF, Liotta LA. CSF proteome: A protein repository for potential biomarker identification. *Expert Rev Proteomics*. 2005;2:57–70.
42. Zheng PP, Luider TM, Pieters R, Avezaat CJ, Van den Bent MJ, Sillevius Smitt PA, et al. Identification of tumor-related proteins by proteomic analysis of cerebrospinal fluid from patients with primary brain tumors. *J Neuropathol Exp Neurol*. 2003;62:855–62.
43. Wulfschuhle JD, Sgroi DC, Krutzsch H, McLean K, McGarvey K, Knowlton M, et al. Proteomics of human breast ductal carcinoma in situ. *Cancer Res*. 2002;62:6740–9.
44. Instituto Nacional de Proteómica-ProteoRed. [consultado 2008] Disponible en: <http://www.proteored.org/Default.asp>
45. Sociedad Española de Proteómica (SEProt). [consultado 2008] Disponible en: <http://www.cbm.uam.es/seprot/index.htm>
46. Morrison RS, Kinoshita Y, Johnson MD, Uo T, Ho JT, McBee JK, et al. Proteomic analysis in the neurosciences. *Mol Cell Proteomics*. 2002;1:553–60.
47. González-Buitrago JM, Ferreira L, Isidoro-García M, Sanz C, Lorente F, Davila I. Proteomic approaches for identifying new allergens and diagnosing allergic diseases. *Chim Acta*. 2007;385:21–7.
48. González-Buitrago JM. Proteómica clínica. Barcelona: SEQC; 2007.