

# PROTEÓMICA EN AMÉRICA LATINA: AVANCES Y DESAFÍOS

Gabriel Padrón, Patricia Cuervo

Laboratório de Pesquisa em Leishmaniose, Instituto Oswaldo Cruz, FIOCRUZ. Av. Brasil 4365, Manguinhos, 21045-900, Rio de Janeiro, RJ, Brazil. Phone: (55-21) 3865-8224; Fax: (55-21) 2290-0479. pcuervo@fiocruz.br

## RESUMEN

El impetuoso desarrollo tecnológico de las últimas décadas en la espectrometría de masas, la cromatografía líquida y la computación, ha convertido a la proteómica en una herramienta imprescindible para el estudio de los sistemas biológicos. En el presente artículo de opinión se ilustran algunas de las características de la proteómica y sus aplicaciones. Igualmente, se hace un análisis de su desarrollo en América Latina y sus perspectivas en la región.

## PROTEÓMICA

Hace apenas poco más de 20 años que el término “proteoma” fue acuñado por Marc Wilkins para denominar el concepto equivalente al genoma para proteínas, y que se define como todas las proteínas presentes en una célula o tejido en un momento dado y en unas condiciones específicas. En esta definición está implícito el carácter altamente dinámico del proteoma, a diferencia del genoma. No es posible hablar del proteoma de una célula si no se especifican las condiciones y el tiempo en que se determinan y cuantifican las proteínas presentes. Algunas proteínas pueden aparecer o desaparecer, o como generalmente ocurre, modificar la concentración en que están presentes al cambiar las condiciones (temperatura, pH, estrés de cualquier tipo, etc.), o simplemente al transcurrir un determinado tiempo en el ciclo de vida normal de una célula (Aebersold & Patterson, 2003). Por otra parte, se denomina “proteómica” al estudio del proteoma y comprende la identificación, cuantificación, localización, función e interacciones de las proteínas.

El desarrollo tecnológico ocurrido a partir de la segunda mitad de la década de los 90 ha provocado un cambio irreversible en la forma de enfrentar las investigaciones biológicas. Estas se desarrollaban de la forma tradicional, es decir, se hacían determinaciones basadas en alguna hipótesis previa, los datos obtenidos eran analizados y se definían nuevas investigaciones o hipótesis, se obtenían nuevos datos y así sucesivamente se iba acumulando la información y el conocimiento.

Pero desde mediados de los 90 es posible obtener información de interés biológico de forma masiva, lo que ha dado lugar a las llamadas “ómicas”: genómica, transcriptómica, proteómica y metabolómica. La necesidad de analizar este enorme flujo de información dio lugar también a la bioinformática. En la actualidad se hace necesaria la aplicación de las ómicas, junto a la bioinformática, para lograr avances acelerados en el conocimiento de los sistemas biológicos. A continuación, vamos a referirnos con más detalle a la proteómica.

Aunque existen algunas alternativas, las dos herramientas fundamentales de la proteómica son la cromatografía líquida de alta eficacia (HPLC) y la espectrometría de masas (MS) y su combinación (LC-MS). El procedimiento más generalizado consiste en extraer las proteínas celulares, digerirlas con una proteinasa, usualmente tripsina, y los péptidos resultantes son analizados y secuenciados por LC-MS. Las proteínas son identificadas y cuantificadas a partir de los espectros de masas de sus péptidos y con el auxilio de varios programas y bases de datos.

### Espectrometría de masas

Para estudiar un compuesto por MS es imprescindible que se encuentre en forma iónica. Por tanto, el primer evento que ocurre en un espectrómetro de masas es la ionización de los compuestos que serán analizados. El método de ionización más usado para compuestos orgánicos ha sido la ionización electrónica (EI), que consiste en bombardear la muestra con electrones con alta energía que la ceden para provocar la ionización. Sin embargo, para que esto ocurra, la muestra tiene que estar en estado gaseoso, por lo que la EI no puede ser usada con compuestos que no sean fácilmente volatilizables, como es el caso de las biomoléculas: péptidos, proteínas, carbohidratos, etc.

No fue hasta 1981 que dos investigadores ingleses, Michael Barber y Richard Bordoli, desarrollaron el método llamado Fast Atom Bombardment (FAB), que causaba simultáneamente la ionización y desorción de la muestra disuelta en una gota de glicerina y bombardeada con átomos de argón o xenón previamente acelerados. Se obtuvieron los espectros de péptidos y proteínas pequeñas como la insulina, y a

partir de los espectros de masas era posible la secuenciación de los péptidos, algo que hasta ese momento solo era posible a través de la secuenciación por Edman. Sin embargo, la eficiencia de ionización era muy baja, por lo que se necesitaban grandes cantidades de péptidos, y solo era posible estudiar proteínas recombinantes o naturales relativamente abundantes.

A mediados de los años 90, aparecieron las primeras fuentes de ionización que mejoraron considerablemente la eficiencia en biomoléculas: Electro spray (ESI) y MALDI (Matrix Assisted Laser Desorption Ionization). Estos dos métodos revolucionaron el análisis de péptidos, proteínas y carbohidratos. Sus creadores, John Fenn y Keuchi Tanaka respectivamente, recibieron el premio Nobel en 2002 (Karas & Hillenkamp, 1988; Tanaka et al., 1988; Fenn et al., 1989).

En el ESI, la muestra en solución es impulsada por una bomba (que puede ser un equipo de HPLC) y se hace pasar a través de un capilar en cuya punta se aplica un potencial de 3 o 4 KV. A la salida se forma un *spray* de gotas cargadas que contienen el solvente y la muestra. El solvente es evaporado y queda la muestra ionizada, que se introduce al analizador del espectrómetro de masas.

En el MALDI, la muestra se encuentra en estado sólido mezclada con una matriz adecuada, y es ionizada por un rayo láser en la región UV. La matriz es esencial, pues la acción directa del láser causa la destrucción de la muestra. Las principales matrices empleadas son ácidos orgánicos insaturados que absorben en el UV, como los ácidos sinapínico, dihidroxi benzoico o alfa-ciano-4-hidroxi-cinámico.

En ambos métodos, la ionización tiene lugar por captura o pérdida de protones, lo que produce iones positivos o negativos, respectivamente. La gran mayoría de los espectros de masas de péptidos y proteínas se obtienen en el modo positivo. En el MALDI, los iones generados son fundamentalmente monocargados, o sea, capturan un solo protón, pero en el Electro spray son multicargados. En péptidos y proteínas, la adición de protones ocurre en el extremo N-terminal y en las cadenas laterales de arginina, lisina e histidina.

En ambos casos solo se produce la ionización de los péptidos sin que ocurra fragmentación de la cadena peptídica. Por tanto, al analizar una mezcla de péptidos solo se observan los iones moleculares, y la única información que se obtiene (aunque muy valiosa) es la masa molecular de cada péptido. Por esta razón es preciso utilizar un procedimiento adicional para estimular la fragmentación de los péptidos y poder deducir la secuencia. Los dos métodos más conocidos son la fragmentación inducida por colisiones (CID) y la HCD (High Energy Collision Dissociation) (Wells & McLuckey 2005). La combinación de HPLC y espectrometría de masas se ha

convertido en la herramienta más poderosa para el análisis de proteínas. Permite estudiar mezclas complejas de proteínas, sin previa separación, e identificar y cuantificar varios miles de proteínas en un solo experimento (Mann et al., 2013). Algunas aplicaciones de la proteómica están resumidas en el Cuadro 1.

En el caso de levaduras es posible identificar y cuantificar más del 80 % del proteoma y en células de mamíferos se pueden identificar y cuantificar más de 7 000 proteínas, incluidas las de baja abundancia como las citosinas y factores de transcripción que están presentes en menos de 100 copias por célula (Picotti et al., 2009; Nagaraj et al., 2012). La sensibilidad actual se encuentra en el orden de los pocos femtomoles o incluso atomoles ( $10^{-15}$  y  $10^{-18}$  moles, respectivamente) y la exactitud de la medición de la masa molecular de los péptidos está en el orden de 1 a 10 ppm (Nagaraj et al., 2011; Mann et al., 2013).

Por otra parte, es bien conocida la importancia de las modificaciones postraduccionales (PTM) en los sistemas celulares y su influencia en los mecanismos de regulación, estabilidad y localización de las proteínas, entre otras funciones. La MS es la única herramienta disponible en la actualidad para identificar, localizar y cuantificar las PTM, aunque se requieren procedimientos específicos de enriquecimiento debido a que muchas modificaciones postraduccionales no modifican a las proteínas estequiométricamente y los péptidos modificados pueden estar en muy poca abundancia (Olsen & Mann 2013).

Existen dos tipos de abordajes en los experimentos de proteómica: los de proteómica para el descubrimiento y los experimentos de proteómica dirigida por hipótesis. En el primero, proteómica para el descubrimiento, se trata de identificar y cuantificar la mayor cantidad de proteínas en diferentes condiciones. Este tipo de abordaje tiene un carácter prospectivo y generalmente para su aplicación no se parte de una hipótesis previa definida. El objetivo es explorar e identificar las proteínas más importantes relacionadas con un determinado proceso y, a partir de los resultados, existe el potencial de generar diversas hipótesis.

El segundo es la proteómica dirigida por hipótesis, que busca una información cuantitativa muy confiable de un grupo de proteínas relacionadas con un proceso biológico determinado. En este caso se parte de la selección previa de un grupo de proteínas que pueden ser cuantificadas en un número diverso de condiciones, entre ellas las variaciones temporales. Con esta segunda alternativa es posible alcanzar una sensibilidad de 10 a 100 veces mayor, pues el espectrómetro de masas solo se ocupa de medir los péptidos específicos de este grupo de proteínas y no del total. Además, se puede tener una interpretación biológica de mayor significación del proceso en estudio.

Cuadro 1. Algunas aplicaciones de la proteómica



En la actualidad, estos dos enfoques (proteómica para el descubrimiento y dirigida por hipótesis) son excluyentes, es decir, es necesario realizar experimentos separados, aunque se espera que con el desarrollo tecnológico, así como el incremento cualitativo y cuantitativo de las bases de datos de espectros y la creación de programas computacionales más eficientes sea posible reunir los dos enfoques en un solo experimento.

#### DESARROLLO DE LA PROTEÓMICA EN AMÉRICA LATINA Y SITUACIÓN ACTUAL

En América Latina, Cuba y Brasil son los dos países que han alcanzado mayor desarrollo en proteómica, gracias a la inversión tecnológica en las últimas dos décadas. En 1987, Cuba fue el primer país latinoamericano donde se instaló un espectrómetro de masas para el análisis de péptidos y proteínas, varios años antes de que apareciera el término "proteoma". Era un equipo de doble sector con fuente de ionización por FAB.

La espectrometría de masas se integró rápidamente al análisis y caracterización de las proteínas recombinantes producidas en el Centro de Ingeniería Genética y Biotecnología (CIGB). Un ejemplo sobresaliente fue el aporte de esta técnica al establecimiento del proceso de producción del interferón alfa-2 recombinante humano. Con el auxilio de la EM se pudo identificar que el extremo N-terminal de la proteína (alrededor del 40 %) estaba acetilado, y que una parte considerable de la proteína no tenía los enlaces de disulfuro correctamente formados. Los procesos de fermentación, re-naturalización y purificación del interferón alfa 2 fueron monitoreados por HPLC y EM hasta que se logró reducir la acetilación a menos del 10 % y tener más del 90 % de la proteína con los puentes de disulfuro correctos. Además, se diseñó un proceso de purificación que elimina estos subproductos

y que permite obtener el interferón alfa 2 con más del 99 % de pureza. De esta forma quedó establecido el uso de la EM para la caracterización de todas las proteínas recombinantes producidas en el CIGB. Estos estudios fueron la base para el desarrollo de la proteómica en Cuba.

Brasil es el país latinoamericano con un número mayor de grupos dedicados a la proteómica y donde están instalados la mayoría de los espectrómetros de masas para estudios de proteínas en la región. México es el tercer país con mayor desarrollo y también existen grupos en Venezuela, Uruguay, Argentina y Costa Rica. En otros países hay grupos que han realizados estudios de proteómica en colaboración con laboratorios de fuera o dentro de la región. No obstante, el desarrollo de la proteómica en América Latina ha sido lento y desigual. El número de publicaciones es muy bajo en comparación al total de publicaciones en este campo, y la gran mayoría corresponden a grupos de los tres países con mayor desarrollo, principalmente a Brasil (Revisado por Padron & Domont, 2014).

#### PRINCIPALES APLICACIONES DE LA PROTEÓMICA EN AMÉRICA LATINA

A diferencia de lo que ocurre en otras regiones, en América Latina las aplicaciones en células y tejidos humanos y de mamíferos no son preponderantes. Solo en Cuba constituyen casi la totalidad de las aplicaciones, donde sobresalen los estudios sobre los mecanismos de acción de nuevos medicamentos (Padrón et al., 1989). En el resto de los países los estudios de proteómica están distribuidos entre plantas, parásitos y sus vectores, bacterias, humanos y otros mamíferos, y venenos (fundamentalmente de serpientes).

Veamos esta distribución en los tres países de más desarrollo, considerando solamente las publicaciones de 2015 don-

de los investigadores tienen una participación sobresaliente, sin tener en cuenta los artículos de revisión. El objetivo no es hacer una evaluación detallada, sino dar una idea de cuáles son los principales campos de aplicación.

En Cuba, todas las publicaciones tratan de aplicaciones en humanos o mamíferos, con la excepción de trabajos dedicados al desarrollo metodológico. En Brasil aproximadamente el 47 % están dedicadas a trabajos en humanos y otros mamíferos, mientras que el 53 % restante tiene una distribución relativamente uniforme entre plantas, bacterias y hongos, parásitos y sus vectores y venenos, con una ligera ventaja en las aplicaciones en plantas. En México los trabajos están igualmente distribuidos entre aplicaciones en plantas y humanos y otros mamíferos con un único ejemplo en venenos y otro en parásitos. Aunque Brasil lidera ampliamente las aplicaciones en el estudio de los parásitos en la región, como se mencionó anteriormente, el número de publicaciones es bajo con respecto al total mundial, y a grupos brasileños corresponden más del 50 % de todas las publicaciones de la región.

Las causas del poco desarrollo de la proteómica en la región son variadas, aunque la insuficiencia tecnológica parece ser la principal. Hay grupos e investigadores altamente calificados, pero el acceso a espectrómetros de masas de última generación es muy limitada. Esta situación puede ser extrapolada a otras ómicas y por tanto al avance en las investigaciones biológicas.

Delante de las limitaciones en la inversión en ciencia y tecnología en la región, los grupos interesados en explotar las potencialidades de la proteómica para responder sus preguntas biológicas necesitan articularse con grupos que detengan las más recientes tecnologías en países más desarrollados. Aunque esta es una vía para avanzar individualmente, es deseable que, para fortalecer este desarrollo, los grupos interactúen entre sí en la propia región, mejorando las capacidades locales, intercambiando información y tecnologías para que podamos responder a los desafíos locales en las distintas áreas de investigación y contribuir al conocimiento y avance en el entendimiento de problemas que afectan exclusivamente a la población de nuestra región y que han sido históricamente relegados.

## BIBLIOGRAFIA

1. Aebersold RH, Patterson SD. Proteomics: the first decade and beyond. *Nat Genet* 2003;33(Suppl):311–23.
2. Fenn J, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989;246:64–71.
3. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10000 kDa. *AnalChem* 1988;60:2299–301.
4. Mann M, Kulak NA, Nagaraj N, Cox J. The coming age of complete, accurate, and ubiquitous proteomes. *Mol Cell*. 2013 Feb 21;49(4):583-90.
5. Nagaraj N, Kulak NA, Cox J, Neuhauser N, Mayr K, Hoerning O, Vorm O, Mann M. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol Cell Proteomics*. 2012 Mar;11(3):M111.013722.
6. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M. Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology*. 2011; 7:548. doi:10.1038/msb.2011.81
7. Olsen JV, Mann M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol Cell Proteomics*. 2013 Dec;12(12):3444-52.
8. Padrón G, Besada V, Agraz A, Quiñones Y, Herrera L, Shimonishi Y, Takao T. Mass Spectrometric Analysis of Recombinant Human Alpha-2 Interferon. *Anal Chim Acta*. 1989; 223:361-369. doi: 10.1016/S0003-2670(00)84101-X.
9. Padrón G, Domont GB. Two decades of proteomics in Latin America: A personal view. *J Proteomics*. 2014;107C:83-92.
10. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. Full Dynamic Range Proteome Analysis of *S. cerevisiae* by Targeted Proteomics. *Cell*. 2009; 138:795-806. doi: 10.1016/j.cell.2009.05.051.
11. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T, et al. Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 1988;2:151–3.
12. Wells JM, McLuckey SA. Collision-Induced Dissociation (CID) of Peptides and Proteins. *Methods in Enzymology*. 2005; 402:148–185.