

# Desarrollo de nuevos métodos para el estudio del péptido N-terminal de proteínas y sus aplicaciones en la industria biotecnológica

✉ Aniel Sánchez, Lázaro Betancourt, Luis J González, Yassel Ramos, Jeovanis Gil, Yanni Solano, Vivian Morera, Yairet García, Félix Álvarez, Galina Moya, Jorge Fernández de Cossio, Gabriel Padrón, Vladimir Besada

División de Química Física. Centro de Ingeniería Genética y Biotecnología, CIGB Ave. 31 e/ 158 y 190, AP 6162, CP 10600, Cubanacán, Playa, Ciudad de La Habana, Cuba  
E-mail:aniel.sanchez@cigb.edu.cu

## Introducción

El desarrollo acelerado de la ingeniería genética y la biotecnología ha posibilitado obtener proteínas cuyo aislamiento es prácticamente imposible porque en condiciones naturales se expresan en muy bajas concentraciones. La técnica de recombinación de ADN permite insertar un gen que codifica para una proteína de interés dentro del genoma de otro microorganismo, empleando una construcción genética adecuada. Al favorecer el crecimiento del organismo hospedero, se puede lograr la proteína recombinante en cantidades suficientes. La localización celular puede ser predeterminada, y de esta forma implementar una estrategia para su purificación.

Sin embargo, durante la traducción, expresión, purificación o almacenamiento de las proteínas, pueden ocurrir modificaciones químicas que alteren su estructura primaria y con ello su función biológica. De manera frecuente también puede haber procesos de degradación por la acción de exopeptidasas que modifican los extremos N y C.

Una de las modificaciones postraduccionales más comunes de péptidos y proteínas es el bloqueo del grupo amino terminal. Se estima que más del 80% de las proteínas de eucariotas se encuentran N<sup>o</sup>-bloqueadas [1]. Tal modificación imposibilita la secuenciación por degradación de Edman. Por esas razones, los organismos reguladores internacionales exigen la información de la secuencia de este extremo, como control de la calidad de las proteínas producidas por esta vía.

La mayoría de los métodos alternativos a la degradación de Edman surgen con la utilización de la espectrometría de masas (MS) como técnica analítica más efectiva para el estudio de la estructura primaria de las proteínas. Ellos se basan en la combinación de digestiones enzimáticas y/o en la modificación química de los grupos aminos primarios, con algún paso cromatográfico previo.

Nuestro trabajo consiste en tres métodos que permiten el estudio del péptido N-terminal, bloqueado o no, de las proteínas. Uno de ellos se basa en el aislamiento selectivo del péptido (método 1) mientras los otros (métodos 2 y 3) permiten reconocerlo en un espectro de masas, partiendo de la proteína separada en una banda de gel para su identificación más directa.

El método de aislamiento selectivo del péptido N-terminal bloqueado (método 1), optimizado para proteínas en solución, es útil para el estudio de proteínas N-bloqueadas, y permite aislar el péptido

N-terminal bloqueado de una manera rápida y efectiva, con el uso de digestiones sucesivas y un paso de intercambio catiónico (Figura 1).

Sin embargo, a pesar de ser muy eficaz para estos casos, no lo es para el estudio de proteínas expresadas en bajas cantidades, ni para el análisis de proteínas con el extremo N libre, ni en estudios de estabilidad para identificar productos de degradación.

Los otros dos métodos (métodos 2 y 3) se optimizaron para el estudio de proteínas separadas en geles de poliacrilamida, lo cual es muy importante porque estas se obtienen en bajas concentraciones. Ambos permiten reconocer al péptido N-terminal de los restantes péptidos en el espectro de masas de la digestión enzimática, a diferencia de otros procedimientos reportados.

Este reconocimiento se logra con el análisis del espectro ESI-MS, donde según la distribución isotópica es posible identificar al péptido N-terminal. El método 2 permite el estudio de proteínas N-bloqueadas y de proteínas con el N-terminal libre, y ambos métodos se pueden emplear en el estudio de productos de degradación de proteínas, blanco importante de análisis en los estudios de estabilidad (Figura 2).

El método 3 combina la posibilidad de identificar al péptido N-terminal en un espectro de masas, reduciendo al mínimo el riesgo de falsas asignaciones, con la posibilidad de diferenciar las series de fragmentación en el espectro de masas ESI-MS/MS, y con ello aumentar la confiabilidad en la asignación de la secuencia del péptido, por lo que es recomendable para el estudio del extremo N-terminal de proteínas de organismos con genomas por secuenciar (Figura 3).

## Resultados y discusión

### Diseño de los métodos

El diseño de los métodos propuestos se basa en el uso de reacciones químicas sencillas, combinadas con la utilización de enzimas muy utilizadas en la química de proteínas junto con la cromatografía convencional.

El primero de los métodos permite el aislamiento selectivo del péptido N-bloqueado mediante la combinación de digestiones enzimáticas con tripsina y carboxipeptidasa B de la proteína de interés. Este tratamiento genera un péptido N-terminal bloqueado,

1. Brown JL, Roberts WK. Evidence that approximately eighty per cent of the soluble proteins from Ehrlich ascites cells are Nalpha-acetylated. *Biol Chem* 1976; 251:1009-14.

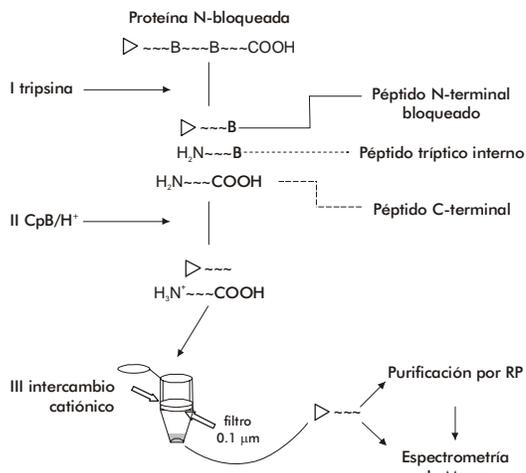


Figura 1. Esquema general de la estrategia propuesta para la identificación del péptido N-terminal bloqueado (método 1). La digestión triptica es tratada con carboxipeptidasa B. Luego del intercambio catiónico, el único péptido no retenido en la matriz es el péptido bloqueado del extremo N.

que a diferencia de los péptidos internos de la proteína no es retenido en un intercambiador catiónico fuerte (Figura 1).

Los otros dos métodos utilizan modificaciones comunes en la química de los grupos aminos primarios, tales como la acilación y la guanidincación. Todos los análisis se realizan mediante espectrometría de masas (Figura 2 y 3).

### Optimización

Las condiciones específicas y la optimización de los diferentes pasos y reacciones en los métodos se lograron inicialmente empleando muestras patrones, como los péptidos casomorfina-7, neuromedin K y la proteína mioglobina equina.

### Análisis del péptido N-terminal en diferentes muestras de proteínas

Los métodos diseñados se aplicaron con éxito en la caracterización e identificación de la secuencia N-terminal de diferentes muestras de proteínas. Con el método de aislamiento selectivo del péptido N-terminal de proteínas N-bloqueadas (método 1) se estudiaron las muestras siguientes [2].

#### Citocromo C

El citocromo C es una proteína que se utiliza con frecuencia como modelo para establecer nuevos métodos de aislamiento selectivo de péptido N-terminal bloqueado. Se logró aislar su péptido N-terminal.

#### Parvalbumina

El aislamiento del péptido N-terminal bloqueado de la proteína parvalbumina permitió establecer y estudiar el comportamiento de este tipo de péptidos cuando son analizados por espectrometría de masas con fuente de ionización MALDI. Se alcanzó el objetivo de una asignación inequívoca de la masa molecular del péptido aislado utilizando espectrómetros con diferentes fuentes de ionización.

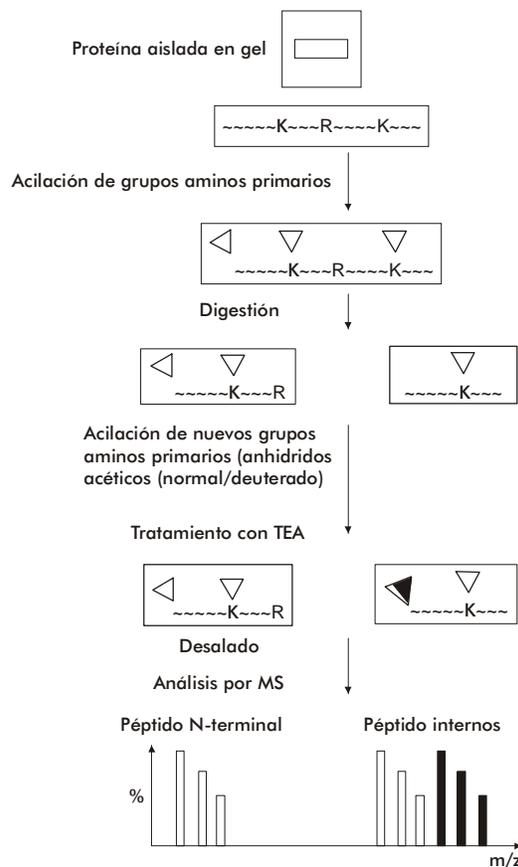


Figura 2. Esquema general del método propuesto. La proteína separada en el gel de poliacrilamida es acilada con anhídrido succínico o acético. La mezcla de péptidos tripticos generados es reacilada con una mezcla equimolar de anhídrido acético normal y deuterado. El péptido N-terminal es reconocido en el espectro de masas debido a que es la única especie que aparece con una distribución isotópica normal.

#### Interferón $\alpha$ -2b humano

Se logró aislar el péptido N-terminal del interferón  $\alpha$ -2b humano parcialmente bloqueado. Este resultado indicó las potencialidades del método en la caracterización de proteínas con heterogeneidades en el extremo N-terminal.

#### Troponina C

El aislamiento selectivo e identificación del péptido N-terminal acetilado de la proteína troponina C demostró que el procedimiento propuesto puede ser aplicado eficientemente en la detección de especies de proteínas bloqueadas dentro de mezclas de proteínas.

#### Levanasacarasa de *Gluconacetobacter diazotrophicus*

Se determinó la secuencia N-terminal de la proteína levanasacarasa en *Gluconacetobacter diazotrophicus* madura, y se detectó la presencia de un grupo piroglutámico en el extremo N-terminal, cuyo bloqueo ocurrió como consecuencia de la hidrólisis del péptido señal antes de la secreción de la proteína al medio.

2. Betancourt L, Besada V, González LJ, Morera V, Padrón G, Takao T, et al. Selective isolation and identification of N-terminal blocked peptides from tryptic protein digests. *J Pept Res* 2001;57:345-53.

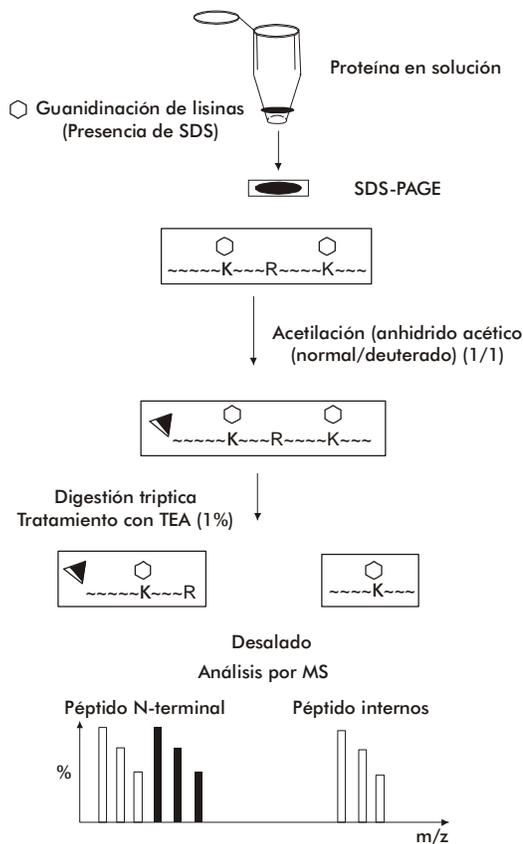


Figura 3. Esquema general de la estrategia propuesta para la identificación del péptido N-terminal. La proteína intacta es guanidada, luego de la digestión triptica, los péptidos son acetilados con una mezcla equimolar de anhídridos acéticos normal y deuterado. El péptido N-terminal es reconocido en el espectro de masas debido a que aparece con una distribución isotópica compleja a diferencia de los péptidos internos.

### Antígeno de superficie del virus de la hepatitis B

El antígeno de superficie del virus de la hepatitis B es el ingrediente farmacéutico activo de la vacuna contra este virus. La identificación y secuenciación del péptido N-terminal bloqueado, permitió por primera vez: 1) determinar la naturaleza química del grupo modificador y 2) localizar la fracción que contiene a este péptido en los mapas peptídicos que se obtienen durante los controles de calidad periódicos a la preparación vacunal.

Los métodos 2 y 3 que se basan en el reconocimiento de la señal correspondiente al péptido N-terminal en el espectro de masas [3, 4], se pueden aplicar al estudio de proteínas con el extremo libre o bloqueado. En estos, el péptido N-terminal adquiere una distribución isotópica diferente con respecto a los péptidos internos de la proteína. Se aplicaron en las muestras siguientes.

**Proteínas lisozima C (14kDa), inhibidor de tripsina (24 kDa), p64K (62kDa), fosforilasa B (97kDa, proteína acetilada en el extremo N) y mezcla de proteínas de *Escherichia coli***

Son muestras de proteínas de diferente peso molecular y composición, cuya hidrólisis enzimática dio lugar a

mezclas cada vez más complejas de péptidos. No obstante, se logró la identificación inequívoca del péptido N-terminal para cada una.

### Fragmento de anticuerpo multivalente tipo scFv para el antígeno carcinoembrionario

Esta proteína se obtuvo por la vía recombinante con vistas a su uso en el radioinmuno diagnóstico o la radioinmunoterapia de diversos tumores [5]. Una de las estrategias de identificación del péptido N-terminal se aplicó a las dos bandas de gel producto de la degradación del fragmento. El resultado permitió definir por cuál de los extremos era degradada la proteína, a la vez que se pudo verificar la estructura primaria de esta proteína.

### Cadena ligera del anticuerpo pHR-3

El anticuerpo R-3, comercialmente denominado TheraCIM® se emplea en el tratamiento de tumores de cabeza y cuello [6]. Esta molécula se expresó de forma transitoria en hojas de plantas de tabaco, y conservó su actividad biológica. Con el objetivo de favorecer su estabilidad biológica en la construcción génica de las cadenas ligera y pesada, se introdujo el péptido señal de esporamina del boniato (*Ipomoea batatas* L.). Debido a la gran importancia de este producto y a las diferencias de péptidos señales entre el plánculo y su homólogo expresado en el fluido ascítico de ratones, TheraCIM®, se realizó la verificación de la secuencia aminoacídica de ambas proteínas resultantes. La aplicación de del método 2 para la verificación del extremo N de la cadena ligera, permitió la identificación del péptido N-terminal y la verificación de la estructura primaria.

### Relevancia del estudio

Los métodos descritos tienen aplicaciones en los estudios de la estructura primaria de las proteínas. Nuestro grupo de trabajo tiene entre sus objetivos principales brindar servicios de análisis de estructura primaria de las proteínas producidas en el país, especialmente con fines terapéuticos. Muchos de los productos del Centro de Ingeniería Genética y Biotecnología tienen alto contenido proteico, al igual que los productos en la etapa de desarrollo e investigación.

El desarrollo de métodos para el estudio del péptido N-terminal ha posibilitado el análisis del extremo amino de proteínas con elevada importancia. El análisis de la levanasacarasa ayudó a comprender el mecanismo de secreción de esta enzima. El estudio de la estructura primaria del antígeno de hepatitis B (componente del producto principal del Centro de Ingeniería Genética y Biotecnología) ha permitido cumplir con las normas de calidad para su uso, exigidas por las entidades reguladoras nacionales e internacionales.

Además, con estas herramientas se pudieron estudiar proteínas en bajas concentraciones, separadas en geles de poliácridamida, tales como el fragmento del anticuerpo específico para el antígeno carcinoembrionario, donde el estudio del extremo N-terminal permitió corroborar que los productos de degradación de la proteína contenían el extremo amino terminal. También se pudo verificar y analizar el péptido N-terminal del

3. Sánchez A, Ramos Y, Solano Y, González LJ, Besada V, Betancourt L, et al. Double acylation for identification of amino-terminal peptides of proteins isolated by polyacrylamide gel electrophoresis. *Rapid Commun Mass Spectrom* 2007;21:2237-44.

4. Sánchez A, Ramos Y, Solano Y, González LJ, Betancourt L, Gil J, et al. Specific isotope labeling for the identification of free N-terminal peptide of proteins separated by polyacrylamide gel electrophoresis. *Eur J Mass Spectrom* (Chichester, Eng) 2007;13:307-9.

5. Tormo BR, Gavilondo JV, Domínguez C, Freyre M, Rodríguez T, Biberfeld P. CEA in colonic adenocarcinomas and precancerous lesions. An immunohistochemical study with a novel monoclonal antibody. *APMIS* 1989 Dec;97(12):1073-80.

6. Mateo C, Moreno E, Amour K, Lombardero J, Harris W, Pérez R. Humanization of a mouse monoclonal antibody that blocks the epidermal growth factor receptor: recovery of antagonistic activity. *Immunotechnology* 1997;3(1):71-81.

planticuerpo pH-3 obtenido de manera transitoria en hojas de tabaco. Estos métodos, además de contribuir al desarrollo de los servicios que brinda el Centro de Ingeniería Genética y Biotecnología, son científicamente novedosos, y se incluyen en la lista de los métodos publicados en revistas internacionales.

### **Reconocimientos**

Queremos expresar nuestro agradecimiento a Meilyn Rodríguez, Merardo Pujol y Lincidio Pérez, del Centro de Ingeniería Genética y Biotecnología, y a Toshifumi Takao y Yasutsugu Shimonishi de la Universidad de Osaka, Japón, por su colaboración en este trabajo.