

Estrategias de obtención de proteínas recombinantes en *Escherichia coli*

José García,^{1*} Zeila Santana,¹ Lourdes Zumalacárregui,² Marisel Quintana,¹ Diamilé González,¹ Gustavo Furrázola,¹ Oscar Cruz¹

¹Centro de Ingeniería Genética y Biotecnología. Ave 31 e/ 158 y 190. La Habana, Cuba. CP 10600.

²Facultad de Ingeniería Química Instituto Superior Politécnico "José Antonio Echeverría". Calle 114 No 11901 entre Ciclovía y Rotonda. La Habana, Cuba. CP 19390.

email: jose.garcia@cigb.edu.cu

La expresión de proteínas recombinantes se ha favorecido con el uso de *Escherichia coli* debido a su relativo bajo costo, alta densidad de cultivo, su fácil manipulación genética y a las diversas herramientas biotecnológicas disponibles que son compatibles. En este artículo se presentan algunas estrategias para la expresión de *Escherichia coli*; se destacan factores genéticos y fisiológicos que incluyen: número de copias del vector de expresión, características del gen, estabilidad del ácido ribonucleico mensajero, promotor empleado, cepa utilizada, composición del medio de cultivo, parámetros de operación en el fermentador y también se abordan la conservación de cepas y la estrategia de cultivo y purificación.

Palabras clave: proteína recombinante, *Escherichia coli*, factores genéticos, factores fisiológicos, fermentación.

Introducción

La bacteria *Escherichia coli* (*E. coli*) es ampliamente utilizada en la industria biotecnológica para la producción de proteínas recombinantes con fines diagnósticos, terapéuticos o vacunales, por procedimientos relativamente baratos, ya que ha sido muy bien caracterizada desde el punto de vista genético y fisiológico (1). Es un bacilo corto gramnegativo que se localiza en las heces fecales; habita en el conducto intestinal normal del hombre y de los animales y en algunos casos puede ocasionar enfermedades como: hepatitis, septicemia, cistitis y otras (2).

Entre las ventajas que este microorganismo brinda como hospedero están: mayor velocidad específica de crecimiento que las levaduras y células de mamíferos; fácil manipulación genética, no posee requerimientos costosos asociados a medios de cultivo o equipamiento a diferencia de las células de organismos superiores; existencia de gran variedad de vectores de expresión estables, y ser un microorganismo aprobado por las entidades reguladoras para su utilización como hospedero en la producción de biofármacos (3,4).

Es conocido que las proteasas producidas en el microorganismo pueden destruir las proteínas recombinantes obtenidas. Su empleo para producir proteínas útiles en la industria alimentaria no está permitido, segrega toxinas, por lo que se considera un patógeno moderado; es incapaz de realizar muchas de las modificaciones postraduccionales que son comunes en proteínas de origen eucariota; no es un eficiente secretor de proteínas al medio de cultivo, es una fuente potencial de pirógenos y su habilidad para promover la formación correcta de numerosos puentes disulfuro es

limitada. Pese a estas desventajas, numerosas proteínas complejas se obtienen en este hospedero con la actividad biológica requerida para ser utilizadas (5).

En las últimas décadas se han desarrollado cepas de *E. coli* modificadas genéticamente para solucionar algunas de sus desventajas. Entre ellas se encuentran varias cepas mutantes de los genes *lonA* y *ompT*, para disminuir la degradación proteolítica de los productos recombinantes; también se han obtenido cepas deficientes, o bien de la tioredoxina reductasa (*trxB*), o de la glutatión reductasa (*gor*), o de ambas a la vez, por lo cual en ellas se favorece la correcta formación de puentes disulfuro en el citoplasma. En otras cepas desarrolladas pueden obtenerse proteínas complejas, con modificaciones postraduccionales propias de células eucariotas, como es el caso de las que portan el gen de la tirosina kinasa para producir proteínas fosforiladas por la tirosina (6, 7).

Para aprovechar las ventajas que ofrece el periplasma, en comparación con el citosol, para la expresión de proteínas recombinantes se han desarrollado vectores plasmídicos en los cuales es posible clonar los genes recombinantes fusionados a señales de secreción que posibilitan el transporte de la proteína de interés a este compartimento (8).

Expresión de proteínas recombinantes en *Escherichia coli*

Entre los parámetros más importantes en la producción exitosa de proteínas recombinantes en *E. coli* se encuentran: la eficiencia transcripcional y traduccional, la estabilidad del

* Ingeniero Químico, Máster en Ciencias, Profesor Auxiliar, Investigador Auxiliar.

vector de expresión y de los ácidos ribonucleicos (ARN) transcritos, la estabilidad de las moléculas ante el ambiente proteolítico del hospedero y la localización y plegamiento de la proteína (3). También es necesario hacer un estudio con diferentes cepas del organismo hospedero, con vistas a seleccionar la más adecuada para la expresión de la proteína recombinante deseada, ya que estudios muestran que esta puede variar al usar una cepa determinada u otra (9).

Con el objetivo de alcanzar altos niveles de expresión de la proteína de interés es necesario estudiar la influencia de diferentes factores en el proceso de producción, los cuales pueden agruparse en genéticos y fisiológicos.

Factores genéticos

Entre los factores genéticos a considerar se encuentran: número de copias del vector de expresión, características del gen, estabilidad del ácido ribonucleico mensajero (ARNm), promotor empleado y cepa utilizada.

• Número de copias del vector de expresión o plasmidio

En la selección del sistema de expresión que permita obtener altos niveles de producción de la proteína recombinante, un aspecto importante es el número de copias del vector dentro de la cepa hospedera. El número de copias del plasmidio en *E. coli* puede variar desde 1 hasta 2 000 por célula, en dependencia de su origen de replicación en la bacteria. Generalmente la síntesis proteica aumenta linealmente con el número de copias del plasmidio hasta aproximadamente 600 copias por célula (10). Este aumento depende de la naturaleza de la proteína y del ambiente celular, en el que se incluye el balance de síntesis/degradación que puede ser afectado por elevadas cargas génicas del vector.

Por otro lado, un alto número de copias incrementa las complicaciones relacionadas con la replicación y expresión intensivas como son las mutaciones, disminución de la estabilidad del vector, problemas con el control de la regulación del promotor y dificultades en la purificación de la proteína de interés al incrementar también los niveles de proteínas codificadas por otros genes plasmídicos, por ejemplo, genes de resistencia a antibióticos (11).

• Características del gen heterólogo

Entre las características del gen heterólogo que pueden influir en los niveles de expresión se menciona el uso de determinados codones específicos en el gen que son poco usados en *E. coli*, denominados codones raros. Se ha observado que la presencia de estos, que codifican el aminoácido arginina en genes heterólogos, afectan considerablemente los niveles de síntesis de la proteína (12). Además, pueden existir determinados codones, fundamentalmente, los que codifican arginina y prolina, en los genes que favorecen el corrimiento del marco de lectura

hasta en un 50% del total de proteína sintetizada. Esto puede originar una terminación prematura de la síntesis de la proteína recombinante o alterar completamente la información genética (11).

Otra característica del gen heterólogo determinante de los niveles de expresión es la presencia de estructuras secundarias en el ARNm, que pueden afectar el inicio de la transcripción y de la traducción al bloquear la unión de la ARN polimerasa al promotor, o de las subunidades ribosómicas al sitio de inicio de la traducción, respectivamente (13).

• Estabilidad del ARN mensajero (ARNm)

Una vez sintetizado el ARNm correspondiente al gen de interés, la estabilidad del transcripto permite un control flexible y muy eficiente de los niveles de síntesis proteica. Por lo general los ARNm bacterianos tienen una vida media corta. En células bacterianas que crecen y se dividen cada 20 o 30 min, algunos ARNm deben renovarse de 5 a 10 veces en cada generación. La vida media del ARNm está regulada por el control del inicio de la transcripción y la velocidad de degradación del mismo. La velocidad de degradación de los ARNm provenientes de algunos genes heterólogos es muy alta, siendo muy bajos o nulos los niveles de síntesis de las proteínas recombinantes (14).

• Tipo de promotor bajo el cual se encuentra situado el gen de interés

El sistema de expresión es la unidad transcripcional que se inserta en un vector de ácido desoxirribonucleico (ADN) extracromosomal. Si el mismo es inducible, las células pueden crecer hasta altas densidades y obtener cantidades del producto hasta de un 30% de la proteína total de la células (15).

Algunos de los promotores más empleados son:

- ◇ Promotor triptófano: Es el responsable de la transcripción de los genes *trpE*, *D*, *C*, *B*, *A*, los cuales codifican para las enzimas que intervienen en la síntesis del triptófano. Se reprime por la presencia de este aminoácido en el medio de cultivo y se induce por su ausencia o al añadir ácido 3-indolacrílico. Constituye un promotor fuerte y ampliamente utilizado (16, 17).
- ◇ Promotor pL (Brazo izquierdo del fago lambda): Es inducible por cambios de temperatura, ya que existe un represor termolábil que a temperaturas entre 28-30 °C se une al promotor impidiendo la transcripción del gen. Su utilización ha confrontado problemas en algunos hospederos, ya que ante un cambio súbito de temperatura, hasta aproximadamente 42 °C, el represor sufre un cambio estructural que impide su unión al promotor, permitiendo la transcripción y la expresión de la proteína de interés (18). Este sistema tiene dos

desventajas fundamentales: la primera, es que solamente puede ser empleado en cepas de *E. coli* que contengan en su genoma el gen que codifica para el represor λ 857 del fago lambda y, la segunda, es que al trabajar a los niveles de temperatura de inducción es posible la degradación de la proteína heteróloga. Una de las alternativas para disminuir el efecto de la primera desventaja es el empleo de cepas comerciales que contengan el gen del represor o cotransformar con un vector que lo incluya.

- ◇ Promotor Lac (lactosa): Este promotor se reprime en presencia de glucosa y se induce en presencia de lactosa o su análogo isopropil-tiogalactosido (IPTG) (19). La primera puede ser empleada solamente en cepas que contengan en su genoma los genes que codifican para las tres enzimas involucradas en el metabolismo de la lactosa (permeasa, beta galactosidasa y transacetilasa). IPTG es un reactivo caro, por lo que su empleo se reduce fundamentalmente a las escalas de laboratorio y piloto.
- ◇ Promotor Tac: Es un híbrido del promotor Lac y el promotor Trp y su función es similar a la del promotor Lac (20).

• Ceba hospedera

La ceba hospedera, la cual es transformada con el vector de expresión, debe ser deficiente en la mayoría de las proteasas naturales, mantener la estabilidad plasmídica y contener los elementos genéticos necesarios, de acuerdo con el sistema de expresión utilizado.

Se recomienda el empleo de cepas robustas capaces de crecer en medios de cultivo con un mínimo de nutrientes y que no sean patogénicas (21).

Una vez que se han establecido los factores genéticos adecuados, se debe garantizar la estabilidad del clon productor, evitando que existan mutaciones, contaminaciones y pérdida de información genética. Por lo anterior se debe garantizar una forma correcta de conservación del microorganismo productor. Para ello se desarrollan los bancos de células.

Estrategias de clonaje

A lo largo de más de dos décadas se han ido mejorando las estrategias genéticas, con el objetivo de aumentar los niveles de expresión de proteínas, así como simplificar los procesos de recobrado. Estas abarcan desde el diseño de vectores, multimerización de genes, mejoramiento de cepas basadas en diferentes principios como la delección de genes que codifican para proteasas, sobreexpresión de codones raros para *E. coli*.

También es posible adaptar la estructura de la molécula recombinante para potenciar una propiedad específica que facilite su purificación, ya sea modificando su punto isoeléctrico o sus constantes de afinidad (22).

Conservación de cepas

• Bancos celulares / Sistema de lotes de semilla

Los bancos de células son elementos fundamentales en el desarrollo y producción de productos biotecnológicos, ya que constituyen los seres vivos que transforman los sustratos y en condiciones predeterminadas llevan a cabo la síntesis del producto de interés (23).

Una de las ventajas en la producción de biotecnológicos/biológicos es la aplicación de un sistema de lote semilla, donde se utilizan subcultivos de células seriadas, las cuales tienen una fuente de partida caracterizada y común para cada lote de producción. Este sistema se utiliza para prevenir los cambios posibles de las propiedades bioquímicas y genéticas que se producen con subcultivos repetidos o de generaciones múltiples (24). Los fabricantes pueden preparar sus propios bancos u obtenerlos de fuentes externas. Ellos son responsables de asegurar la calidad de su preparación y de la ejecución de los ensayos de control.

◇ Bancos de células primarias (BCP) / Bancos de reserva

El banco de células primarias es una suspensión homogénea de células iniciales ya transformadas por el vector de expresión, almacenadas en envases individuales y en condiciones que garanticen su estabilidad genética.

Para llevar a cabo la preparación de un BCP se deberán realizar los pasos siguientes: seleccionar el clon de mayor expresión; tomar el cultivo en la fase logarítmica de crecimiento; lavar las células, resuspender en el medio de cultivo y adicionar crioprotectores (glicerol, dimetil-sulfóxido); distribuir en viales estériles y herméticos; almacenar las alícuotas a -70 °C y mantenerlas bajo custodia y control continuo de la temperatura. Además, se debe trabajar en ambientes controlados, bajo gabinetes de seguridad biológica y evidenciar los estudios de estabilidad y de recobrado de semillas y bancos, los cuales deben estar documentados.

◇ Control de calidad a los bancos de células primarias

Para la caracterización de los BCP se deben realizar diferentes ensayos como: pureza microbiana, viabilidad, análisis de restricción, secuencia nucleotídica del gen, estabilidad plasmídica, reconocimiento inmunológico, actividad biológica y nivel de expresión de la proteína a escala de banco.

◇ Banco de células de trabajo

El Banco de células de trabajo (BCT) se prepara a partir del BCP bajo condiciones definidas de cultivo. Estas células se utilizan directamente en el proceso productivo (24). La

cantidad de viales va a depender del uso, tiempo de vida y condiciones de almacenamiento. Cada vez que se utilice un vial, este debe ser desechado, ya que los cambios de temperatura pueden afectar la estabilidad genética del banco, así como aumentar los riesgos de contaminación.

◇ Control de calidad a los bancos de células de trabajo

Para el control de los BCT se deben realizar los mismos ensayos que a los BCP, pero el nivel de expresión de la proteína se realiza a la escala productiva y se adiciona el cumplimiento de las especificaciones de calidad del ingrediente farmacéutico activo (IFA) y la integridad de la proteína (secuencia de los extremos amino y carboxilo).

En este caso se mantienen los lotes en cuarentena y se aprueban después de la revisión del expediente por parte de la Dirección de Aseguramiento de la Calidad. Después de liberados los lotes se termina el expediente del banco y se entrega al productor para su custodia.

Se recomienda dividir los lotes de semillas y bancos celulares y almacenar las partes en lugares diferentes, lo cual minimiza los riesgos de pérdida totales de los bancos en caso de catástrofes.

Durante la preparación de los bancos primarios o bancos de reserva, los bancos de trabajo o cultivos de trabajo, se debe evitar o minimizar los riesgos de contaminación o alteración, confirmar su viabilidad, pureza e identidad antes de su uso y conformar el expediente de la cepa que debe reflejar la historia de las mismas, así como su control sistemático según los procedimientos establecidos (25).

Factores fisiológicos

Los factores fisiológicos asociados a la expresión y crecimiento celular de la proteína de interés se ponen de manifiesto en el proceso de fermentación. Entre los más importantes se encuentran: composición del medio de cultivo, parámetros de operación en el fermentador y estrategia de cultivo.

• Proceso de fermentación

Los microorganismos y las células en general presentan una identidad propia, especificada por su integración genética y manifestada en su funcionamiento con las enzimas que contienen. Para que los individuos y las especies que ellos forman sobrevivan y sean perpetuados esta especificación genética y su sistema de funcionamiento se deben mantener y reproducir a través del crecimiento.

El primer objetivo de una fermentación industrial es la obtención de un producto de calidad a un costo de producción tan bajo como sea posible. La necesidad de minimizar los costos de producción es el factor que controla la selección

de las materias primas y que conduce al mejoramiento de las tecnologías y a la búsqueda de nuevas cepas (9).

• Composición del medio de cultivo

◇ Requerimientos nutricionales del crecimiento

Fuente de carbono y energía

Aproximadamente el 50% del peso seco del material celular es carbono, por lo que en todo medio de cultivo debe existir un sustrato que proporcione el esqueleto carbonado básico para la síntesis de las unidades estructurales que dan origen a las macromoléculas y estructuras de la célula. En la mayoría de los casos este compuesto es de origen orgánico y entre los más usados se encuentran los azúcares como: glucosa, sacarosa, fructosa, lactosa, entre otros. Por lo general en los procesos degradativos o catabólicos de estos compuestos se produce también la energía necesaria para los procesos biosintéticos y el funcionamiento general de la célula.

La respiración de las bacterias es un proceso complejo que se acompaña del desprendimiento de energía que es indispensable para la síntesis de diferentes compuestos orgánicos. Según el tipo de respiración los microorganismos se dividen en: aerobios estrictos, microaerófilos, anaerobios facultativos y anaerobios estrictos.

Un representante típico de los anaerobios facultativos es *E. coli*, la cual en los medios que contienen fuentes de carbono, se desarrolla como bacteria anaerobia al comienzo del proceso; desintegra los hidratos de carbono mediante la fermentación, para después comenzar a asimilar el oxígeno; crece como aerobia y oxida los productos de la fermentación hasta formar CO₂ y H₂O. Este tipo de microorganismo posee importantes ventajas, puesto que puede vivir tanto en ausencia como en presencia de oxígeno.

Fuente de nitrógeno

En orden consecutivo es el segundo elemento en importancia que debe formar parte del medio de cultivo, ya que juega un papel fundamental en la composición de las estructuras principales de la célula. Al nitrógeno le corresponde del 8 al 15% del residuo seco.

Es un factor esencial de las proteínas y forma parte de los ácidos nucleicos, la pared celular, la membrana celular y otros componentes importantes (26). En la mayoría de los casos se suministra en forma de amonio o de algunas de sus sales, aunque existen otros que requieren aminoácidos, purinas, pirimidinas y vitaminas.

Los requerimientos orgánicos de nitrógeno son relativamente caros, por lo que a veces se trata de encontrar la satisfacción de esta demanda en subproductos industriales (suero de leche) o en productos más baratos (soya, extracto de

levadura), los cuales pueden aportar algunos factores del crecimiento que se requieren en cantidades muy reducidas (27).

• Velocidad específica de crecimiento

La velocidad específica de crecimiento (μ) es función de: tipo de microorganismo, composición del medio de cultivo, temperatura, pH, concentración de sustrato, velocidad de agitación y aireación. En general se puede plantear que:

$$\mu = f(T, \text{pH}, S, \text{variables de operación}).$$

Con el objetivo de mejorar la productividad de la proteína de interés, obtenida al usar una cepa de *E. coli*, se puede correlacionar la velocidad específica de crecimiento con la velocidad de producción del producto de interés. De esta manera se conoce que manteniendo la velocidad específica de crecimiento cerca del valor óptimo para la formación del producto después de la inducción, se obtienen altos niveles de expresión de proteína. Utilizando la alimentación escalonada o constante, después de la inducción del cultivo por elevación de la temperatura, se obtiene una extensión del período de producción de la proteína de manera significativa.

• Efecto de la temperatura y el pH

La temperatura de crecimiento óptima para diferentes microorganismos varía en un amplio intervalo. Por ejemplo: microorganismos psicrófilos: 0-25 °C; microorganismos mesófilos: 25-45 °C, y microorganismos termófilos: > 45 °C. La dependencia de la temperatura de crecimiento microbiano se podría comprender como el efecto de la temperatura sobre las distintas reacciones enzimáticas que ocurren en las células. Este efecto se describe por la ecuación de Arrhenius:

$$\mu = A \cdot e^{-E/R \cdot T}$$

donde:

- A : Constante de Arrhenius (h^{-1}).
- E : energía de activación (kJ kmol^{-1})
- T : temperatura absoluta (K)
- R : constante de los gases ($\text{kJ kmol}^{-1}\text{K}^{-1}$).

El efecto del pH es similar al efecto de la temperatura. La velocidad de las reacciones enzimáticas depende del valor del pH y varía de un microorganismo a otro.

La mayoría de las bacterias crecen a una temperatura alrededor de 37 °C y un pH neutro. Manteniendo el cultivo a estas condiciones se logra que la velocidad específica de crecimiento dependa solamente de la concentración de sustrato y de las variables de operación.

• Efecto de la concentración de sustrato

La dependencia de la velocidad específica de crecimiento (μ) sobre la concentración de sustrato limitante se obtiene por primera vez por J. Monod (28).

$$\mu = \mu_{\text{máx}} \cdot S / (K_s + S)$$

donde:

$\mu_{\text{máx}}$: Velocidad específica máxima de crecimiento que se alcanza en la fase exponencial y depende del tipo de sustrato limitante empleado y de las variables de operación (h^{-1}).

S: Concentración de sustrato limitante (g.L^{-1}).

K_s : Constante de saturación (g.L^{-1}). Esta constante da una medida de la afinidad del microorganismo hacia el sustrato. Mientras más pequeño sea K_s , más afín es el microorganismo al sustrato en cuestión.

Actualmente existen numerosos modelos que relacionan la velocidad específica de crecimiento con la concentración de sustrato, pero el de Monod sigue teniendo gran utilidad.

• Estrategias de cultivo

Existen diferentes tipos de estrategias para las fermentaciones microbianas: fermentación discontinua, fermentación discontinua incrementada y fermentación continua.

◇ Fermentación discontinua

La fermentación discontinua sigue siendo actualmente el método más usado en la industria farmacéutica y biotecnológica. Esta operación consiste en que se cargan el medio de cultivo y el inóculo en un tanque con agitación y aireación, si el proceso lo requiere, donde ocurre una transformación biológica. Teniendo esto se está en presencia de un fermentador o un biorreactor. La fermentación discontinua tiene entre sus ventajas: el ser una operación simple, presentar bajo riesgo de contaminación y facilidad para ser utilizada como patrón para estudios previos. Como desventaja se encuentra que con su uso, se obtiene baja productividad.

Limitaciones de la operación discontinua

Existe una concentración de sustrato limitante inhibitoria al crecimiento. Esto se conoce comúnmente como efecto Crab Tree y consiste en que el microorganismo en vez de consumir la fuente de carbono para crecer, produce ácido acético y otros metabolitos que son tóxicos para las células, por lo que se limita la obtención de altas densidades celulares (29).

En la fermentación discontinua existe una contradicción, al inicio habitan pocas células y existe alta concentración de sustrato, mientras que al final ocurre lo contrario.

En numerosas ocasiones en la producción de proteínas recombinantes la velocidad específica máxima de crecimiento no es la óptima para la formación del producto. Se conoce como $\mu_{\text{crítica}}$ a la velocidad específica a la que se elimina casi totalmente la producción de ácido acético, la cual perjudica tanto al crecimiento como a la formación del producto. Esto puede variar de una cepa a otra (si es auxotrófica o protótrofa), como también influye el medio en que se cultivan (si es medio complejo o definido). Kleman reporta que en un medio definido o salino: $\mu_{\text{crit}} \cong 0,35 \text{ h}^{-1}$ mientras en un medio complejo: $\mu_{\text{crit}} \cong 0,2 \text{ h}^{-1}$ (29).

Medio complejo se le llama a los medios ricos en extractos, suplementados con fuentes orgánicas, aminoácidos, vitaminas y trazas. Son medios muy ricos por lo que los microorganismos tienden a producir más metabolitos que en un medio salino definido.

Por otra parte, se requiere conocer la velocidad específica de crecimiento para la formación del producto (μ_{prod}). Existen tres formas de determinarla:

- √ En un cultivo discontinuo manipulando la aireación y la agitación.
- √ En un cultivo discontinuo con alimentación variando el flujo de incremento.
- √ En un cultivo continuo variando la velocidad de dilución, siempre que no se pierda la información genética.

◇ Fermentación discontinua incrementada

La producción de proteínas heterólogas en *E. coli* puede incrementarse significativamente mediante la utilización de sistemas de cultivo incrementado, el cual permite obtener densidades celulares superiores a 100 g/L de peso seco (30) y disminuir drásticamente los costos de producción de proteínas recombinantes (31).

Este sistema de fermentación presupone resolver algunos inconvenientes, como son: la inhibición del crecimiento por acumulación de sustrato; la capacidad limitada de transferencia de oxígeno y la formación de subproductos inhibitorios para el crecimiento como el ácido acético.

Para cultivar células de *E. coli* a altas densidades es necesario el diseño de un medio de cultivo balanceado, que contenga todos los componentes necesarios para el crecimiento celular y evite la inhibición por altas concentraciones de nutrientes. Con este objetivo se ha propuesto ir incrementando lentamente el sustrato limitante, y con ello evitar la acumulación de nutrientes y obligar a crecer al microorganismo a una velocidad baja (32).

Métodos de adición del incremento

Los flujos de adición del incremento pueden ser: constante, escalonados y exponenciales, y constituyen un elemento

fundamental para la obtención de buenos resultados, ya que no solo afecta a la concentración celular sino también a la productividad del proceso y a la cantidad de productos de interés a obtener.

A flujos constantes la adición se realiza a una velocidad predeterminada y el volumen del cultivo y la densidad celular aumentan, mientras la velocidad específica de crecimiento disminuye constantemente (33).

El flujo escalonado o gradual permite que se incremente la velocidad de adición en la medida que aumenta la concentración celular. En este caso las células pueden crecer exponencialmente durante todo el tiempo de cultivo siempre que la velocidad de adición del sustrato limitante se aumente proporcionalmente al crecimiento celular.

El método de adición mediante un flujo exponencial permite el crecimiento a una velocidad específica constante (generalmente, entre 0,1 y 0,3 h^{-1}). A través de este control se puede minimizar la producción de acetato (34).

En la actualidad se han desarrollado métodos más sofisticados de adición que presuponen esquemas de control por retroalimentación, asociados a parámetros físicos medibles como: el oxígeno disuelto (35), el pH y la velocidad de evolución de CO_2 (36).

Efecto inhibitorio del ácido acético y el oxígeno en el crecimiento de los cultivos incrementados

Es bien conocido que durante el proceso de crecimiento *E. coli* produce ácido acético como subproducto (36). La cantidad de acetato producido depende fundamentalmente de la cepa (37), la velocidad específica de crecimiento (29, 34), el medio utilizado, pH (38), así como el tipo y la concentración de la fuente de carbono (30).

La formación de ácido acético y su actividad inhibitoria puede ser controlada mediante estrategias de alimentación exponencial (38) y mediante sistemas de monitoreo de la fuente de carbono y de oxígeno disuelto (39, 40). Existen procedimientos en los que se combina el proceso de fermentación con métodos de filtración y diálisis (41), así como la utilización de cepas con la enzima fosfotransacetilasa mutadas (42). Otro de los elementos que puede provocar la inhibición del crecimiento, así como la estabilidad y calidad del producto de interés es la concentración de oxígeno disuelto en el medio. La escasez de oxígeno puede causar limitaciones en la producción de aminoácidos y problemas con la estabilidad de los plásmidos (39).

Desventajas del cultivo incrementado

La limitación de este cultivo va a estar dada por el mayor riesgo de contaminación que presenta en comparación con los cultivos discontinuos, así como por la inestabilidad

plasmídica que puede originarse debido al alto número de duplicaciones que se generan durante el proceso. Este fenómeno puede influir significativamente en la disminución de los niveles de expresión de la proteína de interés.

• Estrategias de purificación

El desarrollo de estos procesos ha requerido a la par el desarrollo de procesos de purificación. Las proteínas terapéuticas obtenidas por esta vía deben cumplir no solo rigurosos requerimientos de calidad y seguridad establecidos por las agencias regulatorias, sino también su camino hacia el mercado debe ser corto. La estrategia fundamental de la purificación de proteínas es simple: remover todos los contaminantes mientras se trata de retener en el mayor grado posible la proteína de interés.

De forma general, la purificación de proteínas se puede dividir a partir de su forma de expresión en dos grandes grupos que determinan la estrategia a seguir: purificación de proteínas solubles e insolubles. Las proteínas recombinantes en *E. coli* pueden ser expresadas de dos formas fundamentales: secretadas al medio extracelular (solubles) o expresadas en el interior de la célula, donde pueden ser localizadas en el espacio periplasmático o en el citoplasma celular (Tabla 1). Muchos de los polipéptidos recombinantes producidos en altos niveles en *E. coli* se localizan en el citoplasma en forma insoluble y compacta, formando cuerpos de inclusión, aunque en ocasiones puede encontrarse de forma soluble (43).

Tabla 1. Formas de expresión de las proteínas intracelulares en *E. coli*

Ubicación dentro de la célula	Forma de expresión
Citoplasma	Soluble
	Insoluble (cuerpo de inclusión)
Periplasma	Soluble
	Insoluble (cuerpo de inclusión)

Las extracelulares son generalmente más estables (a menudo como resultado de los puentes disulfuro) y por lo general son moléculas pequeñas. Una alternativa importante para la purificación de estas proteínas es la cromatografía de lecho o cama expandida, una técnica cromatográfica en la cual la proteína es purificada a partir de un extracto crudo (la muestra aplicada es particulada: medio de fermentación que contiene células), donde se combina la preparación de la muestra y captura en un solo paso, muy adecuada para la purificación a gran escala.

Otro método empleado para la purificación de estas proteínas es el uso de “marcadores” de afinidad. Esta técnica consiste en la adición a las proteínas (generalmente en su extremo carboxilo-terminal) de una cola de polihistidina que permite la purificación de estas por medio de una cromatografía de afinidad por iones metálicos (IMAC). La inserción de un sitio de corte de proteasa entre la “cola” de histidinas y la proteína de interés permite la eliminación de dicha cola antes de

cualquier paso posterior. Este método de purificación por afinidad permite obtener elevados niveles de pureza empleando un solo paso cromatográfico (44).

En el caso de las proteínas solubles expresadas intracelularmente es necesaria la ruptura de las células que la contienen. Existen varios métodos para lograr esto, como son: ciclos de congelación/descongelación repetidos, ultrasonido, homogenización por alta presión o permeabilización con solventes orgánicos. El método a escoger dependerá de cuán frágil sea la proteína y de cuán resistente sea la célula que se requiere romper. Los contaminantes en los extractos proteicos pueden incluir una variedad de macromoléculas (ácidos nucleicos, lípidos, polisacáridos y otras proteínas) así como una serie de pequeñas moléculas. Estas últimas se pueden fácilmente separar de las proteínas por métodos basados en la talla molecular, tales como la diálisis, la ultrafiltración o la filtración en gel.

Las macromoléculas son más difíciles de eliminar, especialmente cuando están presentes en grandes cantidades. Las bacterias lisadas liberan grandes cantidades de ADN, originando extractos con alta viscosidad y alta absorbancia a 280 y 260 nm. Este se puede eliminar por precipitación con sulfato de protamina (una proteína natural con propiedades de unión al ADN, la cual contiene un 70% de arginina y lisina) o por tratamiento con compuestos de poliaminas sintéticos. También se emplea el tratamiento con enzimas que ayuden a depolimerizar al ADN y ARN.

◇ Cuerpos de inclusión

Las proteínas recombinantes que se acumulan intracelularmente en *E. coli* frecuentemente se depositan en forma de cuerpos de inclusión. Estos son agregados insolubles de proteínas incorrectamente renaturalizadas, las que carecen de actividad biológica y pueden observarse por microscopía de fase como partículas intracelulares oscuras. El tamaño de los cuerpos de la inclusión varía entre 0,2–0,6 μm (45).

La proteína recombinante desnaturalizada es el componente principal de los cuerpos de inclusión y la separación inicial de los mismos por centrifugación es un paso eficaz de purificación. La molécula recombinante nativa se solubiliza usando agentes caotrópicos a través de un proceso conocido como extracción. Posteriormente se restablece la estructura terciaria de la proteína mediante un paso de renaturalización, que consiste en la disminución de la concentración de este agente por diálisis, dilución directa o empleando una columna cromatográfica (46, 47). Mientras la renaturalización puede ser un proceso sencillo para las proteínas pequeñas que no contienen puentes disulfuro, este es un proceso complejo y de bajos recobrados en muchos casos. Las condiciones óptimas para la renaturalización son dependientes de cada proteína y por tanto no se pueden predecir. Sin embargo,

existe mucha información en la literatura, disponible en bases de datos (48).

Antes de la extracción, los cuerpos insolubles colectados se pueden lavar con soluciones que contienen desnaturalizantes (1-4 mol/L de urea), sacarosa o detergentes para reducir los niveles de contaminantes, pero a costa de incrementar los costos y la complejidad de los procesos.

◊ Empleo de los métodos cromatográficos

La introducción de la cromatografía a comienzos de los años 1960, principalmente la filtración en gel y el intercambio iónico, proporcionaron nuevas oportunidades para la purificación.

Si bien en esta etapa los procesos se enfocaron en mejorar la precisión de los estudios estructura-función de las proteínas y los procesos de escalado se vieron limitados por las características físicas de los geles, a partir de los años 1970, con la aparición de la tecnología del ADN recombinante, el desarrollo de las fermentaciones bacterianas y el cultivo de células de mamíferos, los productos biofarmacéuticos se ajustaron a procesos de purificación basados en los procesos cromatográficos como la herramienta fundamental y con tecnologías de membrana que proporcionaron materiales limpios clarificados, cambio de soluciones, concentración del producto, eliminación viral y filtración estéril (49). La estandarización permitió un enfoque más sistemático al desarrollo de estos procesos y fue el basamento para la introducción del paradigma *captura-purificación-pulido*, ahora extendido en todos los diseños de los esquemas de purificación de proteínas.

La industria desarrolló un nuevo y sistemático diseño integrado de procesos, ingeniería y control, economía de proceso e higiene y aspectos regulatorios. Los sistemas de bioprocesos se introdujeron y el control computarizado se impuso en esta tecnología, previamente dominada por operaciones manuales. La expansión de la cromatografía como la herramienta fundamental en la purificación de las biomoléculas se aprecia claramente en el incremento de los ingresos en bioprocesos para la división de Life Science de la firma GE Healthcare de aproximadamente 36 millones en 1986 a 461 millones en 2006 (50).

◊ Eliminación de endotoxinas

Otro problema confrontado durante la purificación de proteínas terapéuticas expresadas en *E. coli* es la eliminación de las endotoxinas bacterianas.

El término pirógeno microbiano en oposición a endotoxinas de las bacterias gramnegativas se ha convertido en un término descriptivo general para muchas sustancias diferentes. Sin embargo, las sustancias pirogénicas pueden

ser producidas por algunas bacterias grampositivas, micobacterias, hongos y también algunos virus, pero los pirógenos producidos por las bacterias gramnegativas, las endotoxinas, son un asunto a considerar en la industria farmacéutica.

Las endotoxinas bacterianas, encontradas en la membrana externa de las bacterias gramnegativas son miembros de una clase de fosfolípidos llamados lipopolisacáridos (LPS). Los LPS no son productos exógenos de las bacterias gramnegativas. La liberación de estos ocurre después de la muerte o de la lisis celular. Algunos ejemplos de bacterias gramnegativas productoras de pirógenos son la *E. coli* y miembros de los géneros *Proteus*, *Pseudomonas*, *Enterobacter* y *Klebsiella*.

Pueden existir varias fuentes de pirógenos en los productos parenterales y en el equipamiento médico. Entre estas encontramos el agua usada como solvente o en el procesamiento, componentes de envase, reactivos, materias primas o equipamiento usado en la preparación del producto. Las buenas prácticas incluyen el control de los niveles microbiológicos y de endotoxinas en las fuentes potenciales mencionadas.

Teniendo en cuenta la composición de la molécula de LPS (una porción de la molécula es hidrofóbica y la molécula posee carga negativa a pH neutro) también la cromatografía de intercambio aniónico ha sido reportada para la eliminación de endotoxinas.

Chen y colaboradores reportan la reducción de hasta dos mil veces en los niveles de endotoxinas en un paso cromatográfico, empleando una resina de intercambio aniónico: Q XL, de la firma GE Healthcare (50), la cual posee un grupo intercambiador amino cuaternario.

◊ Pureza versus recobrado

La mayoría de los protocolos de purificación de proteínas a escala preparativa involucran varios pasos. Cada paso permite alcanzar un grado de pureza pero un intervalo de selección estrecho. En el caso de las fracciones con mayor pureza, puede reducir grandemente el recobrado.

Otras pérdidas pueden venir de la unión irreversible de la proteína a la matriz cromatográfica, de la desnaturalización, de la oxidación de la proteína, de condiciones drásticas durante la unión o durante la elución de las matrices cromatográficas, exposición a iones metálicos pesados y a la separación de cofactores, grupos prostéticos y agentes estabilizantes de la proteína de interés.

Conclusiones

Por último, uno de los objetivos de los investigadores fue alcanzar altos rendimientos en la producción de productos

biotecnológicos expresados en *E. coli*, con vistas a hacer estos procesos más factibles económicamente y cumplir con las regulaciones vigentes en esta rama de la industria.

Las estrategias han estado encaminadas desde el mejoramiento genético de la cepa y los vectores de expresión hasta los métodos de cultivo y purificación de estas proteínas. No obstante, existen retos que deben abordarse en un futuro, entre los que se encuentran: superar los niveles de expresión de proteínas alcanzados; realizar un adecuado estudio del espacio de diseño durante la etapa de desarrollo de los procesos de fermentación y recobrado, para establecer los parámetros que influyen significativamente en su eficiencia; realizar una adecuada selección del medio de cultivo que facilite tanto la expresión como el crecimiento celular.

Aunque en este artículo se han dado consideraciones generales, el desarrollo de cada producto requerirá un estudio de las características particulares de las moléculas recombinantes y de las cepas productoras.

Referencias

- Ferrer N, Domingo J, Corchero J, Vázquez E, Villaverde A. Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories* 2009;8(1):17-8.
- Perry Chou C. Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. *Appl Microbiol Biotechnol* 2007;76:521-32.
- Jonasson P, Liljeqvist S, Nygren PA, Ståhl S. Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli*. *Biotechnol Appl Biochem* 2002;35(2):91-105.
- Choi JH, Keum KC, Lee SY. Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chemical Engineering Science* 2006;61:876-85.
- Santana H, Martínez E, Sánchez J C, Moya G, Sosa, R, Hardy E, et al. Molecular characterization of recombinant human interferon alpha-2b produced in Cuba. *Biociencia Aplicada* 1999;16(3):154-9.
- Lehmann K, Hoffmann S, Neudecker P, Suhr M, Becker W M, Rosch P. High-yield expression in *Escherichia coli*, purification, and characterization of properly folded major peanut allergen Ara h2. *Protein Expr Purif* 2003;31:250-9.
- Bessette PH, Aslund F, Beckwith J, Georgiou G. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci USA* 1999;96:13703-08.
- Schatz G, Dobberstein B. Common principles of protein translocation across membranes. *Science* 1996;271:1519-26.
- Weikert C, Sauer U, Bailer J. An *E. coli* Host strain useful for efficient overproduction of secreted recombinant protein. *Biotech Bioeng* 1997;59(3):386-91.
- Nordström K, Uhlin B E. Runaway replication plasmids as tools to produce large quantities of proteins from cloned genes in bacteria. *Biotechnology* 1992;10:661-6.
- Bentley WE, Mirjalili N, Andersen DC, Davis RH, Kompala DS. Plasmid-encoded protein: the principal factor in the "metabolic burden" associate with recombinant bacteria. *Biotechnol Bioeng* 1990;35:668-81.
- Lee CP, Li P, Inouye H, Brickman ER, Beckwith J. Genetic studies on the instability of b-galactocidase to be translocated across the *Escherichia coli* cytoplasmic membrane. *J Bacteriol* 1989;171:4609-16.
- Kane JF. Effects of rare codon clusters on high-level expressions of heterologous protein in *Escherichia coli*. *Curr Opin. Biotechnol.* 1995;6:494-500.
- Bechhofer D. 5' m RNA stabilizers. In Belasco JG, Brawerman G, eds. *Control of messenger RNA stability*. San Diego: Academic Press; 1993. p. 31-52.
- Hanning G, Makrides S C. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol* 1998; 16:54-60.
- Espinosa R, Caballero E, Musacchio A, Silva R. Production of a recombinant, immunogenic protein, P64k, of *Neisseria meningitidis* in *Escherichia coli* in fed-batch fermenters. *Biotechnology Letters* 2002;24(5):343-6.
- Chevalet L, Robert A, Gueneau F, Bonnefoy, JY y Ngugen T. Recombinant protein production driven by the tryptophan promoter is tightly controlled in ICONE 200, a new genetically engineered *E. coli* mutant. *Biotechnology and Bioengineering* 2000;69(4):351-8.
- Götting C, Thierbach G, Pühler A, Kalinowski J. Versatile low-copy- number plasmids for temperature inducible overexpression of bacterial genes in *Escherichia coli*. *BioTechniques* 1998;24:362-6.
- Gronenborn B. Overproduction of phage lambda repressor under control of the lac promoter of *Escherichia coli*. *Mol. Gen. Genet.* 1976;148:243-50.
- Chermojovsky Y, Mory Y, Vaks B, Funstein SI, Segev D, Revel M. Production of human interferon in *E. coli* under lac and tryplac promoter control. *Ann NY Acad. Sci* 1983;413:88-96.
- Sørensen HP, Mortensen KK. Advanced genetics strategy for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology* 2005;115:113-28.
- Jonasson P, Liljeqvist S, Nygren P, Stahl S. Genetic design for facilitated production and recovery of recombinant proteins in *E coli* *Biotechnol. Appl. Biochem* 2002;35:91-105.
- Organización Mundial de la Salud. Serie de Informes Técnicos. Anexo 1 "Prácticas adecuadas para la fabricación de productos biológicos". Ginebra: OMS: 1992.
- ICH. Quality of Biotechnological Products: Analysis of the expression construct in cells used for production of rDNA derived products (Q5B) 1997. Disponible en: <http://www.ich.org/cache/compo/276-254-1.html>.
- ISO/TS 11133-1. Part 2 "Practical guidelines on performance testing of culture media". Sydney, Australia: ISO; 2009.
- Tran QH, Uden G. Changes in the proton potential and the cellular energetics of *Escherichia coli* during growth by aerobic and anaerobic respiration or by fermentation. *Eur J Biochem* 1998;251(1-2):538-43.
- Konstantin B, Naoki N, Toshiomi Y. Glucose feeding strategy accounting for the decreasing oxidative capacity of recombinant *E. coli* in fed-batch cultivation for phenylalanine production. *Journal of Fermentation and Bioengineering* 1990;70(4):253-60.

28. Monod J. La technique de culture continue, théorie et applications. *Ann Inst Past* 1950;79:390-410.
29. Kleman G, Strohl W. Acetate metabolism by *Escherichia coli* in high-cell-density fermentation. *Appl Environ. Microbiol.* 1994;60(11) 3952-8.
30. Lee SY. High cell density culture of *Escherichia coli*. *Tibtech* 1996;14:98-105.
31. Makrides SC. Strategies for achieving high level expression of genes in *Escherichia coli*. *Microbiol Rev* 1996;60(3):512-38.
32. Shimizu N, Fukusono S, Harada Y, Fujimori K, Gotoh K, Yamasaki Y. Mass production of human epidermal growth factor using fed-batch cultures of recombinant *Escherichia coli*. *Biotechnol Bioeng* 1991;38(1):37-42.
33. Markl H, Dubach AC, Ogbonna JC. Cultivation of *Escherichia coli* to high cell densities in a dialysis reactor. *Appl Microbiol Biotechnol* 1993;39(1):48-52.
34. Babu KR. Production of interferon alpha in high cell density cultures of recombinant *Escherichia coli* and its single step purification from refolded inclusion body proteins. *Appl Microbiol Biotechnol* 2000;53(6):655-60.
35. Manderson D, Robert A, Dempster A, Chisti Y. A recombinant vaccine against hydatidosis: production of the antigen in *Escherichia coli*. *J Ind Microbiol Biotechnol* 2006;33:173-82.
36. Luli G W, Strohl WR. Comparison of growth, acetate production and acetate inhibition of *Escherichia coli* strains in batch and fed batch fermentations. *Appl Environ Microbiol* 1990;56: 1004-11.
37. Tripathi NK, Sathyaseelan K, Jana AM, Rao PVL, Kang WK, Park TH. High Yield Production of Heterologous Proteins with *Escherichia coli* Defence. *Science Journal* 2009;59(2):137-46.
38. Lin HY. Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited-fed batch cultivation of *Escherichia coli*. *Biotechnol Bioeng* 2001;75(5):347-57.
39. Akesson M, Hagander P, Axelsson J. Avoiding acetate accumulation in *Escherichia coli* cultures using feedback control of glucose feeding. *Biotechnol Bioeng* 2001;73(3): 223-30.
40. Du P, Ye Q, Yu JT. Cultivation integrated with acetate filtration on *Escherichia coli*. *Sheng Wu Gong Cheng Xue Bao;* 2000;16(4):528-30.
41. Zhang WC. High cell density culture of phosphotransacetylase mutants of *Escherichia coli* BL21(DE3). *Sheng Wu Gong Cheng Xue Bao* 2001;17(1):59-63.
42. Beckley K, Verhaert P, van der Wielen LAM, Hubbuch J, Ottens M. Rational and systematic protein purification process development: the next generation. *Trends Biotechnol* 2009;27(12):673-79.
43. Gómez R, Madrazo J, González L, China G, Musacchio A, Rodríguez A, et al. Caracterización estructural y funcional de la proteína recombinante P64k de *Neisseria meningitidis*. *Biología Aplicada* 1999;16(2):83-7.
44. Ward W, Swiatek G. Protein Purification. *Current Analytical Chemistry* 2009;5(2):1-21.
45. Purifying Challenging Proteins. Principles and Methods. Uppsala, Sweden: GE Healthcare Bio-Sciences; 2007.
46. Middelberg APJ. Preparative protein refolding. *Trends in Biotechnology* 2002; 20(10):437-43.
47. Geng X, Wang C. Protein folding liquid chromatography and its recent developments. *Journal of Chromatography B.* 2007;849(1):69-80.
48. Lilie H, Schwarz E, Rudolph R. Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology* 1998; 9(5):497-501.
49. Curling J. Process Chromatography: Five Decades of Innovation. *BioPharm Int* 2007;20 (Suppl. 1):10-20.
50. Chen R, Huang C, Newton B, Ritter G, Old L, Batt C. Factors affecting endotoxin removal from recombinant therapeutic proteins by anion exchange chromatography. *Prot Exp Purif* 2009;64:76-81.

Strategies to obtain recombinant proteins in *Escherichia coli*

Abstract

Recombinant protein expression has been favored by the use of *Escherichia coli* due to its relatively low costs, high-density culture, easy genetic manipulation and also to compatible biotechnological tools. Strategies for recombinant expression in *E. coli* are presented in this paper. Genetic and physiologic factors are presented: number of copies of the expression vector, characteristic of the gene, ribonucleic acid (messenger) stability, promoter, host strain, clone stability, composition of the cultivation media, operation parameters in the bioreactor and cultivation and purification strategies.

Key words: recombinant protein, *Escherichia coli*, genetic factors, physiologic factors, fermentation.
