



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

**Caracterización de la expresión de chaperonas y propiedades estructurales
de los cuerpos de inclusión de la proteína rHuGM-CSF producida en un
sistema termoinducibile en *Escherichia coli***

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTORA EN CIENCIAS

PRESENTA:

SARA RESTREPO PINEDA

TUTOR PRINCIPAL DE TESIS: DR. MAURICIO ALBERTO TRUJILLO ROLDÁN
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

COMITÉ TUTOR: DR. ENRIQUE GARCÍA HERNÁNDEZ
INSTITUTO DE QUÍMICA, UNAM
DR. GEORGES DREYFUS CORTÉS (†)
INSTITUTO DE FISIOLÓGÍA CELULAR, UNAM

Ciudad Universitaria, CD. MX., 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

**Caracterización de la expresión de chaperonas y propiedades estructurales
de los cuerpos de inclusión de la proteína rHuGM-CSF producida en un
sistema termoinducible en *Escherichia coli***

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTORA EN CIENCIAS

PRESENTA:

SARA RESTREPO PINEDA

TUTOR PRINCIPAL DE TESIS: DR. MAURICIO ALBERTO TRUJILLO ROLDÁN
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

COMITÉ TUTOR: DR. ENRIQUE GARCÍA HERNÁNDEZ
INSTITUTO DE QUÍMICA, UNAM
DR. GEORGES DREYFUS CORTÉS (†)
INSTITUTO DE FISIOLÓGÍA CELULAR, UNAM

Ciudad Universitaria, CD. MX., 2023



COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

OFICIO CPCB/242/2023

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **28 de noviembre de 2022** se aprobó el siguiente jurado para el examen de grado de **DOCTORA EN CIENCIAS** de la estudiante **RESTREPO PINEDA SARA** con número de cuenta **514353063** con la tesis titulada "**Caracterización de la expresión de chaperonas y propiedades estructurales de los cuerpos de inclusión de la proteína rHuGM-CSF producida en un sistema termoinducible en *Escherichia coli***", realizada bajo la dirección del **DR. MAURICIO ALBERTO TRUJILLO ROLDÁN**, quedando integrado de la siguiente manera:

Presidente: DR. ALEJANDRO ALAGÓN CANO
Vocal: DRA. NURIA VICTORIA SÁNCHEZ PUIG
Vocal: DR. OCTAVIO TONATIUH RAMÍREZ REIVICH
Vocal: DRA. MARCELA AYALA ACEVES
Secretario: DR. ENRIQUE GARCÍA HERNÁNDEZ

Sin otro particular, me es grato enviarle un cordial saludo.

A T E N T A M E N T E
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 08 de marzo de 2023

COORDINADOR DEL PROGRAMA

DR. ADOLFO GERARDO NAVARRO SIGÜENZA



Agradecimientos Institucionales

✦ Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México, UNAM.

✦ Al apoyo económico recibido por el Consejo Nacional de Ciencia y Tecnología (CONACyT) para realizar mis estudios (número de beca: 589949). Así como a los apoyos económicos recibidos por el Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica de la Universidad Nacional Autónoma de México (PAPITT)-UNAM: IN211422, IN210822, IV201220.

✦ A mi tutor principal, el Dr. Mauricio Alberto Trujillo Roldán y a los estimados miembros del Comité Tutor, el Dr. Enrique García Hernández y el Dr. Georges Dreyfus Cortés (†).

Agradecimientos personales

- ☆ Al Dr. Mauricio Trujillo Roldán, por permitirme ser parte de su grupo de investigación; por ser mi mentor en este proyecto, por confiar en mi trabajo, por inspirarme y ayudarme a crecer como profesional y como persona.

- ☆ A la Dra. Adriana Valdez Cruz, por sus consejos y sugerencias para el desarrollo de este trabajo, por estar siempre dispuesta a ayudarme y por su constante apoyo en este camino.

- ☆ A los miembros de mi comité tutorial por su tiempo, su paciencia y sus valiosos comentarios que enriquecieron mi trabajo. Gracias al Dr. Enrique García por ser parte de mi proceso de formación, por convertirse en un gran amigo y colega. Gracias al Dr. Georges Dreyfus por siempre tener las palabras correctas para hacerme ver más allá de lo obvio...en donde quiera que esté, lo llevaré en mis memorias.

- ☆ A los miembros del jurado, la Dra. Marcela Ayala Aceves, a la Dra. Nuria Sánchez Puig, al Dr. Enrique García Hernández, al Dr. Octavio Tonatiuh Ramírez y al Dr. Alejandro Alagón Cano por su amabilidad, su atención y su tiempo para evaluar mi tesis.

- ☆ Al apoyo técnico del Dr. Axel Luviano en los experimentos de dicroísmo circular; al Dr. Luis Pablo Ávila Barrientos y Diego Rosiles Becerril por su ayuda en los ensayos de replegamiento y HPLC.

- ☆ A mi familia, que lo son todo para mí. Su presencia es mi aliento y mi aliciente.
A mis papás: por su amor y cuidado, por su sabiduría y fortaleza, por enseñarme a ser una persona responsable y disciplinada que nunca abandona lo que se propone.
A mis hermanas: por sus consejos, por creer en mí y por alentarme a perseguir mis sueños.

- ☆ A mis amigos incondicionales de siempre: Sara, Stephy, Puppy, Karla y todos aquellos que a pesar de la distancia y el tiempo siguen estando presentes en mi vida.

- ☆ A mis queridos amigos del laboratorio: Giro, Ram Gamboa, Ram García, Greta, Dani, Bandita, Sau, Diegui, Richie, Pato...porque se convirtieron en mi familia mexicana, me llenaron de cariño e hicieron más felices mis días. Al Inge Abel por su noble corazón y su amistad.

CONTENIDO

1. INTRODUCCIÓN	- 3 -
1.1. Producción de proteínas recombinantes en <i>E. coli</i>	- 3 -
2. ANTECEDENTES	- 5 -
2.1. Cuerpos de inclusión	- 5 -
2.2. Sistema de expresión termoinducible $\lambda pL/pR-cl857$	- 10 -
2.3. Respuesta de choque térmico en <i>E. coli</i>	- 13 -
2.3.1. Mecanismos de las chaperonas moleculares	- 16 -
2.3.2. Convergencia de las respuestas de estrés asociadas con la producción de proteínas recombinantes termoinducidas	- 18 -
2.4. Proteína modelo: Factor Estimulante de Colonias de Granulocitos y Macrófagos	- 20 -
3. JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA	- 23 -
4. HIPÓTESIS	- 24 -
5. OBJETIVOS	- 24 -
5.1. Objetivo general	- 24 -
5.2. Objetivos específicos	- 24 -
6. METODOLOGÍA	- 25 -
6.1. Materiales	- 25 -
6.1.1. Cepas bacterianas	- 25 -
6.1.2. Plásmido	- 25 -
6.1.3. Medio de cultivo	- 25 -
6.2. Métodos	- 27 -
6.2.1. Generación del banco maestro de la cepa <i>E. coli</i> W3110 productora de rHuGM-CSF	- 27 -
6.2.2. Generación del banco de trabajo de la cepa <i>E. coli</i> W3110 productora de rHuGM-CSF	- 28 -
6.2.3. Cinética de crecimiento de <i>E. coli</i> W3110 productora de rHuGM-CSF en matraces agitados	- 28 -
6.2.4. Preparación de pre-inóculos para los cultivos en biorreactor	- 28 -
6.2.5. Cinética de crecimiento de <i>E. coli</i> W3110 productora de rHuGM-CSF en biorreactores	- 29 -
6.2.6. Medición de la absorbancia del cultivo	- 29 -

6.2.7.	Determinación de la biomasa como peso seco _____	- 29 -
6.2.8.	Cuantificación de glucosa _____	- 30 -
6.2.9.	Cuantificación de acetato _____	- 30 -
6.2.10.	Recuperación y purificación de los cuerpos de inclusión _____	- 30 -
6.2.11.	Cuantificación de proteína _____	- 31 -
6.2.12.	Análisis de proteína por electroforesis en geles de poliacrilamida _____	- 31 -
6.2.13.	Inmunodetección de chaperonas por Western Blot _____	- 31 -
6.2.14.	iTRAQ: Marcaje isobárico para la cuantificación absoluta y relativa de proteínas _____	- 32 -
6.2.15.	Espectrometría de masas acoplada a HPLC _____	- 33 -
6.2.16.	Espectroscopia infrarroja por transformada de Fourier (FTIR) _____	- 33 -
6.2.17.	Ensayo de fluorescencia con Tioflavina T _____	- 34 -
6.2.18.	Digestión proteolítica con proteinasa K _____	- 34 -
6.2.19.	Estabilidad en cloruro de guanidinio (GndHCl) _____	- 34 -
6.2.20.	Solubilización, plegamiento y purificación _____	- 34 -
6.2.21.	Espectroscopia de dicroísmo circular (DC) _____	- 35 -
6.2.22.	Análisis estadísticos _____	- 35 -
7.	RESULTADOS _____	- 36 -
7.1.	Cinética de crecimiento de <i>E. coli</i> recombinante en matraces agitados _____	- 36 -
7.2.	Expresión de la proteína rHuGM-CSF en matraces agitados _____	- 38 -
7.3.	Cinética de crecimiento de <i>E. coli</i> recombinante en biorreactores _____	- 39 -
7.4.	Cinética de crecimiento de la cepa silvestre <i>E. coli</i> W3110 _____	- 44 -
7.5.	Expresión de la proteína rHuGM-CSF en biorreactores _____	- 45 -
7.5.1.	Comparación de la expresión en las fracciones de proteína total _____	- 45 -
7.5.2.	Comparación de la expresión de la proteína en cuerpos de inclusión _____	- 46 -
7.6.	Identificación de chaperonas DnaK, DnaJ, GroEL y GroES _____	- 50 -
7.7.	Proteómica parcial de los CI de rHuGM-CSF usando iTRAQ _____	- 53 -
7.8.	Caracterización estructural de los CI de la proteína rHuGM-CSF _____	- 55 -
7.8.1.	Análisis de la estructura secundaria por ATR-FTIR _____	- 55 -
7.8.2.	Ensayo de fluorescencia con tioflavina T (Th-T) _____	- 57 -
7.8.3.	Degradación proteolítica con proteinasa K (PK) _____	- 60 -
7.8.4.	Solubilización con cloruro de guanidinio _____	- 61 -
7.9.	Purificación de rHuGM-CSF plegada y análisis por dicroísmo circular _____	- 61 -

8. DISCUSIÓN	- 64 -
8.1. Cinética de crecimiento de <i>E. coli</i> recombinante en matraces agitados	- 64 -
8.2. Expresión de la proteína rHuGM-CSF en matraces agitados	- 64 -
8.3. Cinética de crecimiento de <i>E. coli</i> recombinante en biorreactores	- 65 -
8.4. Expresión de la proteína rHuGM-CSF en biorreactores	- 68 -
8.5. Identificación de chaperonas DnaK, DnaJ, GroEL y GroES	- 69 -
8.6. Caracterización estructural de los CI de la proteína rHuGM-CSF	- 71 -
8.7. Purificación de rHuGM-CSF plegada y análisis por dicroísmo circular	- 73 -
9. CONCLUSIONES	- 75 -
10. REFERENCIAS BIBLIOGRÁFICAS	- 76 -
11. ANEXOS	- 96 -

LISTA DE FIGURAS

Figura 1. Cuerpos de inclusión de *E. coli*. **(A)** Secciones transversales de la bacteria produciendo un factor estimulante de colonias de granulocitos (G-CSF) bajo microscopio electrónico de transmisión. **(B)** Cuerpos de inclusión aislados por disrupción celular en un homogeneizador y lavados en agua pura bajo microscopio electrónico de barrido (modificado de Peternel *et al.*, 2008).

Figura 2. Producción de cuerpos de inclusión y sus aplicaciones. La sobre-expresión de proteínas recombinantes en *E. coli* conduce comúnmente a la formación de cuerpos de inclusión (CI). Los CI tienen una composición heterogénea de proteína (intermediarios parcialmente plegados, agregados amorfos, conformaciones de tipo amiloide y proteína de tipo nativa) y se encuentran interaccionando con diversas chaperonas citosólicas. Se ha estudiado su potencial como materiales para proliferación celular en ingeniería de tejidos, como medios para inmovilizar enzimas, como agentes de liberación de proteínas terapéuticas y, principalmente, como materia prima para la recuperación de proteínas biológicamente activas.

Figura 3. Regulación molecular del sistema de expresión termoinducible $\lambda pL/pR$ -cl857 de *E. coli*. **(A)** Durante el crecimiento celular a temperaturas por debajo de los 37 °C, el represor cl857 (en morado) se encuentra unido a través de su dominio amino terminal a las regiones operadoras (*oL1/oR1*, *oL2/oR2* y *oL3/oR3*) de los promotores *pL* y *pR*, reprimiendo la transcripción del gen de la proteína recombinante (en amarillo). **(B)** Un aumento en la temperatura por encima de los 37 °C, provoca la liberación del represor cl875, permitiendo el reconocimiento de la ARN polimerasa (en verde) para activar la transcripción del gen recombinante (adaptado de Restrepo-Pineda *et al.*, 2021).

Figura 4. Modulación de la respuesta de choque térmico (HSR) en *E. coli*. **(A)** A bajas temperaturas de crecimiento (30 °C), se mantienen en las células niveles basales del factor σ^{32} , a través de diferentes mecanismos: 1. La regulación traduccional del ARNm del gen *rpoH*, 2. La regulación por chaperonas o 3. La degradación por proteasas. **(B)** Por encima de 34 °C, hay un aumento rápido en la traducción del factor σ^{32} al liberarse la estructura del ARNm de *rpoH*, desencadenando el incremento en las velocidades de síntesis de las proteínas de choque térmico y el encendido de la respuesta de choque térmico.

Figura 5. Participación de las chaperonas moleculares en la regulación de la homeostasis de las proteínas en el citoplasma de *Escherichia coli*. Los polipéptidos nacientes del ribosoma interactúan con el factor TF y con el sistema de chaperonas de plegamiento DnaK/DnaJ/GrpE para alcanzar una conformación nativa. Si esto no ocurre, los intermediarios parcialmente plegados son rescatados por un segundo sistema de plegamiento compuesto por GroEL y GroES o enviados a desagregación por ClpB. En el caso de la sobreexpresión de proteínas recombinantes (líneas rojas punteadas) o condiciones severas de estrés (líneas rojas sólidas), la capacidad de las chaperonas es limitada, requiriendo la ayuda de un sistema de proteínas de estabilización que incluye a IbpA e IbpB (modificada de Kwon *et al.*, 2016).

Figura 6. Relación entre los eventos moleculares que ocurren en *E. coli* durante la producción de una proteína recombinante usando el sistema de expresión termoinducible $\lambda pL/pR-cl857$ (modificado de Restrepo-Pineda *et al.*, 2021).

Figura 7. Representación de superficie **(A)** y caricatura **(B)** de la proteína humana GM-CSF con base en la estructura disponible en el Protein Data Bank (código PDB: 2GMF; Rozwarski *et al.*, 1996). Los colores representan los resultados del análisis de perturbación de desplazamiento químico, en inglés chemical shift perturbation (csp). Los residuos perturbados significativamente se colorean en gradiente de amarillo a rojo (modificado de Blech *et al.*, 2012).

Figura 8. Mapa del plásmido pV3/HuGM-CSF y ubicación de sus elementos genéticos.

Figura 9. Cinética de crecimiento de la cepa *E. coli* W3110 + rHuGM-CSF en matraces agitados de 250 ml con 50 ml de medio definido. Se presenta el cambio de la densidad óptica (DO) a 600 nm en escala exponencial **(A)** y en escala logarítmica **(B)** con respecto al tiempo. Cultivos sin inducir a 30 °C (●) y cultivos inducidos a 42 °C (○). La línea punteada indica el momento de la termoinducción, a una DO_{600} de ~0.9-1.2 u.a. (13 h).

Figura 10. Cinética de crecimiento de la biomasa de *E. coli* W3110 + rHuGM-CSF en matraces agitados de 250 ml con 50 ml de medio definido. Se presenta el cambio de la biomasa en términos del peso seco celular para el cultivo sin inducir a 30 °C (●) y cultivo inducido a 42 °C (○). La línea punteada indica el momento de la termoinducción.

Figura 11. Expresión de la proteína rHuGM-CSF (~14-15 kDa) en *E. coli* W3110 usando matraces agitados a las 24 h de cultivo. Las proteínas se separaron en un gel de poliacrilamida al 15 % teñido con Azul de Coomassie. Carril 1: Marcador de peso molecular (MW); carril 2: Cuerpos de inclusión (CI) del cultivo termoinducido a 42 °C; carril 3: Proteína total soluble (S) del cultivo termoinducido a 42 °C; carril 4: Proteína insoluble (I) del cultivo sin inducir y carril 5: Proteína soluble (S) del cultivo sin inducir.

Figura 12. Cinética de crecimiento de la cepa *E. coli* W3110 + rHuGM-CSF sin termoinducción en biorreactores de 1.2 L con 800 ml de medio definido. Se presenta el cambio en la DO a 600 nm tanto en escala lineal **(A)** como en escala logarítmica **(B)**. Los cultivos se mantuvieron a 30 °C (círculos negros) o a 34 °C (círculos blancos) a lo largo de todo cultivo.

Figura 13. Cinética del crecimiento bacteriano **(A, B)**, consumo de glucosa **(C,D)**, y la producción de acetato **(E, F)** de *E. coli* W3110 + rHuGM-CSF en biorreactores de 1.2 L termoinducidos. Dos temperaturas de pre-inducción fueron evaluadas: 30 °C (círculos llenos) y 34 °C (círculos abiertos) con posterior termoinducción a 42 °C. Las líneas punteadas indican el inicio del aumento de temperatura (7 h para 30 °C y 4 h para 34 °C), que corresponden a una DO de ~2.0 u.a. Se muestra el promedio y la desviación estándar de tres experimentos independientes.

Figura 14. Comportamiento de la tensión de oxígeno disuelto **(A, B)** y pH **(C, D)** de cultivos de *E. coli* W3110 + rHuGM-CSF en biorreactores de 1.2 L. Dos temperaturas de pre-

inducción fueron evaluadas: 30 °C (panel izquierda) y 34 °C (panel derecha) con posterior termoinducción a 42 °C. Las líneas punteadas indican el inicio del aumento de temperatura (7 h para 30 °C y 4 h para 34 °C), que corresponden a una DO de ~2.0 u.a. Se muestra el promedio y la desviación estándar de tres experimentos independientes.

Figura 15. Cinética de crecimiento de la cepa silvestre *E. coli* W3110 en biorreactores de 1.2 L con 800 ml de medio definido a 30 °C. Se presenta el cambio en la DO a 600 nm tanto en escala lineal **(A)** como en escala logarítmica **(B)**, representando el promedio de tres réplicas biológicas con sus respectivas desviaciones estándar.

Figura 16. Lisados celulares totales de *E. coli* W3110 + rHuGM-CSF obtenidos en biorreactores a 30 °C **(A)** y 34 °C **(B)** a diferentes tiempos después de la inducción a 42 °C. Las proteínas separaron en un gel de poliacrilamida al 15 % teñido con azul de Coomassie, que fue cargado volumétricamente. Las flechas señalan una banda entre ~14-15 kDa, que corresponde presumiblemente a rHuGM-CSF.

Figura 17. Expresión de la proteína rHuGM-CSF (~14-15 kDa) en la proteína soluble y los cuerpos de inclusión de *E. coli* W3110 recombinante usando biorreactores. Dos temperaturas de pre-inducción diferentes fueron evaluadas: 30 °C **(A)** y 34 °C **(B)** en un gel de poliacrilamida al 15 % teñido con azul de Coomassie. Las muestras corresponden a la proteína soluble (S) y los cuerpos de inclusión (CI) a diferentes tiempos después de la termoinducción (1 h, 5 h y 18/20 h). Se usaron como controles negativos la cepa sin inducir (*rE. coli* a 30 °C o 34 °C) y la cepa silvestre (WT *E. coli*). El gel se cargó volumétricamente. Las flechas señalan una banda entre ~14-15 kDa, que corresponde presumiblemente a rHuGM-CSF.

Figura 18. Expresión de la proteína rHuGM-CSF (~14-15 kDa) en los cuerpos de inclusión de *E. coli* W3110 recombinante usando biorreactores. Dos temperaturas de pre-inducción diferentes fueron evaluadas: 30 °C (A) y 34 °C (B) en un gel de poliacrilamida al 15 % teñido con azul de Coomassie. Se presentan los cuerpos de inclusión (CI) con 20 µg de proteína a diferentes tiempos después de la termoinducción (1 h, 3 h, 5 h y 18/20 h). Se usaron como controles negativos, la cepa sin inducir (*rE. coli* a 30 °C o 34 °C) y la cepa silvestre (WT *E. coli*). El control positivo es el estándar de referencia para Molgramostim (Molgra).

Figura 19. Concentración de la proteína celular total **(A)**, la proteína en los CI **(B)** y rHuGM-CSF en los CI **(C)** de cultivos de *E. coli* W3110 recombinante bajo diferentes temperaturas de pre-inducción: 30 °C (barras negras) y 34 °C (barras blancas) con posterior termoinducción a 42 °C. Se muestra el promedio y la desviación estándar para tres repeticiones biológicas.

Figura 20. Inmunodetección de las chaperonas GroEL, GroES, DnaK y DnaJ en los CI de la proteína rHuGM-CSF provenientes de cultivos bajo diferentes temperaturas de crecimiento: 30 °C **(A)** y 34 °C **(B)**, termoinducidos a 42 °C. Carriles 1A y 1B: lisados de la cepa recombinante sin inducir a 30 °C o 34 °C, respectivamente. Carriles 2A y 2B: lisados de la cepa silvestre. Carril MW: marcador de peso molecular. Se presentan los CI a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h) bajo las dos condiciones evaluadas.

Las flechas indican las bandas correspondientes a DnaK (~70 kDa, panel 1), DnaJ (~41 kDa, panel 2), GroEL (~60 kDa, panel 3) y GroES (~15 kDa, panel 4).

Figura 21. Inmunodetección de GroES y DnaJ en la proteína total soluble proveniente de cultivos bajo diferentes temperaturas de crecimiento: 30 °C (**A, C**) y 34 °C (**B, D**), termoinducidos a 42 °C. Carril 1: lisados de *E. coli* W3110 recombinante sin inducir a 30 °C o 34 °C, respectivamente; Carril 2: lisados de la cepa *E. coli* W3110 silvestre; Carril MW: marcador de peso molecular. Se muestra la proteína total soluble a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h) a las dos condiciones evaluadas. Las flechas indican las bandas correspondientes a GroES en ~15 kDa (**A, B**) y DnaJ en ~41 kDa (**C, D**).

Figura 22. Segunda derivada de los espectros de absorción obtenidos por ATR-FTIR para los CI provenientes de cultivos pre-inducidos a 30 °C (**A**) y 34 °C (**B**). Los CI fueron cosechados en diferentes momentos después de la termoinducción a 42 °C: 1 h (línea sólida), 3 h (línea guión), 5 h (línea punto), 10 h (línea guión-punto) y 18/20 h (línea doble guión-punto). Los espectros representan el promedio de tres réplicas biológicas y fueron normalizados con respecto al pico de tirosina (~1508 cm⁻¹). Se señalan los picos correspondientes a las estructuras hélice α "random coil" (~1654 cm⁻¹), hojas β (~1637 cm⁻¹) y agregados (~1625 cm⁻¹).

Figura 23. Comparación de los mínimos de la segunda derivada correspondientes a hélices α "random coil" (**A**), hojas β (**B**) y agregados amiloides (**C**) de los CI provenientes de cultivos preinducidos a 30 °C (barras negras) y a 34 °C (barras blancas). Los CI fueron recolectados a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h). El promedio y la desviación estándar son mostrados para tres experimentos independientes.

Figura 24. Espectros de fluorescencia de la unión de Th-T a los CI de rHuGM-CSF, obtenidos bajo dos temperaturas de pre-inducción: 30 °C (**A**) y 34 °C (**B**) después de la termoinducción. Los CI se recolectaron a diferentes tiempos post-inducción: 1 h (línea sólida delgada), 3 h (línea guión-punto), 5 h (línea punto), 10 h (línea guión) y 18/20 h (línea sólida gruesa). El espectro del colorante Th-T sin proteína se utilizó como control (línea punteada cercana a cero) y se presenta el promedio de dos réplicas biológicas por condición.

Figura 25. Cinética de la digestión proteolítica con proteínasa K (PK) de los CI de rHuGM-CSF obtenidos a partir de cultivos pre-inducidos a 30 °C (**A**) y 34 °C (**B**) y posterior termoinducción a 42 °C. Los IB se recolectaron en diferentes momentos después de la inducción: 1 h (cuadrados negros), 3 h (cuadrados blancos), 5 h (triángulos), 10 h (círculos blancos) y 18/20 h (círculos negros). Para la degradación se empleó PK a una concentración final de 25 μ g/ml. Se presentan los promedios normalizados de dos réplicas biológicas.

Figura 26. Perfiles de solubilización de los CI de rHuGM-CSF obtenidos de cultivos preinducidos a 30 °C (barras negras) y 34 °C (barras grises) y posterior termoinducción a 42 °C (1, 3, 5, 10 y 18/20 h). La gráfica muestra la cantidad de proteína solubilizada (mg/ml) después de 24 h de incubación con diferentes concentraciones de GndHCl: 1 M (**A**), 3 M (**B**) y 5 M (**C**). Se presentan los promedios y la desviación estándar de tres experimentos independientes.

Figura 27. Purificación de rHuGM-CSF por HPLC de fase reversa **(A)**. La solución de proteína plegada (~ 10 µg) proveniente de cultivos creciendo a 30 °C (línea continua) o a 34 °C (línea punteada) después la termoinducción a 42 °C fue analizada por HPLC con una longitud de onda del detector de 214 nm. Curva estándar (en el recuadro superior) del estándar de referencia para el GM-CSF humano de la Farmacopea Europea (Molgramostim). Espectro de dicroísmo circular en el UV lejano de la proteína rHuGM-CSF purificada **(B)**. Se muestra el promedio de tres escaneos, tanto para la proteína derivada de CI obtenidos de cultivos creciendo a 30 °C (círculos negros) o a 34 °C (círculos blancos) después de la termoinducción a 42 °C.

LISTA DE TABLAS

Tabla 1. Efecto de los parámetros del cultivo en las propiedades de los cuerpos de inclusión.

Tabla 2. Elementos genéticos del plásmido pV3.

Tabla 3. Composición del medio de cultivo definido (modificado de Caspeta *et al.*, 2013).

Tabla 4. Anticuerpos anti-chaperonas de *E. coli* usados para realizar Western blot de cultivos termoinducidos.

Tabla 5. Comparación de los parámetros cinéticos de *E. coli* W3110 + rHuGM-CSF en biorreactores a 30 °C y 34 °C con posterior termoinducción a 42 °C.

Tabla 6. Análisis densitométrico de la proteína rHuGM-CSF en la proteína total y en los CI a diferentes tiempos después de la termoinducción.

Tabla 7. Comparación de los rendimientos de la proteína rHuGM-CSF bajo diferentes temperaturas de pre-inducción, 30 °C y 34 °C, con posterior termoinducción a 42 °C.

Tabla 8. Proteínas identificadas en los CI usando Proteome Discoverer™ 2.1 SP1 (*Thermo Fisher Scientific, USA*).

ABREVIATURAS

AC	Acetato
ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
ARNm	Ácido ribonucleico mensajero
ATCC	American Type Culture Collection
ATR	Reflexión total atenuada
BSA	Albúmina de suero bovino
CI	Cuerpos de inclusión
CDW	Cell dry weight (Biomasa en peso seco)
DO	Densidad Óptica
DTT	Ditiotreitol
FRET	Transmisión de energía de resonancia de fluorescencia
FTIR	Espectrometría infrarroja por transformada de Fourier
GLC	Glucosa
GndHCl	Cloruro de guanidinio
HCl	Ácido clorhídrico
hFGF-2	Factor de crecimiento de fibroblastos humano 2
HPLC	Cromatografía líquida de alta resolución
HSPs	Proteínas de choque térmico
HSR	Respuesta de choque térmico
HuGM-CSF	Factor estimulante de colonias de granulocitos y macrófagos humano
IL	Interleucina
IPTG	Isopropil- β -D-1-tiogalactopiranosido
iTRAQ	Etiquetas isobáricas para cuantificación relativa y absoluta
kb	Kilobases
kDa	Kilodaltones
L	Litro
LB	Medio Luria Bertani
M	Molar
mg	Miligramos
min	Minutos
ml	Mililitros
mM	Milimolar
N°	Número
nm	Nanómetros
NaCl	Cloruro de Sodio
NaOH	Hidróxido de Sodio
N-terminal	Amino terminal de una cadena peptídica
qP	Velocidad específica de formación de acetato

q_s	Velocidad específica de consumo de sustrato
PBS	Amortiguador de fosfatos y sal
pH	Potencial de hidrógeno
PVDF	Fluoruro de polivinilideno
rpm	Revoluciones por minuto
SDS-PAGE	Electroforesis en gel de poliacrilamida con dodecilsulfato sódico
TBS	Amortiguador Tris-boratos
TOD	Tensión de oxígeno disuelto
u.a.	Unidades de absorbancia
vvm	Volumen de aire por volumen de líquido por minuto
Y_{AC/X}	Rendimiento de acetato por biomasa
Y_{X/S}	Rendimiento de biomasa por sustrato

RESUMEN

Cuando se sobre-expresan proteínas recombinantes en *Escherichia coli* existe una tendencia generalizada de la bacteria a formar agregados proteicos llamados cuerpos de inclusión. Son partículas insolubles y electrodensas que están enriquecidas de proteína recombinante. La formación de los cuerpos de inclusión se ha convertido en un obstáculo en aquellos casos donde el propósito es obtener proteína soluble y, por lo tanto, diferentes estrategias se han formulado para evitarlos. Sin embargo, en las últimas décadas se ha descubierto la posibilidad de recuperar proteína con actividad biológica a partir de estos agregados, lo que ha incrementado el interés científico en comprender su organización molecular y su potencial uso en diversos campos de la biomedicina y biotecnología. El efecto de las condiciones del bioproceso en los rendimientos de recuperación y en las características estructurales de los cuerpos de inclusión también se ha estudiado.

A nivel industrial, una de las estrategias ampliamente desarrolladas para la producción de proteínas recombinantes, es el sistema de expresión termoinducible $\lambda pL/pR-cl857$ de *E. coli*, que se basa en una etapa de crecimiento celular a 30 °C y una etapa posterior de inducción por encima de 37 °C. Las ventajas de este sistema incluyen la regulación fina de los promotores, se evita la adición de inductores químicos y, se minimizan riesgos de contaminación. A pesar de esto, el incremento en la temperatura junto con la producción de una proteína recombinante, pueden provocar cambios en el metabolismo y fisiología celular, desencadenar la respuesta de choque térmico y conducir a la formación de cuerpos de inclusión. Dichos eventos moleculares que ocurren en un sistema inducido por temperatura se han estudiado en contextos separados, pero coinciden en la participación de chaperonas y proteasas, que se encargan de mantener la homeostasis de las proteínas.

Basándonos en que un aumento de la temperatura por encima de los 34 °C induce en *E. coli* la respuesta de choque térmico, aquí nos planteamos estudiar el efecto de la temperatura de pre-inducción (30 °C vs 34 °C) en los rendimientos de proteína recombinante, la expresión de chaperonas y la estructura de los cuerpos de inclusión. La proteína modelo fue el factor estimulante de colonias de granulocitos y macrófagos humano (rHuGM-CSF) de gran relevancia terapéutica y con tendencia a la agregación. Se llevaron a cabo cultivos de *E. coli* recombinante en biorreactores a 34 °C y 30 °C, los cuales fueron inducidos posteriormente a 42 °C. La temperatura de pre-inducción de 34 °C aceleró el crecimiento bacteriano, presentando un incremento de hasta el 69% en la velocidad específica de crecimiento y favoreciendo la acumulación de rHuGM-CSF en los cuerpos de inclusión, en comparación con los cultivos que crecieron a 30 °C. Posteriormente, los estudios de expresión de las chaperonas usando inmunodetección, permitieron identificar la presencia de las dos principales chaperonas de plegamiento, DnaK y GroEL, en los agregados, mientras que sus co-chaperonas, DnaJ y GroES, se encontraron en la fracción soluble. Tras la evaluación de las propiedades estructurales de los cuerpos de inclusión usando el ensayo de unión a tioflavina-T y espectrometría infrarroja por transformada de Fourier (FTIR), se encontró que los cuerpos de inclusión de cultivos que crecieron a una temperatura de pre-inducción de 30 °C presentaron mayor contenido estructural de tipo amiloide y fueron más resistentes a la proteólisis por proteinasa K. Del trabajo realizado se concluyó que una mayor temperatura de pre-inducción (34 en vez de 30 °C) en un sistema termoinducido, permite obtener mayor cantidad de proteína recombinante así como reducir el contenido amiloide de los cuerpos de inclusión.

ABSTRACT

When recombinant proteins are overexpressed in *Escherichia coli*, there is a tendency of the bacterium to form protein aggregates called inclusion bodies. They are insoluble and electron-dense particles that are enriched for recombinant protein. The formation of inclusion bodies has become an obstacle in those cases where the purpose is to obtain soluble protein and, therefore, different strategies have been formulated to avoid them. However, in recent decades the possibility of recovering biologically active proteins from these aggregates has been discovered, which has increased scientific interest in understanding their molecular organization and their potential use in biomedicine and biotechnology. The effect of bioprocess conditions on recovery yields and structural characteristics of inclusion bodies has also been studied.

At the industrial level, a widely used strategy for the production of recombinant proteins in *E. coli* is the $\lambda pL/pR$ -cl857 thermoinducible expression system, which is based on a cell growth stage at 30 °C and a subsequent induction stage above 37 °C. The advantages of this system include the fine regulation of promoters, the addition of chemical inducers is avoided and contamination risks are minimized. Despite this, the increase in temperature together with the production of a recombinant protein can cause changes in metabolism and cell physiology, triggering the heat shock response and the formation of inclusion bodies. These molecular events that occur in a temperature-induced system have been studied in separate contexts, but they coincide in the participation of chaperones and proteases, which are responsible for maintenance of protein homeostasis.

Based on the fact that an increase in temperature above 34 °C induces the heat shock response in *E. coli*, we study the effect of the pre-induction temperature (30 °C vs 34 °C) on the recombinant protein yields, the expression of chaperones and the structure of inclusion bodies. The model protein was the human granulocyte-macrophage colony-stimulating factor (rHuGM-CSF) of great therapeutic relevance and with tendency to aggregation. Recombinant *E. coli* cultures were carried out in bioreactors at 34 °C and 30 °C, which were subsequently induced at 42 °C. The pre-induction temperature of 34 °C accelerated bacterial growth, presenting an increase of up to 69 % in the specific growth rate and promoting the accumulation of rHuGM-CSF in inclusion bodies, compared to cultures that grew at 30 °C. Subsequently, chaperone immunodetection studies allowed us to identify the presence of the two main folding chaperones, DnaK and GroEL, in the aggregates; while their co-chaperones, DnaJ and GroES, were found in the soluble fraction. Evaluation of the structural properties of inclusion bodies was performed using the thioflavin-T binding assay and Fourier transform infrared (FTIR) spectrometry. Inclusion bodies from pre-induced cultures at 30 °C presented higher amyloid-type structure and were more resistant to proteolysis by proteinase K in comparison with inclusion bodies harvested with pre-induction of 34 °C. From this work, we concluded that a higher pre-induction temperature (34 instead of 30 °C) in a thermoinducible system allows obtaining greater amount of recombinant protein as well as reducing the amyloid content of the inclusion bodies.

1. INTRODUCCIÓN

1.1. Producción de proteínas recombinantes en *E. coli*

Desde el nacimiento de la tecnología del ADN recombinante, una enorme variedad de proteínas de interés farmacéutico han sido producidas en grandes cantidades y a bajos costos, convirtiéndose en una estrategia ampliamente usada tanto a nivel de laboratorio como a nivel de desarrollo industrial (Neubauer y Winter, 2001; Modak *et al.*, 2022). La bacteria *Escherichia coli* es una de las plataformas biológicas mejor caracterizadas para la producción de proteínas recombinantes, puesto que ofrece innumerables ventajas, incluyendo su rápido y abundante crecimiento en sustratos económicos, el amplio conocimiento de su genética y metabolismo, la posibilidad de realizar fermentaciones de alta densidad celular y, la vasta disponibilidad de promotores inducibles (Chen, 2012; Baeshen *et al.*, 2015; Sánchez-García *et al.*, 2016; Rosano *et al.*, 2019).

A pesar de la importancia de las células de mamífero por su capacidad para realizar modificaciones postraduccionales en las proteínas, *E. coli* sigue siendo un hospedero de continuo uso para la generación de diversos productos biológicos terapéuticos (Castiñeiras *et al.*, 2018). Hasta finales de julio de 2018, 316 productos biofarmacéuticos fueron aprobados por agencias regulatorias de Estados Unidos y de la Unión Europea, de los cuales ~22 % fueron producidos en *E. coli*, siendo el doble de la cantidad (~17,900 Kg) obtenida a partir de células de mamíferos (Walsh, 2014, 2018; Restrepo-Pineda *et al.*, 2021). Además, en el mercado de los biofármacos antitumorales, el ~69 % son proteínas procedentes de *E. coli*, específicamente destinadas a minimizar los efectos secundarios de la quimioterapia y radioterapia (Sánchez-García *et al.*, 2016).

Una aparente desventaja de la producción de proteínas heterólogas en *E. coli*, es la tendencia de la bacteria a formar agregados proteicos conocidos como cuerpos de inclusión (De Marco *et al.*, 2019). Por muchos años, los cuerpos de inclusión se consideraron un obstáculo en la obtención de proteínas recombinantes solubles y activas. Actualmente, diversos estudios han demostrado que tanto polipéptidos con características amiloides como proteínas con estructura similar a la nativa, coexisten dentro de los cuerpos de inclusión, otorgándoles funcionalidad biológica y estabilidad mecánica (González-Montalbán *et al.*, 2007; Cano-Garrido *et al.*, 2013; Rinas *et al.*, 2017; Singhvi *et al.*, 2020). Este progreso en la percepción de los cuerpos de inclusión bacterianos, ha incrementado el interés científico en el estudio de su organización molecular y su aplicación en diversas áreas como biotecnología, biomedicina, biocatálisis, entre otras (Rinas *et al.*, 2017; Pesarrodonna *et al.*, 2019; Jäger *et al.*, 2020).

Con el objetivo de mejorar la producción de proteínas recombinantes en *E. coli*, han surgido en las últimas décadas novedosas herramientas de ingeniería genética y de biología molecular, que han contribuido con la creación de diferentes cepas hospederas, plásmidos, promotores y, demás elementos regulatorios de un sistema de expresión (Huang *et al.*, 2012; Riggs, 2018). Uno de los pasos críticos para maximizar los niveles de expresión génica en una célula huésped, es la elección óptima del promotor, el cual debe permitir un control estricto y eficiente de la transcripción, sin efectos negativos sobre la viabilidad celular o la calidad del producto esperado (Jajesniak y Wong, 2015; Duzenli y Okay, 2020).

La actividad de un promotor suele estar regulada por un factor físico o químico, por ejemplo, la adición/ausencia de un nutriente (como arabinosa, ramnosa o triptófano; Marschall *et al.*, 2017), la modificación de la temperatura (Valdez-Cruz *et al.*, 2010), el pH (Chou *et al.*, 1995) o la tensión de oxígeno disuelto (Baez *et al.*, 2014). El promotor del fago T7, así como los promotores *lac*, *tac* y *trc* se encuentran dentro de los más usados para producir proteínas recombinantes en *E. coli* y son inducidos por la adición de isopropil- β -D-1-tiogalactopiranosido (IPTG) en el medio de cultivo (Donovan *et al.*, 1996; Browning *et al.*, 2019). Aunque altos niveles de expresión pueden lograrse con el uso de este sistema, en ciertos estudios se ha relacionado al IPTG con inestabilidad del plásmido, pérdida de viabilidad celular, disminución de la velocidad específica de crecimiento y estrés fisiológico (Kosinski *et al.*, 1992; Miroux y Walker, 1996; Dvorak *et al.*, 2015). Sumado a esto, el IPTG o cualquier otro compuesto químico, debe ser retirado de los productos para fines terapéuticos, lo que aumenta los pasos de purificación y los costos del bioproceso (Donovan *et al.*, 1996). Por otro lado, están los promotores *araBAD* y *rhaBAD* inducidos por la adición de l-arabinosa o l-ramnosa, respectivamente (Giacalone *et al.*, 2006), y el promotor *trp* inducido por la limitación de triptófano (Bass y Yansura, 2000). Este grupo de promotores pueden ocasionar limitaciones metabólicas y de crecimiento en las células, debido a las heterogeneidades durante la alimentación del inductor en biorreactores a gran escala (Siegele y Hu, 1997; Sørensen y Mortensen, 2005).

Para aliviar estos inconvenientes, una alternativa son los promotores *pL* y *pR* derivados del bacteriófago λ , los cuales son finamente regulados por el represor termolábil cI857 e inducibles por un aumento de la temperatura del cultivo (Villaverde *et al.*, 1993; Valdez-Cruz *et al.*, 2010; Riggs, 2018). En este sistema de expresión, no es necesaria la adición de inductores químicos como el IPTG, se alcanzan niveles de proteína similares a los promotores T7 y *tac*; además, se reducen los riesgos de contaminación por manipulación del cultivo (Makrides, 1996; Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011). Sin embargo, elevar la temperatura durante la inducción puede ser complicado para grandes volúmenes de cultivo y, el estrés generado por las altas temperaturas, conduce a menudo a la insolubilidad de las proteínas (Caspeta *et al.*, 2009). A pesar de la extensa aplicación del sistema termoinducible $\lambda pL/pR$ -cI857 a nivel industrial, las respuestas fisiológicas, metabólicas y moleculares de *E. coli* durante la sobreproducción de proteínas recombinantes usando inducción térmica, han sido escasamente estudiadas (Valdez-Cruz *et al.*, 2010; Restrepo-Pineda *et al.*, 2021). Un desafío actual es comprender la relación entre la modificación de las condiciones de cultivo, la respuesta de choque térmico y el impacto en la agregación de las proteínas recombinantes. Así como también, desarrollar estrategias de termoinducción que permitan ajustar las características estructurales de los cuerpos de inclusión para recuperar proteínas activas biológicamente y mejorar los rendimientos de producto.

Con este trabajo buscamos comprender cómo la temperatura durante la etapa de crecimiento del cultivo (antes de la termoinducción) puede impactar en la respuesta de choque térmico de *E. coli*, evaluada en términos de la expresión de algunas chaperonas. Evaluamos el efecto de diferentes temperaturas de pre-inducción (30 °C y 34 °C) en el crecimiento de *E. coli* W3110 produciendo un factor estimulante de colonias de granulocitos y macrófagos (rHuGM-CSF) y su efecto en la estructura de los cuerpos de inclusión usando el sistema termoinducible $\lambda pL/pR$ -cI857. El diferencial de 34 a 42 °C en comparación con el de 30 a 42 °C puede generar cambios importantes en la dinámica molecular de la bacteria,

ya que a 34 °C se enciende la repuesta al estrés térmico, los procesos biológicos se aceleran y hay sobre expresión de muchas moléculas. Esto a su vez, puede modificar la producción de las proteínas recombinantes y las propiedades de los cuerpos de inclusión.

2. ANTECEDENTES

2.1. Cuerpos de inclusión

Los cuerpos de inclusión se conocen comúnmente como agregados insolubles, generados durante la producción de proteínas recombinantes en *E. coli* (Williams *et al.*, 1982). Son complejos macromoleculares que contienen proteína heteróloga en diferentes conformaciones (de tipo nativa y no nativa), así como también en menor proporción proteínas celulares, restos de lípidos y ácidos nucleicos (Valax y Georgiou, 1993; Jürgen *et al.*, 2010). Estas partículas electrodensas pueden observarse fácilmente bajo el microscopio electrónico distribuidas en el citoplasma o periplasma de *E. coli* (**Figura 1**), presentando formas esféricas/pseudo-esféricas y diámetros entre 50 a 800 nm (Bowden *et al.*, 1991; Peternel *et al.*, 2008; Margreiter *et al.*, 2008; De Marco *et al.*, 2019).

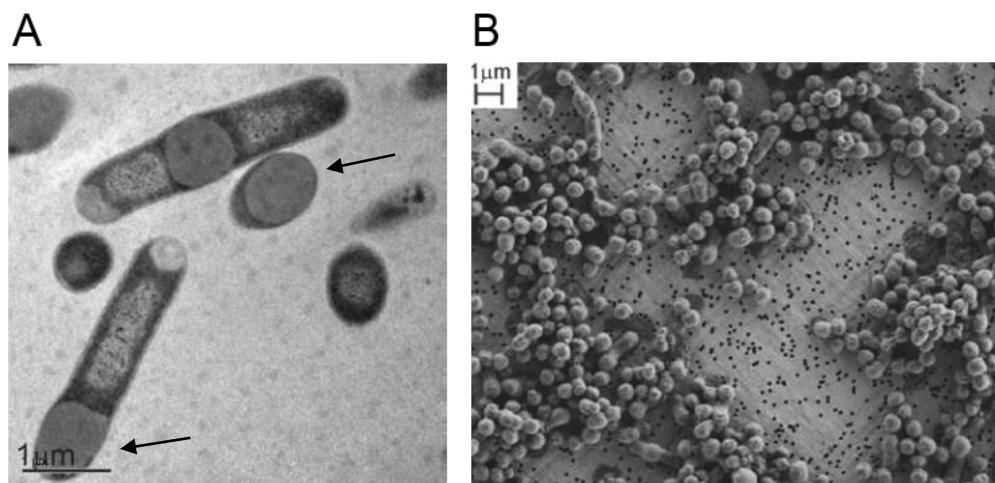


Figura 1. Cuerpos de inclusión de *E. coli*. **(A)** Secciones transversales de la bacteria produciendo un factor estimulante de colonias de granulocitos (G-CSF) bajo microscopio electrónico de transmisión. **(B)** Cuerpos de inclusión aislados por disrupción celular en un homogeneizador y lavados en agua pura bajo microscopio electrónico de barrido (modificado de Peternel *et al.*, 2008).

La agregación proteica es un evento biológico frecuente que tienen en común las bacterias, las levaduras y los animales (Tyedmers *et al.*, 2010). En bacterias, ocurre principalmente durante la sobre expresión de proteínas recombinantes, pues la alta concentración de polipéptidos en el citosol favorece las interacciones específicas entre superficies hidrofóbicas expuestas, incrementando las posibilidades de agregación (Baneix y Mujacic,

2004; Wang *et al.*, 2010; Carratalá *et al.*, 2021). También las condiciones de estrés pueden generar un desequilibrio de la homeostasis celular, conduciendo a una escasez de chaperonas y proteasas y, consecuentemente, al plegamiento ineficiente de las proteínas (Carrió y Villaverde, 2003; Hartl *et al.*, 2011; Schramm *et al.*, 2020).

Durante muchos años, se pensaba que los cuerpos de inclusión estaban formados exclusivamente por proteínas mal plegadas o desplegadas, las cuales constituían un “cluster” proteico no funcional excluido del control de la calidad celular (Rinas *et al.*, 2017). Por lo tanto, se consideraban un obstáculo para tener proteínas solubles y activas, lo que impulsó la creación de diferentes estrategias para evitar la agregación proteica (Bhatwa *et al.*, 2021). Sin embargo, en las últimas décadas, diversos grupos de investigación se han dedicado a ampliar el conocimiento sobre los cuerpos de inclusión, buscando entender la relación entre su origen, sus propiedades y sus posibles aplicaciones (Villaverde y Carrió, 2003; Rinas *et al.*, 2017; De Marco *et al.*, 2019).

Actualmente, los cuerpos de inclusión se definen como una mezcla entre diferentes subpoblaciones de proteínas: proteínas plegadas de forma similar a la nativa, que les brindan actividad biológica y especies proteicas con características amiloides, que les confieren resistencia mecánica (González-Montalban *et al.*, 2007; Cano-Garrido *et al.*, 2013; Rinas *et al.*, 2017; Singhvi *et al.*, 2020). Numerosos estudios han demostrado que los cuerpos de inclusión no sólo tienen un alto grado de organización interna, sino que también comparten características estructurales y biológicas con las fibrillas amiloides (Carrió *et al.*, 2005; Wang *et al.*, 2008; De Groot *et al.*, 2009). Los amiloides son depósitos celulares de proteínas que se ensamblan formando fibras insolubles resistentes a la degradación, están compuestos predominantemente por una estructura rica en hojas β entrecruzadas y se asocian con enfermedades neurodegenerativas (Rambaran y Serpell, 2008; Willbold *et al.*, 2021). Análisis de los cuerpos de inclusión usando FTIR han revelado la coexistencia tanto de una fracción de hojas β altamente ordenadas, similar a los amiloides, como también de una fracción significativa de estructura secundaria de tipo nativo y conformaciones desordenadas (Ami *et al.*, 2005; Singh *et al.*, 2020; Restrepo-Pineda *et al.*, 2022a). Datos de dicroísmo circular (CD) y difracción de rayos X, también respaldan la formación de estructuras de tipo amiloide en los cuerpos de inclusión (Wang *et al.*, 2008; De Groot *et al.*, 2009). Finalmente, ensayos de unión de los cuerpos de inclusión bacterianos a diferentes colorantes que son específicos para amiloides, como la tioflavina T (Th-T) y el rojo congo, prueban la estructura altamente ordenada de los mismos (Carrió *et al.*, 2005; Wang *et al.*, 2008; Upadhyay *et al.*, 2012; Castellanos-Mendoza *et al.*, 2014; Singh *et al.*, 2020). Dichas evidencias sobre el contenido de estructuras de tipo amiloide en los cuerpos de inclusión, impulsan el uso de modelos bacterianos para explorar los determinantes moleculares de la agregación de proteínas en sistemas biológicos simples (Caballero *et al.*, 2022).

La presencia de actividad biológica entre las subpoblaciones de proteínas que forman parte de los cuerpos de inclusión ha cambiado por completo el concepto de este tipo de material, pasando de ser un remanente del proceso a convertirse en un material biocompatible de gran interés en diferentes áreas de biomedicina y biotecnología (Rinas *et al.*, 2017; De Marco *et al.*, 2019; Belková *et al.*, 2022). Como ilustra la **Figura 2**, los cuerpos de inclusión se han usado como materiales de soporte para proliferación celular y regeneración de tejidos (García-Fruitós *et al.*, 2010; Díez-Gil *et al.*, 2010). Otros enfoques han explorado su potencial en biocatálisis como medios para inmovilizar enzimas (Jäger *et al.*, 2020) e

inclusive como nanopartículas de liberación de proteínas terapéuticas (Pesarrodonna *et al.*, 2019; Bargailla *et al.*, 2022).

Al mismo tiempo, el estudio de la relación entre la estructura y la función de las proteínas en los cuerpos de inclusión bacterianos, ha impulsado el desarrollo de métodos de aislamiento, solubilización, replegamiento y purificación más eficientes para la recuperación de proteínas bioactivas (**Figura 2**). Desde métodos para lisar a las células bacterianas manteniendo la integridad de los cuerpos de inclusión, hasta el desarrollo de enfoques combinatorios basados en agentes para la desnaturalización suave, condiciones de replegamiento personalizadas y esquemas de cromatografía multimodal, que han permitido obtener altos rendimientos de recuperación de proteína sin afectar su calidad (Singhvi *et al.*, 2020; Seras-Franzoso *et al.*, 2022). En muchos bioprocesos, los cuerpos de inclusión no necesariamente son un obstáculo, al contrario son deseables para la purificación inicial de la proteína de interés, ya que son fácilmente recuperables por rompimiento celular y centrifugación.

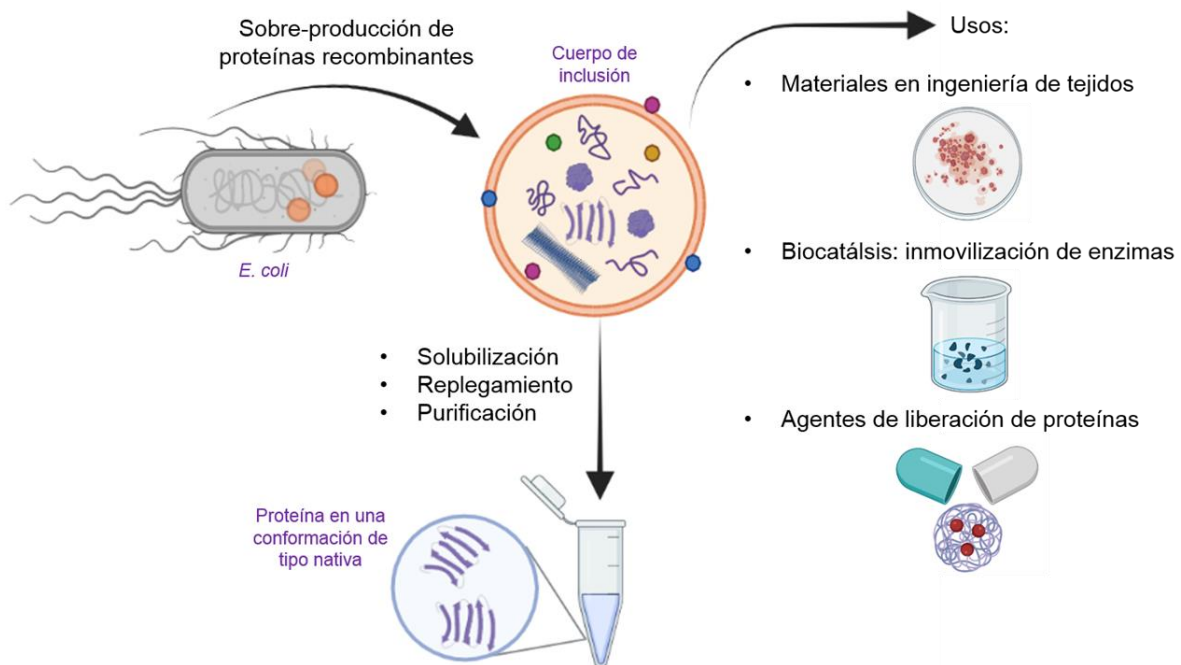


Figura 2. Producción de cuerpos de inclusión y sus aplicaciones. La sobre-expresión de proteínas recombinantes en *E. coli* conduce comúnmente a la formación de cuerpos de inclusión (CI). Los CI tienen una composición heterogénea de proteína (intermediarios parcialmente plegados, agregados amorfos, conformaciones de tipo amiloide y proteína de tipo nativa) y se encuentran interaccionando con diversas chaperonas citosólicas. Se ha estudiado su potencial como materiales para proliferación celular en ingeniería de tejidos, como medios para inmovilizar enzimas, como agentes de liberación de proteínas terapéuticas y, principalmente, como materia prima para la recuperación de proteínas biológicamente activas.

En diversos reportes se ha asociado la actividad y calidad de los cuerpos de inclusión con las condiciones de los bioprocesos (De Groot y Ventura, 2006; Margreiter *et al.*, 2008, Upadhyay *et al.*, 2012; Castellanos-Mendoza *et al.*, 2014; Valdez-Cruz *et al.*, 2017; Calcines-Cruz *et al.*, 2018; De Marco *et al.*, 2019; Restrepo-Pineda *et al.*, 2019, 2022a, 2022b). Durante la producción de proteínas recombinantes, diferentes parámetros del cultivo como el pH, la temperatura, la agitación, el tiempo de cultivo, entre otros, pueden tener un impacto considerable en la composición y en las características físico-químicas de los cuerpos de inclusión, por ejemplo en su tamaño, estructura interna, contenido de proteína activa y potencial para ser solubilizados (**Tabla 1**).

Se ha descrito que el tamaño de los cuerpos de inclusión incrementa con respecto al tiempo de cultivo después de la inducción (Margreiter *et al.*, 2008, Upadhyay *et al.*, 2012). También, la estructura interna de los agregados se modifica, a pesar de que no se han reportado cambios morfológicos importantes (Upadhyay *et al.*, 2012). Con el tiempo de cultivo, los cuerpos de inclusión se vuelven más resistentes a la actividad de la proteinasa K y aumenta su unión a colorantes amiloidogénicos, lo que está indicando una progresión temporal en la estructuración amiloide (Castellanos-Mendoza *et al.*, 2014).

En el mismo sentido, el pH del cultivo tiene un impacto en la formación de los cuerpos de inclusión y en sus características fisicoquímicas. Castellanos-Mendoza y colaboradores (2014) reportaron que los cuerpos de inclusión obtenidos en biorreactores bajo condiciones controladas de pH (pH=7.5) presentaron mayor resistencia a agentes desnaturalizantes, menor degradación por proteinasa K y mayor contenido de estructura de tipo amiloide, en comparación con aquellos formados bajo condiciones de pH no controladas (pH>8.5) (Castellanos-Mendoza *et al.*, 2014). Resultados similares se encontraron cuando se produjeron cuerpos de inclusión bajo un pH constante de 8.5, utilizando una proteína modelo y un fondo genético diferentes (Calcines-Cruz *et al.*, 2018).





El tipo de agitación así como las condiciones de la aireación en matraces agitados pueden modificar el tamaño de los agregados, obteniendo cuerpos de inclusión más pequeños y más fáciles de solubilizar cuando se usa un sistema de mezclado por resonancia acústica a altas velocidades (20 x g), en comparación con el sistema tradicional de agitación orbital a 200 rpm (Valdez-Cruz *et al.*, 2017).

En cuanto a la influencia de la temperatura en la formación y calidad de los cuerpos de inclusión, los primeros estudios defendían el uso de bajas temperaturas (entre 16 °C y 30 °C) durante el crecimiento de *E. coli* para minimizar la producción de cuerpos de inclusión o en algunos casos para enriquecerlos con proteínas parcial o totalmente plegadas, permitiendo así mantener su actividad biológica (Jevsevar *et al.*, 2005; De Groot y Ventura, 2006; Peternel *et al.*, 2008). De la misma forma, reportaron que el porcentaje de recuperación de proteína recombinante de los cuerpos de inclusión obtenidos a bajas temperaturas, era mayor en comparación a los cultivos mantenidos por encima de 37 °C, luego de los pasos de solubilización y renaturalización (Strandberg y Enfors, 1991; Jevsevar *et al.*, 2005; Peternel *et al.*, 2008). Más recientemente, diferentes investigaciones han demostrado que a temperaturas elevadas de inducción (>37 °C) es posible recolectar cuerpos de inclusión enriquecidos de proteínas recombinantes, las cuales presentan estructura similar a la nativa y mantienen su actividad biológica (Restrepo-Pineda *et al.*,

2019; Singh *et al.*, 2020; Restrepo-Pineda *et al.*, 2022a, 2022b). Particularmente, cuerpos de inclusión de la asparaginasa II producidos a altas temperaturas se caracterizaron por una mayor actividad enzimática y menor contenido amiloide (Singh *et al.*, 2020). Asimismo, cuerpos de inclusión de proteínas de uso terapéutico como ESAT-6 y GM-CSF obtenidos de cultivos termoinducidos a 39°C, fueron más susceptibles a procesos de degradación por proteasas K y mostraron un contenido enriquecido de hélices alfa, en comparación con aquellos obtenidos a 42 °C (Restrepo-Pineda *et al.*, 2019; Restrepo-Pineda *et al.*, 2022b). De este modo, el impacto de la temperatura de inducción en las propiedades de los cuerpos de inclusión depende de la naturaleza de la proteína recombinante, lo que implica que se debe evaluar el efecto de este parámetro de cultivo para cada caso en particular.

Actualmente, cuando una proteína de interés no puede expresarse en gran cantidad en forma soluble, los cuerpos de inclusión se consideran una valiosa opción como fuente de proteína recombinante, abarcando en algunos casos hasta el 90 % de su contenido (Carrió y Villaverde, 2003; Belková *et al.*, 2022). El hecho de tener a las proteínas reunidas en estos reservorios biológicos les brinda estabilidad química, resistencia a ataques proteolíticos por proteasas celulares y, su recuperación llega a ser rentable económicamente en bioprocesos a gran escala (Carrió y Villaverde, 2002; Lee *et al.*, 2006; Upadhyay *et al.*, 2012; Kaur *et al.*, 2018).

Tabla 1. Efecto de los parámetros del cultivo en las propiedades de los cuerpos de inclusión.

Parámetros	Características de los CI	Referencias
<ul style="list-style-type: none"> • Tiempo de cultivo 	A mayor tiempo de cultivo luego de la inducción, aumenta el diámetro de los CI, así como su contenido amiloide (mayor unión a colorantes amiloidogénicos). Llegan a ser más resistentes a proteólisis y desnaturalización.	Margreiter <i>et al.</i> , 2008; Upadhyay <i>et al.</i> , 2012; Castellanos-Mendoza <i>et al.</i> , 2014.
<ul style="list-style-type: none"> • pH 	CI formados a pH alcalino presentan mayor contenido estructural de tipo helicoidal, son menos resistentes a proteólisis y disminuye su unión a colorantes amiloidogénicos, en comparación con los CI obtenidos a pH ácido.	Strandberg y Enfors 1991; Castellanos-Mendoza <i>et al.</i> , 2014; Calcines-Cruz <i>et al.</i> , 2018.
<ul style="list-style-type: none"> • Temperatura 	A bajas temperaturas de inducción: CI con proteína activa, menos impurezas, rápida solubilización. A mayor temperatura de inducción: aumenta la tasa de agregación; también puede encontrarse proteína activa.	De Groot y Ventura, 2006; Restrepo-Pineda <i>et al.</i> , 2019; Singh <i>et al.</i> , 2020.
<ul style="list-style-type: none"> • Agitación 	En cultivos de matraces bajo agitación con resonancia acústica (20 x g), se forman CI más pequeños y son más susceptibles a la degradación proteolítica en comparación con los CI formados bajo agitación orbital (200 rpm).	Valdez-Cruz <i>et al.</i> , 2017.

2.2. Sistema de expresión termoinducible $\lambda pL/pR$ -cI857

Para obtener altos rendimientos durante la producción de proteínas recombinantes, una de las estrategias más comunes es el uso de los promotores termoinducibles pL y pR derivados del bacteriófago λ (Villaverde *et al.*, 1993; Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011; Restrepo-Pineda *et al.*, 2021). Estos promotores fuertes y finamente regulados, son controlados por el represor termolábil cI857, el cual surgió a partir de una mutación de la Ala66 por Thr en la secuencia del represor cI del bacteriófago λ , que provocó la inactivación del represor entre 40-45 °C y la inducción del profago (Lieb, 1966). Este hallazgo abrió la posibilidad de regular físicamente la unión del represor λ cI a los promotores pL/pR y darle un uso biotecnológico para la producción de proteínas recombinantes (Villaverde *et al.*, 1993; Valdez-Cruz *et al.*, 2010).

El sistema termoinducible $\lambda pL/pR$ -cI857 de *E. coli* es llevado a cabo usualmente en dos etapas de cultivo: la etapa de crecimiento celular a 28-30 °C (**Figura 3A**) y la etapa de producción donde la temperatura se incrementa por encima de los 37 °C (**Figura 3B**) (Restrepo-Pineda *et al.*, 2021). La expresión del gen es eficientemente reprimida por cI857 a temperaturas por debajo de 37 °C, el cual interactúa a través de su dominio amino terminal con tres regiones operadoras ($oL3/oR3$, $oL2/oR2$ y $oL1/oR1$) de los promotores pL y pR , formando un “loop” de ADN que bloquea la transcripción (Dodd *et al.*, 2005; Lewis *et al.*, 2016). Al elevar la temperatura del cultivo, por arriba de 37 °C, se libera el represor cI857 permitiendo la transcripción del gen de la proteína de interés a partir de los promotores pL/pR (Caulcott y Rhodes, 1986; Villaverde *et al.*, 1993; Dodd *et al.* 2004; Valdez-Cruz *et al.*, 2010).

Con este sistema de expresión se evita el uso de medios especiales o la adición de inductores químicos como el IPTG, el cual además de ser costoso y tóxico para las células, debe ser eliminado del producto final, implicando más pasos y costos en los procesos de purificación (Donovan *et al.*, 1996; Palomares *et al.*, 2004). Asimismo, se alcanzan niveles similares de proteína a los reportados para sistemas basados en los promotores T7 o *tac* y, los riesgos de contaminación por manipulación del cultivo se reducen al controlar la temperatura externamente (Makrides, 1996; Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011).

En la industria biotecnológica, algunos trabajos han destacado el riguroso control de la expresión génica en sistemas termoinducibles, en comparación con otros modelos de expresión (Caspeta *et al.*, 2009; Valdez-Cruz *et al.*, 2010; Caspeta *et al.*, 2013; Restrepo-Pineda *et al.*, 2019). Aunque en sistemas que contienen el promotor *lac* o *trp* los rendimientos son altos, la cantidad de proteína represora sintetizada puede no ser suficiente, por tanto hay un control ineficiente que conduce a fugas en la expresión génica (O'Farrell *et al.*, 1978; Caulcott y Rhodes, 1986). La combinación de los promotores pL/pR con el represor termolábil cI857 da como resultado una poderosa herramienta de expresión, que permite la fina represión de los promotores por debajo de 37 °C (Caulcott y Rhodes, 1986; Villaverde *et al.*, 1993). Hasta la fecha, se han producido diferentes proteínas recombinantes mediante termoinducción, siendo la mayoría ellas de interés terapéutico humano, como por ejemplo, el factor de necrosis tumoral (Menart *et al.*, 2003; Singha *et al.*, 2018), la insulina (Schmidt *et al.*, 1999), la preproinsulina (Caspeta *et al.*, 2009; Caspeta *et*

al., 2013), la hormona del crecimiento humano (Tabandeh *et al.*, 2004; Ortiz-Enriquez *et al.*, 2016), el antígeno ESAT-6 de *Mycobacterium tuberculosis* (Restrepo-Pineda *et al.*, 2019), el factor estimulante de colonias de granulocitos y macrófagos (Restrepo-Pineda *et al.*, 2022a, Restrepo-Pineda *et al.*, 2022b), entre otras.

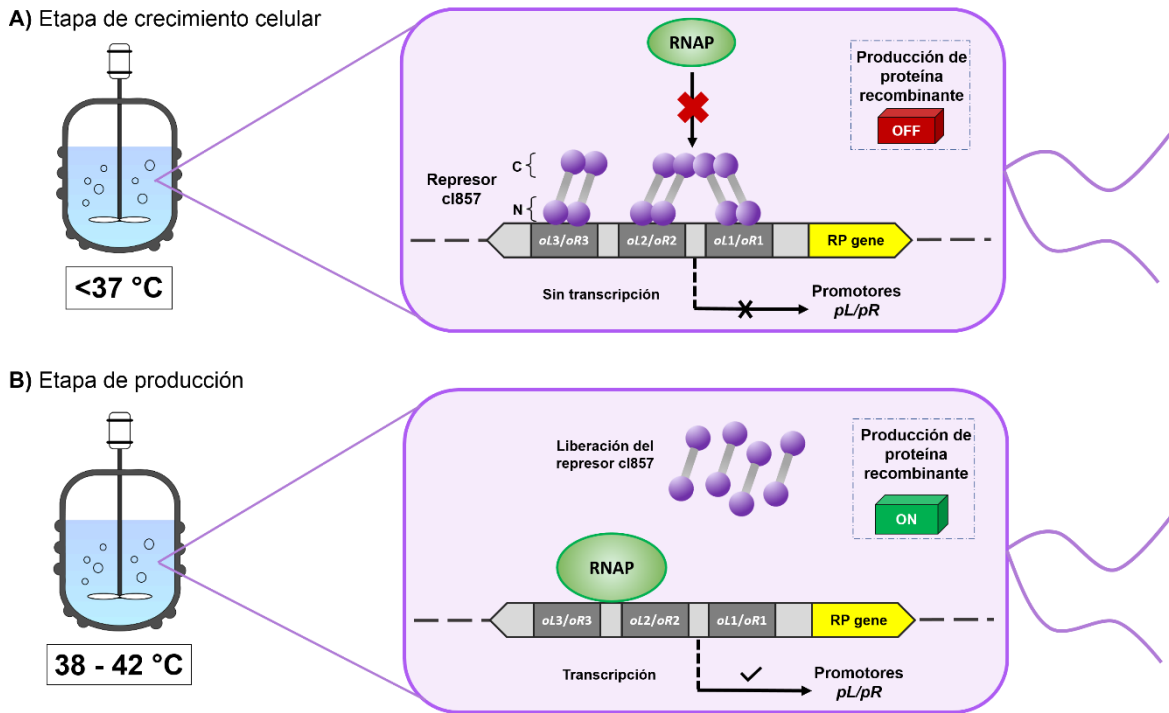


Figura 3. Regulación molecular del sistema de expresión termoinducible $\lambda pL/pR-cl857$ de *E. coli*. **(A)** Durante el crecimiento celular a temperaturas por debajo de los $37\text{ }^{\circ}\text{C}$, el represor cl857 (en morado) se encuentra unido a través de su dominio amino terminal a las regiones operadoras ($oL1/oR1$, $oL2/oR2$ y $oL3/oR3$) de los promotores pL y pR , reprimiendo la transcripción del gen de la proteína recombinante (en amarillo). **(B)** Un aumento en la temperatura por encima de los $37\text{ }^{\circ}\text{C}$, provoca la liberación del represor cl875, permitiendo el reconocimiento de la ARN polimerasa (en verde) para activar la transcripción del gen recombinante (adaptado de Restrepo-Pineda *et al.*, 2021).

Durante la producción de proteínas recombinantes usando un sistema de inducción por temperatura, una serie de eventos moleculares y cambios fisiológicos se desencadenan en *E. coli* en respuesta a la exposición prolongada al calor y a la alta tasa de síntesis de proteínas (Hoffmann y Rinas, 2000; Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011; Restrepo-Pineda *et al.*, 2021). Las células bacterianas activan ciertos procesos celulares tales como la respuesta de choque térmico, la respuesta SOS y la respuesta estricta para contrarrestar las condiciones estresantes (Straus *et al.*, 1987; Gill *et al.*, 2000; Hoffmann y Rinas, 2004). En las mismas circunstancias, ocurren ajustes en los flujos del metabolismo central del carbono debido a la carga metabólica y a la acumulación de acetato, lo que va

acompañado de cambios fisiológicos como relajación en el ADN y modificación en las membranas celulares (Weber *et al.*, 2002; Hoffmann *et al.*, 2002; López *et al.*, 2000; Mejía *et al.*, 1995; Restrepo-Pineda *et al.*, 2021).

A través de los años se han desarrollado diferentes estrategias para mejorar los rendimientos de proteína en el sistema $\lambda pL/pR$ -cI857 y para evitar deterioros drásticos en las células debido al incremento de la temperatura (Restrepo-Pineda *et al.*, 2021). Las primeras investigaciones reportaban un 2.0-6.6 % de proteína recombinante con respecto a la proteína total celular usando el promotor pL (Bernard *et al.*, 1979). Pero más adelante, algunas modificaciones como la reducción en el número de nucleótidos entre el promotor pL y el gen recombinante, o la inserción del sitio Nut para evitar la terminación dentro del promotor pL , condujeron a un enriquecimiento de casi el 30 % de proteína recombinante (Remaut *et al.*, 1981). Posteriormente, como alternativa para la inducción por calor, se diseñó un plásmido conteniendo el promotor pL y el represor nativo λcI , en lugar del represor termolábil cI857 (Mott *et al.*, 1985). El uso de agentes químicos como mitomicina C o ácido nalidíxico activaron la transcripción del promotor mediante la inactivación de la proteína RecA, lo que resultó en sobreproducción de proteína recombinante a niveles de ~40 % de la proteína total celular (Mott *et al.*, 1985).

A pesar de los altos rendimientos reportados, también es ampliamente conocido que la exposición de las células a temperaturas elevadas por períodos de tiempo prolongados, pueden causar una disminución en el crecimiento, segregación de plásmidos, acumulación de chaperonas y proteasas y formación de cuerpos de inclusión (Hoffmann y Rinas, 2001; Hoffmann y Rinas, 2004; Harcum y Haddadin, 2006). Por lo tanto, algunos grupos de investigación han diseñado estrategias de termoinducción que permitan mantener cultivos bacterianos viables y mejorar la fase productiva (Chao *et al.*, 2004; Aucoin *et al.*, 2006; Caspeta *et al.*, 2009; Caspeta *et al.*, 2013; Mansey *et al.*, 2014). En matraces agitados, un choque térmico de 30 a 42 °C durante 2 min y posterior disminución de la temperatura a 37 °C, resultó ser favorable para extender la fase exponencial de crecimiento y aumentar hasta 18 veces la concentración y actividad de la β -galactosidasa bajo la regulación del promotor pL (Gupta *et al.*, 1999). Sin embargo, un choque térmico de 2 min puede ser improbable para los procesos a gran escala, debido a las limitaciones de la transferencia de calor en biorreactores (Aucoin *et al.*, 2006; Caspeta *et al.*, 2009). En este sentido, algunos grupos de investigación han llevado a cabo cultivos termoinducidos usando dos biorreactores interconectados, donde el primer biorreactor se mantiene por debajo de 37 °C, reprimiendo los promotores pL/pR para acumular biomasa, mientras que el segundo biorreactor funciona a 40 °C para inducir la producción de la proteína recombinante (Hortacsu y Ryu, 1991; Chao *et al.*, 2004). Usando esta configuración, se produjo un 30 % más de proteína en comparación con el cultivo tradicional, mejorando la productividad y la estabilidad del plásmido (Chao *et al.*, 2004).

Los cultivos por lote alimentado o “fed-batch” también se ha implementado en sistemas termoinducibles, con el objetivo de alcanzar altas densidades celulares y prolongar el período de producción después de la termoinducción (Aucoin *et al.*, 2006; Caspeta *et al.*, 2009; Caspeta *et al.*, 2013). Un cultivo alimentado de 20 L basado en dos ciclos de inducción de 37 a 42 °C o 30 a 42 °C durante 1 h, resultó en una producción volumétrica tres veces mayor de la proteína GFP en comparación con la inducción simple y, en un incremento en la velocidad específica de crecimiento (Aucoin *et al.*, 2006). Similarmente,

Caspeta y colaboradores (2013) plantearon un esquema de inducción oscilatoria con pulsos entre 37 a 42 °C o 30 a 42 °C por períodos de 2 h, en combinación con una estrategia de alimentación tanto exponencial como lineal. En el primer caso se alcanzó una biomasa 2.3 veces mayor y 1.3 veces más proteína recombinante, mientras que con una alimentación lineal, se obtuvo 4.0 veces más biomasa y 1.7 más proteína recombinante, en contraste con un sistema en lote usando inducción convencional a 42 °C (Caspeta *et al.*, 2013).

A pesar del extenso uso del sistema de expresión $\lambda pL/pR$ -cI857 a nivel industrial y de los avances en las estrategias de cultivo, aún queda por estudiar a profundidad la comunicación entre las respuestas moleculares, fisiológicas y metabólicas en *E. coli* durante la termoinducción (Restrepo-Pineda *et al.*, 2021). Si bien es cierto que no hay un protocolo único o estándar que indique cuál es el mejor esquema de inducción, es importante continuar estudiando la influencia de diversos factores como la naturaleza de la proteína recombinante, las condiciones de operación de cultivo, la escala de cultivo, la concentración de glucosa, el esquema de alimentación, la estrategia de termoinducción (punto, temperatura y duración de inducción), entre otros, para obtener altos rendimientos de producto sin afectar el crecimiento celular (Makrides, 1996; Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011).

2.3. Respuesta de choque térmico en *E. coli*

Las bacterias cuentan con un complejo sistema de procesos celulares para sensor y responder a condiciones de estrés, con el fin de mantener la integridad y viabilidad de sus componentes (Kortmann y Narberhaus, 2012; Schumann, 2016). Particularmente, la exposición de las células a altas temperaturas puede afectar el correcto plegamiento de las proteínas y conducir a su agregación, lo que activa una extensa red de chaperonas moleculares y proteasas (Gomes y Simão, 2014).

La respuesta de choque térmico (HSR, por sus siglas en inglés) es la forma en la que las células se manifiestan ante un aumento de la temperatura para proteger a las proteínas de la desnaturalización (Yamamori *et al.*, 1978). La HSR es una estrategia universal caracterizada por la rápida inducción de un conjunto de proteínas conservadas evolutivamente, que se conocen como proteínas de choque térmico (HSPs), las cuales monitorean y mantienen la homeostasis del proteoma celular (Guisbert *et al.*, 2004; Balchin *et al.*, 2016). Precisamente, la homeostasis de las proteínas se refiere al trabajo cooperativo de chaperonas y proteasas para mantener la calidad conformacional, estabilidad y características funcionales de las proteínas en todos los compartimentos celulares (Clausen *et al.*, 2019). En *E. coli*, la producción de las HSPs es controlada por el factor σ^{32} (gen *rpoH*), que regula la expresión de al menos 120 genes a través de su unión a la ARN polimerasa (Straus *et al.*, 1987; Grossman *et al.*, 1987; Nonaka *et al.*, 2006). La mayoría de estos genes codifican para chaperonas, promoviendo el plegamiento correcto de las proteínas y otras actúan como proteasas, degradando proteínas dañadas o no plegadas; el resto de genes está relacionado con proteínas que desempeñan un papel importante frente a condiciones adversas como daños en el ADN, estrés oxidativo, escasez de nutrientes, entre otros (Rosen y Ron, 2002; Baneyx y Mujacic, 2004; Nonaka *et al.*, 2006; Richter *et al.*, 2010).

La concentración celular del factor σ^{32} es estrictamente regulada por una compleja cascada de mecanismos transcripcionales, traduccionales y postraduccionales (Lim y Gross, 2011). A bajas temperaturas de crecimiento (30 °C), la cantidad de σ^{32} en las células se mantiene en un nivel basal (~50 moléculas por célula), a través de diferentes procesos (Straus *et al.*, 1987; Schumann, 2016). El control traduccional de σ^{32} es mediado por la estructura secundaria del ARN mensajero (ARNm) del gen *rpoH*, que actúa como un termómetro de ARN (Morita *et al.*, 1999; Kortmann y Narberhaus, 2012). A 30 °C (**Figura 4A**) se forma una estructura por emparejamiento intramolecular entre dos segmentos de la región codificante de *rpoH*, muy cerca del codón de inicio AUG, inhibiendo casi por completo el acceso al sitio principal de acoplamiento del ribosoma y, por tanto, la iniciación de la traducción (Grossman *et al.*, 1987; Nagai *et al.*, 1991; Morita *et al.*, 1999).

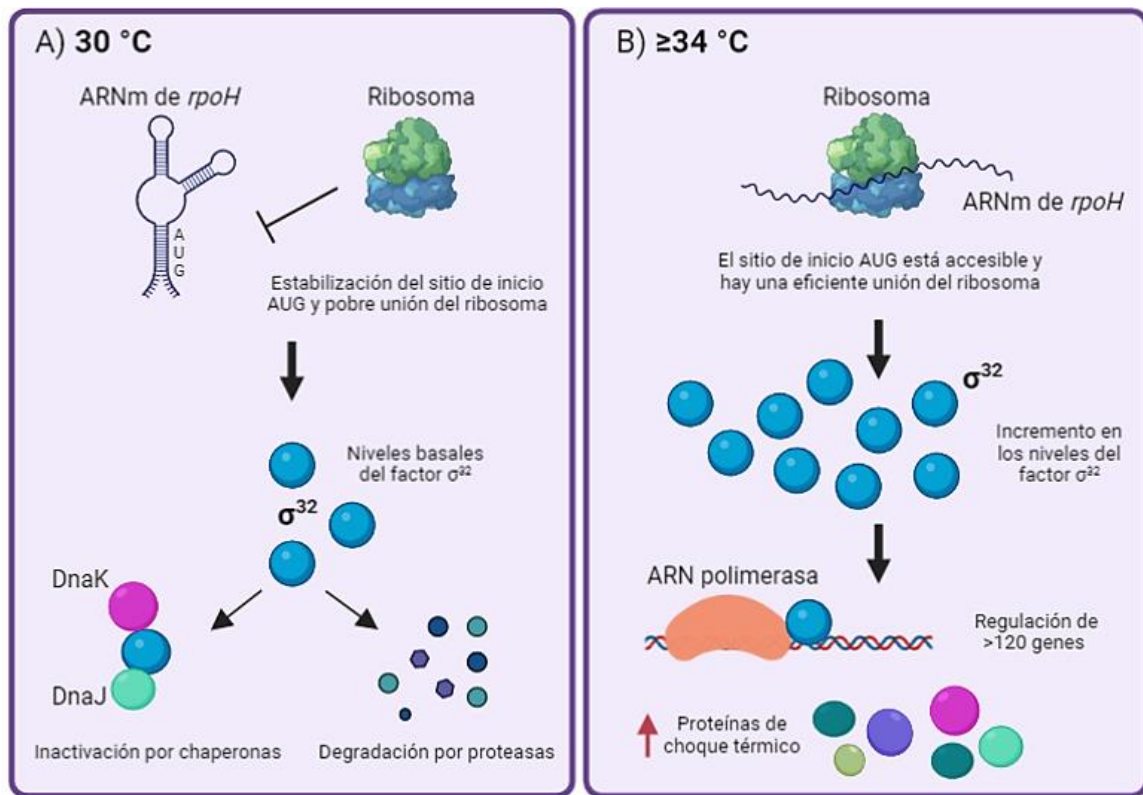


Figura 4. Modulación de la respuesta de choque térmico (HSR) en *E. coli*. **(A)** A bajas temperaturas de crecimiento (30 °C), se mantienen en las células niveles basales del factor σ^{32} , a través de diferentes mecanismos: 1. La regulación traduccional del ARNm del gen *rpoH*, 2. La regulación por chaperonas o 3. La degradación por proteasas. **(B)** Por encima de 34 °C, hay un aumento rápido en la traducción del factor σ^{32} al liberarse la estructura del ARNm de *rpoH*, desencadenando el incremento en las velocidades de síntesis de las proteínas de choque térmico y el encendido de la respuesta de choque térmico.

Por otro lado, cuando la concentración de proteínas mal plegadas es poca, las chaperonas libres DnaK y DnaJ pueden unirse e inactivar a σ^{32} , mediante la formación de un complejo estable σ^{32} -DnaK-DnaJ limitando su unión a la ARN polimerasa (Liberek *et al.*, 1992; Bukau, 1993; Arsene *et al.*, 2000). También, pueden conducir a su degradación por proteasas localizadas en la membrana como FtsH (Blaszczak *et al.*, 1999; Guisbert *et al.*, 2004) o proteasas citosólicas como ClpAP, ClpQY y Lon (Kanemori *et al.*, 1997).

Tras un aumento de la temperatura por encima de 34 °C (**Figura 4B**), los niveles de síntesis y la actividad de σ^{32} incrementan rápidamente, ya que la estructura secundaria del ARNm de *rpoH* se relaja, permitiendo su traducción (Morita *et al.*, 1999; Yura y Nakahigashi, 1999). Paralelamente, las chaperonas se unen a sus sustratos (proteínas desplegadas o mal plegadas) y el factor σ^{32} queda libre para unirse a la ARN polimerasa y favorecer la producción de más HSPs (Chakraborty *et al.*, 2014). Ambos mecanismos resultan en el aumento vertiginoso de los niveles de σ^{32} de hasta 17 veces y el incremento en la velocidad de síntesis de las HSPs, lo que señala el encendido de la respuesta de choque térmico (Straus *et al.*, 1987). Se ha demostrado que el desencadenamiento de la respuesta de choque térmico a temperaturas superiores a los 34 °C es esencial para el crecimiento de *E. coli*, pues cepas con mutaciones en el gen *rpoH* experimentan una deficiencia en el crecimiento a partir de los 34 °C (Yura *et al.*, 1984; Yano *et al.*, 1990).

Hasta el día de hoy, la respuesta de choque térmico sigue siendo un proceso crítico para la supervivencia de los microorganismos expuestos a estrés por calor y, a pesar, de décadas de estudio y de ser ampliamente descrita, siguen llevándose a cabo investigaciones sobre su potencial aplicación en diversas áreas. Por ejemplo, enfoques de biología sintética, han propuesto el uso de los promotores nativos de choque térmico de *E. coli* para la inducción de moléculas en respuesta a condiciones estresantes, ya sea para detectar un contaminante en el agua, para regular y optimizar la producción de un compuesto con relevancia industrial o para administrar un agente terapéutico in vivo (Rodrigues y Rodrigues, 2018).

En adición a la respuesta de choque térmico, las células bacterianas activan otras respuestas moleculares como la respuesta estricta y la respuesta SOS, para contrarrestar las condiciones estresantes causadas por la elevada síntesis de proteína recombinante, la prolongada exposición a altas temperaturas y la entrada a la fase estacionaria (Valdez-Cruz *et al.*, 2010; Restrepo-Pineda *et al.*, 2021). La alta tasa de producción de proteínas puede causar el rápido agotamiento de precursores de biosíntesis como esqueletos de carbono y aminoácidos (Hoffmann y Rinas, 2004). En la fase estacionaria, cuando los aminoácidos están presentes en concentraciones limitantes, se activa la producción de alarmones, guanosina pentafofato o tetrafofato (pppGpp y ppGpp) (Atkinson *et al.*, 2011; Steinchen y Bange, 2016). Estos alarmones regulan la traducción a través de EF-2 y promueven la respuesta estricta, deteniendo la síntesis de tRNA y rRNA, y causando la degradación de ribosomas (Traxler *et al.*, 2008; Bennison *et al.*, 2019).

Del mismo modo, se ha revelado que la expresión de genes recombinantes a partir de los promotores *pL/pR* activa el sistema SOS en las células bacterianas (Arís *et al.*, 1998). La exposición de *E. coli* a agentes o condiciones que dañan el ADN o interfieren con la replicación del ADN da como resultado la inducción de la respuestas SOS (Janion, 2008).

Bajo termoinducción, la elevada síntesis de proteínas y las altas temperaturas, pueden inducir tanto a la agregación de intermediarios proteicos como a una excesiva proteólisis de los productos (Kosinsky y Bailey, 1991; Arís *et al.*, 1998). Estos eventos tienen efectos nocivos sobre las células y provocan la respuesta SOS, con una notable inhibición de la división celular (Arís *et al.*, 1998; Hoffmann y Rinas, 2004).

2.3.1. Mecanismos de las chaperonas moleculares

Las chaperonas moleculares constituyen un conjunto de proteínas que interactúan, estabilizan y ayudan a otras proteínas a adquirir su conformación funcionalmente activa, formando redes de cooperación para mantener la homeostasis celular (Hartl y Hayer-Hartl, 2009; Hartl *et al.*, 2011). Los miembros de esta familia de proteínas hacen parte de las HSPs y, a menudo se conocen como proteínas de estrés, ya que se regulan a la alza en condiciones de estrés, como por ejemplo durante la producción de proteínas recombinantes, donde el aumento en la concentración de los intermediarios proteicos incrementa la posibilidad de agregación (Becker y Craig, 1994; Tyedmers *et al.*, 2010; Fatima *et al.*, 2021). Lo que discrimina a una proteína nativa de su contraparte no nativa o no plegada, es un incremento en la exposición de amino ácidos hidrofóbicos, característica que es reconocida por la chaperonas moleculares, que pueden unirse ya sea a parches hidrofóbicos, secuencias específicas o elementos estructurales (Richter *et al.*, 2010).

Las chaperonas están involucradas en una multitud de funciones de mantenimiento del proteoma, incluido el plegamiento *de novo*, el replegamiento de proteínas desnaturalizadas por estrés, el ensamblaje oligomérico, el tráfico de proteínas y la asistencia en la degradación proteolítica (Voth y Jakob, 2017). Según su mecanismo de acción, las chaperonas pueden clasificarse como de plegamiento, de estabilización o de desagregación (Baneyx y Mujacic, 2004). El primer grupo incluye a los complejos moleculares DnaK/DnaJ/GrpE y GroEL/GroES, que participan en el plegamiento y replegamiento de proteínas *de novo*, a través de ciclos de unión y liberación a sus sustratos regulados por ATP (Langer *et al.*, 1992; Ellis, 2006). Luego están las “small heat shock proteins” (sHSPs), que incluyen a IbpA/IbpB y funcionan como amortiguadores de la agregación durante situaciones de estrés severo (Mogk y Bukau, 2017). Finalmente, están las desagregasas como ClpB, miembro de la familia de Clp ATPasas y cuya función principal es solubilizar proteínas dañadas que no pudieron ser rescatadas por las chaperonas, mediante la interacción con las sHSPs y el sistema DnaK/DnaJ (Doyle *et al.*, 2015).

En *E. coli*, el “Trigger factor” o TF es la primer chaperona en contacto con las cadenas polipeptídicas nacientes del ribosoma (**Figura 5**). La mayoría de las proteínas pequeñas (~70 % del total) pueden plegarse rápidamente sin requerir ayuda adicional, sin embargo, las cadenas más largas interactúan posteriormente con DnaK y DnaJ (Wegrzyn y Deuerling, 2005). DnaK trabaja en conjunto con su co-chaperona DnaJ, la cual activa la unión al sustrato al desencadenar la hidrólisis del ATP unido a DnaK; mientras que la liberación del sustrato está controlada por el intercambio de ADP/ATP catalizado por el factor GrpE (Bhandari y Houry, 2015).

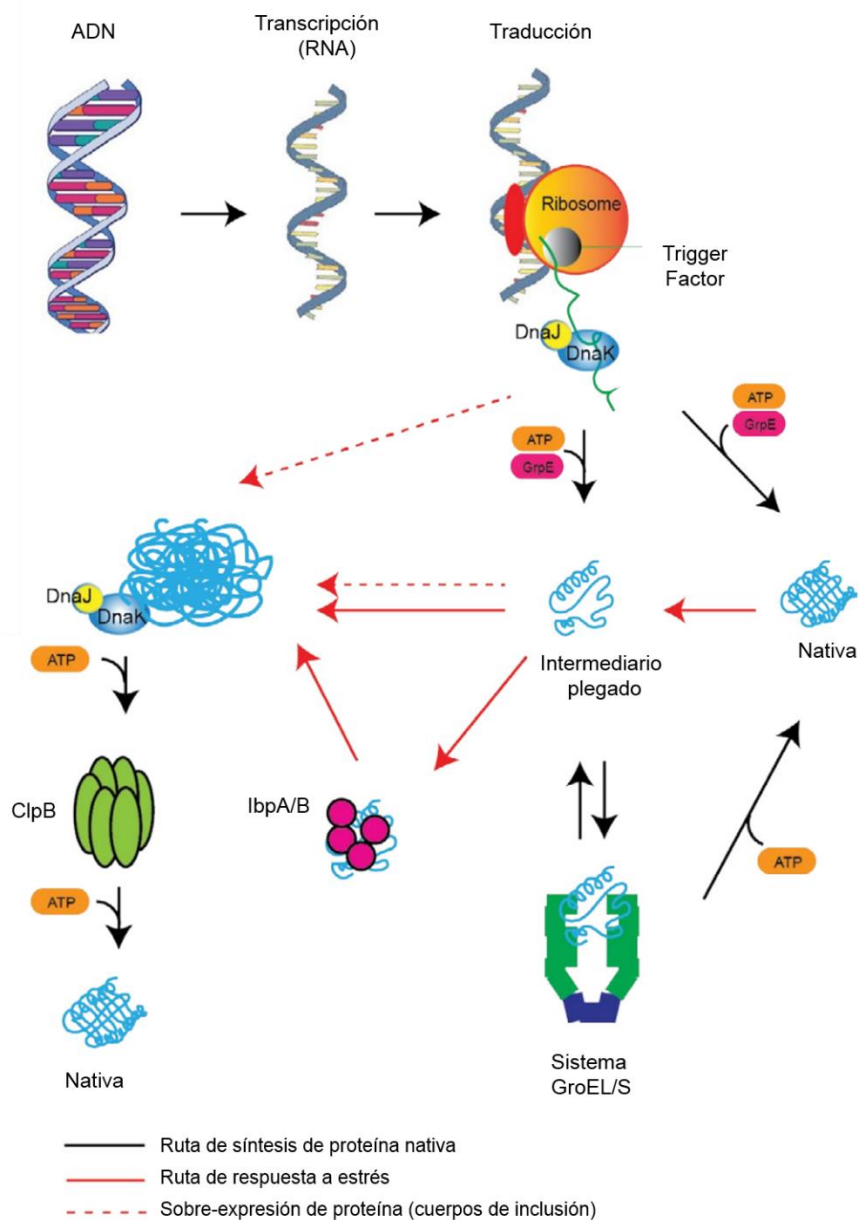


Figura 5. Participación de las chaperonas moleculares en la regulación de la homeostasis de las proteínas en el citoplasma de *Escherichia coli*. Los polipéptidos nacientes del ribosoma interactúan con el factor TF y con el sistema de chaperonas de plegamiento DnaK/DnaJ/GrpE para alcanzar una conformación nativa. Si esto no ocurre, los intermediarios parcialmente plegados son rescatados por un segundo sistema de plegamiento compuesto por GroEL y GroES o enviados a desagregación por ClpB. En el caso de la sobreexpresión de proteínas recombinantes (líneas rojas punteadas) o condiciones severas de estrés (líneas rojas sólidas), la capacidad de las chaperonas es limitada, requiriendo la ayuda de un sistema de proteínas de estabilización que incluye a IbpA e IbpB (modificada de Kwon *et al.*, 2016).

Después de múltiples ciclos de plegamiento, la proteína puede alcanzar una conformación nativa, transferirse al sistema GroEL/GroES que maneja alrededor del 10 % de las proteínas recién sintetizadas o, en definitiva, ser dirigida al sistema de desagregación de ClpB (Mayer, 2010). Durante este proceso, una línea de defensa adicional que se despliega cuando las células experimentan estrés intenso, es la presencia de proteínas estabilizadoras, como IbpA e IbpB, que exhiben una actividad independiente de ATP, previniendo la formación de grandes agregados sin promover activamente su remodelación (Mogk y Bukau, 2017; Mayer, 2010).

Una de las aplicaciones que se le ha dado a las chaperonas moleculares, ha sido la co-expresión con algunas proteínas recombinantes para mejorar su solubilidad y actividad, pero esto dependerá de la naturaleza y propiedades de cada proteína recombinante (Martínez-Alonso *et al.*, 2010). Al mismo tiempo, diferentes trabajos han reportado que las chaperonas se asocian a los cuerpos de inclusión, ejerciendo una función de remodelación de los agregados, ya sea mediante su disolución o la prevención de su agregación (Rinas *et al.*, 2007). Incluso, la morfología de los cuerpos de inclusión también se puede modificar dependiendo de la ausencia o presencia de HSPs específicas (García-Fruitós *et al.*, 2010).

2.3.2. Convergencia de las respuestas de estrés asociadas con la producción de proteínas recombinantes termoinducidas

En un sistema de expresión termoinducible, las respuestas celulares provocadas por el aumento de la temperatura y por la acumulación de proteínas recombinantes, son eventos que a pesar de ocurrir simultáneamente, se han estudiado durante muchos años en contextos separados. El aumento de temperatura requerido para inducir la producción de una proteína recombinante activará la HSR y, por lo tanto, la sobreproducción de chaperonas, que a su vez son necesarias para asistir en el plegamiento correcto de proteínas heterólogas, previniendo la formación de cuerpos de inclusión (Hoffmann y Rinas, 2000; Carrió y Villaverde, 2003; Valdez-Cruz *et al.*, 2010).

La comunicación intracelular entre las respuestas moleculares que se producen durante la termoinducción es una red compleja que incluye simultáneamente la respuesta al choque térmico, la sobreexpresión de proteínas recombinantes y la formación de cuerpos de inclusión. La **Figura 6** ilustra la relación de estos eventos moleculares en *E. coli*. En cultivos a 30 °C, la expresión de la proteína recombinante es inhibida por la interacción del represor cl857 con los promotores *pL/pR*, y los niveles basales de proteínas de choque térmico controlan la actividad de σ^{32} a través de una regulación negativa (Straus *et al.*, 1990; Valdez-Cruz *et al.*, 2010). Por encima de 34 °C, se libera la estructura secundaria del ARN mensajero (ARNm) de *rpoH* (Morita *et al.*, 1999), activando la HSR y, por lo tanto, la producción de genes que codifican para proteínas de choque térmico (Yamamori y Yura, 1980; Caspeta *et al.*, 2009; Valdez-Cruz *et al.*, 2010).

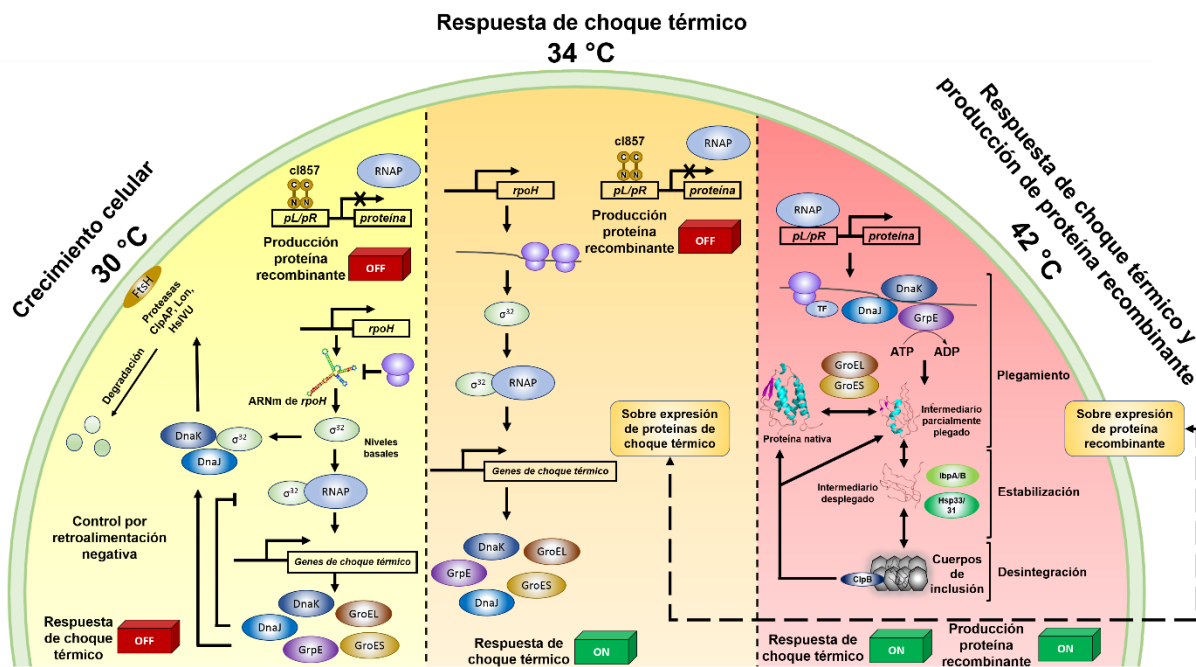


Figura 6. Relación entre los eventos moleculares que ocurren en *E. coli* durante la producción de una proteína recombinante usando el sistema de expresión termoinducible $\lambda pL/pR$ -cl857 (modificado de Restrepo-Pineda *et al.*, 2021).

En un sistema termoinducible, la temperatura continúa aumentando a 37 °C, donde la inactivación del represor cl857 da inicio a la transcripción del gen recombinante a partir de los promotores pL/pR (Villaverde *et al.*, 1993). Después de la inducción, las altas concentraciones de proteína recombinante que se alcanzan en este sistema pueden favorecer la formación de cuerpos de inclusión (Hoffmann y Rinas, 2004; Caspeta *et al.*, 2009), en cuyo proceso van a participar las chaperonas moleculares también sobre expresadas (Gill *et al.*, 2000). En consecuencia, ambas respuestas convergen en la activación de genes que codifican para proteínas de choque térmico (Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011).

En trabajos previos se ha reportado que en cultivos termoinducidos aumentan significativamente los niveles de ARNm de genes que codifican para algunas chaperonas (Harcum y Haddadin, 2006; Caspeta *et al.*, 2009). En los primeros 5-6 minutos después de la inducción térmica incrementan gradualmente hasta 9 veces los niveles de ARNm de DnaK, DnaJ y GroEL y después de 30-40 minutos, los niveles tienden a normalizarse (Valdez-Cruz *et al.*, 2011; Straus *et al.*, 1987).

Aunque hay diferentes reportes que han descrito la respuesta molecular de un sistema termoinducible a nivel del comportamiento de los ARNm de algunos genes de proteínas asociadas a choque térmico, son pocos los trabajos que analizan el comportamiento de los niveles de proteína. Hoffmann y Rinas (2000) estudiaron la cinética de la respuesta de choque térmico y la formación de cuerpos de inclusión después de la termoinducción de la

proteína recombinante hFGF-2 en *E. coli*, encontrando que a los 30 minutos después de la inducción a 42 °C se alcanza la velocidad máxima de síntesis de proteínas de choque térmico y que este punto, corresponde con el momento en el que la proteína recombinante comienza a presentarse mayoritariamente en forma soluble. En la proteómica de los agregados, 1 h post-inducción a 42 °C, observaron la presencia de DnaK y GroEL junto con las proteínas pequeñas IbpA e IbpB (Hoffmann y Rinas, 2000). Esta asociación se ha relacionado con la función de las chaperonas en el proceso de ensamble y desensamble de los cuerpos de inclusión durante la inducción térmica de la proteína hFGF-2. DnaK es esencial para la disolución de agregados, mientras que GroEL se ha sugerido que está involucrada en promover plegamiento de proteínas mal o parcialmente plegadas (Rinas *et al.*, 2007).

Recientemente en nuestro laboratorio, se encontró en geles de electroforesis de una dimensión que las chaperonas DnaK, DnaJ, GroEL y GroES se asocian diferencialmente a la fracción soluble y a los CI de la proteína recombinante dependiendo de las condiciones de cultivo. DnaK se detectó en las fracciones de la proteína soluble, mientras que GroEL se acumuló principalmente dentro de los agregados, tanto en matraces como en biorreactores (Restrepo-Pineda *et al.*, 2019). Sin embargo, hasta ahora no existe un reporte que describa el comportamiento cinético de los niveles de proteínas de choque térmico y su interacción con los cuerpos de inclusión en un sistema termoinducido, así como el posible papel de las chaperonas en las características estructurales y funcionales de los cuerpos de inclusión, e incluso en la obtención, rendimiento y calidad de la proteína recombinante de interés.

2.4. Proteína modelo: Factor Estimulante de Colonias de Granulocitos y Macrófagos

La proteína que se empleó como modelo durante el desarrollo de este proyecto es el Factor Estimulante de Colonias de Granulocitos y Macrófagos humano (HuGM-CSF). Esta molécula pertenece a un grupo de factores estimulantes de colonias (CSF), que son esenciales para la proliferación, diferenciación, maduración y activación funcional de células madre hematopoyéticas (Francisco-Cruz *et al.*, 2014; Hamilton, 2020). El gen que codifica para la proteína GM-CSF está localizado en el brazo largo del cromosoma 5 (5q31), junto con los genes para algunas interleucinas como IL-4, IL-5, IL-9, IL-12 y otros CSF (Metcalf, 2008). La estrecha disposición de los genes codificando para diversas citocinas y sus receptores, ha dilucidado el papel de estas moléculas en el desarrollo de células hematopoyéticas y en la regulación de una gran variedad de respuestas inflamatorias (Dougan *et al.*, 2019).

GM-CSF se sintetiza como una proteína precursora con 144 aminoácidos, la cual es procesada para producir una proteína madura de 127 aminoácidos con una masa molecular predicha de 14.4 kDa (Ghosh *et al.*, 2007). Está formada por cuatro hélices α y dos hojas β antiparalelas (**Figura 7**; Walter *et al.*, 1992; Rozwarski *et al.*, 1996). Contiene dos sitios consenso de N-glicosilación (44 y 54) y cuatro sitios de O-glicosilación (22, 24, 26 y 27) cerca al extremo N-terminal de la molécula (Schrimsher *et al.*, 1987). Además, tiene cuatro cisteínas involucradas en la formación de dos enlaces disulfuro: Cys54-Cys96 y Cys88-Cys121 (Schrimsher *et al.*, 1987; Walter *et al.*, 1992).

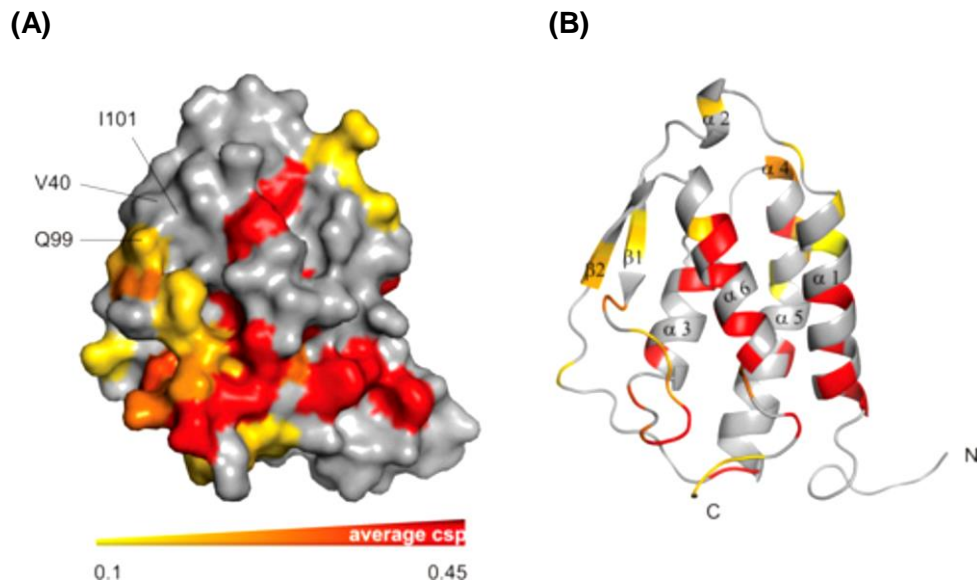


Figura 7. Representación de superficie **(A)** y caricatura **(B)** de la proteína humana GM-CSF con base en la estructura disponible en el Protein Data Bank (código PDB: 2GMF; Rozwarski *et al.*, 1996). Los colores representan los resultados del análisis de perturbación de desplazamiento químico, en inglés chemical shift perturbation (csp). Los residuos perturbados significativamente se colorean en gradiente de amarillo a rojo (modificado de Blech *et al.*, 2012).

La neutropenia es un condición inmunológica que consiste en la reducción del recuento de neutrófilos viables en sangre, la cual aumenta el riesgo y la gravedad de infecciones bacterianas y micóticas en pacientes que están sometidos a tratamiento por quimio/radioterapia, que sufren síndrome mielodisplásico o que han tenido trasplantes de médula ósea (Klastersky y Meert, 2016). Un estudio de Sobrevilla-Calvo y colaboradores (2016) en México, reporta que la mayoría de los esquemas de quimioterapia para varios tipos de cáncer son altamente mielotóxicos y generan un riesgo mayor del 20% para desencadenar neutropenia, por lo que se recomienda administrar al mismo tiempo un CSF de manera profiláctica (Sobrevilla-Calvo *et al.*, 2016). Ya se ha demostrado que la administración de G-CSF (filgrastim) y GM-CSF (molgramostim) recombinantes reduce la incidencia, duración y severidad de la neutropenia inducida por quimioterapia en diferentes tipos de cáncer, disminuyendo el riesgo de infecciones bacterianas (Cebon y Morstyn, 1990; Sung *et al.*, 2007). De ahí que la proteína GM-CSF se considere una molécula de relevancia terapéutica.

GM-CSF se ha expresado y purificado de manera recombinante en diferentes plataformas biológicas, incluyendo células de mamífero (CHO), levaduras (*Pichia pastoris* y *Saccharomyces cerevisiae*) y bacterias como *E. coli*. Se ha observado que la forma no glicosilada que se produce en *E. coli*, conocida como molgramostim es biológicamente activa (Okamoto *et al.*, 1991; Chen *et al.*, 2006). Incluso, fibroblastos y células endoteliales humanas pueden sintetizar y secretar naturalmente GM-CSF funcional que carece de todas las formas de modificación de carbohidratos (Kaushansky *et al.*, 1992). A pesar de las modificaciones post-traduccionales que contiene la proteína nativa, el rHuGM-CSF

producido en *E. coli* conserva su actividad como inmunomodulador (Burgess *et al.*, 1987). Recientemente, la forma recombinante (rHuGM-CSF) se ha utilizado como terapia en proteinosis alveolar pulmonar autoinmune (Trapnell *et al.*, 2020) y se está probando en ensayos clínicos contra COVID-19 (Lang *et al.*, 2020; Bonaventura *et al.*, 2020).

3. JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA

Algunos trabajos han demostrado que es posible controlar las condiciones de cultivo para aumentar los rendimientos de producción de una proteína recombinante de interés. En los sistemas termoinducidos se han modificado la temperatura y el tiempo de inducción, e incluso se han diseñado esquemas de oscilaciones de temperatura que han mejorado de 2 a 4 veces la acumulación de biomasa y disminuido en ~50% la acumulación de ácidos orgánicos en comparación a la inducción convencional (Caspeta *et al.*, 2013). Sin embargo, hasta ahora no se ha descrito cómo la estrategia de termoinducción afecta la dinámica de la respuesta de choque térmico y, cómo ésta a su vez puede afectar la composición proteica y la estructura de los CI de *E. coli* en un sistema inducido por temperatura. Tampoco se ha hecho énfasis en el cambio a nivel proteómico de los CI con respecto al tiempo del cultivo termoinducido.

Desde que se comenzó a estudiar la respuesta de choque térmico en *E. coli*, se demostró que el incremento en los niveles del factor σ^{32} ocurre tras un choque térmico de 30 °C a 34 °C (Straus *et al.*, 1987; Nagai *et al.*, 1991). A temperaturas de crecimiento por debajo de 34 °C (generalmente 30 °C), la región codificante 5' del ARNm de *rpoH* forma una estructura secundaria que oculta el sitio de unión del ribosoma y el codón de inicio, afectando la eficiencia de la traducción (Yuzawa *et al.*, 1993). Por encima de los 34 °C, la desnaturalización parcial de la estructura secundaria del ARNm permite la entrada del ribosoma y mejora el inicio de la traducción del factor σ^{32} y de las proteínas de choque térmico (Morita *et al.*, 1999).

Según esto, nos preguntamos si al cultivar una cepa de *E. coli* recombinante a 34 °C (y no a 30 °C) e inducir a 42 °C, se va a modificar la HSR y, esto a su vez, cambiará la composición y la arquitectura de los CI de la proteína rHuGM-CSF. Se esperaría que los resultados proteómicos y estructurales de los CI obtenidos a 30 °C contra 34 °C sean significativamente diferentes, con el fin de traducirlos en una propuesta metodológica útil para las industrias que producen proteínas biofarmacéuticas en forma de CI usando sistemas termoinducidos. Simplemente variando la temperatura de crecimiento o pre-inducción, se podrían tener CI con características deseadas, por ejemplo, enriquecidos de proteína recombinante y chaperonas y, de manera más importante, más fáciles de solubilizar, lo que disminuiría los costos del bioproceso.

En el presente trabajo, se evaluó el crecimiento, la expresión de las principales chaperonas de plegamiento y las propiedades estructurales de los CI a diferentes horas después de la termoinducción, variando el parámetro de la temperatura de pre-inducción: 30 °C vs 34 °C. La importancia de encontrar parámetros que permitan obtener una mayor cantidad de proteína recombinante activa de los cuerpos de inclusión, podría permitir el desarrollo de un bioproceso mejorado para producir proteínas terapéuticas en sistemas termoinducibles.

4. HIPÓTESIS

La temperatura de pre-inducción (30 °C vs 34 °C) modifica la expresión de chaperonas y las características estructurales de los cuerpos de inclusión del factor estimulante de colonias de granulocitos y macrófagos recombinante humano (rHuGM-CSF) en un sistema termoinducible de *E. coli*.

5. OBJETIVOS

5.1. Objetivo general

Evaluar cinéticamente el efecto de la temperatura de pre-inducción (30 °C vs 34 °C) sobre el crecimiento celular, la expresión de chaperonas de plegamiento y las características estructurales de los cuerpos de inclusión durante la termoinducción de la proteína rHuGM-CSF en *E. coli*.

5.2. Objetivos específicos

- Caracterizar cinéticamente el crecimiento de la cepa *E. coli* W3110 productora de la proteína rHuGM-CSF en biorreactores con medio de cultivo definido, variando la temperatura de pre-inducción (30 y 34 °C) con posterior inducción a 42 °C.
- Cuantificar el consumo de la fuente de carbono y la producción de acetato durante los cultivos termoinducidos.
- Comparar los rendimientos de producción de la proteína recombinante rHuGM-CSF bajo ambas temperaturas de pre-inducción, tanto en los lisados celulares, como en las fracciones de proteína soluble y en los cuerpos de inclusión.
- Identificar la presencia de las principales chaperonas de plegamiento DnaK, DnaJ, GroEL y GroES en los cuerpos de inclusión, evaluando su cinética de expresión durante la termoinducción.
- Determinar las características estructurales de los cuerpos de inclusión de la proteína rHuGM-CSF, tras su recuperación de cultivos con diferentes temperaturas de pre-inducción.

6. METODOLOGÍA

La estrategia metodológica que se plantea para cumplir cada uno de los objetivos propuestos se divide en tres etapas. La etapa inicial consiste en la caracterización de la cinética de crecimiento de la cepa recombinante, la evaluación del consumo de la fuente de carbono y la acumulación de ácidos orgánicos durante los cultivos. La siguiente parte consta del uso de las técnicas de biología molecular para analizar la expresión de la proteína recombinante y algunas proteínas de choque térmico. A la par de la segunda etapa, se llevará a cabo la última parte del proyecto, que incluye las técnicas de análisis estructural de los cuerpos de inclusión obtenidos bajo dos temperaturas de pre-inducción diferentes: 30 °C y 34 °C con posterior inducción a 42 °C.

6.1. Materiales

6.1.1. Cepas bacterianas

Cepa recombinante: Los experimentos se realizaron con una cepa recombinante de *Escherichia coli* W3110 (ATCC® 27325™) productora de la proteína rHuGM-CSF, la cual tiene una copia del gen del represor termolábil cl857 en el cromosoma y lo expresa de forma constitutiva (Product sheet: *Escherichia coli* Castellani and Chalmers ATCC® 27325™). Esta cepa recombinante fue donada generosamente para este proyecto por el Dr. Néstor O. Pérez de la empresa mexicana PROBIOMED S.A. de C.V.

Cepa silvestre: Como control negativo, se utilizó la cepa silvestre de *Escherichia coli* W3110 (ATCC® 27325™), proveniente del laboratorio del Dr. Guillermo Gosset del Instituto de Biotecnología de la UNAM.

6.1.2. Plásmido

La cepa de *E. coli* W3110 (ATCC 27325) fue transformada con el plásmido de expresión pV3/HuGM-CSF que contiene el gen codificante para la proteína GM-CSF humana. Este plásmido posee 5,123 pares de bases (pb), cuenta con el promotor fuerte pL del bacteriófago λ y con un gen de selección de resistencia a la kanamicina (**Figura 8, Tabla 2**).

6.1.3. Medio de cultivo

Se empleó un medio de cultivo definido con glucosa como principal fuente de carbono, el cual fue modificado a partir de la composición descrita anteriormente por Caspeta *et al.* (2013), también para cultivos de *E. coli* recombinante usando un sistema termoinducido.

La composición del medio se presenta en la **Tabla 3**, donde se describe la solución base (KH_2PO_4 , $(\text{NH}_4)_2\text{HPO}_4$ y ácido cítrico), que se prepara y esteriliza antes de cada cultivo. Además, se presentan las soluciones de elementos traza (1000X), tiamina (1000X), kanamicina (1000X), MgSO_4 (400 g/L), glucosa (500 g/L) y casaminoácidos (500 g/L), que fueron preparadas previamente y adicionadas a la solución base antes de inocular. Se agregó 1 ml por L de medio de la solución stock de elementos traza 1000X conteniendo (g/L): Citrato de hierro III (100.8); $\text{ZnSO}_4 \cdot 2\text{H}_2\text{O}$ (22.5); $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ (15.0); EDTA (14.1); H_3BO_3 (3.0); $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ (2.5); $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ (2.1); $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ (1.5), descrita en Caspeta

et al., 2013. Las soluciones de tiamina, kanamicina, glucosa y casaminoácidos se esterizaron por filtración usando membranas de acetato de celulosa de 0.22 μm (*Millipore*). La solución base fue ajustada a un pH de 7.0 con NaOH 1M y se esterilizó a 121°C y 22 psig por 30 minutos en un autoclave Tomy ES-215 (*Sterilizers*).

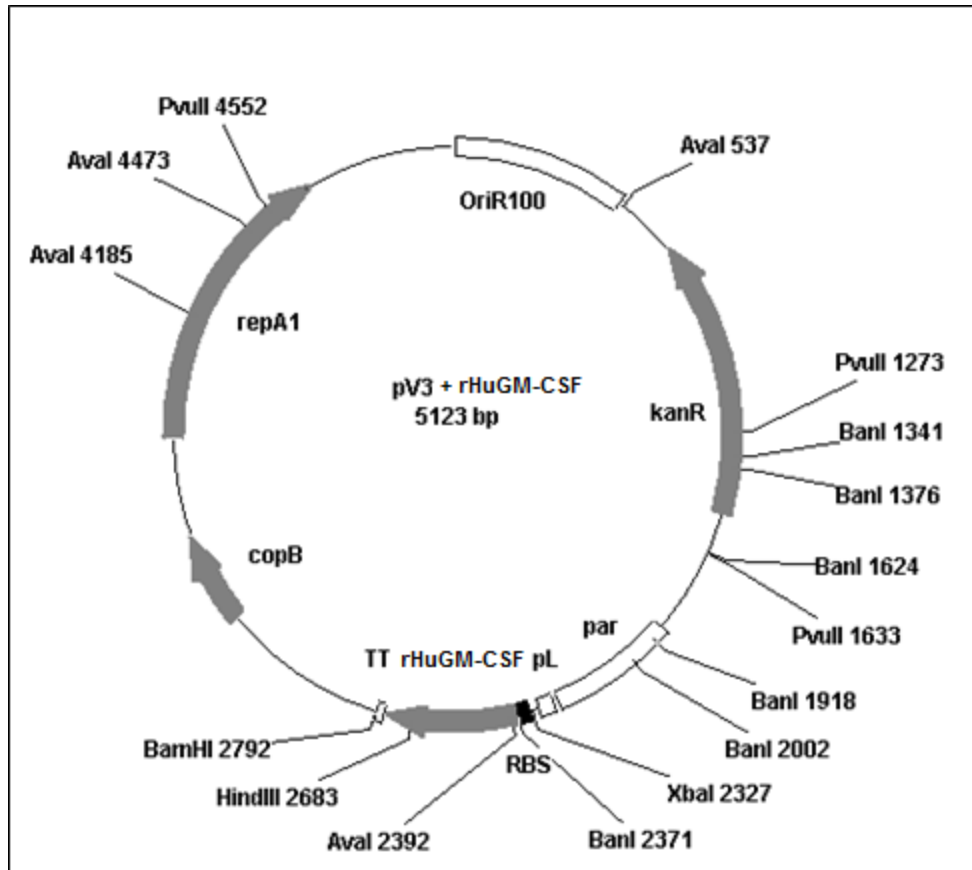


Figura 8. Mapa del plásmido pV3/HuGM-CSF y ubicación de sus elementos genéticos.

La idea de adicionar casaminoácidos al medio se basó en algunos reportes que los proponen como una buena fuente de nitrógeno, proveen nutrientes y cofactores requeridos para el correcto crecimiento de las cepas recombinantes, además poseen un alto grado de digestión y una rápida asimilación por parte de las bacterias (*Fishov et al.*, 1995). Jung (2006) reportó que al utilizar casaminoácidos como fuente de nitrógeno orgánico a una concentración entre 3.0 y 3.2 g/L, se pueden obtener rendimientos de interferón- α recombinante de hasta 10 mg/g usando un reactor de 1.0 L en cultivos de *E. coli* (Jung., 2006). Además, se puede incrementar la velocidad específica de crecimiento de las células (*Marr y Ingraham*, 1962).

Tabla 2. Elementos genéticos del plásmido pV3.

Elemento genético	Función	Posición
<i>OriR100</i>	Origen de replicación	6-520
<i>kanR</i>	Resistencia a kanamicina para seleccionar bacterias recombinantes	714-1508
<i>par</i>	Secuencia de partición	1878-2245
<i>pL</i>	Promotor termoinducible controlado por la proteína cl857	2262-2307
<i>RBS</i>	Sitio de unión a ribosoma	2334-2368
<i>rHuGM-CSF</i>	Codifica para el gen optimizado de HuGM-CSF	2369-2755
<i>TT</i>	Terminador transcripcional	2765-2786
<i>copB</i>	Mantiene bajo el número de copias del plásmido	3283-3543
<i>repA1</i>	Proteína iniciadora de la replicación	3833-4690

Tabla 3. Composición del medio de cultivo definido (modificado de Caspeta *et al.*, 2013).

Solución base (g/L)		Soluciones Stock (concentración final en el medio)	
(NH ₄) ₂ HPO ₄	4.0	MgSO ₄ .7H ₂ O	1.2 g/L
KH ₂ PO ₄	13.3	Tiamina	0.045 g/L
Ácido cítrico	1.7	Kanamicina	0.03 g/L
		Glucosa	17.5 g/L
		Elementos traza	1 X
		*Casaminoácidos	3.0 g/L

*Casaminoácidos: Modificación de la formulación original del medio.

6.2. Métodos

6.2.1. Generación del banco maestro de la cepa *E. coli* W3110 productora de rHuGM-CSF

Se realizó un banco maestro con el fin de tener una colección de células genéticamente homogéneas. De la cepa recombinante *E. coli* W3110 + HuGM-CSF crecida en una caja de Petri con agar LB (proveniente de la empresa PROBIOMED S.A. de C.V.), se tomaron directamente dos colonias y se inocularon por separado en matraces Erlenmeyer de 250 ml con 50 ml de medio LB y 50 µL de kanamicina (concentración final: 30 µg/ml). Se dejaron en incubación a 30 °C y 250 rpm en una incubadora *New Brunswick Scientific Classic C25*, hasta alcanzar una densidad óptica (DO) de 1.3 unidades de absorbancia (u.a.) a 600 nm. Para conservar las células se utilizó 40 % de glicerol puro estéril, haciendo 50 alícuotas de 1000 µL, las cuales se almacenaron a -75 °C.

6.2.2. Generación del banco de trabajo de la cepa *E. coli* W3110 productora de rHuGM-CSF

Para el banco de trabajo, se inocularon 500 μ L de un criovial del banco maestro en tres matraces Erlenmeyer de 250 ml con 50 ml de medio LB y 50 μ L de kanamicina (concentración final: 30 μ g/ml). Se dejaron en incubación a 30 °C y 250 rpm en una incubadora *New Brunswick Scientific Classic C25*, hasta alcanzar una DO de 1.09 unidades de absorbancia (u.a.) a 600 nm. Para conservar las células se utilizó 40 % de glicerol puro estéril, haciendo 50 alícuotas de 1000 μ L, las cuales se almacenaron a -75 °C.

6.2.3. Cinética de crecimiento de *E. coli* W3110 productora de rHuGM-CSF en matraces agitados

Se hicieron cultivos de la cepa recombinante *E. coli* W3110 + rHuGM-CSF en matraces agitados con medio de cultivo definido. Se utilizaron matraces Erlenmeyer convencionales de 250 ml con 50 ml de volumen de llenado. A cada matraz conteniendo solución base se le agregaron las cantidades correspondientes de las soluciones stock que se especificaron en la **Tabla 3** y se inocularon con 500 μ L provenientes de un criovial del banco de trabajo. Los matraces se incubaron a 30 °C (*New Brunswick Scientific Classic C25, USA*) y se hizo la inducción subiendo la temperatura hasta 42 °C al alcanzar la fase pre-estacionaria (DO de ~0.9-1.2. u.a.). Como control se mantuvieron matraces sin inducir (30 °C). En todos los casos la agitación fue constante a 200 rpm y se tomaron muestras para la medición de la DO y análisis de la expresión de la proteína recombinante. Todas las cinéticas se realizaron por triplicado.

6.2.4. Preparación de pre-inóculos para los cultivos en biorreactor

Para que todos los cultivos en biorreactor comenzaran con una misma DO (0.1 u.a.), se emplearon matraces agitados como pre-inóculos. Cada pre-inóculo se preparó con 500 μ L provenientes directamente de un criovial del banco de trabajo de *E. coli* W3110 + rHuGM-CSF más 50 μ L de kanamicina usando matraces Erlenmeyer de 250 ml con 50 ml de medio. Los pre-inóculos se dejaron durante toda la noche (~14 h) a 30 °C y 200 rpm (*New Brunswick Scientific Classic C25*). Se midió la DO y se calculó el volumen que debía tomarse para alcanzar la concentración deseada, usando la ecuación [1]:

$$(C1 * V1) = (C2 * V2) \quad [1]$$

donde C1 es el valor de la DO medida, C2 corresponde a 0.1 u.a., V2 es el volumen final del cultivo y V1 es el volumen que debe tomarse desde el pre-inóculo para inocular el cultivo.

Todos los pre-inóculos se prepararon utilizando el mismo medio en el que se iba a poner el cultivo, para evitar cualquier tipo de variación metabólica por adaptación a un medio nuevo (Trujillo-Roldán *et al.*, 2003). En cada caso, después de inocular se midió la DO (muestra del tiempo cero), para verificar que se comenzaba con valores cercanos a 0.1 u.a.

6.2.5. Cinética de crecimiento de *E. coli* W3110 productora de rHuGM-CSF en biorreactores

Se realizaron cultivos de la cepa recombinante *E. coli* W3110 + rHuGM-CSF en biorreactores de 1.2 L (*Applikon, Netherlands*) con un volumen de operación de 800 ml de medio de cultivo definido. Todos los cultivos se iniciaron a una DO de ~0.1 u.a. usando cultivos en matraces como pre-inóculos. Los biorreactores se instrumentaron con sensores de pH, oxígeno disuelto y temperatura, haciendo un seguimiento en línea con el programa de adquisición de datos BioXpert® de *Applikon, Netherlands*. La tensión de oxígeno disuelto (TOD) fue controlada a 35 % con respecto a la saturación del aire, usando cascada de agitación entre 100-1000 rpm y aireación de 1.0 vvm. El pH del medio fue mantenido en ~7.0 mediante la adición de NaOH 3 M o HCl 3 M y la temperatura fue controlada ya sea con una camisa de calentamiento o por medio de un sistema circulación de agua. Para el control del nivel de espuma se adicionaron 100 µL de un agente antiespumante estéril cuando fue necesario durante el cultivo.

Los cultivos en biorreactor se dejaron en crecimiento a las dos temperaturas evaluadas, ya sea 30 °C o 34 °C, hasta alcanzar la fase pre-estacionaria (DO de ~2.0-3.0 u.a.). En este punto, se hizo la inducción de la proteína recombinante aumentando la temperatura hasta 42 °C para ambos casos. Se mantuvo la misma aireación y se continuó con el control de TOD y pH. Se hizo un seguimiento del crecimiento midiendo la DO del cultivo. Al mismo tiempo se tomaron muestras de 1 ml, se centrifugaron a 4 °C/10,000 x g/10 min y se guardó tanto sobrenadante como pellet a -20°C para la determinación de la expresión de la proteína recombinante, medición de la glucosa y ácidos orgánicos. Todas las cinéticas se realizaron por triplicado.

6.2.6. Medición de la absorbancia del cultivo

El crecimiento del cultivo de *E. coli* W3110 + rHuGM-CSF se determinó midiendo la densidad óptica (DO) a 600 nm en un espectrofotómetro Spectronic Genesys™ 5 (*Thermo, USA*). Las muestras se midieron directamente en el equipo hasta alcanzar concentraciones de 0.7-0.9 u.a. Por encima de estos valores se hicieron diluciones (1:10 o de 1:20) con el mismo medio de cultivo para que la absorbancia medida estuviera dentro del intervalo de respuesta lineal del equipo, siguiendo la ley de Beer-Lambert (Koch, 1970; Toennies y Gallant, 1949; Myers *et al.*, 2013).

6.2.7. Determinación de la biomasa como peso seco

Se secaron membranas nuevas de acetato de celulosa de 0.22 µm de tamaño de poro (*Millipore*) a 90°C en un horno Precision (*Equipar*), luego se pesaron en una balanza analítica Sartorius CP225D y se almacenaron en desecador. Se realizó una cinética de crecimiento de la cepa recombinante en matraces agitados de 250 ml con 50 ml de medio de cultivo definido a 30 °C y 200 rpm, por duplicado. Se hizo seguimiento de la DO y se tomaron muestras de 10 ml que se fueron centrifugaron a 4 °C/8,000 rpm/10 min. El sobrenadante se descartó cuidadosamente y el pellet se resuspendió en 5 ml de PBS 1X, filtrando cada vez. La pasta celular húmeda sobre los filtros pre-pesados se llevó a secar a 90 °C y se pesó periódicamente hasta obtener un peso constante. La concentración de la biomasa (g/L) se calculó como la diferencia entre el peso de los filtros vacíos y los filtros con la biomasa seca, teniendo en cuenta el volumen de muestra. Finalmente, se hizo una correlación lineal entre el peso seco de la biomasa y la DO a 600 nm, para encontrar un

factor de equivalencia entre ambas medidas (Lawrence y Maier, 1977; Toennies y Gallant, 1949).

6.2.8. Cuantificación de glucosa

Las muestras obtenidas en cada punto de la cinética se centrifugaron a 4 °C/10,000 x g/10 min. Los sobrenadantes provenientes de 1 ml de cultivo se usaron para medir la glucosa en un equipo Biochemistry Analyzer YSI 2900D. Para los lavados y la medición se empleó el buffer de fosfatos YSI N°2357, la membrana de glucosa oxidasa YSI N°2365 y una solución calibradora con 2.5 g/L de glucosa preparada en buffer de fosfatos (40 g/L de NaH₂PO₄ y 10 g/L de Na₂HPO₄) como indica el proveedor.

La medición que realiza este analizador bioquímico está basada en un proceso enzimático acoplado a un proceso electroquímico, en el cual la enzima (glucosa oxidasa) se inmoviliza en una membrana que se encuentra entre dos películas, una de policarbonato y otra de acetato de celulosa. El sustrato es oxidado al pasar por la membrana de la oxidasa inmovilizada, produciendo peróxido de hidrógeno, el cual a su vez pasa a través de la membrana de acetato de celulosa para interactuar con el ánodo de platino, generando electrones. El flujo de electrones es proporcional a la concentración de peróxido de hidrógeno y, por tanto, a la concentración de sustrato (*YSI Life Sciences, USA*).

6.2.9. Cuantificación de acetato

La concentración de acetato se cuantificó mediante cromatografía líquida de alta resolución (HPLC, por sus siglas en inglés). Se empleó un equipo *Shimadzu*® compuesto de una unidad de bombas LC-20AT, un detector UV a 210 nm SPD-20A y un regulador de temperatura por circulación de aire CTO-104. Los sobrenadantes provenientes de cada cultivo se pasaron por filtros de 0.20 µm y luego se inyectaron 100 µL de muestra en el equipo. La separación se llevó a cabo usando una columna cromatográfica Aminex® HPX-87H (300 x 7.8 mm) de *BIO-RAD* y para la cuantificación de los metabolitos se empleó una solución estándar comercial (No. catálogo 125-0586 *BIO-RAD*). Como fase móvil se usó H₂SO₄ 0.008N a una velocidad de flujo de 0.6 ml/min y 50°C. Los datos obtenidos fueron procesados en el programa LC Solution (*Shimadzu, Kioto, Japón*).

6.2.10. Recuperación y purificación de los cuerpos de inclusión

Las muestras obtenidas en cada punto del cultivo se centrifugaron a 4 °C/10,000 x g/10 min. El pellet proveniente de 1 ml de cultivo se resuspendió en Buffer de lisis celular a pH= 7.5 (Tris-HCl 50 mM, NaCl 100 mM, EDTA 1mM e inhibidor de proteasas 0.1mM). La suspensión celular fue sonicada en un equipo Soniprep150 (*SANYO-GallenKamp, UK*) a 8 micrones de amplitud por 1 minuto en 3-5 ciclos, dejándola en frío (4 °C) entre cada ciclo. El lisado se centrifugó a 4 °C/14,000 x g/15 min, pero en este caso se recuperó tanto el sobrenadante, que contiene la proteína total soluble (S) como el pellet celular correspondiente a la proteína insoluble (I).

La fracción insoluble se incubó en buffer de lisis con 1 % del detergente no iónico IGEPAL (octylphenoxypolyethoxyethanol) durante 30 min en agitación a 4 °C (cuarto frío). Posteriormente, se centrifugó a 4 °C/14,000 x g/15 min, descartando el sobrenadante. El pellet se lavó con buffer de Tritón X-100 al 0.5 % por 1 hora a 4 °C. Nuevamente, se centrifugó a 4 °C/14,000 x g/15 min, se descartó el sobrenadante y se lavó el pellet entre 3

a 5 veces consecutivas con agua desionizada, centrifugando entre cada lavado. Después de la última centrifugación se eliminó el sobrenadante y se almacenó el pellet a -20 °C, el cual corresponde a los cuerpos de inclusión (CI) purificados (Rodríguez-Carmona *et al.*, 2010; Castellanos-Mendoza *et al.*, 2014; Calcines-Cruz *et al.*, 2018).

6.2.11. Cuantificación de proteína

La concentración de proteína se determinó por el método estándar de Bradford (N°5000006 de *BIO-RAD*), que se basa en la unión del colorante azul brillante de Coomassie G-250 a los aminoácidos básicos y aromáticos de las proteínas (Bradford, 1976). Se realizaron curvas de calibración con estándares de albúmina de suero bovino (BSA N°BAC65 de *Equitech-Bio*), en un intervalo de concentración de 8 a 500 µg/ml. Las muestras se solubilizaron en IEF 1:5. Placas de 96 pozos de fondo plano se usaron para medir la absorbancia a 600 nm en el lector Stat Fax 4200 de *Awareness Technology, Inc.* Tanto las muestras como los estándares se prepararon por duplicado.

6.2.12. Análisis de proteína por electroforesis en geles de poliacrilamida

Para confirmar la expresión de la proteína de interés y obtener información sobre su tamaño molecular y pureza, se llevó a cabo una electroforesis en geles de poliacrilamida o SDS-PAGE (Sambrook *et al.*, 1989). Los CI se disolvieron en un agente desnaturante, SDS (dodecilsulfato sódico) al 2.5%, en una proporción 1:4 durante 12 h a temperatura ambiente. El gel separador se preparó al 15% de poliacrilamida y fue precedido por un gel concentrador al 4%. La cantidad de la muestra cargada a cada pozo fue volumétrica (30 µL). Para la corrida del gel se utilizó una cámara de electroforesis (*Hoefler Inc.*), a corriente constante (80-100 voltios). Los geles fueron teñidos con azul de Coomassie durante toda la noche y posteriormente el exceso de colorante se retiró con una solución de desteñido que contenía 7 % de ácido acético, 70 % de metanol y se aforó con agua. Las imágenes del gel se digitalizaron en un equipo Gel Doc™ EZ Imager (*BIO-RAD*), utilizando el programa *Image Lab 5.2*. Con este mismo programa, se realizó un análisis densitométrico de los geles para determinar los niveles de expresión de proteína en cada condición. Este es un método semi-cuantitativo y sencillo que mide las intensidades de bandas específicas. La densidad de cada banda se representa como un área bajo la curva y se le asigna un porcentaje con respecto al total de las densidades de un carril. De esta manera, fue posible determinar el porcentaje que representa la proteína recombinante con respecto a la cantidad de proteína total y a la proteína en cuerpos de inclusión.

6.2.13. Inmunodetección de chaperonas por Western Blot

Los geles SDS-PAGE fueron transferidos a una membrana de PVDF (Immobilon™ de *Millipore*) previamente activada con metanol puro, usando un sistema semi-húmedo con buffer de transferencia (Tris-Glicina 1X; Metanol 20%) durante 40 minutos a 23 voltios (Transblot, *BIO-RAD*).

A cada membrana transferida se le realizó un bloqueo de 40 min en agitación suave a temperatura ambiente con una solución de leche descremada en polvo al 5 % disuelta en TBS. Después del bloqueo, se hicieron dos lavados consecutivos de la membrana con una solución de lavado (TBS + Tween 20 al 0.06%) durante 10 min cada uno. Luego se incubó con el anticuerpo primario respectivo (anti-DnaK/DnaJ/GroEL/GroES) preparado a la concentración indicada en la Tabla 4. La membrana se lavó tres veces y se incubó con el

anticuerpo secundario. Finalmente, se hicieron dos lavados más y la membrana se trató con luminol acoplado a peroxidasa (Kit Supersignal West PICO N°34080+ FEMTO N°34095/Chemiluminescent Substrate de *Pierce, Thermo-Scientific*). El revelado se hizo por quimioluminiscencia en un equipo *C-Digit* acoplado al programa *Image Studio Digits 3.1*. Se usó un marcador de quimioluminiscencia que es visible en las imágenes.

Tabla 4. Anticuerpos anti-chaperonas de *E. coli* usados para realizar Western blot de cultivos termoinducidos.

Anticuerpo primario	Concentración (en TBS 1X)	Tiempo de incubación	Anticuerpo secundario
Anti-DnaK de Enzo Life Sciences N°ADI-SPA-880	1:7,500	50 min	Anti-ratón de <i>Sigma-Aldrich</i> N°A3562
Anti-DnaJ de Enzo Life Sciences N°ADI-SPA-410.	1:2,000	12 h a 4 °C	Anti-conejo de <i>Sigma-Aldrich</i> N°A3687.
Anti-GroEL de Enzo Life Sciences N°ADI-SPS-875.	1:7,500	50 min	Anti-conejo de <i>Sigma-Aldrich</i> N°A3687.
Anti-GroES de Enzo Life Sciences N°ADI-SPA-210.	1:7,500	50 min	Anti-conejo de <i>Sigma-Aldrich</i> N°A3687.

Con el fin de poder usar la misma membrana para buscar diferentes proteínas, es necesario remover los anticuerpos primario y secundario de la membrana de Western blot. Se empleó el protocolo de “stripping for reprobing” de Abcam, usando un buffer de “stripping” suave a pH=2.2 (Glicina, SDS, Tween 20).

6.2.14. iTRAQ: Marcaje isobárico para la cuantificación absoluta y relativa de proteínas

El iTRAQ es una estrategia de proteómica “shotgun” para la identificación y cuantificación de péptidos usando espectrometría de masas en tándem, la cual permite el análisis proteómico en paralelo de más de dos muestras (Füller y Morris, 2012). Este método utiliza reactivos isobáricos específicos para etiquetar las aminas de péptidos provenientes de cuatro u ocho muestras biológicas diferentes, dependiendo de kit que se elija para trabajar. Los péptidos marcados de cada muestra se mezclan, se separan usando cromatografía líquida bidimensional y se analizan utilizando espectrometría de masas en tándem (Aggarwal *et al.*, 2006).

Para llevar a cabo esta técnica se empleó el kit iTRAQ™ Reagent Multi-Plex (N°94404 de *Applied Biosystems, USA*) diseñado para analizar cuatro muestras simultáneamente. Este kit contiene los reactivos o banderas para el marcaje (114, 115, 116 y 117), un buffer de disolución, un agente bloqueador de cisteínas, un agente reductor, un agente desnaturizante (SDS 2 %) y etanol. El protocolo fue suministrado por el proveedor. Las muestras seleccionadas para el marcaje fueron los cuerpos de inclusión obtenidos de los cultivos a 30 °C (1 h y 5 h después de inducir) y 34 °C (1 h y 5 h después de inducir). En resumen, el primer paso fue solubilizar y desnaturizar 100 µg de los CI con 0.1% de SDS durante 12 horas. Luego, las proteínas fueron reducidas con ditioneitol (DTT) y tratadas con yodoacetamina, un agente bloqueador de las cisteínas. Se digirieron con la enzima tripsina a 37 °C durante toda la noche. Las muestras digeridas fueron marcadas con los cuatro reactivos del kit, incubando por mínimo 2 horas a temperatura ambiente. Finalmente, se

hizo la combinación de los péptidos marcados en un mismo tubo para el posterior análisis por espectrometría de masas. Todo el procedimiento de iTRAQ fue llevado a cabo en el Instituto Nacional de Medicina Genómica (INMEGEN) de la Ciudad de México.

6.2.15. Espectrometría de masas acoplada a HPLC

La mezcla que se obtuvo al final del procedimiento de iTRAQ conteniendo los péptidos marcados de las cuatro condiciones experimentales, se trató con puntas ZipTip (*Millipore, MA*). Estas puntas tienen acoplada internamente una columna C18, la cual permite una extracción en fase sólida para la concentración de la muestra y la remoción de sales que pudieron quedar después del proceso de digestión y marcaje.

Luego, la muestra limpia y concentrada se resuspendió en la solución móvil del HPLC (5 % de acetonitrilo con 0.1 % de ácido fórmico) que está acoplado al espectrómetro de masas. El equipo que se usó para la separación por fase reversa fue un UHPLC (Ultra High Performance Liquid Chromatography) Ultimate 3000 acoplado a un espectrómetro de masas Q Exactive Plus Hybrid Quadrupole-Orbitrap (*Thermo Fisher Scientific, USA*) de la Unidad de Proteómica del Centro de Ciencias Genómicas (CCG) de la UNAM. El espectrómetro de masas Q Exactive Plus es un híbrido que fusiona el alto poder de resolución de un analizador de masas Orbitrap y la selectividad del Quadrupolo. Por lo tanto, el equipo arroja datos reproducibles de alta calidad para la caracterización de una amplia variedad de compuestos con mayor confianza; además mejora la cuantificación de iones de baja abundancia en muestras complejas (Sollicet *et al.*, 2015).

Los datos obtenidos a partir del espectrómetro de masas fueron analizados en el programa Proteome Discoverer™ 2.1 SP1 (*Thermo Fisher Scientific, USA*). Este programa utiliza flujos de trabajo para procesar datos de espectrometría de masas y compararlos con la información de una base de datos, logrando identificar proteínas a partir de los fragmentos digeridos. El flujo de trabajo se fue construyendo según las características del experimento realizado (ver resultados). Este programa es manejado en conjunto por la Unidad de Proteómica del INMEGEN y por la Unidad de Proteómica del CCG, los cuales nos facilitaron amablemente el acceso para el procesamiento de los datos. Como alternativa, se usó Scaffold (*Proteome Software, USA*), que es una herramienta bioinformática más simple de entender y manipular que el Proteome Discoverer para visualizar y validar experimentos de proteómica por espectrometría de masas. Igualmente, el acceso a la aplicación fue facilitado por la Unidad de Proteómica del INMEGEN.

6.2.16. Espectroscopia infrarroja por transformada de Fourier (FTIR)

Para el análisis de la estructura secundaria se utilizaron CI purificados en fresco, los cuales se secaron previamente a temperatura ambiente en un rotoevaporador durante 30-40 minutos. Se usó un espectrofotómetro infrarrojo IRAffinity-1S (*Shimadzu, Kioto, Japón*) acoplado a un dispositivo de reflexión total atenuada (ATR) para obtener los espectros de las muestras en estado sólido. Un total de 40 interferogramas fueron recolectados y promediados con una resolución de 2 cm^{-1} en el rango de 1500-1700 cm^{-1} . Se realizó la segunda derivada de los espectros de absorbancia de la región amida I después de un suavizado de 13 puntos, usando el programa IR LabSolutions (*Shimadzu, Kioto, Japón*).

Esta segunda derivada se dividió por el valor absoluto del pico de tirosina ($\sim 1508 \text{ cm}^{-1}$) para evitar variaciones por diferentes cargas de proteínas sobre el ATR (Ami *et al.*, 2005). Finalmente, las frecuencias en las que se ubicaron los diferentes componentes espectrales se utilizaron para la asignación de la estructura secundaria en los CI (Valdez-Cruz *et al.*, 2017).

6.2.17. Ensayo de fluorescencia con Tioflavina T

El contenido de estructura de tipo amiloide en los CI, se evaluó mediante un ensayo de unión al colorante tioflavina T (Th-T). Se resuspendió una concentración de CI de $50 \mu\text{g/ml}$ en PBS, pH 7.5 y se mezclaron con Th-T $75 \mu\text{M}$ (*Sigma-Aldrich, USA*) incubando durante 30 minutos a $25 \text{ }^\circ\text{C}$. La longitud de onda de excitación fue de 440 nm , y los espectros de emisión se registraron entre 450 y 560 nm en un espectrofotómetro de fluorescencia Cary Eclipse (Agilent Technologies, USA). Como control se midió la fluorescencia base del Th-T ($75 \mu\text{M}$) en PBS (Calcines-Cruz *et al.*, 2018; Castellanos-Mendoza *et al.*, 2014; Upadhyay *et al.*, 2012).

6.2.18. Digestión proteolítica con proteinasa K

Los CI fueron digeridos usando Proteínasa K (*BIOLINE, UK*) a una concentración final de $25 \mu\text{g/ml}$. La digestión se llevó a cabo con $50 \mu\text{g/ml}$ de proteína de CI purificados, que fueron diluidos en un buffer amortiguador (50 mM de Tris-HCl y 150 mM de NaCl, pH 8.0) para un volumen final de 1 ml . Los cambios en la absorbancia se monitorearon a 350 nm durante 100 min en un espectrofotómetro UV/Vis DU®730 (*Beckman coulter, USA*). Los resultados de absorbancia se normalizaron respecto al valor inicial de DO_{350} justo después de la mezcla con la proteínasa K (Calcines-Cruz *et al.*, 2018; Castellanos-Mendoza *et al.*, 2014; Upadhyay *et al.*, 2012; De Groot y Ventura, 2006).

6.2.19. Estabilidad en cloruro de guanidinio (GndHCl)

Se determinó la estabilidad de los CI en presencia de diferentes concentraciones de GndHCl ($0, 1, 3, 5 \text{ M}$) durante 24 h a temperatura ambiente en agitación. Las soluciones del agente caotrópico fueron preparadas en buffer Tris-HCl 10 mM a pH 7.5 y se almacenaron a $4 \text{ }^\circ\text{C}$. La concentración final de proteína en las soluciones de desnaturante fue de 0.20 mg/ml , para un volumen final de $100 \mu\text{L}$. Después de 24 h de incubación, la suspensión fue centrifugada a 8000 g durante 10 min y los sobrenadantes conteniendo la proteína solubilizada se cuantificaron por el método de Bradford (Castellanos-Mendoza *et al.*, 2014; De Groot y Ventura, 2006; Espargaró *et al.*, 2008).

6.2.20. Solubilización, plegamiento y purificación

Los CI purificados de rHuGM-CSF se resuspendieron en una solución de solubilización (6 M de GndHCl, 50 mM de 2-mercaptoetanol en Tris base de 100 mM , pH 8.2, ajustado con HCl 3 M) a 700 rpm , 1 h y $25 \text{ }^\circ\text{C}$ usando una concentración final de proteína de 1.0 mg/ml . Luego, se llevó a cabo el proceso de plegamiento a $25 \text{ }^\circ\text{C}$, 100 rpm y 4 h adicionando gota a gota (1:10) el producto solubilizado en una solución tampón de Tris base (20 mM , pH 8.2). La reacción se detuvo agregando ácido acético 4 M h (Burgess *et al.*, 1987; Belew *et al.*, 1994; Thomson *et al.*, 2012). La solución de proteína plegada se purificó por HPLC de fase

reversa en un Shimadzu LC-20AT (*Shimadzu, Kioto, Japón*) utilizando una columna Zorbax Eclipse XDB-C8 (*Agilent Technologies, Santa Clara, CA, EE. UU.*). Se realizó una curva estándar con la molécula GM-CSF de referencia de la Farmacopea Europea (Molgramostim, Y0000251, *Sigma Aldrich, St. Louis, MO, EE. UU.*), usando concentraciones de 1.345 mg/ml, 0.672 mg/ml y 0.336 mg/ml. La muestra de proteína plegada (~ 10 µg) fue cargada a la columna del HPLC a 50°C con una longitud de onda del detector de 214 nm y una presión de columna máxima de 3000 psig. Para la elución en gradiente lineal a una velocidad de 1.2 ml/min se utilizaron soluciones de ácido trifluoroacético (TFA) al 0,1% v/v en agua (fase móvil A) y TFA al 0.1 % v/v en acetonitrilo al 90 % (fase móvil B) (Nicola *et al.*, 1983; Das *et al.*, 2011).

6.2.21. Espectroscopia de dicroísmo circular (DC)

Se registraron espectros de dicroísmo circular de la proteína rHuGM-CSF en la región de UV lejano a 37 °C con un espectropolarímetro JASCO J-720 (*Jasco Inc., Easton, MD*). Soluciones de proteínas de ~ 0,05 mg/ml se pusieron en una celda de cuarzo de 0.1 cm de longitud. Cada espectro correspondió al promedio de tres escaneos. Las elipticidades se expresan como elipticidad residual media, $[\theta]_{m_{rw}}$ y el contenido de estructura secundaria se calculó a partir de espectros de DC utilizando el programa de deconvolución K2D3 (Louis-Jeune *et al.*, 2012).

6.2.22. Análisis estadísticos

Para determinar si los parámetros de cultivo calculados para cada condición eran o no significativamente diferentes, se llevó a cabo un análisis de varianza (ANOVA), seguido por la prueba de Tukey. Los resultados cuantitativos se expresan como media ± error estándar de la media. Un valor $P < 0.05$ fue considerado estadísticamente significativo.

7. RESULTADOS

7.1. Cinética de crecimiento de *E. coli* recombinante en matraces agitados

Se caracterizó el crecimiento de la cepa recombinante *E. coli* W3110 productora de la proteína rHuGM-CSF en matraces Erlenmeyer de 250 ml con 50 ml de medio definido. Los cultivos se mantuvieron a 30 °C hasta alcanzar la fase pre-estacionaria (DO de ~0.9-1.2 u.a.; **Figura 9A**). En este punto, se aumentó la temperatura hasta 42 °C para inducir la expresión de la proteína recombinante, con una velocidad de calentamiento de 0.26 °C/min. Como control, se creció la cepa recombinante en matraces sin inducción (30 °C).

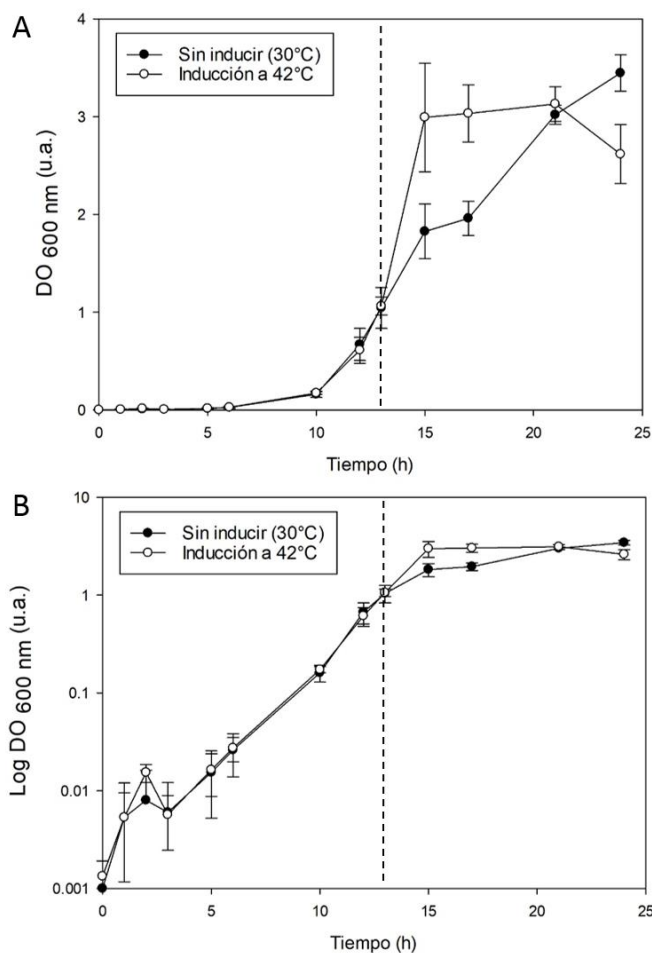


Figura 9. Cinética de crecimiento de la cepa *E. coli* W3110 + rHuGM-CSF en matraces agitados de 250 ml con 50 ml de medio definido. Se presenta el cambio de la densidad óptica (DO) a 600 nm en escala exponencial (**A**) y en escala logarítmica (**B**) con respecto al tiempo. Cultivos sin inducir a 30 °C (●) y cultivos inducidos a 42 °C (○). La línea punteada indica el momento de la termoinducción, a una DO_{600} de ~0.9-1.2 u.a. (13 h). Las barras de error representan la desviación estándar de un triplicado.

El aumento de la temperatura fue llevado a cabo en la fase pre-estacionaria de crecimiento, con el objetivo de tener gran cantidad de células viables para la producción de la proteína. La inducción a una biomasa alta permite la separación de las fases de crecimiento y de producción, para favorecer que los recursos metabólicos se canalicen hacia la acumulación de la proteína recombinante (Overton, 2014). Además, en este punto del crecimiento no se han desencadenado las respuestas de estrés asociadas a la fase estacionaria, por escasez de nutrientes o presencia de metabolitos secundarios (Craig y Gross, 1991; Hengge-Aronis, 1993).

En la **Figura 9A** se muestra el comportamiento de la DO de la cepa recombinante en matraces agitados con respecto al tiempo. La medición en cada punto se hizo por triplicado y se muestran sus respectivas desviaciones estándar. A las 21 h de la cinética, se observó que tanto los cultivos sin inducir como los termoinducidos a 42 °C alcanzaron una DO similar (3.02 ± 0.09 u.a. y 3.13 ± 0.17 u.a., respectivamente). Sin embargo, los cultivos termoinducidos presentaron al final de la cinética una caída en la DO del ~16 % (2.62 ± 0.30 u.a.) comparados con los cultivos no inducidos que continuaron creciendo (3.45 ± 0.19 u.a.).

Las velocidades específicas de crecimiento (μ) fueron similares para ambos cultivos, alcanzando valores de 0.49 ± 0.09 h⁻¹ en los matraces sin inducir y 0.46 ± 0.05 h⁻¹ para los matraces termoinducidos. Este resultado era de esperarse, pues la μ se calculó antes de la termoinducción, tomando los primeros puntos de la cinética donde ambos cultivos se encontraban a 30 °C (**Figura 9B**).

Como se mencionó en la metodología, se hizo una curva de calibración que correlaciona la medición de la DO y la biomasa en peso seco, resultando que 1 unidad de DO_(600 nm) fue equivalente a 0.32 ± 0.04 g/L de peso seco celular para *E. coli* W3110 productora de rHuGM-CSF. Dicha correlación fue lineal durante toda la cinética de crecimiento, obteniendo un R cuadrado de 0.9982. Es importante mencionar que esta correlación se hizo únicamente con la cepa recombinante creciendo a 30 °C, no para la cepa creciendo a 34 °C o bajo inducción a 42 °C. En otros trabajos donde se realizaron cultivos con diferentes cepas de *E. coli* recombinante, se alcanzaron valores semejantes: para *E. coli* W3110 recombinante, 1 DO_(600 nm) correspondió a 0.32 g/L (Sandoval-Basurto *et al.*, 2004); para *E. coli* BL21 (DE3) recombinante, 1 DO_(600 nm) correspondió a 0.39 g/L (Glazyrina *et al.*, 2010) y para una cepa derivada de *E. coli* K-12 en un sistema termoinducido, 1 DO_(600 nm) fue equivalente a 0.28 g/L (Mansey *et al.*, 2014).

De esta manera, con base en la correlación encontrada, se tiene que en matraces sin inducir (30 °C) se alcanzó una biomasa máxima ($X_{m\acute{a}x}$) de 1.10 ± 0.04 g/l a las 24 h de cultivo, mientras que para los matraces termoinducidos a 42 °C, se llegó a una $X_{m\acute{a}x}$ de 0.84 ± 0.07 g/l a las 21 h (**Figura 10**). La misma correlación se empleó para obtener la biomasa de los cultivos en biorreactores.

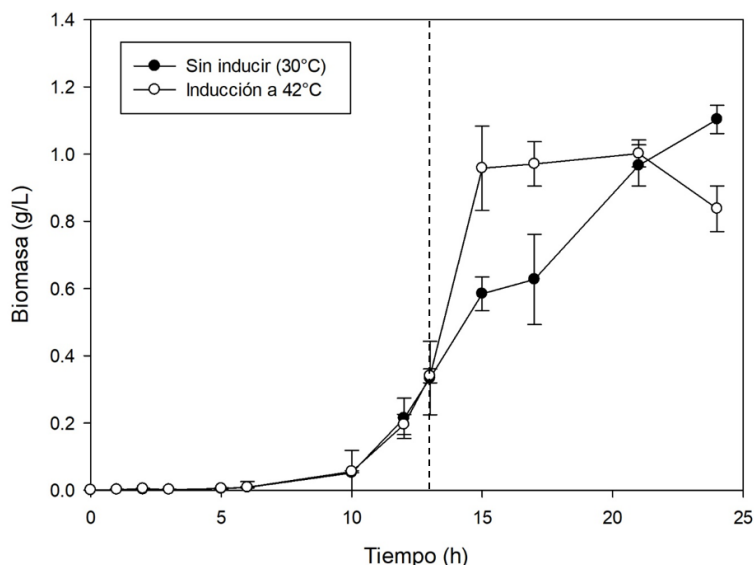


Figura 10. Cinética de crecimiento de la biomasa de *E. coli* W3110 + rHuGM-CSF en matraces agitados de 250 ml con 50 ml de medio definido. Se presenta el cambio de la biomasa en términos del peso seco celular para el cultivo sin inducir a 30 °C (●) y cultivo inducido a 42 °C (○). La línea punteada indica el momento de la termoinducción. Las barras de error representan la desviación estándar de un triplicado.

7.2. Expresión de la proteína rHuGM-CSF en matraces agitados

Se determinó la expresión de la proteína rHuGM-CSF en la fracción soluble (S), insoluble (I) y cuerpos de inclusión (CI) de los cultivos termoinducidos a 42 °C. Como control negativo, se evaluó si había expresión en los cultivos sin inducir (mantenidos a 30 °C). La fracción insoluble (I) se refiere a las proteínas endógenas agregadas sin haber inducido la expresión de una proteína recombinante. Mientras que los cuerpos de inclusión (CI) son los agregados aislados y purificados, que se forman tras la inducción y que se espera contengan en su mayoría a la proteína heteróloga. Las muestras fueron recolectadas al final de cada cultivo (24 h) y las proteínas fueron separadas por electroforesis en geles de poliacrilamida al 15 % (**Figura 11**).

En el carril de los CI obtenidos tras la inducción a 42 °C en matraces, se observa una banda mayoritaria de un peso molecular entre los 14-15 kDa (Carril 2, **Figura 11**), que presumiblemente corresponde a la proteína rHuGM-CSF. Esta banda no está presente en la fracción soluble (Carril 3, **Figura 11**), lo que demuestra que la proteína recombinante se está agregando durante la termoinducción, acumulándose exclusivamente en los CI.

Por otro lado, en las fracciones I y S de los cultivos no inducidos (Carriles 4 y 5 respectivamente, **Figura 11**) no se observó la banda correspondiente a la proteína recombinante. Esto demuestra que el sistema está finamente regulado y que el represor

cl857 es suficiente para mantener bloqueada la expresión de la proteína a temperaturas por debajo de 37 °C.

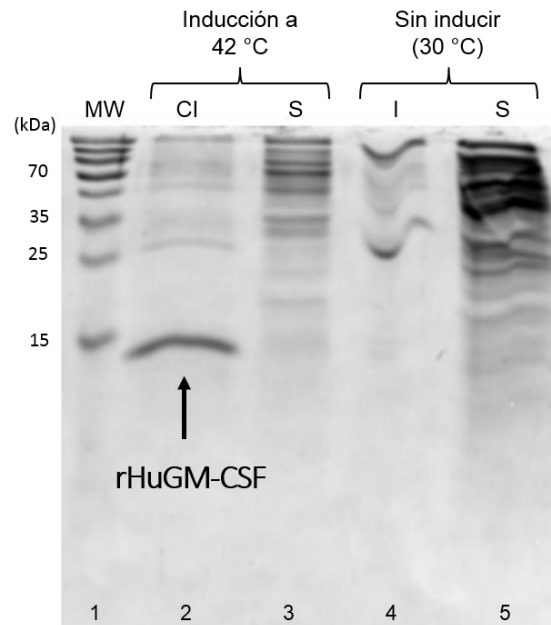


Figura 11. Expresión de la proteína rHuGM-CSF (~14-15 kDa) en *E. coli* W3110 usando matraces agitados a las 24 h de cultivo. Las proteínas se separaron en un gel de poliacrilamida al 15 % teñido con Azul de Coomassie. Carril 1: Marcador de peso molecular (MW); carril 2: Cuerpos de inclusión (CI) del cultivo termoinducido a 42 °C; carril 3: Proteína total soluble (S) del cultivo termoinducido a 42 °C; carril 4: Proteína insoluble (I) del cultivo sin inducir y carril 5: Proteína soluble (S) del cultivo sin inducir.

7.3. Cinética de crecimiento de *E. coli* recombinante en biorreactores

Previo a realizar la termoinducción en biorreactores de 1.2 L, se evaluó el cambio en la DO de la cepa recombinante en cultivos sin inducir, es decir, manteniéndolos a 30 °C o a 34 °C, para luego comparar y evaluar el impacto del aumento de la temperatura en el crecimiento (**Figura 12**). Los parámetros de oxígeno disuelto, pH y temperatura se mantuvieron controlados durante todo el cultivo.

En la cinética de crecimiento a 30 °C sin inducción, se alcanzó una DO máxima de 14.49 ± 0.01 u.a. (**Figura 12A**) y una velocidad específica de crecimiento (μ) de 0.56 ± 0.02 h⁻¹. Mientras que en los cultivos a 34 °C sin inducción, la DO máxima fue de 17.42 ± 1.18 u.a. (**Figura 12A**) y la μ de 0.81 ± 0.03 h⁻¹, esta última aumentando en un ~45% con respecto al crecimiento a 30 °C.

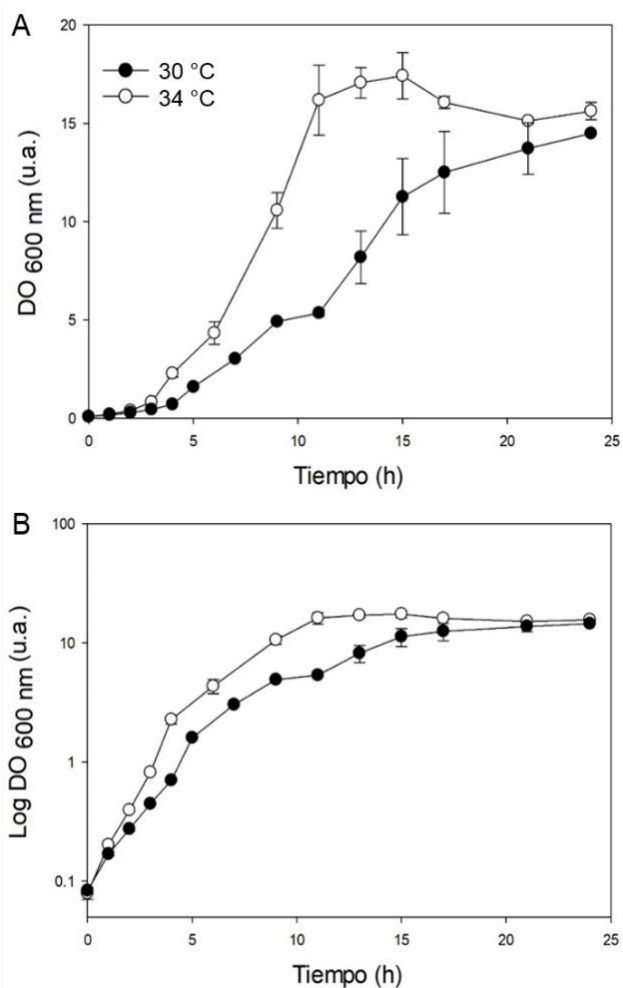


Figura 12. Cinética de crecimiento de la cepa *E. coli* W3110 + rHuGM-CSF sin termoinducción en biorreactores de 1.2 L con 800 ml de medio definido. Se presenta el cambio en la DO a 600 nm tanto en escala lineal (**A**) como en escala logarítmica (**B**). Los cultivos se mantuvieron a 30 °C (círculos negros) o a 34 °C (círculos blancos) a lo largo de todo cultivo. Las barras de error representan la desviación estándar de un triplicado.

Posteriormente, se realizó la caracterización del crecimiento de *E. coli* W3110 productora de rHuGM-CSF en biorreactores probando las dos temperaturas de crecimiento propuestas: 30 °C vs 34 °C con posterior inducción a 42 °C (**Figura 13**). Se tomaron muestras cada hora para la medición del peso seco celular, glucosa, acetato y proteína. Para todos los casos, la termoinducción se llevó a cabo durante la fase pre-estacionaria de crecimiento, cuando las células alcanzaron una DO de ~2.0-3.0 u.a.

En la **Figura 13** se comparan la acumulación de biomasa, el consumo de la fuente de carbono y la producción de acetato bajo termoinducción. En primer lugar, se observa que los cultivos a una temperatura de pre-inducción de 30 °C con posterior inducción a 42 °C, alcanzaron una biomasa máxima ($X_{m\acute{a}x}$) de 3.48 ± 0.19 g/l (21 h; **Figura 13A**), mientras que

para la condición de crecimiento a 34 °C con inducción a 42 °C, $X_{m\acute{a}x}$ fue de 3.48 ± 0.31 g/l (15 h; **Figura 13B**). A pesar de que las biomazas máximas alcanzadas en ambas condiciones de cultivo fueron similares, las velocidades específicas de crecimiento (μ) antes de la termoinducción presentaron diferencias significativas ($P > 0.05$). Para el caso de 30 \rightarrow 42 °C, se calculó una μ de 0.53 ± 0.01 h $^{-1}$ y para los cultivos de 34 \rightarrow 42 °C, la μ fue de 0.90 ± 0.07 h $^{-1}$, indicando que hubo un aumento del ~69 % cuando es mayor la temperatura de pre-inducción (**Tabla 5**).

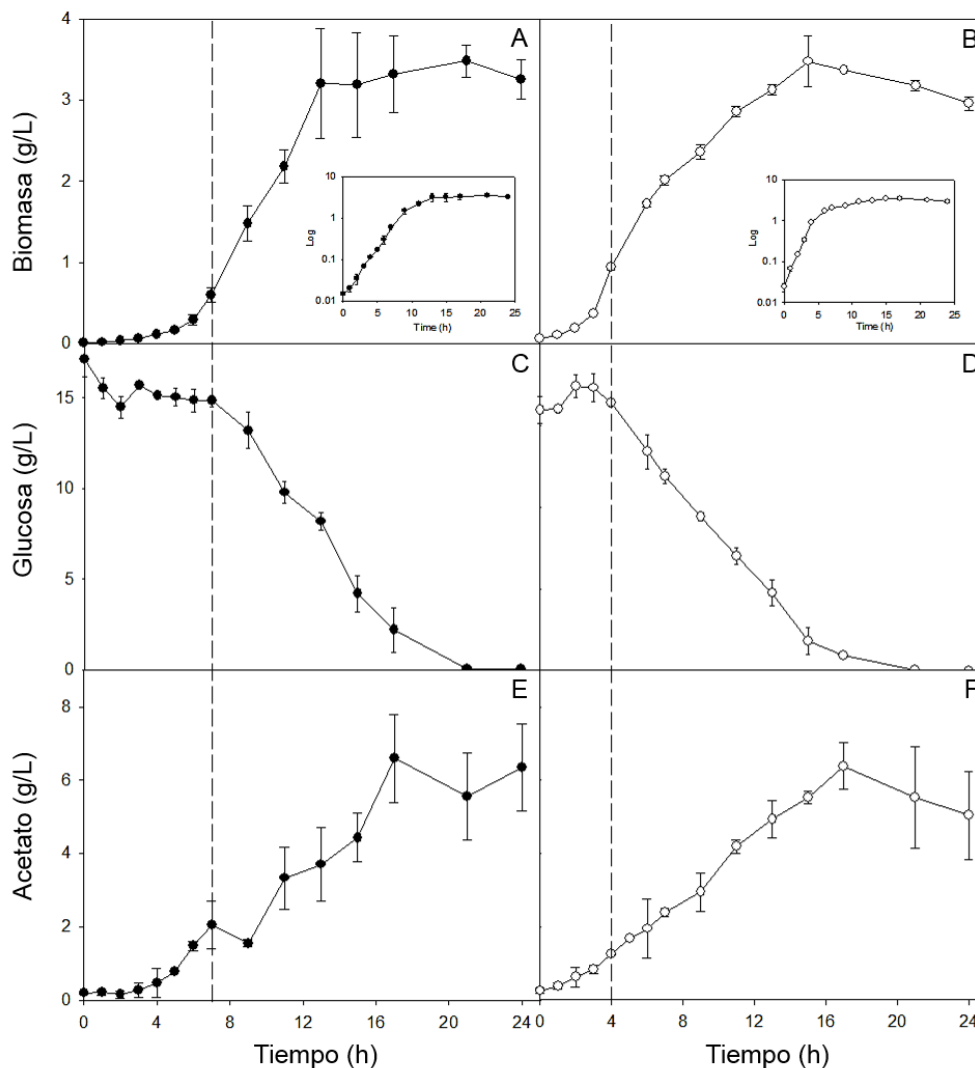


Figura 13. Cinética del crecimiento bacteriano (**A, B**), consumo de glucosa (**C,D**), y la producción de acetato (**E, F**) de *E. coli* W3110 + rHuGM-CSF en biorreactores de 1.2 L termoinducidos. Dos temperaturas de pre-inducción fueron evaluadas: 30 °C (círculos llenos) y 34 °C (círculos abiertos) con posterior termoinducción a 42 °C. Las líneas punteadas indican el inicio del aumento de temperatura (7 h para 30 °C y 4 h para 34 °C), que corresponden a una DO cercana a 2.0 u.a. Se muestra el promedio y la desviación estándar de tres experimentos independientes.

Como se observa en las **Figuras 13C y 13D**, la glucosa se consumió completamente alcanzando valores cercanos a cero después de las 18 horas del cultivo. De igual manera, el acetato se acumuló en el medio, con concentraciones máximas de 6.59 ± 0.69 g/L (**Figura 13E**) y 6.38 ± 0.63 g/L (**Figura 13F**) para los cultivos con temperaturas de pre-inducción de 30 °C y 34 °C, respectivamente. De manera interesante, el rendimiento de biomasa por glucosa ($Y_{X/S}$) y el de acetato por biomasa ($Y_{AC/X}$) no presentaron diferencias significativas entre los cultivos bajo ambas temperaturas de pre-inducción (30 °C vs 34 °C, **Tabla 5**). No obstante, cultivos creciendo a 34 °C consumieron la fuente de carbono más rápido, lo que se vio reflejado en una velocidad específica de consumo de glucosa (q_s) 1.5 veces mayor (3.74 ± 0.34 g/g·h) comparada con los cultivos a una temperatura de pre-inducción de 30 °C (q_s de 2.64 ± 0.20 g/g·h; **Tabla 5**). Estos resultados indican que a pesar de que las células a 34 °C consumieron más rápido la glucosa, no implica que necesariamente la utilicen para la formación de acetato, sino que puede derivarse a la producción de proteína recombinante o proteínas de choque térmico.

Por otro lado, la velocidad específica de producción de acetato (q_p) fue ~65 % más alta en cultivos creciendo a 34 °C, en comparación con aquellos que crecieron a 30 °C (**Tabla 5**). Si bien es cierto que las concentraciones máximas de acetato y biomasa fueron similares para ambas condiciones, es válido mencionar que la q_p se calculó con la μ antes de la inducción, que fue significativamente mayor a 34 °C en comparación con 30 °C (**Tabla 5**), por eso los valores de q_p variaron aunque los rendimientos $Y_{AC/X}$ hayan sido similares.

Tabla 5. Comparación de los parámetros cinéticos de *E. coli* W3110 + rHuGM-CSF en biorreactores a 30 °C y 34 °C con posterior termoinducción a 42 °C.

Parámetros ^{1,2}	Crecimiento a 30 °C→42 °C		Crecimiento a 34 °C→42 °C	
	Antes de inducir	Después de inducir ³	Antes de inducir	Después de inducir ³
μ (h ⁻¹)	0.53 ± 0.01 ^a	0.14 ± 0.03 ^b	0.90 ± 0.07 ^c	0.11 ± 0.01 ^b
td (h)	1.30 ± 0.03 ^a		0.78 ± 0.08 ^b	
$X_{\text{máx}}$ (g _{DCW} /L)	3.48 ± 0.19 ^a		3.48 ± 0.31 ^a	
$Y_{X/S}$ (g _{DCW} /g _{GLC}) ⁴	0.20 ± 0.02 ^a		0.25 ± 0.03 ^a	
q_s (g _{GLC} /g _{DCW} ·h) ⁵	2.64 ± 0.20 ^a		3.74 ± 0.34 ^b	
$Y_{AC/X}$ (g _{AC} /g _{DCW}) ⁴	1.61 ± 0.76 ^a		1.56 ± 0.20 ^a	
q_p (g _{AC} /g _{DCW} ·h) ⁵	0.86 ± 0.41 ^a		1.41 ± 0.20 ^b	

¹ μ : velocidad específica de crecimiento; td: tiempo de duplicación; $X_{\text{máx}}$: concentración de biomasa celular máxima; CDW: Cell Dry Weight (biomasa en peso seco); GLC: Glucosa; AC: Acetato; $Y_{X/S}$: Rendimiento de biomasa por sustrato; $Y_{AC/X}$: Rendimiento de acetato por biomasa; q_s : Velocidad específica de consumo de sustrato; q_p : Velocidad específica de formación de acetato.

²Se presenta el promedio y la desviación estándar para tres experimentos independientes. Un resultado no estadísticamente significativo ($P > 0.05$) se representa con la misma letra, y un resultado estadísticamente significativo ($P < 0.05$) se representa con una letra diferente.

³ μ después de la inducción se calculó a partir del punto de la inducción hasta el comienzo de la fase estacionaria.

⁴ $Y_{X/S}$ y $Y_{AC/X}$ se calcularon utilizando las concentraciones de glucosa y acetato en la $X_{m\acute{a}x}$.

⁵ q_s y q_p se calcularon utilizando la μ antes de la inducci3n y los rendimientos obtenidos en la $X_{m\acute{a}x}$.

La tensi3n de ox3geno disuelto en el medio (TOD) fue controlada a trav3s de un sistema de control PI (proporcional integral), ya que el derivativo se mantuvo en cero. Se observa en las **Figuras 14A y 14B** que los valores de la TOD siempre estuvieron oscilando alrededor del "set point" de 35 %, manteniendo el ox3geno en un rango entre 60-20 % de saturaci3n de aire. Aunque el control no fue estricto, se confirma que no hubo limitaci3n de ox3geno en los cultivos en biorreactor. Para mejorar el control, se podr3an ajustar los par3metros del sistema a los valores 3ptimos, disminuyendo el valor del proporcional y a3nadiendo una acci3n derivativa, para hacerlo PID. En el caso del pH del medio, se mantuvo controlado cercano a 7.0 mediante un sistema *on/off* de adici3n autom3tica de NaOH 3 N o HCl 3 N (**Figuras 14C y 14D**).

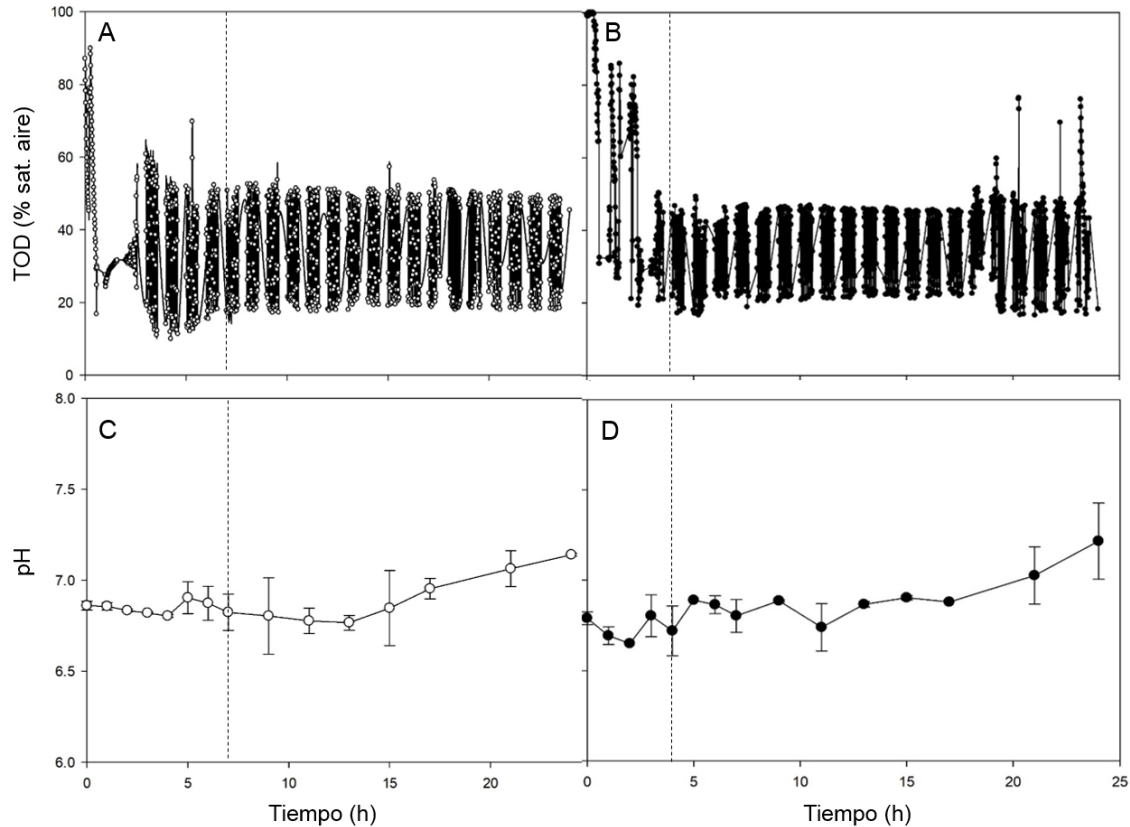


Figura 14. Comportamiento de la tensi3n de ox3geno disuelto (**A, B**) y pH (**C, D**) de cultivos de *E. coli* W3110 + rHuGM-CSF en biorreactores de 1.2 L. Dos temperaturas de pre-inducci3n fueron evaluadas: 30 °C (panel izquierda) y 34 °C (panel derecha) con posterior termoinducci3n a 42 °C. Las l3neas punteadas indican el inicio del aumento de temperatura (7 h para 30 °C y 4 h para 34 °C), que corresponden a una DO de ~2.0 u.a. Se muestra el promedio y la desviaci3n est3ndar de tres experimentos independientes.

7.4. Cinética de crecimiento de la cepa silvestre *E. coli* W3110

Se caracterizó el crecimiento de la cepa silvestre *E. coli* W3110 (“*Wild type*”) en biorreactores de 1.2 L con 800 ml de medio de cultivo definido (**Figura 15**).

Los cultivos se llevaron a cabo a 30 °C, manteniendo además el control de los parámetros de pH en 7.0 y la TOD en 35 %. El objetivo principal de realizar estos cultivos fue tomar muestras para la recuperación de la biomasa y así obtener los lisados celulares. Dichos lisados celulares se usaron como control negativo de la expresión de la proteína recombinante en geles SDS-PAGE, junto con los lisados de la cepa recombinante sin inducir (a 30 °C o 34 °C).

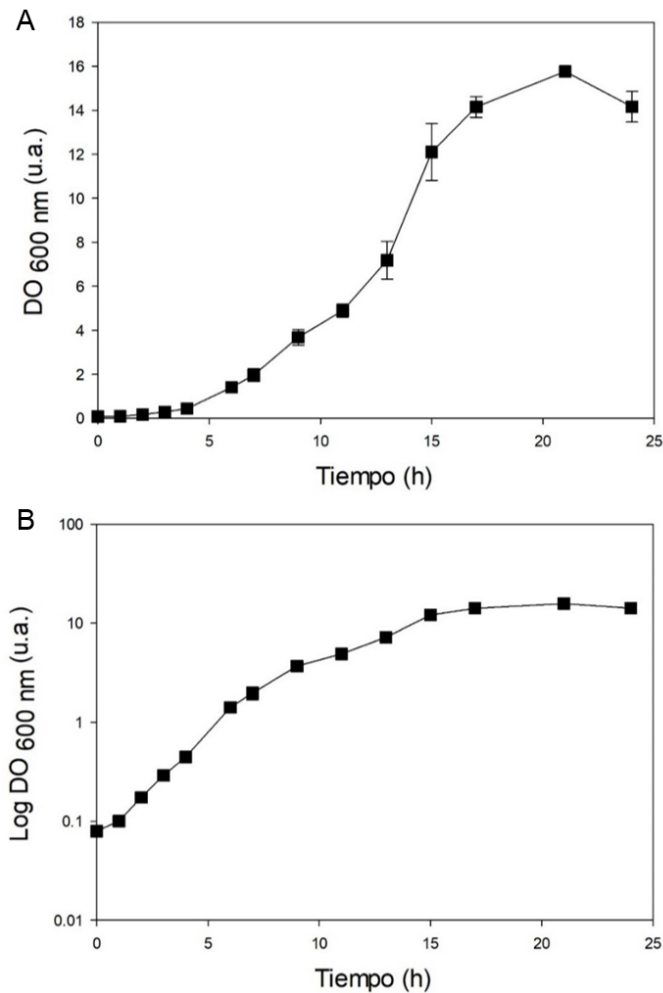


Figura 15. Cinética de crecimiento de la cepa silvestre *E. coli* W3110 en biorreactores de 1.2 L con 800 ml de medio definido a 30 °C. Se presenta el cambio en la DO a 600 nm tanto en escala lineal (**A**) como en escala logarítmica (**B**), representando el promedio de tres réplicas biológicas con sus respectivas desviaciones estándar.

La DO máxima fue de 15.16 ± 0.08 u.a. a las 21 h de cultivo (**Figura 15**), levemente mayor a la obtenida en los cultivos en biorreactores a 30 °C (no inducidos) de la cepa recombinante (14.49 ± 0.01 u.a, **Figura 12**). Además, la velocidad específica de crecimiento de la cepa de *E. coli* W3110 silvestre fue $\mu=0.48 \pm 0.02$ h⁻¹.

7.5. Expresión de la proteína rHuGM-CSF en biorreactores

7.5.1. Comparación de la expresión en las fracciones de proteína total

Para determinar la expresión de la proteína recombinante, se hizo en primer lugar una electroforesis en geles de poliacrilamida al 15 %, usando las fracciones de proteína total. Para esto se cargaron los lisados celulares totales de la cepa recombinante *E. coli* W3110 productora de rHuGM-CSF, creciendo a dos temperaturas de pre-inducción diferentes, 30 °C o 34 °C, con posterior inducción a 42 °C (**Figura 16**). Como controles negativos se emplearon: la muestra justo antes del aumento de la temperatura (0 h), los lisados celulares de la cepa *E. coli* W3110 recombinante sin termoinducir, creciendo a 30 °C o 34 °C y, los lisados de la cepa *E. coli* W3110 silvestre (WT) que creció a 30 °C.

Este fue uno de los primeros resultados que nos confirmó la producción de la proteína de interés tras la termoinducción a 42 °C en biorreactores, tanto en cultivos creciendo a 30 °C (**Figura 16A**) como a 34 °C (**Figura 16B**). En los cultivos que crecieron a 34 °C (**Figura 16B**), se observa una banda intensa a la altura esperada de la proteína rHuGM-CSF (entre ~14-15 kDa) justo 1 h después de la termoinducción a 42 °C, mientras que a 30 °C (**Figura 16A**), esta misma banda comienza a ser visible a las 3 h de termoinducción.

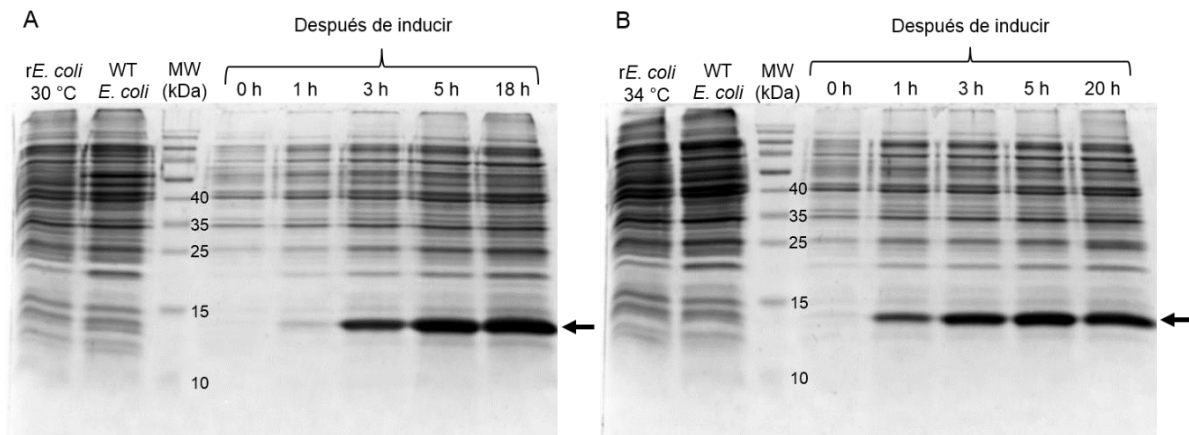


Figura 16. Lisados celulares totales de *E. coli* W3110 + rHuGM-CSF obtenidos en biorreactores a 30 °C (**A**) y 34 °C (**B**) a diferentes tiempos después de la inducción a 42 °C. Las proteínas separaron en un gel de poliacrilamida al 15 % teñido con azul de Coomassie, que fue cargado volumétricamente. Las flechas señalan una banda entre ~14-15 kDa, que corresponde presumiblemente a rHuGM-CSF.

Como era de esperarse, no se observó la banda de la proteína recombinante en la hora cero justo antes de la termoinducción, confirmando la estricta regulación del sistema por el represor cI857 antes del cambio de la temperatura. Lo cual es acorde con los resultados observados en la evaluación de la expresión de la proteína en matraces agitados (**Figura 11**).

El análisis densitométrico del gel correspondiente a la proteína celular total (**Tabla 6**) indica que 1 h después de inducir a 30 °C, la banda de la proteína rHuGM-CSF representa un ~11 % del carril, mientras que a 34 °C representa el ~22 %. Conforme pasó el tiempo de la termoinducción, fue aumentando el tamaño de la banda, indicando la acumulación de la proteína de interés rHuGM-CSF (**Figura 16**). Al final del cultivo, la proteína recombinante fue ~34 % del carril a una temperatura de pre-inducción de 30 °C y ~30 % para 34 °C. Esto se traduce en una producción de 0.34 ± 0.10 mg/ml de rHuGM-CSF a 30 °C y 0.26 ± 0.04 mg/ml de rHuGM-CSF a 34 °C. Nuestros resultados corresponden con lo reportado por trabajos anteriores, donde se indica que usando un sistema de expresión regulado por temperatura, es posible alcanzar altos rendimientos cercanos al 30 % de la proteína recombinante con respecto a la proteína celular total (Remaut *et al.*, 1981; Makrides, 1996; Valdez-Cruz *et al.*, 2010; Restrepo-Pineda *et al.*, 2021).

7.5.2. Comparación de la expresión de la proteína en cuerpos de inclusión

Después de analizar la producción de proteína total en los lisados celulares, el objetivo fue contrastar la expresión en las fracciones soluble e insoluble, para identificar en cuál de ellas se estaba acumulando la proteína rHuGM-CSF. Se obtuvieron por separado las fracciones de proteína soluble (S) y proteína insoluble. Esta última fracción fue procesada para la recuperación y purificación de los cuerpos de inclusión (CI). Se compararon las muestras obtenidas a diferentes tiempos después de la termoinducción (1 h, 5 h y tiempo final) en un gel de poliacrilamida al 15 % (**Figura 17**). En este caso se usaron dos controles negativos, los lisados celulares de la cepa *E. coli* W3110 recombinante sin termoinducir, creciendo a 30 °C o 34 °C y los lisados de la cepa *E. coli* W3110 silvestre (WT) cultivada a 30 °C.

En la **Figura 17**, se observa una banda cercana a los 14-15 kDa, únicamente en los carriles de los CI después de la termoinducción, tanto a 30 °C como a 34 °C. Esta misma banda que puede corresponder al factor rHuGM-CSF, no aparece en ningún carril de las fracciones solubles, lo que significa que la proteína recombinante inevitablemente forma cuerpos de inclusión después de la termoinducción. Este mismo resultado se obtuvo para los matraces agitados (**Figura 11**), es decir, que independiente de la escala de cultivo, la proteína tiende a agregarse.

Para ambas condiciones evaluadas, la expresión de la proteína rHuGM-CSF fue mayor a las 5 h después del aumento de la temperatura a 42 °C. En congruencia con los resultados de la proteína total, la expresión de la proteína recombinante fue levemente más intensa 1 h después de inducir bajo la temperatura de pre-inducción de 34 °C (**Figura 17B**) en comparación con 30 °C (**Figura 17A**). La banda de interés no se presentó en las muestras

de los lisados provenientes del cultivo bacteriano sin termoinducir (*rE. coli* a 30 °C o 34 °C) o en el cultivo de la cepa sin transformar (WT *E. coli*) para ambos casos.

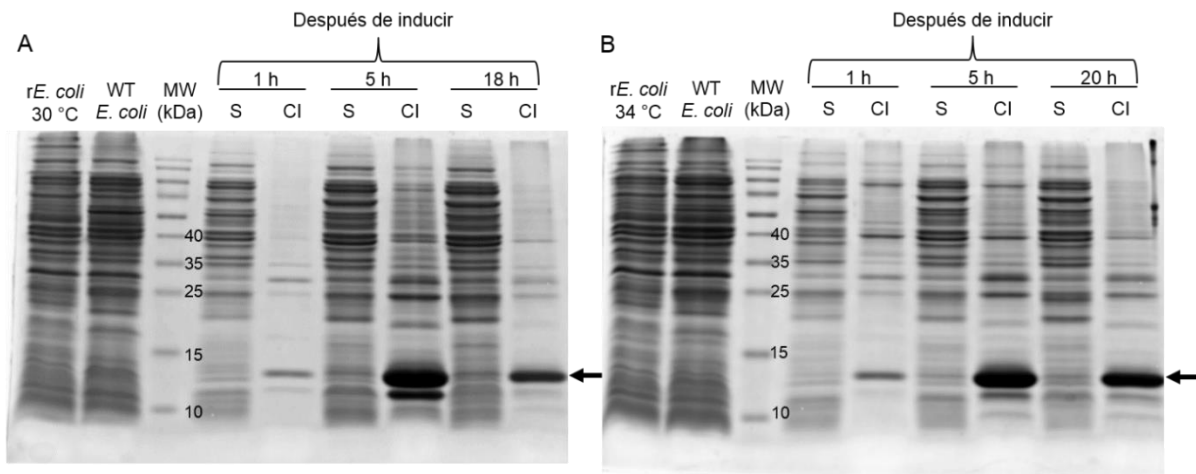


Figura 17. Expresión de la proteína rHuGM-CSF (~14-15 kDa) en la proteína soluble y los cuerpos de inclusión de *E. coli* W3110 recombinante usando biorreactores. Dos temperaturas de pre-inducción diferentes fueron evaluadas: 30 °C (**A**) y 34 °C (**B**) en un gel de poliacrilamida al 15 % teñido con azul de Coomassie. Las muestras corresponden a la proteína soluble (S) y los cuerpos de inclusión (CI) a diferentes tiempos después de la termoinducción (1 h, 5 h y 18/20 h). Se usaron como controles negativos la cepa sin inducir (*rE. coli* a 30 °C o 34 °C) y la cepa silvestre (WT *E. coli*). El gel se cargó volumétricamente. Las flechas señalan una banda entre ~14-15 kDa, que corresponde presumiblemente a rHuGM-CSF.

Posteriormente, para deducir los rendimientos de producción de rHuGM-CSF en los CI, se realizaron geles cargados a la misma cantidad de proteína (cuantificada por el método de Bradford). Se emplearon 20 µg de los CI obtenidos cinéticamente después de la inducción a 42 °C (1 h, 3 h, 5 h 10 h y 18/20 h), provenientes de los cultivos que crecieron tanto a 30 °C como a 34 °C (**Figura 18**). Para los controles negativos nuevamente se utilizaron los lisados provenientes del cultivo bacteriano sin termoinducir (*rE. coli* a 30 °C o 34 °C) y la cepa sin transformar (WT *E. coli*). El estándar de referencia de la farmacopea europea para el GM-CSF humano o Molgramostim (catálogo N°Y0000251), se utilizó como control positivo. Dicho estándar es una preparación pura de la proteína, que no contiene albúmina y es producido de forma recombinante en *E. coli*, usando un inserto que contiene la secuencia del GM-CSF humano. Se cargó 1 µg de dicho estándar en cada gel.

En la **Figura 18**, se puede ver que la banda señalada cerca a los ~14-15 kDa coincide con la banda que representa al estándar de referencia para Molgramostim (forma no glicosilada de GM-CSF producida en *E. coli*), lo que comprueba electroforéticamente que esta banda es la proteína recombinante de interés: rHuGM-CSF. También hay una banda inferior que está presente cuando se visualizan los CI, la cual podría corresponder, con base a su peso molecular, a las proteínas de choque térmico pequeñas IbpA e IbpB. Dichas proteínas suelen agregarse en los cuerpos de inclusión junto con la proteína recombinante. Además, se ha reportado ampliamente que el complejo IbpA-IbpB se asocia con proteínas agregadas

para estabilizarlas y protegerlas de la desnaturalización irreversible y la proteólisis extensa durante el choque térmico (Allen *et al.*, 1992; Kuczyńska-Wiśnik *et al.*, 2002; Lethanh *et al.*, 2005). En nuestro caso, el complejo lbpA-lbpB puede resultar indispensable para contener el estrés generado debido a la termoinducción prolongada.

Adicionalmente, se observó la acumulación de rHuGM-CSF dentro de los CI conforme aumentó el tiempo de termoinducción, con algunas diferencias notables según la temperatura de pre-inducción. En los CI obtenidos de cultivos que crecieron a 30 °C (**Figura 18A**), la máxima acumulación de rHuGM-CSF en los CI ocurrió a las 5 h después de la inducción a 42 °C y luego disminuyó levemente con el tiempo de cultivo. Mientras que para los CI obtenidos de los cultivos que crecieron a 34 °C (**Figura 18B**), la cantidad de proteína recombinante fue incrementando hora tras hora hasta finalizar la cinética. Un gel de poliacrilamida al 15 % se realizó con las muestras correspondientes a las réplicas biológicas de los CI (**Anexo 2**), con el fin de calcular los rendimientos de producción de la proteína recombinante.

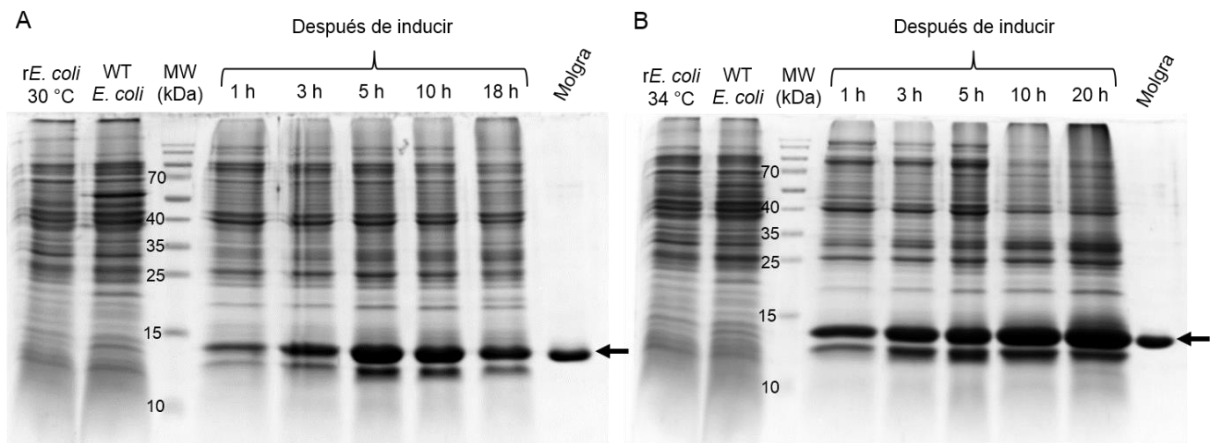


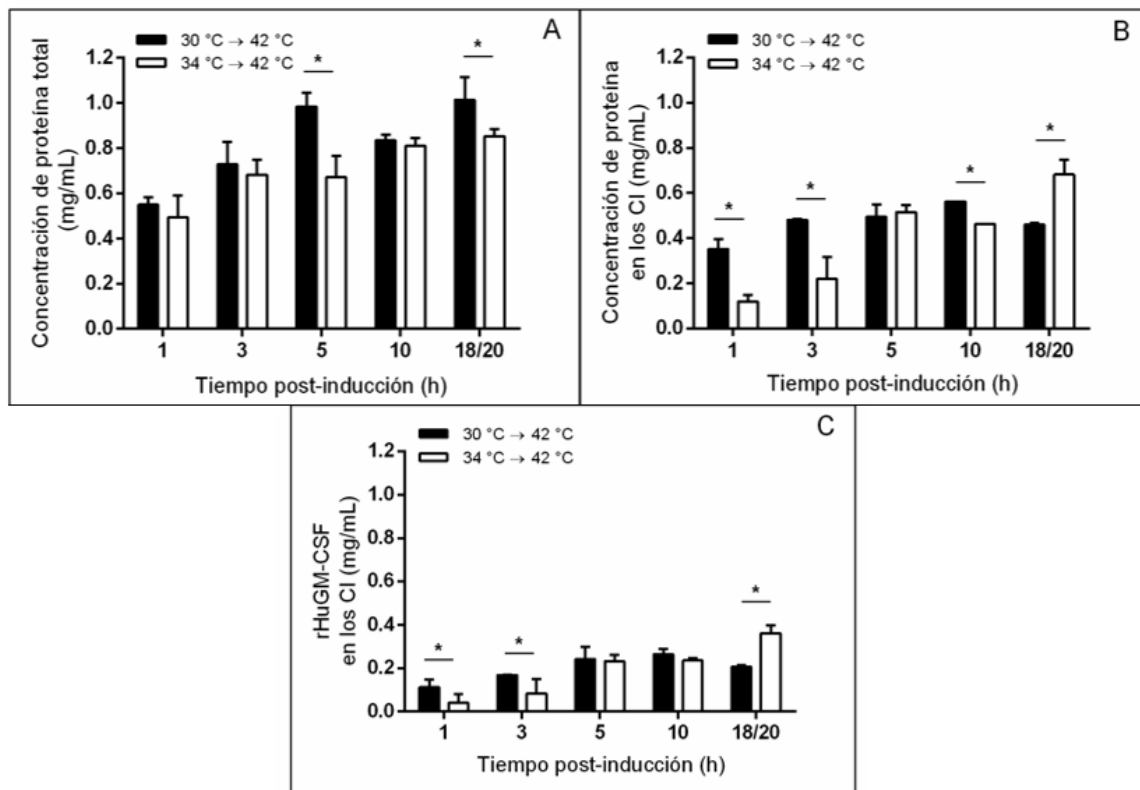
Figura 18. Expresión de la proteína rHuGM-CSF (~14-15 kDa) en los cuerpos de inclusión de *E. coli* W3110 recombinante usando biorreactores. Dos temperaturas de pre-inducción diferentes fueron evaluadas: 30 °C (**A**) y 34 °C (**B**) en un gel de poliacrilamida al 15 % teñido con azul de Coomassie. Se presentan los cuerpos de inclusión (CI) con 20 µg de proteína a diferentes tiempos después de la termoinducción (1 h, 3 h, 5 h 10 h y 18/20 h). Se usaron como controles negativos, la cepa sin inducir (*rE. coli* a 30 °C o 34 °C) y la cepa silvestre (WT *E. coli*). El control positivo es el estándar de referencia para Molgramostim (Molgra).

Para determinar los porcentajes correspondientes al factor rHuGM-CSF, tanto en la proteína total como en los CI, se llevaron a cabo análisis densitométricos en función del tiempo (**Tabla 6**) usando el programa *Image Lab* (Versión 5.2 de *BIO-RAD*). Se presentan los porcentajes de banda respecto a cada carril, que fueron obtenidos de la densitometría de los geles de proteína total y proteína en cuerpos de inclusión. Se comparan los contenidos de proteína proveniente de cultivos a dos temperaturas de pre-inducción diferentes: 30 °C y 34 °C.

Tabla 6. Análisis densitométrico de la proteína rHuGM-CSF en la proteína total y en los CI a diferentes tiempos después de la termoinducción.

Tiempo post-inducción (h)	rHuGM-CSF en la proteína total (%)		rHuGM-CSF en los CI (%)	
	30 °C→42 °C	34 °C→42 °C	30 °C→42 °C	34 °C→42 °C
1	11 ± 1	22 ± 3	32 ± 2	34 ± 2
3	28 ± 2	31 ± 3	35 ± 2	38 ± 5
5	35 ± 4	33 ± 2	49 ± 1	45 ± 2
10	-	-	47 ± 3	51 ± 2
18/20	34 ± 3	30 ± 4	45 ± 2	53 ± 3

La influencia de la temperatura de pre-inducción en las concentraciones de la proteína celular total, proteína agregada en los CI y rHuGM-CSF en los CI durante la termoinducción, se ilustra en las **Figuras 19A, 19B y 19C**, respectivamente.



Cuando las células crecen a una temperatura de pre-inducción de 30 °C, la concentración de la proteína total al final del cultivo termoinducido fue ~18 % mayor que la concentración de la proteína total obtenida a 34 °C (**Figura 19A**). Sin embargo, la cantidad de la proteína total agregada fue mayor en aquellos cultivos bajo una temperatura de pre-inducción de 34 °C (**Figura 19B**). De igual manera, la acumulación de la proteína recombinante rHuGM-CSF (**Figura 19C**), fue significativamente más alta a 34 °C (0.36 ± 0.05 mg/ml) que a 30 °C (0.21 ± 0.01 mg/ml), si se compara la última hora del cultivo.

En la **Tabla 7**, se indica la proporción de rHuGM-CSF en la proteína total y en los CI al final de cada cultivo, bajo ambas temperaturas de pre-inducción, así como los rendimientos de proteína total por biomasa ($Y_{PT/X}$) y el rendimiento de proteína recombinante por biomasa ($Y_{PR/X}$). Según los resultados de la densitometría, la proteína rHuGM-CSF representa el 45 ± 2 % y 53 ± 3 % de los CI obtenidos al final de cultivos termoinducidos que crecieron a 30 °C y 34 °C, respectivamente (**Tabla 6**). A pesar de que el rendimiento de la proteína total con respecto a la biomasa ($Y_{PT/X}$) fue similar bajo ambas condiciones evaluadas, el rendimiento de rHuGM-CSF por biomasa ($Y_{PR/X}$) fue dos veces superior para los cultivos a una temperatura de pre-inducción de 34 °C→42 °C (0.12 ± 0.02 g/g) comparados con los de 30 °C→42 °C (0.06 ± 0.03 g/g) (**Tabla 7**).

Tabla 7. Comparación de los rendimientos de la proteína rHuGM-CSF bajo diferentes temperaturas de pre-inducción, 30 °C y 34 °C, con posterior termoinducción a 42 °C.

Parámetros	Crecimiento a 30 °C→42 °C	Crecimiento a 34 °C→42 °C
Proteína total, PT (mg/ml)	1.01 ± 0.10^a	0.85 ± 0.03^b
rHuGM-CSF en PT (mg/ml)	0.34 ± 0.10^a	0.26 ± 0.04^a
$Y_{PT/X}$ (g_{TP}/g_{DCW})	0.29 ± 0.08^a	0.24 ± 0.05^a
rHuGM-CSF en CI (mg/ml)	0.21 ± 0.01^a	0.36 ± 0.05^b
$Y_{PR/X}$ (g_{RP}/g_{DCW})	0.06 ± 0.03^a	0.12 ± 0.02^b

PT: proteína total; CI: cuerpos de inclusión; PR: proteína recombinante; $Y_{PT/X}$: rendimiento de proteína total por biomasa; $Y_{PR/X}$: rendimiento de proteína recombinante, rHuGM-CSF, por biomasa. Se presenta el promedio y la desviación estándar para tres experimentos independientes. Un resultado no estadísticamente significativo ($P > 0.05$) se representa con la misma letra, y un resultado estadísticamente significativo ($P < 0.05$) se representa con una letra diferente.

$Y_{PT/X}$ y $Y_{PR/X}$ se calcularon usando los valores máximos de las concentraciones de proteína.

7.6. Identificación de chaperonas DnaK, DnaJ, GroEL y GroES

Interesados en analizar la expresión diferencial de las principales proteínas de choque térmico de *E. coli* bajo las dos temperaturas de pre-inducción, 30 °C y 34 °C, se realizaron inmunodetecciones de las chaperonas de plegamiento DnaK, DnaJ, GroEL y GroES en los CI obtenidos de cultivos termoinducidos en biorreactores (**Figura 20**).

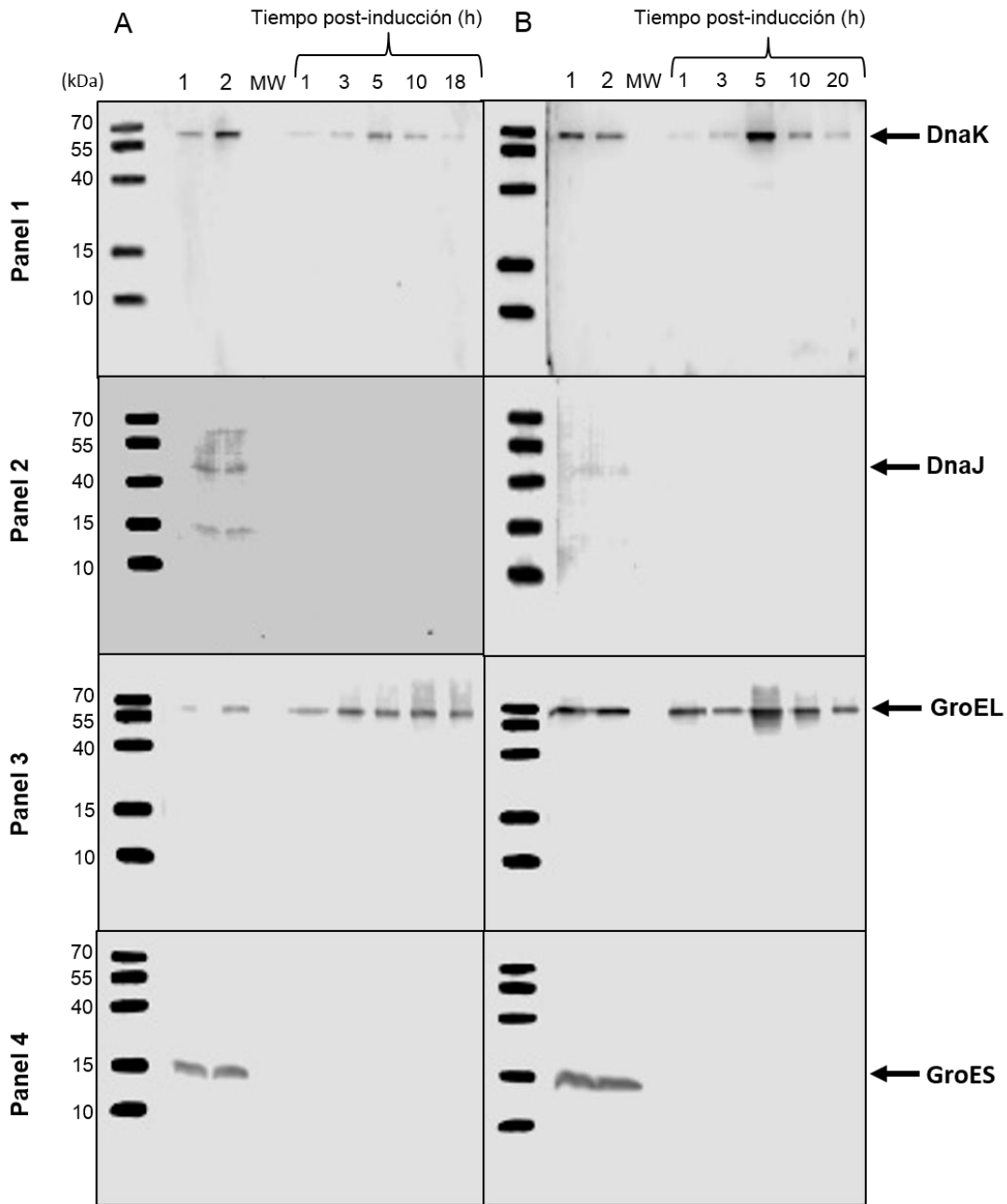


Figura 20. Inmunodetección de las chaperonas GroEL, GroES, DnaK y DnaJ en los CI de la proteína rHuGM-CSF provenientes de cultivos bajo diferentes temperaturas de crecimiento: 30 °C (**A**) y 34 °C (**B**), termoinducidos a 42 °C. Carriles 1A y 1B: lisados de la cepa recombinante sin inducir a 30 °C o 34 °C, respectivamente. Carriles 2A y 2B: lisados de la cepa silvestre. Carril MW: marcador de peso molecular. Se presentan los CI a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h) bajo las dos condiciones evaluadas. Las flechas indican las bandas correspondientes a DnaK (~70 kDa, panel 1), DnaJ (~41 kDa, panel 2), GroEL (~60 kDa, panel 3) y GroES (~15 kDa, panel 4).

El gel SDS-PAGE cargado con la misma cantidad de proteína (20 µg) en la **Figura 18**, se transfirió a una membrana de PVDF para llevar a cabo los Western blot correspondientes de cada chaperona. También se realizaron los Western blot usando los geles correspondientes a las réplicas biológicas (**Anexo 3**).

Todas las chaperonas se lograron identificar en los controles positivos, es decir, en los lisados de la cepa *E. coli* W3110 recombinante creciendo a 30 o 34 °C sin termoinducción y en los lisados de la cepa silvestre, como era lo esperado (Carriles 1 y 2, **Figura 20**). La banda de la chaperona DnaK fue observada cerca a los 70 kDa, lo cual corresponde con su peso molecular reportado (Bardwell y Craig, 1984). DnaK se encontró débilmente expresada en las primeras horas (1 y 3 h) después de la termoinducción a 42 °C, pero la intensidad de la banda aumentó significativamente a las 5 h y volvió a su nivel inicial, tanto en los cuerpos de inclusión provenientes de cultivos a 30 °C (**Figura 20A**, Panel 1) como a 34 °C (**Figura 20B**, Panel 1). Es notable que la intensidad de la banda de DnaK fue mayor en los CI después de las 5 h de termoinducción en los cultivos a 34 °C→42 °C, en comparación con los del 30 °C→42 °C. En el caso de GroEL, cuyo peso molecular es de aproximadamente 60 kDa (Hayer-Hartl *et al.*, 2015), se observó una banda en los CI obtenidos tanto a una temperatura de pre-inducción de 30 °C (**Figura 20A**, Panel 3) como a 34 °C (**Figura 20A**, Panel 3). Dicha banda fue de intensidad similar durante todo el tiempo después de la termoinducción a 42 °C.

Seguidamente, se realizó la detección de las co-chaperonas DnaJ y GroES, que se conocen por actuar como acompañantes de DnaK y GroEL en el plegamiento, respectivamente. De manera interesante, las bandas correspondientes a DnaJ con ~41 kDa (**Figura 20**, Panel 2) y GroES con ~15 kDa (**Figura 20**, Panel 4) únicamente en los controles positivos. Es decir, que DnaJ y GroES no están asociadas a los CI bajo ninguna de las dos temperaturas de pre-inducción probadas, sin embargo, podríamos suponer que pueden quedar en la fracción de proteína soluble después de la termoinducción a 42 °C.

Para corroborar esta suposición, incubamos las membranas obtenidas a partir de los geles de la proteína total soluble con los anticuerpos específicos para GroES y DnaJ. GroES fue identificada entre los 10-15 kDa, en la fracción de proteína total soluble de cultivos creciendo tanto a 30 °C como a 34 °C (**Figuras 21A y 21B**). Asimismo, se detectó una banda a una altura aproximada de 41 kDa, que corresponde a la proteína DnaJ en la proteína total soluble (**Figura 21C y 21D**). La presencia de estas bandas, como era de esperarse, no se restringió a los controles de la cepa sin inducir o a la cepa silvestre, sino que también se observan en las muestras de los cultivos termoinducidos a 42 °C. El comportamiento fue similar tanto para los cultivos a 30 °C como a 34 °C siendo un poco más intensas las bandas bajo la temperatura de pre-inducción de 34 °C→42 °C.

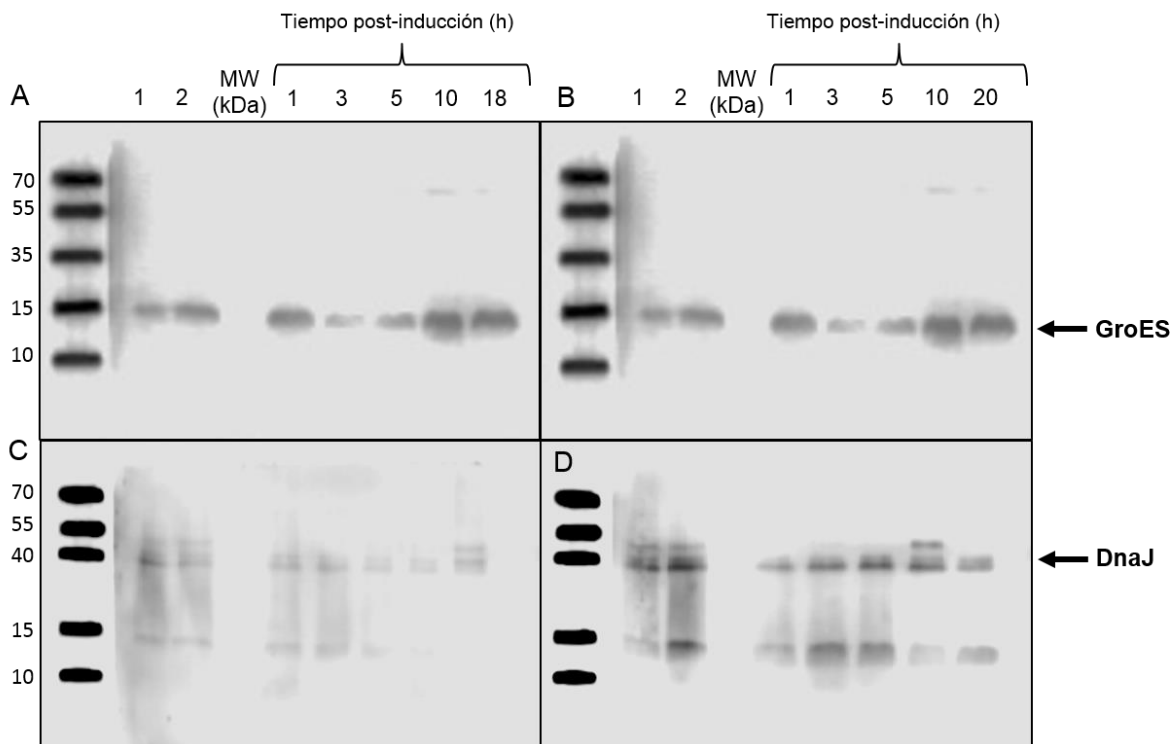


Figura 21. Inmunodetección de GroES y DnaJ en la proteína total soluble proveniente de cultivos bajo diferentes temperaturas de crecimiento: 30 °C (**A, C**) y 34 °C (**B, D**), termoinducidos a 42 °C. Carril 1: lisados de *E. coli* W3110 recombinante sin inducir a 30 °C o 34 °C, respectivamente; Carril 2: lisados de la cepa *E. coli* W3110 silvestre; Carril MW: marcador de peso molecular. Se muestra la proteína total soluble a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h) a las dos condiciones evaluadas. Las flechas indican las bandas correspondientes a GroES en ~15 kDa (**A, B**) y DnaJ en ~41 kDa (**C, D**).

7.7. Proteómica parcial de los CI de rHuGM-CSF usando iTRAQ

Un acercamiento sobre la composición proteica de los cuerpos de inclusión de rHuGM-CSF obtenidos en cultivos de *E. coli*, fue llevado a cabo usando una técnica de proteómica “shotgun” conocida como iTRAQ (etiquetas isobáricas para la cuantificación relativa y absoluta). Esta técnica implica una serie de procesos experimentales que van desde la digestión y el marcaje de péptidos, hasta el uso de cromatografía líquida (LC), espectrometría de masas en tándem (MS/MS) y algoritmos de búsqueda en bases de datos, para la identificación y cuantificación simultáneas de proteínas (Aggarwal *et al.*, 2006).

Para la identificación de las proteínas se trabajó con la plataforma Proteome Discoverer™ 2.1 SP1 (Thermo Fisher Scientific, USA), la cual requiere que el usuario ajuste los parámetros del flujo de trabajo para la búsqueda. En resumen, se comienza seleccionando la carpeta (archivos.raw) que contiene la información resultante del espectrómetro de

masas, así como la base de datos contra la cual se va a comparar. En este caso usamos la base de datos de *E. coli* de Uniprot en formato FASTA. A continuación, se indica la enzima utilizada para escindir las proteínas (tripsina) y se pueden especificar una serie de modificaciones estándar a los péptidos que resultan de la digestión trípica y el tratamiento de la muestra. Los protocolos generalmente aconsejan seleccionar la oxidación de la metionina porque el ambiente no es inerte, así como la acetilación y la carbamilación de la cisteína cuando se emplea un protocolo iTRAQ. Es importante indicarle al programa los denominadores que se van a usar, es decir, con qué reactivos o banderas se marcaron las muestras (114, 115, 116 y 117). Finalmente, se puede ingresar otras especificaciones que van a permitir hacer la búsqueda más específica y eliminar falsos positivos. Por ejemplo, se puede indicar una tolerancia de masa deseada para controlar qué tan cerca están las coincidencias entre los valores m/z experimentales y los m/z calculados, que en este caso fue de 0.9 Da para las masas de los fragmentos. Luego del análisis en el programa Proteome Discoverer™, se identificaron únicamente 12 proteínas que coincidían con la base de datos de *E. coli* (**Tabla 8**). Por supuesto este no era el resultado esperado, pues es un número muy bajo de identificaciones para la cantidad de datos que se pueden obtener con esta técnica.

Como alternativa se propuso hacer una nueva búsqueda usando otro de los programas disponibles, Scaffold (Proteome Software, USA), insertando los mismos datos del programa anterior. El procesamiento de la información arrojó la identificación de 1,028 proteínas, de las cuales 23 presentaron un porcentaje de cobertura en la secuencia de 100 %. La cobertura se refiere al porcentaje de la secuencia de una proteína que fue cubierto por los péptidos identificados; es decir, qué tanto de la secuencia de una proteína se logró identificar con confianza luego de la proteómica. Dentro de las chaperonas identificadas se encontró nuevamente a IbpA con 100 % de cobertura, ClpB con 68 %, FimC (chaperona periplásmica) con 54 % y ProQ (chaperona de unión a RNA) con 20 %. Sin embargo, no se encontró a DnaK, DnaJ, GroEL ni GroES.

Tabla 8. Proteínas identificadas en los CI usando Proteome Discoverer™ 2.1 SP1 (Thermo Fisher Scientific, USA).

Descripción de proteínas	Cobertura (%)
<i>30S ribosomal protein S7 OS=Escherichia coli (strain K12 / DH10B)</i>	20.67
<i>30S ribosomal protein S10 OS=Escherichia coli O157</i>	14.56
<i>50S ribosomal protein L2 OS=Escherichia coli</i>	8.79
<i>50S ribosomal protein L27 OS=Escherichia coli</i>	17.65
<i>Threonine synthase OS=Escherichia coli</i>	3.27
<i>50S ribosomal protein L6 OS=Escherichia coli</i>	5.65
<i>50S ribosomal protein L13 OS=Escherichia coli</i>	6.59
<i>Small heat shock protein IbpA OS=Escherichia coli</i>	8.03
<i>Uncharacterized protein OS=Escherichia coli</i>	8.77
<i>50S ribosomal protein L30 OS=Escherichia coli</i>	25.42
<i>Transketolase OS=Escherichia coli</i>	1.04
<i>Phosphate acetyltransferase OS=Escherichia coli</i>	2.09

De las pocas proteínas que se lograron identificar, se resalta la presencia de IbpA, la cual se ha asociado ampliamente con el crecimiento microbiano a temperaturas elevadas y ClpB involucrada en la desagregación de los CI (Laskowska *et al.*, 1996; Mogk *et al.*, 2019). IbpA junto con IbpB hacen parte de la primera línea de defensa contra las condiciones de estrés que pueden afectar la estabilidad del proteoma (Allen *et al.*, 1992; Carra *et al.*, 2017). En *E. coli*, se ha observado antes que IbpA, IbpB y ClpB hacen parte de los CI tras la sobreproducción de proteínas recombinantes (Hoffmann y Rinas, 2000; Jürgen *et al.*, 2010), lo que justifica su presencia en nuestros resultados.

Para mejorar la cantidad y la calidad de los datos proteómicos, se modificaron varias veces los parámetros de búsqueda de ambos programas, pero en general, los resultados fueron similares. Esto nos llevó a pensar que tal vez no eran los parámetros del programa, sino la preparación de la muestra antes de la digestión y el marcaje. Según la búsqueda realizada, en la mayoría de los trabajos proteómicos basados en el uso de iTRAQ, se utiliza el conjunto de proteínas solubles de alguna muestra; en ningún caso se han utilizado cuerpos de inclusión. Durante el procesamiento de nuestra muestra, 100 µg de los CI se incubaron con 0.1 % de SDS para su solubilización. La adición de este detergente posiblemente no es compatible con los pasos posteriores de digestión y marcaje, afectando la resolución de la identificación. Para un próximo experimento de proteómica cuantitativa se sugiere no sólo encontrar los parámetros que más se ajusten a las condiciones experimentales utilizadas, sino también buscar alternativas para el tratamiento adecuado de la muestra, especialmente en el paso de la solubilización de los CI. La metodología de los geles 2D acoplada a MALDI-TOF es una opción viable para reforzar los resultados de la proteómica de los CI.

7.8. Caracterización estructural de los CI de la proteína rHuGM-CSF

7.8.1. Análisis de la estructura secundaria por ATR-FTIR

El efecto de la temperatura de pre-inducción (30 °C o 34 °C) en el contenido de la estructura secundaria de los CI de rHuGM-CSF obtenidos después de la termoinducción, fue analizado por ATR-FTIR (**Figuras 22 y 23**).

La espectroscopía infrarroja por transformada de Fourier (ATR-FTIR) es una técnica sensible para determinar la estructura secundaria de las proteínas y estudiar la formación de agregados (Miller *et al.*, 2013). FTIR funciona incidiendo radiación infrarroja sobre una muestra. Cada molécula tiene un conjunto característico de bandas de absorción en el infrarrojo. Las bandas características que se encuentran en los espectros infrarrojos de proteínas y polipéptidos incluyen la amida I y la amida II (Gallagher, 2009). Particularmente, la frecuencia de la banda amida I (1500-1700 cm⁻¹) informa sobre la vibración del enlace C=O, que está involucrado en el enlace de hidrógeno presente en diferentes elementos de la estructura secundaria como hélices α , hojas β y conformaciones desordenadas (Miller *et al.*, 2013).

Las segundas derivadas de los espectros se utilizaron para identificar las bandas principales y asignarlas a los principales componentes de la estructura secundaria de la proteína. Dicha transformación también sirve para hacer una cuantificación relativa ya que el mínimo de la segunda derivada de cada componente es proporcional al máximo de su absorción (Ami *et al.*, 2005).

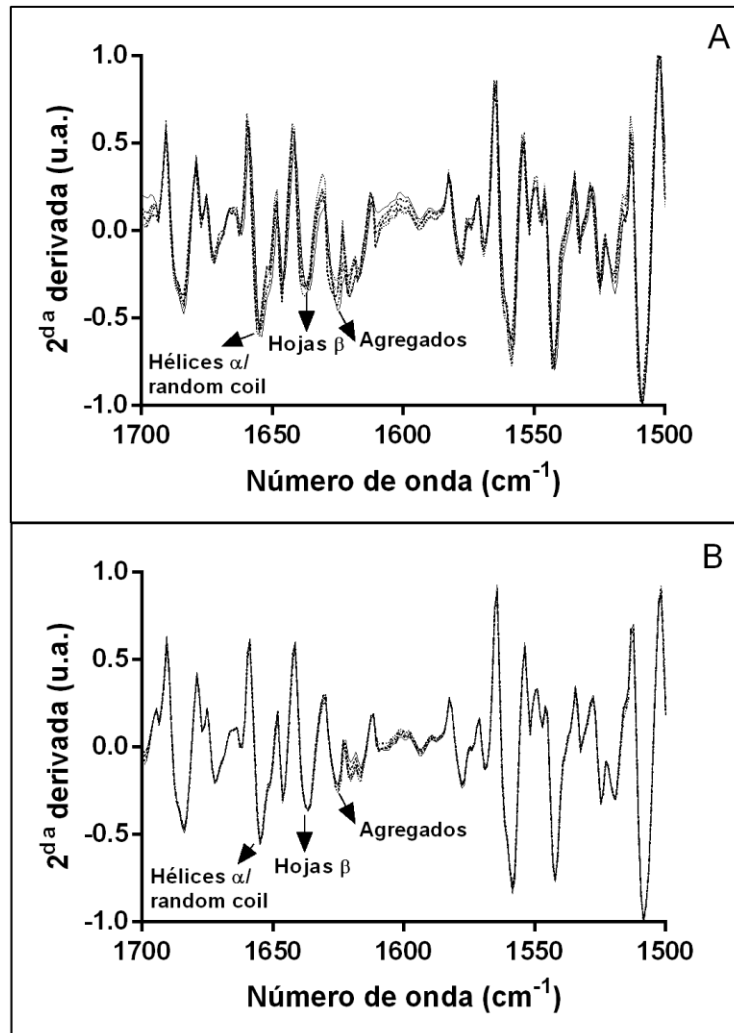


Figura 22. Segunda derivada de los espectros de absorción obtenidos por ATR-FTIR para los CI provenientes de cultivos pre-inducidos a 30 °C (A) y 34 °C (B). Los CI fueron cosechados en diferentes momentos después de la termoinducción a 42 °C: 1 h (línea sólida), 3 h (línea guión), 5 h (línea punto), 10 h (línea guión-punto) y 18/20 h (línea doble guión-punto). Los espectros representan el promedio de tres réplicas biológicas y fueron normalizados con respecto al pico de tirosina (~1508 cm⁻¹). Se señalan los picos correspondientes a las estructuras hélice α/"random coil" (~1654 cm⁻¹), hojas β (~1637 cm⁻¹) y agregados (~1625 cm⁻¹).

Con base en la literatura, se asignaron los elementos de la estructura secundaria a las bandas mayoritarias obtenidas del ATR-FTIR. La banda principal a 1654 cm^{-1} fue asignada como hélices α -"random coil", mientras que las bandas en 1636 cm^{-1} y 1625 cm^{-1} fueron designadas como hojas β en estructura nativa (hojas β) y estructuras β intermoleculares relacionadas con la conformación amiloide (agregados), respectivamente (Ami *et al.*, 2005; Singh *et al.*, 2020). En la **Figura 22** se comparan las segundas derivadas de los espectros de absorbancia en la región amida I ($1500\text{-}1700\text{ cm}^{-1}$) para los CI formados a 42°C provenientes de cultivos con pre-inducciones de 30°C (**Figura 22A**) y a 34°C (**Figura 22B**) a diferentes tiempos después de la termoinducción (1, 3, 5, 10 h y 18/20 h). Ambas figuras están normalizadas con respecto al pico de tirosina ($\sim 1508\text{ cm}^{-1}$).

En los espectros obtenidos para los CI de rHuGM-CSF recuperados a ambas temperaturas de pre-inducción (**Figura 22**), se identifica la presencia de los tres componentes estructurales asignados. Se observa que el tamaño de los picos correspondientes a las hélices α -"random coil" en 1654 cm^{-1} y a las hojas β en 1636 cm^{-1} no difieren entre las temperaturas de pre-inducción. Sin embargo, los CI de rHuGM-CSF de cultivos creciendo a $30^\circ\text{C}\rightarrow 42^\circ\text{C}$ (**Figura 22A**) exhibieron una mayor intensidad espectral de la banda a 1625 cm^{-1} asignada para los agregados, en comparación con los CI de cultivos que crecieron a $34^\circ\text{C}\rightarrow 42^\circ\text{C}$ (**Figura 22B**). Es decir, que para una mayor temperatura de pre-inducción, el contenido de agregados que tienen una conformación de tipo amiloide es menor. Es importante mencionar que en este tipo de gráfica no es posible observar diferencias claras entre los espectros con respecto al tiempo post-inducción.

Con el objetivo encontrar cambios en la estructura secundaria bajo las condiciones evaluadas con respecto al tiempo, se muestran los mínimos de la segunda derivada para cada componente estructural (**Figura 23**). El contenido de agregados amiloides fue mayor para los CI obtenidos de cultivos creciendo a 30°C desde la primera hora después de la termoinducción (**Figura 23C**), comparado con los CI a 34°C que mostraron poco enriquecimiento en estas estructuras durante todo el cultivo. El contenido de hélices α -"random coil" (**Figura 23A**) y hojas β (**Figura 23B**) fue similar en todos los CI de rHuGM-CSF, independientemente de la temperatura previa a la inducción y del tiempo de post-inducción.

7.8.2. Ensayo de fluorescencia con tioflavina T (Th-T)

Se empleó un método alternativo para la determinación del contenido estructural de los CI, que se basa en el uso de colorantes de diagnóstico amiloide, como por ejemplo, la tioflavina T (Carrió *et al.*, 2005; Wang *et al.*, 2008; De Groot *et al.*, 2009; Singh *et al.*, 2020). La tioflavina T (Th-T) ha sido descrita como un marcador específico para estudiar la conformación laminar de la estructura amiloide, ya que se une particularmente a la superficie de los canales formados por hojas β entrecruzadas en los agregados (Biancalana *et al.*, 2008; Wu *et al.*, 2009). La intercalación de Th-T en una hoja extendida de estructuras amiloides en los CI, se mide a través del incremento en la emisión máxima de fluorescencia comparada con el colorante libre. De esta manera, una mayor fluorescencia indica un mayor contenido de estructura amiloide (LeVine, 1993; Castellanos-Mendoza *et al.*, 2014).

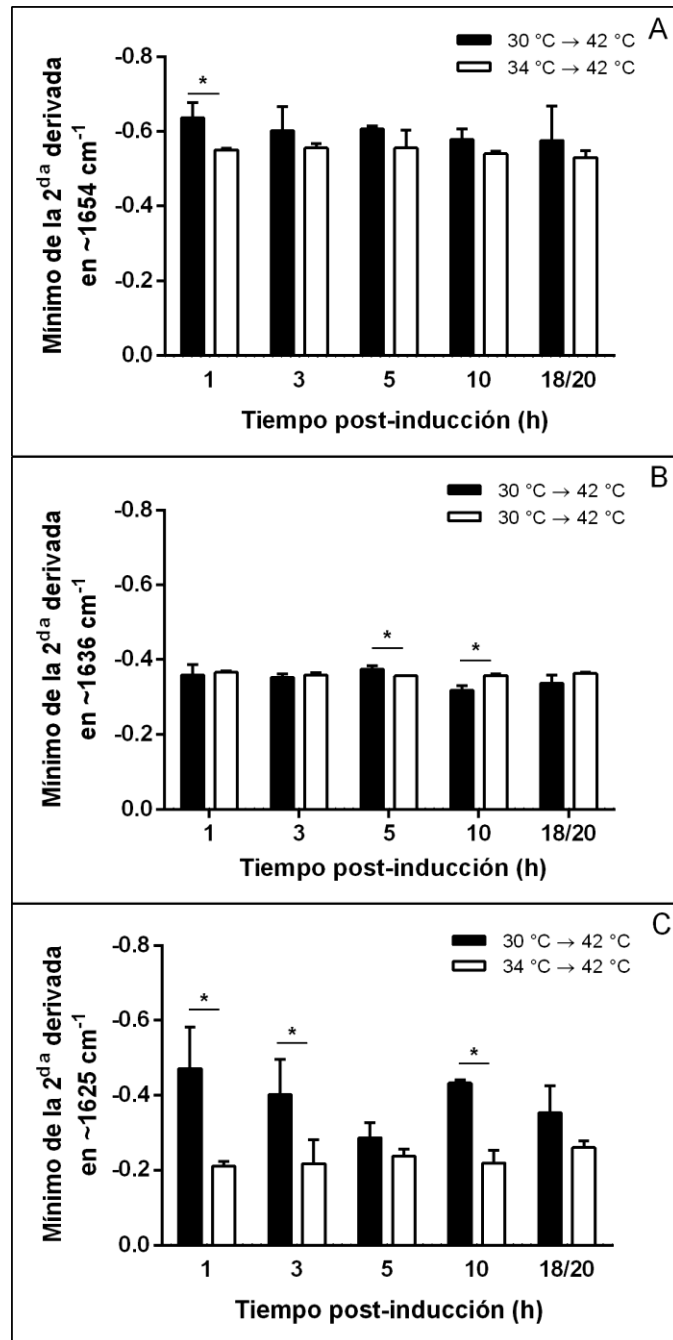


Figura 23. Comparación de los mínimos de la segunda derivada correspondientes a hélices α /"random coil" (A), hojas β (B) y agregados amiloides (C) de los CI provenientes de cultivos preinducidos a 30 °C (barras negras) y a 34 °C (barras blancas). Los CI fueron recolectados a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h). El promedio y la desviación estándar son mostrados para tres experimentos independientes.

Se evaluó el cambio en los espectros de fluorescencia de Th-T después de la incubación con los CI de rHuGM-CSF, obtenidos en diferentes momentos post-inducción (**Figura 24**). En todos los casos, el control utilizado fue una solución de Th-T en buffer sin adición de proteína. La emisión máxima de fluorescencia fue de alrededor de 485 nm, tanto en las muestras de CI obtenidos de cultivos preinducidos a 30 °C como a 34 °C, lo cual es una característica típica de los agregados amiloides (Singh *et al.*, 2020).

Los CI recolectados a partir de los cultivos que crecieron a 30 °C→42 °C mostraron un aumento gradual en la intensidad de fluorescencia con respecto al tiempo de inducción (**Figura 24A**), indicando que la cantidad de estructura amiloide fue mayor al final del cultivo (18 h de termoinducción). Mientras que para los CI obtenidos a 34 °C→42 °C estos cambios en relación con las horas de cultivo no fueron evidentes (**Figura 24B**).

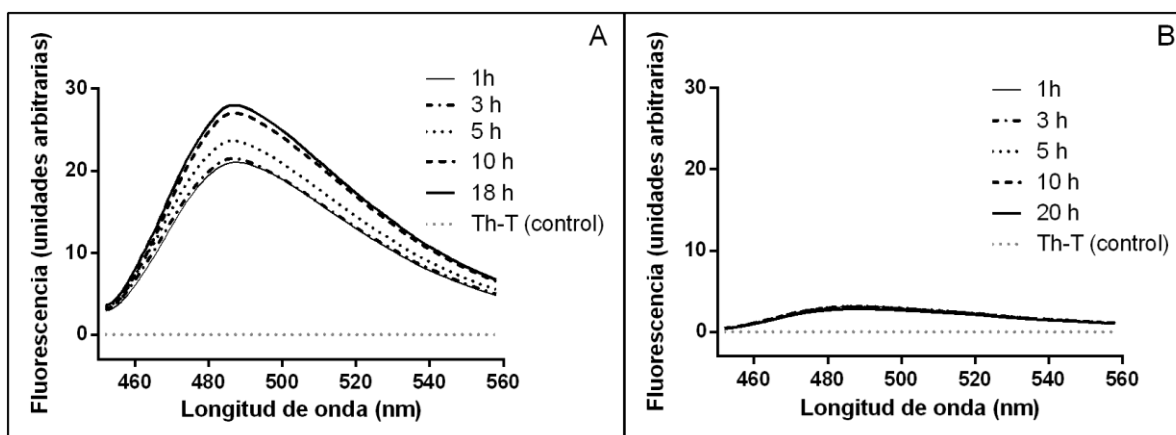


Figura 24. Espectros de fluorescencia de la unión de Th-T a los CI de rHuGM-CSF, obtenidos bajo dos temperaturas de pre-inducción: 30 °C (**A**) y 34 °C (**B**) después de la termoinducción. Los CI se recolectaron a diferentes tiempos post-inducción: 1 h (línea sólida delgada), 3 h (línea guión-punto), 5 h (línea punto), 10 h (línea guión) y 18/20 h (línea sólida gruesa). El espectro del colorante Th-T sin proteína se utilizó como control (línea punteada cercana a cero) y se presenta el promedio de dos réplicas biológicas por condición.

Al mismo tiempo, es posible observar que las intensidades de fluorescencia fueron significativamente mayores para los CI a 30 °C→42 °C (**Figura 24A**), con valores máximos cercanos a 28-30 unidades arbitrarias de fluorescencia, en comparación con los CI a 34 °C→42 °C (**Figura 24B**). Esta menor señal de fluorescencia de Th-T para los CI de cultivos a 34 °C, sugiere una pobre estructuración amiloide.

7.8.3. Degradación proteolítica con proteínasa K (PK)

La evaluación de la resistencia de los CI a la degradación proteolítica con proteínasa K (PK) ha sido útil para caracterizar la organización molecular y la estabilidad de los agregados, así como dar una indicación de su contenido de amiloide (Castellanos-Mendoza *et al.*, 2014; Calcines-Cruz *et al.*, 2018; Restrepo-Pineda *et al.*, 2019; Singh *et al.*, 2020). La PK es una serín-proteasa que exhibe poca actividad en regiones estructuralmente dominadas por hojas β , típicas de las fibrillas amiloides, pero es altamente activa en dominios hidrofílicos enriquecidos por hélices α o regiones desordenadas (De Groot *et al.*, 2009; Vázquez-Fernández *et al.*, 2012; Macedo *et al.*, 2015). Mientras menor contenido de estructura de tipo amiloide tenga un núcleo de proteína, será menos resistente a la PK y se verá una disminución significativa en la absorbancia (Upadhyay *et al.*, 2012).

Para entender el efecto de las diferentes temperaturas de pre-inducción en la organización molecular de los CI, se realizó el seguimiento de la actividad enzimática de la PK durante 100 min de reacción, midiendo la absorbancia a 350 nm (**Figura 25**). Se presenta el valor medio de dos réplicas con sus desviaciones estándar. Todos los experimentos comenzaron con la misma cantidad de proteína en los CI (50 $\mu\text{g/ml}$) y se normalizaron los valores de las absorbancias obtenidas.

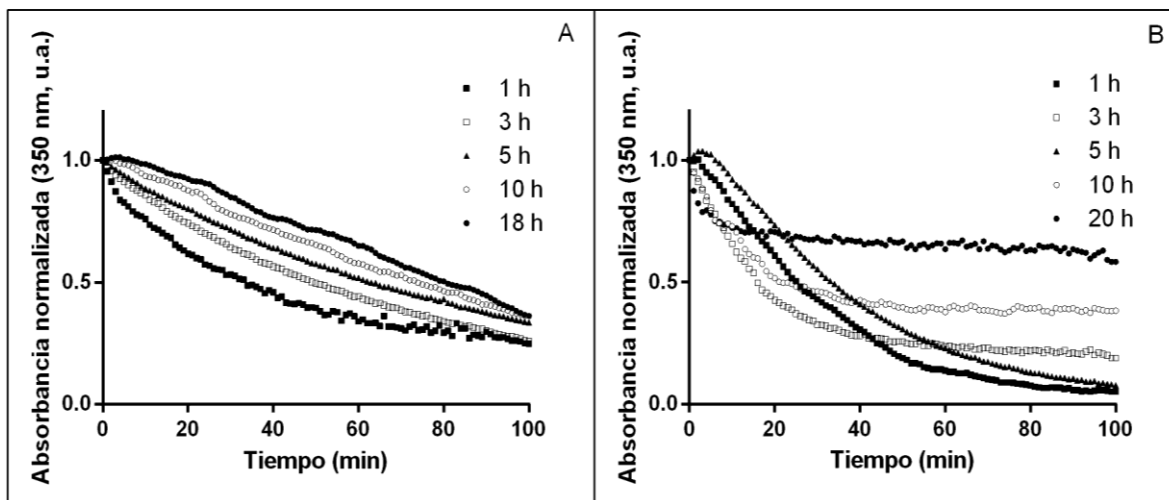


Figura 25. Cinética de la digestión proteolítica con proteínasa K (PK) de los CI de rHuGM-CSF obtenidos a partir de cultivos pre-inducidos a 30 °C (**A**) y 34 °C (**B**) y posterior termoinducción a 42 °C. Los IB se recolectaron en diferentes momentos después de la inducción: 1 h (cuadrados negros), 3 h (cuadrados blancos), 5 h (triángulos), 10 h (círculos blancos) y 18/20 h (círculos negros). Para la degradación se empleó PK a una concentración final de 25 $\mu\text{g/ml}$. Se presentan los promedios normalizados de dos réplicas biológicas.

Se observó que los CI provenientes de cultivos a una temperatura de pre-inducción de 34 °C (**Figura 25B**) fueron más propensos al ataque proteolítico por PK que los CI de cultivos

a 30 °C (**Figura 25A**), ya que presentan una pendiente más pronunciada ante la digestión con PK. Además, en las primeras horas después de la termoinducción, los CI de cultivos a 30 °C→42 °C presentaron mayor susceptibilidad a la digestión con PK que aquellos recolectados durante las últimas horas del cultivo (**Figura 25A**). El mismo comportamiento se observó bajo la condición de 34 °C→42 °C, donde los CI recuperados a las 1, 3 y 5 h después de la termoinducción, mostraron menor resistencia a la actividad PK que los cosechados a las 10 y 20 h (**Figura 25B**), demostrando que los CI de rHuGM-CSF se van haciendo más resistentes frente a la PK conforme aumenta el tiempo de inducción.

7.8.4. Solubilización con cloruro de guanidinio

Para continuar con la caracterización estructural, se determinaron los perfiles de solubilización de los CI de rHuGM-CSF obtenidos en cultivos inducidos a 42°C con pre-inducciones a 30 °C como a 34 °C, usando concentraciones crecientes de cloruro de guanidinio (GndHCl) para la misma concentración inicial de proteína (**Figura 26**).

El GndHCl es un agente caotrópico fuerte cuya naturaleza iónica provoca la desnaturalización de las proteínas globulares y, proporciona una estimación de la estabilidad conformacional de los CI (Monera *et al.*, 1994; Del Vecchio *et al.*, 2002; Castellanos-Mendoza *et al.*, 2014). En la **Figura 26**, se observa que los agregados de la proteína rHuGM-CSF obtenidos tras la termoinducción fueron sensibles a la solubilización química. Dichos perfiles de solubilización de los CI exhiben una tendencia similar para las dos temperaturas de pre-inducción evaluadas (30 °C vs 34 °C).

La cantidad de proteína solubilizada a bajas concentraciones de GndHCl (**Figura 26A**) fue muy poca. Sin embargo, al incrementar las concentraciones del agente caotrópico a 3 y 5 M (**Figuras 26B y 26C**, respectivamente), la cantidad de proteína solubilizada aumentó entre 4 y 6 veces en comparación con los valores que se obtuvieron a 1 M. Además, se puede notar que para el tratamiento a 3 y 5 M de GndHCl, aquellos CI producidos durante las últimas horas después de la termoinducción (10 y 18/20 h), fueron ligeramente más susceptibles comparados con los CI obtenidos en las primeras horas post-inducción (1 y 3 h). Finalmente, los valores máximos de proteína solubilizada se alcanzaron a una concentración de 3 M de agente caotrópico (0.555 ± 0.088 mg/ml para 30 °C y 0.451 ± 0.104 mg/ml para 34 °C).

7.9. Purificación de rHuGM-CSF plegada y análisis por dicroísmo circular

Finalmente, se presentan los resultados obtenidos por HPLC (**Figura 27A**) y dicroísmo circular (DC, **Figura 27B**) para estudiar si los procesos de solubilización de los CI, así como el plegamiento y purificación afectan la estructura secundaria de rHuGM-CSF. El DC es una herramienta muy utilizada para identificar cambios en la estructura secundaria de las proteínas (Kelly y Price 2000; Kelly *et al.*, 2005).

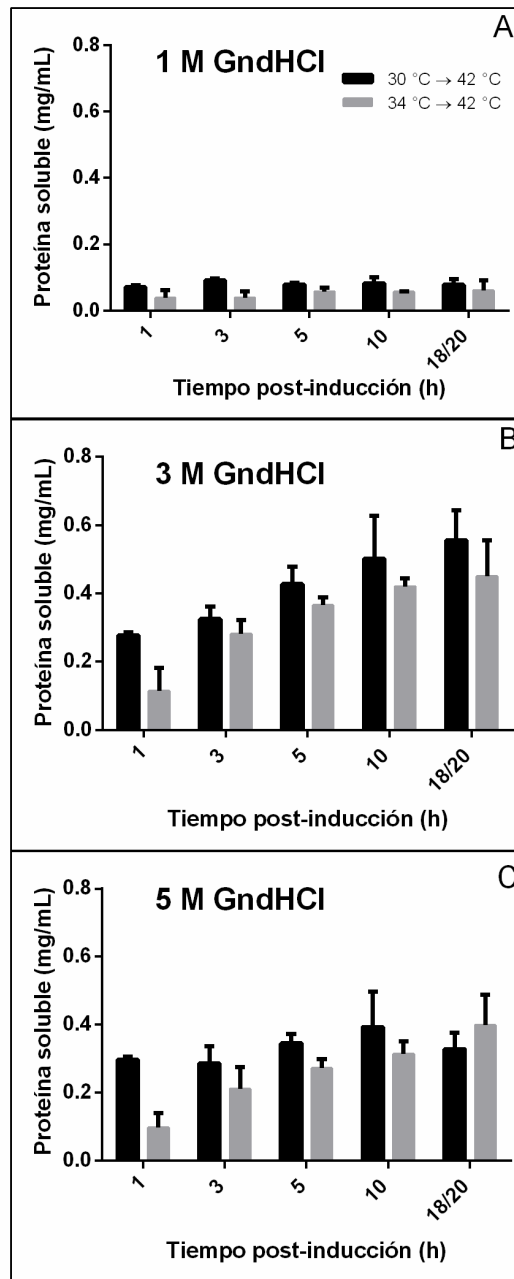


Figura 26. Perfiles de solubilización de los CI de rHuGM-CSF obtenidos de cultivos preinducidos a 30 °C (barras negras) y 34 °C (barras grises) y posterior termoinducción a 42 °C (1, 3, 5, 10 y 18/20 h). La gráfica muestra la cantidad de proteína solubilizada (mg/ml) después de 24 h de incubación con diferentes concentraciones de GndHCl: 1 M (A), 3 M (B) y 5 M (C). Se presentan los promedios y la desviación estándar de tres experimentos independientes.

Anteriormente, el DC se ha aplicado para estudiar la estructura de proteínas después de la desnaturalización y el repliegamiento y, saber bajo qué rango de condiciones son retenidas sus propiedades biológicas (Kelly y Price, 1997).

Los resultados demostraron efectivamente que no hay cambios en el contenido de estructura secundaria de la proteína tras el plegamiento y purificación, pues los espectros de DC de la proteína rHuGM-CSF bajo pre-inducción a 30 o 34 °C fueron similares (**Figura 27B**). Una conformación helicoidal típica con dos hombros negativos a ~208 y ~220 nm fue observada y los cálculos de la estructura secundaria sugieren un contenido de 36 % de hélices α y 15 % de hojas β para ambas condiciones, lo cual concuerda con valores reportados anteriormente de ~30 % de hélices α para GM-CSF recombinante en el UV lejano (Malekian *et al.*, 2019; Wingfield *et al.*, 1988). Es decir, que usando ambas temperaturas de pre-inducción con posterior inducción a 42 °C, es posible recuperar proteínas plegadas con estructuras secundarias similares a las nativas a partir de los CI.

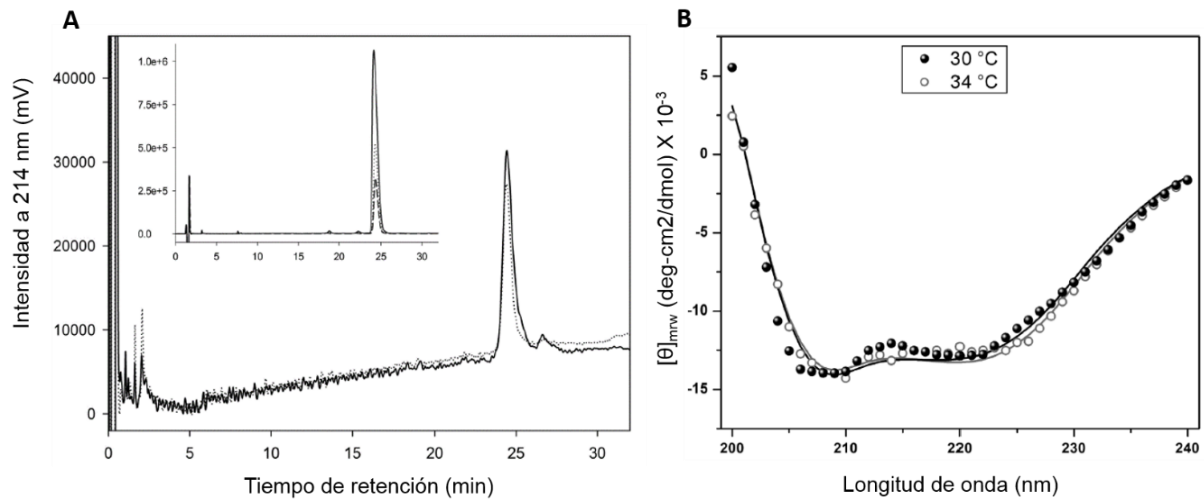


Figura 27. Purificación de rHuGM-CSF por HPLC de fase reversa (**A**). La solución de proteína plegada (~ 10 μ g) proveniente de cultivos creciendo a 30 °C (línea continua) o a 34 °C (línea punteada) después la termoinducción a 42 °C fue analizada por HPLC con una longitud de onda del detector de 214 nm. Curva estándar (en el recuadro superior) del estándar de referencia para el GM-CSF humano de la Farmacopea Europea (Molgramostim). Espectro de dicroísmo circular en el UV lejano de la proteína rHuGM-CSF purificada (**B**). Se muestra el promedio de tres escaneos, tanto para la proteína derivada de CI obtenidos de cultivos creciendo a 30 °C (círculos negros) o a 34 °C (círculos blancos) después de la termoinducción a 42 °C.

8. DISCUSIÓN

8.1. Cinética de crecimiento de *E. coli* recombinante en matraces agitados

El crecimiento de la cepa recombinante *E. coli* W3110 productora de la proteína rHuGM-CSF fue caracterizado en matraces Erlenmeyer de 250 ml con 50 ml de medio definido (**Figura 9A**). Durante la fase de crecimiento exponencial, los cultivos se mantuvieron a 30 °C con una agitación orbital de 200 rpm; posteriormente, fueron termoinducidos a 42 °C en la fase pre-estacionaria (**Figura 9B**). Los cultivos de la cepa recombinante creciendo en matraces sin inducir (30 °C), se usaron como control.

Tanto los cultivos sin inducir como los termoinducidos a 42 °C alcanzaron una DO similar (3.02 ± 0.09 u.a. y 3.13 ± 0.17 u.a., respectivamente) en la hora 21 de la cinética (**Figura 9A**). Sin embargo, los cultivos termoinducidos experimentaron una disminución del ~16 % en la DO (2.62 ± 0.30 u.a.) comparados con los cultivos que continuaron creciendo a 30 °C (3.45 ± 0.19 u.a.). Se ha reportado previamente que el incremento de la temperatura hasta 42 °C en cultivos termoinducidos, desencadena no sólo la producción de proteína recombinante, sino que también activa la respuesta de choque térmico y la síntesis de chaperonas y proteasas (Hoffmann y Rinas, 2000; Rinas *et al.*, 2007). Estas respuestas moleculares resultan en un incremento de la demanda energética celular, que a su vez puede generar una disminución en el crecimiento y alteraciones en el metabolismo central del carbono (Wittmann *et al.*, 2007; Hoffmann y Rinas, 2004; Valdez-Cruz *et al.*, 2010; Valdez-Cruz *et al.*, 2011; Restrepo-Pineda *et al.*, 2021). Este comportamiento indica que los recursos celulares son re-direccionados desde la formación de biomasa hacia la producción de proteína recombinante tras la termoinducción (Hoffmann *et al.*, 2002; Weber *et al.*, 2002).

8.2. Expresión de la proteína rHuGM-CSF en matraces agitados

La expresión de la proteína rHuGM-CSF se comparó en la fracción soluble (S) y los cuerpos de inclusión (CI) de los cultivos termoinducidos a 42 °C. Para esto, se evaluaron las muestras obtenidas en el tiempo final de cada cultivo (24 h) y se separaron las proteínas en geles de poliacrilamida al 15 % (**Figura 11**). Como control, se evaluó si había expresión en los cultivos sin inducir (mantenidos a 30 °C) tanto en la fracción soluble (S) como insoluble (I).

En los CI de los cultivos termoinducidos a 42 °C, se destaca una banda de aproximadamente 14-15 kDa (Carril 2, **Figura 11**) que presumiblemente corresponde a la proteína rHuGM-CSF, la cual no se observó en la fracción soluble (Carril 3, **Figura 11**). Esto indica que la proteína recombinante de interés se agrega formando cuerpos de inclusión tras la termoinducción. La formación de cuerpos de inclusión de la proteína GM-CSF es un fenómeno ampliamente documentado, que se ha observado cuando se expresa en diferentes cepas de *E. coli* y bajo diferentes sistemas de inducción. Los primeros estudios sobre la producción del factor GM-CSF en bacterias, se basaron en plásmidos inducibles por un cambio de temperatura de 28 a 42-48 °C (DeLamarter *et al.*, 1985; Burgess *et al.*, 1987). Se reportó que tanto el GM-CSF murino (DeLamarter *et al.*, 1985) como el GM-CSF

humano (Burgess *et al.*, 1987) formaron agregados cuando se expresaron en *E. coli*, lo que concuerda con nuestros resultados. Este conocimiento fue confirmado posteriormente usando inducción con IPTG (Schwanke *et al.* 2009; Thomson *et al.* 2012) o autoinducción metabólica (Malekian *et al.*, 2019). La tendencia a la agregación convierte al factor GM-CSF en una proteína ideal para estudiar el comportamiento estructural de los CI obtenidos en cultivos termoinducidos.

En el caso del control, es decir, en las fracciones I y S de los cultivos no inducidos (Carriles 4 y 5 respectivamente, **Figura 11**), no se observó la banda correspondiente a la proteína recombinante. La ausencia de esta banda en el cultivo sin inducir, corrobora que el sistema se encuentra regulado de manera estricta, pues el represor *ci857* está inhibiendo de manera eficaz la expresión de la proteína a temperaturas por debajo de 37 °C.

Una gran variedad de estudios han demostrado que los promotores *pL* y *pR* son finamente regulados a través de la termosensibilidad del represor *ci857* (Remaut *et al.*, 1981; Villaverde *et al.*, 1993; Nauta *et al.*, 1997). En el bacteriófago λ , la estricta regulación del represor *ci* depende de la formación de un “loop” de ADN a través de la octamerización de dímeros de λ ci (Dodd *et al.*, 2001). Un par de dímeros de λ ci unidos cooperativamente a las regiones *oR1* y *oR2* pueden interactuar con otro par de dímeros unidos a *oL1* y *oL2*, formando un “loop” de 2.4 kb (**Figura 3**) para mantener el estado lisogénico (Dodd *et al.*, 2001, 2004; Stayrook *et al.*, 2008; Lewis *et al.*, 2016; Restrepo-Pineda *et al.*, 2021). Sumado a esto, aunque *ci857* es inactivado de manera eficiente a 42 °C, las moléculas de represor que se encuentran parcialmente desestabilizadas a temperaturas inferiores a 42 °C, son capaces de auto-activar su propia transcripción para contribuir a una fuerte represión (Jechlinger *et al.*, 2005).

8.3. Cinética de crecimiento de *E. coli* recombinante en biorreactores

La caracterización cinética del crecimiento de la cepa *E. coli* W3110 productora de rHuGM-CSF en biorreactores se llevó a cabo bajo dos temperaturas de pre-inducción: 30 °C y 34 °C con posterior inducción a 42 °C, monitoreando los cambios en las concentraciones de la biomasa, glucosa y acetato durante todo el cultivo (**Figura 13**).

Para ambas temperaturas de pre-inducción, las biomásas máximas fueron similares con una $X_{m\acute{a}x}$ de 3.48 ± 0.19 g/l a 30 °C (**Figura 13A**) y de 3.48 ± 0.31 g/l a 34 °C (**Figura 13B**). Sin embargo, dicha biomasa máxima cayó levemente al final de la fermentación para ambas condiciones, posiblemente a causa de la carga metabólica generada tras la termoinducción.

Las velocidades específicas de crecimiento (μ) antes de inducir fueron significativamente diferentes ($P > 0.05$; **Tabla 5**), con un incremento de ~69 % en la μ de los cultivos cuando crecieron a 34 °C (0.90 ± 0.07 h⁻¹), en comparación con los que crecieron a 30 °C (0.53 ± 0.01 h⁻¹), lo que demuestra que una mayor temperatura de pre-inducción permite obtener un crecimiento celular más rápido. Previamente se reportó la influencia de diferentes temperaturas (30, 37 y 40 °C) en la velocidad específica de crecimiento de *E. coli* W3110IQ productora de un antígeno viral. Cultivos a 37 °C alcanzaron la mayor μ con un valor de 0.93 h⁻¹, que es 45 % y 30 % más alto que a 30 °C y 40 °C, respectivamente (Tey *et al.*, 2004). Dicha correspondencia entre la μ bacteriana y la temperatura ha sido descrita

mediante la ecuación de Arrhenius, que permite relacionar la temperatura con las velocidades de las reacciones químicas a través de su acción sobre las enzimas celulares (Huang *et al.*, 2012; Noll *et al.*, 2020). De esta manera, se ha señalado que bajas temperaturas inhiben la actividad enzimática y disminuye la velocidad del metabolismo celular y, consecuentemente, del crecimiento; mientras que temperaturas demasiado altas, pueden provocar la desnaturalización irreversible de las enzimas intracelulares (Tey *et al.*, 2004; Daniel *et al.*, 2008; Noll *et al.*, 2020).

También se compararon los valores de las DO máximas promedio obtenidas en los cultivos sin termoinducción (14.49 ± 0.01 u.a. a $30\text{ }^{\circ}\text{C}$ y 17.42 ± 1.18 u.a. a $34\text{ }^{\circ}\text{C}$; **Figura 12**) versus las DO máximas promedio de los cultivos termoinducidos (9.95 ± 1.73 u.a. para $30\text{ }^{\circ}\text{C} \rightarrow 42\text{ }^{\circ}\text{C}$ y 10.87 ± 1.19 u.a. para $34\text{ }^{\circ}\text{C} \rightarrow 42\text{ }^{\circ}\text{C}$; **Anexo 1**). Se encontró que en los biorreactores termoinducidos, hubo una disminución de ~40 % a $30\text{ }^{\circ}\text{C}$ y de ~60 % a $34\text{ }^{\circ}\text{C}$ en la acumulación de biomasa, en comparación con los biorreactores que no fueron inducidos. Las células bacterianas modificadas para la producción de proteínas recombinantes sufren diversos cambios fisiológicos y metabólicos importantes. Después de la inducción de una proteína recombinante, se ha reportado un descenso en la velocidad específica de crecimiento y en la biomasa, así como un aumento en la producción de dióxido de carbono (Jensen y Carlsen, 1990; Hofmann y Rinas, 2004). La reducción en la velocidad de crecimiento se ha relacionado con la carga metabólica, la cual ocurre cuando los nutrientes y la energía se desvían desde el metabolismo de la célula hospedera hacia la síntesis de la proteína recombinante (Bentley *et al.*, 1990). Específicamente, la carga metabólica en sistemas inducidos por temperatura es el resultado de las altas tasas de síntesis de proteína recombinante y proteínas de choque térmico, lo que aumenta los requerimientos energéticos en *E. coli* (Hoffmann y Rinas 2001). Para aliviar la demanda energética elevada, hay un reajuste en los flujos de las rutas metabólicas que permite mantener las reservas de ATP (Hoffmann y Rinas, 2004; Hoffman *et al.*, 2002; Wittmann *et al.*, 2007; Restrepo-Pineda *et al.*, 2021). En un sistema termoinducido produciendo la proteína hFGF-2 en *E. coli*, se encontró un aumento en el flujo de rutas catabólicas como la vía Embden-Meyerhof-Parnas (EMP), el ciclo de los ácidos tricarbónicos (TCA) y la fosforilación oxidativa (Weber *et al.*, 2002). Al mismo tiempo, hubo una reducción en el flujo de la vía de las pentosas fosfato, síntesis de aminoácidos y metabolismo de los lípidos (Weber *et al.*, 2002). Adicionalmente, un estudio del fluxoma y metaboloma de *E. coli* en respuesta a la termoinducción, reveló que el flujo a través de la glucólisis fue ~20% mayor durante la fase de producción que en la fase de crecimiento, lo que condujo a una acumulación significativa de los subproductos acetato, piruvato y lactato (Wittmann *et al.*, 2007).

Dichas alteraciones metabólicas, junto con los eventos moleculares asociados al estrés, pueden estar generando no sólo la diferencia de la DO en los cultivos termoinducidos comparados con los no inducidos, sino también la caída que se observa en la biomasa al final de los biorreactores termoinducidos (**Figuras 13A y 13B**).

Para las dos temperaturas de pre-inducción evaluadas, la fuente de carbono se fue agotando conforme aumentaba el crecimiento celular (**Figuras 13C y 13D**) y, simultáneamente, el acetato se fue acumulando (**Figuras 13E y 13F**). Debido a que las concentraciones máximas de acetato y biomasa fueron similares para ambas condiciones, los rendimientos de biomasa por glucosa ($Y_{X/S}$) y el de acetato por biomasa ($Y_{AC/X}$) no presentaron diferencias significativas (**Tabla 5**). Mientras que la velocidad específica de consumo de glucosa (q_s) y la velocidad específica de producción de acetato (q_p), fueron

significativamente mayores para los cultivos a una temperatura de pre-inducción de 34 °C (**Tabla 5**).

El comportamiento proporcional entre el crecimiento, el consumo de la glucosa y la acumulación de acetato en cultivos de *E. coli*, se ha reportado en numerosos trabajos. La velocidad a la que se forma el acetato está directamente relacionada con la velocidad a la que crecen las células y la velocidad a la que consumen la fuente de carbono (Eiteman y Altman, 2006; Pinhal *et al.*, 2019). En este trabajo, encontramos que la temperatura de pre-inducción tiene un importante efecto en dichas velocidades, pues los cultivos bacterianos a una temperatura de pre-inducción de 34 °C no sólo crecen a una velocidad mayor, sino que también aceleran los procesos metabólicos. Esto se ve reflejado en unos valores de q_s y q_p más altos en comparación con los cultivos a una temperatura de pre-inducción de 30 °C.

En cultivos por lotes bajo condiciones aeróbicas con glucosa como única fuente de carbono, *E. coli* acumula acetato (Basan *et al.*, 2015; Han y Eiteman, 2019). A altas concentraciones de glucosa (~20 g/L), el consumo de sustrato se da a una alta velocidad, generando un desbalance de la vía glicolítica y la vía de los ácidos tricarbónicos, lo cual lleva a la producción de acetato; a este fenómeno se le conoce como sobreflujo metabólico (Xu *et al.*, 1999; Shiloach y Rinas, 2009). El sobreflujo metabólico ocurre debido a que la velocidad de síntesis de la acetil coenzima A (AcCoA) sobrepasa la capacidad del ciclo de los ácidos tricarbónicos (TCA) para consumir el metabolito (Luli y Strohl, 1990; Bernal *et al.*, 2016). Debido a esto, la cantidad de AcCoA no consumida por el TCA es llevada a la ruta de fosfoacetilasa-acetato cinasa (Pta-Ack) en donde se sintetiza acetato (Dittrich *et al.*, 2005; Eiteman y Altman, 2006).

Se ha sugerido que la formación de acetato correlaciona con la velocidad específica de consumo de glucosa (Han y Eiteman, 2019). Al iniciar la transición entre el crecimiento exponencial y la fase estacionaria, la fuente de carbono es consumida rápidamente hasta llegar a un umbral, donde la formación de acetato es máxima (Wolfe, 2005). En este trabajo, la acumulación de acetato fue consistente con el agotamiento de la glucosa en el medio (**Figura 13**). Como era de esperarse, hubo una disminución en la velocidad específica de crecimiento en la entrada a la fase estacionaria para ambos cultivos (**Figuras 13A y 13B, Tabla 5**), lo cual coincide con el inicio de la disminución de la glucosa (**Figura 13C y 12D**) y el aumento en las concentraciones de acetato (**Figura 13E y 13F**). Para los cultivos que crecieron a 30 °C, la entrada a la fase estacionaria ocurrió entre las 8-10 h, mientras que para los cultivos a una temperatura de pre-inducción de 34 °C, la fase estacionaria comenzó entre las 6-8 h (ver escala logarítmica, **Figuras 13A y 13B**); esto se debe justamente a que a mayor temperatura de pre-inducción se presentó un crecimiento bacteriano acelerado, consumiendo más rápido los nutrientes (mayor q_s), deteniendo la producción de biomasa y generando a una mayor velocidad el acetato (mayor q_p).

En trabajos anteriores, se ha reportado la acumulación de acetato durante la producción de proteínas recombinantes bajo el sistema termoinducible $\lambda pL/pR-cl857$, obteniendo valores cercanos a 4.0 g/l en *E. coli* BL21 (Caspeta *et al.*, 2009, 2013); 0.2 g/l en *E. coli* K-12 (Mansey *et al.*, 2014); 3.5–7.4 g/l en *E. coli* 53606 (Restrepo-Pineda *et al.*, 2019) y ~6.5 g/l en *E. coli* W3110 (este trabajo). Se ha reportado que bajo termoinducción hay un incremento en el flujo del ciclo del TCA acompañado por un aumento drástico en el flujo glucolítico y en la producción de ácidos orgánicos, tras el cambio de la temperatura (Hoffmann *et al.*, 2002;

Wittmann *et al.*, 2007). Harcum y Haddadin (2006) hallaron que en cultivos fed-batch de 4 L sometidos a doble estrés por choque térmico e inducción química, algunos genes implicados en la síntesis de acetato, pertenecientes a la familia *aldH* (aldehído deshidrogenasa) son sobre expresados entre 2.1 y 3.8 veces comparados con cultivos sin estrés. Estos resultados demuestran que además del sobreflujo metabólico, el choque térmico también contribuye al aumento en la acumulación de acetato en sistemas termoinducidos.

Es importante mencionar que la producción de acetato es un evento indeseado en cultivos bacterianos, ya que representa un derroche de esqueletos carbonados, incrementa la presión osmótica interna e interfiere con la biosíntesis de la metionina; además se ha asociado con la disminución en el contenido de plásmido en *E. coli* (Roe *et al.*, 2002; Eiteman y Altman, 2006; Han y Eiteman, 2019). Al mismo tiempo, el acetato a concentraciones mayores a 2.0 g/L puede inhibir la producción de proteína recombinante (Shiloach *et al.*, 1996; Jensen y Carlsen, 1990). Por lo tanto, se han explorado diversas estrategias de optimización de las condiciones del proceso, el diseño de medios de cultivo y la modificación genética de cepas, con el objetivo de disminuir la producción de acetato y así, mejorar la producción de proteínas recombinantes (Eiteman y Altman, 2006; De Mey *et al.*, 2007).

8.4. Expresión de la proteína rHuGM-CSF en biorreactores

En primer lugar, se evaluó la expresión de la proteína recombinante en los lisados totales de *E. coli* W3110 productora de rHuGM-CSF mediante geles de poliacrilamida al 15 %, observando una banda cercana a los 14-15 kDa (**Figura 16**), que puede corresponder a la proteína de interés. En los cultivos que crecieron a 34 °C (**Figura 16B**), esta banda se aprecia de manera intensa justo 1 h después de la termoinducción a 42 °C, mientras que a 30 °C (**Figura 16A**) apenas comienza a ser perceptible, aumentando su intensidad a partir de la tercera hora de termoinducción. Se puede deducir entonces, que un cultivo a 34 °C que viene creciendo más rápido, con un estado metabólico diferente al cultivo a 30 °C y con una respuesta de choque térmico encendida, va a desencadenar la expresión de la proteína recombinante inmediatamente después de la termoinducción a 42 °C. Mientras que en un cultivo que viene creciendo a 30 °C, la expresión de la proteína heteróloga se retrasa y llega a ser significativa a partir de la tercera hora después de inducir.

Consecutivamente, se analizó la acumulación de la proteína rHuGM-CSF en los CI provenientes de cultivos que crecieron tanto a 30 °C como a 34 °C. Se realizaron geles de poliacrilamida al 15 % con 20 µg de los CI obtenidos a diferentes tiempos (1 h, 3 h, 5 h 10 h y 18/20 h) después de la inducción a 42 °C (**Figura 18**). Para ambas condiciones, se observó una banda alrededor de los ~14-15 kDa que corrió electroforéticamente a la misma altura del estándar de referencia para Molgramostim. Además, se observó la acumulación de rHuGM-CSF dentro de los CI conforme aumentó el tiempo de termoinducción, con algunas diferencias según la temperatura de pre-inducción. A una temperatura de pre-inducción de 30 °C, la proteína rHuGM-CSF representó el 45 ± 2 % de los CI al final del cultivo, mientras que a 34 °C correspondió al 53 ± 3 % de los CI (**Tabla 6**). En términos de concentración, se obtuvo un ~70 % más de rHuGM-CSF en los CI de cultivos que crecieron

a 34 °C (0.36 ± 0.05 mg/ml a las 20 h), en comparación con los cultivos que crecieron a 30 °C (0.21 ± 0.01 a las 18 h). Esto se vio reflejado en un aumento del doble en el rendimiento de rHuGM-CSF por biomasa ($Y_{PR/X}$) a 34 °C (**Tabla 7**).

Según estos resultados, el crecimiento bacteriano a 34 °C favoreció la acumulación de la proteína rHuGM-CSF en los cuerpos de inclusión durante la termoinducción, en comparación con una pre-inducción de 30 °C. Las células bacterianas cultivadas a 34 °C experimentaron una aceleración en sus procesos celulares, lo cual fue confirmado por una mayor velocidad específica de crecimiento en contraste con 30 °C (**Tabla 5**). Esto puede estar asociado con un incremento en la tasa de traducción de la proteína recombinante, una mayor concentración de polipéptidos nacientes en el citosol y, por lo tanto, aumenta la probabilidad de interacciones estereo-específicas que conducen a la agregación de rHuGM-CSF a mayor temperatura (Singh *et al.*, 2020; Adachi *et al.*, 2015). Al mismo tiempo, es posible que las células bajo pre-inducción a 30 °C requieran más recursos para hacer frente al estrés por el aumento de la temperatura. Es decir, que pueden estar consumiendo las reservas energéticas y la biomasa preferiblemente para la generación de proteínas de choque térmico (Hoffmann y Rinas, 2004). Mientras que las células que crecieron a una temperatura de pre-inducción de 34 °C, de alguna manera tienen una pre-adaptación fisiológica y metabólica al estrés térmico, favoreciendo la productividad (Bukau, 1993; Sandberg *et al.*, 2014; Kim *et al.*, 2020). En consecuencia, la temperatura de pre-inducción está modificando no sólo la velocidad específica de crecimiento, sino también la velocidad de síntesis de la proteína recombinante y su acumulación en los CI.

Anteriormente, la velocidad de crecimiento celular se ha relacionado con la cantidad de proteína recombinante contenida en los CI y su actividad biológica. Iafolla y colaboradores (2008) encontraron que a una velocidad de crecimiento rápida (tiempo de duplicación de 45 min), la correlación entre la cantidad de CI y la proteína EGFP activa fue casi 5 veces mayor que a una velocidad de crecimiento lenta (80 min). A bajas velocidades de crecimiento los CI fueron menos abundantes y con menos proteína activa (Iafolla *et al.*, 2008; De Marco *et al.*, 2019).

En un sistema termoinducido basado en el promotor *pL* y el represor *cl857*, se encontró que la expresión de la proteína recombinante depende de la velocidad específica de crecimiento antes de la inducción. En este trabajo, se evaluó el efecto de la velocidad específica de crecimiento sobre la síntesis del interferón α (IFN- α) en *E. coli* K-12. Se utilizó un quimiostato para controlar la μ previa a la inducción y, así, diferentes tasas de dilución (D) que van desde 0.025 h^{-1} a 0.2 h^{-1} fueron investigadas. Los rendimientos del IFN- α con respecto a la biomasa fueron 4 veces más altos cuando la 'D' pre-inducción se incrementó de 0.025 h^{-1} a 0.2 h^{-1} (Curless *et al.*, 1994).

8.5. Identificación de chaperonas DnaK, DnaJ, GroEL y GroES

Para la identificación de las principales chaperonas moleculares de *E. coli* (DnaK, DnaJ, GroEL y GroES) en los CI obtenidos de cultivos a temperaturas de pre-inducción de 30 y 34 °C, se realizaron inmunodetecciones con anticuerpos específicos. Todas las chaperonas se presentaron en los lisados de la cepa recombinante sin termoinducir y en los lisados de la cepa silvestre (Carriles 1 y 2, **Figura 20**), que se usaron como controles de expresión.

La expresión de DnaK en los CI fue aumentando desde las primeras horas de termoinducción (1 y 3 h) hasta alcanzar su máxima intensidad a las 5 h y volvió a su nivel inicial al final del cultivo, tanto en los CI provenientes de cultivos a 30 °C (**Figura 20A**, Panel 1) como a 34 °C (**Figura 20B**, Panel 1). Además, parece que su expresión fue mayor en los CI a 34 °C. Esto significa que la expresión de DnaK en los CI puede verse modificada por la temperatura de pre-inducción y por el tiempo de termoinducción. En el caso de GroEL, se observó una expresión constante durante todo el tiempo post-inducción, tanto a una temperatura de pre-inducción de 30 °C (**Figura 20A**, Panel 3) como de 34 °C (**Figura 20A**, Panel 3). DnaJ y GroES no fueron identificadas en los CI, pero se hallaron en las fracciones de proteína total soluble de cultivos creciendo tanto a 30 °C como a 34 °C (**Figura 21**).

E. coli cuenta con un robusto sistema de control para ayudar al plegamiento de las proteínas recién sintetizadas en el citosol y mantener la homeostasis celular, incluyendo a la chaperona DnaK con sus acompañantes DnaJ y GrpE, y a la chaperona GroEL con su co-chaperona GroES (Bhandari y Houry, 2015). En este trabajo, las chaperonas se asociaron diferencialmente a las fracciones de proteína tras la termoinducción, detectando a DnaK y GroEL en los CI (**Figura 20**), mientras que DnaJ y GroES se encontraron en la proteína soluble total (**Figura 21**). Esto nos demuestra que la interacción de las chaperonas con los agregados es un proceso dinámico, que puede ir fluctuando según los mecanismos de acción de las familias de chaperonas y las condiciones del ambiente citosólico de *E. coli*.

Como una manera de relacionar nuestros resultados con la información reportada en la literatura, se propone que la localización o asociación de las chaperonas, ya sea a la fracción soluble de proteína o a los agregados, depende de la secuencia en el procesamiento de los sustratos. Durante el plegamiento existe una cooperación dentro de la red de chaperonas, que implica la formación y disociación constante de complejos proteicos (Hartl *et al.*, 2011; Kumar y Sourjik, 2012). La comprensión actual de la actividad de chaperonas intracelulares indica que DnaJ y GrpE se unen secuencialmente a DnaK (Langer *et al.*, 1992; Doyle *et al.*, 2015). DnaJ promueve la afinidad al sustrato de DnaK al estimular la hidrólisis de ATP, luego la liberación es facilitada por GrpE, que reduce la afinidad de DnaK por el sustrato al mediar el intercambio de ADP a ATP (Genevaux *et al.*, 2007). Así, DnaK queda libre para cooperar con otras chaperonas, como IbpA/B y ClpB, que se han encontrado asociadas a los agregados de proteína (Hoffman y Rinas 2000; Jürgen *et al.*, 2010), pudiendo ser ésta la razón de encontrar a DnaK asociada a los cuerpos de inclusión y no a DnaJ.

Previamente, experimentos de FRET en *E. coli* expresando chaperonas fusionadas con proteínas fluorescentes para dilucidar la red de interacciones durante el plegamiento, revelaron que GroES es una proteína de poca interacción. Además de la interacción consigo misma, sólo se reportó unión cooperativa con GroEL. Esto significa que funciona en gran medida de manera independiente de otras chaperonas (Kumar y Sourjik, 2012). Además, GroES se recluta en una etapa posterior del ciclo de las chaperoninas y no puede participar en la transferencia de sustrato a GroEL (Kerner *et al.*, 2005), lo que puede explicar su preferencia por la fracción soluble, sin vínculo con los CI.

Las diferencias en la expresión de DnaK debido a la temperatura de pre-inducción sugieren que DnaK se requiere preferentemente a temperaturas elevadas para mantener el crecimiento celular viable y otras múltiples funciones bacterianas (Mayer, 2021). La importancia de DnaK para el crecimiento de las células a altas temperaturas ya se ha

reportado. Bajo condiciones de estrés como choque térmico a 42 °C, células de *E. coli* con mutaciones nulas en el gen de *dnak* pueden crecer lentamente a 30 °C y 37 °C, pero a 42 °C pierden la capacidad de formar colonias después de 2 h de exposición, debido a una rápida inhibición de la síntesis de ADN y ARN y a un bloqueo en la división celular (Paek *et al.*, 1987). Lo que coincide con el hecho de que la expresión de DnaK aumente con el tiempo de exposición a la temperatura de inducción (**Figura 20A y 20B**, Panel 1), manteniendo viable el crecimiento celular y la expresión de la proteína recombinante.

Por el contrario, la producción constante de GroEL durante la etapa de termoinducción, bajo las dos temperaturas de pre-inducción evaluadas (**Figura 20A y 20B**, Panel 3), corrobora su importancia como regulador central de la viabilidad celular bajo todas las condiciones de crecimiento (Hartl *et al.*, 2011; Hayer-Hartl *et al.*, 2015). Algunos estudios han reportado que GroEL cumple un trabajo de protección del crecimiento bacteriano en una amplia gama de temperaturas, desde bajas (17 a 30 °C; Fayet *et al.*, 1989) hasta temperaturas normales/altas (20 a 40 °C; Kusukawa y Yura, 1988). En resumen, GroEL parece actuar como una pieza clave para apoyar el crecimiento a temperaturas fisiológicas normales, mientras que DnaK podría ser esencial principalmente a temperaturas altas.

Haciendo una revisión bibliográfica, algunos trabajos han analizado la composición de los cuerpos de inclusión aislados de *E. coli* después de la inducción por IPTG o tras el aumento de la temperatura. Se ha revelado la presencia de chaperonas como DnaK, GroEL, GrpE, HtpG, IbpA e IbpB, pero no de DnaJ ni GroES (Valax y Georgiou, 1993; Hoffmann y Rinas, 2000; Carrió y Villaverde 2005; Jürgen *et al.*, 2010). La ausencia de éstas dos últimas no fue discutida, probablemente porque se supone que permanecen en la fracción soluble como la mayoría de las proteínas de choque térmico (> 90 %; Hoffmann y Rinas, 2000).

Además de las proteínas de choque térmico, en el proteoma insoluble de *E. coli* después de un cambio de 30 °C a 42 °C, también se encuentra el represor cl857, el factor de elongación Tu (EFTu), proteínas de membrana externa como OmpA, OmpC y OmpF y proteínas relacionadas con el metabolismo, como la succinato deshidrogenasa (Sdh), la dihidrolipoamida deshidrogenasa (LpdA), la D-tagatosa-1,6-bisfosfato aldolasa (GatY) y la triptofanasa (TnaA) (Rinas *et al.*, 2007). Recientemente, en nuestro laboratorio se reportó la presencia de la chaperona DnaK destacadamente en la fracción de proteína soluble de cultivos a dos temperaturas diferentes de inducción (39°C y 42 °C). Mientras que GroEL se presentó tanto en la fracción soluble como insoluble (Restrepo-Pineda *et al.*, 2019). En otras palabras, tanto la temperatura de pre-inducción como la de inducción, pueden impactar en la interacción de las chaperonas de plegamiento con los cuerpos de inclusión cuando se usa un sistema termoinducible en *E. coli*.

8.6. Caracterización estructural de los CI de la proteína rHuGM-CSF

En los últimos años, diversos estudios han revelado que los CI no son simples depósitos amorfos de proteínas agrupadas (Wang *et al.*, 2008; De Groot *et al.*, 2009; De Marco *et al.*, 2019). La proteína recombinante dentro de los CI puede existir en diferentes estados de estructuración, encontrando tanto especies proteicas amiloides como especies que conservan estructura de tipo nativa (García-Fruitós *et al.*, 2005; De Marco *et al.*, 2019). Se ha sugerido que el grado de organización interna de los CI va a impactar en los rendimientos de los protocolos de replegamiento y solubilización, los cuales a su vez, son importantes

para la recuperación de una proteína recombinante biológicamente activa (Rodríguez-Carmona *et al.*, 2010; Singhvi *et al.*, 2020).

Los cambios estructurales de los CI dependen no sólo de la dinámica de agregación y la naturaleza de la proteína recombinante, sino también de las condiciones del bioproceso como la temperatura, la agitación, el pH del cultivo, entre otros (De Groot y Ventura, 2006; Castellanos-Mendoza *et al.*, 2014; Valdez-Cruz *et al.*, 2017; Calcines-Cruz *et al.*, 2018; Restrepo-Pineda *et al.*, 2019; De Marco *et al.*, 2019). Por tanto, la caracterización estructural de los CI bajo diversos parámetros del cultivo ha ganado interés en los últimos años.

Con el objetivo de estudiar el efecto de la temperatura de pre-inducción (30 y 34 °C) en el contenido de la estructura secundaria de los CI de rHuGM-CSF se realizaron ensayos de ATR-FTIR. Los espectros de absorbancia y los mínimos de la segunda derivada para cada componente estructural (**Figuras 22 y 23**), indican un contenido similar de hélices α /"random coil" en 1654 cm^{-1} y hojas β en 1636 cm^{-1} para ambas temperaturas de pre-inducción. Sin embargo, los CI de rHuGM-CSF de cultivos creciendo a 30 °C exhibieron una mayor intensidad espectral de la banda a 1625 cm^{-1} asignada para agregados amiloides, revelando un mayor contenido amiloide en comparación con los CI a 34 °C (**Figura 23C**). El ensayo de unión de Th-T a los CI de rHuGM-CSF (**Figura 24**) confirmó estos resultados, pues se demostró que las intensidades de fluorescencia fueron significativamente mayores para los CI a 30 °C (**Figura 24A**), en comparación con los CI a 34 °C (**Figura 24B**), lo que sugiere mayor estructuración amiloide. Sin embargo, es importante mencionar que en algunos casos, este colorante amiloidogénico puede ser incapaz de unirse a los agregados cuando éstos son demasiado compactos y pequeños (Carrió *et al.*, 2005). Sería necesario complementar estos resultados con la información que ofrecen algunos ensayos de caracterización de tamaño de los CI, tales como microscopía electrónica de transmisión y dispersión dinámica de luz (DLS), para obtener una idea del tamaño de los CI bajo ambas temperaturas de pre-inducción.

El estudio de resistencia a la proteólisis usando PK, demostró que los CI provenientes de cultivos a una temperatura de pre-inducción de 34 °C (**Figura 25B**) fueron menos resistentes al ataque proteolítico por PK que los CI de cultivos a 30 °C (**Figura 25A**). Si relacionamos este resultado con el obtenido en el ATR-FTIR (**Figura 23**) y en el ensayo con Th-T (**Figura 24**), podríamos plantear que los CI a 34 °C tienen un núcleo con menor proporción de estructura de tipo amiloide y, por tanto, son más sensibles a la digestión con PK, en comparación con los CI a 30 °C. Adicionalmente, se encontró una relación directamente proporcional entre la resistencia de los CI de rHuGM-CSF frente a la PK y el tiempo de inducción (**Figura 25**), lo cual puede deberse a la estructuración progresiva que van adquiriendo los CI durante el cultivo. Para los CI de la hormona del crecimiento humano se describió un comportamiento similar, observando un aumento gradual de la resistencia de los CI frente a la PK con el tiempo de termoinducción (Upadhyay *et al.*, 2012).

Estudios previos han reportado que la temperatura de inducción también afecta la disposición estructural de los agregados proteicos y, al mismo tiempo, su resistencia a la digestión enzimática. Para dos sistemas diferentes, se obtuvieron efectos contrastantes. Los CI de la asparaginasa II obtenidos en un sistema inducido por IPTG a altas temperaturas (37 y 42 °C) fueron más susceptibles a la digestión con PK en comparación

con los producidos a bajas temperaturas (20 y 30 °C) (Singh *et al.*, 2020). Por otro lado, en nuestro laboratorio, los CI de la proteína rHuGM-CSF obtenidos tras inducción térmica a 39°C fueron menos resistentes a la proteólisis en comparación con los CI formados en cultivos inducidos a 42 °C (Restrepo-Pineda *et al.*, 2022b). En otras palabras, es importante mencionar que no sólo los parámetros de cultivo (en este caso la temperatura de inducción) pueden impactar en las propiedades estructurales de los cuerpos de inclusión, sino que las respuestas dependen también de la naturaleza de la proteína recombinante y el sistema de expresión.

Para finalizar el análisis estructural de los CI de rHuGM-CSF, se realizaron pruebas de solubilización con el agente caotrópico, GndHCl, a diferentes concentraciones (1, 3 y 5 M). Los CI de rHuGM-CSF se solubilizaron de manera efectiva en una concentración intermedia (3 M) de GndHCl (**Figura 26**), sin observar diferencias significativas entre las dos temperaturas de pre-inducción estudiadas. Por lo general, altas concentraciones de urea o GndHCl (6-8 M) se utilizan en la solubilización de los CI, lo que puede causar una alteración importante en la estructura nativa de la proteína recombinante y genera bajos rendimientos de recuperación de proteínas bioactivas (Upadhyay *et al.*, 2012, 2016; Singhvi *et al.*, 2020). Los resultados de la solubilización obtenidos en este estudio, se pueden considerar como una ventaja para los procesos posteriores de purificación a nivel industrial, evitando el uso de altas concentraciones de agente caotrópico. Cuando los CI se pueden solubilizar con las menores concentraciones posibles de desnaturizante, las proteínas recuperadas a partir de ellos conservan parte de su estructura plegada, haciendo más eficientes los procesos de recuperación después de la solubilización (Singh *et al.*, 2015).

8.7. Purificación de rHuGM-CSF plegada y análisis por dicroísmo circular

Con el propósito de demostrar la influencia de la temperatura de pre-inducción (34 y 30 °C) en la bioactividad de la proteína obtenida tras la solubilización de los CI y el plegamiento, se realizó el estudio de la estructura secundaria de la proteína rHuGM-CSF purificada mediante espectroscopia de dicroísmo circular (DC). Los espectros de DC de rHuGM-CSF bajo pre-inducción a 30 o 34 °C fueron similares (**Figura 27B**), comprobando que no hay cambios en el contenido de estructura secundaria de la proteína tras el plegamiento y purificación. Estos resultados sugieren que se conserva la actividad biológica de rHuGM-CSF usando ambas temperaturas de pre-inducción con posterior inducción a 42 °C, pues fue posible recuperar proteína plegada con estructura secundaria similar a la nativa a partir de los CI.

Algunos trabajos han demostrado que la proteína recombinante GM-CSF producida en *E. coli* bajo diferentes sistemas de expresión conserva su actividad biológica después de la recuperación y replegamiento de los CI (Burgess *et al.*, 1987; DeLamarter *et al.*, 1985; Schwanke *et al.*, 2009; Thomson *et al.*, 2012). El GM-CSF humano producido en *E. coli* bajo la regulación de un promotor termoinducible presentó una actividad específica de 2.9×10^7 unidades/mg con células de médula ósea (Burgess *et al.*, 1987). El GM-CSF recombinante derivado de murino estimuló el crecimiento de colonias de granulocitos y macrófagos de células de la médula ósea de ratón (DeLamarter *et al.*, 1985). Finalmente, el rHuGM-CSF purificado a partir de CI promovió el crecimiento celular en una línea de células

hematopoyéticas humanas de manera similar a la proteína comercialmente disponible (Schwanke *et al.*, 2009; Thomson *et al.*, 2012). Los resultados anteriores confirman que la ausencia de glicosilación no afecta la bioactividad de la proteína rHuGM-CSF obtenido en una plataforma bacteriana.

9. CONCLUSIONES

Las condiciones de los bioprocesos influyen en las propiedades de los CI y en el rendimiento de recuperación de las proteínas recombinantes. En este trabajo, se analizó el efecto de la temperatura de pre-inducción (30 °C vs 34 °C) sobre el crecimiento, la producción de proteína recombinante, la expresión de chaperonas y la estructura de los cuerpos de inclusión, usando el sistema termoinducible $\lambda pL/pR-cl857$ en *E. coli*.

Se encontró que la temperatura de pre-inducción de 34 °C aceleró el crecimiento bacteriano, observando un incremento de hasta el 69 % en la velocidad específica de crecimiento, con respecto a aquellos que crecieron a 30 °C. A pesar de que ambos cultivos alcanzaron biomásas máximas similares, hubo un aumento significativo en las velocidades de consumo de glucosa y de producción de acetato (q_S y q_P , respectivamente) a mayor temperatura de pre-inducción. Al mismo tiempo, el crecimiento bacteriano a 34 °C favoreció la acumulación de la proteína recombinante en los CI durante la termoinducción a 42 °C, obteniendo al final del cultivo un ~70 % más de rHuGM-CSF en los CI, en comparación con los cultivos que crecieron a 30 °C.

Las chaperonas DnaK y GroEL se asociaron a los CI durante la termoinducción, mientras que sus co-chaperonas correspondientes, DnaJ y GroES, se encontraron en la fracción de la proteína soluble.

Durante la evaluación de las propiedades estructurales usando el ensayo de fluorescencia con Tioflavina-T y ATR-FTIR, se encontró que los CI derivados de cultivos que crecieron a una temperatura de pre-inducción de 30 °C, presentaron mayor contenido estructural de tipo amiloide. En congruencia con lo anterior, los CI obtenidos a una mayor temperatura de pre-inducción (34 °C) fueron más sensible a la degradación proteolítica por proteinasa K, que aquellos obtenidos en cultivos que crecieron a 30 °C. También se encontró que los CI de rHuGM-CSF fueron solubilizados de manera efectiva con bajas concentraciones de cloruro de guanidinio (3 M y 5 M), lo cual es importante en procesos de purificación a nivel industrial.

Finalmente, los espectros de dicroísmo circular de los cuerpos de inclusión provenientes de ambas temperaturas de pre-inducción fueron comparables con la conformación helicoidal típica de la proteína, dando noción sobre la relación entre su estructura y actividad biológica.

Del trabajo realizado se concluye que un aumento de la temperatura de pre-inducción de 30 °C a 34 °C, permite obtener mayor cantidad de proteína recombinante, así como reducir el contenido estructural de tipo amiloide de los agregados. Esto conduce no sólo a mejorar los rendimientos y costos de recuperación de las proteínas a partir de los CI, sino que también brinda la posibilidad de tener una proteína recombinante activa usando el sistema termoinducible en *E. coli*. Es preciso insistir en que los CI han dejado de ser un inconveniente en muchos bioprocesos, para convertirse en una materia prima valiosa para la obtención de proteínas recombinantes. Las afirmaciones obtenidas en este estudio podrían servir como una propuesta metodológica para producir proteínas de uso farmacéutico en cuerpos de inclusión bacterianos, utilizando sistemas termoinducidos.

10. REFERENCIAS BIBLIOGRÁFICAS

Adachi M, So M, Sakurai K, Kardos J, Goto Y (2015) Supersaturation- limited and unlimited phase transitions compete to produce the pathway complexity in amyloid fibrillation. *J Biol Chem* 290(29):18134–18145.

Aggarwal K, Choe LH, Lee KH (2006). Shotgun proteomics using the iTRAQ isobaric tags. *Briefings in Functional Genomics*, 5(2):112-120.

Allen SP, Polazzi JO, Gierse JK, Easton AM (1992) Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *J Bacteriol*, 174: 6938-6947.

Ami D, Natalello A, Gatti-Lafranconi P, Lotti M, Doglia SM (2005) Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Lett* 579:3433–3436.

Arís A, Corchero JL, Benito A, Carbonell X, Viaplana E, Villaverde, A (1998) The expression of recombinant genes from bacteriophage lambda strong promoters triggers the SOS response in *Escherichia coli*. *Biotechnol Bioeng*, 60(5), 551-559.

Arsene F, Tomoyasu T, Bukau B (2000) The heat shock response of *Escherichia coli*. *J Food Microbiol*, 55(1-3), 3-9.

Atkinson GC, Tenson T, Haurlyuk V (2011). The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PloS one*, 6(8), e23479.

Aucoin MG, McMurray-Beaulieu V, Poulin F (2006) Identifying conditions for inducible protein production in *E. coli*: combining a fed-batch and multiple induction approach. *Microb Cell Fact*, 5:27.

Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MM, Ramadan HA, Saini KS, Redwan EM (2015) Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives. *J Microbiol Biotechnol*, 25(7):953-962

Baez A, Majdalani N, Shiloach J (2014) Production of recombinant protein by a novel oxygen-induced system in *Escherichia coli*. *Microb Cell Fact*, 13(1), 1-7.

Balchin D, Hayer-Hartl M, Hartl FU (2016) In vivo aspects of protein folding and quality control. *Science*, 353:aac4354.

Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol*, 22(11):1399.

Bardwell JC, Craig EA (1984) Major heat shock gene of *Drosophila* and the *Escherichia coli* heat-inducible dnaK gene are homologous. *PNAS*, 81(3), 848-852.

Barguilla I, Unzueta U, Carratalá JV, Cano-Garrido O, Villaverde A, Hernández A, Ferrer-Miralles N (2022) Toxicity Profiling of Bacterial Inclusion Bodies in Human Caco-2 Cells. *Front Bioeng Biotechnol.* 10:842256.

Basan M, Hui S, Okano H (2015) Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* 528, 99–104.

Bass SH, Yansura DG (2000) Application of the *E. coli* trp promoter. *Mol Biotechnol*, 16:253–60.

Becker J, Craig EA (1994) Heat-shock proteins as molecular chaperones. *Eur J Biochem*, 219(1-2), 11-23.

Belew M, Zhou Y, Wang S, Nyström LE, Janson JC (1994). Purification of recombinant human granulocyte-macrophage colony-stimulating factor from the inclusion bodies produced by transformed *Escherichia coli* cells. *J Chromatogr A*, 679(1):67-83.

Belková M, Koszagova R, Nahálka J (2022) Active inclusion bodies: the unexpected journey. *J Microbiol Biotechnol Food Sci*, 12(1), e5951-e5951.

Bennison DJ, Irving SE, Corrigan RM (2019). The impact of the stringent response on TRAFAC GTPases and prokaryotic ribosome assembly. *Cells*, 8(11), 1313.

Bentley WE, Mirjalili N, Andersen DC, Davis RH, Kompala DS (1990) Plasmid-encoded protein: the principal factor in the “metabolic burden” associated with recombinant bacteria. *Biotechnol Bioeng*, 35(7): 668-681.

Bernal V, Castaño-Cerezo S, Cánovas M (2016) Acetate metabolism regulation in *Escherichia coli*: carbon overflow, pathogenicity, and beyond. *Appl Microbiol Biotechnol* 100, 8985–9001.

Bernard HU, Remaut E, Hershfield MV (1979) Construction of plasmid cloning vehicles that promote gene expression from the bacteriophage lambda pL promoter. *Gene*, 5:59–76.

Biancalana M, Koide S (2010) Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *BBA-Proteins Proteomics*, 1804(7), 1405-1412.

Bhandari V, Houry WA (2015) Substrate interaction networks of the *Escherichia coli* chaperones: trigger factor, DnaK and GroEL. *Prokaryotic Syst Biol*, 271-294.

Bhatwa A, Wang W, Hassan YI, Abraham N, Li XZ, Zhou T (2021) Challenges associated with the formation of recombinant protein inclusion bodies in *Escherichia coli* and strategies to address them for industrial applications. *Frontiers Bioeng Biotechnol*, 9, 630551.

Blaszczak A, Georgopoulos C, Liberek K (1999) On the mechanism of FtsH-dependent degradation of the σ^{32} transcriptional regulator of *Escherichia coli* and the role of the DnaK chaperone machine. *Mol Microbiol*, 31(1), 157-166.

Blech M, Seeliger D, Kistler B, Bauer MM, Hafner M, Hörer S, Park JE (2012). Molecular structure of human GM-CSF in complex with a disease-associated anti-human GM-CSF autoantibody and its potential biological implications. *Biochemical Journal*, 447(2), 205-215.

Bonaventura A, Vecchié A, Wang TS, Lee E, Cremer PC, Carey B, Rajendram P, Hudock KM, Korbee L, Van Tassell BW, Dagna L and Abbate A (2020) Targeting GM-CSF in COVID-19 Pneumonia: Rationale and Strategies. *Front. Immunol.* 11:1625.

Bowden GA, Paredes AM, Georgiou G (1991) Structure and morphology of protein inclusion bodies in *Escherichia coli*. *Biotechnol*, 9:725-730.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal biochem*, 72:248-254.

Browning DF, Godfrey RE, Richards KL, Colin Robinson C, Busby SJW (2019) Exploitation of the *Escherichia coli* lac operon promoter for controlled recombinant protein production. *Biochem Soc Trans*, 47:755–63.

Bukau B (1993) Regulation of the *Escherichia coli* heat shock response. *Mol Microbiol*, 9(4), 671-680.

Caballero AB, Gamez P, Sabate R, Espargaró A (2022). Anti-Amyloid Drug Screening Methods Using Bacterial Inclusion Bodies. In: Arluison V, Wien F, Marcoleta A (eds) *Bacterial Amyloids. Methods in Molecular Biology*, vol 2538. Humana, New York, NY.

Calcines-Cruz C, Olvera A, Castro-Acosta RM, Zavala G, Alagón A, Trujillo-Roldán MA, Valdez-Cruz NA (2018) Recombinant-phospholipase A2 production and architecture of inclusion bodies are affected by pH in *Escherichia coli*. *Int J Biol Macromol*, 108:826-836.

Cano-Garrido O, Rodríguez-Carmona E, Díez-Gil C, Vázquez E, Elizondo E, Cubarsi R, Seras-Franzoso J, Corchero JL, Rinas U, Ratera I, Ventosa N, Veciana J, Villaverde A, García-Fruitós E (2013) Supramolecular organization of protein-releasing functional amyloids solved in bacterial inclusion bodies. *Acta biomaterialia*, 9(4), 6134-6142.

Carra S, Alberti S, Arrigo PA, Benesch JL, Benjamin IJ, Boelens W, Carver JA (2017). The growing world of small heat shock proteins: from structure to functions. *Cell Stress Chaperon*, 22(4), 601-611.

Carratalá, J.V., Cisneros, A., Hellman, E, Villaverde A, Ferrer-Miralles N (2021) Insoluble proteins catch heterologous soluble proteins into inclusion bodies by intermolecular interaction of aggregating peptides. *Microb Cell Fact*, 20:30.

Carrió MM, Villaverde A (2002) Construction and deconstruction of bacterial inclusion bodies. *J Biotechnol*, 96(1): 3-12.

Carrió M, Villaverde A (2003) Role of molecular chaperones in inclusion body formation. *FEBS letters*, 537(1-3):215-221.

Carrió M, González-Montalbán N, Vera A, Villaverde A, Ventura S (2005) Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol*, 347(5):1025-1037.

Caspeta L, Flores N, Pérez NO, Bolívar F, Ramírez OT (2009) The effect of heating rate on *Escherichia coli* metabolism, physiological stress, transcriptional response, and production of temperature-induced recombinant protein: A scale-down study. *Biotechnol Bioeng*, 102(2):468–482.

Caspeta L, Lara A, Pérez NO, Flores N, Bolívar F, Ramírez OT (2013) Enhancing thermo-induced recombinant protein production in *Escherichia coli* by temperature oscillations and post-induction nutrient feeding strategies. *J Biotechnol*, 167:47– 55.

Castellanos-Mendoza A, Castro-Acosta RM, Olvera A, Zavala G, Mendoza-Vera M, García-Hernández E, Valdez-Cruz NA (2014) Influence of pH control in the formation of inclusion bodies during production of recombinant Sphingomyelinase-D in *Escherichia coli*. *Microb Cell Fact*, 13(1):137.

Castiñeiras T, Williams SG, Hitchcock AG, Smith DC (2018) *E. coli* strain engineering for the production of advanced biopharmaceutical products. *FEMS Microbiol Lett*. 365(15), fny162.

Caulcott CA, Rhodes M (1986) Temperature-induced synthesis of recombinant proteins. *Trends Biotechnol*, 4:142–6.

Cebon JS, Morstyn G (1990) The potential role of granulocyte-macrophage colony stimulating factor (GM-CSF) in cancer chemotherapy. *Cancer surveys*, 9(1):131-155.

Chao YP, Wen CS, Wang JY (2004) A facile and efficient method to achieve LacZ overproduction by the expression vector carrying the thermoregulated promoter and plasmid copy number. *Biotechnol Prog*, 20:420–5.

Chakraborty A, Mukherjee S, Chattopadhyay R, Roy S, Chakrabarti S (2014) Conformational adaptation in the *E. coli* sigma 32 protein in response to heat shock. *J Phys Chem B*, 118(18), 4793-4802.

Chen J, Wu XF, Zhang YZ (2006) Expression, purification and characterization of human GM-CSF using silkworm pupae (*Bombyx mori*) as a bioreactor. *J Biotechnol*, 123(2):236-247.

Chen R (2012) Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol Adv*, 30(5): 1102-1107.

- Chou CH, Aristidou AA, Meng SY, Bennett GN, Sanet KY (1995) Characterization of a pH-inducible promoter system for high-level expression of recombinant proteins in *Escherichia coli*. *Biotechnol Bioeng*, 47:186–92.
- Clausen L, Abildgaard AB, Gersing SK, Stein A, Lindorff-Larsen K, Hartmann-Petersen R (2019) Protein stability and degradation in health and disease. *Advances in protein chemistry and structural biology* (114): 61-83.
- Craig EA, Gross CA (1991) Is hsp70 the cellular thermometer? *Trends Biochem Sci*, 16:135-140.
- Curless CE, Pope J, Loredó L, Tsai LB (1994) Effect of preinduction specific growth rate on secretion of granulocyte macrophage colony stimulating factor by *Escherichia coli*. *Biotechnol Prog*, 10(5), 467-471.
- Daniel RM, Danson MJ, Eisinger R (2008) The effect of temperature on enzyme activity: new insights and their implications. *Extremophiles* 12, 51–59.
- Das KM, Banerjee S, Shekhar N, Damodaran K, Nair R, Somani S, Raiker VP, Jain S, Padmanabhan S (2011) Cloning, soluble expression and purification of high yield recombinant hGM-CSF in *Escherichia coli*. *Int J Mol Sci* 12(3):2064–2076.
- De Groot NS, Ventura S (2006) Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Lett* 2006, 580(27):6471-6476.
- De Groot NS, Sabate R, Ventura S (2009) Amyloids in bacterial inclusion bodies. *Trends Biochem Sci*, 34(8), 408-416.
- DeLamarter JF, Mermod JJ, Liang CM, Eliason JF, Thatcher DR (1985) Recombinant murine GM-CSF from *E. coli* has biological activity and is neutralized by a specific antiserum. *EMBO J*, 4(10):2575-2581.
- Del Vecchio P, Graziano G, Granata V, Barone G, Mandrich L, Rossi M, Manco G (2002) Denaturing action of urea and guanidine hydrochloride towards two thermophilic esterases. *Biochem J*, 367(3): 857-863.
- De Marco A, Ferrer-Miralles N, Garcia-Fruitós E, Mitraki A, Peternel S, Rinas U, Trujillo-Roldán MA, Valdez-Cruz NA, Vázquez E, Villaverde A (2019) Bacterial inclusion bodies are industrially exploitable amyloids. *FEMS Microbiology Reviews*, 43(1):53-72.
- De Mey M, Maeseneire SD, Soetaert W, Vandamme E (2007) Minimizing acetate formation in *E. coli* fermentations. *J Ind Microbiol Biotechnol*, 34(11):689–700,
- Díez-Gil C, Krabbenborg S, García-Fruitós E, Vázquez E, Rodríguez-Carmona E, Ratera I, Veciana J (2010) The nanoscale properties of bacterial inclusion bodies and their effect on mammalian cell proliferation. *Biomaterials*, 31(22), 5805-5812.
- Dittrich CR, Bennett GN, San KY (2005) Characterization of the acetate-producing pathways in *Escherichia coli*. *Biotechnol Prog*, 21(4), 1062-1067.

Dodd IB, Perkins AJ, Tsemitsidis D (2001) Octamerization of λ CI repressor is needed for effective repression of PRM and efficient switching from lysogeny. *Gene Dev*, 15:3013–22.

Dodd IB, Shearwin KE, Perkins AJ (2004) Cooperativity in long-range gene regulation by the λ CI repressor. *Gene Dev*, 18:344–54.

Dodd IB, Shearwin KE, Egan JB (2005) Revisited gene regulation in bacteriophage λ . *Curr Opin Genet Dev*, 15(2):145-152.

Donovan RS, Robinson CW, Glick BR (1996) Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. *J Ind Microbiol*, 116:145–54.

Dougan M, Dranoff G, Dougan SK (2019) GM-CSF, IL-3, and IL-5 family of cytokines: regulators of inflammation. *Immunity*, 50(4), 796-811.

Doyle SM, Shastry S, Kravats AN, Shih YH, Miot M, Hoskins JR, Wickner S (2015) Interplay between *E. coli* DnaK, ClpB and GrpE during protein disaggregation. *J Mol Biol*, 427(2), 312-327.

Duzenli OF, Okay S (2020) Promoter engineering for the recombinant protein production in prokaryotic systems. *AIMS Bioeng*, 7:62.

Dvorak P, Chrast L, Nikel PI (2015) Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21 (DE3) carrying a synthetic metabolic pathway. *Microb Cell Fact*, 14:201.

Eiteman MA, Altman E (2006) Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol*, 24(11):530-6.

Ellis RJ (2006) Molecular chaperones: assisting assembly in addition to folding. *Trends Biochem ci*, 31(7), 395-401.

Espargaró A, Sabaté R, Ventura S (2008) Kinetic and thermodynamic stability of bacterial intracellular aggregates. *FEBS letters*, 582(25-26), 3669-3673.

Fatima K, Naqvi F, Younas H (2021) A review: Molecular chaperone-mediated folding, unfolding and disaggregation of expressed recombinant proteins. *Cell Biochem Biophys*, 79(2), 153-174.

Fayet O, Ziegelhoffer T, Georgopoulos C (1989) The groES and groEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J Bacteriol* 171(3):1379–1385.

Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Vázquez E, Villaverde A (2009) Microbial factories for recombinant pharmaceuticals. *Microb Cell Fact*, 8(1).

Fishov I, Zaritsky A, Grover NB (1995) On microbial states of growth. *Mol Microbiol*, 15(5):789-794.

Francisco-Cruz A, Aguilar-Santelises M, Ramos-Espinosa O, Mata-Espinosa D, Marquina-Castillo B, Barrios-Payan J, Hernandez-Pando R (2014) Granulocyte–macrophage colony-stimulating factor: not just another haematopoietic growth factor. *Medical oncology*, 31(1):774.

Füller HR, Morris GE (2012) Quantitative proteomics using iTRAQ labeling and mass spectrometry. *Integrative Proteomics*.

Gallagher W (2009) FTIR analysis of protein structure. *Course manual Chem*, 455.

García-Fruitós E, González-Montalbán N, Morell M, Vera A, Ferraz R M, Arís A, Villaverde A (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Fact*, 4(1), 27.

García-Fruitós E, Seras-Franzoso J, Vazquez E, Villaverde A (2010) Tunable geometry of bacterial inclusion bodies as substrate materials for tissue engineering. *Nanotechnol*, 21(20), 205101.

Ghosh PK, Bhardwaj D, Karnik R (2007) Human granulocyte-macrophage colony-stimulating factor: The protein and its current y emerging applications. *Indian J Biotechnol* (6):435-448.

Giacalone MJ, Gentile AM, Lovitt BT (2006) Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *Bio Techniques*, 40:355–64.

Gill RT, Valdes JJ, Bentley WE (2000) A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*. *Metab Eng*, 2(3):178-189.

Gomes SL, Simão RCG (2014) Stress responses: Heat. In *Reference Module in Biomedical Research*. Elsevier, Inc.

González-Montalbán N, García-Fruitós E, Villaverde A (2007) Recombinant protein solubility-does more mean better? *Nat Biotechnol*, 25(7):718.

Grossman AD, Straus DB, Walter WA, Gross CA (1987) Sigma 32 synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. *Genes Dev*, 1:179–84.

Guisbert E, Herman C, Lu CZ, Gross CA (2004) A chaperone network controls the heat shock response in *E. coli*. *Gene Dev*, 18:2812–21.

Gupta JC, Jaisani M, Pandey G (1999) Enhancing recombinant protein yields in *Escherichia coli* using the T7 system under the control of heat inducible lambda PL promoter. *J Biotechnol*, 68:125–34.

- Glazyrina J, Materne EM, Dreher T (2010) High cell density cultivation and recombinant protein production with *Escherichia coli* in a rocking-motion-type bioreactor. *Microb Cell Fact* 9, 42.
- Hamilton JA (2020) GM-CSF in inflammation. *The Journal of experimental medicine*, 217(1).
- Han Q, Eiteman MA (2019) Acetate formation during recombinant protein production in *Escherichia coli* K-12 with an elevated NAD(H) pool. *Eng Life Sci*, 19: 770– 780
- Harcum SW, Haddadin FT (2006) Global transcriptome response of recombinant *Escherichia coli* to heat-shock and dual heat-shock recombinant protein induction. *J Ind Microbiol Biotechnol*, 33(10):801-814.
- Hartl FU, Hayer-Hartl M (2009). Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol*, 16(6), 574.
- Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature*, 475(7356), 324-332.
- Hayer-Hartl M, Bracher A, Hartl FU (2015) The GroEL–GroES chaperonin machine: a nanocage for protein folding. *Trends Biochem Sci*, 41(1):62-76.
- Hengge-Aronis R (1993) Survival of hunger and stress: the role of rpoS in early stationary phase gene regulation in *E. coli*. *Cell*, 72(2):165-168.
- Hoffmann F, Rinas U (2000) Kinetics of heat-shock response and inclusion body formation during temperature-induced production of basic fibroblast growth factor in high-cell-density cultures of recombinant *Escherichia coli*. *Biotechnol Prog*, 16(6):1000-1007.
- Hoffmann F, Rinas U (2001) On-line estimation of the metabolic burden resulting from the synthesis of plasmid-encoded and heat-shock proteins by monitoring respiratory energy generation. *Biotechnol Bioeng*, 76(4): 333-340.
- Hoffmann F, Weber J, Rinas U (2002) Metabolic adaptation of *Escherichia coli* during temperature-induced recombinant protein production: 1. Readjustment of metabolic enzyme synthesis. *Biotechnol Bioeng*, 80(3):313-319.
- Hoffmann F, Rinas U (2004) Stress induced by recombinant protein production in *Escherichia coli*. *Adv Biochem Eng Biotechnol*, 89:73-92.
- Hortacsu A, Ryu DD (1991) Optimization of a two-stage recombinant fermentation process: the dilution rate effect. *Biotechnol Bioeng*, 38:831–7.
- Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol*, 39(3): 383-399.
- Iafolla MAJ, Mazumder M, Sardana V, Velauthapillai T, Pannu K, McMillen DR (2008) Dark proteins: effect of inclusion body formation on quantification of protein expression. *Proteins*, 72:1233–1242.

- Jäger VD, Lamm R, Küsters K, Ölçücü G, Oldiges M, Jaeger KE, Büchs J, Krauss, U (2020) Catalytically-active inclusion bodies for biotechnology—general concepts, optimization, and application. *Appl Microbiol Biotechnol*, 104(17), 7313-7329.
- Jajesniak P, Wong TS (2015) From genetic circuits to industrial-scale biomanufacturing: bacterial promoters as a cornerstone of biotechnology. *AIMS Bioeng*, 2:277–96.
- Janion C (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int J Biol Sci*, 4(6), 338.
- Jechlinger W, Glocker J, Haidinger W, Matis A, Szostak MP, Lubitz W (2005) Modulation of gene expression by promoter mutants of the λ cl857/pRM/pR system. *J Biotechnol*, 116(1), 11-20.
- Jensen EB, Carlsen S (1990) Production of recombinant human growth hormone in *Escherichia coli*: expression of different precursors and physiological effects of glucose, acetate, and salts. *Biotechnol Bioeng*, 36(1): 1-11.
- Jevsevar S, Gaberc-Porekar V, Fonda I, Podobnik B, Grdadolnik J, Menart V (2005) Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnol Prog*, 21(2):632-9.
- Jung KH (2006) Continuous production of recombinant interferon- α in *Escherichia coli* via the derepression of trp promoter using casamino acid. *Process Biochem*, 41(4):809-814.
- Jürgen B, Breitenstein A, Urlacher V, Büttner K, Lin H, Hecker M, Neubauer P (2010) Quality control of inclusion bodies in *Escherichia coli*. *Microb Cell Fact*, 9(1), 41.
- Kanemori M, Nishihara K, Yanagi H, Yura T (1997). Synergistic roles of HslVU and other ATP-dependent proteases in controlling in vivo turnover of sigma32 and abnormal proteins in *Escherichia coli*. *J Bacteriol*, 179(23), 7219-7225.
- Kaur J, Kumar A, Kaur J (2018) Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. *Int J Biol Macromol*, 106, 803-822.
- Kaushansky K, Lopez J., Brown CB (1992) Role of carbohydrate modification in the production and secretion of human granulocyte macrophage colony-stimulating factor in genetically engineered and normal mesenchymal cells. *Biochem*, 31(6):1881-1886.
- Kelly SM, Price NC (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochim Biophys Acta* 1338(2):161–185.
- Kelly SM, Price NC (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr Protein Pept Sci* 1(4):349–384.
- Kelly SM, Thomas JJ, Price NC (2005) How to study proteins by circular dichroism. *BBA-Proteins Proteom* 1751(2):119–139.

- Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, Stines AP, Hartl FU (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell*, 122(2), 209-220.
- Kim S, Kim Y, Suh DH, Lee CH, Yoo SM, Lee SY, Yoon SH (2020). Heat-responsive and time-resolved transcriptome and metabolome analyses of *Escherichia coli* uncover thermo-tolerant mechanisms. *Sci Rep* 10, 17715
- Klastersky JA, Meert AP (2016) Understanding the risk for infection in patients with neutropenia. *Intens Care Med*, 42(2):268-270.
- Koch AL (1970) Turbidity measurements of bacterial cultures in some available commercial instruments. *Anal Biochem*, 38(1):252-259.
- Kortmann J, Narberhaus F (2012) Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Microbiol*, 10(4), 255-265.
- Kosinski MJ, Rinas U, Bailey JE (1992) Isopropyl- β -dthiogalactopyranoside influences the metabolism of *Escherichia coli*. *Appl Microbiol Biotechnol*, 36:782-4.
- Krauss U, Jäger VD, Diener M, Pohl M, Jaeger KE (2017) Catalytically-active inclusion bodies carrier-free protein immobilizates for application in biotechnology and biomedicine. *J Biotechnol*, 258:136-147.
- Kuczyńska-Wiśnik D, Kędzierska S, Matuszewska E, Lund P, Taylor A, Lipinska B, Laskowska E (2002). The *Escherichia coli* small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock. *Microbiol*, 148(6), 1757-1765.
- Kumar M, Sourjik V (2012) Physical map and dynamics of the chaperone network in *Escherichia coli*. *Mol Microbiol*, 84(4), 736-747.
- Kusukawa N, Yura T (1988) Heat shock protein GroE of *Escherichia coli*: key protective roles against thermal stress. *Genes Dev* 2(7):874-882.
- Kwon HY, Park SS, Farid M, Rhee DK (2016). Protein expression via the molecular chaperone ClpL. *J Microb Biochem Technol*, 8, 2.
- Lang FM, Lee KMC, Teijaro JR, Becher B, Hamilton JA (2020) GM-CSF-based treatments in COVID-19: reconciling opposing therapeutic approaches. *Nat Rev Immun*, 20(8), 507-514.
- Langer T, Lu C, Echols H, Flanagan J, Hayer MK, Hartl FU (1992) Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature*, 356(6371), 683-689.
- Laskowska E, Wawrzynow A, Taylor A (1996) IbpA and IbpB, the new heat-shock proteins, bind to endogenous *Escherichia coli* proteins aggregated intracellularly by heat shock. *Biochimie*, 78(2): 117-122

- Lawrence JV, Maier S (1977) Correction for the inherent error in optical density readings. *Appl Env Microbiol*, 33(2):482-484.
- Lee GH, Cooney D, Middelberg APJ, Choe WS (2006) The economics of inclusion body processing. *Bioprocess Biosyst Eng* 29, 73–90.
- LeVine H (1995) Thioflavine T interaction with amyloid β -sheet structures. *Amyloid* 2(1):1–6.
- Lewis DE, Gussin GN, Adhya S (2016) New Insights into the Phage Genetic Switch: Effects of Bacteriophage Lambda Operator Mutations on DNA Looping and Regulation of PR, PL, and PRM. *J Mol Biol*, 428(22):4438-4456.
- Lethanh H, Neubauer P, Hoffmann F (2005) The small heat-shock proteins IbpA and IbpB reduce the stress load of recombinant *Escherichia coli* and delay degradation of inclusion bodies. *Microb Cell Fact*, 11;4(1):6.
- Liberek K, Galitski TP, Zylicz M, Georgopoulos C (1992) The DnaK chaperone modulates the heat shock response of *Escherichia coli* by binding to the sigma 32 transcription factor. *PNAS*, 89(8), 3516-3520.
- Lieb M (1966) Studies of heat-inducible lambda bacteriophage, I. Order of genetics sites and properties of mutant prophages. *J Mol Biol*, 16:149–63.
- Lim B, Gross CA (2010) Cellular response to heat shock and cold shock. In *Bacterial Stress Responses*, Storz G, Hengge R (eds), 91-114.
- López-García P, Forterre P (2000) DNA topology and the thermal stress response, a tale from mesophiles and hyperthermophiles. *Bioessays* 22:738–46.
- Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins*, 80: 374-381.
- Luli GW, Strohl WR (1990) Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl Env Microbiol*, 56(4), 1004-1011.
- Macedo B, Sant'Anna R, Navarro S, Cordeiro Y, Ventura S (2015) Mammalian prion protein (PrP) forms conformationally different amyloid intracellular aggregates in bacteria. *Microb Cell Fact* 14(1):174.
- Malekian R, Sima S, Jahanian-Najafabadi A, Moazen F, Akbari V (2019) Improvement of soluble expression of GM-CSF in the cytoplasm of *Escherichia coli* using chemical and molecular chaperones. *Protein Expr Purif*, 160:66-72.
- Mansey MS, Ghareeb KA, Moghazy AN, Tawfick MM, Fouda MM, El Marzugi N, El Enshasy HA (2014) Glucose concentration affects recombinant interferon α -2b production in *Escherichia coli* using thermo-induction system. *J Applied Pharm Sci*, 4(05): 001-005.

- Makrides SC (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev*, 60(3): 512-538.
- Margreiter G, Messner P, Caldwell K, Bayer K (2008) Size characterization of inclusion bodies by sedimentation field-flow fractionation. *J Biotechnol*, 138:67–73.
- Marr AG, Ingraham JL (1962) Effect of temperature on the composition of fatty acids in *Escherichia coli*. *J Bacteriol*, 84(6):1260-1267.
- Marschall L, Sagmeister P, Herwig C (2017) Tunable recombinant protein expression in *E. coli*: promoter systems and genetic constraints. *Appl Microbiol Biotechnol*, 101(2), 501-512.
- Martínez-Alonso M, García-Fruitós E, Ferrer-Miralles N, Rinas U, Villaverde A (2010) Side effects of chaperone gene co-expression in recombinant protein production. *Microb Cell Fact*, 9(1), 1-6.
- Mayer MP (2010) Gymnastics of molecular chaperones. *Mol Cell*, 39(3), 321-331.
- Mayer MP (2021) The Hsp70-chaperone machines in bacteria. *Frontiers Mol Biosci*, 512.
- Mejía R, Gómez-Eichelmann MC, Fernández MS (1995) Membrane fluidity of *Escherichia coli* during heat-shock. *Biochim Biophys Acta Biomembr*, 1239:195–200.
- Menart V, Jevsevar S, VilarMet (2003) Constitutive versus thermoinducible expression of heterologous proteins in *Escherichia coli* based on strong PR/PL promoters from phage lambda. *Biotechnol Bioeng*, 83:181–90.
- Metcalf D (2008) Hematopoietic cytokines. *Blood*, 111(2):485-491.
- Miller LM, Bourassa MW, Smith RJ (2013) FTIR spectroscopic imaging of protein aggregation in living cells. *Biochim Biophys Acta Biomembr* 1828(10):2339–2346.
- Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol*, 260:289–98.
- Modak M, Nyayanit N, Sivaram A, Patil N (2022). *The Recombinant DNA Technology Era. In: A Complete Guide to Gene Cloning: From Basic to Advanced. Techniques in Life Science and Biomedicine for the Non-Expert.* Springer, Cham.
- Mogk A, Bukau B (2017) Role of sHsps in organizing cytosolic protein aggregation and disaggregation. *Cell Stress Chaperon*, 22(4), 493-502.
- Mogk A, Ruger-Herreros C, Bukau B (2019). Cellular functions and mechanisms of action of small heat shock proteins. *Annu Rev Microbiol*, 73, 89-110.

- Monera OD, Kay CM, Hodges RS (1994) Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Sci* 3(11):1984–1991.
- Morita MT, Tanaka Y, Kodama TS, Kyogoku Y, Yanagi H, Yura T (1999) Translational induction of heat shock transcription factor sigma 32: evidence for a built-in RNA thermosensor. *Genes Dev*, 13(6):655-65.
- Mott JE, Grant RA, Ho YS (1985) Maximizing gene expression from plasmid vectors containing the lambda pL promoter: strategies for overproducing transcription termination factor rho. *Proc Natl Acad Sci USA*, 82:88–92.
- Myers JA, Curtis BS, Curtis WR (2013) Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC biophysics*, 6 (4):1-16.
- Nagai H, Yuzawa H, Yura T (1991) Regulation of the heat shock response in *E. coli*: involvement of positive and negative cis-acting elements in translation control of sigma 32 synthesis. *Biochimie*, 73(12):1473-1479.
- Nauta A, Burg B, Karsens H (1997) Design of thermolabile bacteriophage repressor mutants by comparative molecular modeling. *Nat Biotechnol* 15, 980–983 (1997).
- Neubauer P, Winter J (2001) Expression and Fermentation Strategies for Recombinant Protein Production in *Escherichia coli*. In *Recombinant Protein Production with prokaryotic and eukaryotic cells. A comparative view on host physiology*. Springer Netherlands, 195-258.
- Nicola NA, Metcalf D, Matsumoto M, Johnson GR (1983) Purification of a factor inducing differentiation in murine myelomonocytic leukemia cells. Identification as granulocyte colony-stimulating factor. *J Biol Chem*, 258(14), 9017-9023.
- Noll P, Lilge L, Hausmann R, Henkel M (2020) Modeling and Exploiting Microbial Temperature Response. *Processes*, 8(1):121.
- Nonaka G, Blankschien M, Herman C, Gross CA, Rhodius VA (2006) Regulon and promoter analysis of the *E. coli* heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. *Genes Dev*, 20:1776–89.
- O'Farrell P, Polisky B, Gelfand DH (1978) Regulated expression by readthrough translation from a plasmid-encoded betagalactosidase. *J Bacteriol*, 134:645–54.
- Ortiz-Enriquez C, Romero-Díaz ADJ, Hernández-Moreno AV, Cueto-Rojas HF, Miranda-Hernández MP, López-Morales CA, Pérez NO, Salazar-Ceballos R, Cruz-García N, Flores-Ortiz LF, Medina-Rivero E (2016) Optimization of a recombinant human growth hormone purification process using quality by design. *Prep Biochem Biotech*, 46:815–21.
- Okamoto M, Nakai M, Nakayama C, Yanagi H, Matsui H, Noguchi H, Hara H (1991) Purification and characterization of three forms of differently glycosylated recombinant human granulocyte-macrophage colony-stimulating factor. *Arch Biochem Biophys*, 286(2):562-568.

Overton TW (2014) Recombinant protein production in bacterial hosts. *Drug discovery today*, 19(5), 590-601.

Paek KH, Walker GC (1987) *Escherichia coli* dnaK null mutants are inviable at high temperature. *J Bacter* 169(1):283–290.

Palomares LA, Estrada-Mondaca S, Ramirez OT (2004) Production of recombinant proteins: challenges and solutions. *Methods Mol Biol*, 267:15-52.

Pesarrodona M, Jauset T, Díaz-Riascos ZV, Sánchez-Chardi A, Beaulieu ME, Seras-Franzoso J, Sánchez-García L, Baltà-Foix R, Mancilla S, Fernández Y, Rinas U, Schwartz S, Soucek L, Villaverde A, Abasolo I, Vázquez E (2019) Targeting Antitumoral Proteins to Breast Cancer by Local Administration of Functional Inclusion Bodies. *Advanced Science*, 6(18), 1900849.

Peternel Š, Grdadolnik J, Gaberc-Porekar V, Komel R (2008) Engineering inclusion bodies for non-denaturing extraction of functional proteins. *Microb Cell Fact*, 7(1), 34:1-9.

Pinhal S, Ropers D, Geiselmann J, De Jong H (2019) Acetate metabolism and the inhibition of bacterial growth by acetate. *J Bacteriol*, 201(13), e00147-19.

Phue JN, Shiloach J (2004) Transcription levels of key metabolic genes are the cause for different glucose utilization pathways in *E. coli* B (BL21) and *E. coli* K (JM109). *J Biotechnol*, 109:21–30.

Rambaran RN, Serpell LC (2008) Amyloid fibrils: abnormal protein assembly. *Prion*, 2(3), 112-117.

Remaut E, Stanssens P, Fiers W (1981) Plasmid vectors for high-efficiency expression controlled by the PL promoter of coliphage lambda. *Gene*, 15(1):81-93.

Restrepo-Pineda S, Bando-Campos CG, Valdez-Cruz NA, Trujillo-Roldán MA (2019) Recombinant production of ESAT-6 antigen in thermoinducible *Escherichia coli*: the role of culture scale and temperature on metabolic response, expression of chaperones, and architecture of inclusion bodies. *Cell Stress Chaperon*, 24(4): 777-792.

Restrepo-Pineda S, Pérez NO, Valdez-Cruz NA, Trujillo-Roldán MA (2021) Thermoinducible expression system for producing recombinant proteins in *Escherichia coli*: advances and insights. *FEMS Microbiol Rev*, 45(6), fuab023.

Restrepo-Pineda S, Sánchez-Puig N, Pérez NO, García-Hernández E, Valdez-Cruz NA, Trujillo-Roldán MA (2022a) The pre-induction temperature affects recombinant HuGM-CSF aggregation in thermoinducible *Escherichia coli*. *Appl Microbiol Biotechnol*, 106(8):2883-2902.

Restrepo-Pineda S, Rosiles-Becerril D, Vargas-Castillo AB, Ávila-Barrientos LP, Luviano A, Sánchez-Puig N, García-Hernández E, Pérez NO, Valdez-Cruz NA, Trujillo-Roldán MA, Valdez-Cruz NA (2022b) Induction temperature impacts the structure of recombinant HuGM-CSF inclusion bodies in thermoinducible *E. coli*. *Electron J Biotechnol*, 59, 94-106.

Richter K, Haslbeck M, Buchner J (2010). The heat shock response: life on the verge of death. *Mol cell*, 40(2), 253-266.

Riggs PD (2018) Overview of protein expression vectors for *E. coli*. *Current Protocols Essential Laboratory Techniques*, e23.

Rinas U, Hoffmann F, Betiku E, Estapé D, Marten S (2007) Inclusion body anatomy and functioning of chaperone-mediated in vivo inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. *J Biotechnol*, 127(2):244-257.

Rinas U, García-Fruitós E, Corchero JL, Vázquez E, Seras-Franzoso J, Villaverde A (2017) Bacterial Inclusion Bodies: Discovering Their Better Half. *Trends Biochem Sci*, 1334:1-12.

Rodrigues JL, Rodrigues LR (2018) Potential applications of the *Escherichia coli* heat shock response in synthetic biology. *Trends Biotechnol* 36.2: 186-198.

Rodríguez-Carmona E, Cano-Garrido O, Seras-Franzoso J, Villaverde A, García-Fruitós E (2010). Isolation of cell-free bacterial inclusion bodies. *Microb Cell Fact*, 9(1):71.

Roe AJ, O'Byrne C, McLaggan D, Booth IR (2002) Inhibition of *Escherichia coli* growth by acetic acid: a problem with methionine biosynthesis and homocysteine toxicity. *Microbiol*, 148(7), 2215-2222.

Ron EZ, Segal G, Sirkis R, Robinson M, Graur D (2000) Regulation of heat-shock response in bacteria. *Ann NY Acad Sci*, 649-655.

Rosano GL, Morales ES, Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Sci*, 28(8): 1412-1422.

Rosen R, Ron EZ (2002) Proteome analysis in the study of the bacterial heat-shock response. *Mass Spectrom Rev* 21.4: 244-265.

Rozwarski DA, Diederichs K, Hecht R, Boone T, Karplus PA (1996). Refined crystal structure and mutagenesis of human granulocyte-macrophage colony-stimulating factor. *Proteins: Structure, Function, and Bioinformatics*, 26(3), 304-313.

Sambrook J, Fritschh EF, Maniatis T (1989) *Molecular cloning*. 2da Edition. Cold Spring Harbor Lab Press NY, 2: 9-14.

Sánchez-García L, Martín L, Mangues, R, Ferrer-Miralles N, Vázquez E, Villaverde A (2016) Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact*, 15(1):33.

Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, Palsson BO, Sommer M, Feist AM (2014). Evolution of *Escherichia coli* to 42 °C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. *Mol Biol Evol*, 31(10):2647-62.

Sandoval-Basurto EA, Gosset G, Bolívar F, Ramírez OT (2005) Culture of *Escherichia coli* under dissolved oxygen gradients simulated in a two-compartment scale-down system: Metabolic response and production of recombinant protein. *Biotechnol Bioeng*, 89: 453-463

Schmidt M, Babu KR, Khanna N (1999) Temperature-induced production of recombinant human insulin in high-cell density cultures of recombinant *Escherichia coli*. *J Biotechnol*, 68:71–83.

Schramm FD, Schroeder K, Jonas K (2020) Protein aggregation in bacteria. *FEMS Microbiol Rev*, 44:54–72.

Schrimsher JL, Rose K, Simona MG, Wingfield P (1987) Characterization of human and mouse granulocyte-macrophage-colony-stimulating factors derived from *Escherichia coli*. *Biochem J*, 247(1):195-199.

Schumann W (2016) Regulation of bacterial heat shock stimulons. *Cell Stress Chaperon*, 21.6: 959-968.

Schwanke RC, Renard G, Chies JM, Campos MM, Junior ELB, Santos DS, Basso LA (2009) Molecular cloning, expression in *Escherichia coli* and production of bioactive homogeneous recombinant human granulocyte and macrophage colony stimulating factor. *Int J Biol Macromol*, 45(2): 97-102.

Seras-Franzoso J, Cano-Garrido O, Peternel S, Arís A, Garcia-Fruitós E (2022) Purification of Inclusion Bodies Produced in Bacteria and Yeast. In: Garcia Fruitós E, Arís Giralt A (eds) *Insoluble Proteins. Methods in Molecular Biology*, vol 2406. Humana, New York, NY.

Shiloach J, Kaufman J, Guillard AS, Fass R (1996) Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by *Escherichia coli* BL21 (λ DE3) and *Escherichia coli* JM109. *Biotechnol Bioeng*, 49(4), 421-428.

Shiloach J, Rinas U (2009) Glucose and acetate metabolism in *E. coli*–system level analysis and biotechnological applications in protein production processes. In *Systems biology and biotechnology of Escherichia coli* (pp. 377-400). Springer, Dordrecht.

Siegele DA, Hu JC (1997) Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc Natl Acad Sci USA*, 94:8168–72.

Singh A, Upadhyay V, Upadhyay AK, Singh SM, Panda AK (2015) Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microb Cell Fact*, 14(1), 1-10.

Singh A, Upadhyay V, Singh A, Panda AK (2020) Structure-function relationship of inclusion bodies of a multimeric protein. *Frontiers Microbiol*, 11, 876.

Singha TK, Gulati P, Kumar S (2018) Nonconventional induction strategies for production of recombinant human tumor necrosis factor-alpha in *Escherichia coli*. *J App Biol Biotechnol*, 6:23–7.

- Singhvi P, Saneja A, Srichandan S, Panda AK (2020) Bacterial inclusion bodies: a treasure trove of bioactive proteins. *Trends Biotechnol*, 38(5), 474-486.
- Sobrevilla-Calvo PDJS, Moreno NS, Carrillo FJO (2016) Neutropenia inducida por quimioterapia: el punto de vista del oncólogo. *Gaceta Mexicana de Oncología*, 15(6):344-349.
- Sollic M, Roy-Lachapelle A, Sauvé S (2015) Quantitative performance of liquid chromatography coupled to Q-Exactive high resolution mass spectrometry (HRMS) for the analysis of tetracyclines in a complex matrix. *Anal Chim Acta*, 853:415-424.
- Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol*, 115:113–28.
- Stayrook S, Jaru-Ampornpan P, Ni J (2008) Crystal structure of the λ repressor and a model for pairwise cooperative operator binding. *Nature* 452, 1022–1025.
- Steinchen W, Bange G (2016) The magic dance of the alarmones (p) ppGpp. *Mol Microbiol*, 101(4), 531-544.
- Strandberg L, Enfors SO (1991) Factors influencing inclusion body formation in the production of a fused protein in *Escherichia coli*. *Appl Environ Microbiol*, 57(6):1669-74.
- Straus DB, Walter WA, Gross CA (1987) The heat shock response of *E. coli* is regulated by changes in the concentration of sigma 32. *Nature*, 329(6137):348-351.
- Straus D, Walter W, Gross CA (1990) DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. *Genes Dev*, 4(12a), 2202-2209.
- Sung L, Nathan PC, Alibhai SM, Tomlinson GA, Beyene J (2007) Meta-analysis: Effect of Prophylactic Hematopoietic Colony-Stimulating Factors on Mortality and Outcomes of Infection. *Ann Intern Med*, 147:400-411.
- Tabandeh F, Shojaosadati SA, Zomorodipour A (2004) Heat-induced production of human growth hormone by high cell density cultivation of recombinant *Escherichia coli*. *Biotechnol Lett*, 26:245–50.
- Taylor G, Hoare M, Gray DR, Marston FAO (1986). Size and density of protein inclusion bodies. *Biotechnol*, 4(6):553.
- Tey BT, Yong KH, Ong HP (2004) Optimal conditions for hepatitis B core antigen production in shaken flask fermentation. *Biotechnol Bioproc E* 9, 374–378.
- Thomson CA, Olson M, Jackson LM, Schrader JW (2012). A simplified method for the efficient refolding and purification of recombinant human GM-CSF. *PLoS One*, 7(11): e49891.
- Toennies G, Gallant DL (1949) The relation between photometric turbidity and bacterial concentration. *Growth*, 13(1):7-20.

- Trapnell BC, Inoue Y, Bonella F, Morgan C, Jouneau S, Bendstrup E, Jouhikainen T (2020) Inhaled molgramostim therapy in autoimmune pulmonary alveolar proteinosis. *N Engl J Med*, 383(17), 1635-1644.
- Traxler MF, Summers SM, Nguyen HT, Zacharia VM, Hightower GA, Smith JT, Conway T (2008). The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol Microbiol*, 68(5), 1128-1148.
- Trujillo-Roldán MA, Peña C, Galindo E (2003) Components in the inoculum determine the kinetics of *Azotobacter vinelandii* cultures and the molecular weight of its alginate. *Biotechnol Lett*, 25(15):1251-1254.
- Tyedmers J, Mogk A, Bukau B (2010) Cellular strategies for controlling protein aggregation. *Nat Rev Mol Cell Biol*, 11(11), 777-788.
- Upadhyay AK, Murmu A, Singh A, Panda AK (2012) Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. *PLOS ONE* 7(3):e33951.
- Valax P, Georgiou G (1993) Molecular characterization of betalactamase inclusion bodies produced in *Escherichia coli*. 1. Composition. *Biotechnol Prog*, 9:539–47.
- Valdez-Cruz NA, Caspeta L, Pérez NO, Ramírez OT, Trujillo-Roldán MA (2010) Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. *Microb Cell Fact*, 9:1
- Valdez-Cruz NA, Ramírez OT, Trujillo-Roldán MA (2011) Heat shock molecular responses caused by the production of recombinant proteins in the heat inducible expression system: *E. coli* with lambda phage pL and/or pR promoters. *Bioeng Bugs*, 2:2.
- Valdez-Cruz NA, Reynoso-Cereceda GI, Pérez-Rodríguez S, Restrepo-Pineda S, González-Santana J, Olvera A, Trujillo-Roldán MA (2017). Production of a recombinant phospholipase A2 in *Escherichia coli* using resonant acoustic mixing that improves oxygen transfer in shake flasks. *Microb Cell Fact*, 16(1):129.
- Vázquez-Fernández E, Alonso J, Pastrana MA, Ramos A, Stitz L, Vidal E, Dynin I, Petsch B, Silva CJ, Requena JR (2012) Structural organization of mammalian prions as probed by limited proteolysis. *PLoS ONE* 7(11):e50111.
- Villaverde A, Benito A, Viaplana E, Cubarsi R (1993) Fine regulation of cl857- controlled gene expression in continuous culture of recombinant *Escherichia coli* by temperature. *Appl Environ Microb*, 59(10):3485-3487.
- Voth W, Jakob U (2017) Stress-activated chaperones: a first line of defense. *Trends Biochem Sci*, 42(11), 899-913.
- Walsh G (2014) Biopharmaceutical benchmarks 2014. *Nat Biotechnol*, 32(10):992-1000.
- Walsh G (2018) Biopharmaceutical benchmarks 2018. *Nat Biotechnol* 36(12):1136-1145.

Walter MR, Cook WJ, Ealick SE, Nagabhushan TL, Trotta PP, Bugg CE (1992) Three-dimensional structure of recombinant human granulocyte-macrophage colony-stimulating factor. *J Mol Biol*, 224(4):1075-1085.

Wang L, Maji SK, Sawaya MR, Eisenberg D, Riek R (2008) Bacterial inclusion bodies contain amyloid-like structure. *PLoS biology*, 6(8), e195.

Wang W, Nema S, Teagarden D (2010) Protein aggregation—pathways and influencing factors. *Int J Pharm*, 390(2), 89-99.

Weber J, Hoffmann F, Rinas U (2002) Metabolic adaptation of *Escherichia coli* during temperature-induced recombinant protein production: 2. Redirection of metabolic fluxes. *Biotechnol Bioeng*, 80(3):320-330.

Wegrzyn RD, Deuerling E (2005) Molecular guardians for newborn proteins: ribosome-associated chaperones and their role in protein folding. *Cellular and Molecular Life Sciences CMLS*, 62(23), 2727-2738.

Willbold D, Strodel B, Schröder GF, Hoyer W, Heise H (2021) Amyloid-type protein aggregation and prion-like properties of amyloids. *Chem Rev*, 121(13), 8285-8307.

Williams DC, Van Frank RM, Muth WL, Burnett JP (1982) Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Science*, 215(4533), 687-689.

Wingfield P, Graber P, Moonen P, Craig S, Pain R (1988) The conformation and stability of recombinant-derived granulocyte-macrophage colony stimulating factors. *Eur J Biochem* 173(1):65–72.

Wittmann C, Weber J, Betiku E, Krömer J, Böhm D, Rinas U (2007) Response of fluxome and metabolome to temperature-induced recombinant protein synthesis in *Escherichia coli*. *J Biotechnol*, 132:375-384.

Wolfe AJ (2005) The acetate switch. *Microbiol Mol Biol Rev*, 69(1), 12-50.

Xu B, Jahic M, Blomsten G (1999) Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with *Escherichia coli*. *Appl Microbiol Biotechnol* 51, 564–571.

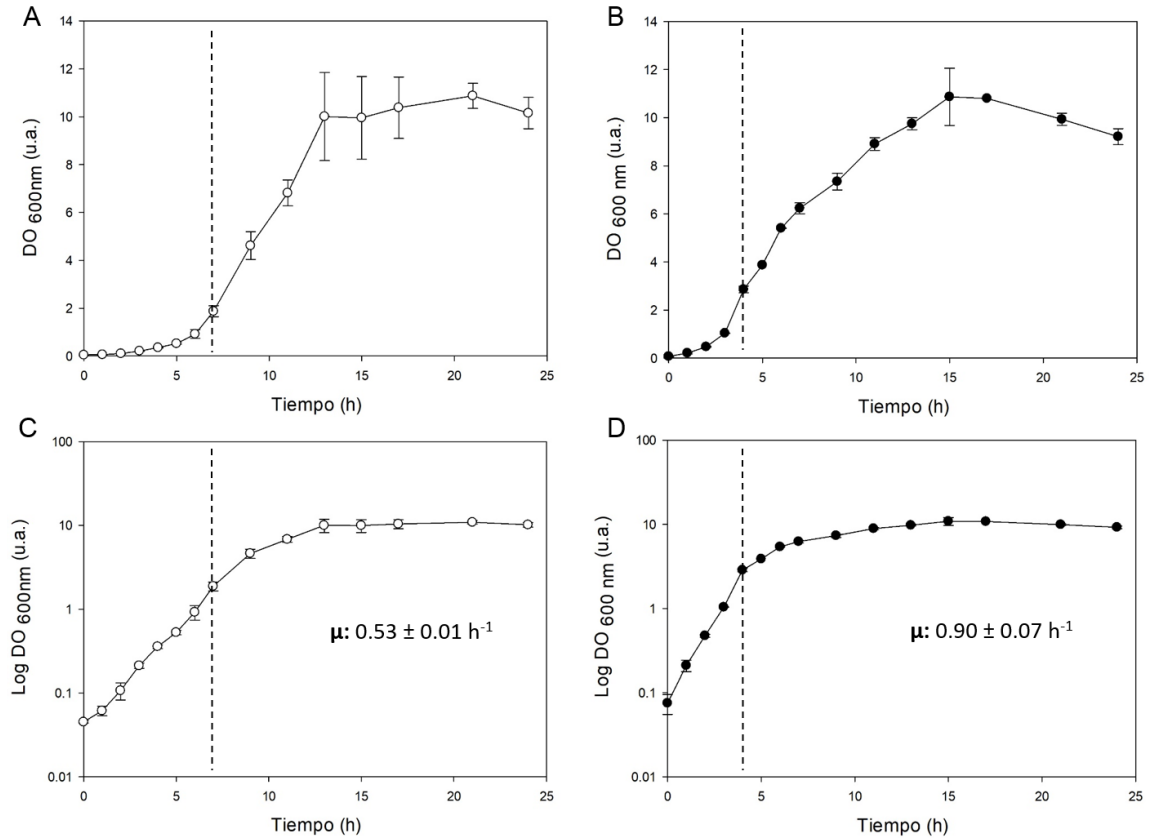
Yamamori T, Ito K, Nakamura Y, Yura T (1978) Transient regulation of protein synthesis in *Escherichia coli* upon shift-up of growth temperature. *J Bacteriol*, 134(3):1133-1140.

Yamamori T, Yura T (1980). Temperature-induced synthesis of specific proteins in *Escherichia coli*: evidence for transcriptional control. *J Bacteriol*, 142(3), 843-851.

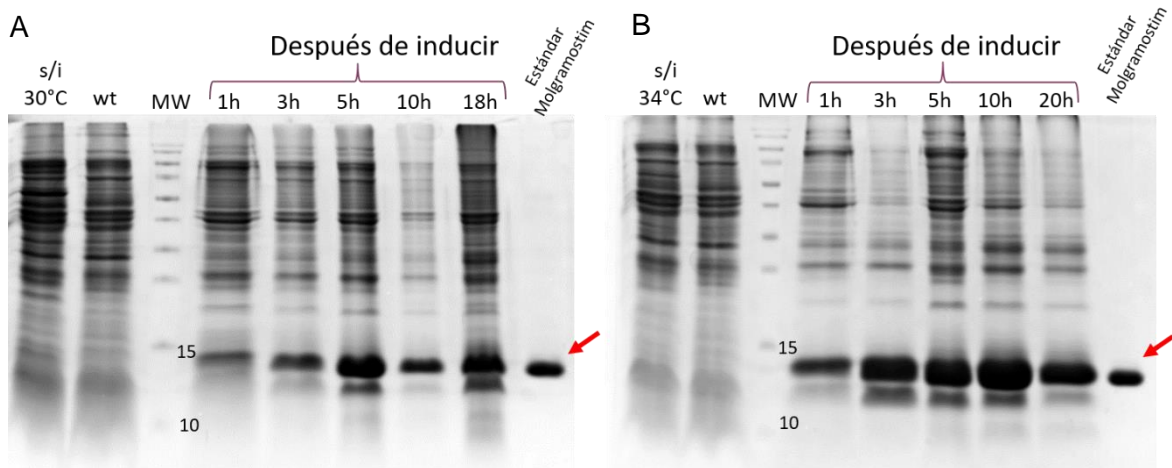
Yura T, Nakahigashi K (1999) Regulation of the heat-shock response. *Curr Opin Microbiol*, 2(2), 153-158.

Yuzawa H, Nagai H, Mori H, Yura T (1993) Heat induction of σ^{32} synthesis mediated by mRNA secondary structure: a primary step of the heat shock response in *Escherichia coli*. *Nucleic Acids Res*, 21(23):5449-5455.

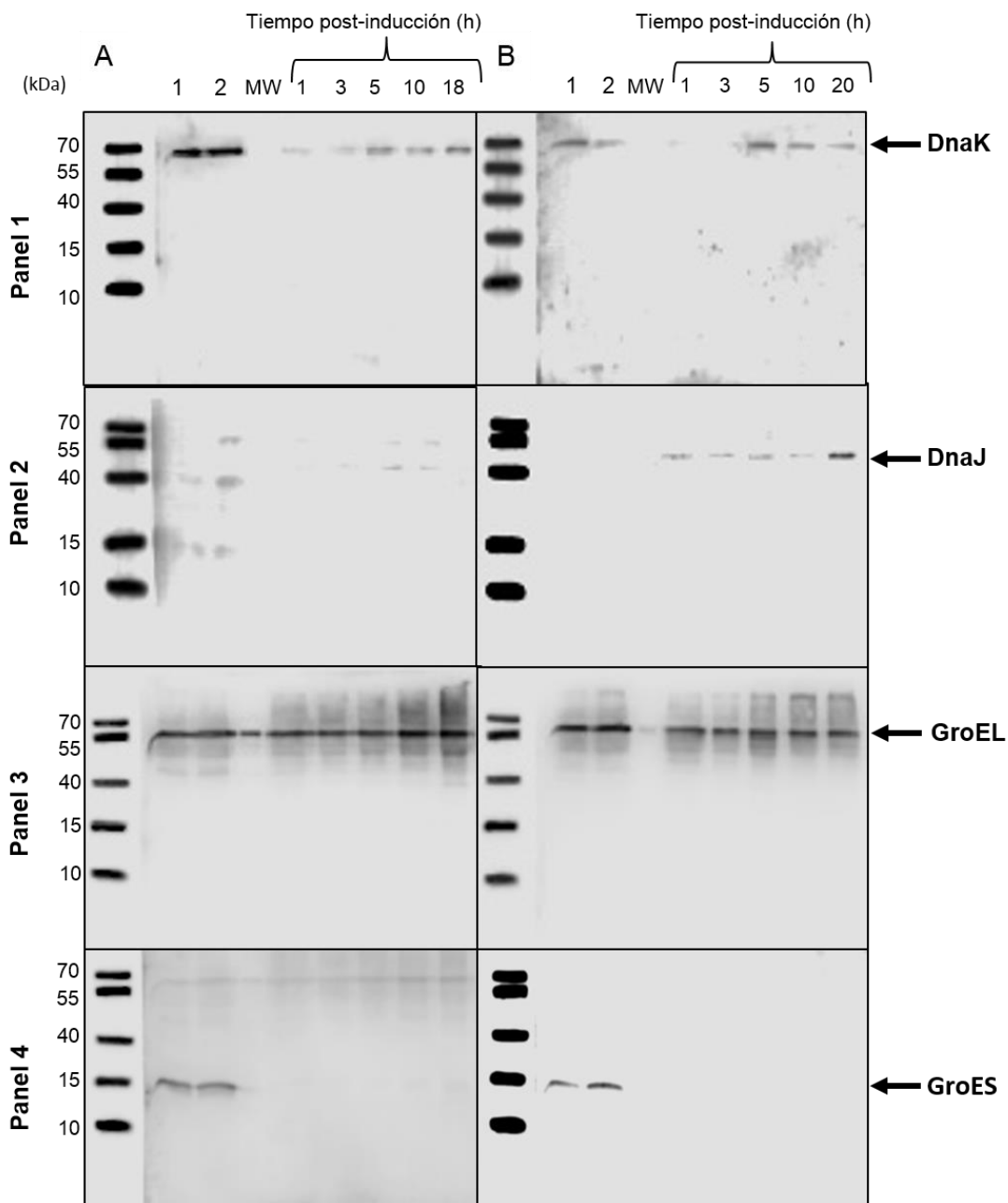
11. ANEXOS



Anexo 1. Cinética del crecimiento bacteriano de *E. coli* W3110 + rHuGM-CSF en biorreactores de 1.2 L termoinducidos. Dos temperaturas de pre-inducción fueron evaluadas: 30 °C (**A**) y 34 °C (**B**) con posterior termoinducción a 42 °C. Las líneas punteadas indican el inicio del aumento de temperatura (7 h para 30 °C y 4 h para 34 °C). Las escalas logarítmicas para el cultivo pre-inducido a 30 °C (**C**) y a 34 °C (**D**) también se presentan. Se muestra el promedio y la desviación estándar de tres experimentos independientes.




Anexo 2. Réplica biológica de la expresión de la proteína rHuGM-CSF (~14-15 kDa) en *E. coli* W3110 recombinante usando biorreactores. Dos temperaturas de pre-inducción diferentes fueron evaluadas: 30 °C (**A**) y 34 °C (**B**) en un gel de poliacrilamida al 15 % teñido con azul de Coomasie. Se presentan los cuerpos de inclusión (CI) con 20 µg de proteína a diferentes tiempos después de la termoinducción (1h, 3 h, 5 h y tiempo final). Se usaron como controles negativos, la cepa sin inducir (*rE. Coli* a 30 °C o 34 °C) y la cepa silvestre (WT *E. coli*). El control positivo es el estándar de referencia para Molgramostim (Molgra).



Anexo 3. Réplica biológica de la inmunodetección de las chaperonas GroEL, GroES, DnaK y DnaJ en los CI de la proteína rHuGM-CSF provenientes de cultivos bajo diferentes temperaturas de crecimiento: 30 °C (**A**) y 34 °C (**B**), termoinducidos a 42 °C. Carriles 1A y 1B: lisados de la cepa recombinante sin inducir a 30 °C o 34 °C, respectivamente. Carriles 2A y 2B: lisados de la cepa silvestre. Carril MW: marcador de peso molecular. Se presentan los CI a diferentes tiempos post-inducción (1, 3, 5, 10 y 18/20 h) bajo las dos condiciones evaluadas. Las flechas indican las bandas correspondientes a DnaK (~70 kDa, panel 1), DnaJ (~41 kDa, panel 2), GroEL (~60 kDa, panel 3) y GroES (~15 kDa, panel 4).



The pre-induction temperature affects recombinant HuGM-CSF aggregation in thermoinducible *Escherichia coli*

Sara Restrepo-Pineda¹ · Nuria Sánchez-Puig² · Néstor O. Pérez³ · Enrique García-Hernández² · Norma A. Valdez-Cruz¹ · Mauricio A. Trujillo-Roldán^{1,4} 

Received: 2 November 2021 / Revised: 28 March 2022 / Accepted: 30 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The overproduction of recombinant proteins in *Escherichia coli* leads to insoluble aggregates of proteins called inclusion bodies (IBs). IBs are considered dynamic entities that harbor high percentages of the recombinant protein, which can be found in different conformational states. The production conditions influence the properties of IBs and recombinant protein recovery and solubilization. The *E. coli* growth in thermoinduced systems is generally carried out at 30 °C and then recombinant protein production at 42 °C. Since the heat shock response in *E. coli* is triggered above 34 °C, the synthesis of heat shock proteins can modify the yields of the recombinant protein and the structural quality of IBs. The objective of this work was to evaluate the effect of different pre-induction temperatures (30 and 34 °C) on the growth of *E. coli* W3110 producing the human granulocyte–macrophage colony-stimulating factor (rHuGM-CSF) and on the IBs structure in a $\lambda pL/pR$ -cI857 thermoinducible system. The recombinant *E. coli* cultures growing at 34 °C showed a ~69% increase in the specific growth rate compared to cultures grown at 30 °C. The amount of rHuGM-CSF in IBs was significantly higher in cultures grown at 34 °C. Main folding chaperones (DnaK and GroEL) were associated with IBs and their co-chaperones (DnaJ and GroES) with the soluble protein fraction. Finally, IBs from cultures that grew at 34 °C had a lower content of amyloid-like structure and were more sensitive to proteolytic degradation than IBs obtained from cultures at 30 °C. Our study presents evidence that increasing the pre-induction temperature in a thermoinduced system allows obtaining higher recombinant protein and reducing amyloid contents of the IBs.

Key Points

- Pre-induction temperature determines inclusion bodies architecture
- In pre-induction (above 34 °C), the heat shock response increases recombinant protein production
- Inclusion bodies at higher pre-induction temperature show a lower amyloid content

Keywords Growth temperature · Recombinant protein · Thermoinduction · Inclusion body · Chaperones, rHuGM-CSF

✉ Mauricio A. Trujillo-Roldán
maurotru@gmail.com; maurotru@biomedicas.unam.mx

¹ Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad de México, CP 04510, México

² Universidad Nacional Autónoma de México, Instituto de Química, Ciudad Universitaria, Ciudad de México 04510, México

³ Probiomed S.A. de C.V. Planta Tenancingo, Cruce de Carreteras Acatingo-Zumpahuacan SN, Tenancingo CP 52400, Estado de México, México

⁴ Departamento de Biología Molecular y Biotecnología, Unidad de Bioprocesos, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad de México, CP 04510, México

Introduction

The production of recombinant proteins for therapeutic use has become of great interest in current biotechnology and biomedicine. To achieve high product yields, make bio-processes efficient, and minimize production costs, different biological platforms, expression systems, culture medium designs, and purification strategies have been explored (Assenberg et al. 2013; Terol et al. 2021; Kumar et al. 2020; Kaur et al. 2018). *Escherichia coli* as a host for recombinant protein production has numerous advantages compared to other bacteria, yeasts, or mammalian cells (Baeshen et al. 2015; Sánchez-García et al. 2016). *E. coli* can grow on inexpensive substrates in a short time, it is a well-known

organism, high cell density cultures can be easily reached, and the accumulation of the recombinant protein can represent 30% or more of the total cellular protein (Huang et al. 2012; Rosano and Ceccarelli 2014; Rosano et al. 2019). Additionally, within the biopharmaceuticals approved through 2018, ~22% were produced in *E. coli* (Walsh 2018).

When recombinant proteins are overexpressed in *E. coli*, the bacteria tend to form protein aggregates called inclusion bodies (IBs) (Baneyx and Mujacic 2004; Williams et al. 1982). IBs are recombinant protein-enriched deposits with spherical or pseudo-spherical shapes and diameters between 50–800 nm (De Marco et al. 2019; Margreiter et al. 2008; Castellanos-Mendoza et al. 2014). Its formation is a dynamic phenomenon that occurs mainly due to the interactions between partially folded or misfolded polypeptide chains during the overexpression of recombinant proteins, stressful conditions, or the shortage of heat shock proteins (HSPs), which are responsible for maintaining the cell proteome homeostasis (Baneyx and Mujacic 2004; Fahner et al. 2004; Carrió and Villaverde 2003; Hartl and Hayer-Hartl 2009). Although IBs were considered an obstacle in obtaining soluble and active recombinant proteins, recent studies have shown that proteins with native-like structure and polypeptides with amyloid characteristics coexist within IBs, giving them both biological functionality and mechanical stability (González-Montalbán et al. 2007; Cano-Garrido et al. 2013; Rinas et al. 2017; Singhvi et al. 2020). This progress in the perception of IBs has increased the scientific interest in the study of their molecular organization and application in various areas as materials for cell proliferation (García-Fruitós et al. 2010), drug release agents (Villaverde et al. 2012; Pesarrodonna et al. 2019) and biocatalysts (Jäger et al. 2020). Nowadays, it is known that bioprocess conditions such as pH (Castellanos-Mendoza et al. 2014; Calcines-Cruz et al. 2018), temperature (de Groot and Ventura 2006; Peternel et al. 2008; Restrepo-Pineda et al. 2019), agitation (Valdez-Cruz et al. 2017), among others, can determine the formation, composition, and structure of IBs (De Marco et al. 2019). However, the effect of different induction strategies on its properties, for example, the use of a temperature-induced system, is a subject that remains poorly explored.

The $\lambda pL/pR$ -cI857 thermoinducible expression system is a widely applied strategy in the industry to produce recombinant proteins in *E. coli* (Valdez-Cruz et al. 2010, 2011; Restrepo-Pineda et al. 2021). The use of this system avoids the addition of chemical inducers such as isopropyl- β -D-1-thiogalactopyranoside (IPTG), the possibilities of contamination are minimized through the external control of temperature, and target recombinant protein yields of at least 30% can be obtained regarding the total protein (Remaut et al. 1981; Caspeta et al. 2013; Singha et al. 2018). Below 37 °C, recombinant protein expression is regulated by the binding of the cI857 thermolabile repressor to the operator regions

of the pL and pR promoters derived from bacteriophage λ , facilitating the formation of a DNA loop that inhibits RNA polymerase activity (Dodd et al. 2004; Lewis et al. 2016). By raising to 37 °C, cI857 is released and allows the transcription of the gene of interest (Caulcott and Rhodes 1986; Villaverde et al. 1993; Valdez-Cruz et al. 2010). However, it has been reported that above 34 °C, *E. coli* initiates the heat shock response (HSR) to cope with heat stress and maintain cellular homeostasis (Yamamori et al. 1978; Morita et al. 1999; Yano et al. 1990; Yura 2019). HSR is controlled by the transcription factor σ^{32} (RpoH), which regulates the expression of an extensive network of chaperones and proteases involved in folding of nascent proteins and the removal of damaged/unfolded proteins (Guisbert et al. 2004, 2008; Baneyx and Mujacic 2004; Balchin et al. 2016). The temperature upshift during recombinant protein synthesis also involves a metabolic adaptation, reflected in the reduction of the specific growth rate, accumulation of organic acids, mainly acetate, and a readjustment in metabolic fluxes (Hoffmann and Rinas 2004; Hoffman et al. 2002; Wittmann et al. 2007; Restrepo-Pineda et al. 2021). Recombinant protein production typically follows a two-phase strategy; initially, cells are grown at 30 °C (in batch or feed-batch cultures), and then recombinant protein expression is induced at a constant temperature between 38 °C and 42 °C (Caspeta et al. 2009, 2013; Valdez-Cruz et al. 2010; Restrepo-Pineda et al. 2021).

In a thermoinduced system in addition to HSR, recombinant protein overexpression and the IBs formation co-occur (Valdez-Cruz et al. 2010, 2011; Restrepo-Pineda et al. 2021). Also, HSPs are overexpressed in response to elevated temperature and recombinant protein accumulation (Gill et al. 2000; Carrió and Villaverde 2003). During thermoinduction, the mRNA levels of heat shock genes (*dnaK*, *dnaJ*, and *groEL*) increase between 2 and 9 times compared with 30 °C (Valdez-Cruz et al. 2011). At the proteomic level, Hoffmann and Rinas (2000) found that just 30 min after thermoinduction at 42 °C, HSPs reach their maximum synthesis rate, and IBs analyzed one-hour post-induction contains chaperones such as DnaK, GroEL, IbpA, and IbpB. Previous findings in our laboratory revealed that using a thermoinducible system, GroEL is present in both the soluble protein fraction and in the IBs. At the same time, DnaK predominated in the soluble fraction (Restrepo-Pineda et al. 2019). The presence of these chaperones in IBs has been related to their function in the dissolution of these aggregates or preventing their aggregation during thermoinduction (Rinas et al. 2007). Moreover, the assembly and morphology of IBs can also be modified depending on the absence or presence of specific HSPs (García-Fruitós et al. 2010). Furthermore, IBs harvested from *E. coli* cultures induced by IPTG at high temperatures (39 or 42 °C) were more accessible to be solubilized in urea than those obtained at lower temperatures (20 and 30 °C) (Singh et al. 2020).

Granulocyte–macrophage colony-stimulating factor (GM-CSF) is a protein whose primary function is to stimulate germinal hematopoietic cells for the formation of the differentiated myeloid lineage and takes part in regulating a wide variety of inflammatory responses (Francisco-Cruz et al. 2014; Wicks and Roberts 2016; Hamilton 2019; Dougan et al. 2019). GM-CSF has been recombinantly expressed in mammalian cells, yeasts, and bacteria, receiving FDA (Food and Drug Administration) approval in 1991 for the treatment of neutropenia (Mehta et al. 2015; Dougan et al. 2019). Even though the human GM-CSF contains *O*- and *N*-glycosylation sites (Walter et al. 1992), the non-glycosylated form produced in *E. coli* is biologically active and has therapeutic relevance (Okamoto et al. 1991; Cumming 1991). Recent studies have elucidated the protective role of GM-CSF in autoimmune diseases such as pulmonary alveolar proteinosis (Trapnell et al. 2020; Zhang et al. 2020), its use in oncolytic immunotherapies, and adjuvant in cancer vaccines (Kaufman et al. 2014) as well as its possible administration in treatments against COVID-19 (Lang et al. 2020; Bonaventura et al. 2020).

Although some studies revealed the influence of induction temperature and induction time on IBs structure using a thermoinducible system (Caspeta et al. 2009, 2013; Restrepo-Pineda et al. 2019), this report aimed to elucidate the effect of growth (or pre-induction) temperature on the subsequent formation process, protein composition and structural characteristics of IBs from an *E. coli* bacterial culture producing the recombinant human GM-CSF (rHuGM-CSF). Here, bioreactor cultures of *E. coli* W3110 grown at either 30 or 34 °C were subsequently thermoinduced at 42 °C. Growth curves, carbon source consumption, and acetate production were analyzed. SDS-PAGE evaluated the expression kinetics of the rHuGM-CSF protein in the total protein fraction and IBs. Changes in the expression of some HSPs (DnaK/J, GroEL/ES) were identified by western blot. The structural analysis of the aggregates was carried out using attenuated total reflection-Fourier transform infrared spectroscopy (ATR-FTIR), binding to an amyloidogenic dye, proteolytic digestion, and denaturation with a chaotropic agent. Finally, the secondary structure of the refolded and purified rHuGM-CSF from IBs was studied by circular dichroism (CD) to elucidate its biological activity indirectly.

Materials and methods

Strain, media composition, and bioreactor conditions

E. coli W3110 (ATCC® 27325™) was used as a host to produce the recombinant human granulocyte–macrophage colony-stimulating factor (rHuGM-CSF). The coding

sequence for the rHuGM-CSF (GenBank accession number OL419360) was cloned in the pV3 plasmid containing the gene of the cI857 thermolabile repressor and the *pL* promoter from bacteriophage λ (Lowman and Bina 1990). pV3 is a low copy number plasmid based on the RepA-CopB from R100 plasmid (Olsson et al. 2004) and the *par* sequence to increase plasmid stability from plasmid pSD101 (Miller et al. 1983). A working cell bank containing aliquots of 1 ml at an optical density at 600 nm (OD_{600}) of 1.09 absorbance units (AU) was generated with 40% (v/v) glycerol and stored at -75 °C (Restrepo-Pineda et al. 2019).

For inoculums and bioreactor cultures, a defined culture medium was prepared as described in Restrepo-Pineda et al. (2019), as follows (in g/l): 4.0 (NH₄)₂HPO₄; 13.3 KH₂PO₄; 1.7 citric acid; 1.2 MgSO₄·7H₂O; 0.045 thiamine; 0.1 kanamycin; 17.5 glucose; 3.0 casamino acids and trace elements (2.0 ml/l of 500X stock). Glucose, MgSO₄, and trace elements stocks were separately sterilized at 121 °C and 22 *psig* for 30 min (ES-215 sterilizer, Tomy Digital Biology, Tokyo, Japan). Thiamine, casamino acids, and kanamycin solutions were sterilized using 0.22- μ m-pore-size filters (Merck Millipore, Billerica, MA, USA) and added before inoculation. The pH of the culture medium was adjusted to 7.0 \pm 0.1 with 3 N NaOH and 8 N HCl.

For preparing the inoculum, 500 μ l of the working cell bank was cultured in conventional 250-ml Erlenmeyer flasks with 50 ml filling volume. Cells were incubated overnight (~14 h) at 30 °C and 200 rpm (New Brunswick Scientific Classic C25, Enfield, CT, USA). The inoculum volume was decided based on the OD_{600} measurement to start the bioreactor experiments with the same cell concentration (~0.1 AU). The batch cultures were carried out in 1.2-l bioreactors (Applikon, Delft, Netherlands) with a working volume of 800 ml. Dissolved oxygen tension (DOT) was controlled at 35% with respect to air saturation by a cascade of agitation between 100–1000 rpm and constant airflow rate (1 vvm, volume of air per volume of culture medium). Medium pH in bioreactors was maintained at 7.0 \pm 0.1 by adding 3 N NaOH or 3 N HCl. The temperature was controlled to either 30 or 34 °C (42 °C for induction) using a heating/cooling circulating water bath (PolyScience, Niles, IL, USA). To avoid foaming, a sterile antifoam agent was added manually when necessary. pH, DOT, and temperature were monitored and controlled online with the BioXpert software (Applikon, Delft, Netherlands).

Cell concentration estimation

Growth of the strain *E. coli* W3110 producing rHuGM-CSF was determined by following the OD_{600} (Spectronic Genesys 5, Thermo Electron Corporation, Westmont, IL, USA). OD_{600} measurements were converted to dry cell weight (DCW) through a linear correlation standard curve.

Briefly, 10 ml samples from three independent cultures were centrifuged at $8,000 \times g$ for 10 min. The cell pellet was washed with 1X PBS (pH 7.5) twice and filtered using 0.22- μm -pore-size filters (Merck Millipore, Billerica, MA, USA). The wet cell paste was dried at 90°C for 48 h. After complete drying, the filters were weighed again. The difference in mass was used to calculate the DCW. 1.0 AU was equivalent to 0.33 ± 0.04 g/l of DCW.

Recombinant protein thermoinduction

Bioreactor cultures were grown at either 30 or 34°C until reaching the pre-stationary phase (OD_{600} of 2.0–3.0 AU). At this point, the thermoinduction of rHuGM-CSF production was carried out by increasing to 42°C , maintaining DOT and pH control (Restrepo-Pineda et al. 2019). The heating rate was the same in both conditions ($0.5^\circ\text{C}/\text{min}$). Samples of 1.0 ml were taken at different post-induction times (1, 3, 5, 10, and 18/20 h) and centrifuged at $10,000 \times g$ for 10 min. Supernatants were used for glucose and acetate estimation, while pellets were stored at -20°C for further analysis. All experiments were performed in triplicate.

Glucose and acetate quantification

Supernatants were filtered using sterile syringe filters with a 0.22- μm -pore-size before injection. Glucose concentration was determined in a biochemistry analyzer YSI 2900D (YSI Inc, Yellow Springs, OH, USA) equipped with a glucose oxidase membrane (YSI 2365), a buffer solution (YSI 237), and a standard calibrator solution (2.5 g/l of glucose). The concentrations of acetate were determined by high-performance liquid chromatography (HPLC) in a Shimadzu LC-20AT (Shimadzu, Kyoto, Japan) using an Aminex HPX-87H column (300 \times 7.8 mm; 9- μm internal diameter, Bio-Rad, Hercules, CA, USA). The mobile phase consisted of 0.008 N NH_2SO_4 with 0.6 ml/min at 50°C and 215 nm UV absorbance. A commercial standard solution was used for acetate (No. 125–0586, Bio-Rad), and data obtained were processed in the LC Solution software (Shimadzu, Kyoto, Japan).

Recovery and purification of IBs

For IBs isolation, the method described in Calcines-Cruz et al. (2018) was followed with minor modifications. Briefly, the cell biomass pellets from 1 ml culture were diluted in lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 7.5) containing protease inhibitor (0.1 mM PMSF, phenylmethylsulfonyl fluoride). Each sample was disrupted by sonication using a Soniprep150 (Sanyo Gallenkamp PLC, Loughborough, UK) at 8 μm amplitude in 3–10 cycles of 30 s, keeping on ice. The lysed cell mixture was centrifuged at $14,000 \times g$ for 15 min at 4°C , and both the supernatant

with the total soluble protein and the pellet with the insoluble protein (IBs) were recovered. Insoluble protein fraction was incubated in lysis buffer with 1% (v/v) IGEPAL (Sigma Aldrich, St. Louis, MO, USA) for 30 min under agitation at 4°C and centrifuged at $14,000 \times g$ for 5 min. The pellet was resuspended in lysis buffer containing 0.5% (v/v) Triton X-100 and centrifuged at $14,000 \times g$ for 15 min. Finally, the pellet was washed 3 to 5 times with deionized water, centrifuged between each wash ($14,000 \times g$ for 15 min), and the purified IBs were stored at -20°C .

Protein quantification, rHuGM-CSF identification, and chaperones immunodetection

The concentration of total soluble protein and protein in the IBs was measured by Bradford assay (Bio-Rad, Hercules, California, USA) according to supplier recommendations. Insoluble proteins were previously solubilized in isoelectric focusing (IEF) buffer (final 1:5 dilution) at room temperature for at least 3 h. Calibration curves with bovine serum albumin (BSA, Equitech-Bio, Kerrville, TX, USA) were made. Both samples and standards were prepared in triplicate, and OD_{600} was measured on a Stat Fax 2100 Microplate Reader (Awareness Technology Inc., Palm City, FL, USA).

Samples collected were used to analyze the production of the total soluble protein and rHuGM-CSF accumulation in IBs on 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). Previously, equal amounts of protein were solubilized using 2.5% SDS for 12 h at room temperature, and 20 μg of protein was loaded in each lane. Gels were stained with Coomassie Brilliant Blue R-250 (Bio-Rad, Hercules, CA, USA), and the percentage of rHuGM-CSF in IBs was determined by densitometry using the Image-Lab software and Gel Doc EZ Imager (Bio-Rad, Hercules, CA, USA).

The specific detection of chaperones was carried out by western blot, as reported by Restrepo-Pineda et al. (2019). Proteins in the polyacrylamide gels were transferred to a polyvinylidene difluoride (PVDF) membrane (Immobilon, Millipore, Bedford, MA, USA) by a semi-wet approach via Trans-Blot system (Bio-Rad, Hercules, California, USA). Membranes were blocked for 40 min with 1X TBS plus 5% skim-milk at room temperature and gentle shaking. Two washes for 10 min each were made with wash buffer (1X TBS plus 0.06% Tween-20), and primary antibodies (mouse DnaK ADI-SPA-880 dilution 1:7500, rabbit GroEL ADI-SPA-875 dilution 1:7500, rabbit DnaJ ADI-SPA-410 dilution 1:2000, and rabbit GroES ADI-SPA-210 dilution 1:7500; Enzo Life Sciences, Farmingdale, NY, USA) were added for 1 h at room temperature. Later, membranes were washed three times, followed by incubation with the corresponding secondary antibodies (goat anti-mouse IgG A9044 dilution 1:2500 and goat anti-rabbit IgG A0545

dilution 1:2500, HRP-conjugated, Merck-Sigma-Aldrich, St. Louis, MI, USA) for 1 h at room temperature. Three more washes were done, and the proteins were detected by chemiluminescence using SuperSignal West Pico and Femto substrates (Thermo Fisher Scientific, Waltham, MA, USA) in a C-DIGIT blot scanner (LI-COR, Lincoln, NE, USA). A homemade chemiluminescence marker was used to reveal the images, and the membranes were incubated with mild stripping buffer (glycine, SDS, Tween-20, pH 2.2) to remove the antibodies.

ATR-FTIR spectroscopy

The components of the secondary structure of IBs were identified by attenuated total reflection (ATR)–Fourier transform infrared (FTIR) spectroscopy (Valdez-Cruz et al. 2017; Calcines-Cruz et al. 2018; Singh et al. 2020; Gil-Garcia et al. 2020). Freshly purified IBs were dried at room temperature for 1 h using a speed vacuum concentrator and placed on Specac Quest ATR diamond accessory (Specac Limited, Slough, UK) coupled to an infrared spectrometer IRAffinity-1S (Shimadzu, Kyoto, Japan). Each sample consisted of 40 acquisitions with a resolution of 2 cm^{-1} in the range of $1500\text{--}1700\text{ cm}^{-1}$ and subsequently averaged. After 13 points of smoothing, the second derivatives of the spectrum of the amide I region were determined with the IR LabSolutions program (Shimadzu, Kyoto, Japan) and normalized with respect to the absolute value of the tyrosine peak ($\sim 1508\text{ cm}^{-1}$) (Ami et al. 2005). The analysis of the location frequencies indicated the abundance of the different elements of the secondary structure in the IBs.

Thioflavin T binding assay

The amyloidogenic properties of IBs were evaluated by measuring the fluorescence after binding to the thioflavin T (Th-T) dye (Castellanos-Mendoza et al. 2014; Calcines-Cruz et al. 2018; Singh et al. 2020; Gil-Garcia et al. 2020). 50 mg/ml of protein in IBs were resuspended in phosphate buffer (pH 7.5) containing $75\text{ }\mu\text{M}$ of Th-T (Sigma Aldrich, St. Louis, MO, USA) and incubated for 1 h at $25\text{ }^\circ\text{C}$. The fluorescence signal was measured in a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Palo Alto, CA, USA) with an excitation wavelength of 440 nm and 5 nm bandwidth. The emission spectrum was recorded from 450 to 560 nm with a spectral resolution of 5 nm. Each spectrum was acquired five times, averaged, and smoothed. The spectrum of Th-T without protein was obtained as a control.

Proteinase K digestion

The resistance of the IBs to enzymatic degradation was determined by incubating 50 mg/ml of sample with $25\text{ }\mu\text{g/ml}$

of proteinase K (PK, Sigma Aldrich, St. Louis, MO, USA). The proteolytic digestion was carried out in 1 ml of buffer (50 mM Tris–HCl and 150 mM NaCl, pH 8.0) at room temperature. Changes in absorbance were monitored at 350 nm for 100 min in a UV/Vis DU@730 spectrophotometer (Beckman Coulter Inc., Brea, CA, USA) mixing by pipetting every minute. Data were normalized with respect to the initial absorbance value (Upadhyay et al. 2012; Castellanos-Mendoza et al. 2014; Valdez-Cruz et al. 2017; Calcines-Cruz et al. 2018; Singh et al. 2020).

Stability in guanidinium chloride

The stability of IBs against chemical solubilization was examined by adding 1.0 mg/ml of protein in IBs to $100\text{ }\mu\text{l}$ of 10 mM Tris–HCl buffer (pH 7.5) containing different concentrations (0, 1, 3 and 5 M) of guanidinium hydrochloride (GndHCl, Sigma-Aldrich, St. Louis, MO, USA). After 24 h of gentle shaking at room temperature, the samples were centrifuged at $8000\times g$ for 10 min. The supernatants were recovered, and the solubilized protein was quantified by the Bradford method described above. The solubility profiles of the IBs harvested at different times after induction were obtained by plotting the soluble protein concentration against GndHCl concentration (De Groot and Ventura 2006; Espargaró et al. 2008; Castellanos-Mendoza et al. 2014).

IBs solubilization, refolding, and purification of rhGM-CSF

Purified IBs of rHuGM-CSF were resuspended in a solubilization buffer (6 M GndHCl, 50 mM 2-mercaptoethanol in 100 mM Tris base, pH 8.2, adjusted with 3 M HCl) at 700 rpm , 1 h and $25\text{ }^\circ\text{C}$ to a final protein concentration of 1.0 mg/ml . Then, the protein folding process was carried out at $25\text{ }^\circ\text{C}$, 100 rpm , and 4 h by drop-by-drop dilution (1:10) of the solubilized product in a Tris base buffer (20 mM , pH 8.2). The reaction was stopped by adding 4 M acetic acid to a final concentration of 0.2 M (Burgess et al. 1987; Belew et al. 1994; Thomson et al. 2012).

The folded protein solution was purified by reverse-phase HPLC in a Shimadzu LC-20AT (Shimadzu, Kyoto, Japan) using a Zorbax Eclipse XDB-C8 column (Agilent Technologies, Santa Clara, CA, USA). A standard curve of the European Pharmacopoeia reference standard for human GM-CSF (Molgramostim, Y0000251, Sigma Aldrich, St. Louis, MO, USA) was carried out at 1.345 mg/ml , 0.672 mg/ml , and 0.336 mg/ml . The folded protein sample ($\sim 10\text{ }\mu\text{g}$) was loaded in the HPLC at $50\text{ }^\circ\text{C}$ with a detector wavelength of 214 nm and a maximum column pressure of 3000 psig . Solutions of 0.1% v/v trifluoroacetic acid (TFA) in water (mobile phase A) and 0.1% v/v TFA in 90% acetonitrile (mobile

phase B) were used for the linear gradient elution at a flow rate of 1.2 ml/min (Nicola et al. 1983; Das et al. 2011).

Circular dichroism (CD) spectroscopy

CD spectra of purified rHuGM-CSF in water were recorded at 37 °C in the far-UV region with a JASCO J-720 spectropolarimeter (Jasco Inc., Easton, MD) described elsewhere (Luviano et al. 2019). Protein solutions of ~0.05 mg/ml were loaded into a quartz cell of 0.1-cm length path. Each spectrum corresponded to the average of three repetitive scans and was corrected by the buffer signal. Ellipticities are reported as mean residue ellipticity, $[\theta]_{\text{mrw}}$. Secondary structure content was calculated from CD spectra using the deconvolution software K2D3 (Louis-Jeune et al. 2012).

Statistical analysis

Statistically significant differences between the data were calculated using analysis of variance (ANOVA), followed by Tukey's test. The quantitative results are expressed as mean \pm standard error of the mean. A p -value < 0.05 was considered statistically significant.

Results

Pre-induction temperature affects the specific growth rate (μ) in thermoinduced *E. coli* cultures.

Cultures of *E. coli* W3110 were grown at either 30 or 34 °C until reaching the pre-stationary phase of growth (OD_{600} nm of 2.0–3.0 AU). Thermoinduction (at 42 °C with a heating rate of 0.50 °C/min) was done in the pre-stationary phase to have a large number of viable cells and take advantage of the fact that the stress responses associated with entering the stationary phase, due to nutrients limitation or the presence of secondary metabolites, has not been triggered (Hengge-Aronis, 1993; Overton, 2014). The dry cell weight (DCW) determination indicated that 1.0 AU was equivalent to 0.33 ± 0.04 g/l of DCW, similar to that reported in other works: 0.32 g/l for recombinant *E. coli* W3110 (Sandoval-Basurto et al. 2004); 0.31 ± 0.05 g/l for recombinant *E. coli* 53,606 (Restrepo-Pineda et al. 2019) and 0.28 g/l for a recombinant strain derived from *E. coli* K-12 (Mansey et al. 2014). Figure 1 compares the kinetics of biomass, the consumption of the carbon source, and the production of acetate. The maximum biomass (X_{max}) for the cultures grown at 30 °C \rightarrow 42 °C was 3.48 ± 0.19 g/l at 21 h (Fig. 1A, Table 1) and for those grown at 34 °C \rightarrow 42 °C was 3.48 ± 0.31 g/l at 15 h (Fig. 1B, Table 1).

A significant difference in the specific growth rate (μ) before thermoinduction is observed between cultures of

30 °C \rightarrow 42 °C (0.53 ± 0.01 h⁻¹) and cultures of 34 °C \rightarrow 42 °C (0.90 ± 0.07 h⁻¹) as shown in Table 1. This represents an increase of ~69% in the μ of recombinant *E. coli* growing at a higher temperature (34 °C). Glucose was consumed entirely in both cases, reaching values close to zero after 20 h of culture (Fig. 1C, D). Interestingly, no significant differences were observed in the biomass per glucose yield ($Y_{X/S}$) and the acetate per biomass yield ($Y_{AC/X}$) between cultures with different pre-induction temperatures (30 and 34 °C; Table 1). However, cultures growing at 34 °C \rightarrow 42 °C consumed the carbon source faster, yielding a specific glucose consumption rate (q_s) of 3.74 ± 0.34 g/g·h, which is 1.5 higher than for cultures grown at 30 °C \rightarrow 42 °C with q_s of 2.64 ± 0.20 g/g·h (Table 1). Similarly, the specific acetate production rate (q_p) was ~65% higher in the cultures at 34 °C \rightarrow 42 °C than at 30 °C \rightarrow 42 °C (Table 1). Acetate reached similar maximum concentrations of 6.58 ± 0.39 g/l at 30 °C \rightarrow 42 °C and 6.38 ± 0.63 g/l at 34 °C \rightarrow 42 °C after 17 h of culture (Fig. 1E, F).

The DOT was controlled in the bioreactors through a proportional–integral–derivative (PID) control algorithm (Trujillo-Roldán et al. 2001). DOT oscillated around the setpoint of 35%, confirming no oxygen limitation in the cultures (Supplemental Fig. S1A, B). Likewise, the pH of the medium was kept close to 7.0 ± 0.1 using an automatic addition system of 3 N NaOH or 3 N HCl (Supplemental Fig. S1C, D).

rHuGM-CSF is preferentially accumulated in IBs using a thermoinducible system, and the amount of recombinant protein within IBs increased to 34 °C \rightarrow 42 °C

To determine the amount of rHuGM-CSF produced after thermoinduction, fractions of total protein obtained from cultures growing at either 30 °C \rightarrow 42 °C or 34 °C \rightarrow 42 °C were analyzed on 15% SDS-PAGE (Supplemental Fig. S2). Samples of different post-induction times (1, 3, 5, and 18/20 h) were loaded in gels, and a sample before thermoinduction (0 h) was used as a negative control. A band corresponding to the molecular weight of the rHuGM-CSF protein (~14–15 kDa) was seen after thermoinduction for both conditions (Supplemental Fig. S2). The densitometric analysis of the gel 1 h after thermoinduction revealed that rHuGM-CSF corresponds to ~22% of the total protein for the 34 °C \rightarrow 42 °C cultures (Supplemental Fig. S2B), while for the 30 °C \rightarrow 42 °C cultures, rHuGM-CSF band represented ~11.5% of the total protein (Supplemental Fig. S2A). The amounts of recombinant protein increased with the induction time, reaching at the end of the culture percentages of ~34% of rHuGM-CSF at 30 °C \rightarrow 42 °C and ~30% of rHuGM-CSF at 34 °C \rightarrow 42 °C (Supplemental Fig. S2).

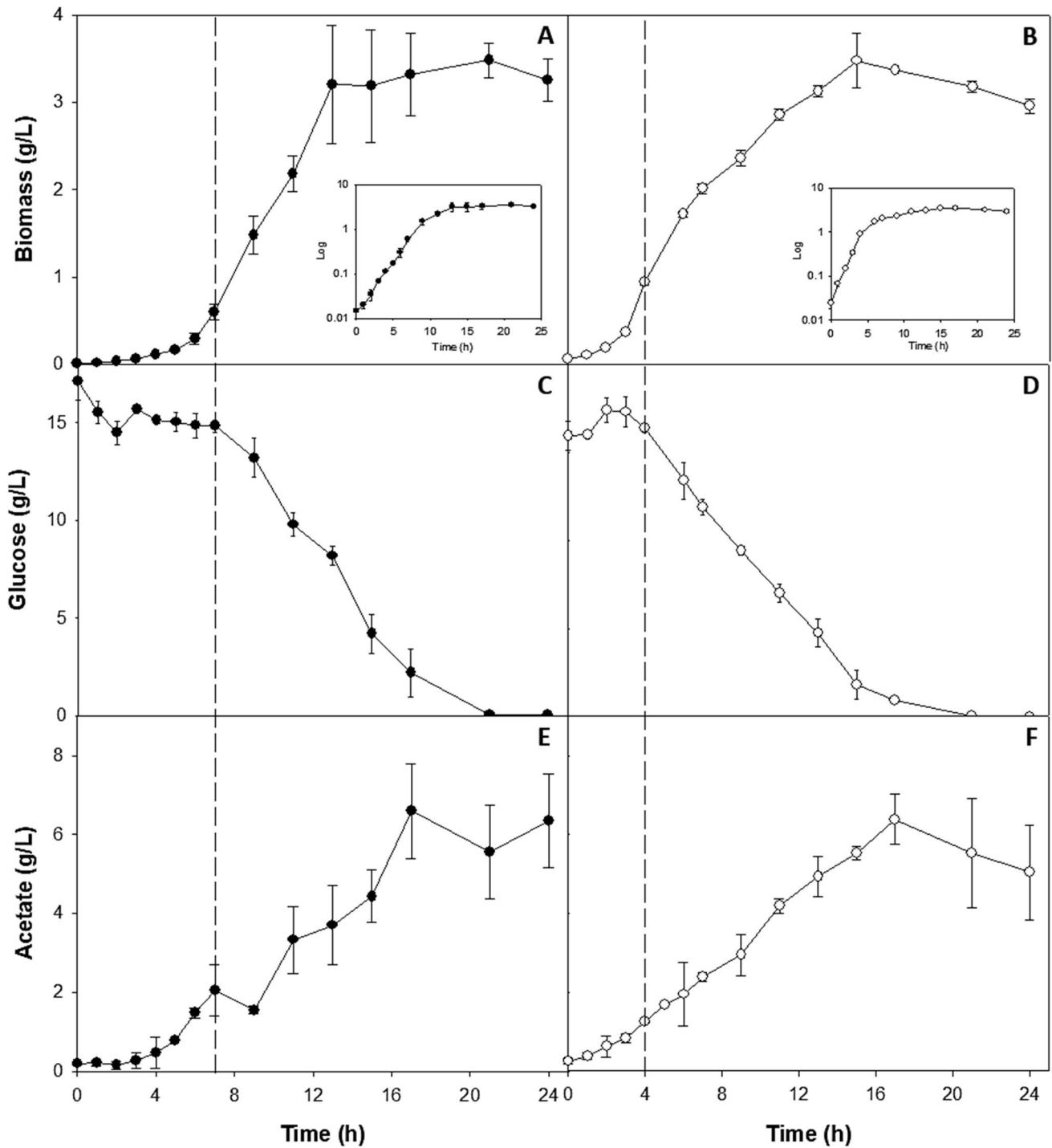


Fig. 1 Kinetics of bacterial growth (A, B), glucose consumption (C, D), and acetate production (E, F) of *E. coli* W3110 producing the rHuGM-CSF protein in 1.2-l bioreactors. Two pre-induction temperatures were evaluated: 30 °C (filled circles) and 34 °C (open circles) with subsequent thermoinduction at 42 °C. Vertical dotted lines indi-

cate the start of the temperature increase (7 h for 30 °C and 4 h for 34 °C), which correspond to an OD_{600} of ~2.0 AU. The graph presents the mean with their respective standard deviation of three independent experiments

Similar final concentrations of 0.34 ± 0.10 g/l at 30 °C → 42 °C and 0.26 ± 0.04 g/l at 34 °C → 42 °C of rHuGM-CSF were obtained (Table 1). Our results agree with previous

reports, where recombinant protein yields close to 30% of the total protein were achieved in a thermoinducible expression system (Remaut et al. 1981; Valdez-Cruz et al. 2010).

Table 1 Comparison of the kinetic parameters of *E. coli* W3110 bioreactor cultures growing at two different temperatures: 30 °C and 34 °C with subsequent thermoinduction of rHuGM-CSF at 42 °C. The mean and standard deviation of three biological replicates per condition are presented

	Growth at 30 °C Induction at 42 °C		Growth at 34 °C Induction at 42 °C	
	Before induction	After induction	Before induction	After induction
^A μ (h ⁻¹)	0.53 ± 0.01 ^a	0.14 ± 0.03 ^b	0.90 ± 0.07 ^c	0.11 ± 0.01 ^b
td (h)		1.30 ± 0.03 ^a		0.78 ± 0.08 ^b
^B X_{\max} (g _{DCW} /L)		3.48 ± 0.19 ^a		3.48 ± 0.31 ^a
^C $Y_{X/S}$ (g _{DCW} /g _{GLC})		0.20 ± 0.02 ^a		0.25 ± 0.03 ^a
^D q_s (g _{GLC} /g _{DCW} ·h)		2.64 ± 0.20 ^a		3.74 ± 0.34 ^b
^E $Y_{AC/X}$ (g _{AC} /g _{DCW})		1.61 ± 0.76 ^a		1.56 ± 0.20 ^a
^F q_p (g _{AC} /g _{DCW} ·h)		0.86 ± 0.41 ^a		1.41 ± 0.20 ^b
Total protein, TP (g/l)		1.01 ± 0.10 ^a		0.85 ± 0.03 ^b
^G rHuGM-CSF in TP (%)		34 ± 3 ^a		30 ± 4 ^a
rHuGM-CSF in TP (g/l)		0.34 ± 0.10 ^a		0.26 ± 0.04 ^a
^H $Y_{TP/X}$ (g _{TP} /g _{DCW})		0.29 ± 0.08 ^a		0.24 ± 0.05 ^a
^I rHuGM-CSF in IBs (%)		45 ± 2 ^a		53 ± 3 ^b
^J $Y_{RP/X}$ (g _{RP} /g _{DCW})		0.06 ± 0.03 ^a		0.12 ± 0.02 ^b

Abbreviations: μ , specific growth rate; td: doubling time; X_{\max} , maximum biomass concentration; DCW, dry cell weight; GLC, glucose; AC, acetate; $Y_{X/S}$, biomass per substrate yield; $Y_{AC/X}$, acetate per biomass yield; q_s , specific glucose consumption rate; q_p , specific acetate formation rate; TP, total protein; IBs, inclusion bodies; RP, recombinant protein; $Y_{TP/X}$, total protein per biomass yield; $Y_{RP/X}$, rHuGM-CSF per biomass yield; rHuGM-CSF, recombinant human granulocyte–macrophage colony-stimulating factor

Data are presented as mean ± standard deviation

A non-statistically significant test result ($P > .05$) is represented with the same letter, and a statistically significant test result ($P < .05$) is represented with a different note

A: μ after induction was calculated from cell growth just after thermoinduction until the beginning of steady state

B: X_{\max} was reached at 21 h in cultures grown at 30 °C and 15 h in cultures grown at 34 °C

C, E: $Y_{X/S}$ and $Y_{AC/X}$ were calculated using the glucose and acetate concentrations at X_{\max}

D, F: q_s and q_p were calculated using the μ before induction and the yields obtained at X_{\max}

G, I: $Y_{TP/X}$ and $Y_{RP/X}$ were calculated using maximum values of protein concentrations

H, J: Percentage of rHuGM-CSF in TP and IBs was based on the densitometric analysis from bands identified in SDS-PAGE gels

Afterward, IBs from *E. coli* cultures growing at 30 or 34 °C with induction at 42 °C were purified by multiple washing steps. The band corresponding to rHuGM-CSF was visible only in the IBs after the upshift to 42 °C, regardless of the growth temperature, but not in the soluble protein fraction (Supplemental Fig. S3). Subsequently, to determine the rHuGM-CSF yields in the insoluble fraction, 20 μ g of IBs obtained at different post-induction times (1, 3, 5, 10, and 18/20 h) from *E. coli* cultures that grew at 30 or 34 °C were analyzed on 15% SDS-PAGE (Fig. 2). The European Pharmacopoeia reference standard for human GM-CSF (Molgramostim) was used as a positive control (Molgra St. in Fig. 2). The accumulation of rHuGM-CSF in IBs showed differences depending on the growth (pre-induction) temperature (Fig. 2). Densitometric analysis of the rHuGM-CSF protein in IBs indicated the content of $\sim 45 \pm 2\%$ and $\sim 53 \pm 3\%$ of recombinant protein in thermoinduced cultures, from 5 h to the end of the culture, that grew at 30 and 34 °C, respectively (Table 1). Although the total protein concentration at the end

of cultures was higher in those that follow 30 °C \rightarrow 42 °C, (Fig. 3A), the amount of protein (Fig. 3B) and rHuGM-CSF within the IBs was higher in those 34 °C \rightarrow 42 °C (Table 1). The rHuGM-CSF per biomass yield ($Y_{RP/X}$) was two times higher for cultures of 34 °C \rightarrow 42 °C (0.12 ± 0.02 g/g) than for those of 30 °C \rightarrow 42 °C (0.06 ± 0.03 g/g), demonstrating that the bacterial growth at 34 °C favored the accumulation of rHuGM-CSF in IBs during thermoinduction (Table 1).

Main folding chaperones (DnaK and GroEL) are associated with IBs and their co-chaperones (DnaJ and GroES) to the soluble protein fraction during thermoinduction

Here, immunodetection of the main folding chaperones (DnaK and GroEL) and their co-chaperones (DnaJ and GroES) was carried out (Fig. 4). As positive controls, total protein lysates of the recombinant *E. coli* W3110 growing at 30 or 34 °C without thermoinduction and whole protein

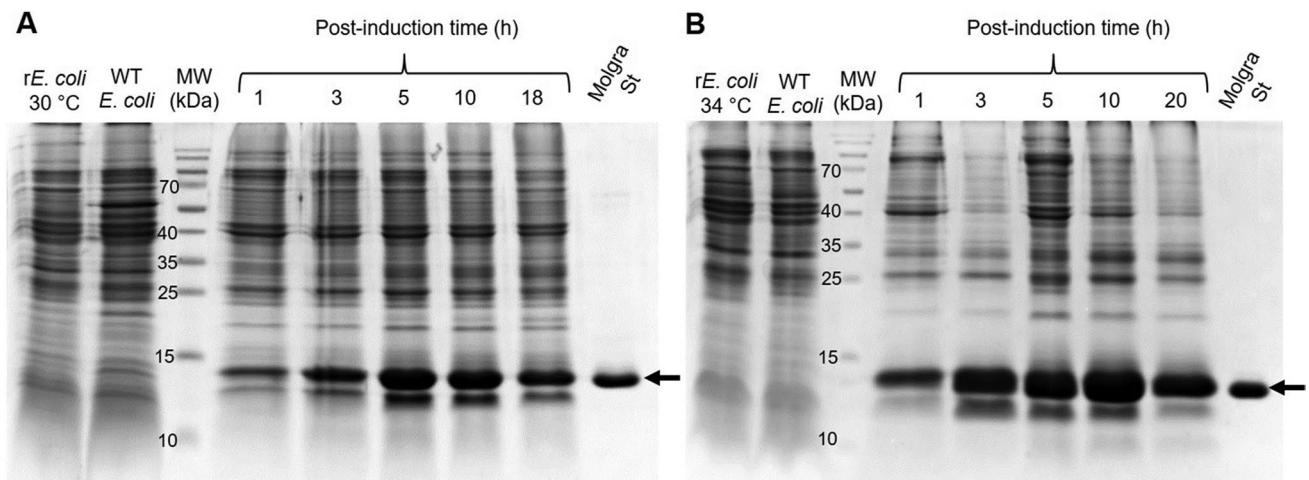
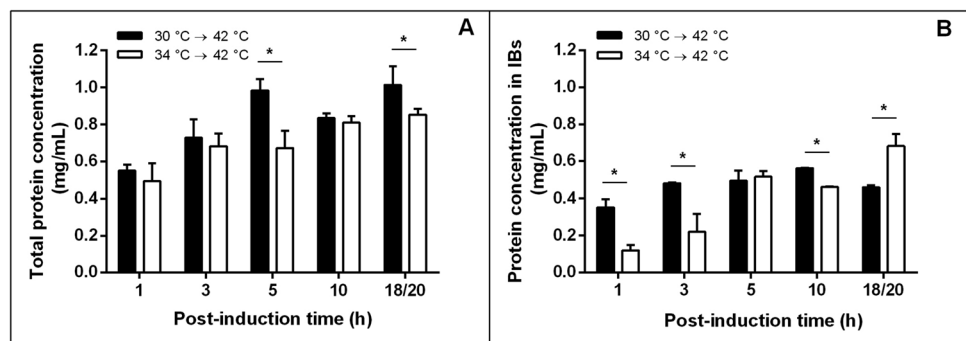


Fig. 2 Analysis of protein in IBs by 15% SDS-PAGE gel stained with Coomassie blue. Purified IBs from *E. coli* W3110 bioreactor cultures growing at 30 °C (A) or 34 °C (B) with subsequent rHuGM-CSF thermoinduction at 42 °C are presented. Lane r*E. coli* 30 °C and lane r*E. coli* 34 °C: total protein of the recombinant *E. coli* W3110 without thermoinduction growing at 30 or 34 °C, respectively; Lane WT

E. coli: total protein from wild-type *E. coli* W3110 strain. Lane MW: molecular weight marker. IBs from different post-induction times (1, 3, 5, 10, and 18/20 h) to the two conditions evaluated are shown. Molgra St: Molgramostim reference standard (2 µg). Arrows indicate the band corresponding to the rHuGM-CSF protein (~14 kDa)

Fig. 3 The concentration of total protein (A) and protein in IBs (B) of *E. coli* W3110 cultures under different pre-induction temperatures: 30 °C (black bars) or 34 °C (white bars) and subsequent rHu-GM-CSF thermoinduction at 42 °C. The mean and standard deviation for three biological replicates per condition are shown



lysates from wild-type *E. coli* W3110 were used. In both control fractions, the HSPs mentioned were identified, as expected (Lanes 1 and 2, Fig. 4). DnaK chaperone with an approximate molecular weight of 70 kDa (Bardwell and Craig, 1984) was found weakly expressed 1 and 3 h after thermoinduction, but the intensity of the band increased at 5 h after thermoinduction and returned to a baseline level at the end of cultivation, both in IBs from *E. coli* that grew at 30 °C → 42 °C (Fig. 4A, Panel 1) and 34 °C → 42 °C (Fig. 4B, Panel 1). However, at 34 °C → 42 °C the band for DnaK (Fig. 4B, Panel 1) was noticeably stronger than at 30 °C → 42 °C (Fig. 4A, Panel 1). In the case of the GroEL chaperone, a band close to 60 kDa was observed in the IBs obtained either at a growth temperature of 30 °C (Fig. 4A, Panel 3) or 34 °C (Fig. 4B, Panel 3), being of similar intensity during all time after thermoinduction at 42 °C.

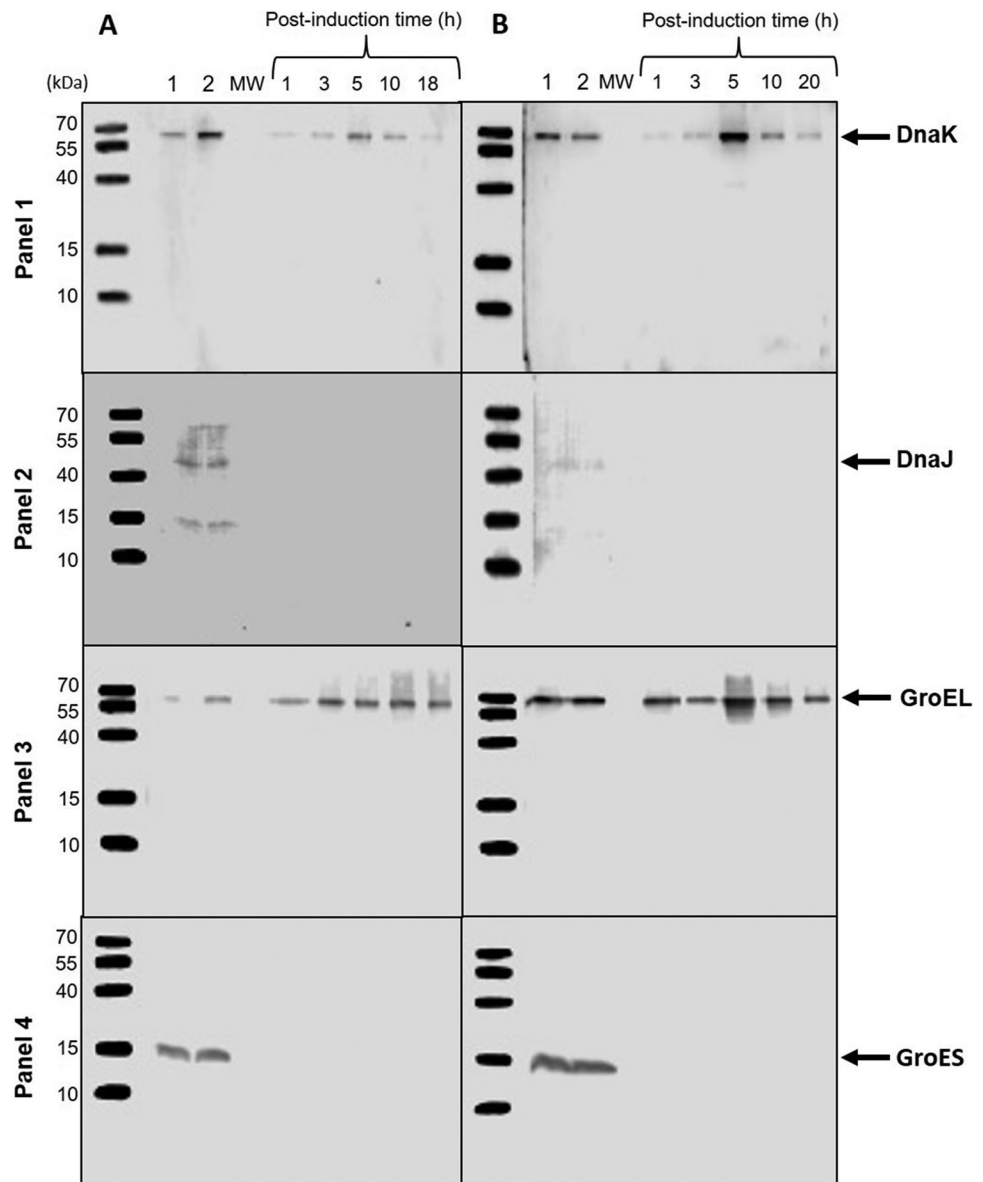
The immunodetection of the co-chaperones DnaJ with ~41 kDa (Fig. 4, Panel 2) and GroES with ~15 kDa (Fig. 4, Panel 4) was performed, and the bands corresponding

to these two proteins were only observed in lanes 1 and 2, which belong to the positive controls. That is, DnaJ and GroES are not associated with IBs under either of the two pre-induction temperatures tested but remain in the soluble fractions of the non-induced recombinant and the wild-type strains. Finally, GroES and DnaJ co-chaperones were immunodetected in the total soluble protein from *E. coli* cultures growing at 30 °C and 34 °C with thermoinduction at 42 °C (Supplemental Fig. S4).

The pre-induction temperature influenced the content of amyloid-like structure in IBs in a thermoinducible system

The effect of pre-induction temperature (30 or 34 °C) on the amyloid content of rHuGM-CSF IBs obtained under thermoinduction at 42 °C (1 h, 3 h, 5 h, 10 h, and 18/20 h) was analyzed by ATR-FTIR (Fig. 5). ATR-FTIR is a sensitive technique to determine the secondary structure of proteins

Fig. 4 Immunodetection of DnaK, DnaJ, GroEL, and GroES chaperones in rHuGM-CSF IBs from cultures under different pre-induction temperatures: 30 °C (A) or 34 °C (B) and subsequent thermoinduction at 42 °C. Lanes 1A and 1B: total protein lysates of the recombinant *E. coli* W3110 without thermoinduction growing at 30 or 34 °C, respectively. Lanes 2A and 2B: whole protein lysate from wild-type *E. coli* W3110 strain. Lane MW: molecular weight marker. IBs from different post-induction times (1, 3, 5, 10, and 18/20 h) to the two conditions evaluated are shown. Arrows indicate the bands corresponding to DnaK (~70 kDa, panel 1), DnaJ (~41 kDa, panel 2), GroEL (~60 kDa, panel 3), and GroES (~15 kDa, panel 4). The amount of protein in IBs (20 µg) loaded on the SDS-PAGE gels was used as a loading control for western blotting



and studying aggregates formation (Miller et al. 2013). In particular, the absorbance spectra were obtained in the amide I region (1,700–1,500 cm^{-1}), and second derivatives were used to identify the major bands and assign them to the protein secondary structure components (Fig. 5A, B). The major band at 1654 cm^{-1} was assigned as α -helices/random coil, while the bands at 1636 cm^{-1} and 1625 cm^{-1} were designated as β -sheets in native structure (β -sheets) and intermolecular β -structures related to amyloid conformation (aggregates), respectively (Ami et al. 2006; Li et al. 2019; Singh et al. 2020). The rHuGM-CSF IBs from cultures growing at 30 °C \rightarrow 42 °C (Fig. 5A) exhibited a higher spectral intensity of the band at 1625 cm^{-1} compared to IBs from cultures growing at 34 °C \rightarrow 42 °C (Fig. 5B), indicating a decrease in the content of amyloid aggregates when the pre-induction temperature is increased. The content of

α -helices and β -sheets in the IBs did not differ significantly for both pre-induction temperatures. Furthermore, no clear differences were observed between the ATR-FTIR spectra of the IBs concerning the post-induction time (Fig. 5A, B). The minima of the second derivative of each structural component allow to observe the differences between the secondary structure content under the evaluated conditions (Fig. 5C, D, E). From the first hour after thermoinduction, the content of amyloid aggregates was higher for IBs from cultures grown at 30 °C than at 34 °C (Fig. 5E). Moreover, the content of α -helices (Fig. 5C) and β -sheets (Fig. 5D) was similar in rHuGM-CSF IBs, regardless of the pre-induction temperature and post-induction time.

Alternatively, amyloid-diagnostic dyes are used to determine amyloid-like structure in IBs (Carrió et al. 2005; De Groot et al. 2009; Singh et al. 2020). Thioflavin T (Th-T)

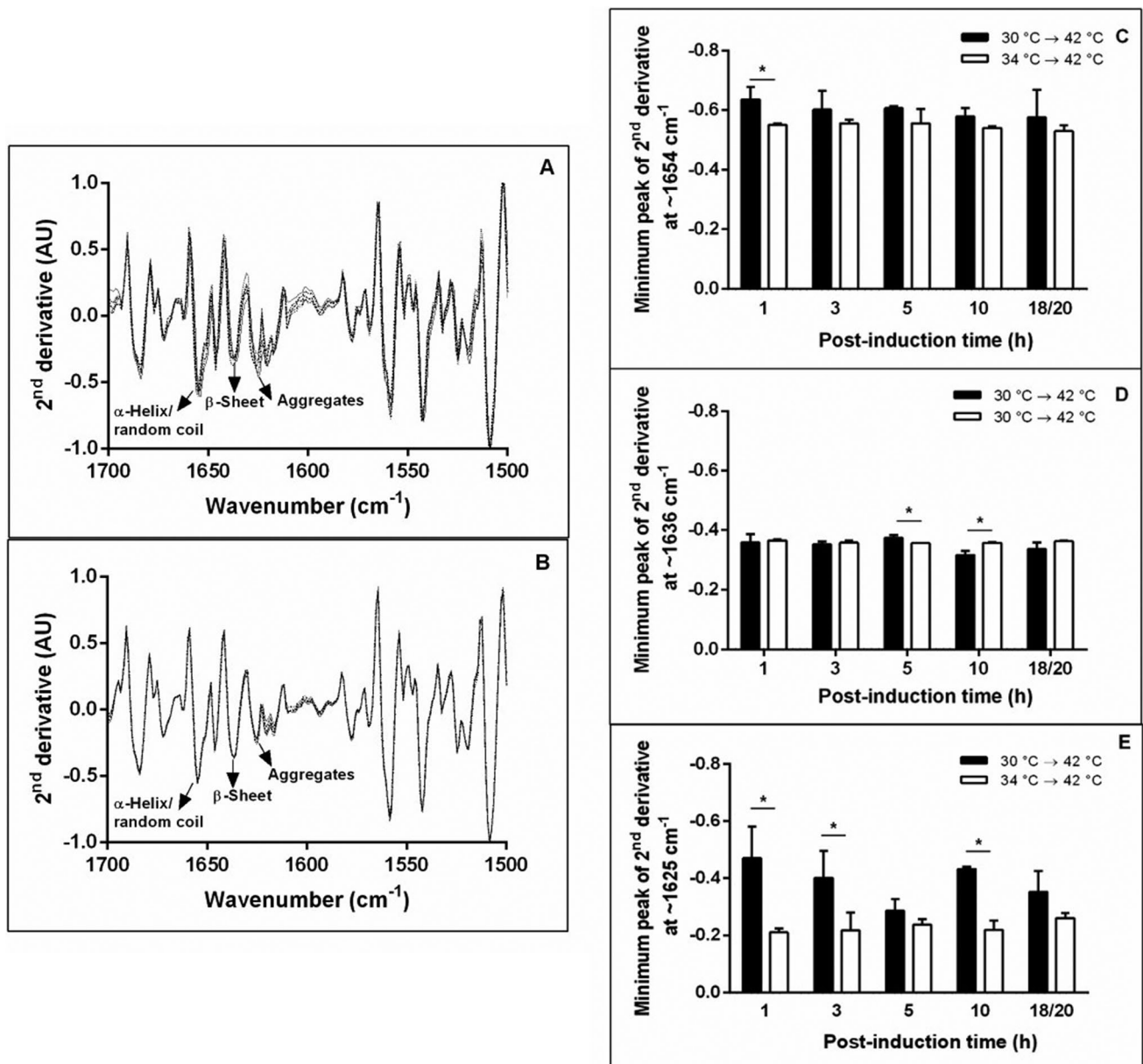


Fig. 5 Amyloid content in rHuGM-CSF IBs by ATR-FTIR. Second derivatives of the absorbance spectra for IBs from cultures growing at 30 °C (A) or 34 °C (B) and harvested at different times after thermoinduction at 42 °C: 1 h (solid line), 3 h (dashed line), 5 h (dotted line), 10 h (dashed-dotted line) and 18/20 h (dashed double-dotted line). Data were normalized with respect to the tyrosine peak ($\sim 1508\text{ cm}^{-1}$), and the major bands were used to identify and assign structural components as α helix/random coil ($\sim 1654\text{ cm}^{-1}$), β sheets

($\sim 1636\text{ cm}^{-1}$), and amyloid aggregates ($\sim 1625\text{ cm}^{-1}$). Spectra represent the average of three biological replicas. Comparison of the second derivatives minima corresponding to α helix/random coil (C), β sheets (D), and amyloid aggregates (E) of IBs from cultures at 30 °C (black bars) or 34 °C (white bars) collected at different times after thermoinduction at 42 °C. The mean and standard deviation are shown for three independent experiments

is a specific marker to study the amyloid conformation in aggregates since it binds to the surface of channels formed by cross-linked β -sheets (Krebs et al. 2005; Wu et al. 2009). A higher fluorescence indicates a higher amyloid content (LeVine 1995; Castellanos-Mendoza et al. 2014). The change in the fluorescence spectra of Th-T was evaluated after incubation with the rHuGM-CSF IBs (Fig. 6). The

maximum emission fluorescence was around 485 nm, both for IBs from cultures 30 °C \rightarrow 42 °C and 34 °C \rightarrow 42 °C, which is a typical feature of amyloid aggregates (Singh et al. 2020). IBs collected from cultures grown at 30 °C \rightarrow 42 °C showed a gradual increase in fluorescence intensity over time, that is, the amount of amyloid structure was greater at 18 h post-induction (Fig. 6A). In contrast, the fluorescence

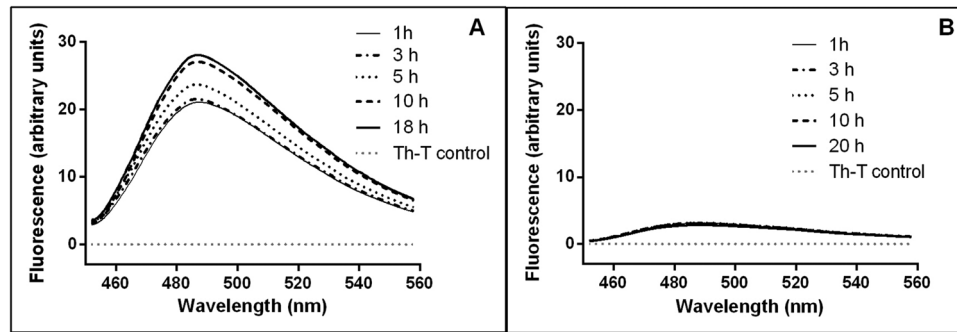


Fig. 6 Fluorescence emission spectra of Th-T binding to rHuGM-CSF IBs obtained under pre-induction temperatures of 30 °C (**A**) or 34 °C (**B**) and thermoinduction at 42 °C. IBs were harvested at different times post-induction: 1 h (thin solid line), 3 h (dashed-dotted

line), 5 h (dotted line), 10 h (dashed line), and 18/20 h (thick solid line). The spectrum of Th-T without protein was used as a control (dotted gray line), and the mean of two biological replicates per condition is shown

signal of Th-T was minimal for the IBs produced in cultures carried out at 34 °C → 42 °C (Fig. 6B), suggesting either a lack of amyloid structures in these aggregates or the inability of binding to them due to greater compaction or a small size (Carrió et al. 2005).

Pre-induction temperature impacts the resistance to proteolytic degradation and solubilization of IBs.

Evaluation of the IBs resistance to proteolytic degradation with proteinase K (PK) has been helpful to characterize the molecular organization and stability of aggregates, as well as an indication of their amyloid content (Castellanos-Mendoza et al. 2014; Calcines-Cruz et al. 2018; Restrepo-Pineda et al. 2019; Singh et al. 2020). PK is a serine protease that exhibits low activity on regions structurally dominated by β -sheets, typical of amyloid fibrils, but it is highly active in hydrophilic domains enriched by loops and α -helices (De Groot et al. 2009; Vázquez-Fernández et al. 2012; Macedo et al. 2015). The enzymatic activity of PK on the IBs was monitored for 100 min at 350 nm (Fig. 7). IBs from *E. coli* cultures of 34 °C → 42 °C were more susceptible to proteolytic attack by PK (Fig. 7B) than IBs from cultures of 30 °C → 42 °C (Fig. 7A). In the first hours after thermoinduction, IBs from 30 °C → 42 °C appeared to be more susceptible to PK digestion than IBs collected during the final hours (Fig. 7A); the same behavior was observed at 34 °C → 42 °C, where IBs obtained at 1, 3 and 5 h post-induction showed less resistance to PK activity than IBs harvested at 10 and 20 h (Fig. 7B). This means that the pre-induction temperature and the post-induction time affect the structural arrangement of the protein aggregates and, at the same time, their resistance to enzymatic digestion.

Solubilization of rHuGM-CSF IBs against increasing concentrations of guanidinium chloride was determined.

GndHCl is a strong chaotropic agent whose ionic nature causes denaturation of globular proteins and provides an estimate of the conformational stability of IBs (Monera et al. 1994; Del Vecchio et al. 2002; Castellanos-Mendoza et al. 2014). In Fig. 8, the solubilization profiles of rHuGM-CSF IBs exhibit a similar trend under the two evaluated pre-induction temperatures (30 and 34 °C). Protein aggregates obtained after thermoinduction were sensitive to chemical solubilization. At 1 M GndHCl, the amount of solubilized protein was small, but it increased at higher concentrations of GndHCl (3 and 5 M) (Fig. 8). Statistical analysis indicated that the concentrations of solubilized protein did not present significant differences with respect to the post-induction time (Fig. 8).

After IBs solubilization in GndHCl (6 M) and folding process, rHuGM-CSF was purified by HPLC showing a retention time of ~24.5 min, in both, 30 °C → 42 °C and 34 °C → 42 °C (Fig. 9). A similar retention time was observed when the European Pharmacopoeia reference standard for human GM-CSF was injected in the HPLC (inset of Fig. 9). All the small peaks observed before and after the rHuGM-CSF peak are surely the host cell proteins from the IBs. Finally, the refolding ability of the solubilized rHuGM-CSF was assayed by CD spectroscopy. In the far-UV region, the asymmetric environments of peptide bonds yield CD signals that are characteristic of each different structural element, allowing the secondary structure of a protein to be estimated. As shown in Fig. 10, protein samples obtained from both thermoinduction regimens showed DC spectra largely overlapping each other, exhibiting two minima centered ~208 and ~220 nm that are typical of helical-type secondary structures. Deconvolution analysis of these spectra yielded α -helix and β -strand contents of 36% and 15%, respectively, which agree with those of the crystallographic structure of the protein, α -helix = 39% and β -strand = 9% (PDB code 1csg; Walter et al. 1992).

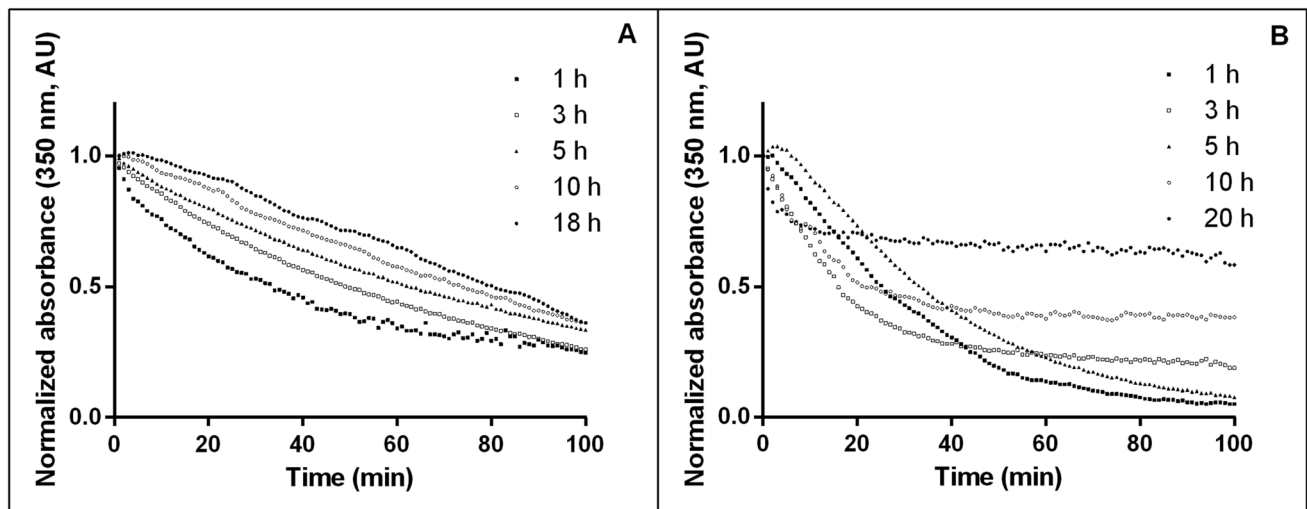


Fig. 7 Kinetics of proteolytic digestion of rHuGM-CSF IBs with proteinase-K from cultures obtained at different pre-induction temperatures: 30 °C (A) or 34 °C (B) and subsequent thermoinduction at 42 °C. IBs were collected at different times post-induction: 1 h (black

circles), 3 h (white circles), 5 h (triangles), 10 h (white squares) and 18/20 h (black squares). The progressive degradation was followed by absorbance at 350 nm for 100 min, and data were normalized. Traces represent the average of at least two independent experiments

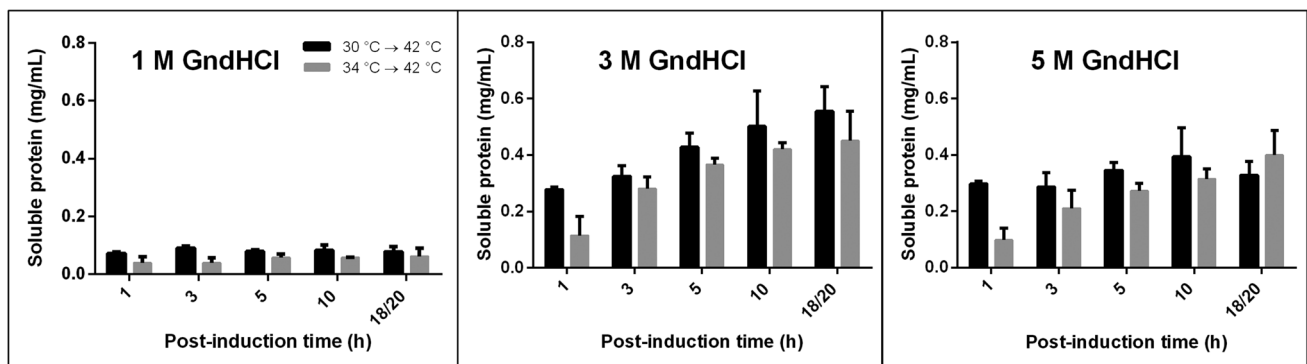


Fig. 8 Solubilization profiles of rHuGM-CSF IBs obtained from pre-induced cultures at 30 °C or 34 °C after different times of thermoinduction at 42 °C (1, 3, 5, 10 and 18/20 h). The amount of solubilized protein (mg/ml) after 24 h of incubation with 1 M (A), 3 M (B), and

5 M (C) concentrations of guanidinium chloride is presented. Bars indicate the mean and standard deviation of the data obtained at each time from three independent experiments

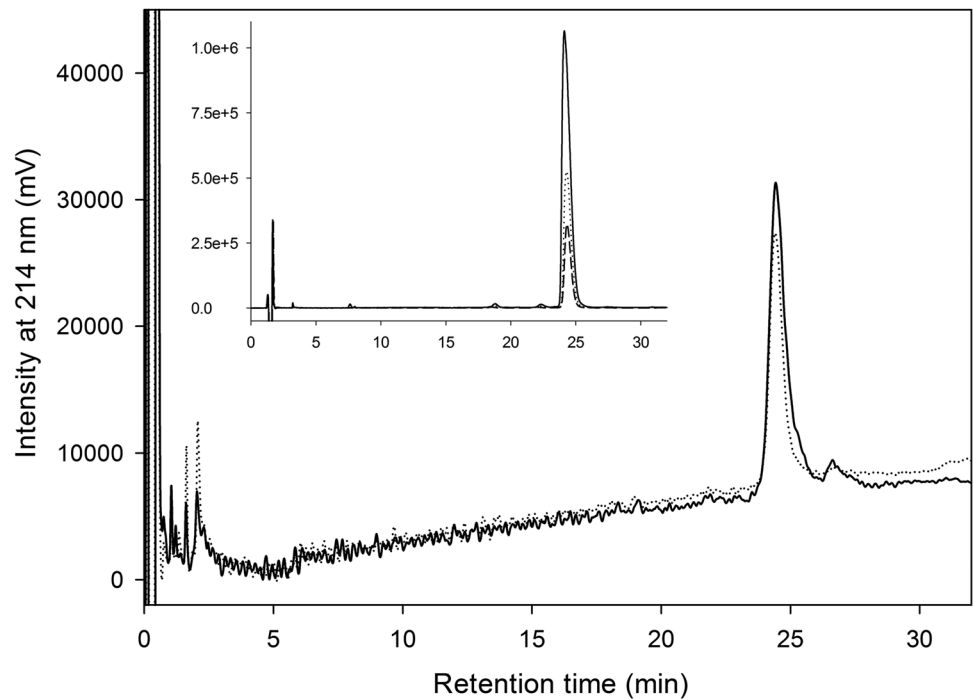
Discussion

Innumerable bioprocesses of recombinant protein production for therapeutic use are carried out in *E. coli*, and the formation of bacterial aggregates has become a common phenomenon that, contrary to being relegated, has gained interest in recent years (García-Fruitós et al. 2010; De Marco et al. 2019; Pesarrodonna et al. 2019; Jäger et al. 2020). Inclusion bodies (IBs) are enriched reservoirs of recombinant protein, which can be considered a previous protein purification step (De Marco et al. 2019; Restrepo-Pineda et al. 2021). In thermoinducible systems, the recombinant protein overexpression, the heat shock response, and IBs formation co-occur in the cells (Valdez-Cruz et al. 2010; Restrepo-Pineda

et al. 2021a). Therefore, understanding the communication between molecular responses and physiological events in this system can be useful to design optimized production bioprocesses that allow higher yields of biologically active recombinant protein with inexpensive and straightforward recovery steps from IBs (Rosano et al. 2019; Restrepo-Pineda et al. 2021).

In thermoinducible systems, the temperature upshifts and the over synthesis of recombinant proteins and HSPs cause an increase in energy demand, metabolic alterations, and a decrease in cell growth (Hoffmann and Rinas 2004; Restrepo-Pineda et al. 2021). Our study shows the effect of pre-induction temperature on cell growth, chaperone composition, and structural characteristics of IBs collected at

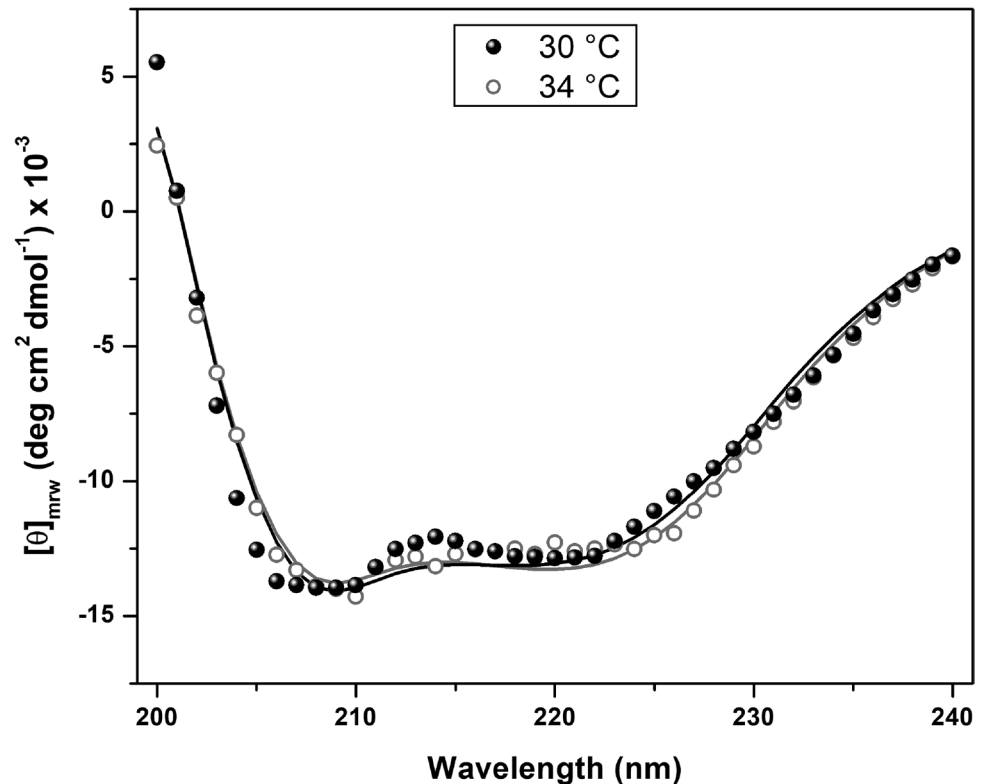
Fig. 9 Purification of the folded rHuGM-CSF by reverse-phase HPLC. The folded protein solution (~10 µg) from pre-induced cultures at 30 °C (continuous line) or 34 °C (dotted line) after thermoinduction at 42 °C was loaded in the HPLC at 50 °C with a detector wavelength of 214 nm. A standard curve (inset) of the European Pharmacopoeia reference standard for human GM-CSF (Molgramostim, Y0000251, Sigma Aldrich, St. Louis, MO, USA) was carried out at 1.345, 0.672, and 0.336 mg/ml



different times of thermoinduction from cultures producing the rHuGM-CSF. *E. coli* W3110 growing either at 30 °C or 34 °C reached similar maximum biomass. However, a significant increase in the specific growth rate (μ), the specific glucose consumption rate (q_g), and the specific acetate

formation rate (q_p) was evidenced by increasing the pre-induction temperature (Fig. 1, Table 1). In previous publications, the production of recombinant proteins under the $\lambda pL/pR-c1857$ thermoinducible system was also accompanied by the acetate accumulation, reporting values near 4.0 g/l in *E.*

Fig. 10 Far-UV circular dichroism spectra of purified rHuGM-CSF in water from IBs obtained from pre-induced cultures at 30 °C (filled circles) or 34 °C (open circles) after thermoinduction at 42 °C. Each spectrum corresponded to the average of three repetitive scans and was corrected by the buffer signal



coli BL21 (Caspeta et al. 2009, 2013); 0.2 g/l in *E. coli* K-12 (Mansey et al. 2014); 3.5–7.4 g/l in *E. coli* 53,606 (Restrepo-Pineda et al. 2019) and 5.9–11.3 g/l in *E. coli* W3110 (this work). Aerobic *E. coli* cultures under conditions of excess glucose (20–40 g/l) are accompanied by metabolic overflow, which can lead to high excretion of acetate and other by-products (Wittmann et al. 2007; Phue and Shiloach 2004). This acetate accumulation is due to a flow redirection from pyruvate dehydrogenase to pyruvate oxidase to remedy the pyruvate node load during induction (Wittmann et al. 2007; Shiloach and Rinas 2009). Concentrations above 2.4 g/l of acetate in the culture medium can generate a decrease in bacterial growth and inhibit recombinant protein production (Dittrich et al. 2005; Eiteman and Altman 2006), which could be associated with the fall in biomass at the end of thermoinduced cultures (Fig. 1) as it can also be due to a complete depletion of the carbon (Restrepo-Pineda et al. 2019).

Bacterial growth at 34 °C favored the accumulation of rHuGM-CSF in IBs during thermoinduction at 42 °C in comparison with a pre-induction of 30 °C, possibly due to those bacterial cells grown at 34 °C have an accelerated rate of cellular processes, reflected in a higher specific growth rate compared to 30 °C. This may be associated with an increased rate of translation, a higher concentration of nascent recombinant polypeptide in the cytosol, and, therefore, the probability of stereo-specific interactions that lead to the increase of rHuGM-CSF aggregation (Singh et al. 2020; Adachi et al. 2015). In addition, it is possible that cells cultivated at 30 °C require more resources to deal with heat stress; while the cells that grew at 34 °C already have a physiological and metabolic pre-adaptation to thermal stress, favoring productivity (Cullum et al. 2001). According to these, the pre-induction temperature is modifying not only the specific growth rate but also the post-induction recombinant protein synthesis rate and its accumulation in IBs. The cellular growth rate has been related to the IBs amount, and biological activity of the recombinant protein produced in them. Iafolla et al. (2008) found that at the fastest growth rate, more active EGFP (enhanced green fluorescent protein) was present in IBs, while at slower growth rate IBs are less abundant and with less active EGFP (Iafolla et al. 2008).

The tendency of GM-CSF to aggregation makes it an ideal candidate protein to study the behavior at the structural and compositional level of IBs obtained from thermoinduced cultures. The first studies about the GM-CSF production in bacteria were based on a temperature-inducible plasmid (from 28 to 42–48 °C). They reported that both murine GM-CSF (DeLamarter et al. 1985) and human GM-CSF (Burgess et al. 1987) accumulated in IBs when produced in *E. coli*, which agrees with our results. This knowledge has been corroborated in subsequent reports using chemical induction with IPTG (Schwanke

et al. 2009; Thomson et al. 2012) or autoinduction (Malekian et al. 2019a). However, those did not analyze the protein secondary structure or the amyloid content of the aggregates, they only demonstrated the bioactivity of the recombinant GM-CSF obtained from IBs. The human GM-CSF produced in *E. coli* under the regulation of a heat-inducible promoter showed a specific activity of 2.9×10^7 units/mg with bone marrow cells (Burgess et al. 1987). Recombinant Murine-derived GM-CSF stimulated the growth of granulocyte and macrophage colonies of mouse bone marrow cells (DeLamarter et al. 1985). The refolded and purified rHuGM-CSF promoted the cell growth in a human hematopoietic cell line in a similar way to the commercially available protein (Schwanke et al. 2009; Thomson et al. 2012). Notably, these results indicate that the recombinant protein reaches an active conformation after recuperation and refolding from IBs and, confirms that the absence of glycosylation or the addition of an extra N-terminal methionine residue does not affect the bioactivity of the rHuGM-CSF obtained on a bacterial platform (Burgess et al. 1987; DeLamarter et al. 1985).

E. coli has robust control systems to assist the folding of newly synthesized proteins in the cytosol, including the DnaK chaperone with its DnaJ and GrpE co-chaperones, and GroEL chaperone with its GroES co-chaperone (Houry 2001; Bhandari and Houry 2015). In this work, molecular chaperones were differentially associated with the protein fractions, detecting DnaK and GroEL in the IBs. At the same time, DnaJ and GroES were found in the total soluble protein, possibly reflecting the order in which they interact with the protein folding intermediates during aggregation. Differences in DnaK expression due to pre-induction temperature suggest that DnaK is preferentially required at elevated temperatures to maintain viable cell growth and other multiple bacterial functions (Mayer 2021). Previous results indicate that *E. coli* cells with *dnak* null mutations can grow slowly at 30 °C and 37 °C, but at 42 °C they lose the ability to form colonies after 2 h of exposure (Paek and Walker 1987). In contrast, the constant production of GroEL throughout the thermoinduction stage and under the two growth temperature conditions evaluated corroborates its importance as a central regulator for protein folding in *E. coli* (Hayer-Hartl et al. 2015). Some studies have reported that GroEL complies a job of protecting the bacterial growth in a wide range of temperatures, from low (17 to 30 °C; Fayet et al. 1989) to normal/high temperatures (20 to 40 °C; Kusukawa and Yura 1988). In brief, GroEL seems to act as a key piece in supporting growth at normal physiological temperatures, whereas DnaK might be essential mostly at higher temperature. Some studies have revealed the presence of chaperones such as DnaK, GrpE, GroEL, GroES, IbpA and IbpB in IBs isolated from *E. coli* after induction by IPTG or temperature increase, but they did not find DnaJ (Hoffmann and Rinas

2000; Rinas et al. 2007; Jürgen et al. 2010). The absence of DnaJ was not discussed in those reports, probably because it is assumed that it remains in the soluble fraction like most HSPs (> 90%; Hoffmann and Rinas 2000).

In recent years, it has been proposed that the structural properties of IBs define the protocols of isolation, solubilization, and refolding, mainly during recovery of a biologically active recombinant protein (Singh and Panda 2005; Singhvi et al. 2020). Structural changes of IBs depend not only on the dynamic of aggregation and nature of the recombinant protein but also on the bioprocess conditions (Castellanos-Mendoza et al. 2014; Calcines-Cruz et al. 2018; Restrepo-Pineda et al. 2019; De Marco et al. 2019). Structural analyses by ATR-FTIR and binding to Th-T indicated that at a higher pre-induction temperature (34 °C), the IBs have a lower proportion of amyloid-like conformations than at 30 °C (Figs. 5 and 6). Moreover, previous studies revealed that regions with higher amyloid-like conformation in IBs are more resistant to proteolysis with by PK (Upadhyay et al. 2012). Therefore, the degradation profiles in Fig. 7 also confirm the results observed in the secondary structure analysis by ATR-FTIR and binding to Th-T, suggesting that the rHuGM-CSF IBs obtained at 30 °C → 42 °C have a higher proportion of amyloid-like structure compared to IBs obtained at 34 °C → 42 °C.

Usually, high concentrations of urea or GndHCl (6–8 M) are used in the solubilization of IBs, which can cause a significant disturbance in the structure of the recombinant protein folded state and result in a low recovery of bioactive protein (Upadhyay et al. 2012, 2016; Singhvi et al. 2020). In this study, the rHuGM-CSF IBs were effectively solubilized in low concentrations of GndHCl (3 M and 5 M, Fig. 8), which is important in the downstream process at the industrial level. When IBs can be solubilized at the lowest possible denaturant concentration, it can result in proteins that retain part of their folded structure, making refolding processes more efficient and consequently recovering biological activity (Singh et al. 2015). Indeed, the removal of the denaturing agent yielded folded proteins with native-like secondary structure contents, as revealed by the HPLC purification recovering the folded rHuGM-CSF (Fig. 9), as also by solution CD spectroscopy (Fig. 10). CD is a widely used tool to identify changes in the secondary structure of proteins, which can impact their mechanism of action or in the regulation of their biological activity (Kelly and Price 2000; Kelly et al. 2005). The application of CD to analyze the protein structure after denaturalization, and to associate these data with the rate of recovery of biological activity is common (Kelly and Price 1997). In this study, CD spectra of rHuGM-CSF from *E. coli* cultures growing either at 30 or 34 °C were similar (Fig. 10) A typical helical conformation with two negative shoulders at ~208 and ~220 nm was observed. Calculations of the secondary structure suggest 36% α -helix and

15% β -strand for both conditions, which agrees with previously reported values of ~30% α -helix content for recombinant GM-CSF by far-UV CD (Malekian et al. 2019b; Wingfield et al. 1988). The refolded rHuGM-CSF characterization by CD allowed obtaining valuable insights to understand the protein structure–function relationship.

The claims obtained in this study could serve as a methodological proposal to produce biopharmaceutical proteins in IBs using thermoinduced systems. By modifying the growth or pre-induction temperature, it is possible to have conformationally different IBs with a high content of recombinant protein, and easier to solubilize, lowering bioprocesses efforts and costs. Moreover, the similarity in the secondary structure content between the recombinant protein isolated from bacterial aggregates and the reported structure of rHuGM-CSF revealed the possibility of obtaining biologically active protein after solubilization and refolding processes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00253-022-11908-z>.

Acknowledgements Sara Restrepo-Pineda is a doctoral student from “Programa de Doctorado en Ciencias Biológicas” of the “Universidad Nacional Autónoma de México” (UNAM) and received fellowship from CONACYT (CVU 589949), authors thank both institutions for the support provided. EG-H, NS-P, NOP, MAT-R and NAV-C are members of the Sistema Nacional de Investigadores, Consejo Nacional de Ciencia y Tecnología. This project was developed under the Institutional Program of the Instituto de Investigaciones Biomédicas UNAM: “La producción de biomoléculas de interés biomédico en bacterias y hongos”. Authors also appreciate the technical support by Dr. Axel Luviano in CD experiments and Diego Rosiles-Becerril and M. Sc. Luis Pablo Ávila-Barrientos in protein refolding and HPLC development.

Author contributions SR-P, NAV-C, EG-H, NS-P contributed to experimentation; SR-P, EG-H, NAV-C, and MAT-R helped in conceptualization; MAT-R and NAV-C contributed to funding acquisition; MAT-R and NAV-C contributed to project administration; MAT-R, NOP, NS-P, EG-H, and NAV-C helped in resources; SR-P, MAT-R, and NAV-C contributed to writing—original draft; SR-P, NOP, NS-P, MAT-R, and NAV-C helped in writing—review and editing.

Funding This work was supported by “Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Universidad Nacional Autónoma de México” (PAPIIT-UNAM IN210822: NAVC, IN211422, IV201220: MATR). The funders had no role in data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability The authors confirm that all relevant data are included in this article and its supplementary information files.

Declarations

Ethics approval This article does not report any studies with human participants or animals performed by the authors.

Conflict of interest NOP works in Probiomed S.A. de C.V., which manufactures recombinant human therapeutic proteins.

References

- Adachi M, So M, Sakurai K, Kardos J, Goto Y (2015) Supersaturation-limited and unlimited phase transitions compete to produce the pathway complexity in amyloid fibrillation. *J Biol Chem* 290(29):18134–18145. <https://doi.org/10.1074/jbc.M115.648139>
- Ami D, Natalello A, Gatti-Lafranconi P, Lotti M, Doglia SM (2005) Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Lett* 579(16):3433–3436. <https://doi.org/10.1016/j.febslet.2005.04.085>
- Ami D, Natalello A, Taylor G, Tonon G, Doglia SM (2006) Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. *Biochim Biophys Acta* 1764(4):793–799. <https://doi.org/10.1016/j.bbapap.2005.12.005>
- Assenberg R, Wan PT, Geisse S, Mayr LM (2013) Advances in recombinant protein expression for use in pharmaceutical research. *Curr Opin Struct Biol* 23(3):393–402. <https://doi.org/10.1016/j.sbi.2013.03.008>
- Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MM, Ramadan HA, Saini KS, Redwan EM (2015) Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives. *J Microbiol Biotechnol* 25(7):953–962. <https://doi.org/10.4014/jmb.1412.12079>
- Balchin D, Hayer-Hartl M, Hartl FU (2016) *In vivo* aspects of protein folding and quality control. *Science* 353(6294):aac4354. <https://doi.org/10.1126/science.aac4354>
- Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* 22:1399–1408. <https://doi.org/10.1038/nbt1029>
- Bardwell JC, Craig EA (1984) Major heat shock gene of *Drosophila* and the *Escherichia coli* heat-inducible *dnaK* gene are homologous. *Proc Natl Acad Sci USA* 81(3):848–852. <https://doi.org/10.1073/pnas.81.3.848>
- Belew M, Zhou Y, Wang S, Nyström LE, Janson JC (1994) Purification of recombinant human granulocyte-macrophage colony-stimulating factor from the inclusion bodies produced by transformed *Escherichia coli* cells. *J Chromatogr A* 679(1):67–83. [https://doi.org/10.1016/0021-9673\(94\)80312-9](https://doi.org/10.1016/0021-9673(94)80312-9)
- Bhandari V, Houry WA (2015) Substrate interaction networks of the *Escherichia coli* chaperones: trigger factor, DnaK and GroEL. *Adv Exp Med Biol* 883:271–294. https://doi.org/10.1007/978-3-319-23603-2_15
- Bonaventura A, Vecchié A, Wang TS, Lee E, Cremer PC, Carey B, Rajendram P, Hudock KM, Korbee L, Van Tassell BW, Dagna L, Abbate A (2020) Targeting GM-CSF in COVID-19 pneumonia: rationale and strategies. *Front Immunol* 11:1625. <https://doi.org/10.3389/fimmu.2020.01625>
- Burgess AW, Begley CG, Johnson GR, Lopez AF, Williamson DJ, Mermod JJ, Simpson RJ, Schmitz A, DeLamararter JF (1987) Purification and properties of bacterially synthesized human granulocyte-macrophage colony stimulating factor. *Blood* 69(1):43–51. <https://doi.org/10.1182/blood.V69.1.43.43>
- Calcines-Cruz C, Olvera A, Castro-Acosta RM, Zavala G, Alagón A, Trujillo-Roldán MA, Valdez-Cruz NA (2018) Recombinant-phospholipase A2 production and architecture of inclusion bodies are affected by pH in *Escherichia coli*. *Int J Biol Macromol* 108:826–836. <https://doi.org/10.1016/j.ijbiomac.2017.10.178>
- Cano-Garrido O, Rodríguez-Carmona E, Díez-Gil C, Vázquez E, Elizondo E, Cubarsi R, Seras-Franzoso J, Corchero JL, Rinas U, Ratera I, Ventosa N, Veciana J, Villaverde A, García-Fruitós E (2013) Supramolecular organization of protein-releasing functional amyloids solved in bacterial inclusion bodies. *Acta Biomater* 9:6134–6142. <https://doi.org/10.1016/j.actbio.2012.11.033>
- Carrió M, Villaverde A (2003) Role of molecular chaperones in inclusion body formation. *FEBS Lett* 537(1–3):215–221. [https://doi.org/10.1016/S0014-5793\(03\)00126-1](https://doi.org/10.1016/S0014-5793(03)00126-1)
- Carrió M, González-Montalbán N, Vera A, Villaverde A, Ventura S (2005) Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* 347(5):1025–1037. <https://doi.org/10.1016/j.jmb.2005.02.030>
- Caspeta L, Flores N, Pérez NO, Bolívar F, Ramírez OT (2009) The effect of heating rate on *Escherichia coli* metabolism, physiological stress, transcriptional response, and production of temperature-induced recombinant protein: a scale-down study. *Biotechnol Bioeng* 102(2):468–482. <https://doi.org/10.1002/bit.22084>
- Caspeta L, Lara AR, Pérez NO, Flores N, Bolívar F, Ramírez OT (2013) Enhancing thermo-induced recombinant protein production in *Escherichia coli* by temperature oscillations and post-induction nutrient feeding strategies. *J Biotechnol* 167(1):47–55. <https://doi.org/10.1016/j.jbiotec.2013.06.001>
- Castellanos-Mendoza A, Castro-Acosta RM, Olvera A, Zavala G, Mendoza-Vera M, García-Hernández E, Alagón A, Trujillo-Roldán MA, Valdez-Cruz NA (2014) Influence of pH control in the formation of inclusion bodies during production of recombinant sphingomyelinase-D in *Escherichia coli*. *Microb Cell Fact* 13:1–14. <https://doi.org/10.1186/s12934-014-0137-9>
- Caulcott CA, Rhodes M (1986) Temperature-induced synthesis of recombinant proteins. *Trends Biotechnol* 4:142–146. [https://doi.org/10.1016/0167-7799\(86\)90164-2](https://doi.org/10.1016/0167-7799(86)90164-2)
- Cullum AJ, Bennett AF, Lenski RE (2001) Evolutionary adaptation to temperature. IX. Preadaptation to novel stressful environments of *Escherichia coli* adapted to high temperature. *Evolution* 55(11):2194–2202. <https://doi.org/10.1111/j.0014-3820.2001.tb00735.x>
- Cumming DA (1991) Glycosylation of recombinant protein therapeutics: control and functional implications. *Glycobiology* 1(2):115–130. <https://doi.org/10.1093/glycob/1.2.115>
- Das KM, Banerjee S, Shekhar N, Damodaran K, Nair R, Somani S, Raiker VP, Jain S, Padmanabhan S (2011) Cloning, soluble expression and purification of high yield recombinant hGMCSF in *Escherichia coli*. *Int J Mol Sci* 12(3):2064–2076. <https://doi.org/10.3390/ijms12032064>
- De Groot NS, Ventura S (2006) Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Lett* 580(27):6471–6476. <https://doi.org/10.1016/j.febslet.2006.10.071>
- De Groot NS, Sabate R, Ventura S (2009) Amyloids in bacterial inclusion bodies. *Trends Biochem Sci* 34(8):408–416. <https://doi.org/10.1016/j.tibs.2009.03.009>
- De Marco A, Ferrer-Mirallas N, Garcia-Fruitós E, Mitraki A, Peternel S, Rinas U, Trujillo-Roldán MA, Valdez-Cruz NA, Vázquez E, Villaverde A (2019) Bacterial inclusion bodies are industrially exploitable amyloids. *FEMS Microbiol Rev* 43:53–72. <https://doi.org/10.1093/femsre/fuy038>
- Del Vecchio P, Graziano G, Granata V, Barone G, Mandrich L, Rossi M, Manco G (2002) Denaturing action of urea and guanidine hydrochloride towards two thermophilic esterases. *Biochem J* 367(3):857–863. <https://doi.org/10.1042/BJ20020695>
- DeLamararter JF, Mermod JJ, Liang CM, Eliason JF, Thatcher DR (1985) Recombinant murine GM-CSF from *E. coli* has biological activity and is neutralized by a specific antiserum. *EMBO J* 4(10):2575–2581. <https://doi.org/10.1002/j.1460-2075.1985.tb03973.x>
- Dittrich CR, Bennett GN, San KY (2005) Characterization of the acetate-producing pathways in *Escherichia coli*. *Biotechnol Prog* 21(4):1062–1067. <https://doi.org/10.1021/bp050073s>
- Dodd IB, Shearwin KE, Perkins AJ, Burr T, Hochschild A, Egan JB (2004) Cooperativity in long-range gene regulation by the lambda CI repressor. *Genes Dev* 18(3):344–354. <https://doi.org/10.1101/gad.1167904>

- Dougan M, Dranoff G, Dougan SK (2019) GM-CSF, IL-3, and IL-5 family of cytokines: regulators of inflammation. *Immunity* 50(4):796–811. <https://doi.org/10.1016/j.immuni.2019.03.022>
- Eiteman MA, Altman E (2006) Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol* 24(11):530–536. <https://doi.org/10.1016/j.tibtech.2006.09.001>
- Espargaró A, Sabate R, Ventura S (2008) Kinetic and thermodynamic stability of bacterial intracellular aggregates. *FEBS Lett* 582:3669–3673. <https://doi.org/10.1016/j.febslet.2008.09.049>
- Fahnert B, Lilie H, Neubauer P (2004) Inclusion bodies: formation and utilization. In: Scheper T (ed) *Physiological stress responses in bioprocesses*. *Adv Biochem Eng* 89:93–142
- Fayet O, Ziegelhoffer T, Georgopoulos C (1989) The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J Bacteriol* 171(3):1379–1385. <https://doi.org/10.1128/jb.171.3.1379-1385.1989>
- Francisco-Cruz A, Aguilar-Santelises M, Ramos-Espinosa O, Mata-Espinosa D, Marquina-Castillo B, Barrios-Payan J, Hernandez-Pando R (2014) Granulocyte–macrophage colony-stimulating factor: not just another haematopoietic growth factor. *Med Oncol* 31(1):774. <https://doi.org/10.1007/s12032-013-0774-6>
- García-Fruitós E, Seras-Franzoso J, Vazquez E, Villaverde A (2010) Tunable geometry of bacterial inclusion bodies as substrate materials for tissue engineering. *Nanotechnol* 21(20):205101. <https://doi.org/10.1088/0957-4484/21/20/205101>
- Gil-García M, Navarro S, Ventura S (2020) Coiled-coil inspired functional inclusion bodies. *Microb Cell Fact* 19:117. <https://doi.org/10.1186/s12934-020-01375-4>
- Gill RT, Valdes JJ, Bentley WE (2000) A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*. *Metab Eng* 2(3):178–189. <https://doi.org/10.1006/mben.2000.0148>
- González-Montalbán N, García-Fruitós E, Villaverde A (2007) Recombinant protein solubility—does more mean better? *Nat Biotechnol* 25(7):718. <https://doi.org/10.1038/nbt0707-718>
- Guisbert E, Herman C, Lu CZ, Gross CA (2004) A chaperone network controls the heat shock response in *E. coli*. *Genes Dev* 18(22):2812–2821. <https://doi.org/10.1101/gad.1219204>
- Guisbert E, Yura T, Rhodius VA, Gross CA (2008) Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev* 72(3):545–554. <https://doi.org/10.1128/MMBR.00007-08>
- Hamilton JA (2019) GM-CSF-dependent inflammatory pathways. *Front Immunol* 10:2055. <https://doi.org/10.3389/fimmu.2019.02055>
- Hamilton JA (2020) GM-CSF in inflammation. *J Exp Med* 217(1). <https://doi.org/10.1084/jem.20190945>
- Hartl FU, Hayer-Hartl M (2009) Converging concepts of protein folding *in vitro* and *in vivo*. *Nat Struct Mol Biol* 16:574. <https://doi.org/10.1038/nsmb.1591>
- Hayer-Hartl M, Bracher A, Hartl FU (2015) The GroEL–GroES chaperonin machine: a nano-cage for protein folding. *Trends Biochem Sci* 41(1):62–76. <https://doi.org/10.1016/j.tibs.2015.07.009>
- Hengge-Aronis R (1993) Survival of hunger and stress: the role of rpoS in early stationary phase gene regulation in *E. coli*. *Cell* 72(2):165–168. [https://doi.org/10.1016/0092-8674\(93\)90655-a](https://doi.org/10.1016/0092-8674(93)90655-a)
- Hoffmann F, Rinas U (2000) Kinetics of heat-shock response and inclusion body formation during temperature-induced production of basic fibroblast growth factor in high-cell density cultures of recombinant *Escherichia coli*. *Biotechnol Prog* 16(6):1000–1007. <https://doi.org/10.1021/bp0000959>
- Hoffmann F, Rinas U (2004) Stress induced by recombinant protein production in *Escherichia coli*. In: Scheper T (ed) *Physiological stress responses in bioprocesses*. *Adv Biochem Eng* 89:73–92. <https://doi.org/10.1007/b93994>
- Hoffmann F, Weber J, Rinas U (2002) Metabolic adaptation of *Escherichia coli* during temperature-induced recombinant protein production: 1. Readjustment of metabolic enzyme synthesis. *Biotechnol Bioeng* 80:313–319. <https://doi.org/10.1002/bit.10379>
- Houry WA (2001) Chaperone-assisted protein folding in the cell cytoplasm. *Curr Protein Pept Sci* 2(3):227–244. <https://doi.org/10.2174/1389203013381134>
- Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol* 39(3):383–399. <https://doi.org/10.1007/s10295-011-1082-9>
- Iafolla MAJ, Mazumder M, Sardana V, Velauthapillai T, Pannu K, McMillen DR (2008) Dark proteins: effect of inclusion body formation on quantification of protein expression. *Proteins* 72:1233–1242. <https://doi.org/10.1002/prot.22024>
- Jäger VD, Lamm R, Küsters K, Ölçücü G, Oldiges M, Jaeger KE, Büchs J, Krauss U (2020) Catalytically-active inclusion bodies for biotechnology—general concepts, optimization, and application. *Appl Microbiol Biotechnol* 104:7313–7329. <https://doi.org/10.1007/s00253-020-10760-3>
- Jürgen B, Breitenstein A, Urlacher V, Büttner K, Lin H, Hecker M, Schweder T, Neubauer P (2010) Quality control of inclusion bodies in *Escherichia coli*. *Microb Cell Fact* 9:41. <https://doi.org/10.1186/1475-2859-9-41>
- Kaufman HL, Ruby CE, Hughes T, Slingluff CL (2014) Current status of granulocyte–macrophage colony-stimulating factor in the immunotherapy of melanoma. *J Immunother Cancer* 2(1):1–13. <https://doi.org/10.1186/2051-1426-2-11>
- Kaur J, Kumar A, Kaur J (2018) Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. *Int J Biol Macromol* 106:803–822. <https://doi.org/10.1016/j.ijbiomac.2017.08.080>
- Kelly SM, Price NC (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochim Biophys Acta* 1338(2):161–185. [https://doi.org/10.1016/s0167-4838\(96\)00190-2](https://doi.org/10.1016/s0167-4838(96)00190-2)
- Kelly SM, Price NC (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr Protein Pept Sci* 1(4):349–384. <https://doi.org/10.2174/1389203003381315>
- Kelly SM, Thomas JJ, Price NC (2005) How to study proteins by circular dichroism. *BBA-Proteins Proteom* 1751(2):119–139. <https://doi.org/10.1016/j.bbapap.2005.06.005>
- Krebs MR, Bromley EH, Donald AM (2005) The binding of thioflavin-T to amyloid fibrils: localisation and implications. *J Struct Biol* 149(1):30–37. <https://doi.org/10.1016/j.jsb.2004.08.002>
- Kumar J, Chauhan A, Shah RL, Gupta JA, Rathore AS (2020) Amino acid supplementation for enhancing recombinant protein production in *E. coli*. *Biotechnol Bioeng* 117(8):2420–2433. <https://doi.org/10.1002/bit.27371>
- Kusukawa N, Yura T (1988) Heat shock protein GroE of *Escherichia coli*: key protective roles against thermal stress. *Genes Dev* 2(7):874–882. <https://doi.org/10.1101/gad.2.7.874>
- Lang FM, Lee KMC, Teijaro JR, Becher B, Hamilton JA (2020) GM-CSF-based treatments in COVID-19: reconciling opposing therapeutic approaches. *Nat Rev Immunol* 20:507–514. <https://doi.org/10.1038/s41577-020-0357-7>
- LeVine H (1995) Thioflavine T interaction with amyloid β -sheet structures. *Amyloid* 2(1):1–6. <https://doi.org/10.3109/13506129509031881>
- Lewis DEA, Gussin GN, Adhya S (2016) New insights into the phage genetic switch: effects of bacteriophage lambda operator mutations on DNA looping and regulation of PR, PL, and PRM. *J Mol Biol* 428(22):4438–4456. <https://doi.org/10.1016/j.jmb.2016.08.027>

- Li H, Lantz R, Du D (2019) Vibrational approach to the dynamics and structure of protein amyloids. *Molecules* 24(1):186. <https://doi.org/10.3390/molecules24010186>
- Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins* 80(2):374–381. <https://doi.org/10.1002/prot.23188>
- Lowman HB, Bina M (1990) Temperature-mediated regulation and downstream inducible selection for controlling gene expression from the bacteriophage λ pL promoter. *Gene* 96(1):133–136. [https://doi.org/10.1016/0378-1119\(90\)90353-S](https://doi.org/10.1016/0378-1119(90)90353-S)
- Luviano A, Cruz-Castañeda R, Sánchez-Puig N, García-Hernández E (2019) Cooperative energetic effects elicited by the yeast Shwachman-Diamond syndrome protein (Sdo1) and guanine nucleotides modulate the complex conformational landscape of the elongation factor-like 1 (Efl1) GTPase. *Biophys Chem* 247:13–24. <https://doi.org/10.1016/j.bpc.2019.02.003>
- Macedo B, Sant'Anna R, Navarro S, Cordeiro Y, Ventura S (2015) Mammalian prion protein (PrP) forms conformationally different amyloid intracellular aggregates in bacteria. *Microb Cell Fact* 14(1):174. <https://doi.org/10.1186/s12934-015-0361-y>
- Malekian R, Jahanian-Najafabadi A, Moazen F, Ghavimi R, Mohammadi E, Akbari V (2019) High-yield production of granulocyte-macrophage colony-stimulating factor in *E. coli* BL21 (DE3) by an auto-induction strategy. *Iran J Pharm Res* 18(1):469–478
- Malekian R, Sima S, Jahanian-Najafabadi A, Moazen F, Akbari V (2019) Improvement of soluble expression of GM-CSF in the cytoplasm of *Escherichia coli* using chemical and molecular chaperones. *Protein Express Purif* 160:66–72. <https://doi.org/10.1016/j.pep.2019.04.002>
- Mansey MS, Ghareeb KA, Moghazy AN, Tawfik MM, Fouda MM, El Marzugi NA, Othman NZ, El Enshasy HA (2014) Glucose concentration affects recombinant interferon α -2b production in *Escherichia coli* using thermo-induction system. *J Appl Pharm Sci* 4:1–5. <https://doi.org/10.7324/JAPS.2014.40501>
- Margreiter G, Messner P, Caldwell KD, Bayer K (2008) Size characterization of inclusion bodies by sedimentation field-flow fractionation. *J Biotechnol* 138(3–4):67–73. <https://doi.org/10.1016/j.jbiotec.2008.07.1995>
- Mayer MP (2021) The Hsp70-chaperone machines in bacteria. *Front Mol Biosci* 8:694012. <https://doi.org/10.3389/fmolb.2021.694012>
- Mehta HM, Malandra M, Corey SJ (2015) G-CSF and GM-CSF in neutropenia. *J Immunol* 195(4):1341–1349. <https://doi.org/10.4049/jimmunol.1500861>
- Miller LM, Bourassa MW, Smith RJ (2013) FTIR spectroscopic imaging of protein aggregation in living cells. *Biochim Biophys Acta Biomembr* 1828(10):2339–2346. <https://doi.org/10.1016/j.bbamem.2013.01.014>
- Miller CA, Tucker WT, Meacock PA, Gustafsson P, Cohen SN (1983) Nucleotide sequence of the partition locus of *Escherichia coli* plasmid pSC101. *Gene* 24(2–3):309–315. [https://doi.org/10.1016/0378-1119\(83\)90091-4](https://doi.org/10.1016/0378-1119(83)90091-4)
- Monera OD, Kay CM, Hodges RS (1994) Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Sci* 3(11):1984–1991. <https://doi.org/10.1002/pro.5560031110>
- Morita MT, Tanaka Y, Kodama TS, Kyogoku Y, Yanagi H, Yura T (1999) Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. *Genes Dev* 13(6):655–665. <https://doi.org/10.1101/gad.13.6.655>
- Nicola NA, Metcalf D, Matsumoto M, Johnson GR (1983) Purification of a factor inducing differentiation in murine myelomonocytic leukemia cells. *J Biol Chem* 258(14):9017–9023. [https://doi.org/10.1016/S0021-9258\(18\)32158-6](https://doi.org/10.1016/S0021-9258(18)32158-6)
- Okamoto M, Nakai M, Nakayama C, Yanagi H, Matsui H, Noguchi H, Hara H (1991) Purification and characterization of three forms of differently glycosylated recombinant human granulocyte-macrophage colony-stimulating factor. *Arch Biochem Biophys* 286(2):562–568. [https://doi.org/10.1016/0003-9861\(91\)90080-3](https://doi.org/10.1016/0003-9861(91)90080-3)
- Olsson J, Paulsson J, Nordström K (2004) Effect of the CopB auxiliary replication control system on stability of maintenance of Par+ plasmid R1. *J Bacteriol* 186(1):207–211. <https://doi.org/10.1128/JB.186.1.207-211.2004>
- Overton TW (2014) Recombinant protein production in bacterial hosts. *Drug Discov Today* 19(5):590–601. <https://doi.org/10.1016/j.drudis.2013.11.008>
- Paek KH, Walker GC (1987) *Escherichia coli* dnaK null mutants are inviable at high temperature. *J Bacter* 169(1):283–290. <https://doi.org/10.1128/jb.169.1.283-290.1987>
- Pesarrodona M, Jauset T, Díaz-Riascos ZV, Sánchez-Chardi A, Beaulieu ME, Seras-Franzoso J, Sánchez-García L, Baltà-Foix R, Mancilla S, Fernández Y, Rinas U Jr, Schwartz S, Soucek L, Villaverde A, Abasolo I, Vázquez E (2019) Targeting antitumoral proteins to breast cancer by local administration of functional inclusion bodies. *Adv Sci (weinh)* 6(18):1900849. <https://doi.org/10.1002/advs.201900849>
- Peternel Š, Grdadolnik J, Gaberc-Porekar V, Komel R (2008) Engineering inclusion bodies for non-denaturing extraction of functional proteins. *Microb Cell Fact* 7(1):34. <https://doi.org/10.1186/1475-2859-7-34>
- Phue JN, Shiloach J (2004) Transcription levels of key metabolic genes are the cause for different glucose utilization pathways in *E. coli* B (BL21) and *E. coli* K (JM109). *J Biotechnol* 109(1–2):21–30. <https://doi.org/10.1016/j.jbiotec.2003.10.038>
- Remaut E, Stanssens P, Fiers W (1981) Plasmid vectors for high-efficiency expression controlled by the pL promoter of coliphage lambda. *Gene* 15(1):81–93. [https://doi.org/10.1016/0378-1119\(81\)90106-2](https://doi.org/10.1016/0378-1119(81)90106-2)
- Restrepo-Pineda S, Bando-Campos CG, Valdez-Cruz NA, Trujillo-Roldán MA (2019) Recombinant production of ESAT-6 antigen in thermoinducible *Escherichia coli*: the role of culture scale and temperature on metabolic response, expression of chaperones, and architecture of inclusion bodies. *Cell Stress Chaperon* 24:777–792. <https://doi.org/10.1007/s12192-019-01006-x>
- Restrepo-Pineda S, Pérez NO, Valdez-Cruz NA, Trujillo-Roldán MA (2021) Thermoinducible expression system for producing recombinant proteins in *Escherichia coli*: advances and insights. *FEMS Microbiol Rev* 45:fuab023. <https://doi.org/10.1093/femsre/fuab023>
- Rinas U, Hoffmann F, Betiku E, Estapé D, Marten S (2007) Inclusion body anatomy and functioning of chaperone-mediated in vivo inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. *J Biotechnol* 127(2):244–257. <https://doi.org/10.1016/j.jbiotec.2006.07.004>
- Rinas U, García-Fruitós E, Corchero JL, Vázquez E, Seras-Franzoso J, Villaverde A (2017) Bacterial inclusion bodies: discovering their better half. *Trends Biochem Sci* 1334:1–12. <https://doi.org/10.1016/j.tibs.2017.01.005>
- Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* 5:172. <https://doi.org/10.3389/fmicb.2014.00172>
- Rosano GL, Morales ES, Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Sci* 28(8):1412–1422. <https://doi.org/10.1002/pro.3668>
- Sánchez-García L, Martín L, Mangués R, Ferrer-Miralles N, Vázquez E, Villaverde A (2016) Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact* 15(1):33. <https://doi.org/10.1186/s12934-016-0437-3>
- Sandoval-Basurto EA, Gosset G, Bolívar F, Ramírez OT (2005) Culture of *Escherichia coli* under dissolved oxygen gradients

- simulated in a two-compartment scale-down system: Metabolic response and production of recombinant protein. *Biotechnol Bioeng* 89(4):453–463
- Schwanke RC, Renard G, Chies JM, Campos MM, Junior ELB, Santos DS, Basso LA (2009) Molecular cloning, expression in *Escherichia coli* and production of bioactive homogeneous recombinant human granulocyte and macrophage colony stimulating factor. *Int J Biol Macromol* 45(2):97–102. <https://doi.org/10.1016/j.ijbmac.2009.04.005>
- Shiloach J, Rinas U (2009) Glucose and acetate metabolism in *E. coli* – system level analysis and biotechnological applications in protein production processes. In: Lee SY (ed) *Systems Biology and Biotechnology of Escherichia coli*. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-9394-4_18
- Singh SM, Panda AK (2005) Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng* 99(4):303–310. <https://doi.org/10.1263/jbb.99.303>
- Singh A, Upadhyay V, Panda AK (2015) Solubilization and refolding of inclusion body proteins. *Methods Mol Biol* 1258:283–291. https://doi.org/10.1007/978-1-4939-2205-5_15
- Singh A, Upadhyay V, Singh A, Panda AK (2020) Structure-function relationship of inclusion bodies of a multimeric protein. *Front Microbiol* 11:876. <https://doi.org/10.3389/fmicb.2020.00876>
- Singha TK, Gulati P, Kumar S (2018) Nonconventional induction strategies for production of recombinant human tumor necrosis factor- α in *Escherichia coli*. *J App Biol Biotechnol* 6:23–27. <https://doi.org/10.7324/JABB.2018.60105>
- Singhvi P, Saneja A, Srichandan S, Panda AK (2020) Bacterial inclusion bodies: a treasure trove of bioactive proteins. *Trends Biotechnol* 38(5):474–486. <https://doi.org/10.1016/j.tibtech.2019.12.011>
- Terol GL, Gallego-Jara J, Martínez RAS, Vivancos AM, Díaz MC, Puente TD (2021) Impact of the expression system on recombinant protein production in *Escherichia coli* BL21. *Front Microbiol* 12:682001. <https://doi.org/10.3389/fmicb.2021.682001>
- Thomson CA, Olson M, Jackson LM, Schrader JW (2012) A simplified method for the efficient refolding and purification of recombinant human GM-CSF. *PLoS ONE* 7(11):e49891. <https://doi.org/10.1371/journal.pone.0049891>
- Trapnell BC, Inoue Y, Bonella F, Morgan C, Jouneau S, Bendstrup E, Campo I, Papisir SA, Yamaguchi E, Cetinkaya E, Ilkovich MM, Kramer MR, Veltkamp M, Kreuter M, Baba T, Ganslandt C, Tarnow I, Waterer G, Joughikainen T, IMPALA Trial Investigators (2020) Inhaled Molgramostim therapy in autoimmune pulmonary alveolar proteinosis. *N Engl J Med* 383(17):1635–1644. <https://doi.org/10.1056/NEJMoa1913590>
- Trujillo-Roldán MA, Peña C, Ramírez OT, Galindo E (2001) Effect of oscillating dissolved oxygen tension on the production of alginate by *Azotobacter vinelandii*. *Biotechnol Prog* 17(6):1042–1048. <https://doi.org/10.1021/bp010106d>
- Upadhyay AK, Murmu A, Singh A, Panda AK (2012) Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. *PLoS ONE* 7:e33951. <https://doi.org/10.1371/journal.pone.0033951>
- Upadhyay V, Singh A, Jha D, Singh A, Panda AK (2016) Recovery of bioactive protein from bacterial inclusion bodies using trifluoroethanol as solubilization agent. *Microb Cell Fact* 15(1):1–13. <https://doi.org/10.1186/s12934-016-0504-9>
- Valdez-Cruz NA, Caspeta L, Pérez NO, Ramírez OT, Trujillo-Roldán MA (2010) Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda *pL* and/or *pR* promoters. *Microb Cell Fact* 9:18. <https://doi.org/10.1186/1475-2859-9-18>
- Valdez-Cruz NA, Ramírez OT, Trujillo-Roldán MA (2011) Molecular responses of *E. coli* caused by heat stress and recombinant protein production during temperature induction. *Bioeng Bugs* 2(2):105–110. <https://doi.org/10.4161/bbug.2.2.14316>
- Valdez-Cruz NA, Reynoso-Cereceda GI, Pérez-Rodríguez S, Restrepo-Pineda S, González-Santana J, Olvera A, Trujillo-Roldán MA (2017) Production of a recombinant phospholipase A2 in *Escherichia coli* using resonant acoustic mixing that improves oxygen transfer in shake flasks. *Microb Cell Fact* 16(1):129. <https://doi.org/10.1186/s12934-017-0746-1>
- Vázquez-Fernández E, Alonso J, Pastrana MA, Ramos A, Stitz L, Vidal E, Dynin I, Petsch B, Silva CJ, Requena JR (2012) Structural organization of mammalian prions as probed by limited proteolysis. *PLoS ONE* 7(11):e50111. <https://doi.org/10.1371/journal.pone.0050111>
- Villaverde A, Benito A, Viaplana E, Cubarsi R (1993) Fine regulation of *c1857*- controlled gene expression in continuous culture of recombinant *Escherichia coli* by temperature. *Appl Environ Microb* 59(10):3485–3487. <https://doi.org/10.1128/aem.59.10.3485-3487.1993>
- Villaverde A, García-Fruitós E, Rinas U, Seras-Franzoso J, Kosoy A, Corchero JL, Vazquez E (2012) Packaging protein drugs as bacterial inclusion bodies for therapeutic applications. *Microb Cell Fact* 11(1):1–5. <https://doi.org/10.1186/1475-2859-11-76>
- Walsh G (2018) Biopharmaceutical benchmarks 2018. *Nat Biotechnol* 36:1136–1145. <https://doi.org/10.1038/nbt.4305>
- Walter MR, Cook WJ, Ealick SE, Nagabhushan TL, Trotta PP, Bugg CE (1992) Three dimensional structure of recombinant human granulocyte-macrophage colony stimulating factor. *J Mol Biol* 224(4):1075–1085. [https://doi.org/10.1016/0022-2836\(92\)90470-5](https://doi.org/10.1016/0022-2836(92)90470-5)
- Wicks IP, Roberts AW (2016) Targeting GM-CSF in inflammatory diseases. *Nat Rev Rheumatol* 12(1):37–48. <https://doi.org/10.1038/nrrheum.2015.161>
- Williams DC, Van Frank RM, Muth WL, Burnett JP (1982) Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Science* 215(4533):687–689. <https://doi.org/10.1126/science.7036343>
- Wingfield P, Graber P, Moonen P, Craig S, Pain R (1988) The conformation and stability of recombinant-derived granulocyte-macrophage colony stimulating factors. *Eur J Biochem* 173(1):65–72. <https://doi.org/10.1111/j.1432-1033.1988.tb13967.x>
- Wittmann C, Weber J, Betiku E, Krömer J, Böhm D, Rinas U (2007) Response of fluxome and metabolome to temperature-induced recombinant protein synthesis in *Escherichia coli*. *J Biotechnol* 132(4):375–384. <https://doi.org/10.1016/j.jbiotec.2007.07.495>
- Wu C, Biancalana M, Koide S, Shea JE (2009) Binding modes of thiopyran-T to the single-layer β -sheet of the peptide self-assembly mimics. *J Mol Biol* 394(4):627–633. <https://doi.org/10.1016/j.jmb.2009.09.056>
- Yamamori T, Ito K, Nakamura Y, Yura T (1978) Transient regulation of protein synthesis in *Escherichia coli* upon shift-up of growth temperature. *J Bacteriol* 134(3):1133–1140. <https://doi.org/10.1128/jb.134.3.1133-1140.1978>
- Yano R, Nagai H, Shiba K, Yura T (1990) A mutation that enhances synthesis of sigma 32 and suppresses temperature-sensitive growth of the *rpoH15* mutant of *Escherichia coli*. *J Bacteriol* 172(4):2124–2130. <https://doi.org/10.1128/jb.172.4.2124-2130.1990>
- Yura T (2019) Regulation of the heat shock response in *Escherichia coli*: history and perspectives. *Genes Genet Syst* 94:103–108. <https://doi.org/10.1266/ggs.19-00005>
- Zhang F, Weng D, Su Y, Yin C, Shen L, Zhang Y, Zhou Y, Li Q, Hu Y, Li H (2020) Therapeutic effect of subcutaneous injection of low dose recombinant human granulocyte-macrophage colony-stimulating factor on pulmonary alveolar proteinosis. *Respir Res* 21(1):1. <https://doi.org/10.1186/s12931-019-1261-1>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.