



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica Superior
d'Enginyeria Agronòmica i del Medi Natural

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Agronómica
y del Medio Natural

Expresión y purificación del factor de iniciación 2 (IF2) de
Mycobacterium tuberculosis

Trabajo Fin de Grado

Grado en Biotecnología

AUTOR/A: Bresó Fernández, Rosendo

Tutor/a: González Pellicer, Ana

Cotutor/a externo: Cortés Méndez, Teresa

Director/a Experimental: Villamayor Belinchón, Laura

CURSO ACADÉMICO: 2023/2024

Expresión y purificación del factor de iniciación 2 (IF2) de *Mycobacterium tuberculosis*

RESUMEN

La tuberculosis es una enfermedad grave causada por *Mycobacterium tuberculosis* (Mtb), y se calcula que una cuarta parte de la población mundial está infectada por esta bacteria de forma latente. Se ha postulado que la regulación de la traducción puede jugar un papel clave en la adaptación de la bacteria dentro del hospedador, por lo que comprender cómo se regulan sus mecanismos de traducción podría facilitar el desarrollo de nuevas estrategias terapéuticas que ayuden a combatir esta enfermedad. El presente trabajo de fin de grado se centra en la expresión y purificación del factor de iniciación 2 (IF2) del sistema de traducción de Mtb, con el propósito de reconstituir un sistema de traducción *in vitro* en el futuro. En el trabajo experimental se emplearon técnicas de bioquímica y biología molecular para producir y purificar esta proteína, que juega un papel crucial en el posicionamiento del tRNA_i en el ribosoma durante el inicio de la traducción. Para la obtención de la proteína se adquirieron vectores comerciales con la secuencia codificante de IF2 en su versión completa, así como una versión truncada (IF2tr) que carece de la región N-terminal desestructurada. Los vectores se transformaron en *Escherichia coli*, consiguiendo la expresión y purificación de IF2 e IF2tr, las cuales se utilizarán posteriormente junto a otros componentes de la maquinaria traduccional para recrear *in vitro* un complejo de iniciación canónico del sistema traduccional de Mtb. Esto permitirá encontrar diferencias estructurales y funcionales entre otras bacterias, con ayuda de herramientas como la microscopía electrónica. Los resultados de tales ensayos proporcionarán información crucial sobre los procesos de traducción en Mtb y podrían tener implicaciones significativas en el diseño de fármacos contra la tuberculosis.

Palabras clave: *Mycobacterium tuberculosis*, traducción, ribosoma, expresión de proteínas, purificación de proteínas, factor de iniciación 2.

ABSTRACT

Tuberculosis (TB) is a serious disease caused by *Mycobacterium tuberculosis* (Mtb), and it is estimated that a quarter of the world's population harbours the bacteria in the form of an asymptomatic latent infection. It has been acknowledged that translational regulation might be key in the adaptation of the bacteria inside the host, so understanding how translation mechanisms are regulated could facilitate the development of new therapeutic strategies to help tackle this disease. The present work focuses on the expression and purification of the initiation factor 2 (IF2), a protein from the Mtb translation system, with the aim of reconstituting an *in vitro* translation system for Mtb. In the experimental work, biochemistry and molecular biology techniques were used to produce and purify IF2, which has a crucial role in the binding of tRNA_i to the ribosome at translation initiation. For that purpose, commercial vectors with the coding sequence of both IF2 and a truncated version of IF2 (IF2tr), lacking the N-terminal unstructured region, were acquired and transformed into *E. coli*. Both IF2 and IF2tr were successfully expressed and purified, for its further *in vitro* assembling along with other transcriptional components. Thus, we would reconstitute a canonical initiation complex of the Mtb translation system that, in combination with tools like electron microscopy, would allow to find structural and functional differences between other bacteria. The results of such assays will provide crucial information about the translation processes in Mtb and could have significant implications in the design of anti-TB drugs.

Key words: *Mycobacterium tuberculosis*, translation, ribosome, protein expression, protein purification, initiation factor 2.

AGRADECIMIENTOS

En primer lugar, me gustaría agradecer a mis tutoras, sin las cuales no hubiera podido realizar este trabajo. A Ana, por enseñarme las bases y principios de la microbiología, fundamentales para cualquier experimento que he tenido que realizar; y a Teresa, por permitirme entrar en su laboratorio y acogerme como a uno más desde el primer día.

Especialmente, debo agradecer a Laura por su dedicación, atención y paciencia conmigo. Gracias por ser la “mamá pato”, explicándome todos los procesos del día a día en el laboratorio y guiándome en mis primeros pasos en este mundillo.

No me olvido del resto del laboratorio, Gerard, Laura F., Pablo, Esther y Carme; que habéis hecho estos meses de trabajo mucho más fáciles, tanto que parecía que no me quería ir a casa. Tampoco me olvido del resto del IBV, que han hecho mi estancia más agradable.

También se lo agradezco a mis amigos más cercanos, por los momentos de desconexión, y a los Cuencheadores: sin vosotros la carrera no hubiera sido lo mismo. Por último, gracias a mis padres por haberme permitido estudiar con total libertad, a mi hermano y a mi abuelo.

ÍNDICE

1. INTRODUCCIÓN	2
1.1 Tuberculosis	2
1.2 Características de <i>Mycobacterium tuberculosis</i>	3
1.3 Regulación del inicio de la traducción en procariontas y en Mtb	3
1.4 Factor de iniciación 2 (IF2)	5
2. OBJETIVOS	6
3. MATERIALES Y MÉTODOS	6
3.1 Diseño de los constructos de IF2 e IF2tr y clonación en pTYB21	6
3.1.1 Constructos y descripción del vector pTYB21	6
3.1.2 Transformación en <i>Escherichia coli</i>	8
3.2 Reclonación de IF2tr en pLIC-SGC1	9
3.2.1 Descripción del vector pLIC-SGC1	9
3.2.2 Clonación independiente de ligación (LIC)	10
3.2.3 Análisis de las colonias transformadas	12
3.3 Expresión y purificación a pequeña escala	13
3.3.1 Pruebas de expresión y solubilidad con pLyss y GroEL/ES	13
3.3.2 Pruebas de expresión, solubilidad y purificación con glucosa 1%	13
3.3.3 Prueba de expresión, solubilidad y purificación de pLIC-IF2tr	14
3.4 Expresión y purificación a gran escala	14
3.4.1 Expresión a gran escala y purificación por el sistema IMPACT	14
3.4.2 Expresión y purificación de IF2tr con chaperonas endógenas	15
3.4.3 Expresión y purificación de pLIC-IF2tr por IMAC	15
3.5 Otras técnicas	16
3.5.1 Preparación de glicerizados	16
3.5.2 Análisis y cuantificación de proteínas por colorimetría (Bradford)	16
3.5.3 Geles de agarosa	16
3.5.4 SDS-PAGE	16
4. RESULTADOS Y DISCUSIÓN	17
4.1 Expresión y purificación a pequeña escala de pTYB21-IF2 y pTYB21-IF2tr	17
4.2 Expresión a gran escala de los constructos pTYB21	18
4.3 Purificación a gran escala de IF2tr	19
4.4 Expresión de IF2tr con chaperonas endógenas de <i>E. coli</i>	21
4.5 Purificación de IF2	24

4.6 Reclonación de IF2tr en pLIC-SGC1	27
4.7 Expresión y purificación a pequeña escala de pLIC-IF2tr	30
4.8 Expresión y purificación a gran escala de pLIC-IF2tr	31
4.9 Consideraciones finales y trabajos futuros.....	34
5. CONCLUSIONES.....	36
6. BIBLIOGRAFÍA.....	37

Índice de figuras

Figura 1. Predicción en Alphafold de la estructura tridimensional del factor de iniciación 2 (IF2) de <i>Mycobacterium tuberculosis</i> en su versión completa (azul) y truncada (morado).....	7
Figura 2. Mapa del vector pTYB21.	7
Figura 3. Flujo de trabajo para la expresión y purificación de proteínas por el sistema IMPACT (NEB).	8
Figura 4. Mapa del plásmido pLIC-SGC1.	10
Figura 5. Geles SDS-PAGE (10% poliacrilamida) de la prueba de expresión a pequeña escala de IF2 e IF2tr en células pLysS y GroEL/ES.....	17
Figura 6. Geles SDS-PAGE (10% poliacrilamida) de las pruebas de expresión y solubilidad a pequeña escala de IF2 y IF2tr en células BL21(DE3) con 1% glucosa.....	18
Figura 7. Gel SDS-PAGE (10% poliacrilamida) de la expresión a gran escala de IF2 e IF2tr en células BL21(DE3) con 1% glucosa.....	19
Figura 8. Gel SDS-PAGE (10% poliacrilamida) de la solubilidad de IF2tr a gran escala y su purificación en columna de quitina.	20
Figura 9. Resultado de la filtración en gel de IF2tr.	21
Figura 10. Gel SDS-PAGE (10% poliacrilamida) de la expresión y solubilidad de IF2tr con chaperonas endógenas.....	22
Figura 11. Gel SDS-PAGE (10% poliacrilamida) de la purificación de IF2tr con chaperonas endógenas en columna de quitina.....	22
Figura 12. Resultado de la purificación de IF2tr con chaperonas endógenas por filtración en gel.	23
Figura 13. Gel SDS-PAGE (10% acrilamida) de la purificación de IF2 e IF2tr a pequeña escala en columna de quitina.....	25
Figura 14. Gel SDS-PAGE (10% acrilamida) de la purificación de IF2 a gran escala en columna de quitina.....	26
Figura 15. Filtración en gel de la purificación de IF2 a gran escala.	26
Figura 16. Gel de agarosa (1% agarosa en TAE 1X) de las pruebas de optimización la PCR de amplificación del inserto.	27
Figura 17. Gel de agarosa (1% agarosa en TAE 1X) de la PCR de inserto optimizada y la digestión de pLIC-SGC1 con BseRI.	28
Figura 18. Gel de agarosa (1% agarosa en TAE 1X) de la PCR de colonia.	28
Figura 19. Gel de agarosa (1% agarosa en TAE 1X) de la segunda PCR de colonia.	29
Figura 20. Gel de agarosa (1% agarosa en TAE 1X) de la tercera PCR de colonia.	29
Figura 21. Geles SDS-PAGE (10% poliacrilamida) de la prueba de expresión y solubilidad a pequeña escala de pLIC-IF2tr.	30
Figura 22. Gel SDS-PAGE (10% acrilamida) de la prueba de purificación de pLIC-IF2tr a pequeña escala en columna de histidina His SpinTrap™ (Cytiva).	31
Figura 23. Gel SDS-PAGE (10% poliacrilamida) de la expresión a gran escala de pLIC-IF2tr.	31
Figura 24. Resultado de la purificación de pLIC-IF2tr en columna de níquel.	32
Figura 25. Resultado de la filtración en gel de los dos lotes de pLIC-IF2tr obtenidos tras la HisTrap.	33
Figura 26. Imagen al microscopio electrónico de transmisión (TEM) de la tinción negativa de complejos de preiniciación 30S.....	36

Relación del trabajo con los Objetivos de Desarrollo Sostenible de la agenda 2030.

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No procede
ODS 1. Fin de la pobreza			X	
ODS 2. Hambre cero				X
ODS 3. Salud y bienestar	X			
ODS 4. Educación de calidad				X
ODS 5. Igualdad de género				X
ODS 6. Agua limpia y saneamiento				X
ODS 7. Energía asequible y no contaminante				X
ODS 8. Trabajo decente y crecimiento económico				X
ODS 9. Industria, innovación e infraestructuras				X
ODS 10. Reducción de las desigualdades		X		
ODS 11. Ciudades y comunidades sostenibles				X
ODS 12. Producción y consumo responsables				X
ODS 13. Acción por el clima				X
ODS 14. Vida submarina				X
ODS 15. Vida de ecosistemas terrestres				X
ODS 16. Paz, justicia e instituciones sólidas				X
ODS 17. Alianzas para lograr objetivos				X

Descripción de la alineación del TFG con los ODS con un grado de relación más alto:

- **ODS 3. Salud y bienestar:** La tuberculosis es una importante enfermedad infecciosa, considerada como una de las 10 mayores causas de mortalidad en el mundo. En este sentido, la investigación en el inicio de la traducción de *Mycobacterium tuberculosis*, y los mecanismos que la diferencian de otros organismos, tiene un claro potencial de transferencia hacia el desarrollo de nuevos agentes terapéuticos que actúen de forma específica contra este patógeno. Esto ayudaría a establecer nuevos tratamientos y vacunas más efectivos, que mejoren la calidad de vida de los pacientes de tuberculosis, y reduzcan la incidencia y mortalidad de esta enfermedad, que en el año 2023 causó 1,3 millones de muertes en todo el mundo. En adición, los resultados derivados de este trabajo también se relacionan directamente con la resistencia antimicrobiana, una creciente amenaza para la salud pública, el crecimiento económico y la seguridad alimentaria, entre otros, que debe atajarse de manera prioritaria.
- **ODS 10. Reducción de las desigualdades:** Según la OMS, alrededor del 50% de los pacientes de tuberculosis, junto a sus familias, afrontan costes totales directos (tratamientos médicos) e indirectos (pérdidas de ingresos, incapacitación para trabajar) considerados como catastróficos (>20% de los ingresos anuales del hogar). Además, 2 de cada 3 casos en todo el mundo se concentran en 8 países, la mayoría pertenecientes a regiones de África y Asia en vías de desarrollo: India, Indonesia, China, Filipinas, Pakistán, Nigeria, Bangladesh y República Democrática del Congo. Esto indica la existencia de barreras económicas y financieras que impiden el acceso y mantenimiento del tratamiento contra la tuberculosis. La investigación en tuberculosis puede, sin duda, ayudar al desarrollo de nuevos tratamientos y métodos de prevención más asequibles y fácilmente accesibles para la sociedad, independientemente de las situaciones socioeconómicas de cada individuo.

1. INTRODUCCIÓN

1.1 Tuberculosis

La tuberculosis (TB) era, hasta la pandemia del COVID-19, la enfermedad infecciosa provocada por un solo microorganismo que más muertes en humanos causaba en el mundo. Actualmente, ocupa el segundo lugar, de acuerdo con el informe mundial de la tuberculosis emitido por la Organización Mundial de la Salud en 2023 (WHO, 2023). El principal organismo responsable de esta infección es la bacteria *Mycobacterium tuberculosis* (Mtb), también conocida como bacilo de Koch, que se transmite de persona a persona mediante los aerosoles generados por una persona infectada con la enfermedad activa. Tras la infección, casi siempre (85-90% de infecciones) la bacteria sobrevive en el hospedador en forma latente, no causando sintomatología alguna, y dando lugar a lo que comúnmente se refiere como tuberculosis latente. Se estima que una cuarta parte de la población mundial está infectada con tuberculosis latente (Houben & Dodd, 2016). En el resto de los casos (5-10% de infecciones), el sistema inmune no es capaz de controlar la bacteria, de forma que se desarrolla la enfermedad conocida como tuberculosis activa, aunque también se han descrito casos minoritarios en los que, tras la infección, el hospedador es capaz de eliminar el patógeno completamente (WHO, 2023). Para añadir mayor complejidad, la tuberculosis latente tiene la posibilidad de reactivarse a tuberculosis activa. Con todo esto, las estimas actuales para la incidencia de tuberculosis en 2022 están en 1,7 billones de casos de tuberculosis latente, 7,5 millones de casos nuevos de tuberculosis activa y 1,3 millones de muertes por TB (WHO, 2023).

La tuberculosis activa suele afectar a los pulmones (tuberculosis pulmonar), aunque en la llamada tuberculosis extrapulmonar puede presentarse en otros órganos diferentes como la piel, tracto digestivo, sistema nervioso central, etc. La tuberculosis pulmonar cursa con síntomas como tos crónica con esputo sanguinolento, fiebre, sudores nocturnos y pérdida de peso y se transmite por vía aérea, mediante gotas expulsadas por los pacientes al toser, estornudar o hablar (Smith, 2003).

Los métodos diagnósticos más comunes incluyen el cultivo de muestras de esputo y tejido, microscopía y amplificación de ácidos nucleicos del patógeno. En los casos positivos, el tratamiento estándar comprende una terapia cuádruple durante 2 meses, que consta de 4 fármacos: la isoniazida, la rifampicina, el etambutol y la pirazinamida; conocidos como fármacos de primera línea. Seguidamente, se pasa a una administración de rifampicina e isoniazida por otros 4 meses. En total, el tratamiento farmacológico dura unos 6 meses, requiere una gran constancia y adherencia por el paciente y tiene un índice de éxito de más del 80% (Suárez et al., 2019; Alsayed & Gunosewoyo, 2023).

Desgraciadamente, el tratamiento y el índice de éxito se complican en los casos de infecciones con cepas de Mtb drogorresistentes. De hecho, más de un cuarto de las muertes globales causadas por bacterias resistentes a antibióticos son causadas por Mtb (Schrader et al., 2021). En los casos en que existen cepas resistentes a rifampicina o isoniazida (los principales fármacos de primera línea), se habla de tuberculosis resistente (DR-TB), y de tuberculosis multirresistente (MDR-TB) cuando existe resistencia a ambos a la vez. En estas situaciones se recurre a fármacos de segunda línea como la bedaquilina, levofloxacino y clofamizina, que conllevan una toxicidad y duración del tratamiento mayores (desde 9 a 18 meses). También existen las cepas ultrarresistentes (XDR), resistentes a fármacos de primera línea, a cualquier fluoroquinolona y al menos a uno de segunda línea lo que complica enormemente las opciones de tratamiento y la prognosis de la enfermedad (WHO, 2023).

1.2 Características de *Mycobacterium tuberculosis*

M. tuberculosis pertenece al género *Mycobacterium*. Las micobacterias son conocidas por su peculiar pared celular rica en ácido micólico y otros lípidos. Esto las hace ácido-alcohol resistentes, lo cual impide la tinción Gram de sus membranas, en favor de otras técnicas de tinción como la de Ziehl-Neelsen (Madison, 2001). Además, les confiere una resistencia intrínseca a muchos antibióticos y condiciones ambientales hostiles.

Mtb en particular, es un patógeno intracelular facultativo, con tamaño entre 0,2-0,7 x 1-10 μm , ligeramente curvado, aerobio estricto, inmóvil, no formador de esporas ni cápsulas. Su genoma es circular, con un tamaño de 4,4 Mb y un alto contenido en GC (Cole et al., 1998). Exhibe un crecimiento especialmente lento (tiempo de generación de 16-20 h aproximadamente), además de que puede entrar en estado latente en huéspedes inmunocompetentes, retrasando su multiplicación hasta décadas (Dutta & Karakousis, 2014; Koch & Mizrahi, 2018). Una vez llega a los pulmones, Mtb es fagocitada por los macrófagos alveolares, pero cuenta con diversos mecanismos para inhibir la maduración del fagosoma y sobrevivir dentro de unas estructuras especializadas llamadas granulomas. Dentro del granuloma, la bacteria entra en un estado no replicativo de latencia que, como ya se ha dicho, puede retrasar su multiplicación durante décadas (Huang et al. 2019). En aquellos casos en los que el granuloma no puede contener la replicación de la bacteria, o se reactiva la multiplicación después de un periodo de latencia, el crecimiento bacteriano provoca la ruptura del granuloma y la progresión de la enfermedad a tuberculosis activa. Durante este ciclo de infección, Mtb se tiene que adaptar a una gran variedad de estreses celulares y alternar entre estados de replicación activos y no activos. Hoy en día, los mecanismos moleculares que regulan la adaptación de Mtb durante las diferentes etapas del ciclo de infección no son totalmente conocidos. Como veremos más adelante, se han descrito diferencias asociadas con el proceso de síntesis de proteínas en Mtb que podrían ser relevantes en este proceso. Además, el ribosoma, que es la maquinaria molecular que permite la síntesis de proteínas, es diana de hasta el 40% de antibióticos (Rodnina, 2018; Kumar et al., 2021).

1.3 Regulación del inicio de la traducción en procariontes y en Mtb

La traducción es el último paso de la expresión génica, en el que la secuencia codificante del ARN mensajero (ARNm) se traduce a la secuencia de aminoácidos de una proteína mediante el ribosoma. Comprende 3 fases principales (iniciación, elongación y terminación), seguidas de una cuarta etapa final de reciclaje del ribosoma. En bacterias, esto ocurre de manera co-transcripcional, de modo que el ARNm naciente de la ARN polimerasa (RNAP) se une al ribosoma, formando un complejo macromolecular conocido como "expresoma" (Kohler et al., 2017).

La iniciación es la etapa más regulada y que más energía consume. A grandes rasgos, la iniciación canónica consiste en el reclutamiento del ARNm al ribosoma, gracias al reconocimiento de la secuencia Shine-Dalgarno (SD) de unión al ribosoma (RBS) (que está presente en la región 5' no traducida (UTR) del ARNm) por parte de la secuencia anti-SD presente en el 16S ARN ribosómico (ARNr) del ribosoma. Pero como veremos a continuación, se han descrito mecanismos alternativos de iniciación de la traducción en procariontes, en función de las características del ARNm.

Entre los distintos tipos de ARNm procariontes, los más estudiados son aquellos que contienen la secuencia Shine-Dalgarno (SD) en su 5'UTR, los cuales se traducen mediante el mecanismo de iniciación canónico arriba mencionado. En mayor detalle, en este proceso el mensajero se alinea correctamente con el ribosoma gracias a la secuencia SD, presente unos 8-10 nucleótidos aguas arriba del codón de inicio (AUG), en la región 5' UTR. De esta manera, la subunidad pequeña del ribosoma 30S (SSU) es reclutada al RBS del mensajero gracias a la complementariedad entre la secuencia SD y la anti-SD, presente en el ARNr 16S (Shine & Dalgarno, 1974). El proceso es

promovido por los factores de iniciación (IF) IF1, IF2 e IF3 que, junto a la SSU y al ARN de transferencia cargado con N-formilmietionina (fMet-ARNt^{fMet}), forman el complejo de preiniciación 30S (PIC). En cuanto se une el ARNm, se estabiliza en el complejo de iniciación 30S, que unirá la subunidad grande del ribosoma 50S (LSU). En ese momento, se disocian los factores de iniciación y se forma el complejo 70S, listo para pasar a la etapa de elongación (Rodnina, 2018).

Por otro lado, se han descrito mecanismos alternativos para la regulación del inicio de la traducción. El primero de ellos está asociado a transcritos que carecen de región 5' UTR y, por tanto, de secuencia SD, conocidos como transcritos *leaderless*. Estos transcritos están presentes en muchas bacterias y arqueas, así como en eucariotas (Zheng et al., 2011). En *E. coli*, estos mensajeros siguen una vía de iniciación diferente, que parte de ribosomas 70S ensamblados a los que se une un ARNm *leaderless* (lARNm) con un grupo fosfato en el extremo 5' y un codón de inicio AUG cercano a este (Moll et al., 2004). Aunque en *E. coli* este tipo de iniciación por parte del ribosoma 70S en transcritos *leaderless* se lleva a cabo con una baja eficiencia, se han descrito situaciones, asociadas a respuesta a estrés o a presión por antibióticos, donde se produce una reprogramación traduccional que permite la traducción selectiva de transcritos *leaderless* (Leiva & Katz, 2022). La traducción *leaderless* también ha demostrado ser importante en otros organismos patógenos, como *Mycoplasma pneumoniae* (Weiner et al., 2000), lo que sugiere que podría ser responsable de ciertos mecanismos de resistencia y adaptación de los patógenos al ambiente hostil del huésped. El otro mecanismo alternativo de iniciación se asocia a transcritos que, a pesar de tener una región 5' UTR, o bien la secuencia SD está ausente o ésta carece de la secuencia consenso (secuencia SD débil). En estos casos, el inicio de la traducción está mediado por la proteína ribosomal bS1 que directamente reconoce regiones ricas en AT en la región 5' UTR permitiendo la formación del PIC (Boni et al., 1991; Komarova et al., 2005). Estos mecanismos de iniciación no canónica podrían ser mucho más protagónicos en Mtb respecto al organismo modelo *Escherichia coli*, siendo que en torno a un cuarto (25%) del transcriptoma de Mtb es expresado en forma de lARNm, frente al 1% de *E. coli* (Cortes et al., 2013; Shell et al., 2015); y un 36% adicional del transcriptoma está formado por transcritos con 5' UTR que carecen de SD conservada. Además, la proteína ribosomal bS1 de Mtb cuenta con un dominio C-terminal único que podría estar involucrado en la regulación de la especificidad de iniciación mediada por bS1 (Salah et al., 2009).

Se ha demostrado que en Mtb tanto los transcritos canónicos como no canónicos se traducen con eficiencias similares, ya sea en condiciones de crecimiento exponencial o en condiciones de escasez de nutrientes, las cuales se usan como modelo *in vitro* de latencia en el huésped (Sawyer et al., 2021). Esto sugeriría que, a diferencia de *E. coli*, Mtb es capaz de traducir con eficiencias similares transcritos canónicos y no canónicos utilizando la maquinaria de traducción presente en la célula. Asimismo, usando la técnica de perfilado de ribosomas, algunos autores han detectado diferencias en los patrones de reclutamiento de ribosomas a codones de inicio asociados con genes canónicos y no canónicos, indicativo de la co-existencia de mecanismos alternativos de iniciación (Sawyer et al., 2021). Los mismos autores también afirman que, bajo condiciones de escasez de nutrientes, los patrones de reclutamiento asociados con iniciación canónica se reducen significativamente, sugiriendo que, en estas condiciones los mecanismos de iniciación alternativos podrían ser más relevantes. Todo esto en conjunto con otras evidencias de esta plasticidad ribosomal como son la hibernación ribosomal, la estabilización de ribosomas 70S como formas de respuestas a estrés (Sawyer & Cortes, 2022), o la presencia de aspectos únicos asociados con el ribosoma de Mtb como el puente B9 entre ambas subunidades o la presencia de dos proteínas ribosomales bS22 y bL27, específicas de micobacterias (Yang et al., 2017), hacen necesario un mayor entendimiento de los mecanismos moleculares que regulan el inicio de la traducción en Mtb.

Sin duda, los avances en este campo podrían tener relevancia directa en el entendimiento de los mecanismos de adaptación al estrés, la búsqueda de nuevas dianas terapéuticas y otras soluciones que permitan atajar el problema mundial de la tuberculosis.

1.4 Factor de iniciación 2 (IF2)

El factor de iniciación 2 (IF2) es uno de los factores de iniciación de la traducción (IFs) en procariotas, junto a IF1 e IF3. Es una proteína GTPasa que se une al “hombro” de la SSU 30S en su forma unida a GTP, durante la formación del complejo de iniciación (30S IC). Además de reclutar el ARNt iniciador (fMet-tRNA^{fMet}), tiene una función homóloga al factor de iniciación eucariota eIF5b, contribuyendo a la formación del complejo de iniciación 30S IC y a la progresión hasta el 70S IC, mediante la unión de la LSU 50S y la disociación de los IFs gracias a la hidrólisis del GTP. Otras funciones descritas incluyen que IF2-GTP, junto al ARNt iniciador, ayudan al alineamiento del ARNm para establecer la pauta de lectura adecuada en la iniciación (Roy et al., 2018). Además, se ha descrito *in vitro* que la proporción entre IF2 e IF3 tiene un papel fundamental en la iniciación de transcritos *leaderless* (Grill et al., 2000).

IF2 está codificado en el genoma bacteriano por el gen *infB*. En Mtb tiene una longitud de 2700 pb, que codifican para 900 aminoácidos, resultando en una proteína con un peso molecular de 94 kDa, siendo el IF de mayor tamaño. El tamaño promedio de IF2 en procariotas oscila entre los 800-900 aminoácidos, aunque hay especies como *Thermus thermophilus* o *Bacillus subtilis* que tienen un tamaño de 500-700 aminoácidos ya que carecen de la extensión N-terminal desestructurada presente en otros géneros como *Escherichia*, *Pseudomonas*, *Streptomyces* o *Mycobacterium*. Es una proteína altamente conservada en todas las bacterias (Margus et al., 2007) compartiendo un 94% identidad de secuencia en Mtb respecto a la de *Mycobacterium smegmatis* (especie cercana, aunque de crecimiento más rápido y no patógena, que se utiliza comúnmente como especie modelo para Mtb) y un 48% de identidad con *E. coli* (Kumar et al., 2021). En ésta última especie, se ha descrito que la actividad catalítica reside principalmente en la mitad C-terminal de la proteína que es la más conservada y en la que se encuentra el dominio GTPasa y el dominio C2, que interactúa con el ARNt iniciador (Spurio et al., 2000; Laalami et al., 1991). El dominio N-terminal, por su parte, es la región con mayor variabilidad de secuencia entre especies y su papel en la traducción no está tan definido. Se cree que puede actuar aumentando la afinidad de IF2 por la unión al ribosoma y estabilizando los complejos 70S (Simonetti et al., 2013). Sin embargo, se ha demostrado que es prescindible *in vitro* e *in vivo* bajo condiciones óptimas de crecimiento y no está presente en organismos superiores (Caserta et al., 2006; Wienk et al., 2012).

Actualmente, se ha determinado la estructura a resolución atómica del IF2 de algunos organismos como *Thermus thermophilus* (Simonetti et al., 2008; Myasnikov et al., 2005), y se han resuelto complejos de iniciación 30S y 70S de varios organismos como *P. aureginosa* o *E. coli* (Basu et al., 2022; López-Alonso et al., 2017; Allen et al., 2005). No obstante, estas estructuras incluyen principalmente el dominio C-terminal y solo una pequeña parte del N-terminal de IF2. A día de hoy sigue sin haber estructuras que muestren al completo el dominio N-terminal incluyendo la extensión desestructurada, y menos todavía en un contexto de iniciación de la traducción.

Respecto a micobacterias, en los últimos años se han resuelto varias estructuras de ribosomas, principalmente unidos a antibióticos (Yang et al., 2017) o factores de hibernación (Kumar et al., 2024), pero por el momento no se han resuelto complejos de iniciación funcionales.

En resumen, parece claro que podrían existir diferencias importantes entre los mecanismos de inicio de la traducción ya descritos en *E. coli* y los que podrían estar operando en Mtb. Y del mismo

modo que en *E. coli*, donde la optimización de sistemas de traducción *in vitro* a partir de componentes purificados ha sido de vital importancia para entender los procesos traduccionales, se necesitan optimizar estos sistemas en Mtb para poder caracterizar a nivel molecular dichos mecanismos. Esto permitirá tener un mejor entendimiento de sus posibles implicaciones en la regulación de los procesos adaptativos de este importante patógeno humano.

2. OBJETIVOS

El presente trabajo de fin de grado se enmarca en el proyecto de investigación titulado *Elucidating translation initiation mechanisms in Mycobacteria* (IHRC22/00001) de la convocatoria Sello de Excelencia ISCIII-Health 2022, que se desarrolla en la Unidad de Control de la Regulación Génica en Patógenos del Instituto de Biomedicina de Valencia (IBV-CSIC), dirigida por la Dra. Teresa Cortés. El objetivo general del proyecto es comprender en profundidad la iniciación de la traducción en micobacterias, caracterizando a nivel molecular los diferentes mecanismos que regulan la iniciación canónica y no canónica. Para ello, se pretende reconstituir un sistema de traducción *in vitro* que permitirá llevar a cabo ensayos funcionales y estructurales con los que obtener más información sobre la traducción en Mtb, ampliando los conocimientos en este campo y abriendo puertas a nuevas posibilidades terapéuticas.

Antes de mi llegada al laboratorio, el equipo ya había producido los factores IF1 e IF3, faltando solo el IF2 para empezar con experimentos de microscopía electrónica, como la preparación de rejillas de tinción negativa con complejos de preiniciación 30S. Así, el objetivo principal de este trabajo de fin de grado es la purificación del factor de iniciación 2 (IF2) de Mtb, tanto en su versión completa como en su versión truncada (IF2tr). Para ello se pretende expresar ambas versiones de forma heteróloga en *E. coli* para proceder con su posterior purificación.

3. MATERIALES Y MÉTODOS

3.1 Diseño de los constructos de IF2 e IF2tr y clonación en pTYB21

3.1.1 Constructos y descripción del vector pTYB21

La secuencia del gen del factor de iniciación 2 (IF2) de *Mycobacterium tuberculosis* se utilizó para diseñar los constructos génicos que permitieron la sobreexpresión y posterior purificación de la proteína en el organismo modelo *Escherichia coli*. Mtb parece tener un dominio N-t desestructurado más largo que otras especies, sin plegamiento secundario prácticamente (Figura 1). Para facilitar los objetivos de expresión de la proteína en este trabajo, se diseñó una versión truncada (IF2tr), que carecía de los aminoácidos 1-280 de este dominio N-terminal, ya que se preveía que la región desestructurada podía originar ciertos problemas de agregación o insolubilidad por un plegamiento más complejo. Esto resultó en una secuencia de 1865 pb, 620 aminoácidos y unos 68 kDa aproximadamente.

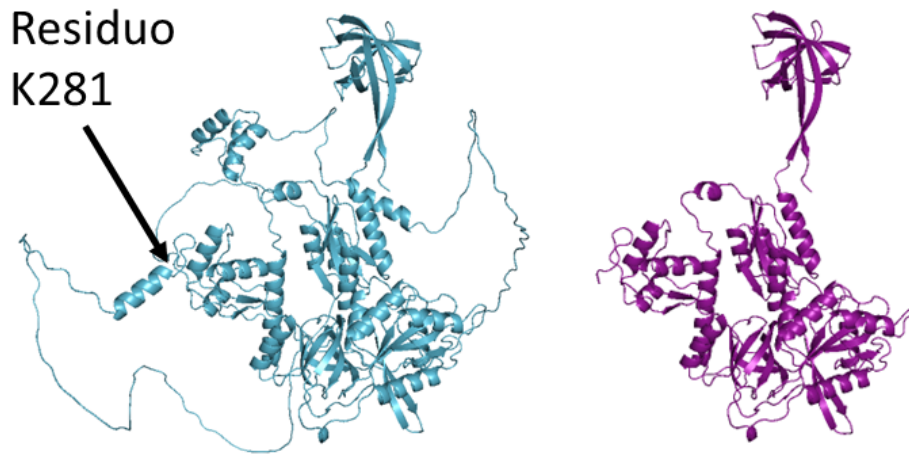


Figura 1. Predicción en AlphaFold de la estructura tridimensional del factor de iniciación 2 (IF2) de *Mycobacterium tuberculosis* en su versión completa (azul) y truncada (morado). La flecha indica la lisina 281, hasta donde se realizó el truncamiento del dominio N-terminal (entrada P9WKK1 UniProt).

Los plásmidos se adquirieron comercialmente a la empresa ProteoGenix (Schiltigheim, Francia), diseñados para contener el gen de IF2 o IF2tr en el vector de expresión pTYB21 y optimizando los codones para su expresión en *E. coli*. Las principales características de este vector son que otorga resistencia a la ampicilina como marcador de selección, que la expresión del constructo está bajo el control del operón *lac* y se inicia a partir del promotor T7 (por lo que la cepa empleada debe expresar la T7 polimerasa de algún modo, como la *E. coli* BL21(DE3)), y que la proteína de fusión contiene una etiqueta para su purificación en el extremo N-terminal que consiste de dos regiones inteína (Sce VMA 5' y 3') con un dominio de unión a quitina (CBD) incluido entre ambas regiones (Figura 2).

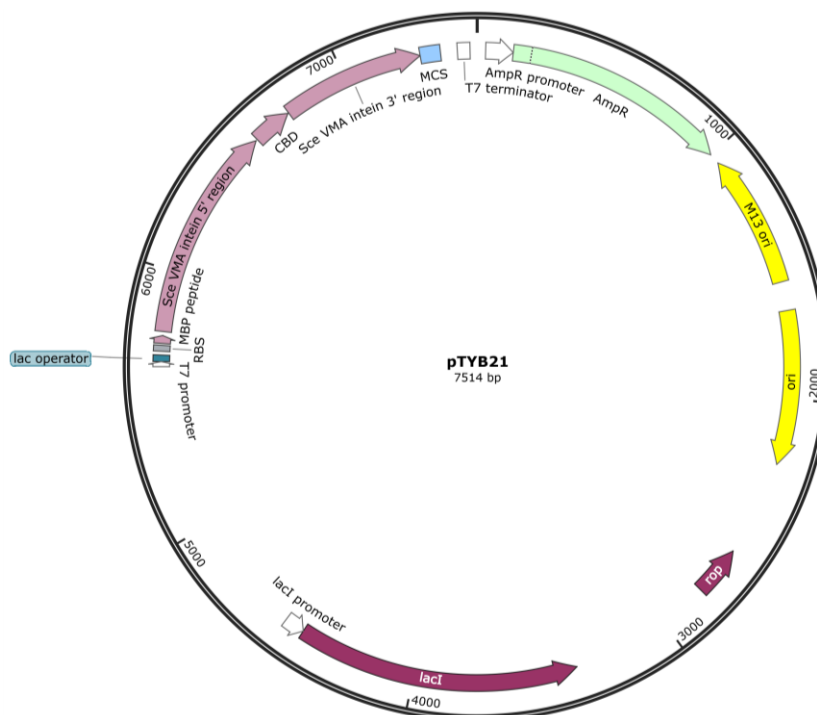


Figura 2. Mapa del vector pTYB21. Se muestran secuencias de interés como la etiqueta inteína, con el dominio de unión a quitina (CBD) incluido, el operón *lac* y el marcador de selección AmpR. Obtenido de SnapGene.

Este tipo de etiqueta permite la purificación de la proteína clonada mediante el sistema IMPACT™ (*Intein Mediated Purification with an Affinity Chitin-binding Tag*) (New England Biolabs®, NEB), basado en una etapa de cromatografía de afinidad con resina de quitina, donde la proteína queda retenida gracias al dominio CBD (*chitin binding domain*), seguida de una incubación en presencia de compuestos tior, como DTT, que activa la autoescisión característica de las inteínas que participan en procesos de *splicing* (ver apartado 3.4.1 para más detalles del protocolo). Esto permite separar la proteína diana de la etiqueta de afinidad y purificarla en un solo paso cromatográfico, sin emplear una proteasa dirigida (Figura 3).

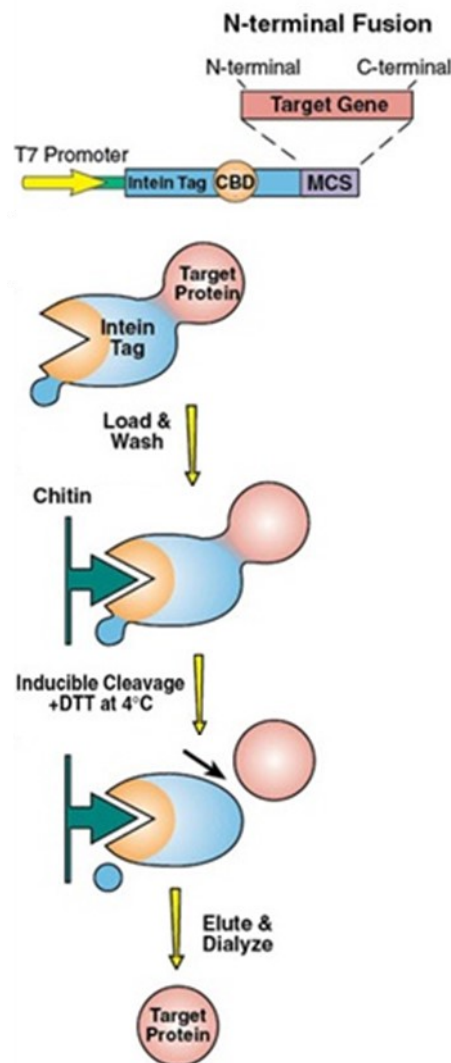


Figura 3. Flujo de trabajo para la expresión y purificación de proteínas por el sistema IMPACT (NEB). Adaptado de New England Biolabs (s.f.).

3.1.2 Transformación en *Escherichia coli*

Con anterioridad a mi llegada al laboratorio, se había intentado expresar IF2 y IF2tr en células de expresión *E. coli* BL21(DE3), sin embargo, en ningún caso se había conseguido obtener las proteínas de forma soluble. Por ello, tras mi incorporación, se probaron otras alternativas como co-expresión con pLysS o pGroESL. El plásmido pLysS viene incluido en determinadas cepas comerciales y codifica para la lisozima del fago T7, que inhibe la ARN polimerasa T7, reprimiendo así la expresión constitutiva del plásmido pTYB21 (Studier, 1991). Esto resulta interesante ya que cabe la posibilidad de que el IF2 de micobacterias resulte tóxico al interferir con el IF2 nativo de *E. coli* por su homología

de secuencia relativamente alta, del 48% (Kumar et al., 2021). Por otro lado, pGroESL permite la sobreexpresión del sistema de chaperoninas GroEL/GroES de *E. coli*, que puede facilitar el correcto plegamiento de la proteína expresada, dado que contiene un dominio desestructurado (Morizono et al., 1997; Ishii, 2017). En conjunto, se transformaron ambas versiones de pTYB21 en células de la cepa BL21(DE3) pLysS quimiocompetentes y también se realizó una cotransformación de pTYB21 y pGroESL (proporcionado por el Dr. Gatenby, DuPont; Wilmington, EE. UU.) en células BL21(DE3) electrocompetentes, manteniendo condiciones de asepsia (trabajo cerca de llama, material estéril, etc.) en todo momento.

Las células electrocompetentes se transformaron por electroporación, añadiendo 100 ng de plásmido (pTYB21) a alícuotas de 50 μ L de células BL21(DE3) descongeladas. En el caso de la cotransformación con GroEL/ES, se añadieron 200 ng de cada vector, al ser menos probable que ambos sean incorporados por la célula. La mezcla se transfirió a cubetas de electroporación, para aplicarle un pulso eléctrico de 1700 V. Por otro lado, 100 μ L de células quimiocompetentes se transformaron con 100 ng de plásmido por choque térmico, incubando la mezcla anterior en hielo por 30 minutos. A continuación, se sometió a un choque térmico de 45 segundos a 42 °C, seguido de dos minutos en hielo. Tras terminar cualquiera de los dos protocolos de transformación, se añadieron inmediatamente 200 μ L de medio LB fresco, pasando las células a un tubo Eppendorf e incubándolas 45-60 minutos a 37 °C en agitación (180 rpm) para su recuperación.

Una vez recuperadas de la transformación, las células transformadas se sembraron en placas de agar LB con los antibióticos correspondientes (100 μ g/ml ampicilina para BL21 y 100 μ g/ml ampicilina + 35 μ g/mL cloranfenicol para BL21 pLysS y BL21 GroEL/ES) y se incubaron en estufa a 37 °C toda la noche.

Adicionalmente, los vectores se transformaron también en células de amplificación *E. coli* DH5 α electrocompetentes, para la expansión clonal de los vectores y su preservación.

3.2 Reclonación de IF2tr en pLIC-SGC1

3.2.1 Descripción del vector pLIC-SGC1

Tal y como se discutirá más adelante (ver Resultados), los resultados obtenidos al purificar el constructo pTYB21-IF2tr no fueron los deseados, por lo que se decidió reclonar el gen a un vector con una etiqueta alternativa, que ofreciera un mejor plegamiento de la proteína de fusión. En este caso, se eligió pLIC-SGC1 como vector de destino, que se caracteriza por tener una etiqueta de afinidad basada en una cola de 6 histidinas en el extremo N-terminal (6xHis tag), seguida de una secuencia de corte de la proteasa TEV (Figura 4). El plásmido tiene un tamaño de 7505 pares de bases, y presenta el operón *lac* para controlar la expresión del constructo, y el gen *ampR* como marcador de selección.

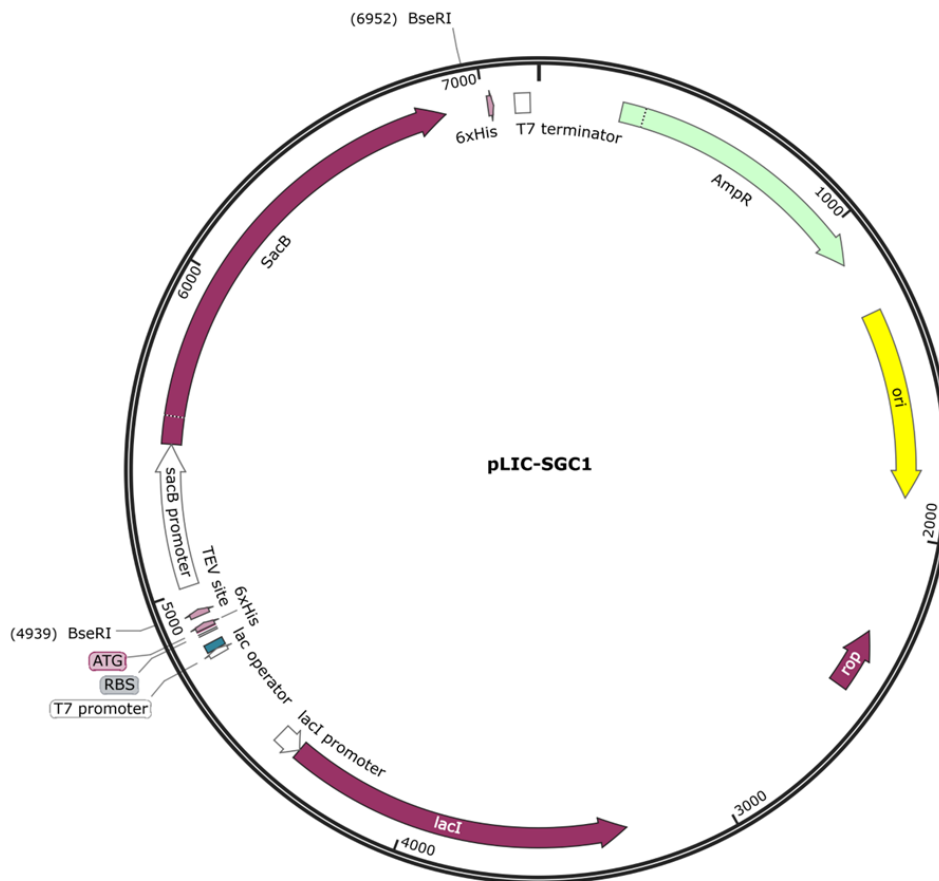


Figura 4. Mapa del plásmido pLIC-SGC1. Se muestran secuencias relevantes como la etiqueta 6xHis, los sitios de corte de BseRI a los lados del gen SacB, y el marcador de selección AmpR. Obtenido de SnapGene.

Este vector permite la purificación del constructo expresado por el sistema IMAC (*Immobilized Metal Ion Affinity Chromatography*). Éste consiste en una cromatografía de afinidad empleando una columna de agarosa con algún ion metálico divalente inmovilizado, generalmente Ni^{2+} , que interacciona con la cola de histidinas y retiene así la proteína de fusión. Tras lavar la columna, se realiza una elución en gradiente aumentando progresivamente la concentración de imidazol, el cual desplaza a la histidina en su unión a la columna. Así, se eluyen las especies en orden de menor a mayor grado de interacción con la resina de níquel, incluyendo la proteína de interés (Lee et al., 2007). La etiqueta 6xHis puede eliminarse por proteólisis aprovechando la secuencia TEV, pero en este caso no se consideró necesario porque no afectaba en las etapas posteriores.

3.2.2 Clonación independiente de ligación (LIC)

Otro aspecto clave del vector pLIC-SGC1 es que permite clonar el gen de IF2tr por el método de clonación independiente de clonación o LIC (Aslanidis & de Jong, 1990). Este sistema resulta más rápido y eficaz comparado con las estrategias de clonación habituales, además de evitar el uso de las enzimas de restricción-ligación propias de tales protocolos (Bonsor et al., 2006). De esta manera, el método LIC aprovecha la actividad exonucleasa 3'→5' de la T4 ADN polimerasa, creando extremos cohesivos entre el inserto y el vector para facilitar la alineación de ambas moléculas de ADN y su posterior ligación, una vez transformadas, en el interior celular (Kroemer, s.f.).

El primer paso fue la amplificación del gen IF2tr usando los cebadores indicados en la Tabla 1, a partir del vector pTYB21 extraído de células DH5α gracias al kit de miniprep QIAprep Spin (QIAGEN), siguiendo las instrucciones del fabricante.

Tabla 1. Secuencias de los oligonucleótidos empleados como cebadores. La pareja de cebadores FW_IF2tr_pLIC y RV_IF2tr_pLIC se diseñaron para la amplificar el inserto. La secuencia en negrita corresponde al codón de inicio o parada de cada cebador. La secuencia subrayada corresponde a los *overhangs* o alas que se añaden a los extremos del inserto amplificado, para su posterior ligación con el vector destino (protocolo LIC). Adicionalmente, se diseñó un tercer cebador (FW_IF2tr_seq), que hibrida en el cuerpo del gen, para permitir la secuenciación del gen. Por otro lado, se utilizó otra pareja de cebadores para el análisis de las colonias transformadas, que consistía en un cebador directo (T7_FW) y reverso (T7_RV) que hibridan en el promotor y terminador T7, respectivamente.

Cebador	Secuencia (5'→3')
FW_IF2tr_pLIC	<u>TACTTCCAATCC</u> ATGA AGCGTCAGAAACGC
RV_IF2tr_pLIC	<u>TATCCACCTTTACTTT</u> ATT TCTCGAACTGCGGATG
FW_IF2tr_seq	CTGGTCCGGAAGAATTCTG
T7_FW	TAATACGACTCACTATAG
T7_RV	CCGCTGAGCAATAACTA

El inserto, con un tamaño de 1800 pb, se amplificó por PCR empleando la Phusion™ ADN polimerasa de alta fidelidad (Thermo Fisher Scientific) y los cebadores FW_IF2tr_pLIC (directo) y RV_IF2tr_pLIC (reverso) (Tabla 1) que contenían unas alas (secuencias subrayadas en la Tabla 1) complementarias a la secuencia del vector. Antes de la PCR final, se optimizaron las condiciones de temperatura, de tampón de reacción (probando tanto el tampón High-Fidelity, de uso común; o el tampón GC, para genes ricos en GC), así como de adición o no de dimetilsulfóxido (DMSO) al 3%. De este modo, se determinaron las condiciones óptimas para realizar la PCR (Tabla 2).

Tabla 2. Composición de la mezcla de reacción optimizada de la PCR de inserto.

Componentes	Concentración final	Volumen (µL)
Tampón GC 5X	1X	10
dNTPs 10 mM	200 µM (cada uno)	1
Cebador FW 10 µM	0,5 µM	2,5
Cebador RV 10 µM	0,5 µM	2,5
DMSO	3%	1,5
Phusion™ ADN polimerasa de alta fidelidad 2 U/µL	0,02 U/µL	0,5
ADN (plásmido pTYB21)	25 ng	2,5
H ₂ O Milli-Q	-	Hasta 50 µL

Para la PCR se utilizó un programa de amplificación en el termociclador que constaba de: etapa inicial de desnaturalización de 30 segundos a 95 °C; 30 ciclos de desnaturalización 10'' a 98 °C, *annealing* 30'' a T_m = 59 °C (mejores resultados en las pruebas de T_m) y elongación 1'30'' a 72 °C; y etapa final de elongación final 1' a 72 °C. Los productos de PCR se cargaron en un gel de agarosa al 1% para comprobar el tamaño resultante.

El siguiente paso fue digerir el vector pLIC-SGC1, obtenido por miniprep de células DH5α, para linealizarlo y eliminar el gen SacB que portaba. Para ello se digirieron 5 µg de ADN plasmídico añadiendo 5U/µg de la enzima de restricción BseRI (NEB). Se prepararon 50 µL de reacción, disolviendo el ADN y la enzima en 5 µL del tampón incluido (concentración final 1X) y H₂O Milli-Q estéril hasta completar el volumen. La reacción tuvo lugar a 37 °C durante 3h, y luego se incubó 20 min a 80 °C para inactivar la enzima de restricción. Se tomó una alícuota para observar el resultado en gel de agarosa al 1%.

Después de la comprobación, se preparó otro gel con un pocillo de gran tamaño, donde se cargó el volumen restante de la reacción de digestión. Tras la migración electroforética, la banda correspondiente al vector pLIC digerido (cerca de 5000 pb) se cortó y purificó mediante el GenElute™ Gel Extraction Kit (Sigma-Aldrich). A su vez, el producto de PCR del gen IF2tr también se purificó con el GenElute™ PCR Clean-Up Kit (Sigma-Aldrich), siguiendo el protocolo recomendado. A continuación, se midió la concentración de ADN de ambas muestras mediante NanoDrop y se procedió a realizar el tratamiento con la T4 polimerasa (NEB). Para ello, se prepararon reacciones en paralelo que contenían tampón r2.1 1X, dNTPs 25 mM (dGTP para el vector y dCTP para el inserto), DTT 50 mM, 3 U/μg ADN de T4 polimerasa, 1 μg ADN (de vector o inserto según corresponda) y H₂O Milli-Q hasta completar 50 o 20 μL según la reacción del vector o el inserto, respectivamente.

El tratamiento se realizó a 22 °C durante 30 min, seguido de una inactivación de 20 min a 75 °C. De esta manera la T4 polimerasa digirió los extremos y añadió un nucleótido colgando, que facilita la ligación de las muescas formadas al alinearse ambas moléculas de ADN. Finalmente, el vector y el inserto tratados se incubaron juntos durante 10 minutos a temperatura ambiente, en un ratio vector:inserto de 1:2 o 1:3 (v/v). La mezcla se transformó por choque térmico en *E. coli* DH5α quimiocompetentes siguiendo el mismo procedimiento descrito anteriormente. Posteriormente, se plaqueó todo el volumen en placas de agar LB+ampicilina y se incubaron 16 horas a 37 °C.

3.2.3 Análisis de las colonias transformadas

Las colonias crecidas en el medio selectivo se picaron con puntas estériles para realizar PCRs de colonia y determinar si habían incorporado el constructo deseado o solo el vector original. Con las mismas colonias, se volvieron a sembrar placas en estría para perpetuar el cultivo y se inocularon en 10 mL medio LB y ampicilina 100 μg/mL para una posible miniprep si la PCR saliera positiva.

La amplificación tuvo lugar empleando los cebadores del promotor y terminador T7, los cuales hibridan a sendos lados del gen insertado (Tabla 1). Se prepararon reacciones de 25 μL totales con 12,5 μL Green MasterMix 2X (NZYTech), 1,25 μL de cada cebador (10 μM), 1 colonia transformada y H₂O Milli-Q hasta completar el volumen. El programa de PCR consistió de: una etapa de desnaturalización inicial de 3 minutos a 95 °C; 30 ciclos de desnaturalización 30'' a 94 °C, *annealing* 30'' a 59 °C y elongación 1'30'' a 72 °C; y etapa final de elongación final 5' a 72 °C.

Como los resultados no eran claros (ver apartado Resultados), se llevó a cabo una segunda PCR de colonia utilizando en este caso los cebadores que se emplearon para amplificar el inserto (FW_IF2tr_pLIC y RV_IF2tr_pLIC, Tabla 1). Debido a que los resultados seguían siendo ambiguos, se optó por realizar minipreps a partir de los cultivos inoculados, extrayendo el ADN plasmídico con el kit QIAprep Spin y repitiendo la PCR. Esta vez, se hicieron dos combinaciones de *primers*: una reacción de PCR con T7_FW y IF2tr_pLIC_RV, y otra reacción con FW_IF2tr_seq (un *primer* diseñado para secuenciar el gen que solo hibrida con la secuencia de IF2tr, no con ninguna zona del vector) y T7_RV (Tabla 1). Los plásmidos de las colonias positivas se mandaron a secuenciar al servicio de secuenciación de MacroGen (Madrid) usando los cebadores de T7 y el cebador FW_IF2tr_seq (Tabla 1) y se analizaron para comprobar que habían incorporado el inserto adecuadamente. Con este fin, se empleó el software de alineamiento de secuencias EMBOSS Needle (Madeira et al., 2024) para alinear las secuencias obtenidas contra la secuencia del plásmido pLIC-SGC1 con el inserto de IF2tr (clonación realizada *in silico* conociendo ambas secuencias). Una vez revisadas las secuencias de los vectores aislados, los clones positivos se transformaron en células de expresión BL21(DE3) mediante electroporación, para comenzar con las pruebas a pequeña escala (apartado 3.3.3).

3.3 Expresión y purificación a pequeña escala

3.3.1 Pruebas de expresión y solubilidad con pLysS y GroEL/ES

En primer lugar, se realizaron pruebas a pequeña escala para determinar las condiciones óptimas de expresión de IF2 en las células BL21(DE3) pLysS y BL21(DE3) GroEL/ES. Para ello, se inocularon precultivos de 5 mL de medio LB con una colonia transformada, a los que se añadió la concentración correspondiente de antibiótico (cloranfenicol a 35 µg/mL y ampicilina a 100 µg/mL respectivamente). Ambos tipos celulares crecieron en incubador a 37 °C por 16 horas (*overnight*). Al día siguiente, se inocularon 20 mL de medio LB más los antibióticos correspondientes con 0,5 mL de precultivo, realizando así una dilución 1/40 para lograr una OD₆₀₀ inicial de entre 0,05-0,1. El cultivo se creció a 37 °C hasta una OD₆₀₀ de 0,6-0,8; momento en que se realizó la inducción con IPTG a una concentración final de 0,3mM. Se probaron tres condiciones de inducción con cada tipo celular: 37 °C durante 3 horas, 25 °C y 16 °C, ambas durante 16 horas. Pasado el tiempo de inducción, se centrifugaron los cultivos a 4000 rpm (Centrífuga 5810 R Eppendorf™, rotor A-4-62) 10 minutos, descartando el sobrenadante y resuspendiendo el sedimento celular o pellet en 3 mL de tampón de lavado (HEPES 20 mM pH 7,4, NaCl 150 mM) para volver a centrifugar en las mismas condiciones. Finalmente, los pellets obtenidos tras descartar el último sobrenadante se congelaron a -20 °C. Durante el proceso se tomaron alícuotas de 600 µL previas a la inducción ("PRE") y posteriores ("POST") para su carga en un gel SDS-PAGE en el que observar la expresión de IF2.

Previamente a la purificación se llevó a cabo una prueba de solubilidad, que consiste en una lisis celular y una separación del extracto celular en fracción soluble e insoluble, para conocer el estado de la proteína expresada. Todos los pasos se realizaron manteniendo condiciones de frío, mediante baños de hielo o trabajo en cámara fría, para evitar la desnaturalización de las proteínas del extracto. Primeramente, se disolvieron los pellets congelados en 750 µL de tampón de lisis para proteína (HEPES-KOH pH 8, 20 mM, KCl 0,5 M, EDTA 1 mM, 0,1% Triton X-100) con 1 pastilla/50 mL de cóctel de inhibidor de proteasas cComplete™ EDTA free Protease Inhibitor (Roche) y se transfirieron a tubos Eppendorf de 1,5 mL. Las células resuspendidas se lisaron por sonicación en un procesador ultrasónico VCX 750 (Sonics & Materials), empleando ciclos de 5 segundos ON y 15 segundos OFF durante 5 minutos. Por último, se centrifugaron a 13000 rpm, 5 minutos a 4 °C (Centrífuga Sorvall Legend Micro 21R, Thermo Fisher Scientific). La solubilidad de IF2 o IF2tr se evaluó de dos maneras. Por un lado, se cargó un gel SDS-PAGE con volúmenes iguales de las muestras de pellet y sobrenadante para observar la intensidad relativa de las bandas del gel. Por otro lado, se calculó el ratio pellet:sobrenadante a partir de las intensidades de absorción a 280 nm medidas en NanoDrop o Bradford (ver apartado 3.5.1).

3.3.2 Pruebas de expresión, solubilidad y purificación con glucosa 1%

Los resultados de las pruebas con pLysS y GROEL/ES no fueron los esperados (ver apartado de Resultados), de manera que se probó otra estrategia diferente basada en añadir glucosa al 1% en los precultivos a pequeña escala, reprimiendo así la expresión constitutiva de IF2 o IF2tr bajo el control del operón *lac*. De este modo, se prepararon precultivos de 5 mL de medio LB, ampicilina 100 µg/mL, 1% glucosa y una colonia BL21(DE3) con pTYB21-IF2 o pTYB21-IF2tr. Tras crecer *overnight* a 37 °C en agitación (180 rpm), se usaron 0,5 mL de precultivo para inocular cultivos de 20 mL LB y ampicilina 100 µg/mL. Se dejaron crecer en las mismas condiciones hasta una OD₆₀₀ de 0,6-0,8 para llevar a cabo la inducción y las pruebas de temperatura según el apartado 3.3.1.

La prueba de solubilidad también se realizó siguiendo el apartado anterior, mejorando los resultados previos, por lo que todos los precultivos se suplementaron con glucosa al 1% a partir de este momento.

Tras comprobar que las proteínas eran solubles, se ejecutó una purificación a pequeña escala en una columna Mini Bio-Spin (Biorad®), con 250 µL de resina de quitina (NEB). El primer paso fue equilibrar la columna con 10 volúmenes de columna o CV (2,5 mL) de tampón de lisis para inteína. A continuación, se añadieron los sobrenadantes resultantes de la sonicación, tomando una alícuota previa para su observación en el gel, y se centrifugaron las columnas 2 minutos a 1000 rpm. Se recogió el líquido no retenido (NR) y se lavaron las columnas con 6 CV de tampón de lavado para inteína o *wash* (HEPES 20 mM pH 8, KCl 1 M, Triton X-100 al 0,1%, EDTA 1 mM,), centrifugando de nuevo. Los siguientes lavados se realizaron con 1 CV de tampón de elución o *cleavage* para inteína (HEPES 20 mM, KCl 0,5 M, EDTA 1 mM) y 1 CV de tampón de elución suplementado con 50 mM DTT. Finalmente, se incubaron las columnas en 2 CV de tampón de elución con 50 mM DTT a 4 °C *overnight* para provocar la autoescisión de la etiqueta inteína y soltar IF2 de la columna. Al día siguiente, se eluyeron las columnas por gravedad y se observó el resultado en SDS-PAGE.

3.3.3 Prueba de expresión, solubilidad y purificación de pLIC-IF2tr

Del mismo modo que con los otros constructos, previamente a la purificación de pLIC-SGC1-IF2tr se llevaron a cabo diversas pruebas a pequeña escala para testar las posibles condiciones.

Para ello, se realizaron pruebas de expresión y solubilidad a pequeña escala tal y como se ha detallado en los apartados anteriores, pero utilizando un tampón de lisis para histidinas (HEPES 20 mM pH 7,5, imidazol 20 mM, KCl 500 mM, Triton X-100 al 0,1%, DTT 1 mM). En este caso, también se añadió 1% de glucosa a los precultivos. Una vez comprobada la solubilidad de IF2tr, se procedió a la prueba de purificación utilizando en este caso columnas His SpinTrap™ (Cytiva). Previamente la columna se equilibró con 600 µL de Tampón His-A (20 mM HEPES pH 7,5, 20 mM imidazol, 500 mM KCl, 1 mM DTT) centrifugando 30 s a 100 g. Después se añadieron los sobrenadantes, se lavó la columna con tampón His-A para eliminar proteínas no unidas y se eluyeron las retenidas en 2 pasos, aplicando 600 µL tampón His-B (20 mM HEPES pH 7,5, 500 mM imidazol, 150 mM KCl, 1 mM DTT) y centrifugando en cada etapa. De nuevo, se observó el resultado de la purificación en un gel SDS-PAGE.

3.4 Expresión y purificación a gran escala

3.4.1 Expresión a gran escala y purificación por el sistema IMPACT

Tras definir las condiciones óptimas de expresión a pequeña escala, se repitió el proceso a gran escala, con mayores volúmenes de cultivo. En cuanto a la expresión, se prepararon precultivos de 25 mL LB, ampicilina 100 µg/mL, 1% glucosa y una muestra de los glicerizados conservados con el constructo de pTYB21-IF2tr. Se dejaron crecer *overnight* a 37 °C en agitación (180 rpm) y al día siguiente, se inocularon en 1 L de cultivo logrando una OD₆₀₀ inicial de 0,05-0,1. Los cultivos crecieron a 37 °C en agitación hasta una OD₆₀₀ = 0,6 aproximadamente, momento en que se tomaron las alícuotas "PRE", se indujeron con IPTG a 0,3 mM y se incubaron toda la noche a 16 °C, tras haber elegido esta temperatura como mejor condición de expresión en la pequeña escala. A la mañana siguiente se recogieron los cultivos (tomando la alícuota "POST") centrifugándolos a 4000 rpm, durante 30 minutos y a una temperatura de 6 °C en la Centrífuga Sorvall Lynx 6000 con el rotor F9-6x1000 LEX (Thermo Fisher Scientific). Los sedimentos celulares se lavaron con 25 mL de tampón lavado y se centrifugaron a 6 °C y 4000 rpm durante 10 min en la centrífuga 5810 R (Eppendorf™).

Más tarde, se llevó a cabo la lisis y purificación de las proteínas, manteniendo la temperatura lo más baja posible en todos los pasos. Los pellets lavados se resuspendieron en 50 mL de tampón de lisis para inteína con inhibidor de proteasas (1,5 pastillas/50 mL) y la muestra se sonicó 30 min (ciclos 15 s ON, 45 s OFF) a una amplitud del 30%. Tras tomar las muestras de pellet y sobrenadante

para el gel de solubilidad, la muestra se centrifugó 45 min a 12000 rpm y una temperatura de 6 °C y la fracción soluble se filtró con filtros Filtropur S 0,2 µm (Sarstedt) para eliminar contaminantes y agregados. Posteriormente, la muestra se cargó por gravedad en columnas que contenían 5ml de resina de quitina. Tras pasar el volumen no retenido (fracción NR), se lavaron con 6 CV (30 mL) de tampón *wash* (W), 1 CV de tampón *cleavage* sin DTT (C) y 1 CV del de tampón *cleavage* suplementado con DTT 50 mM (C_{DTT}). Por último, se incubaron en 2 CV de tampón *cleavage* con DTT durante toda la noche a 4 °C. De cada etapa se tomaron muestras para su observación en gel y su cuantificación relativa por el método de Bradford.

Al día siguiente, se recogieron los eluidos, se mezclaron al ser la misma condición, y se concentraron utilizando concentradores Amicon® Ultra – 15 mL, con tamaño de poro de 50 kDa (Millipore). Tras ello, se centrifugaron las muestras de nuevo (13000 rpm, 10 min) y se filtraron para eliminar agregados e impurezas. Posteriormente se realizaron inyecciones de 0,5 ml de muestra en una columna de cromatografía de exclusión molecular Superdex200 10/300 (GE Healthcare), acoplada a un equipo de cromatografía líquida FPLC NGC Chromatography System (Bio-Rad), configurado con un flujo de 0,5 mL/min de tampón *storage* (HEPES 20 mM pH 7,5, KCl 150 mM, MgCl₂ 5mM, 10% glicerol, DTT 2 mM). Se recogió el eluido en fracciones de 500 µL y, tras analizar el cromatograma, las fracciones de interés se cargaron en un gel SDS-PAGE y se cuantificaron por Bradford. Finalmente, aquellas que contenían la proteína de interés se concentraron hasta el máximo volumen posible, se centrifugaron 10 minutos a 14000 rpm para eliminar posibles agregados, y se congelaron usando nitrógeno líquido. Las alícuotas se almacenaron a -80 °C para su conservación y futuros usos.

3.4.2 Expresión y purificación de IF2tr con chaperonas endógenas

En un primer intento de purificar IF2tr mediante el sistema IMPACT, se vio que la proteína estaba agregada (ver apartado Resultados). Para tratar de evitar la agregación, se probó a expresar los cultivos a gran escala en presencia de chaperonas endógenas para intentar que la proteína se plegara correctamente y aumentara su solubilidad (Kolaj et al., 2009). Para ello, antes de la inducción con IPTG, se sometieron a las células a diferentes tipos de estrés. En primer lugar, se prepararon precultivos de 12,5 mL LB y matraces de 1 L con 500 mL LB de la misma manera que en 3.4.1. Los cultivos se crecieron a 37 °C hasta una OD₆₀₀ de 0,8. En este momento se tomaron alícuotas “PRE” y se procedió a inducir el estrés celular. Todos los matraces se sometieron a un choque de frío, por incubación en hielo durante 1 hora. Seguidamente, se separaron en tres condiciones distintas: un matraz se mantuvo sin más estrés, otro recibió estrés oxidativo por la adición de etanol al 2%, y el último se sometió a un estrés osmótico al añadir sorbitol a una concentración final 0,5 M. Posteriormente, se indujo la expresión de IF2tr con IPTG 0,3 mM y se incubaron todos los matraces a 16 °C, 180 rpm durante 16 horas. Al día siguiente, los cultivos se recogieron siguiendo el protocolo descrito. Los pellets obtenidos se lavaron con 10 mL de tampón de lavado y se procedió a la lisis y purificaron siguiendo el mismo procedimiento descrito en el apartado 3.4.1.

3.4.3 Expresión y purificación de pLIC-IF2tr por IMAC

Para la expresión a gran escala de pLIC-IF2tr, se siguió el mismo procedimiento de expresión que se ha descrito anteriormente (ver apartado 3.4.1), pero en este caso se indujeron los cultivos durante 16 h a 25 °C (mejor condición obtenida tras los experimentos de expresión a pequeña escala).

Tras la centrifugación y lavado, los pellets se resuspendieron en 65 mL de tampón de lisis para histidinas (ver apartado 3.3.3) y una pastilla de inhibidor de proteasas. A continuación, se sonicaron 30 minutos según el procedimiento habitual de gran escala, a 30% de amplitud. Posteriormente, se centrifugaron a 12000 rpm, durante 40 min, y una temperatura de 6 °C. Los sobrenadantes se

filtraron y cargaron en una columna con resina de níquel HisTrap™ FF (GE Healthcare) de 5 mL de volumen (CV), a un flujo de 3 mL/min utilizando la bomba peristáltica Dynamax RP-1 (Rainin™). Una vez recogido no retenido, la columna se acopló al equipo FPLC para realizar el lavado con 10 CV de tampón His-A (ver 3.3.3) y la elución con tampón His-B. Se efectuó una elución en gradiente discontinuo, pasando de 0 a 100% en 6 CV.

Tras observar el gel y el cromatograma de la elución, las fracciones de interés se juntaron y se concentraron en Amicons de 30 kDa. Las muestras se filtraron e inyectaron en la columna de exclusión molecular Superdex200 10/300 GL (GE Healthcare) previamente equilibrada con tampón *storage*. Se recogieron fracciones de 0,5 mL y las de interés se volvieron a juntar y concentrar hasta el máximo volumen posible. Por último, la muestra final se centrifugó (14000 rpm, 5 min, 6 °C) para eliminar posibles agregados y se alicuotó y congeló en N₂ líquido para su conservación a -80 °C.

3.5 Otras técnicas

3.5.1 Preparación de glicerizados

Todas las cepas o líneas celulares con las que se trabajó se preservaron preparando glicerizados al 20% de glicerol, y almacenándolos a -80 °C. Esto permite mantener la uniformidad genética de las células y un acceso más sencillo a ellas en caso de repetir alguna prueba, evitando una nueva transformación cada vez.

3.5.2 Análisis y cuantificación de proteínas por colorimetría (Bradford)

La cantidad de proteína en las muestras representativas de cada ensayo o prueba se determinó de manera colorimétrica, basándonos en su absorbancia a determinada longitud de onda según el método de Bradford. Éste se basa en el cambio de color que sufre el reactivo de Bradford al interaccionar con las proteínas (Bradford, 1976). Para ello, se construyó en primer lugar una recta patrón con seroalbúmina bovina (BSA) a 1 mg/mL, preparando disoluciones de 0, 2, 4, 6 y 8 mg/mL. Para evaluar el cambio colorimétrico, se añadió 1 µL de muestra a 999 µL de reactivo de Bradford comercial (Bio-Rad Protein Assay Dye Reagent Concentrate) en cubetas de plástico. Posteriormente, se determinó la absorbancia a 595 nm (agitando la muestra previamente para que la reacción ocurra de manera homogénea) en un espectrofotómetro y se interpoló el valor obtenido con los valores de la recta patrón, calculando así la concentración de proteínas en la muestra.

3.5.3 Geles de agarosa

Los fragmentos y moléculas de ADN con los que se trabajó en la reclonación de IF2tr en pLIC-SGC1 (apartado 3.2) fueron cargados en geles de agarosa para evaluar su calidad y cantidad. Se prepararon geles de agarosa al 1% en tampón TAE (Tris-Acetato-EDTA) teñidos con 3,5 µL de SafeView (NBS Biologicals). Tras gelificar, se cargaron 100-500ug de cada muestra con 1 µL tampón de carga Gel Loading Dye, Purple 6X (NEB) y, por separado, entre 5-7,5 µL de marcador de pesos moleculares: Gene Ruler 1 kb (Thermo Fisher™), NZYDNA Ladder VIII o III (NZYTech) según el caso. Se corrió la electroforesis en TAE 1X durante 40 minutos a 100 V, observando el resultado en un equipo GelDoc Go (Bio-Rad). Adicionalmente, se midió la concentración del ADN en un dispositivo NanoDrop™2000 (Thermo Fisher Scientific), midiendo la A₂₆₀ de las bases nitrogenadas.

3.5.4 SDS-PAGE

Las proteínas en disolución de las diferentes etapas se analizaron por SDS-PAGE para determinar su pureza, cantidad e integridad. Se prepararon geles de 10% poliacrilamida y dodecilsulfato sódico (SDS) al 2% para evaluar la presencia de proteínas contaminantes y la cantidad de proteína deseada en las etapas de expresión, lisis y purificación. Se prepararon muestras de 10-20 µg de proteína

disueltas en tampón de carga Laemmli 5X (Laemmli, 1970). Tras calentar las muestras 10 minutos a 95 °C en termobloque, se cargaron en el gel junto con 5 µL de marcador de pesos moleculares Spectra™ Multicolor Broad Range Protein Ladder (Thermo Scientific™). Los geles corrieron a 200 V durante unos 60 minutos, y se tiñeron con una solución casera de Azul de Coomassie durante 10 minutos en agitación. Tras ese tiempo, se cambió la solución por una de desteñido (10% etanol, 10% ácido acético glacial) y se procedió a la observación en un equipo GelDoc Go. En las ocasiones que requerían un visionado del gel más rápido, se utilizó el tinte BlueSafe (NZYTech), precalentando el gel en agua ultrapura (Milli-Q®) 1 min al microondas y luego calentando otro minuto junto con el BlueSafe.

4. RESULTADOS Y DISCUSIÓN

4.1 Expresión y purificación a pequeña escala de pTYB21-IF2 y pTYB21-IF2tr

Antes de mi estancia en el laboratorio, en el grupo ya habían intentado expresar los constructos IF2/IF2tr-pTYB21 sin éxito. Por tanto, tras mi incorporación, lo primero que intentamos fue expresar el constructo con estrategias que pudieran mejorar la expresión de la proteína recombinante. Para ello, se transformó y expresó a pequeña escala el constructo en células BL21(DE3) pLysS y BL21(DE3) GroEL/ES. En el caso de IF2tr y la cotransformación con pGroESL, no llegaron a crecer colonias viables, por lo que no pudo realizarse la prueba de expresión. Como se observa en la siguiente figura, en ningún caso hubo expresión con GroEL/ES (Fig. 5a) o con pLysS (Fig. 5b). En cuanto a GroEL/ES, sabemos que la cotransformación e inducción funcionaron, porque se observa una gran banda en el carril "POST" de unos 50 kDa, correspondiente al peso de las chaperoninas expresadas.

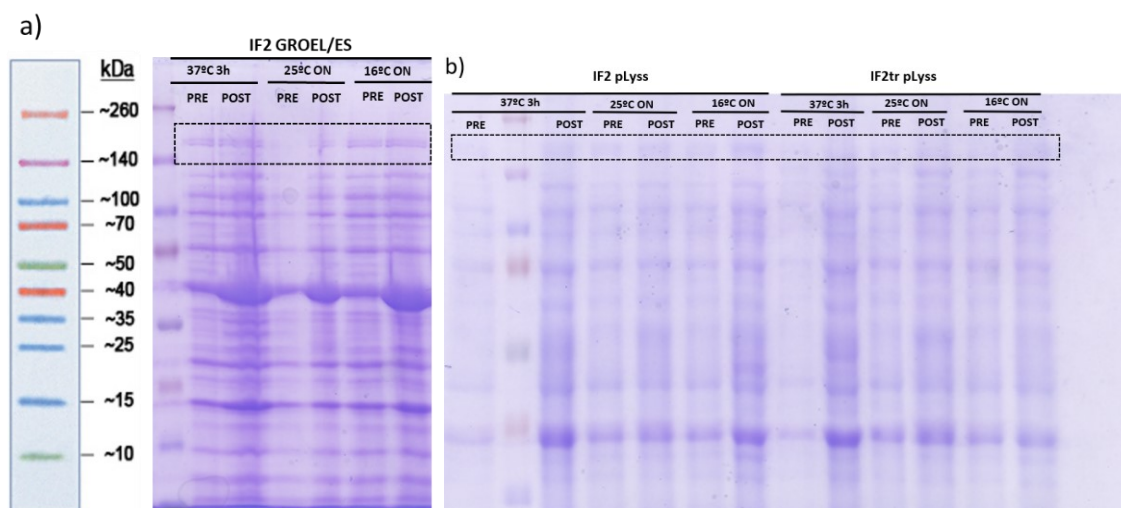


Figura 5. Geles SDS-PAGE (10% poliacrilamida) de la prueba de expresión a pequeña escala de IF2 e IF2tr en células pLysS y GroEL/ES. Se cargaron 10 µg de muestra de las alícuotas tomadas antes ("PRE") y después ("POST") de la inducción y 5 µL de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2-inteína (≈170 kDa) o IF2tr-inteína (≈140 kDa). a) Prueba de expresión de IF2 en BL21(DE3) cotransformada con pGroESL. b) Prueba de expresión de IF2 e IF2tr en BL21(DE3) pLysS.

A continuación, se llevó a cabo otra prueba añadiendo glucosa 1% en los precultivos, para tratar de evitar la expresión basal de los constructos antes de la inducción, ya que se encuentran bajo el control del operón *lac*. En esta ocasión, sí se consiguió expresar ambas proteínas de forma satisfactoria, además de reprimir la transcripción indeseada de las mismas en ausencia de IPTG (Figura 6a), ya que en la carrera "PRE" no se observa expresión. A juzgar por la intensidad de las

bandas “POST”, la mejor condición de expresión resultó ser a 37 °C para ambas proteínas. Por otro lado, IF2tr se expresó en mayor cantidad que IF2, lo cual era de esperar teniendo en cuenta que esta versión, a priori, debería ser más estable y fácil de producir.

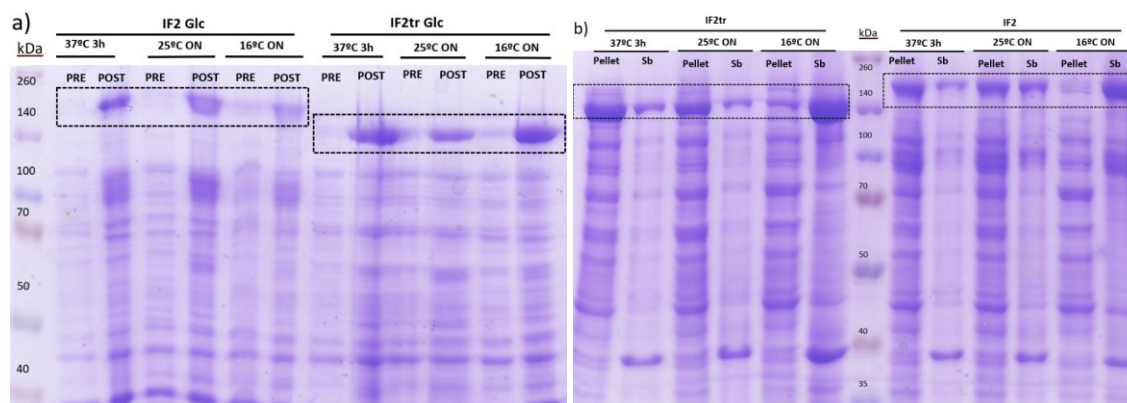


Figura 6. Geles SDS-PAGE (10% poliacrilamida) de las pruebas de expresión y solubilidad a pequeña escala de IF2 y IF2tr en células BL21(DE3) con 1% glucosa. Se cargaron 10 µg de muestra y 5 µL de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2-inteína (≈170 kDa) o IF2tr-inteína (≈140 kDa). a) Gel de expresión de las muestras previas a la inducción (“PRE”) y tras el tiempo de inducción (“POST”) de cada condición de crecimiento: 3h a 37 °C, overnight a 25 °C y overnight a 16 °C. b) Gel de solubilidad de las alícuotas correspondientes a la fracción insoluble (Pellet) y soluble (Sb, sobrenadante) de cada condición.

Dado que en la expresión a pequeña escala todas las condiciones presentaban un nivel relativamente similar de expresión, se realizó la prueba de solubilidad para todas las condiciones (Figura 6b). En ella se vio que la proteína de interés se vuelve más insoluble conforme aumenta la temperatura, ya que la intensidad de las bandas del sobrenadante es menor a 37 °C, a 25 °C la proporción de proteína en el pellet frente al sobrenadante es más equilibrada y a 16 °C la proteína está definitivamente en mayor presencia en la fracción soluble.

De esta manera, la condición de 16 °C *overnight* se eligió como la más idónea para expresión a gran escala.

4.2 Expresión a gran escala de los constructos pTYB21

Tras las pruebas a pequeña escala, se llevó a cabo el crecimiento de los cultivos añadiendo un 1% glucosa a los precultivos, y expresando el constructo toda la noche a 16 °C. En la figura 7 se puede apreciar que la expresión a gran escala en estas condiciones fue exitosa, al haber una gran cantidad de proteína expresada en la alícuota “POST”, con respecto a la “PRE”. De nuevo, la versión truncada se expresa en mayor proporción que la completa, lo cual era previsible teniendo en cuenta que carece de la parte desestructurada, más difícil de expresar.

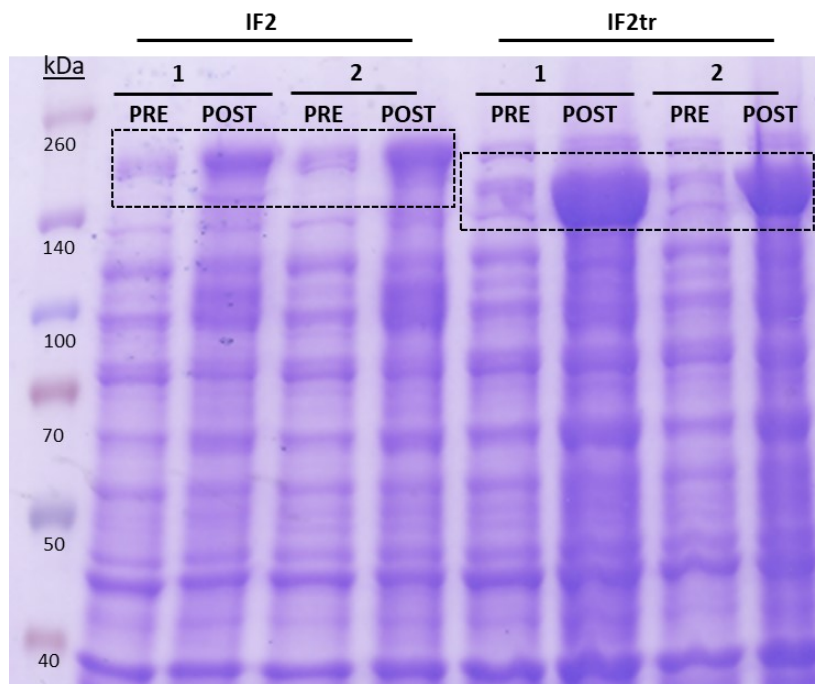


Figura 7. Gel SDS-PAGE (10% poliacrilamida) de la expresión a gran escala de IF2 e IF2tr en células BL21(DE3) con 1% glucosa. Se tomaron alícuotas de los dos cultivos previas a la inducción (“PRE”) y tras el tiempo de inducción (“POST”) *overnight* a 16 °C. Se cargaron 10 µg de muestra y 5 µL de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2-inteína (≈170 kDa) o IF2tr-inteína (≈140 kDa).

4.3 Purificación a gran escala de IF2tr

Tras ver que la expresión de IF2tr era mayor, y presuponiendo que sería menos problemática que IF2, se decidió realizar primero la purificación a gran escala de esta versión. Para ello, las células que habían expresado IF2tr se resuspendieron en tampón de lisis, se sonicaron y la fracción soluble se purificó mediante el sistema IMPACT. El rendimiento de la lisis celular y la cromatografía de afinidad se evaluó mediante SDS-PAGE y cuantificación de proteínas. Como puede observarse en la figura 8, el gel mostró que la proteína IF2tr es soluble, ya que se encuentra en mayor proporción en el sobrenadante que en el pellet y la cuantificación por Bradford respalda este resultado, al haber una concentración casi 8 veces superior en el sobrenadante respecto al pellet. Sin embargo, el gel también evidencia que existe un “goteo” constante de IF2tr en todas las etapas de lavado.

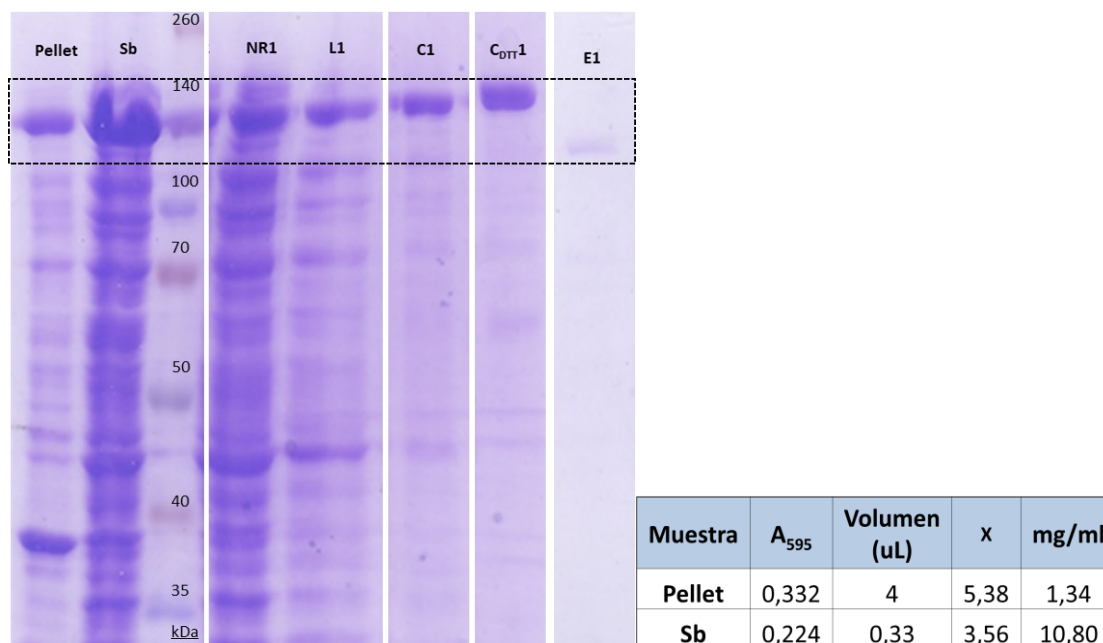


Figura 8. Gel SDS-PAGE (10% poliacrilamida) de la solubilidad de IF2tr a gran escala y su purificación en columna de quitina. Se cargaron alícuotas de la fracción insoluble (Pellet) y soluble (Sb, sobrenadante) tras la centrifugación del lisado celular; y alícuotas de cada etapa de la cromatografía de afinidad de 1 de las 3 columnas de quitina empleadas (1): no retenido (NR1), lavado con tampón *wash* (L1), lavado con tampón *cleavage* sin DTT (C1) y con DTT (C_{DTT}1) y eluido (E1). Se cargaron 10-20 µg de muestra y 5 µL de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-inteína (≈140 kDa). A la derecha se muestra una tabla de la cuantificación por Bradford; el valor “x” se obtuvo interpolando en la pendiente de la recta patrón de BSA 1 mg/mL.

Aunque inicialmente se pensó que este goteo podría deberse a una gran expresión de la proteína que estaría saturando la resina, el gel de los eluidos mostró que en realidad era debido a un problema en el plegamiento del constructo. El motivo que justifica esta idea es que de la columna eluyó muy poca proteína y además ésta tenía el mismo tamaño que el constructo inicial. Si el sistema IMPACT hubiese funcionado, la proteína eluida debería pesar unos 60 kDa menos por la escisión de la etiqueta, dando una banda correspondiente al peso molecular de IF2tr (≈70 kDa).

Pese a que el resultado de la elución no era el esperado, se decidió continuar el protocolo e inyectar la proteína en gel filtración para ver el estado en el que se encontraba. Para ello se juntaron las 3 fracciones de eluido y se concentraron hasta que la proteína empezó a precipitar. En ese momento, se pinchó la muestra en la columna de exclusión molecular acoplada a FPLC, obteniéndose el cromatograma que se observa en la figura 9.

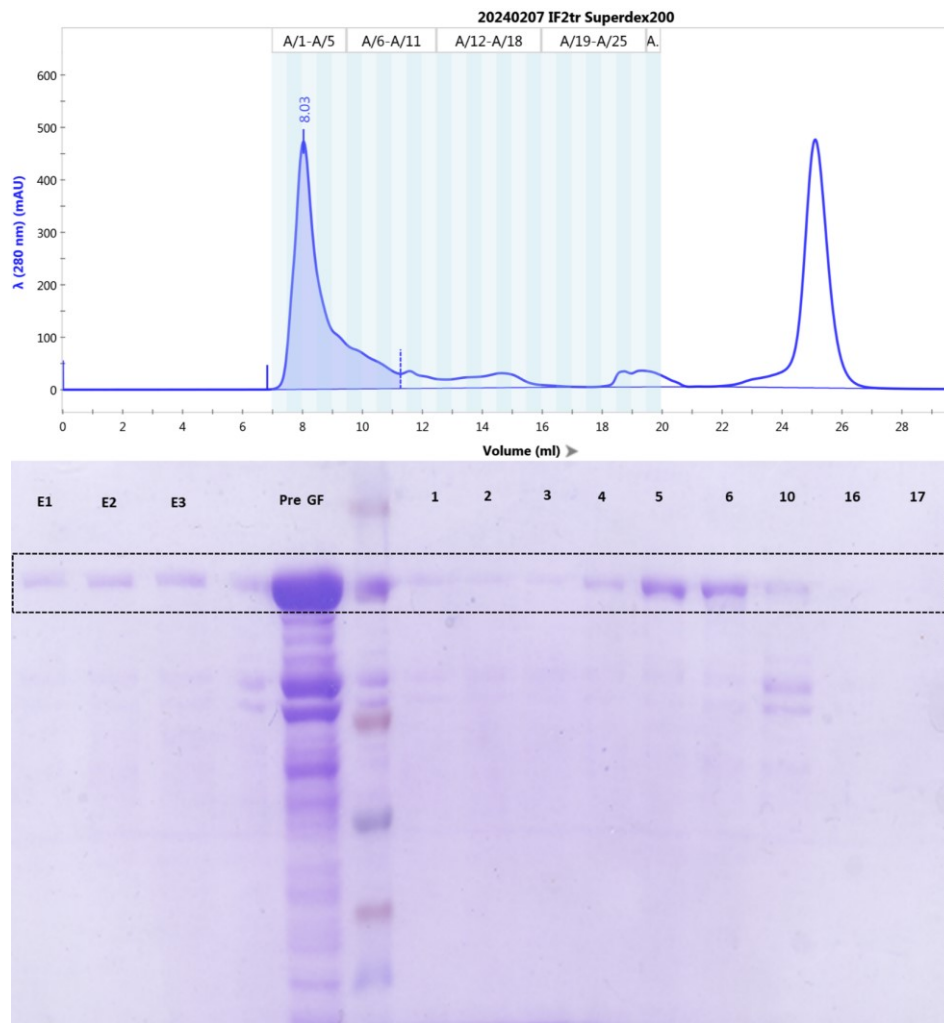


Figura 9. Resultado de la filtración en gel de IF2tr. Se muestra el cromatograma de elución de IF2tr en la columna de filtración en gel Superdex200 (GE Healthcare), donde se observan dos picos: el primero correspondiente a IF2tr y el segundo, correspondiente a especies de bajo peso molecular que se eluyen al final de la columna. Abajo se muestra un gel SDS-PAGE (10% acrilamida) de los eluidos de cada columna (E), la muestra concentrada (Pre GF) y las fracciones de interés.

Analizando el cromatograma, la mayor parte de absorbancia a 280 nm se registra en el pico del principio, lo que indica que la proteína sale de la columna en el “volumen muerto” o *void volume*, junto al resto de especies que no interactúan con la columna. Esto ocurre cuando la molécula es tan grande que no entra por los poros de la resina que compone la fase estacionaria (O’Fágáin et al., 2011), indicando que nuestra proteína puede estar agregada, formando partículas de alto peso molecular. Este resultado confirmó las sospechas anteriores de que el constructo expresado no interactúa correctamente con la columna, no por haber una excesiva cantidad, sino porque probablemente tiene un mal plegamiento que le conduce a formar agregados solubles.

4.4 Expresión de IF2tr con chaperonas endógenas de *E. coli*

Para tratar de revertir los agregados, se planteó expresar IF2tr en las mismas condiciones (0.3 mM IPTG, ON 16 °C) pero induciendo la producción de chaperonas endógenas de *E. coli*, para que ayudaran al plegamiento de la proteína. Como se ha explicado en el apartado 3.4.2, tras someter las células a choque térmico en hielo, un grupo se mantuvo sin recibir más estrés (TA), otro recibió 2% de etanol (EtOH) y otro, sorbitol al 0,5 M (Sorb) previamente a la inducción. En este caso, los resultados de la expresión y la solubilidad también fueron adecuados, como se observa en la figura 10.

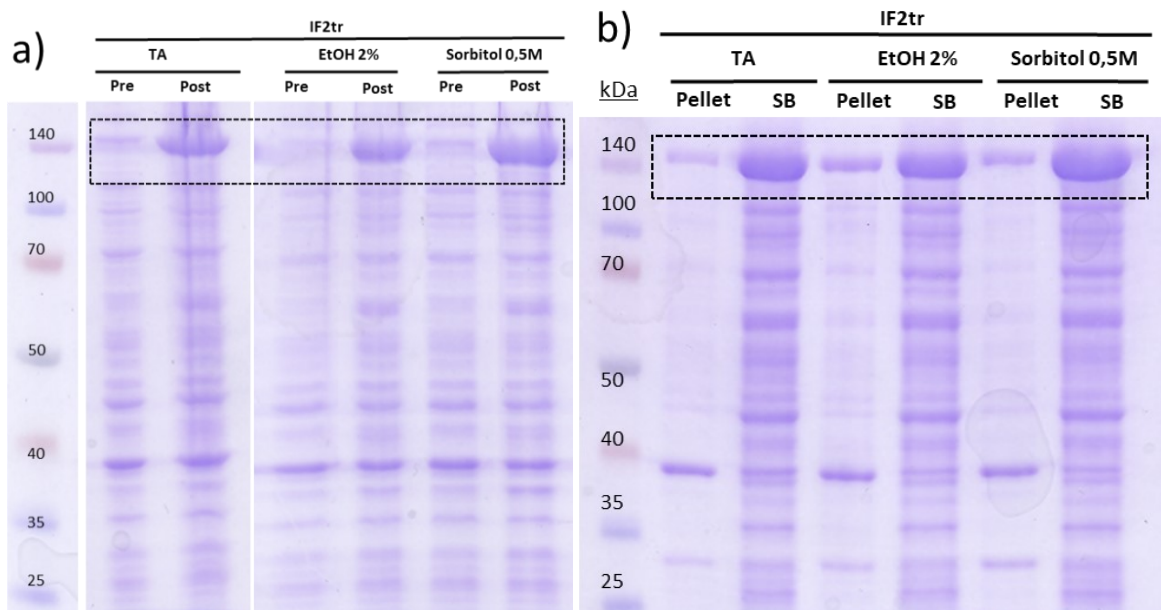


Figura 10. Gel SDS-PAGE (10% poliacrilamida) de la expresión y solubilidad de IF2tr con chaperonas endógenas. Se cargaron 10 μ L de muestra y 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-inteína (\approx 140 kDa). a) Gel de expresión de las alícuotas pre y post inducción de cada tipo de estrés: solo choque térmico (TA), etanol 2% (EtOH) y sorbitol 0,5 M. b) Gel de solubilidad de las alícuotas de la fracción insoluble (Pellet) y soluble (SB, sobrenadante).

Los sobrenadantes filtrados se cargaron en la columna de quitina e IF2tr se purificó en las mismas condiciones observando cada etapa en SDS-PAGE (Figura 11) y cuantificando la proteína presente por Bradford. En este caso, el cultivo que no había sufrido más estrés antes de la inducción (TA) se purificó a temperatura ambiente, mientras que los otros dos (2% de etanol y 0.5M de sorbitol) se purificaron en frío. Tras analizar los resultados del gel, una vez más, se apreció un ligero goteo de la proteína en cada etapa, aunque en menor medida que en el primer intento sin chaperonas.

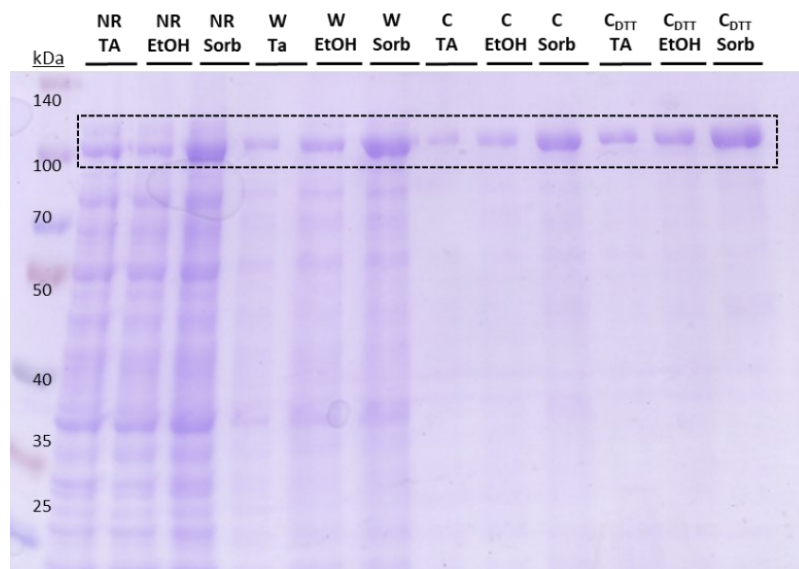


Figura 11. Gel SDS-PAGE (10% poliacrilamida) de la purificación de IF2tr con chaperonas endógenas en columna de quitina. Análisis de cada etapa de la cromatografía de afinidad en las 3 columnas de quitina empleadas: no retenido (NR), lavado con tampón wash (L), lavado con tampón cleavage sin DTT (C) y con DTT (C_{DTT}). En cada columna se purificó un tipo de condición: solo choque térmico (TA), etanol al 2% (EtOH) y sorbitol a 0,5 M. Se cargaron 10-20 μ g de muestra y 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-inteína (\approx 140 kDa).

En este caso, las columnas también se eluyeron obteniendo muy poca proteína (Tabla 3), de modo que, previendo que ninguna condición había mejorado el plegamiento, se decidió analizar también algunas fracciones de los lavados.

Tabla 3. Cuantificación de proteínas en los eluidos por el método de Bradford. El valor "x" se obtuvo interpolando en la pendiente de la recta patrón de BSA 1 mg/mL.

Muestra	A ₅₉₅	Volumen Bradford(μL)	x	mg/ml	Volumen (mL)	mg totales
E1	0,057	20	0,76	0,04	10	0,38
E2	0,025	20	0,28	0,01	10	0,14
E3	0,033	20	0,40	0,02	10	0,20

Así pues, se juntaron los eluidos (E1-3) y, por otro lado, los lavados. En el caso del sorbitol se juntaron los últimos lavados (C+C_{DTT} Sorb) y no se mezclaron con ninguna otra condición al ser los que más intensidad de banda tenían en el gel. Sin embargo, los lavados (C+C_{DTT}) del resto de condiciones (TA+EtOH) se juntaron porque el patrón de bandas era muy similar, y había poca cantidad de proteína. Las muestras se concentraron por separado y se pincharon en la columna de filtración en gel (Figura 12).

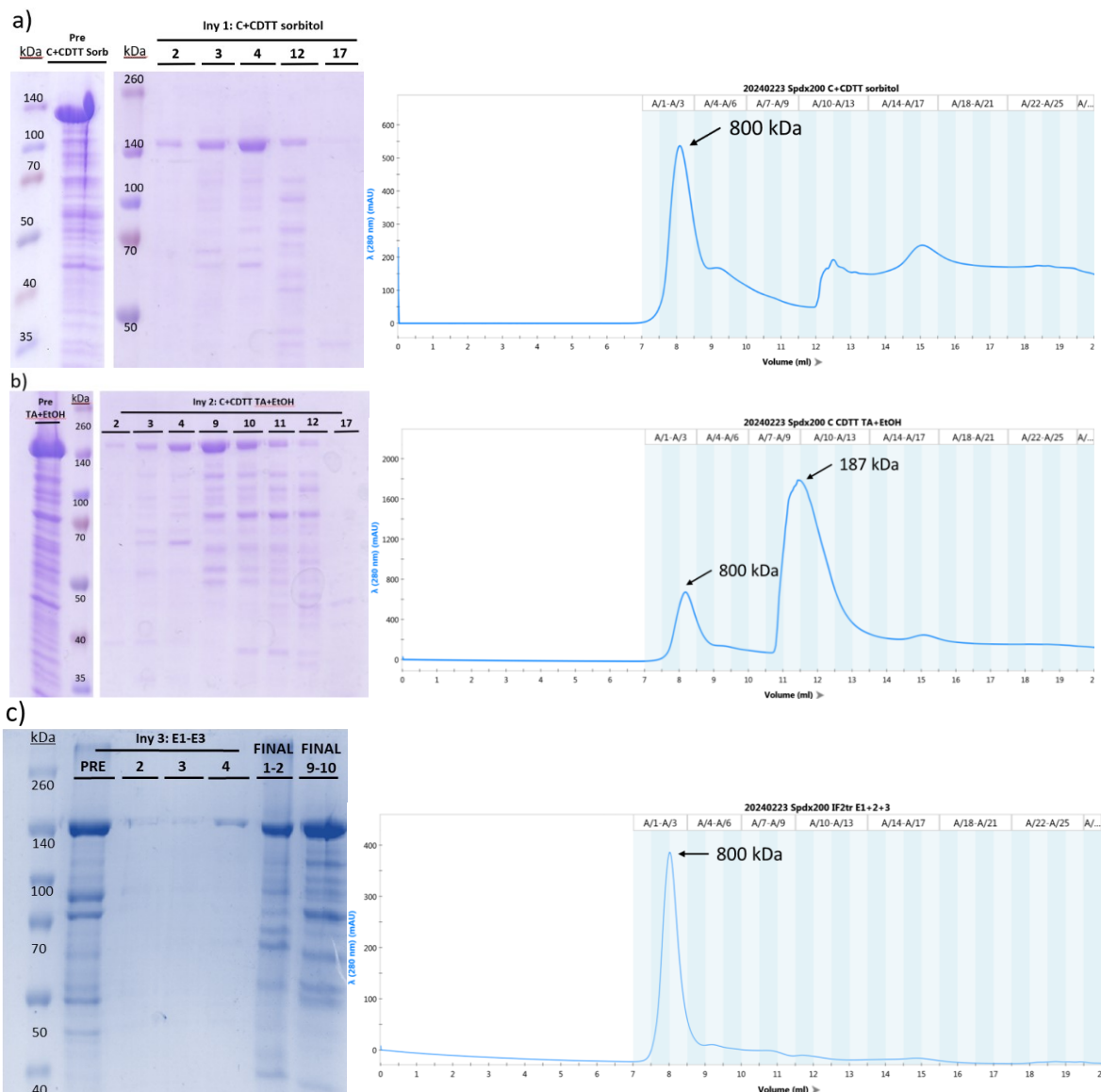


Figura 12. Resultado de la purificación de IF2tr con chaperonas endógenas por filtración en gel. Para cada pinchazo se cargaron 500 μL de muestra (Pre) en la columna de filtración en gel Superdex200 (GE Healthcare) acoplada al FPLC NGC

Cromatography System (Bio-Rad). a) Gel SDS-PAGE (10% poliacrilamida) y cromatograma de las fracciones C+C_{DTT} Sorbitol. b) Gel SDS-PAGE (10% poliacrilamida) y cromatograma de las fracciones C+C_{DTT} de Temperatura ambiente+Etanol. c) Cromatograma de los eluidos mezclados (E1-3) y gel SDS-PAGE (10 % poliacrilamida) de las fracciones recogidas a la salida de la columna (Iny3) y de las muestras finales (FINAL 2-4 o 9-10) para congelar.

Al analizar los cromatogramas y los geles cargados con las fracciones más relevantes, se vio que la expresión con sorbitol seguía mostrando agregados de IF2tr en el volumen muerto de la columna (Fig 12a), indicando que el estrés osmótico no había servido para mejorar el problema de plegamiento. Por su parte, en los lavados de TA+EtOH (Fig. 12b) sí aparecía un pico de absorbancia a una altura de unos 187 kDa, además del pico de agregado. Este pico, aunque tiene un tamaño ligeramente superior, sí que podría encajar con un monómero del constructo IF2tr-inteína (140 kDa), ya que es posible que se comporte como una proteína de mayor peso molecular. Esto puede ser debido, por ejemplo, a un plegamiento fibrilar, pues el FPLC interpola el tamaño de pico con los patrones asumiendo un plegamiento globular de las proteínas (O'Fágáin et al., 2011). No obstante, al haber juntado ambas condiciones (TA y EtOH) no podemos afirmar cuál es la responsable detrás de este pico. De todos modos, al inyectar los eluidos de todas las condiciones juntas, no se registró este pico, sino que aparecía de nuevo el pico en el volumen muerto (Fig. 12c), por lo que se concluyó que en ninguna condición se conseguía revertir la agregación.

Además, cabe destacar que, tras la elución de las columnas de quitina, el tamaño de las bandas en todos los geles es de 140 kDa y no los 70 kDa que debería pesar IF2tr sola. Esto sugiere que la inteína ni siquiera se está escindiendo, y que por tanto lo que recogemos en los eluidos sigue siendo proteína no unida a la resina, es decir, proteína que no acabó de lavarse. Una idea importante que se deduce de aquí, junto al mayor tamaño aparente en los cromatogramas, es que el constructo tiene un plegamiento aberrante que impide, por un lado la interacción adecuada entre el dominio CBD y la columna de quitina y, por otro, la escisión de la inteína; conduciendo a la formación de agregados solubles.

En este momento nos planteamos que el problema en la funcionalidad de la inteína podía ser debido a un mal plegamiento en la propia etiqueta o en IF2tr, lo cual tenía soluciones diferentes. Si IF2tr estuviese mal plegado y este mal plegamiento condujese a la formación de agregados, deberíamos plantearnos como mejorar la expresión. Por el contrario, si el problema estuviera en la etiqueta, de manera que no pudiera plegarse adecuadamente por culpa de la secuencia *linker* entre la inteína y la proteína, tal vez la solución podría ser reclonar IF2tr a otro vector con otra etiqueta.

4.5 Purificación de IF2

Para resolver esta dicotomía sobre el mal plegamiento de IF2tr, decidimos comprobar si IF2 se podía purificar correctamente. De esta manera, aprovechando que el segmento C-terminal era exactamente el mismo, podríamos elucidar si el truncamiento era el factor responsable del mal plegamiento de la etiqueta inteína. Para ello, se realizó una purificación a pequeña escala en la que los sobrenadantes de la prueba de solubilidad (apartado 4.1 o figura 6b) se cargaron en columnas Mini Bio-Spin (Biorad®), con 250 µL de resina de quitina (NEB) siguiendo el protocolo de pequeña escala detallado en 3.3.2. Tras realizar los lavados y la elución, las muestras se cargaron en un SDS-PAGE (Figura 13):

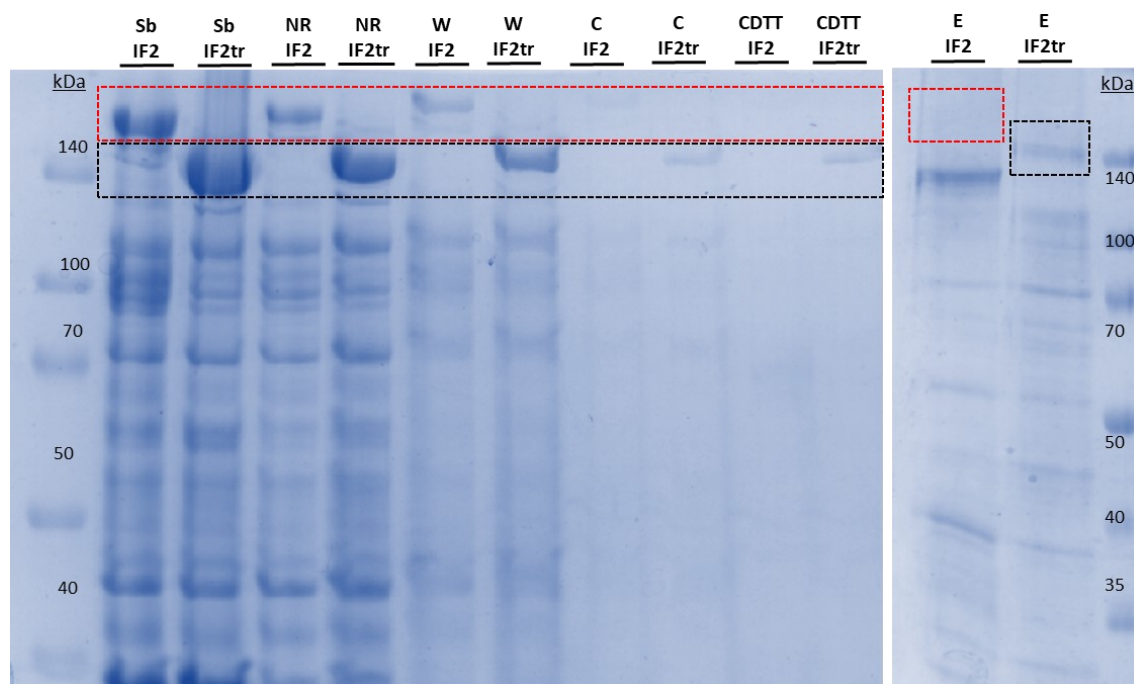


Figura 13. Gel SDS-PAGE (10% acrilamida) de la purificación de IF2 e IF2tr a pequeña escala en columna de quitina. En el gel se cargaron muestras de los sobrenadantes (Sb), el no retenido y lavados de la columna (NR, W, C, C_{DTT}) y la elución (E). Las zonas recuadradas en rojo y negro corresponden al peso molecular de la proteína de fusión IF2-inteína (≈ 170 kDa) e IF2tr-inteína (≈ 140 kDa), respectivamente.

Al analizar los geles vemos que sí hay diferencias en el comportamiento de IF2 con respecto al de IF2tr. En primer lugar, en el no retenido de IF2 se observa una banda menos intensa que en el de IF2tr, partiendo de una intensidad similar en ambos sobrenadantes. Este hecho, junto a que en los lavados hay menos pérdida de IF2; invita a pensar que la versión completa se une mejor a la columna que la troncada, donde sí hay este goteo habitual. Además, en el gel de elución se observa que IF2 eluye con un tamaño menor al del constructo IF2-inteína (≈ 170 kDa), lo que confirma su escisión y por tanto que la etiqueta en este caso es funcional. Asimismo, el eluido de IF2tr presenta la banda correspondiente al constructo sin cortar (≈ 140 kDa), descartando así la posible opción de que en realidad la incubación con DTT fuera subóptima, y confirmando que en este constructo la inteína no puede autoescindirarse.

En consecuencia con los resultados de la pequeña escala de IF2, se procedió a su purificación a gran escala siguiendo los pasos del apartado 3.4.1. La solubilidad de la proteína y su purificación por el protocolo IMPACT se evaluaron mediante SDS-PAGE (Figura 14) y cuantificación por Bradford:

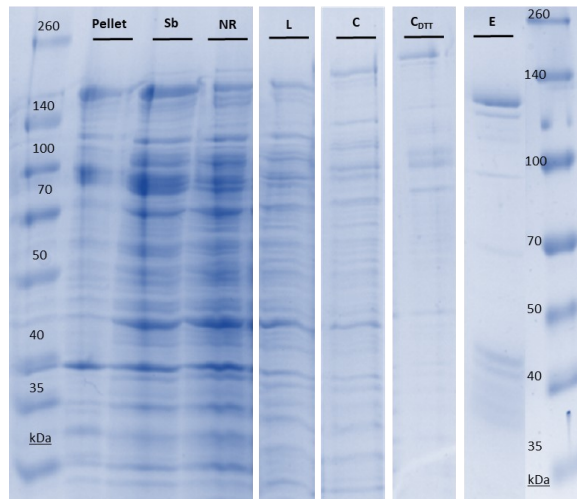


Figura 14. Gel SDS-PAGE (10% acrilamida) de la purificación de IF2 a gran escala en columna de quitina. La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2-inteína (≈ 170 kDa). En el primer carril se cargaron 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). En el resto, se cargaron 10-20 μ g de cada muestra: pellet, sobrenadante (Sb), no retenido (NR), lavado con tampón wash (L), lavado con tampón cleavage sin DTT (C) y con DTT (C_{DTT}), y eluido (E).

Los eluidos se concentraron (Pre GF) y se sometieron a filtración en gel (Figura 15).

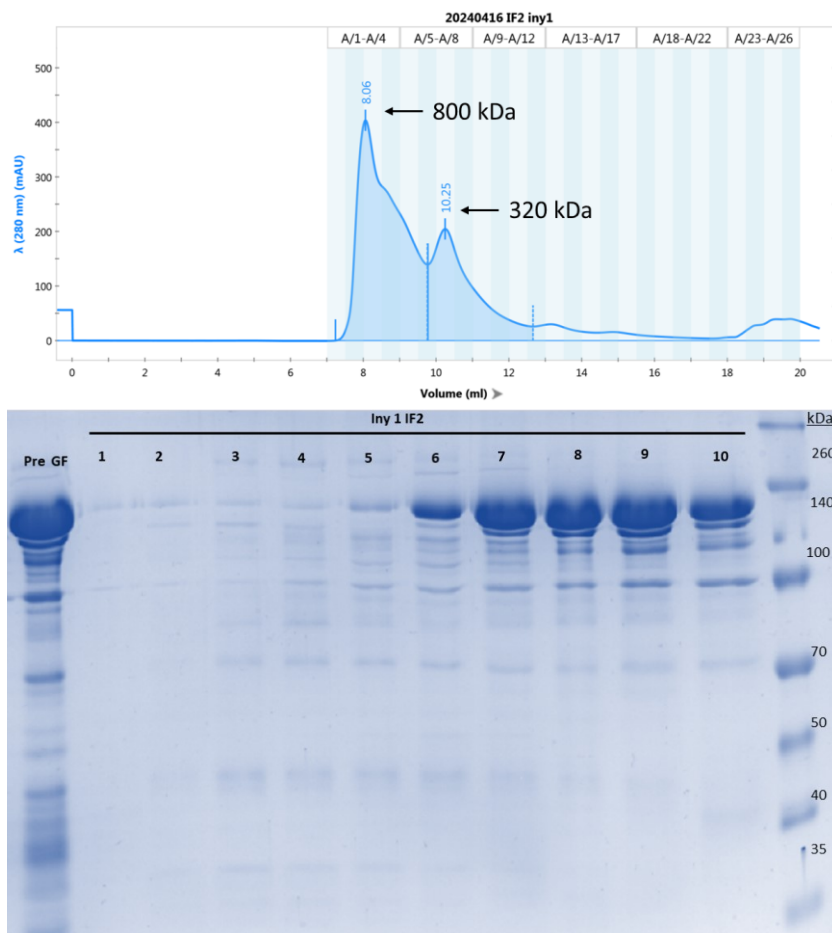


Figura 15. Filtración en gel de la purificación de IF2 a gran escala. a) Cromatograma de la elución de IF2 en la columna de gel filtración Superdex200 (GE Healthcare). b) Gel SDS-PAGE (10% acrilamida) con 10-20 μ g de la muestra concentrada antes de la inyección (Pre GF) y de las fracciones recogidas a la salida; y 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™).

El perfil de elución estaba formado por dos picos, uno inicial en el volumen muerto (≈ 870 kDa) y un segundo pico de unos 320 kDa. Curiosamente, el primer pico de 870 kDa no mostró ninguna banda de proteína al cargarlo en el gel (Fracciones 1-5), por lo que se pensó que podría tratarse de otros metabolitos que absorban a 280 nm o algún agregado proteico muy grande que tenga mucha absorbancia y no entre en el gel. Por su parte, el segundo pico de los cromatogramas (FR6-10) sí contiene IF2 a un buen nivel de pureza. Este pico tiene un tamaño mucho mayor que IF2, por lo que podría estar en alguna forma dimérica/oligomérica. No obstante, no se ha descrito en la literatura que esta proteína forme oligómeros, por lo que simplemente puede tratarse de una forma monomérica, en la que el dominio desestructurado podría hacer que su plegamiento no sea globular y refleje un tamaño aparente mayor al real.

Finalmente se mezclaron las fracciones FR6-11, se concentraron y se congelaron con N_2 líquido. En esta purificación se consiguió obtener 0,6 mL a 0,9 mg/mL, es decir, 0,54 mg totales de IF2.

La obtención de IF2 confirmó que la proteína sí se puede expresar y purificar en *E. coli* sin necesidad de utilizar chaperonas u otras alternativas como detergentes o agentes caotrópicos. Por tanto, llegados a este punto, era necesario reclonar IF2tr en otro vector con otra etiqueta de purificación distinta.

4.6 Reclonación de IF2tr en pLIC-SGC1

La reclonación de IF2tr en pLIC-SGC1 comenzó con la PCR de inserto para amplificar el gen de IF2tr, clonado en pTYB21. Previamente, se hicieron varias pruebas para encontrar las condiciones de T_m , tampones y aditivos óptimas. Finalmente, se eligió una T_m de 59 °C y el tampón GC+3% DMSO, ya que dieron los resultados con mayor intensidad de banda del amplicón (2000 pb) y con menos amplificación inespecífica (Figura 16). Posteriormente, se llevó a cabo la PCR de inserto con las condiciones definitivas (apartado 3.2.2), logrando amplificar el inserto satisfactoriamente (Figura 17).

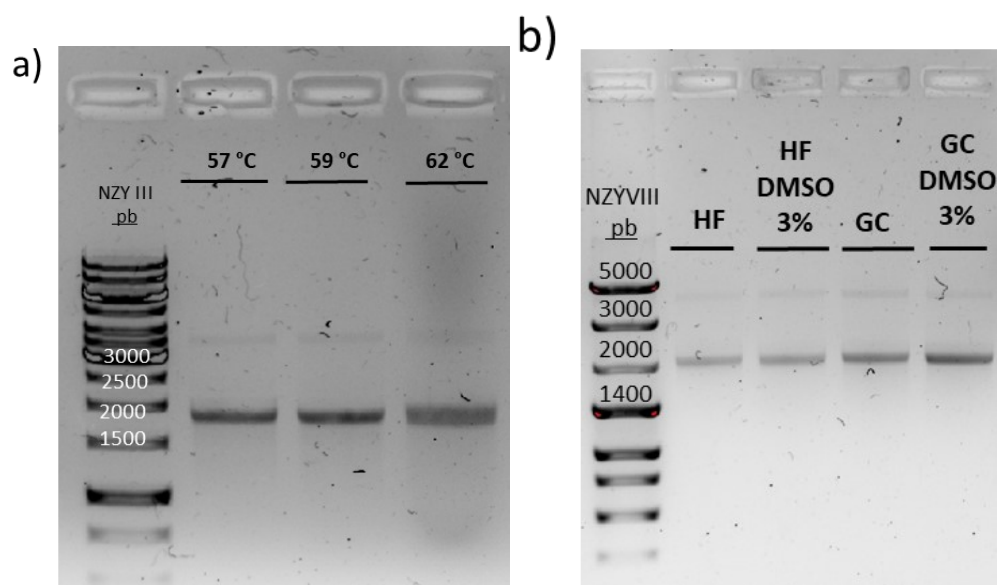


Figura 16. Gel de agarosa (1% agarosa en TAE 1X) de las pruebas de optimización la PCR de amplificación del inserto. Se cargaron 5 μ L de muestra con 1 μ L de Gel Loading Dye, Purple 6X (NEB). a) Prueba de T_m . Se cargaron 5 μ L de marcador de pesos moleculares Gene Ruler 1 kb (Thermo Fisher™) y 5 μ L de cada producto de PCR. b) Prueba de tampón y aditivos. Se cargaron 7,5 μ L de marcador de pesos moleculares NZYDNA Ladder VIII (NZYTech) y 5 μ L de cada producto de PCR.

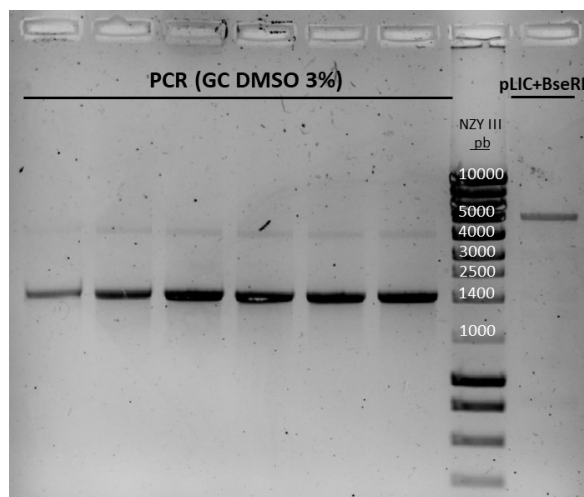


Figura 17. Gel de agarosa (1% agarosa en TAE 1X) de la PCR de inserto optimizada y la digestión de pLIC-SGC1 con BseRI. Se cargaron muestras de los 6 tubos de PCR empleados, junto al marcador de pesos moleculares NZYDNA Ladder III (NZYTech). A la derecha del marcador aparece el resultado de la digestión.

Por otra parte, se digirió el vector destino pLIC-SGC1 con la enzima de restricción BseRI (Fig. 17) obteniendo una banda de 5000 pb en el gel, aunque debería observarse también la banda de 2000 pb del inserto del vector liberado. Pese a ello, el tamaño de dicha banda (5000 pb) corresponde al producto del vector linealizado que esperábamos, por lo que se decidió continuar con el protocolo y purificarlo.

Los productos de la PCR y la digestión se trataron por separado con T4 ADN polimerasa y, tras realizar la clonación independiente de ligación (LIC), se transformaron en *E. coli* DH5 α . Las diferentes colonias crecidas se picaron para realizar una PCR de colonia usando los cebadores del T7 promotor y T7 terminador (Tabla 1). Los resultados se observaron en gel de agarosa (Figura 18).

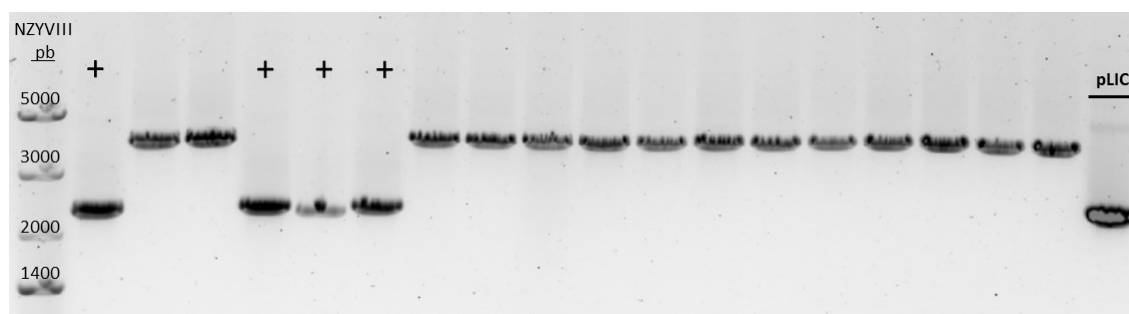


Figura 18. Gel de agarosa (1% agarosa en TAE 1X) de la PCR de colonia. Se cargaron 7,5 μ L de marcador de pesos moleculares NZYDNA Ladder VIII (NZYTech) y 5 μ L de los productos de PCR. Las colonias positivas aparecen marcadas con un "+". El último carril corresponde al control negativo pLIC-SGC1 con el gen SacB.

Se obtuvieron 4 colonias positivas, que amplificaban un segmento de ADN de aproximadamente 2000 pb, mientras que el resto fueron negativas. Las negativas producían una banda de 4000 pb cuya aparición fue desconcertante, pues no hay ninguna región que amplifique entre los T7 de 4000 pb. Por ello, una posible explicación es que los cebadores alinearan de forma inespecífica en otro lugar del plásmido. Por otra parte, el pLIC-SGC1 utilizado como control negativo daba una banda de tamaño muy similar a las positivas. Esto ocurre porque el inserto de los genes IF2tr y SacB tiene una longitud parecida entre ellos (1800 y 2000 pb, respectivamente), por lo que el tamaño de sus productos tras amplificar con T7 también era muy similar (2100 y 2300 pb). Por tanto, aunque el vector había sido purificado de una banda con un tamaño que correspondía al pLIC-SGC1 sin SacB, no podíamos afirmar con seguridad que las colonias positivas habían incorporado IF2tr.

Por ello se repitió la PCR de colonia cambiando los cebadores de T7 por los que se habían empleado para hacer la PCR de amplificación del inserto (FW_IF2tr_pLIC y RV_IF2tr_pLIC) (Figura 19).

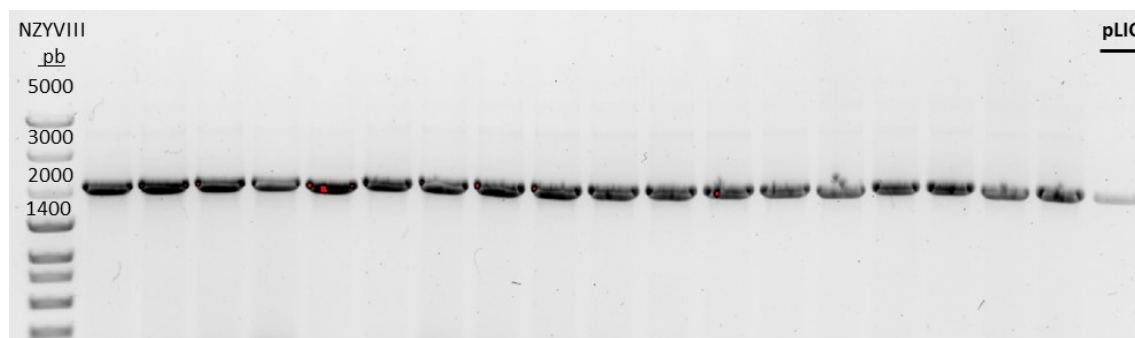


Figura 19. Gel de agarosa (1% agarosa en TAE 1X) de la segunda PCR de colonia. Se cargaron 7,5 μ L marcador de pesos moleculares NZYDNA Ladder VIII (NZYTech) y 5ul del producto de PCR. Las colonias positivas aparecen marcadas con un "+". El último corresponde al control negativo pLIC-SGC1 con el gen SacB.

En la segunda PCR, sospechosamente todas las colonias salieron positivas, con una banda de unos 2100 pb. Esto podía ser porque realmente todas las colonias eran positivas y habían incorporado el inserto o porque los *overhangs* de los cebadores hibridaran en todas las células con pLIC por complementariedad, aunque no deberían amplificar ya que sus 3' teóricamente no estaban alineando.

Como los resultados seguían sin ser concluyentes, se realizó una última PCR de colonia combinando los cebadores: T7_FW con RV_IF2tr_pLIC; y FW_IF2tr_seq con T7_RV (Figura 20). Para simplificar el ensayo, esta PCR de colonia se realizó sólo con algunas de las colonias empleadas anteriormente (colonias 1-8).

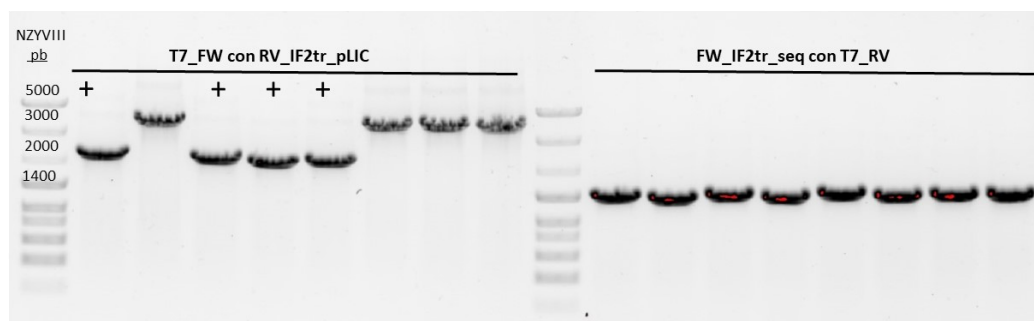


Figura 20. Gel de agarosa (1% agarosa en TAE 1X) de la tercera PCR de colonia. Se cargaron 7,5 μ L de marcador de pesos moleculares NZYDNA Ladder VIII (NZYTech) y 5 ul de los productos de PCR. Las colonias positivas aparecen marcadas con un "+". A la izquierda aparece la reacción con los cebadores T7_FW y RV_IF2tr_pLIC. A la derecha aparece la reacción con FW_IF2tr_seq y T7_RV.

Las reacciones con el primer par de cebadores dieron el mismo resultado que la primera PCR de colonia: 4 colonias positivas con banda de 2100 pb, y las negativas con banda de 4000 pb. Por su parte, el segundo par de *primers* dio lugar a un amplicón de unos 1350 pb en todas las colonias, lo que refuerza la posibilidad de que todas las fueran positivas, ya que FW_IF2tr_seq solo puede hibridar dentro del inserto IF2tr.

Finalmente, se mandaron a secuenciar las 4 colonias que habían resultado positivas en las 3 PCRs, sobre las cuales se tenía más confianza. La secuenciación se realizó usando los cebadores de T7 y el cebador FW_IF2tr_seq. Las secuencias se analizaron utilizando la herramienta EMBOSS Needle (EMBL-EBI) y con ello se reveló que las colonias habían incorporado con éxito el gen IF2tr, sin alteraciones en la pauta de lectura o en la secuencia de nucleótidos.

4.7 Expresión y purificación a pequeña escala de pLIC-IF2tr

Una vez comprobado el éxito de la reclonación, se expresó transformo y expresó el nuevo constructo pLIC-IF2tr en células BL21(DE3), siguiendo los pasos relatados en el apartado 3.3.3 de Materiales y métodos. En el gel de expresión (Figura 21a), se observa que la colonia 1 expresaba la proteína en mayor cantidad, por lo que se continuó con ella en la prueba de solubilidad y el resto del protocolo. En esta ocasión, la expresión resultó mejor a 37 y 25 °C que a 16 °C, obteniéndose mayor cantidad de proteína en el lisado y en un estado más soluble (Fig. 21b).

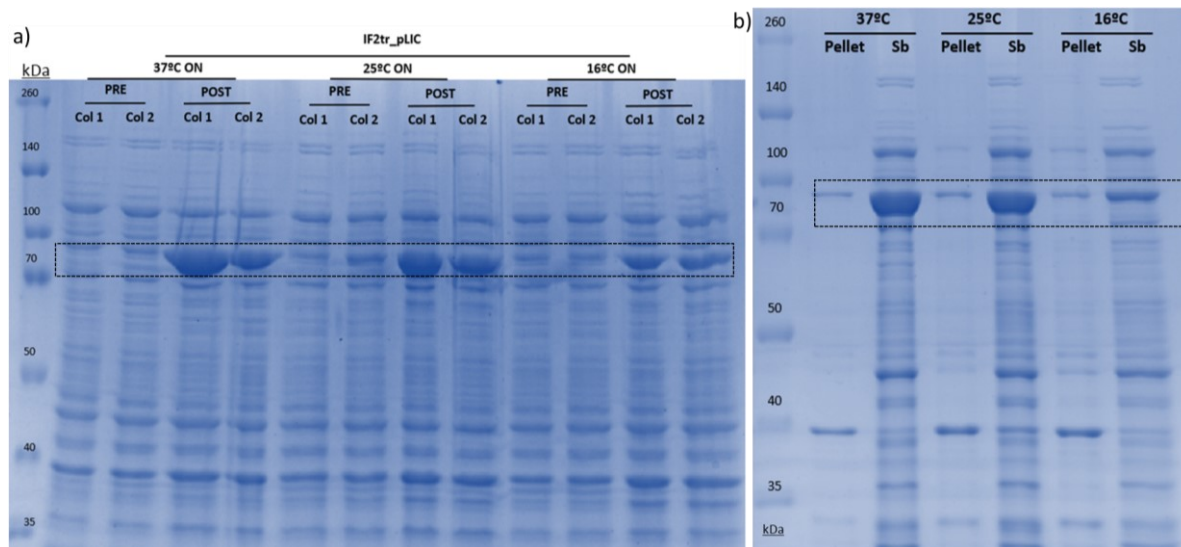


Figura 21. Geles SDS-PAGE (10% poliacrilamida) de la prueba de expresión y solubilidad a pequeña escala de pLIC-IF2tr. La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-6xHis (≈ 70 kDa). a) Gel de expresión de las alícuotas "PRE" y "POST" inducción. Se cargaron 10 μ g de muestra y 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). b) Gel de solubilidad de las alícuotas de la fracción insoluble (Pellet) y soluble (Sb, sobrenadante) tras la sonicación de cada condición de expresión.

Tras analizar el gel del ensayo de solubilidad, se descartó la condición de 16 °C al ser menos soluble y se decidió continuar con la purificación a pequeña escala de los sobrenadantes de 25 y 37 °C. Analizando la figura 22, podemos afirmar que la nueva versión pLIC-IF2tr sí se retiene en la columna cromatográfica, ya que no encontramos apenas proteína en el no retenido (NR) ni el lavado (W), y se eluye una gran cantidad de ella en las dos condiciones probadas, con un buen grado de pureza.

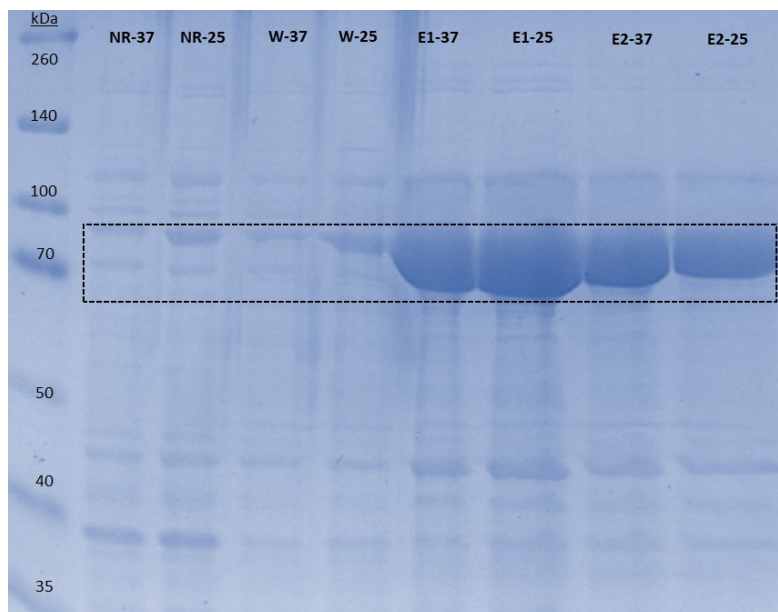


Figura 22. Gel SDS-PAGE (10% acrilamida) de la prueba de purificación de pLIC-IF2tr a pequeña escala en columna de histidina His SpinTrap™ (Cytiva). En el gel se cargaron 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™) y 10-20 μ g de muestra de cada condición (37 y 25 $^{\circ}$ C) en las diferentes etapas cromatográficas: los sobrenadantes (Sb), el no retenido (NR), lavado de la columna (W) y la elución en dos pasos (E1, E2). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-6xHis (\approx 70 kDa).

De las dos condiciones probadas, 37 $^{\circ}$ C 3h o 25 $^{\circ}$ C *overnight*, se eligió la segunda para la expresión a gran escala puesto que IF2 se había expresado a 16 $^{\circ}$ C, y se prefería que las proteínas estuvieran expresadas en las condiciones más similares posibles.

4.8 Expresión y purificación a gran escala de pLIC-IF2tr

Como se ha dicho anteriormente, la expresión a gran escala se llevó a cabo a 25 $^{\circ}$ C *overnight* (Figura 23), tras los resultados de la pequeña escala.

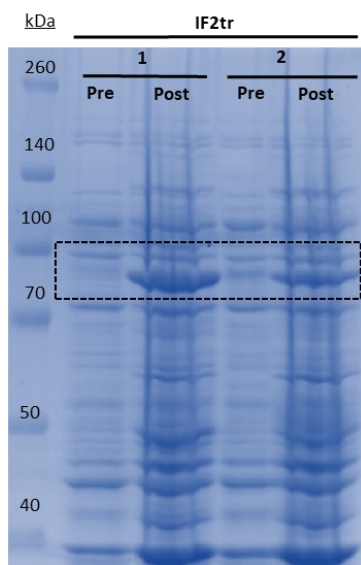


Figura 23. Gel SDS-PAGE (10% poliacrilamida) de la expresión a gran escala de pLIC-IF2tr. Se tomaron alícuotas de los cultivos 1 y 2 previas a la inducción ("PRE") y tras el tiempo de inducción ("POST") *overnight* a 25 $^{\circ}$ C. Se cargaron 10 μ g de muestra y 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-6xHis (\approx 70 kDa).

Tras comprobar que el constructo IF2tr-6xHis se expresó con éxito a gran escala a 25 °C, las células se lisaron y el sobrenadante se purificó por el sistema IMAC, utilizando una columna de níquel HisTrap™ FF (GE Healthcare). Las fracciones se cargaron en un SDS-PAGE (Figura 24).

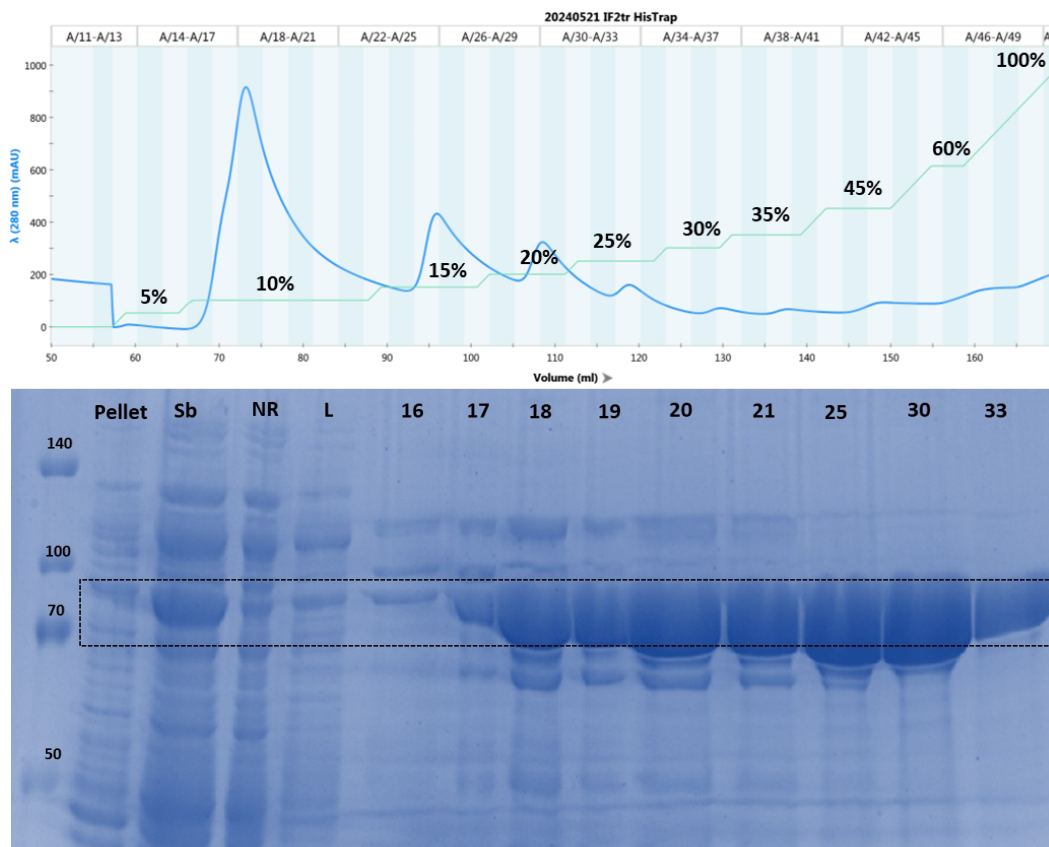


Figura 24. Resultado de la purificación de pLIC-IF2tr en columna de níquel. Arriba se muestra el cromatograma de la columna de níquel, lavada con 100% tampón His-A y eluida con un gradiente de tampón His-B de 0-100% en 6 CV (línea verde). Abajo, el gel SDS-PAGE (10 % poliacrilamida) de la solubilidad (Pellet, Sobrenadante), el no retenido (NR), el lavado (L) y las fracciones de interés. Se cargaron 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-6xHis (\approx 70 kDa).

En el gel se observó un goteo continuo de IF2tr en todas las fracciones desde el 10-30% imidazol, repartido en un pico inicial muy alto, y picos posteriores de menor tamaño. Este goteo podía ser provocado por una saturación de la columna al cargar demasiada cantidad de proteína, a juzgar por el gel y la cuantificación en Bradford. Otra posibilidad es que existieran dos poblaciones de IF2tr diferentes en la muestra, y una de ellas se eluyera de forma más temprana.

Ante la duda, se actuó pensando que podrían existir 2 poblaciones de IF2tr diferentes en la muestra: la población del primer pico (fracciones de la 18 a la 21, F18-21), con una etiqueta 6xHis menos expuesta y que no se retiene bien a la columna; y la población de los picos posteriores, con una etiqueta que se retiene más fuertemente (F25-33). Para no mezclar ambas, cada población se separó en un lote o *batch* diferente, se concentraron por separado y se inyectaron en gel filtración para analizarlas. Las fracciones más representativas se cargaron en un gel SDS-PAGE para su posterior análisis (Figura 25).

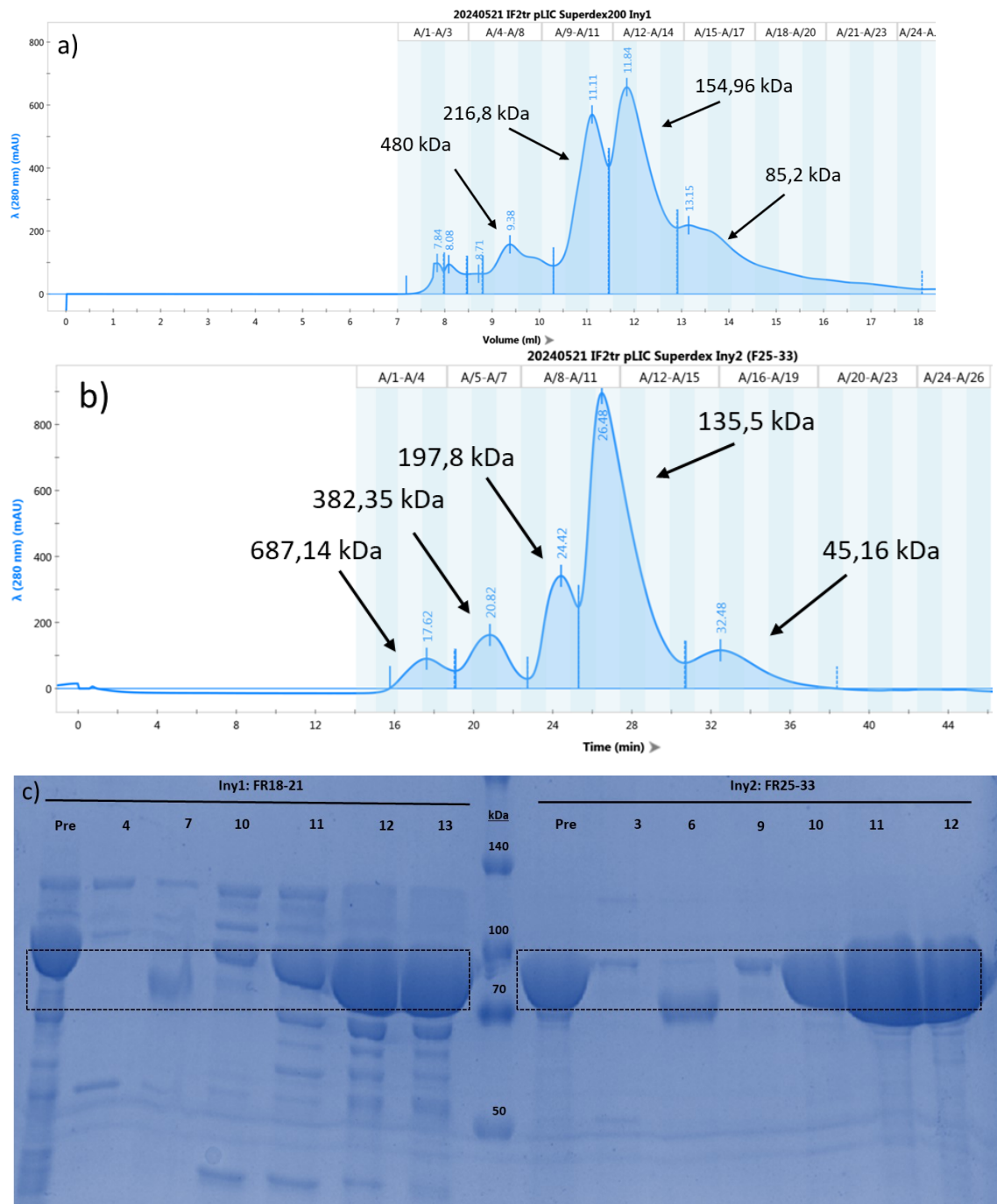


Figura 25. Resultado de la filtración en gel de los dos lotes de pLIC-IF2tr obtenidos tras la HisTrap. a) Cromatograma del primer pinchazo (Iny 1), con las fracciones F18-21 de la columna de níquel. b) Cromatograma del segundo pinchazo (Iny 2), con las fracciones F25-33. c) Gel SDS-PAGE (10% acrilamida) de los concentrados corridos en la Iny1 y 2 (Pre) y las fracciones recogidas a la salida de la columna; junto a 5 μ L de marcador de pesos moleculares Spectra™ Multicolor Broad Range (Thermo Scientific™). La zona recuadrada corresponde al peso molecular de la proteína de fusión IF2tr-6xHis (\approx 70 kDa).

Los dos cromatogramas mostraron un perfil de elución muy similar con un pico principal de 135-155 kDa (las diferencias pueden deberse a que la cantidad de muestra inyectada no era la misma y esto puede afectar al volumen de elución). Sin embargo, el segundo pinchazo, correspondiente al segundo *batch* (F25-33), estaba mucho más puro que el primero (Figura 25c). En la figura 25a, el

cromatograma de FR18-21 (10% de imidazol) mostró también un gran pico de contaminación anterior (216,8 kDa), por lo que se decidió descartar esta muestra y continuar solo con FR25-33.

Así pues, todas las fracciones de interés de la purificación del *batch 2*, se juntaron y concentraron hasta el máximo volumen posible. Finalmente, se centrifugó la muestra y se cuantificó, logrando purificar 10,3 mg totales de proteína IF2tr. Se dividió en alícuotas y se congeló con nitrógeno líquido, almacenándola a -80 °C en tampón *storage*.

El rendimiento final del protocolo de purificación de IF2tr fue mucho mejor que en IF2, dado que se obtuvo una cantidad de proteína total mayor, y también se logró alcanzar una concentración más alta sin que precipitara. Esto indica que la versión truncada sí es más estable que la completa en realidad, cumpliendo el propósito para el cual se diseñó, y que puede purificarse perfectamente con la cola de histidinas. Entonces, los problemas iniciales para obtenerla se debían verdaderamente a un mal plegamiento del constructo, cuya inteína fusionada no era funcional. Esto probablemente se deba a que, en el plegamiento final, no quede expuesta o correctamente organizada, por alguna incompatibilidad con la región que conecta la etiqueta con IF2tr.

4.9 Consideraciones finales y trabajos futuros

Antes de empezar con los experimentos, se comprobó en la bibliografía existente que ya estaba descrita la expresión y purificación de IF2 recombinante de Mtb en *E. coli* (Srivastava et al., 2016). Además, en el diseño de los constructos se tuvo en cuenta el alto contenido en GC del genoma de Mtb respecto a *E. coli*, lo cual conlleva que la mayoría de codones tienen un sesgo hacia terminar en C o G. Esto significa que serán menos usados por *E. coli*, obligando a la optimización de los codones para aumentar la tasa de expresión (Lakey et al., 2000). Con todo esto, pensábamos que la proteína iba a ser fácilmente obtenida desde un principio. Sin embargo, nos encontramos con varios inconvenientes durante el procedimiento, en ambas versiones de la proteína.

El primer impedimento fue expresar las proteínas con éxito. Dado que de forma normal con células BL21 no había buena expresión, se tuvo que recurrir a estrategias alternativas. Sin embargo, con ellas tampoco se consiguió solucionar el problema (Fig. 5). En la cepa pLysS no se obtuvo prácticamente expresión, pese a que la lisozima del fago T7 ha demostrado ser eficaz en la expresión de genes tóxicos para la célula (Studier, 1991), como puede ser IF2 de Mtb para *E. coli*. Tampoco hubo éxito con la co-expresión de chaperonas moleculares, ya sea con el vector pGroESL o induciendo su producción por estrés celular, las cuales pueden mejorar la expresión y reducir la agregación proteica en otros casos (Kolaj et al., 2009). Afortunadamente, al final se consiguió una expresión en cantidades aceptables con la adición de glucosa (Fig. 6), que también ha demostrado ser útil en la optimización de estos procesos (Kaur et al., 2018) por bloquear la expresión basal del operón *lac*.

En segundo lugar, los experimentos de purificación de IF2tr tampoco funcionaron inicialmente. Pese a que en un principio se preveía que trabajar con la versión truncada iba a ser más sencillo que con la completa, resultó que el constructo pTYB21-IF2tr presentaba problemas de agregación. Algunos resultados, como que no se unía a la columna de quitina y tampoco se daba la reacción de autoescisión, nos llevaron a pensar que la etiqueta inteína no era funcional. Tras ver que la proteína de fusión estaba mal plegada, mientras que pTYB21-IF2 no tenía problemas en el plegamiento, se concluyó que debía existir algún tipo de incompatibilidad de la inteína con la región *linker* que la conecta con la proteína. La empresa NEB indica que hay ciertos residuos desfavorables en cuanto a la eficiencia de escisión en este sistema, como P, C, N, D o R (NEB, s.f.). Aunque estos residuos no están en la secuencia *linker* del inicio de IF2tr, sí que es cierto que es una zona con residuos de carácter básico y repetidos entre sí (KRQKRQ). Esto podría ser uno de los motivos por los que la

inteína no funciona, a pesar de que sea una etiqueta de uso muy extendido, y que incluso ha demostrado aumentar la solubilidad de ciertas proteínas (Arnau et al., 2006; Khodarami et al., 2021). Otro aspecto importante a la hora de diseñar la proteína de fusión es decidir si la etiqueta inteína irá en el extremo N-terminal o C-terminal, puesto que la fusión en el C-t ha demostrado ser menos conflictiva en cuanto a expresión ineficiente e insolubilidad con respecto al N-t (Xu et al., 2000). De este modo, se decidió reclonar la proteína en otro vector, pLIC-SGC1, cuya etiqueta de histidinas sí resultó efectiva para purificarla hasta la homogeneidad. Ésta es la etiqueta de afinidad más frecuentemente usada en purificación de proteínas, y destaca por su pequeño tamaño y plegamiento sencillo (Zhao et al., 2013), dos características esenciales detrás del éxito en esta purificación. De hecho, otros autores establecieron un sistema de traducción *in vitro* de Mtb similar a nuestro objetivo, purificando todos los factores de traducción con etiquetas 6xHis (Srivastava et al., 2016).

Por otro lado, aunque la versión completa IF2 sí consiguió purificarse por el sistema IMPACT usando el vector original pTYB21, el rendimiento (cantidad de proteína total y concentración final) fue menor que con IF2tr-pLIC, demostrando que la versión completa sí es más inestable y complicada de manejar en el proceso de purificación, tal y como se predijo en el diseño de ambas versiones.

Algunas limitaciones o mejoras del trabajo incluirían responder a por qué ambas versiones (IF2, IF2tr) muestran un tamaño aparente mayor al real en los cromatogramas de filtración en gel. Esto pasa por determinar el estado oligomérico de la proteína, para ver si forma algún complejo homodimérico/trimérico. Para ello, se podría llevar a cabo un ensayo SEC-MALS, que combina la cromatografía de exclusión molecular (SEC), o filtración en gel, con una prueba de dispersión de luz de ángulo múltiple (MALS). La SEC estima el peso molecular de los picos a partir del volumen de elución y un patrón de proteínas cargado previamente, y supone un plegamiento globular de las proteínas de la muestra que puede llevar a errores; mientras que la técnica MALS permite comprobar el tamaño real y distinguir entre monómeros, oligómeros nativos o agregados (Some et al., 2019). Aun así, lo más probable es que el plegamiento irregular de IF2 (más fibrilar que globular) dé lugar a una estimación imprecisa de su peso molecular, puesto que no se ha descrito en la literatura la formación de oligómeros de esta proteína tan estudiada.

Por último, en el futuro cercano las proteínas purificadas en este trabajo permitirán la realización de estudios estructurales de microscopía electrónica. En este sentido, actualmente ya han servido para preparar rejillas de tinción negativa de complejos de preiniciación 30S (PIC), visibles en la figura 26, y se espera que en los próximos meses puedan ser utilizadas para la preparación de muestras de microscopía crioelectrónica (CryoEM). A medio y largo plazo, también se prevé que estas proteínas puedan ser utilizadas en la reconstitución de un sistema de traducción *in vitro* propio, para desarrollar pruebas de inhibición traduccional con fármacos y desarrollar más investigación básica sobre la iniciación en Mtb.

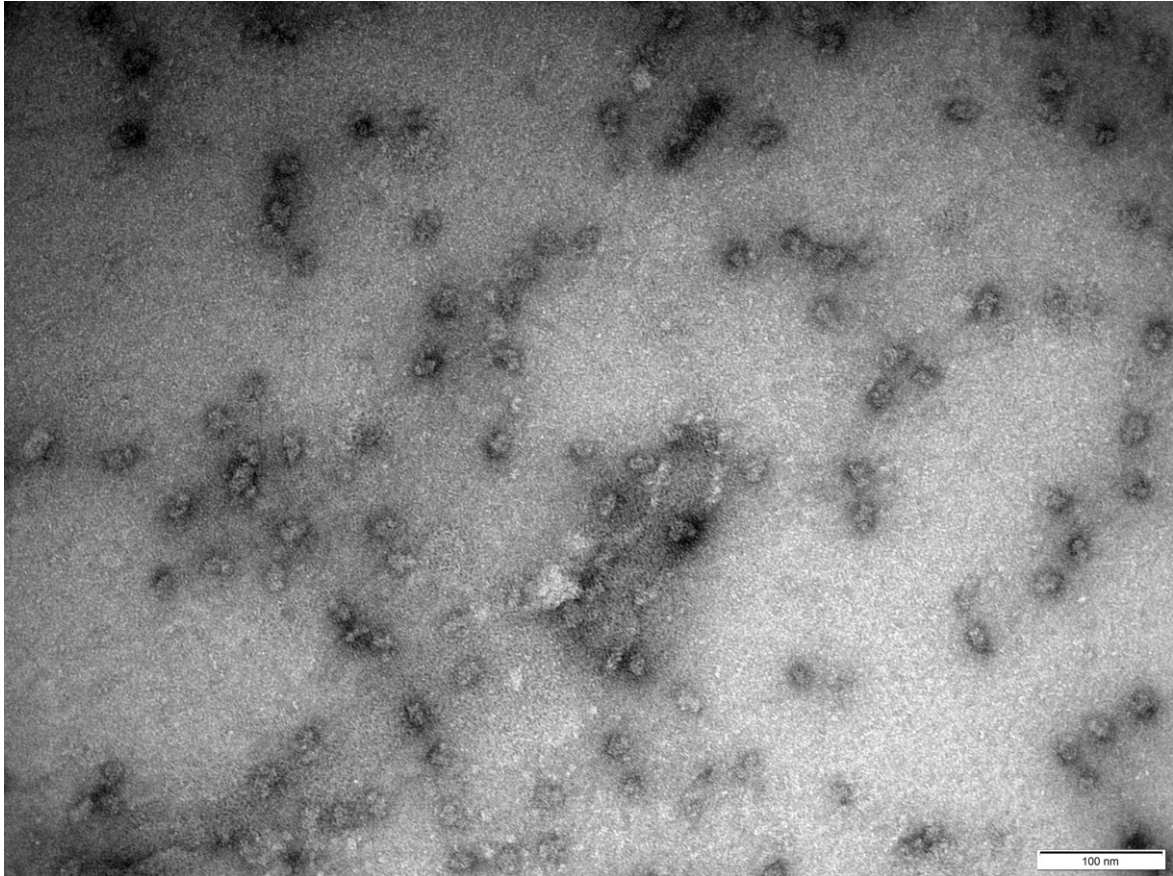


Figura 26. Imagen al microscopio electrónico de transmisión (TEM) de la tinción negativa de complejos de preiniciación 30S. Se mezclaron *in vitro* diferentes componentes purificados de la maquinaria traduccional de micobacterias: ribosomas 30S (de *M. smegmatis*), los factores de iniciación IF1, IF2tr e IF3 (purificados de forma recombinante en *E. coli*) y el ARNt iniciador (tRNA^{fMet}). Como agente de tinción se empleó acetato de uranilo 0,5%.

5. CONCLUSIONES

Durante la realización del presente trabajo de fin de grado se han logrado cumplir los objetivos propuestos al principio, superando los obstáculos encontrados y proponiendo nuevas soluciones a ellos. Las conclusiones extraídas son las siguientes:

- La proteína IF2 de *M. tuberculosis* y su versión truncada IF2tr pueden expresarse de forma heteróloga en *E. coli*, obteniendo en ambos casos la proteína de forma soluble. Además, en el caso de IF2tr, la proteína purificada es funcional ya que permite el ensamblaje de complejos de iniciación tal y como se ha comprobado mediante microscopía electrónica de tinción negativa.
- La elección del método de purificación es esencial para obtener resultados satisfactorios; por ejemplo, IF2 se puede purificar empleando el protocolo IMPACT con una etiqueta inteína, mientras que su versión truncada IF2tr requiere una etiqueta de afinidad 6xHis.
- La realización previa de pruebas de expresión y purificación a pequeña escala es esencial para optimizar las condiciones de estos procesos a gran escala.
- En la expresión heteróloga de proteínas, si se trabaja con proteínas potencialmente tóxicas para la célula, como puede ser IF2 de Mtb en *E. coli*, es fundamental reprimir su expresión basal antes de la inducción. En el caso del operón *lac*, la adición de glucosa al 1% resultó clave para asegurar un crecimiento adecuado de las células.

6. BIBLIOGRAFÍA

Allen, G. S., Zavialov, A., Gursky, R., Ehrenberg, M., & Frank, J. (2005). The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell*, 121(5), 703–712. <https://doi.org/10.1016/j.cell.2005.03.023>

Alsayed, S. S. R., & Gunosewoyo, H. (2023). Tuberculosis: Pathogenesis, Current Treatment Regimens and New Drug Targets. *International journal of molecular sciences*, 24(6), 5202. <https://doi.org/10.3390/ijms24065202>

Arnau, J., Lauritzen, C., Petersen, G. E., & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein expression and purification*, 48(1), 1–13. <https://doi.org/10.1016/j.pep.2005.12.002>

Aslanidis, C., & de Jong, P. J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic acids research*, 18(20), 6069–6074. <https://doi.org/10.1093/nar/18.20.6069>

Basu, R. S., Sherman, M. B., & Gagnon, M. G. (2022). Compact IF2 allows initiator tRNA accommodation into the P site and gates the ribosome to elongation. *Nature communications*, 13(1), 3388. <https://doi.org/10.1038/s41467-022-31129-2>

Boni, I. V., Isaeva, D. M., Musychenko, M. L., & Tzareva, N. V. (1991). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic acids research*, 19(1), 155–162. <https://doi.org/10.1093/nar/19.1.155>

Bonsor, D., Butz, S. F., Solomons, J., Grant, S., Fairlamb, I. J., Fogg, M. J., & Grogan, G. (2006). Ligation independent cloning (LIC) as a rapid route to families of recombinant biocatalysts from sequenced prokaryotic genomes. *Organic & biomolecular chemistry*, 4(7), 1252–1260. <https://doi.org/10.1039/b517338h>

Bradford M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, 72, 248–254. <https://doi.org/10.1006/abio.1976.9999>

Caserta, E., Tomsic, J., Spurio, R., La Teana, A., Pon, C. L., & Gualerzi, C. O. (2006). Translation initiation factor IF2 interacts with the 30 S ribosomal subunit via two separate binding sites. *Journal of molecular biology*, 362(4), 787–799. <https://doi.org/10.1016/j.jmb.2006.07.043>

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., ... Barrell, B. G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 393(6685), 537–544. <https://doi.org/10.1038/31159>

Cortes, T., Schubert, O. T., Rose, G., Arnvig, K. B., Comas, I., Aebersold, R., & Young, D. B. (2013). Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in *Mycobacterium tuberculosis*. *Cell reports*, 5(4), 1121–1131. <https://doi.org/10.1016/j.celrep.2013.10.031>

Dutta, N. K., & Karakousis, P. C. (2014). Latent tuberculosis infection: myths, models, and molecular mechanisms. *Microbiology and molecular biology reviews : MMBR*, 78(3), 343–371. <https://doi.org/10.1128/MMBR.00010-14>

Grill, S., Gualerzi, C. O., Londei, P., & Bläsi, U. (2000). Selective stimulation of translation of leaderless mRNA by initiation factor 2: evolutionary implications for translation. *The EMBO journal*, 19(15), 4101–4110. <https://doi.org/10.1093/emboj/19.15.4101>

Houben, R. M., & Dodd, P. J. (2016). The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS medicine*, 13(10), e1002152. <https://doi.org/10.1371/journal.pmed.1002152>

Huang, L., Nazarova, E. V., & Russell, D. G. (2019). *Mycobacterium tuberculosis*: Bacterial Fitness within the Host Macrophage. *Microbiology spectrum*, 7(2), 10.1128/microbiolspec.BAI-0001-2019. <https://doi.org/10.1128/microbiolspec.BAI-0001-2019>

IMPACT™ affinity tag (s.f.). New England Biolabs® (NEB). Consultado el 13 de junio de 2024. [https://www.neb.com/en/applications/protein-purification/affinity-purification-and-expression-tags/impact-affinity-tag/#:~:text=The%20IMPACT%E2%84%A2%20\(Intein%20Mediated,purification%20using%20a%20chitin%20resin.](https://www.neb.com/en/applications/protein-purification/affinity-purification-and-expression-tags/impact-affinity-tag/#:~:text=The%20IMPACT%E2%84%A2%20(Intein%20Mediated,purification%20using%20a%20chitin%20resin.)

IMPACT™ Vectors and Applications (s.f.). New England Biolabs® (NEB). Consultado el 15 de junio de 2024. <https://www.neb.com/en/tools-and-resources/selection-charts/impact-vectors-and-applications>

Ishii N. (2017). GroEL and the GroEL-GroES Complex. *Sub-cellular biochemistry*, 83, 483–504. https://doi.org/10.1007/978-3-319-46503-6_17

Kaur, J., Kumar, A., & Kaur, J. (2018). Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. *International journal of biological macromolecules*, 106, 803–822. <https://doi.org/10.1016/j.ijbiomac.2017.08.080>

Khodakarami, A., Dabirmanesh, B., Asad, S., & Khaledi, M. (2021). Enhanced Solubility and One-Step Purification of Functional Dimeric Carboxypeptidase G2. *Biochemistry. Biokhimiia*, 86(2), 190–196. <https://doi.org/10.1134/S0006297921020073>

Koch, A., & Mizrahi, V. (2018). *Mycobacterium tuberculosis*. *Trends in microbiology*, 26(6), 555–556. <https://doi.org/10.1016/j.tim.2018.02.012>

Kohler, R., Mooney, R. A., Mills, D. J., Landick, R., & Cramer, P. (2017). Architecture of a transcribing-translating expressome. *Science (New York, N.Y.)*, 356(6334), 194–197. <https://doi.org/10.1126/science.aal3059>

Kolaj, O., Spada, S., Robin, S., & Wall, J. G. (2009). Use of folding modulators to improve heterologous protein production in *Escherichia coli*. *Microbial cell factories*, 8, 9. <https://doi.org/10.1186/1475-2859-8-9>

Komarova, A. V., Tchufistova, L. S., Dreyfus, M., & Boni, I. V. (2005). AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *Journal of bacteriology*, 187(4), 1344–1349. <https://doi.org/10.1128/JB.187.4.1344-1349.2005>

Kroemer, T. (s.f.). Una revisión general sobre clonación molecular. *Goldbio*. Consultado el 15 de junio de 2024. <https://goldbio.com/articles/article/SOBRE-CLONACION-MOLECULAR>

Kumar, N., Sharma, S., & Kaushal, P. S. (2021). Protein synthesis in *Mycobacterium tuberculosis* as a potential target for therapeutic interventions. *Molecular aspects of medicine*, 81, 101002. <https://doi.org/10.1016/j.mam.2021.101002>

- Kumar, N., Sharma, S., & Kaushal, P. S.** (2024). Cryo- EM structure of the mycobacterial 70S ribosome in complex with ribosome hibernation promotion factor RafH. *Nature communications*, 15(1), 638. <https://doi.org/10.1038/s41467-024-44879-y>
- Laalami, S., Putzer, H., Plumbridge, J. A., & Grunberg-Manago, M.** (1991). A severely truncated form of translational initiation factor 2 supports growth of *Escherichia coli*. *Journal of molecular biology*, 220(2), 335–349. [https://doi.org/10.1016/0022-2836\(91\)90017-z](https://doi.org/10.1016/0022-2836(91)90017-z)
- Laemmli U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680–685. <https://doi.org/10.1038/227680a0>
- Lakey, D. L., Voladri, R. K., Edwards, K. M., Hager, C., Samten, B., Wallis, R. S., Barnes, P. F., & Kernodle, D. S.** (2000). Enhanced production of recombinant *Mycobacterium tuberculosis* antigens in *Escherichia coli* by replacement of low-usage codons. *Infection and immunity*, 68(1), 233–238. <https://doi.org/10.1128/IAI.68.1.233-238.2000>
- Lee, J. J., Bruley, D. F., & Kang, K. A.** (2007). Effect of pH and imidazole on protein C purification from Cohn fraction IV-1 by IMAC. *Advances in experimental medicine and biology*, 599, 61–66. https://doi.org/10.1007/978-0-387-71764-7_9
- Leiva, L. E., & Katz, A.** (2022). Regulation of Leaderless mRNA Translation in Bacteria. *Microorganisms*, 10(4), 723. <https://doi.org/10.3390/microorganisms10040723>
- López-Alonso, J. P., Fabbretti, A., Kaminishi, T., Iturrioz, I., Brandi, L., Gil-Carton, D., Gualerzi, C. O., Fucini, P., & Connell, S. R.** (2017). Structure of a 30S pre-initiation complex stalled by GE81112 reveals structural parallels in bacterial and eukaryotic protein synthesis initiation pathways. *Nucleic acids research*, 45(4), 2179–2187. <https://doi.org/10.1093/nar/gkw1251>
- Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R., & Butcher, S.** (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic acids research*, gkae241. Advance online publication. <https://doi.org/10.1093/nar/gkae241>
- Madison B. M.** (2001). Application of stains in clinical microbiology. *Biotechnic & histochemistry: official publication of the Biological Stain Commission*, 76(3), 119–125
- Margus, T., Remm, M., & Tenson, T.** (2007). Phylogenetic distribution of translational GTPases in bacteria. *BMC genomics*, 8, 15. <https://doi.org/10.1186/1471-2164-8-15>
- Moll, I., Hirokawa, G., Kiel, M. C., Kaji, A., & Bläsi, U.** (2004). Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. *Nucleic acids research*, 32(11), 3354–3363. <https://doi.org/10.1093/nar/gkh663>
- Morizono, H., Tuchman, M., Rajagopal, B. S., McCann, M. T., Listrom, C. D., Yuan, X., Venugopal, D., Barany, G., & Allewell, N. M.** (1997). Expression, purification and kinetic characterization of wild-type human ornithine transcarbamylase and a recurrent mutant that produces 'late onset' hyperammonaemia. *The Biochemical journal*, 322 (Pt 2)(Pt 2), 625–631. <https://doi.org/10.1042/bj3220625>
- Myasnikov, A. G., Marzi, S., Simonetti, A., Giuliodori, A. M., Gualerzi, C. O., Yusupova, G., Yusupov, M., & Klaholz, B. P.** (2005). Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. *Nature structural & molecular biology*, 12(12), 1145–1149. <https://doi.org/10.1038/nsmb1012>

- O'Fágáin, C., Cummins, P. M., & O'Connor, B. F.** (2011). Gel-filtration chromatography. *Methods in molecular biology (Clifton, N.J.)*, 681, 25–33. https://doi.org/10.1007/978-1-60761-913-0_2
- Rodnina M. V.** (2018). Translation in Prokaryotes. *Cold Spring Harbor perspectives in biology*, 10(9), a032664. <https://doi.org/10.1101/cshperspect.a032664>
- Roy, B., Liu, Q., Shoji, S., & Fredrick, K.** (2018). IF2 and unique features of initiator tRNA^{fMet} help establish the translational reading frame. *RNA biology*, 15(4-5), 604–613. <https://doi.org/10.1080/15476286.2017.1379636>
- Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., & Bontems, F.** (2009). Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic acids research*, 37(16), 5578–5588. <https://doi.org/10.1093/nar/gkp547>
- Sawyer, E. B., & Cortes, T.** (2022). Ribosome profiling enhances understanding of mycobacterial translation. *Frontiers in microbiology*, 13, 976550. <https://doi.org/10.3389/fmicb.2022.976550>
- Sawyer, E. B., Phelan, J. E., Clark, T. G., & Cortes, T.** (2021). A snapshot of translation in Mycobacterium tuberculosis during exponential growth and nutrient starvation revealed by ribosome profiling. *Cell reports*, 34(5), 108695. <https://doi.org/10.1016/j.celrep.2021.108695>
- Schrader, S. M., Botella, H., Jansen, R., Ehrt, S., Rhee, K., Nathan, C., & Vaubourgeix, J.** (2021). Multifunctional antimicrobial resistance from a metabolic mutation. *Science advances*, 7(35), eabh2037. <https://doi.org/10.1126/sciadv.abh2037>
- Shell, S. S., Wang, J., Lapierre, P., Mir, M., Chase, M. R., Pyle, M. M., Gawande, R., Ahmad, R., Sarracino, D. A., Ioerger, T. R., Fortune, S. M., Derbyshire, K. M., Wade, J. T., & Gray, T. A.** (2015). Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape. *PLoS genetics*, 11(11), e1005641. <https://doi.org/10.1371/journal.pgen.1005641>
- Shine, J., & Dalgarno, L.** (1974). The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), 1342–1346. <https://doi.org/10.1073/pnas.71.4.1342>
- Simonetti, A., Marzi, S., Billas, I. M., Tsai, A., Fabbretti, A., Myasnikov, A. G., Roblin, P., Vaiana, A. C., Hazemann, I., Eiler, D., Steitz, T. A., Puglisi, J. D., Gualerzi, C. O., & Klaholz, B. P.** (2013). Involvement of protein IF2 N domain in ribosomal subunit joining revealed from architecture and function of the full-length initiation factor. *Proceedings of the National Academy of Sciences of the United States of America*, 110(39), 15656–15661. <https://doi.org/10.1073/pnas.1309578110>
- Simonetti, A., Marzi, S., Myasnikov, A. G., Fabbretti, A., Yusupov, M., Gualerzi, C. O., & Klaholz, B. P.** (2008). Structure of the 30S translation initiation complex. *Nature*, 455(7211), 416–420. <https://doi.org/10.1038/nature07192>
- Smith I.** (2003). Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clinical microbiology reviews*, 16(3), 463–496. <https://doi.org/10.1128/CMR.16.3.463-496.2003>
- Some, D., Amartely, H., Tsadok, A., & Lebendiker, M.** (2019). Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *Journal of visualized experiments: JoVE*, (148), 10.3791/59615. <https://doi.org/10.3791/59615>

- Spurio, R., Brandi, L., Caserta, E., Pon, C. L., Gualerzi, C. O., Misselwitz, R., Krafft, C., Welfle, K., & Welfle, H. (2000). The C-terminal subdomain (IF2 C-2) contains the entire fMet-tRNA binding site of initiation factor IF2. *The Journal of biological chemistry*, 275(4), 2447–2454. <https://doi.org/10.1074/jbc.275.4.2447>
- Srivastava, A., Asahara, H., Zhang, M., Zhang, W., Liu, H., Cui, S., Jin, Q., & Chong, S. (2016). Reconstitution of Protein Translation of Mycobacterium Reveals Functional Conservation and Divergence with the Gram-Negative Bacterium Escherichia coli. *PLoS one*, 11(8), e0162020. <https://doi.org/10.1371/journal.pone.0162020>
- Studier F. W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of molecular biology*, 219(1), 37–44. [https://doi.org/10.1016/0022-2836\(91\)90855-z](https://doi.org/10.1016/0022-2836(91)90855-z)
- Suárez, I., Fünfer, S. M., Kröger, S., Rademacher, J., Fätkenheuer, G., & Rybniker, J. (2019). The Diagnosis and Treatment of Tuberculosis. *Deutsches Arzteblatt international*, 116(43), 729–735. <https://doi.org/10.3238/arztebl.2019.0729>
- Weiner, J., 3rd, Herrmann, R., & Browning, G. F. (2000). Transcription in Mycoplasma pneumoniae. *Nucleic acids research*, 28(22), 4488–4496. <https://doi.org/10.1093/nar/28.22.4488>
- WHO (World Health Organization) (2023). Global tuberculosis report 2023. Licence: CC BY-NC-SA 3.0 IGO. <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023>
- Wienk, H., Tishchenko, E., Belardinelli, R., Tomaselli, S., Dongre, R., Spurio, R., Folkers, G. E., Gualerzi, C. O., & Boelens, R. (2012). Structural dynamics of bacterial translation initiation factor IF2. *The Journal of biological chemistry*, 287(14), 10922–10932. <https://doi.org/10.1074/jbc.M111.333393>
- Xu, M. Q., Paulus, H., & Chong, S. (2000). Fusions to self-splicing inteins for protein purification. *Methods in enzymology*, 326, 376–418. [https://doi.org/10.1016/s0076-6879\(00\)26066-7](https://doi.org/10.1016/s0076-6879(00)26066-7)
- Yang, K., Chang, J. Y., Cui, Z., Li, X., Meng, R., Duan, L., Thongchol, J., Jakana, J., Huwe, C. M., Sacchettini, J. C., & Zhang, J. (2017). Structural insights into species-specific features of the ribosome from the human pathogen Mycobacterium tuberculosis. *Nucleic acids research*, 45(18), 10884–10894. <https://doi.org/10.1093/nar/gkx785>
- Zhao, X., Li, G., & Liang, S. (2013). Several affinity tags commonly used in chromatographic purification. *Journal of analytical methods in chemistry*, 2013, 581093. <https://doi.org/10.1155/2013/581093>
- Zheng, X., Hu, G. Q., She, Z. S., & Zhu, H. (2011). Leaderless genes in bacteria: clue to the evolution of translation initiation mechanisms in prokaryotes. *BMC genomics*, 12, 361. <https://doi.org/10.1186/1471-2164-12-361>