

Expresión de proteínas recombinantes

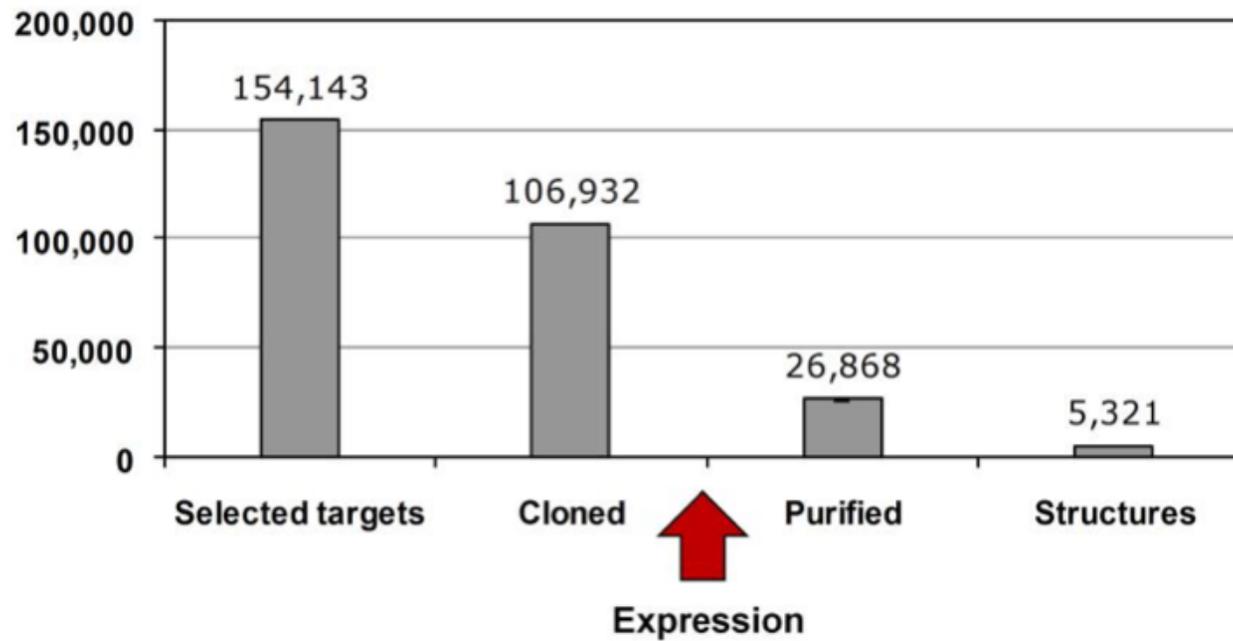
Nuria Sánchez Puig
Departamento de Química de
Biomacromoléculas
Instituto de Química, UNAM.

¿Por qué sobre-expresar proteínas?

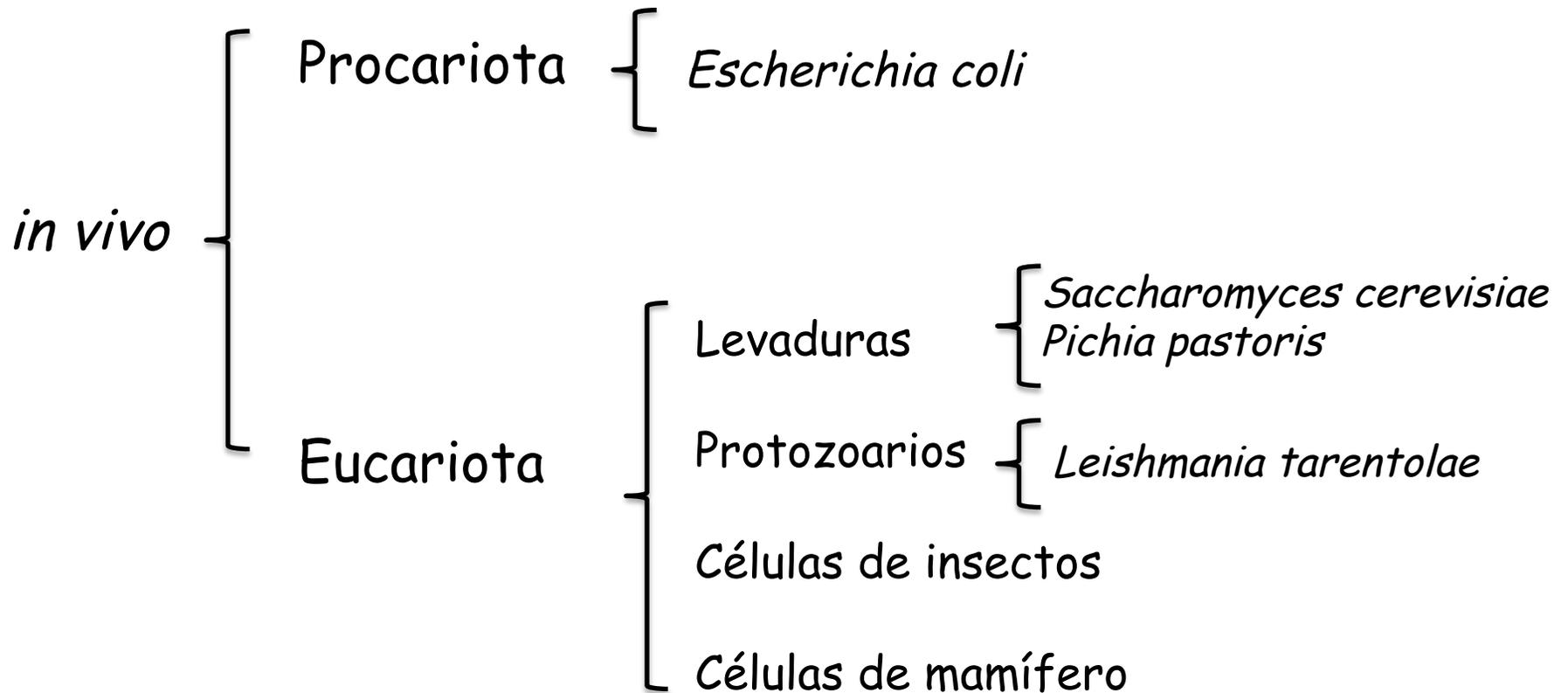
- Para estudios estructurales
Cristalografía, NMR, EM
- Estudios funcionales
Enzimología, inmunoprecipitación, estudios biofísicos
- Como antígenos para la producción de anticuerpos
- Como anticuerpos usados como fármacos
- Como suplementos para cultivos celulares
Factores de crecimiento

Pareciera simple

Attrition rate in Structural Genomics programmes



Sistemas de sobre-expresión de proteínas en sistemas heterólogos



in vitro - sistemas de transcripción-traducción libre de células

Sistemas de transcripción-traducción libre de células

- Extractos celulares involucrados en altos niveles de síntesis proteica.

reticulocitos de ratón, germen de trigo ó *Escherichia coli*

- Componentes macromoleculares para producir RNA:
ribosomas, tRNAs, aminoacil-tRNA sintetasas, factores de iniciación, elongación y terminación, etc,
- Suplementos energéticos
amino ácidos, fuentes de energía ATP y GTP, sistemas de regeneración energética (fosfoenol-piruvato y piruvato cinasa)

Sistemas de transcripción-traducción libre de células

Ventajas

Proteínas tóxicas no son eliminadas por las células

Incorporación de amino ácidos no naturales - ^{15}N , ^{13}C

La reacción se pueden llevar acabo en batch ó en reactores semi-continuos donde los "nutrientes" pueden suplementarse

Desventajas

Variabilidad de lote a lote

Baja cantidad de proteína funcional

¿Por qué sobre-expresar proteínas en sistemas eucariotas?

Modificaciones post-traduccionales

O- y N-glicosilación

Procesamiento de señales: pre- y prepro-péptido

Formación de puentes disulfuro

Presencia de chaperonas que favorecen el plegamiento

En levadura es fácil alcanzar grandes cantidades de biomasa

¿Por qué **NO** sobre-expresar proteínas en células eucariotas?

Células de mamífero

Costo elevado \$\$\$

Requiere de equipo especializado

Rendimiento proteico bajo

Difícil de escalar a volúmenes grandes

Componentes usados en los medios de cultivo inseguros? - suero bovino

Células de levadura

Menos variedad de cepas y plásmidos disponibles

Difíciles de romper

Expresión estable - creación de la cepa o línea celular mediante recombinación

Células de mamífero

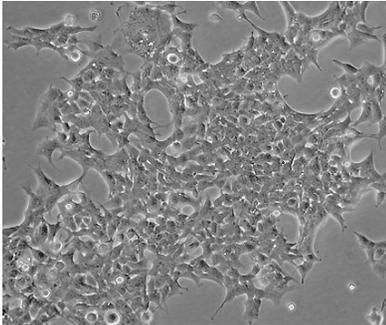
Transfección transitoria

- Gen a la proteína en días
- Verificar la expresión
- Bajo rendimiento
- Usado en estudios de alta densidad

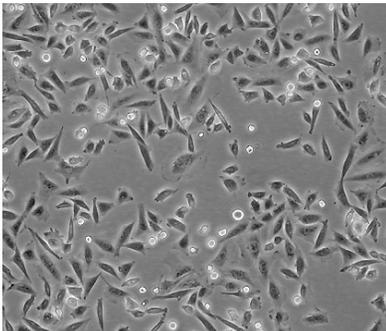
Transfección estable

- Gen a la proteína en días \geq 2 meses
- Proceso complejo
- El gene de interés se integra en el genoma
- Rendimientos altos (1-5 mg/l cultivo)

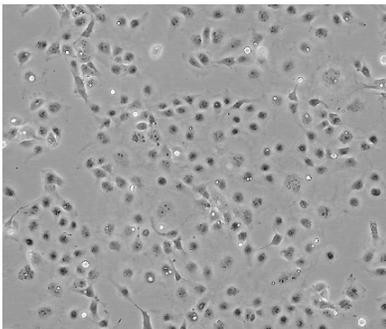
Células de mamífero comúnmente utilizadas



HEK 293: Células de riñón embriónico humano



CHO: Células de ovario de hamster chino



COS: Fibroblastos de simio

Promotores regulables comunmente utilizados para la sobre-expresión de proteínas en levadura

S. cerevisiae

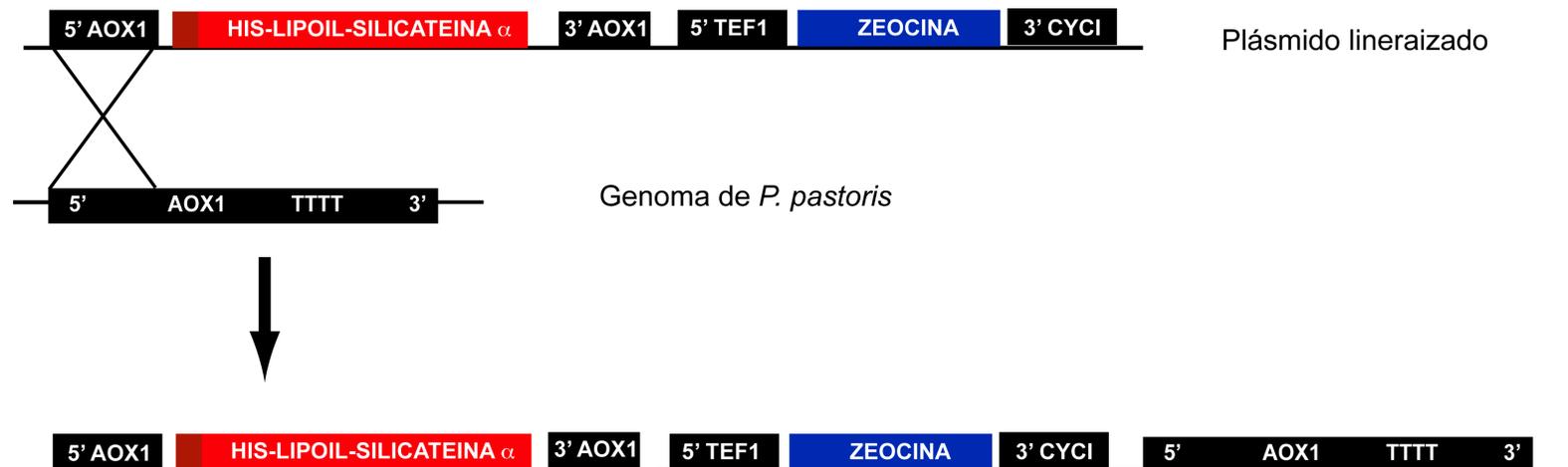
Promotor GAL 1-10

Reprimido por glucosa e
inducible por galactosa

P. pastoris

Promotor AOX1 y AOX2

Inducibles por metanol



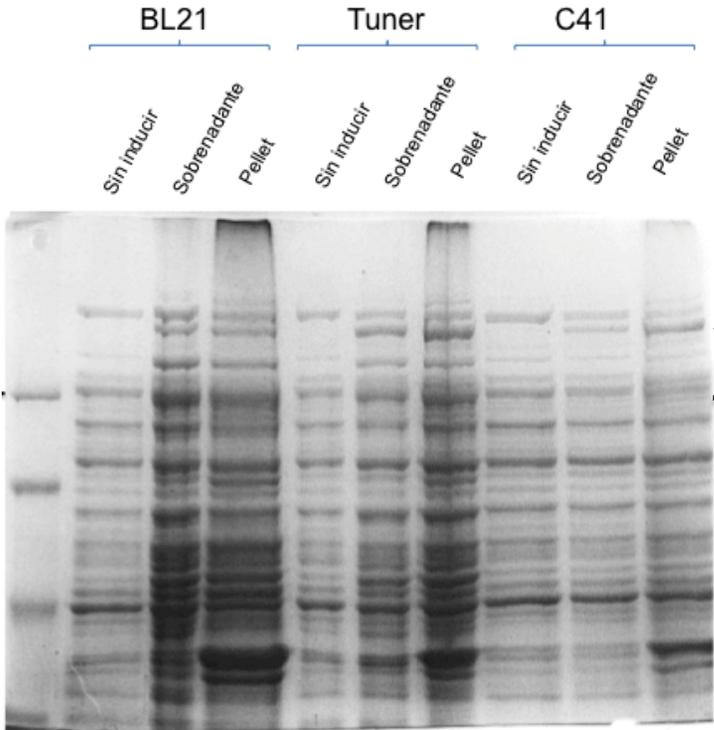
Plásmido linealizado

Genoma de *P. pastoris*

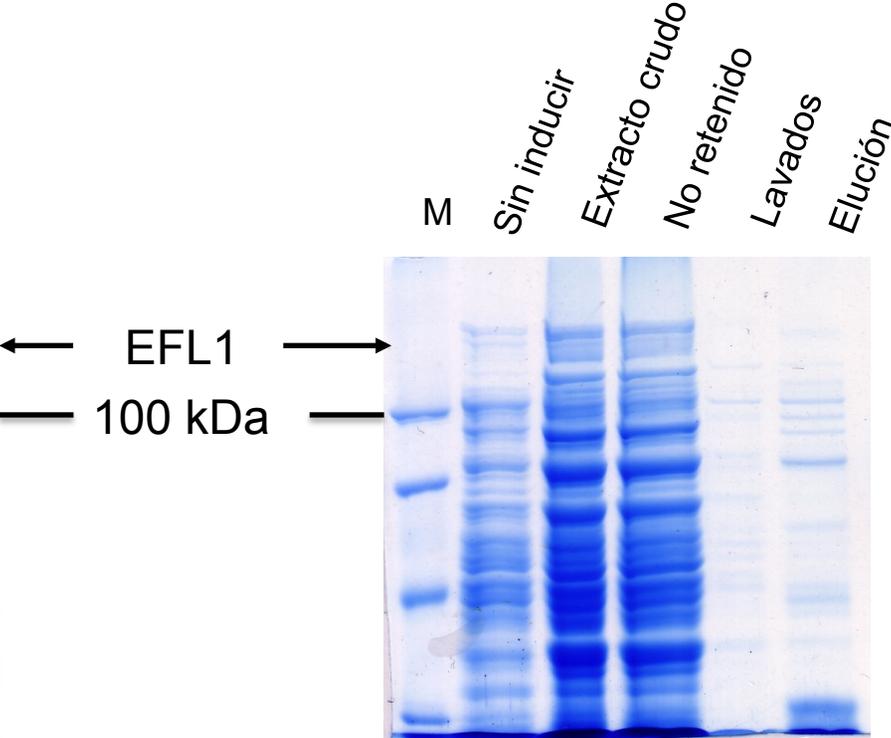
A pesar de las desventajas en ocasiones sólo un sistema de expresión eucarionte puede expresar mi proteína de interés

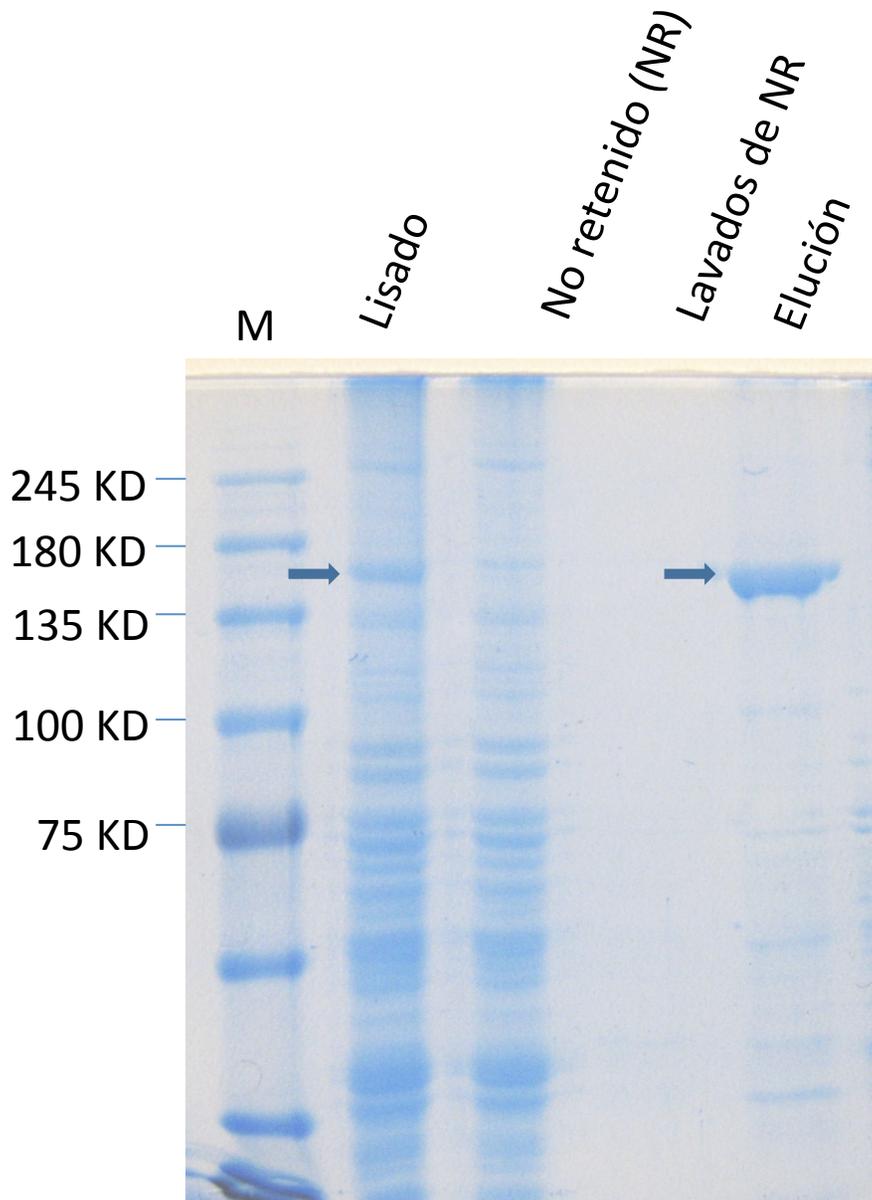
La proteína EFL1 se expresa en células bacterianas pero no parece se funcional

Expresión de EFL1 en *E. coli*



Purificación de EFL1



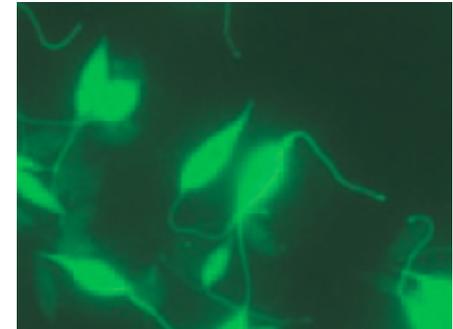


EFL1 expresadas en
S. cerevisiae BCY123 es
funcional

Leishmania expression system - **LEXSY** (Jena Bioscience)

Protozoario *Leishmania tarentolae* - organismo eucarionte unicelular, flagelado.

Aislado del geko *Tarentolae annularis* y *T. mauritanica* - inocuo para mamíferos.



- Simplificada en el manejo - no cultivo de tejidos
- Velocidad de crecimiento mayor que células de mamífero e insecto
- Modificaciones postraduccionales equivalentes a las de células de mamífero
- Altos rendimientos de proteína (mg/L cultivo)
- Caro \$\$\$
- No hay mucha variedad de vectores

¿Por qué sobre-expresar proteínas en *E. coli*?

- Es un sistema bien establecido.
- Fácil de manipular.
- Existe una gran variedad de vectores y cepas de expresión.
- Inocuo, no requiere infraestructura sofisticada, bajo costo.
- Apto para incorporar diferentes amino ácidos
 - seleno-metionina, seleno-cisteína, amino ácidos no naturales
 - Isotopos para NMR: ^{13}C , ^{15}N , ^2H

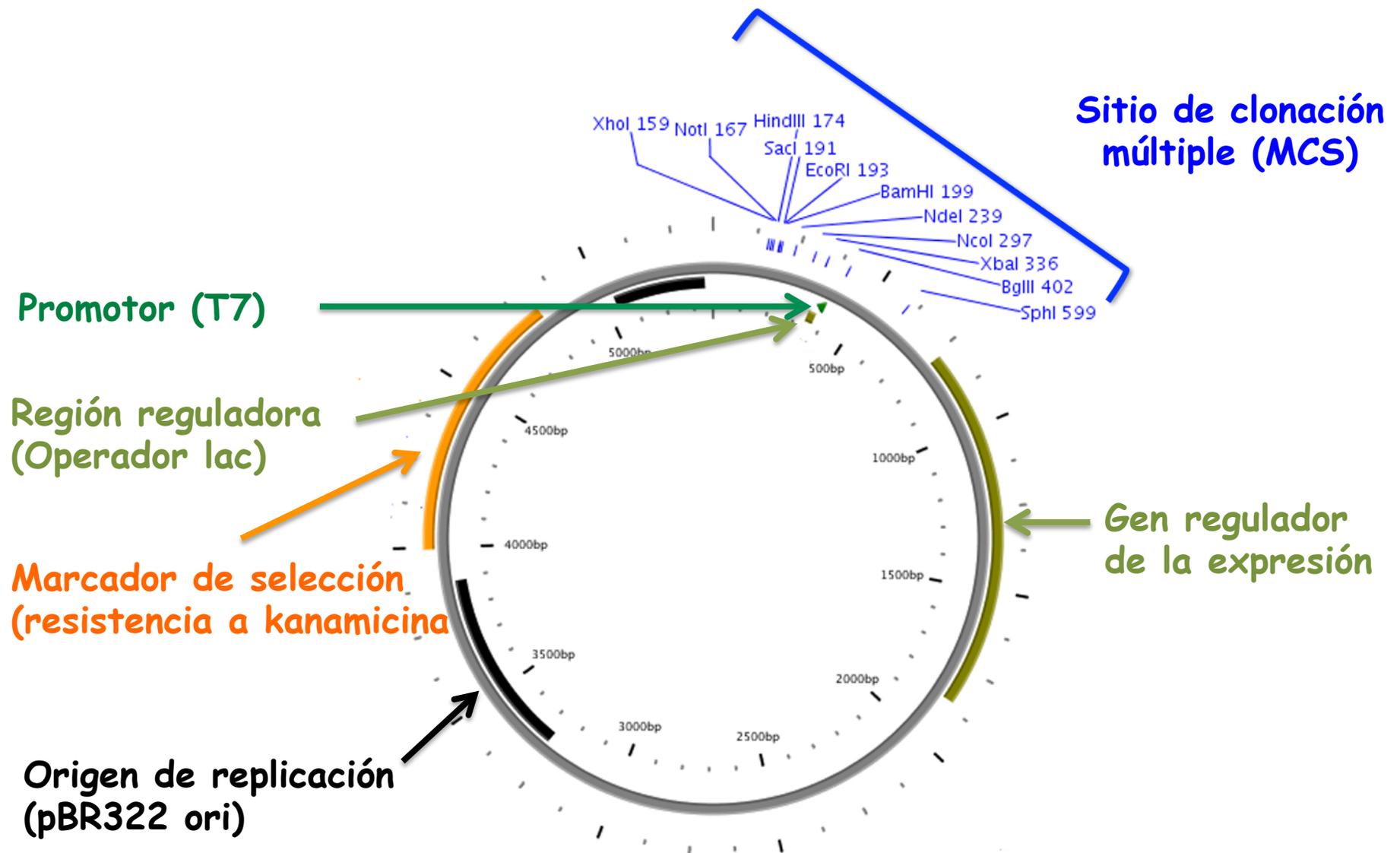
¿Por qué **NO** sobre-expresar proteínas en *E. coli*?

- Las modificaciones post-traduccionales de eucariontes no se llevan acabo.
- Poca o nula expresión
 - Marco de lectura erróneo
 - Inestabilidad en el RNAm ó en la proteína
 - Uso de codones inadecuado
- Problemas de solubilidad
 - Falta de chaperonas.
 - Niveles y velocidad de expresión demasiado alto comparado con la velocidad de plegamiento
 - Falta de ligando natural



Formación de cuerpos de inclusión

Anatomía de un vector de expresión en *E. coli*



Vectores comerciales más usados para sobre-expresar proteínas en *E. coli*

	Promotor	Origen de replicación	Proteína de fusión	Etiqueta	Inductor
pET (Novagen)	T7	pBR322	Trx, DsbA, GST, DsbC, Nus A, pEL B	His, V5, myc	IPTG
pGEX (Pharmacia)	trc	pBR322	GST	His	IPTG
pQE (QIAGEN)	T5	pUC	-	His	IPTG
pBAD (Invitrogen)	<i>ara</i>	pBR322	Trx	myc	L-arabinosa
pMAL (NEB)	<i>tac</i>	pBR322	MBP	His	IPTG

Promotores regulables más usados para sobre-expresar proteínas recombinantes

Promotor *tac* y promotor *trc*

Es un promotor híbrido que combina regiones del promotor triptófano y lactosa.

región -35 del promotor *trp*

región -10 del promotor *lac*

Regulado por el represor *lac*

3x más fuerte que el promotor *trp*

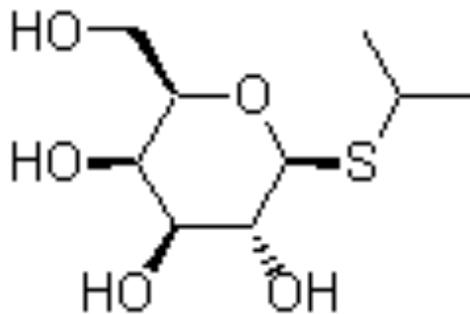
10x más fuerte que el promotor *lac*

Difieren en una bp

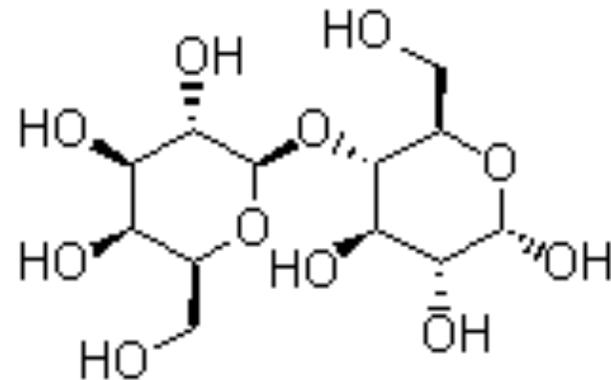
Promotor *lac*

- Regulado positivamente por lactosa o sus análogos, e.g. IPTG análogo no-hidrolizable.
- Represión por catabolito (glucosa).
- Reprimido por el represor *lac* (*lacI*)

IPTG
Isopropil- β -D-tiogalactósido

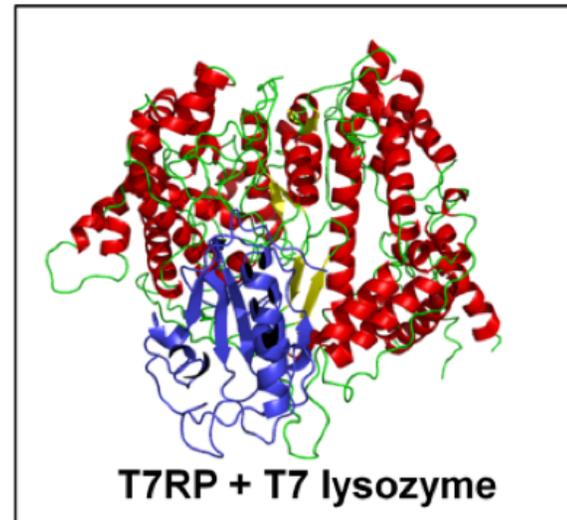
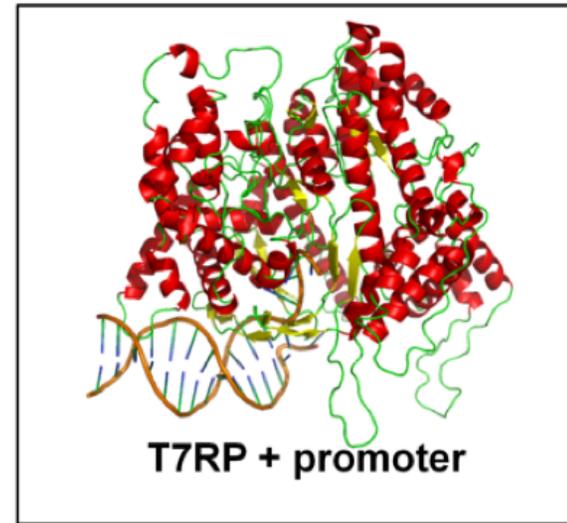


Lactosa

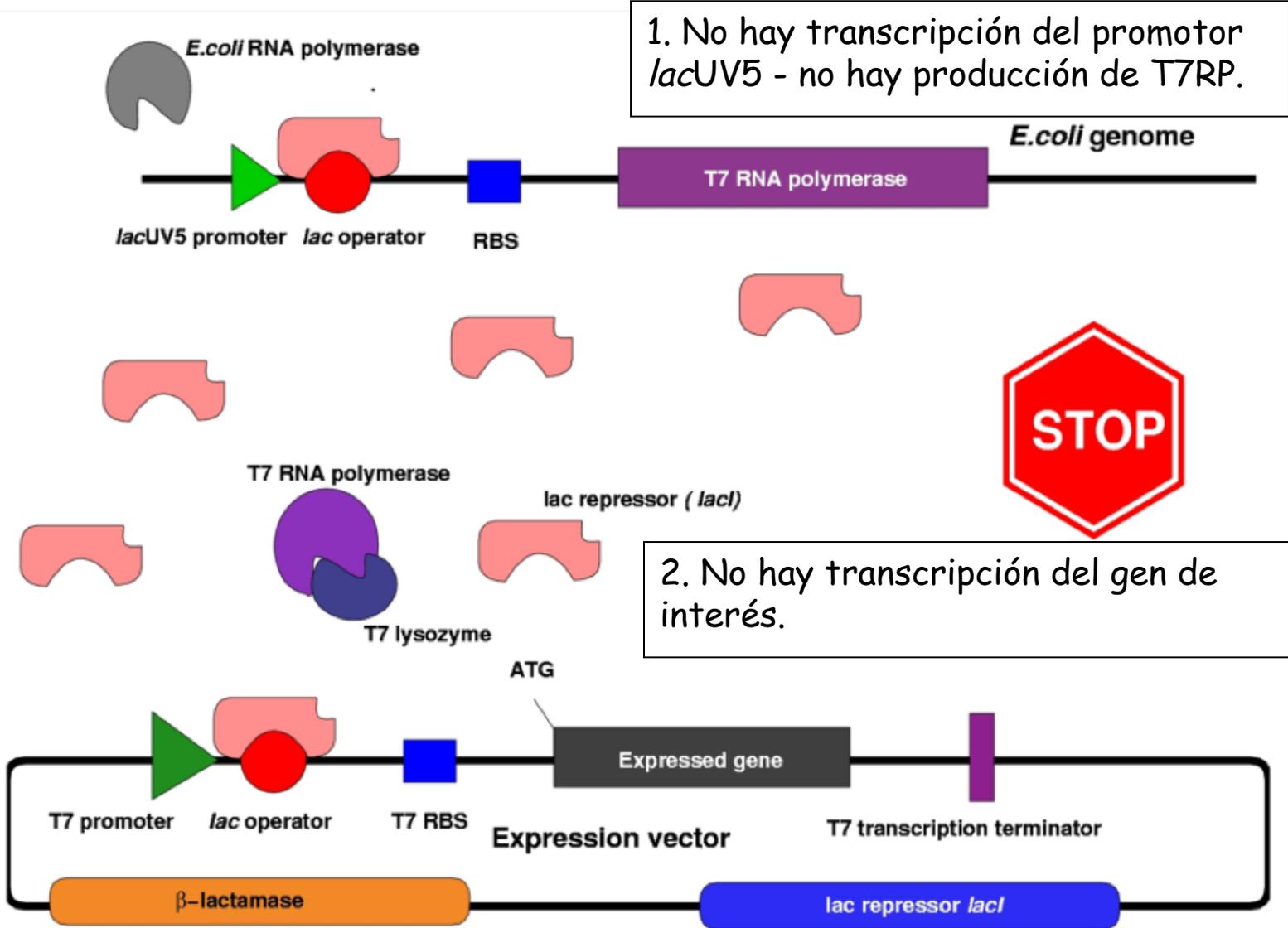


El sistema T7

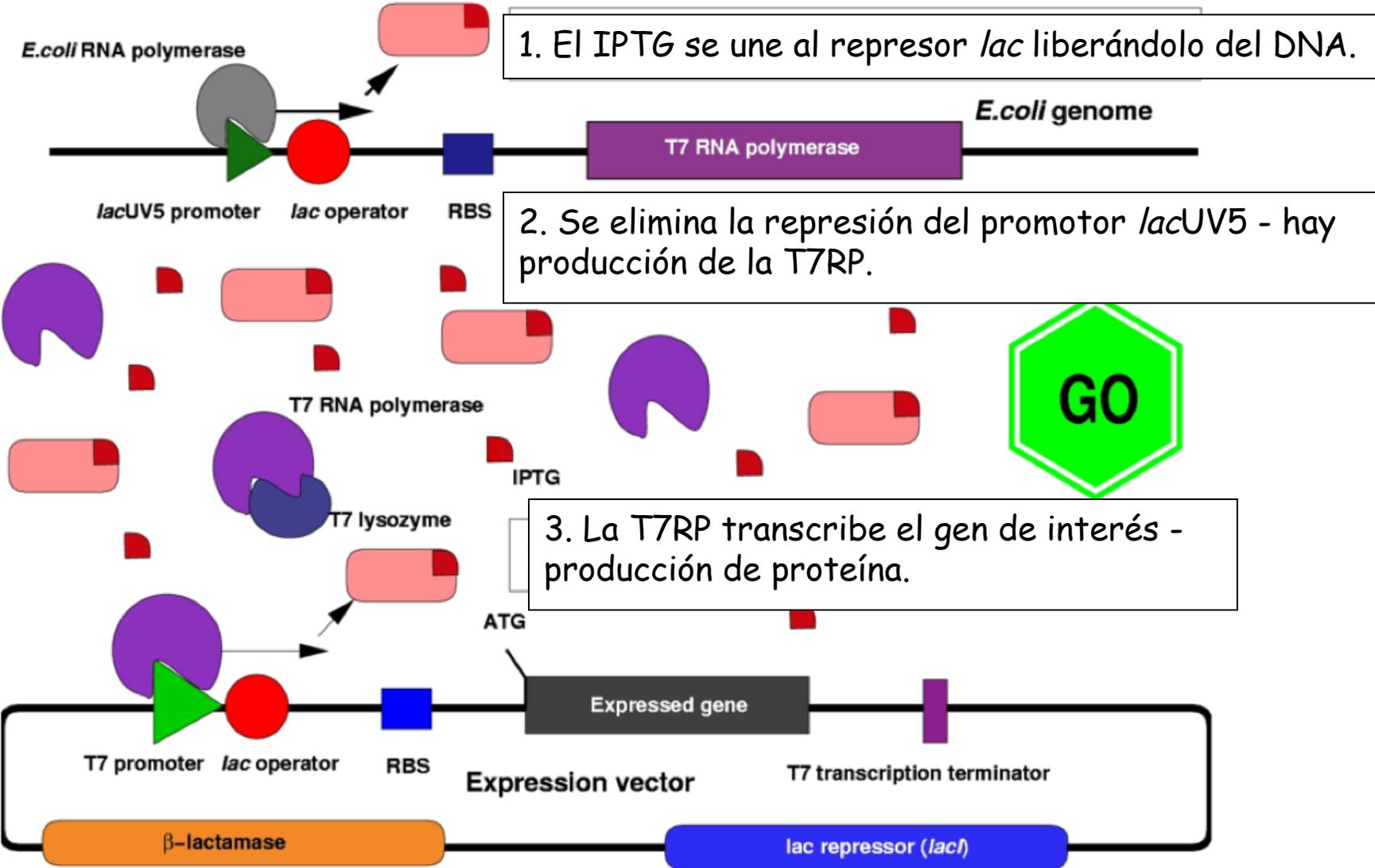
- El promotor T7 es el promotor del *gene 1* del bacteriófago T7.
- Es reconocido únicamente por la RNA polimerasa del bacteriófago T7 (T7RP).
 - Rápida y altamente procesiva - buena para transcritos grandes.
- T7RP se inhibe por la lisozima del bacteriófago T7 (plásmido pLys)



El sistema T7 reprimido



El sistema T7 inducido



E. coli BL21 (DE3)

Cepa de trabajo para expresar en el sistema T7

Contiene una copia de la RNA polimerasa del bacteriófago T7 bajo el control *lac*

Rápido crecimiento

No patogénica

Muchas variantes según la finalidad de uso

Deficiente en proteasas *ompT*⁻

**¿Cómo sobrepasar los problemas de expresión en *E.coli*
antes mencionados?**

Marco de lectura erróneo

Verificar que la secuencia codificante se encuentra en el marco de lectura correcta - SECUENCIAR

		<i>NdeI</i>									
		┌───────────┐									
		CATATGCACCACCACCACCACAGCGGCGCTTTT									
Marco 1	✓	M	H	H	H	H	H	S	G		
Marco 2	✗	C	T	T	T	T	T	T	A	A	
Marco 3	✗	Y	A	P	P	P	P	P	P	Q	R

Considerar cuando se intercambia la secuencia codificante de un vector a otro

e.g., vectores pET A, B, C

Uso de codones inadecuado

Diferentes organismos usan el mismo codon con diferentes frecuencias - verificar tabla de uso de codones

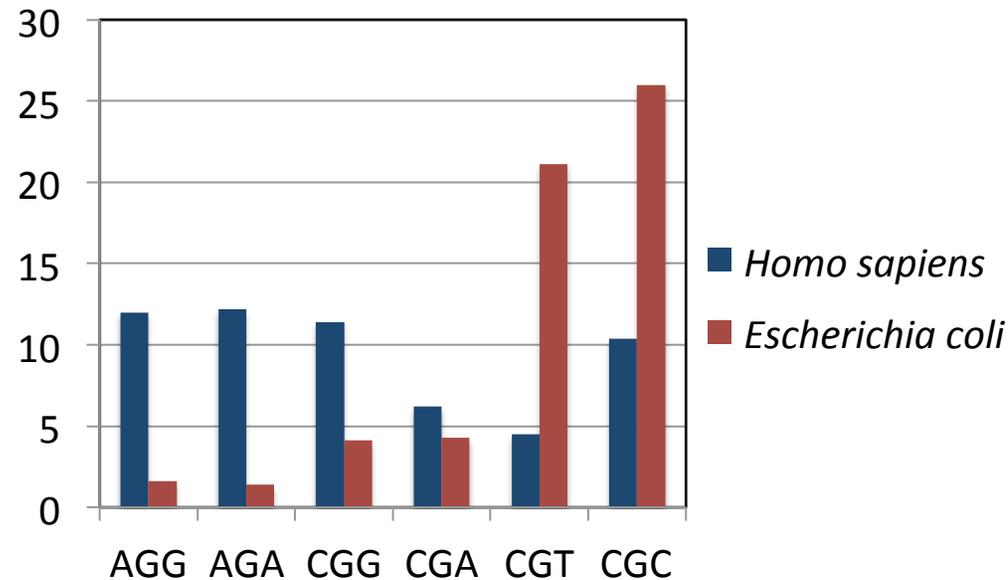
Alto número de codones raros puede reducir considerablemente los niveles de expresión

La bacteria se queda sin el correspondiente tRNA - la traducción se detiene y la proteína no se pliega correctamente

Codones raros en *E. coli*

Arg	AGA, AGG
Ile	CTA, ATA
Pro	CCC
Gly	GGA

Uso de codones de Arginina Humano - vs - *E. coli*



Tablas de uso de codones <http://www.kazusa.or.jp/codon/>

Análisis de codones usados en una secuencia

Graphical Codon Usage Analyser <http://gcu.schoedl.de/>

Optimizer <http://genomes.urv.es/OPTIMIZER/>

¿Cómo sobrepasar el problema de un uso de codones inadecuado?

- Mutar los codones correspondientes mediante PCR - pocos codones
- Sintetizar el gen completo con los codones optimizados - \$
\$\$ 0.4/bp
- Usar cepas especiales con copias extras de los tRNAs inusuales

E. coli Rosetta(DE3)

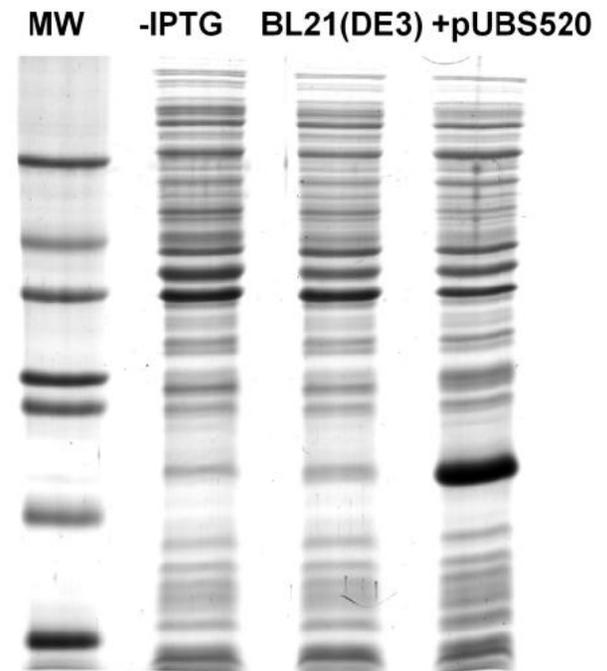
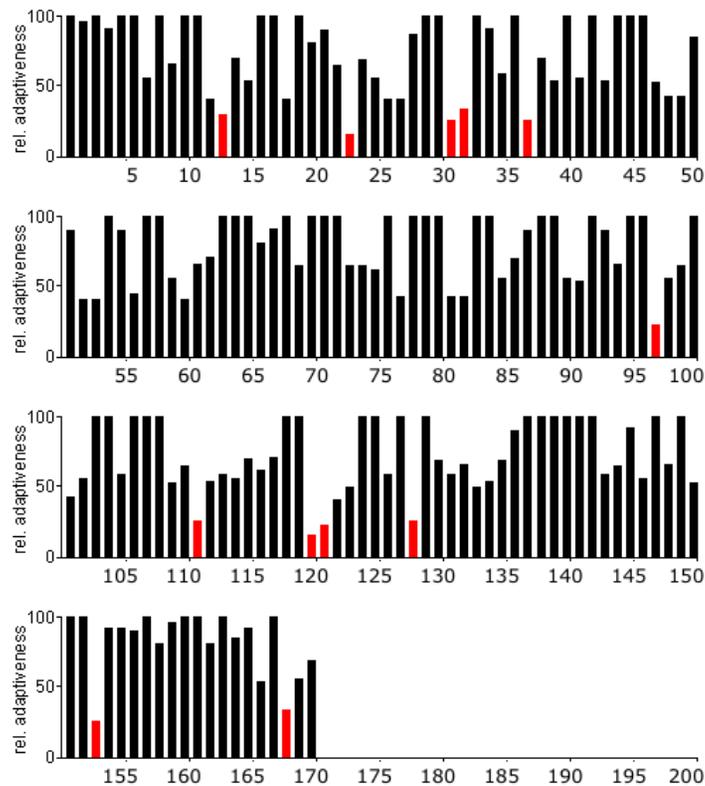
E. coli BL21(DE3) pRIL

Expresión del extremo N-terminal de la Bruton tirosin- cinasa

aa 1-170

Codones *AGA + AGG* = 7 (4%)

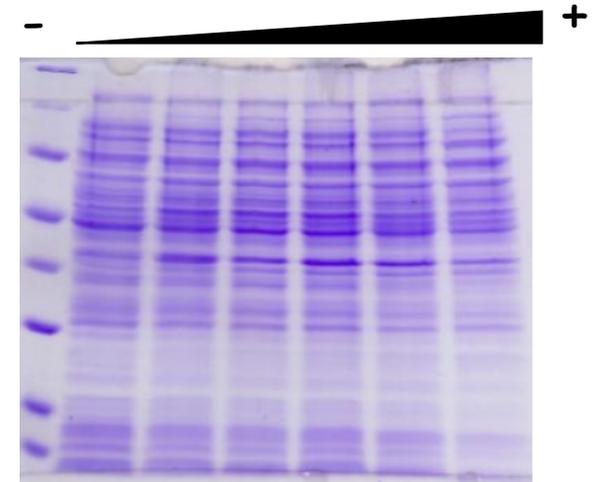
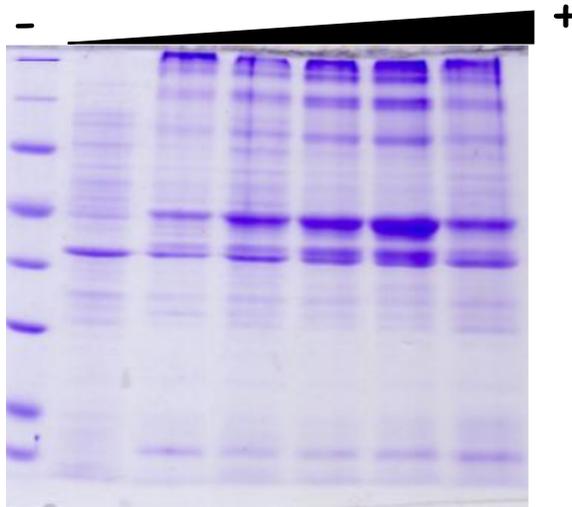
2 tadem



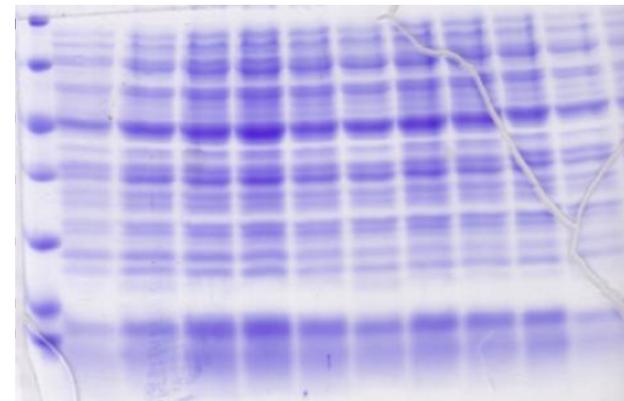
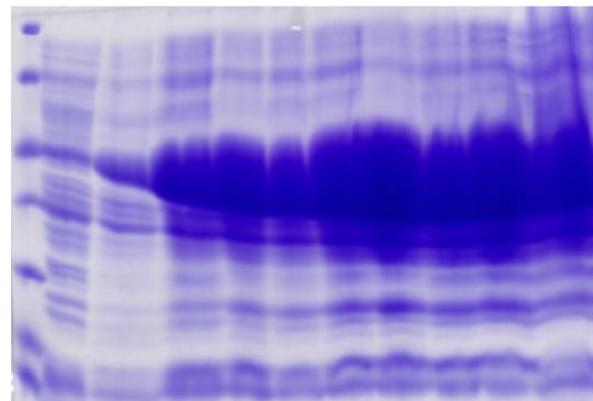
Expresión del dominio catalítico de la silicateína- α

Fracción insoluble

Fracción soluble



E. coli C41

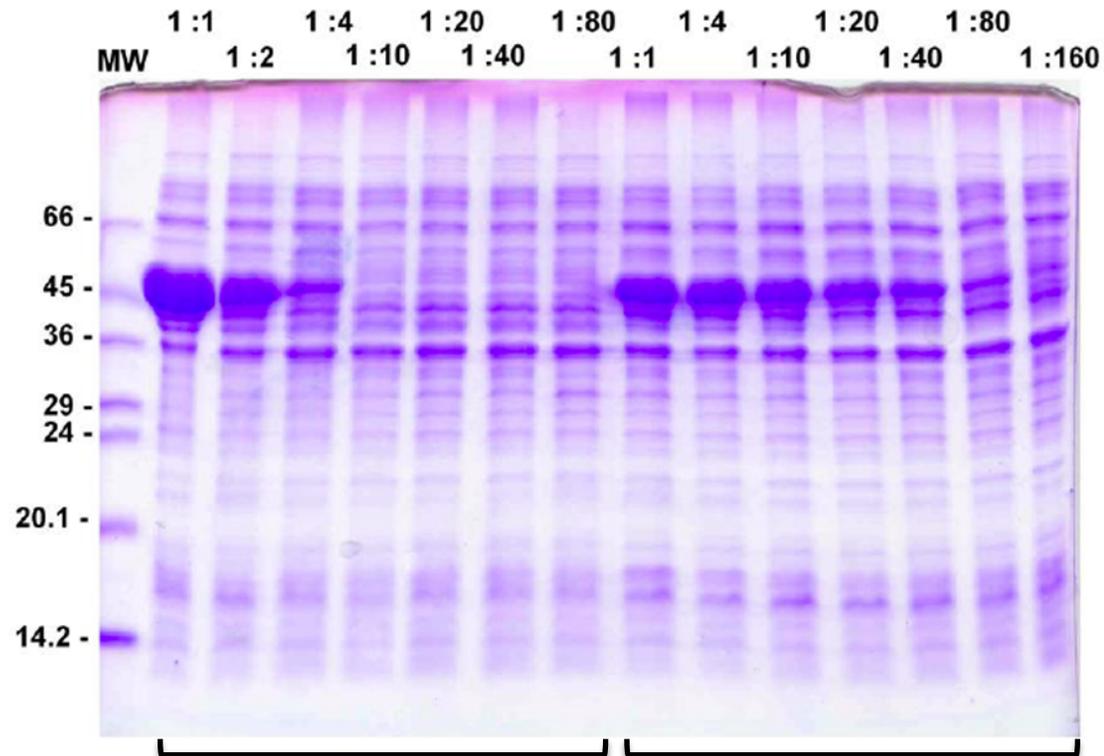


E. coli
Rosseta-gammi

Cuerpos de inclusión

- Son agregados densos de proteína parcialmente plegada o proteína que ha adquirido una conformación no-nativa.
- Se forman durante la sobre-expresión de ciertas proteínas recombinantes en E. coli.
- Localización en citoplasma y periplasma
- Resistentes a la acción de proteasas, detergentes
- Unen colorantes como las fibras amiloides
- Tamaño 1 μm diámetro

Reducción de la concentración de inductor



MBP en un vector T7
(inductor IPTG)

MBP en un vector pBAD
(inductor L-arabinosa)

[IPTG] = 500 – 50 μ M

[L-arabinosa] = 0.2% - 0.002%

Expresión usando diferentes tipos de cepa de *E. coli*

Cepa	Propiedades
Rosetta (DE3)	Copias extra para los genes tRNA Arg, Pro, Ile y Leu en un plásmido adicional.
BL21 (DE3) Star	Mutación en la RNAsaE para mayor estabilidad de los RNAm
Tuner (DE3)	Mutación en la permeasa del operón lactosa (<i>lacY</i>)
BL21-AI	RNA polimerasa T7 inducible por arabinosa
BL21-SI	RNA polimerasa T7 inducible por sal
C41 (DE3)	Tolerante a la sobre-expresión de proteínas tóxicas y de membrana
B834 (DE3)	Derivada de BL21 auxótrofa para metionina
Origami (DE3)	Deleciones en los genes <i>trxB/gor</i> para facilitar la formación de puentes de disulfuro

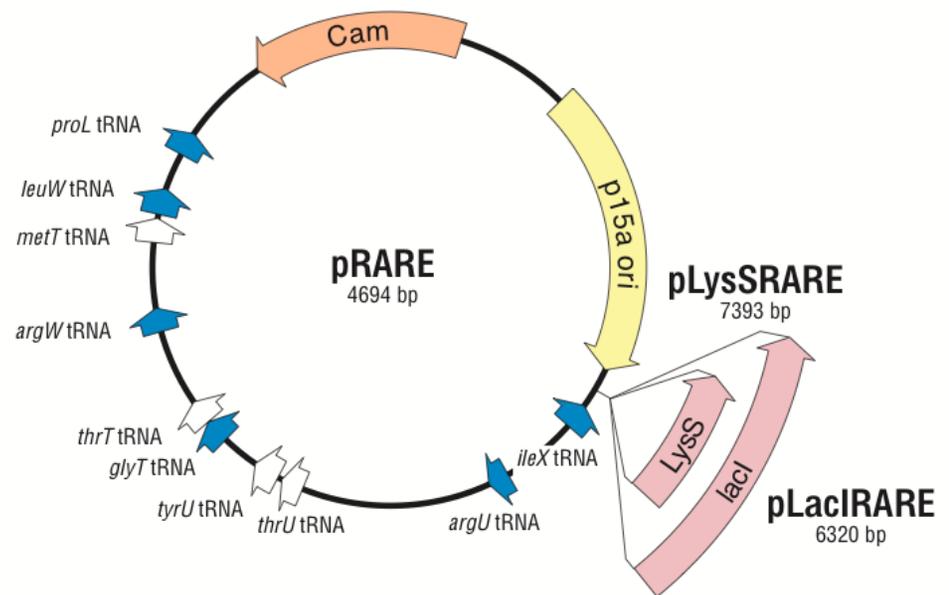
E. coli Rosetta-gami = Rosetta + Origami

Combinación de diferentes cepas de *E. coli* con plásmidos especiales

Rosetta(DE3) pLysS
Rosetta(DE3) placI

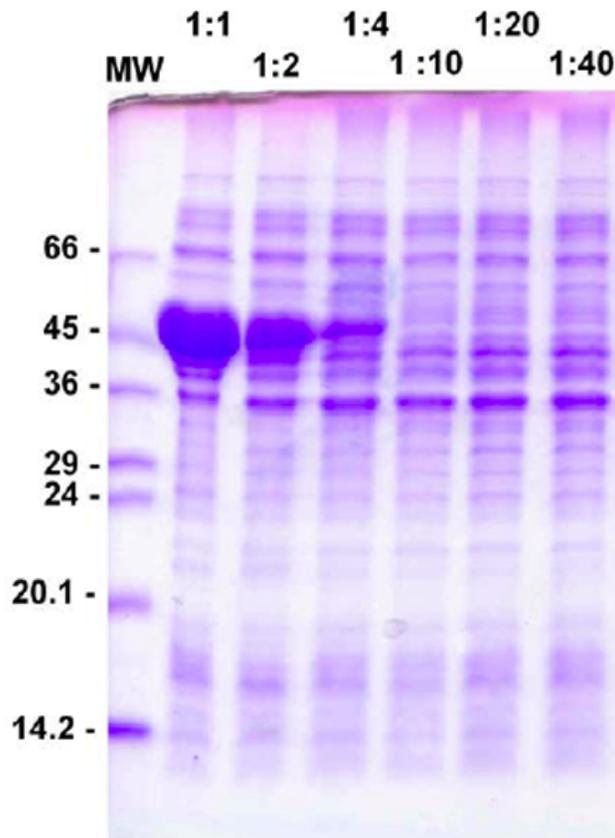
Rosetta-gami(DE3) pLysS
Rosetta-gami(DE3) placI

BL21(DE3) pRARE
BL21(DE3) pLysSRARE

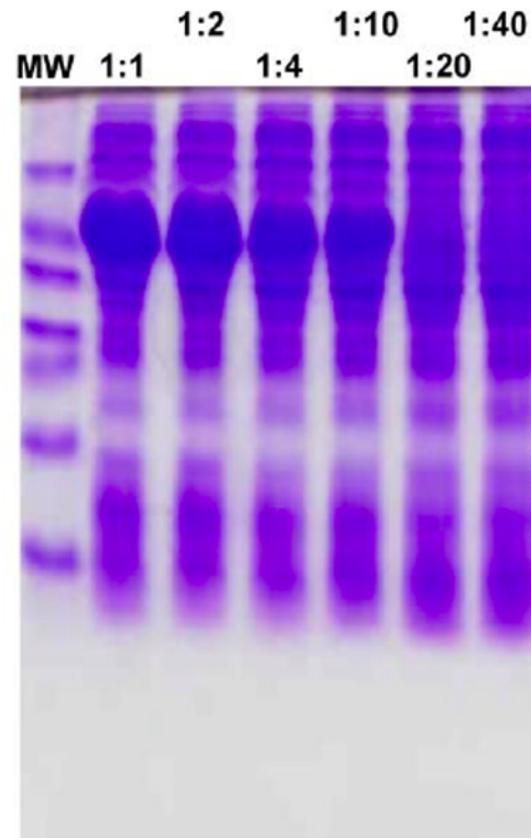


pLysS = plásmido que codifica para
lisozima T7
inhibidor de la RNA pol T7

Expresión usando diferentes tipos de cepa de *E. coli*



MBP expresada en
E. coli Tuner(DE3)



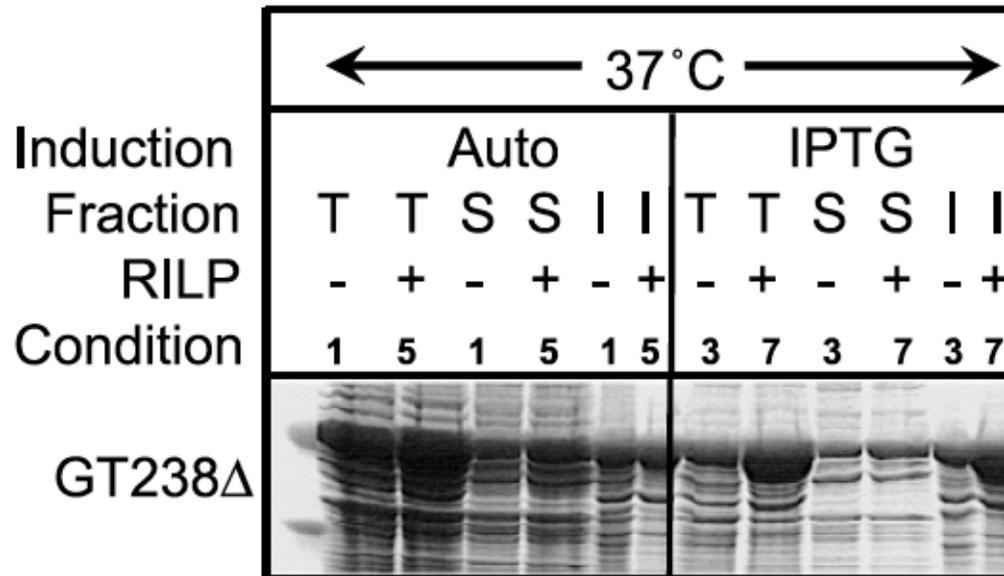
MBP expresada en
E. coli BL21(DE3)

Expresión en diferentes medios de cultivo - Autoinducción en el sistema T7

Caldo Terrífico modificado	
Triptona	1.2%
Extracto de levadura	2.4%
K ₂ HPO ₄	72 mM
KH ₂ PO ₄	17 mM
MgSO ₄	2 mM
Acido aspártico	0.38%
Glicerol	0.8%
Lactosa	0.5%
Glucosa	0.015%

- No adecuado para cepas lacY⁻ y lacZ⁻
- Altas densidades de cultivo OD₆₀₀ ≥ 15
- Una vez consumida la glucosa (represor del promotor T7/lac) la lactosa es utilizada como fuente de carbono e inductor

Expresión de la protease TEV

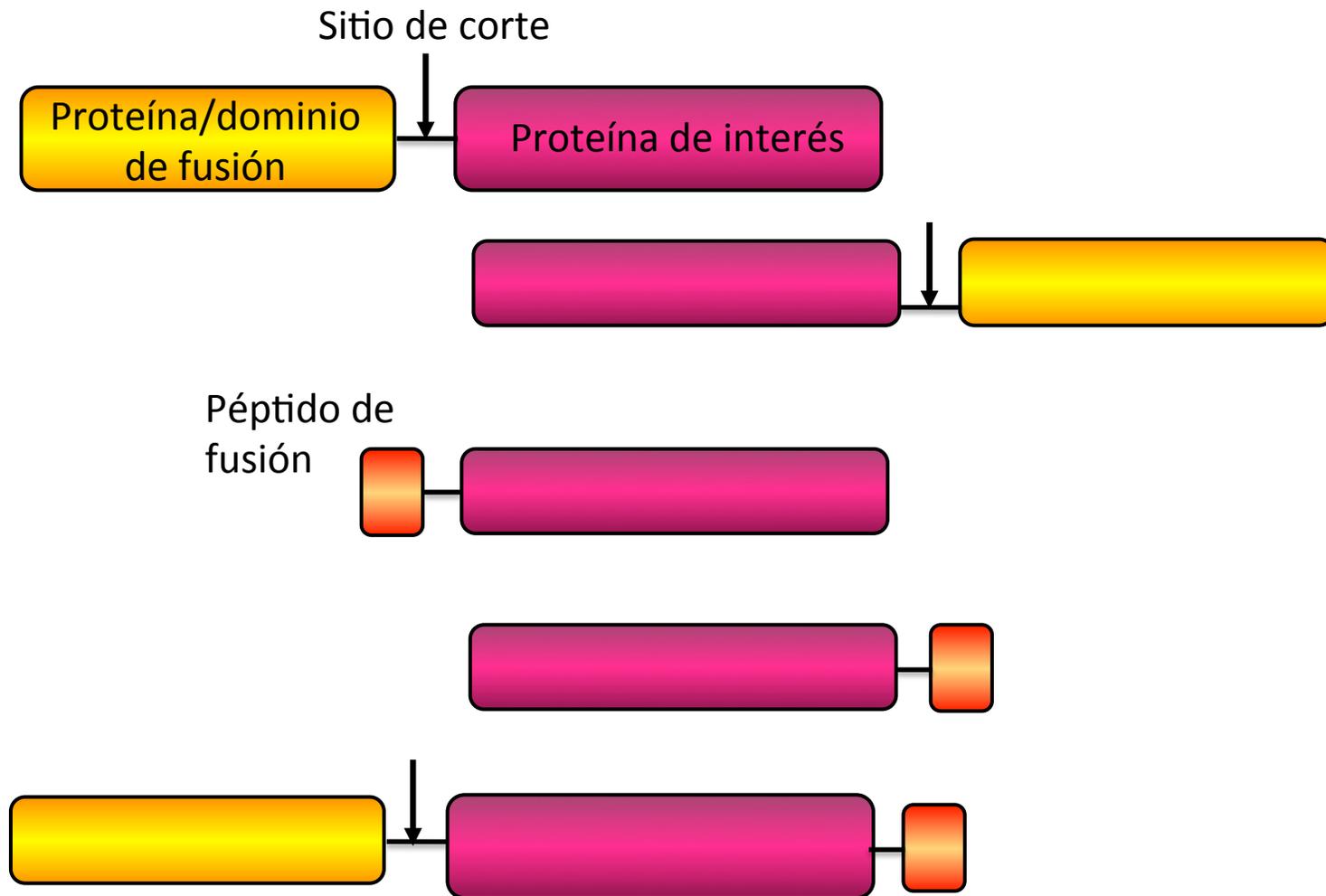


Uso de proteínas y etiquetas fusionadas a la proteína de interés

¿Por qué son buenas?

- Incremento en los niveles de expresión
- Aumentar la solubilidad de mi proteína de interés
- Permiten una detección fácil y rápida de mi proteína e.g. WB
- Ayudan en la purificación usando métodos establecidos
 - alta pureza en un solo paso
 - potencialmente distinguen moléculas plegadas de las incorrectas

Proteínas de fusión



Proteínas de fusión comúnmente utilizadas

Glutación S-transferasa (GST)	<i>S. japonicum</i>
Proteína de unión a maltosa (MBP)	<i>E. coli</i>
Tioredoxina (Trx)	<i>E. coli</i>
Dominio dihidrolipoil acetiltransferasa	<i>B. stearothermophilus</i>

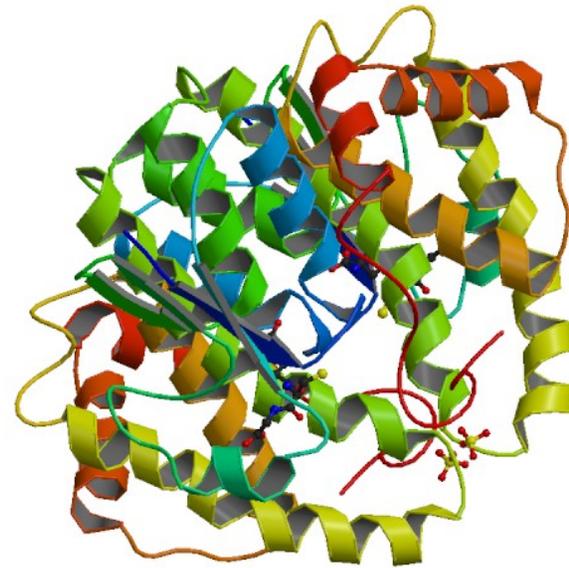
Etiquetas de fusión comúnmente utilizadas

His-tag	6 x Histidinas
---------	----------------

V5

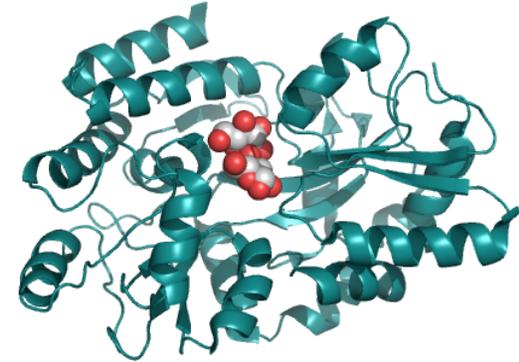
Glutación S-transferasa (GST)

- Dimérica, 30 kDa c/u
- Buena para purificar
- Problemas de agregación - mantener reducida mediante el uso de DTT
- Vectores pGEX y pET

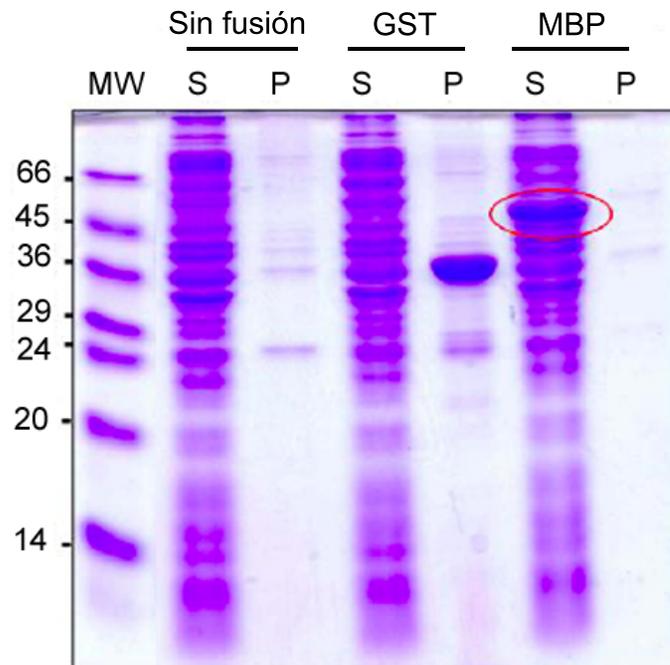


Proteína de unión a maltosa (MBP)

- Proteína monomérica, 40 kDa
- Buena para fomentar la solubilidad
- Vectores pMAL (NEB)
- Buen acarreador para la cristalización
- Versiones citoplásmicas y secretadas a periplasma
- Pobre como medio de purificación - resina de amilosa baja capacidad



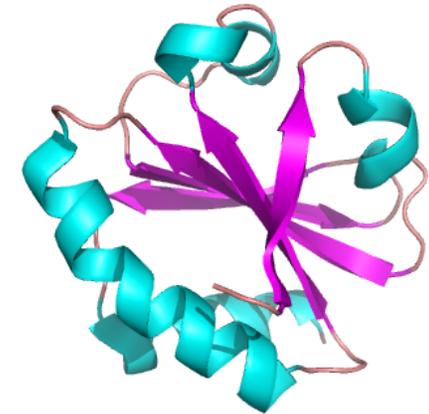
PDB: 1MDQ



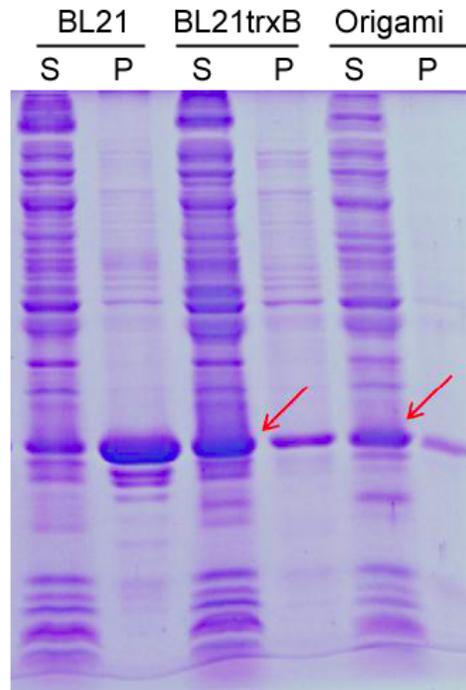
Expresión del dominio CHRD
de la proteína Chordin

Tioredoxina (Trx)

- Disulfuro reductasa pequeña, 14 kDa
- Chaperona
- Altamente soluble, 50 % del total de proteína
- Buen acarreador para la cristalización
- No existe una buena resina de afinidad para su purificación - combinada con His-Tag



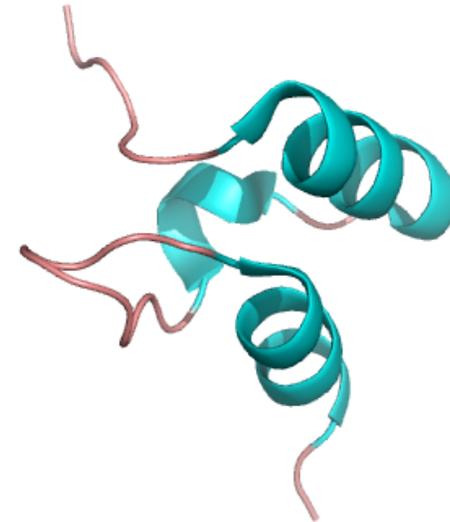
PDB: 2TRX



Expresión de la proteína
Trx-ALK4

Dominio lipoil (dihidrolipoil acetiltransferasa)

- Extremo N-terminal de la proteína E2p de *B. stearothermophilus*.
- Extremadamente resistente a proteasas y desnaturalización química.
- Altamente soluble, previene la formación de cuerpos de inclusión.
- Pequeña (12 kDa)
- Difícil de cristalizar.

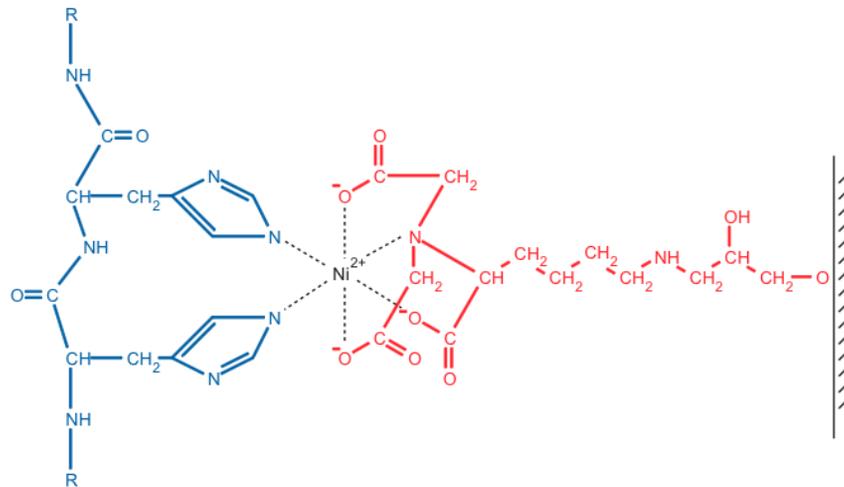


His-Tag como una etiqueta de fusión para su uso en cromatografía de afinidad a metales

- Desarrollada por Hoffman-La Roche en los años 80s
- Basado en la afinidad de histidinas contiguas por metales - Ni^{2+} , Co^{2+} , Cu^{2+}
- Pequeña - no afecta la función de la proteína
- Funciona independiente de su posición

N-terminal, C-terminal, espacial

- Discrimina entre diferentes estados de plegamiento

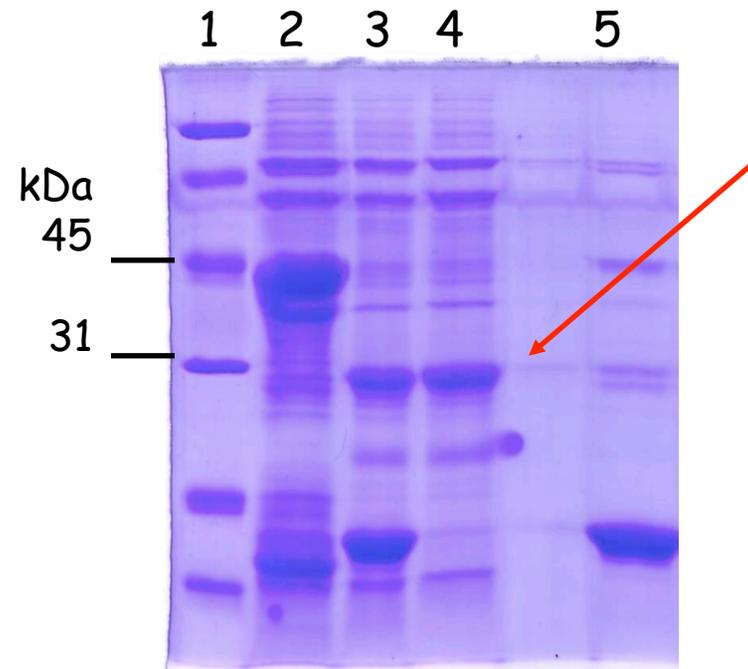


Ultima reflexión

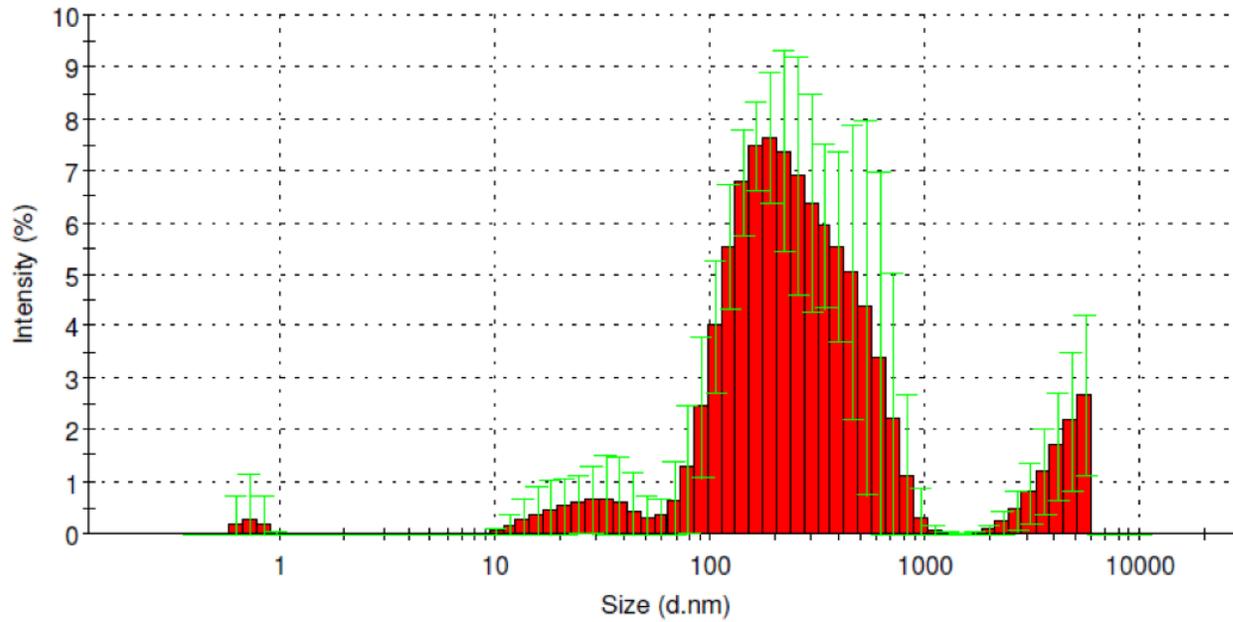
La expresión soluble de una proteína no es garantía que sea funcional
intrínsecamente desordenada, agregados, parcialmente plegada

HisLipoilTEV-silicateína- α

1. Marcador
2. Elución 1a. cromatografía de afinidad Ni^{2+}
3. Digestión con TEV
4. Fracción no retenida 2a. cromatografía afinidad Ni^{2+}
5. Fracción eluída 2a. cromatografía de afinidad Ni^{2+}



La muestra con el dominio catalítico de la silicateína- α se encuentra formando agregados solubles



	M_w (kDa)	R_H (nm)
Chimotripsinógeno	25	2.4
Insulina	34.2	2.7
Ovoalbúmina	43	3.1
Inmunoglobulina G	160	7.1



Conclusiones

- Expresar una proteína de forma heteróloga de forma exitosa muchos parámetros deben de ser ajustados caso-específico.

factores externos { Microorganismo destino y sus diferentes variedades
Temperatura y tiempo de inducción
Cantidad de inductor

Ingeniería de proteína { Expresión de dominios individuales
Optimización de los extremos N- y C-
Uso de codones
Uso de proteínas de fusión

- El uso de proteínas y etiquetas de fusión generalmente aumenta los niveles de expresión, favorece la solubilidad y ayuda en la purificación.