

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 381 433**

51 Int. Cl.:
C12N 15/63 (2006.01)
C12N 15/67 (2006.01)
C12N 9/36 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **09716044 .4**
96 Fecha de presentación: **27.02.2009**
97 Número de publicación de la solicitud: **2268818**
97 Fecha de publicación de la solicitud: **05.01.2011**

54 Título: **Sistema de expresión para proteínas**

30 Prioridad:
29.02.2008 SE 0800483

45 Fecha de publicación de la mención BOPI:
28.05.2012

45 Fecha de la publicación del folleto de la patente:
28.05.2012

73 Titular/es:
Xbrane Bioscience Aktiebolag
Stureplan 15
111 45 Stockholm , SE

72 Inventor/es:
WILLEM DE GIER, Jan y
WAGNER, Samuel

74 Agente/Representante:
Carpintero López, Mario

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

ES 2 381 433 T3

DESCRIPCIÓN

Sistema de expresión para proteínas

Campo de la invención

La invención se refiere a un sistema de expresión para polipéptidos. En concreto, la invención se refiere a plásmidos y células huésped para expresión de polipéptidos, que pueden ajustar la intensidad de la expresión mejorando de este modo los rendimientos de la expresión. Más específicamente, la presente invención se refiere a vectores nuevos que se expresan en un huésped, que comprende un promotor ajustable unido operablemente a una secuencia de ácido nucleico que codifica la lisozima T7. Además, la invención se refiere al uso de estos vectores para la expresión de ácidos nucleicos que codifican la lisozima de T7. La invención también se refiere al uso de la lisozima de Y7 para ajustar la actividad de la ARN polimerasa de T7 de un modo para dar una expresión óptima de ácidos nucleicos que codifican, por ejemplo, polipéptidos, en concreto proteínas con una compleja biogénesis postraducciona, por ejemplo proteínas de membrana, proteínas secretoras, proteínas modificadas postraduccionalmente (p. ej., glicosiladas) y proteínas que requieren la ayuda de chaperonas para su plegamiento.

Antecedentes de la invención

Se han descrito muchos sistemas para la expresión de ácidos nucleicos que codifican, por ejemplo, polipéptidos tales como proteínas recombinantes en sistemas procariotas. Los sistemas de expresión basados en la ARN polimerasa de T7 (T7RNAP, SEC ID N° 17) son los sistemas más usados para producir proteínas recombinantes en *Escherichia coli* (Studier, y col. (1990) Methods Enzymol 185:60-89). Cabe destacar que además de en *E. coli*, la expresión basada en T7RNAP se ha usado en un huésped de diferentes organismos procariotas y eucariotas (p. ej., Kang y col. (2007) Protein Expr Purif. 55(2):325-33; Dower y Robasch (2002) RNA 8:686-697; Nguyen y col. (2004) Plant Biotechnol. J. 2: 301-310; Polkinghorne y Roy (1995) Nucleic Acids Res. 23(1): 188-191; Wang et al. (2007) Analytical Biochemistry, en prensa (doi:10.1016/j.ab.2007.11.037)). La T7RNAP reconoce un promotor muy específico, es decir un promotor de T7 (Chamberlin y col., (1970) Nature 228, 227-231; Oakley y Coleman (1977) PNAS 74, 4266-4270; Rosa (1979) Cell 16, 815-825 ;Panayotatos y Wells (1979) Nucl. Acids Res. 13, 2227-2240; Dunn y Studier (1983) J. Mol. Biol. 166, 477-535/175, 111-112), y transcribe ADN con una tasa de transcripción ocho veces mayor que la ARN polimerasa de *E. coli* (lost, y col. (1992) J Bacteriol 174:619-22). Esto es una gran ventaja para la producción de alto rendimiento de la proteína recombinante en *E. coli*. El inconveniente de la buena eficiencia es, por ejemplo, la gran carga metabólica y la tensión de plegamiento impuesta sobre la célula huésped. Aunque la producción de proteínas en el citoplasma de *E. coli* es relativamente sencilla, la sobreexpresión de proteínas de membrana en *E. coli* sigue suponiendo un reto (Wagner, y col. (2006) Trends Biotechnol 24:364-71). Aunque las proteínas de membrana a menudo pueden expresarse fácilmente en los cuerpos de inclusión, su replegamiento en proteínas funcionales no suele tener éxito. La sobreexpresión de las proteínas de membrana por la acumulación en el sistema de la membrana citoplásmica evita este problema de replegamiento, pero suele ser tóxico, de modo que se reducen mucho los rendimientos. Esto se debe principalmente a los complejos requisitos de la biogénesis de las proteínas de membrana, en las que, tras el inicio de la traducción en el ribosoma, los complejos de las cadenas nacientes de proteína de membrana y ribosoma están dirigidos a la membrana plasmática (Luirink, y col. (2005) Annu Rev Microbiol 59:329-55). Los dominios transmembranales (DTM) de las proteínas de membrana quedan atrapados en el translocón Sec y, después, se dividen en la bicapa lipídica. La sobreexpresión de la proteína de membrana satura fácilmente una o varias etapas de la vía de la biogénesis que conduce a la agregación indeseada y la degradación de proteínas recombinantes en el citoplasma. Además, el bloqueo de la vía secretora tiene como resultado una toxicidad intensa para la célula huésped (Wagner, y col. (2007) Mol Cell Proteomics 6:1527-50). Una menor viabilidad de la célula huésped y el plegamiento incorrecto de la proteína recombinante sobreexpresada tiene como resultado bajos rendimientos del producto deseado. En este caso, el principal problema es que la expresión de los polipéptidos por la T7RNAP es demasiado fuerte. La mayoría de los sistemas de expresión basados en T7RNAP han dependido exclusivamente de una actividad fija de la T7RNAP que tiene como resultado una intensidad fija de expresión de proteínas recombinantes. Las cepas más usadas para este fin son BL21 (DE3) y sus derivados BL21 (DE3)pLysS, BL21 (DE3)pLysE, C41 (DE3) y C43(DE3) (Studier (1991) J Mol Biol 219:37-44) (Miroux (1996) J Mol Biol 260: 289-98). Los huéspedes portadores de los plásmidos pLysS y pLysE expresan la lisozima de T7, , pLysS a un nivel bajo establecido y , pLysE a un nivel alto establecido. Todos estos sistemas fijos no dejan oportunidad a ajustar la intensidad de la expresión a los requisitos individuales de diferentes proteínas diana. Dado que es impredecible qué intensidad de expresión es óptima para obtener los mejores rendimientos de sobreexpresión de una proteína concreta, la expresión se tiene que cribar en diferentes cepas con intensidades de expresión fijas basadas en T7RNAP. Sería ideal que se pudieran someter a detección selectiva diferentes intensidades de expresión en solo un sistema de expresión basado en T7RNAP. Este sistema de expresión "todo en uno" facilitaría enormemente la detección selectiva de la sobreexpresión. Por tanto, existe la necesidad de proporcionar sistemas de expresión procariotas y eucariotas alternativos con la capacidad para ajustar de forma continua la intensidad de la expresión de secuencias de ácido nucleico, en particular de las que codifican proteínas de membrana y otros polipéptidos con complejos requisitos biogénicas postraduccionales; por ejemplo proteínas secretoras, proteínas que se modifican postraduccionalmente (p. ej., glicosiladas (Wacker y col., (2002) Science 298, páginas 1790 - 1793), y proteínas que requieren la ayuda de chaperonas para su plegamiento. Encontrar la cepa óptima basada en T7RNAP para la sobreexpresión de una proteína es cuestión de "ensayo y error" que requiere tiempo, de modo que se usan, por ejemplo, BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pLysS

(Studier (1991) J Mol Biol 219: 37-44), C41(DE3) y C43(DE3) (Miroux (1996) J Mol Biol 260, 289-98). En estas cepas, la actividad de T7RNAP está fijada y se debe realizar detección selectiva de la cepa más adecuada según la sobreexpresión de la proteína deseada. La cepa que será la mejor es algo que no se puede predecir.

Técnica anterior

- 5 Studier ((1991) Journal of Molecular Biology 219:37-44) divulga varios sistemas basados en la ARN polimerasa de T7 para la expresión de proteínas. En los sistemas de expresión divulgados se usan niveles en equilibrio de la lisozima T7 con el fin de inhibir la actividad de fondo de la ARN polimerasa de T7 y, por tanto, prevenir la expresión con fugas y/o reducir los niveles de expresión de las proteínas diana tóxicas. Reduciendo la expresión con fugas de la proteína diana antes de la inducción de la expresión y/o reduciendo los niveles de la expresión inducida de las proteínas diana se mejora la viabilidad de la célula huésped y, por tanto, se incrementan los rendimientos proteicos.

Se han desarrollado varios sistemas de expresión en los que la ARN polimerasa de T7 está bajo el control directo de un promotor inducible:

- 15 Hartnett y col. ((2006) Promega notes 94:27-30) y el documento EP 1 847 611 A2 divulgan sistemas para expresión de proteínas, en los que la ARN polimerasa de T7 está bajo el control directo de un promotor rhaBAD inducible por ramnosa.

McKinney y col. ((2002) Journal of Bacteriology 184: 6056-6059) divulgan una cepa de Salmonella para expresión génica regulada, en la que la ARN polimerasa de T7 está bajo el control de un elemento de control araC-P_{BAD} que es inducible por arabinosa.

- 20 Sistemas similares se divulgan en el documento WO 2008/017073 A, Wycuff y col. ((2000) Analytical Biochemistry 277:67-73 y Chao y col. ((2002) Biotechnology progress 18:8756-7938).

- 25 Otro sistema basado en la ARN polimerasa de T7 inducible o ajustable, en el que la ARN polimerasa de T7 está directamente regulada por, por ejemplo, calor o luz, incluye sistemas divulgados en Chao y col. (2002) Applied Microbiology and Biotechnology 58:446-453), Wang y col. (2004) Biotechnology Process 20:1352-1358), Saida y col. ((2006) Current Protein and Peptide Science 7:1389-2037), Asanuma y col. ((2001) Nucleic Acids Research Suplemento N° 2, páginas 75-76) y Liang y col. ((2004) Nucleic Acids Symposium Series N° 51, páginas 349-350).

Sumario de la invención

- Los inventores han demostrado que un ajuste de la actividad de T7RNAP facilita que la célula huésped acople los procesos de transcripción, traducción y postraducción, y da como resultado rendimientos más altos de proteína expresada funcionalmente y también reduce la agregación indeseada de proteína mal plegada y mal dirigida. En consecuencia, se proporciona un nuevo sistema para la expresión ajustable de ácidos nucleicos que codifican, por ejemplo, polipéptidos tales como proteínas recombinantes en sistemas procariotas. El sistema se basa en la capacidad de la lisozima T7 (T7Lys) para inhibir la actividad de T7RNAP. La expresión de T7Lys se puede ajustar de forma continua a medida que su expresión está bajo el control de un promotor cuya actividad se puede titular, en concreto T7Lys está unido operablemente al promotor rhaBAD inducible por L-ramnosa (Giacalone, y col. (2006) Biotechniques 40:355-64). El núcleo de la invención es un nuevo vector que se puede expresar en un huésped que comprende un promotor ajustable unido operablemente a la secuencia de ácido nucleico que codifica T7Lys. Dicho vector no está en el mismo grupo de incompatibilidad que el sistema pET u otros vectores de expresión basados en T7, ya que comprende orígenes de replicación diferentes de los orígenes colE1 de los vectores derivados de pBR1322, en concreto los siguientes orígenes de replicación. p15A de pACYC, repA de pSC101, y cloDF13 de pCDF. Esto significa que el vector nuevo es capaz de coexistir en la misma célula que los vectores del sistema Pet de otros vectores de expresión basados en T7. Esto permite el uso de una cepa para optimizar las condiciones de expresión para una proteína específica y elimina la necesidad de detectar de forma selectiva la cepa más óptima.

- 45 BL21 de *Escherichia coli* (DE3) transformada con un vector derivado de pACYC que comprende el promotor rhaBAD unido operablemente a T7Lys se denomina Lemo21(DE3). En esta cepa, los niveles de T7Lys se regulan mediante la adición de L-ramnosa al medio de cultivo. T7Lys inhibe la T7RNAP hasta un grado en función de la concentración de L-ramnosa. La T7RNAP transcribe una proteína recombinante diana unida operablemente a un promotor de Y7 de un vector del sistema pET. Los niveles de ARNm de dicha proteína recombinante dependen de la actividad de T7RNAP y, por tanto, de la concentración de L-ramnosa en el medio de cultivo. Los sistemas basados en T7Lys unidos operablemente a otros promotores ajustables funcionan en consecuencia. Esto incluye los promotores de lactosa [lac] (Yanisch-Perron y col. 1985, Gene 33, 103-109), y de triptófano [trp] (Goeddel y col., 1980, Nature (London) 287, 411-416), y los promotores híbridos derivados de estos dos [tac y trc] (Brosius, 1984, Gene 27 : 161 - 172 ; Arnann y Brosius, 1985, Gene 40, 183- 190), el promotor araB inducible por arabinosa (documento WO 86 04356) y el promotor rhaSB de ramnosa inducible por L-ramnosa dos 03068956). Los niveles de ARNm reducidos permite que *E. coli* se enfrente con la tensión por sobreexpresión, que continúe creciendo y que no se vea afectada por la sobreexpresión y, por último, que de más proteína producida. El sistema de acuerdo con la invención permite la inhibición de la ARN polimerasa de T7 por la lisozima T7 de un modo ajustable de forma continua para encontrar las condiciones de expresión óptimas para cada proteína diana. El sistema permite de forma conveniente la detección selectiva de expresión óptima usando solo una cepa en lugar de la detección selectiva de la expresión

óptima en diferentes cepas con intensidades fijas de expresión basada en T7RNAP. La compatibilidad de nuestro sistema con la plataforma de expresión de pET usada y otras plataformas de expresión basadas en T7 ofrecen un inmenso potencial para el uso del producto de los inventores que va desde la producción de proteínas en laboratorio a escala pequeña a aplicaciones en fermentadores a escala industrial.

5 La invención comprende adicionalmente los aspectos siguientes:

1. Una célula huésped capaz de expresar ARN polimerasa de T7 tras inducción, comprendiendo la célula huésped:

10 un primer ácido nucleico que tiene un gen de la lisozima de T7 o un gen variante de la lisozima de T7 que codifica una variante de la lisozima de T7 que tiene la capacidad de inhibir la actividad de la ARN polimerasa de T7, bajo el control de un promotor, y

un segundo ácido nucleico que tiene un promotor de T7 unido operablemente a una secuencia de ácido nucleico que codifica un polipéptido diana,

que se caracteriza porque

15 el gen de la lisozima de T7 o el gen variante de la lisozima de T7 está bajo el control de un promotor ajustable;

el promotor ajustable tiene un efecto sobre el índice de transcripción del gen de la lisozima de T7 que depende de la concentración o intensidad de un inductor del promotor, de forma que el índice de transcripción del gen de la lisozima de T7 aumenta con la concentración o intensidad creciente del inductor y que la expresión de la lisozima de T7 se puede ajustar de forma continua mediante dicho promotor ajustable;

20 y porque

el promotor ajustable se selecciona de un modo tal que la expresión del polipéptido diana se puede ajustar mediante el control de la expresión del gen de la lisozima de T7 o el gen variante de la lisozima de T7 cuando se induce la expresión de la ARN polimerasa de T7.

25 2. La célula huésped de la reivindicación 1, en la que dicho ácido nucleico comprende el promotor ajustable es un primer vector y dicho segundo ácido nucleico que comprende el promotor de T7 es un segundo vector y en el que el primero y el segundo vector son compatibles.

3. La célula huésped de cualquiera de las reivindicaciones 1 o 2, en la que dicho promotor ajustable es ajustable por ramnosa o arabinosa.

30 4. La célula huésped de cualquiera de las reivindicaciones 1 o 2, en la que dicho promotor ajustable es ajustable por luz.

5. La célula huésped de cualquiera de las reivindicaciones 1 o 2, en la que dicho promotor ajustable es ajustable por temperatura.

6. La célula huésped de cualquiera de las reivindicaciones 1 a 5, en la que dicho primer ácido nucleico comprende un marcador de selección.

35 7. La célula huésped de la reivindicación 6, en la que el marcador de selección es un marcador de selección antibiótico.

8. La célula huésped de cualquiera de las reivindicaciones 1 a 7, en la que dicha variante de la lisozima de T7 es LysY.

40 9. La célula huésped de cualquiera de las reivindicaciones 1 a 8, en la que dicho primer ácido nucleico es al menos un 80 % idéntico a un ácido nucleico escogido del grupo que comprende las SEC ID N° 1-5.

10. La célula huésped de la reivindicación 9, en la que dicho primer ácido nucleico es idéntico a un ácido nucleico escogido del grupo que comprende las SEC ID N° 1-5.

45 11. Una célula huésped de las reivindicaciones 1 a 10, seleccionada del grupo de *E. coli*, *Pseudomonas aeruginosa*, *Erwinia carotovora*, *Salmonella choleraesuis*, *Agrobacterium tumefaciens*, *Chromobacterium violaceum*, *Lactococcus lactis*, *Bacillus subtilis*, *Salmonella*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, CHO, NS0, HEK293, HeLa, Sf9, tabaco, arroz y *Leishmania tarentolae*.

12. Un procedimiento para producir un polipéptido diana, que comprende las etapas de:

a) proporcionar una célula huésped de acuerdo con una cualquiera de las reivindicaciones 1 a 11;

b) inducir la expresión del polipéptido diana

c) controlar la expresión del polipéptido diana ajustando dicho promotor ajustable y, por tanto, la expresión de dicha lisozima de T7;

y opcionalmente

d) aislar el polipéptido.

5 13. El procedimiento de la reivindicación 19, en el que la inducción de la expresión del polipéptido diana se realiza induciendo la expresión de la ARN polimerasa de T7 o una variante de la misma.

14. El procedimiento de cualquiera de las reivindicaciones 12 a 13, en el que el ajuste del promotor ajustable, o una variante del mismo, se consigue mediante cualquiera de: inducción ligera, ajuste de la temperatura, adición de un inductor químico, en concreto ramnosa o arabinosa.

10 15. Un ácido nucleico que comprende un gen de la lisozima de T7 o un gen variante de la lisozima de T7 que codifica una variante de la lisozima de T7 que tiene la capacidad de inhibir la actividad de la ARN polimerasa de T7, bajo el control de un promotor, caracterizado porque

15 el gen de la lisozima de T7 o el gen variante de la lisozima de T7 está bajo el control de un promotor rhaBAD ajustable y que el promotor ajustable tiene un efecto sobre el índice de transcripción del gen de la lisozima de T7 que depende de la concentración de un inductor del promotor rhaBAD, de forma que el índice de transcripción del gen de la lisozima de T7 aumenta con la concentración creciente del inductor y que la expresión de la lisozima de T7 se puede ajustar de forma continua mediante dicho promotor rhaBAD ajustable.

20 16. Un ácido nucleico de acuerdo con la reivindicación 15, que es al menos un 80 % idéntico con un ácido nucleico escogido del grupo de las SEC ID Nº 1-5.

17. Un ácido nucleico de acuerdo con la reivindicación 16 escogido del grupo de las SEC ID Nº 1-5.

18. Uso del ácido nucleico de cualquiera de las reivindicaciones 15 a 17 para controlar la expresión de un polipéptido diana mediante la lisozima de T7.

Breve descripción de las figuras

25 La invención se explicará adicionalmente con referencia a las Figuras adjuntas, en las que:

Las Figuras 1A-E muestra los mapas de plásmidos de acuerdo con diferentes realizaciones de la invención, que expresan lisozima de T7 o variantes de la misma a partir de un promotor inducible por L-ramnosa.

La Figura 2 muestra un modelo del control de la actividad de la ARN polimerasa de T7 mediante titulación de la expresión de la lisozima de T7 de acuerdo con una realización de la presente invención.

30 La Figura 3 muestra los niveles de expresión de la lisozima de T7 en células Lemo21(DE3) de la presente invención, cultivadas en presencia de concentraciones diferentes de L-ramnosa.

La Figura 4 muestra los niveles de ARNm de la proteína recombinante YidC-GFP, expresada en células Lemo21(DE3) de la presente invención, cultivadas en presencia de concentraciones diferentes de L-ramnosa.

35 La Figura 5 muestra A, crecimiento y B, expresión proteica de las células Lemo21(DE3) de la presente invención, que expresan la proteína recombinante YidC-GFP y cultivadas en presencia de concentraciones diferentes de L-ramnosa.

La Figura 6 muestra la fracción de las células Lemo21 (DE3) que no expresan YidC-GFP 4 horas y 18 horas, respectivamente, después de la inducción con IPTG, en presencia de concentraciones diferentes de L-ramnosa.

40 La Figura 7 muestra A, crecimiento y B, expresión proteica de las células Lemo21(DE3) de la presente invención, que expresan la proteína recombinante YidC-GFP y cultivadas en presencia de concentraciones diferentes de L-ramnosa.

45 La Figura 8 muestra la agregación de proteínas en el citoplasma de células Lemo21(DE3) de la presente invención, que expresan la proteína recombinante YidC-GFP durante 4 horas en presencia de concentraciones diferentes de L-ramnosa.

La Figura 9 muestra A, crecimiento y B, expresión proteica de la proteína recombinante YidC-GFP en un cultivo semicontinuo, en una comparación entre cepas diferentes.

La Figura 10 muestra la expresión de proteínas de membrana de fusión-GFP en células Lemo21(DE3) de la presente invención en comparación con la expresión en otras cepas de *E.coli*.

La Figura 11 muestra los niveles de expresión de la chaperona soluble que requiere luciferasa-HiS₈ en células Lemo21(DE3) de la presente invención, cultivadas en presencia de concentraciones diferentes de L-ramnosa.

- 5 La Figura 12 muestra A, niveles de expresión de la uracilo ADN glicosilasa de *E. coli*, una proteína citoplásmica soluble, de expresión en células Lemo21(DE3) cultivadas en presencia de concentraciones diferentes de L-ramnosa, así como en comparación con la expresión en otras cepas de *E. coli*.

Descripción detallada de las realizaciones

Como se usa en el presente documento, las definiciones siguientes se proporcionan para facilitar la comprensión de la presente invención.

- 10 Un "vector que se puede expresar en un huésped" o "vector de expresión" es una construcción de ácido polinucleico generado de forma recombinante o sintética, con una serie de elementos de ácido polinucleico especificados que permiten la transcripción de una secuencia de ácido nucleico concreto en una célula huésped.

- Normalmente, este vector incluye una unidad transcripcional que comprende una secuencia de ácido nucleico concreto que se va a transcribir unido operablemente a un promotor. Un vector que se puede expresar en un
15 huésped puede ser, por ejemplo, un plásmido de replicación autónoma o autorreplicación, un cósmico, un fago, un virus o un retrovirus.

- Los términos "huésped", "célula huésped" y "célula huésped recombinante" se usan de forma intercambiable en el presente documento para indicar una célula procariota o eucariota en la cual se han introducido uno o más vectores o secuencias de ácido nucleico de la invención aisladas y purificadas. Debe entenderse que dichos términos no sólo
20 se refieren a la célula sujeto concreta, sino también a la progenie, o potencial progenie, de dicha célula. Dado que se pueden producir ciertas modificaciones en las generaciones posteriores debido a mutaciones o influencias ambientales, dicha progenie puede, de hecho, no ser idéntica a la célula parental, pero todavía está incluida en el alcance del término tal como se usa en la presente memoria descriptiva.

- "Promotor", como se usa en el presente documento, se refiere a una secuencia de ácido nucleico que regula la expresión de una unidad transcripcional. Una "región promotora" es una región reguladora de ARN capaz de unir la ARN polimerasa en una célula e iniciar la transcripción de una secuencia codificadora (dirección 3') posterior. Dentro de la región promotora se encontrará un sitio de inicio de la transcripción (convenientemente definido mediante mapeo con la nucleasa S1), así como los dominios de unión a proteína (secuencias consenso) responsables de la unión de la ARN polimerasa, como la región -35 putativa y la caja Pribnow.

- 30 Un "promotor ajustable" es un promotor cuyo efecto sobre la expresión de una unidad de transcripción depende de la concentración de un inductor (si el inductor es un compuesto químico) o la intensidad de un inductor (si el inductor es la luz o la temperatura). Concentraciones menores del inductor dan como resultado un menor índice de transcripción (índice de transcripción; número de moléculas de ARNm sintetizadas por unidad de tiempo) y concentraciones mayores del inductor dan como resultado un índice de transcripción mayor.

- 35 Los términos "regulable", "controlable", "titulable", "ajustable" y "ajustable" se tienen que usar de forma sinónima.

- "Operón de L-ramnosa" se refiere al operón rhaSR-rhaBAD como se ha descrito para *E. coli* en Holcroft y Egan, 2000, J. Bacteriol. 182 (23), 6774-6782. El operón rhaBAD es un operón catabólico regulado positivamente que transcribe RhaB, RhaA y RhaD de forma divergente de otro operón rha, rhaSR, con aproximadamente 240 pb de ADN separando sus sitios de inicio de la transcripción respectivos. El operón rhaSR codifica los dos activadores
40 específicos de L-ramnosa RhaS y RhaR. RhaR regula la transcripción de rhaSR, mientras que RhaS se une al ADN cadena arriba en -32 a -81 respecto al sitio de inicio de la transcripción de rhaBAD. Además, el operón intergénico rhaSR-rhaBAD contiene los sitios de unión de CRP en las posiciones -92,5 (CRP1) respecto al sitio de inicio de la transcripción de rhaBAD y los sitios de unión de CRP en las posiciones -92,5 (CRP 2), -115,5 (CRP 3) y 116,5 (CRP 4) respecto al sitio de inicio de la transcripción de rhaSR así como un sitio de unión para RhaR que abarca de -32 a -82 respecto al sitio de inicio de la transcripción de rhaSR. Con "región promotora rhaBAD del operón de L-ramnosa" se quiere decir el operón rhaBAD que comprende esencialmente el sitio de inicio de la transcripción de rhaBAD, la
45 supuesta región -35, la caja Pribnow, el sitio de unión de CRP CPR1, el sitio de unión para RhaS respecto al sitio de inicio de la transcripción de rhaBAD, así como los sitios de unión de CRP CRP 2-4 y los sitios de unión para RhaR respecto al sitio de inicio de la transcripción de rhaSR. Con "región promotora rhaBAD" se quiere decir el promotor del operón rhaBAD que comprende esencialmente el sitio de inicio de la transcripción de rhaBAD, la
50 supuesta región -35, la caja Pribnow, el sitio de unión para RhaS y el sitio de unión CRP1 respecto al sitio de inicio de la transcripción de rhaBAD, y el sitio de unión de CRP CRP4 o una parte del mismo respecto al sitio de inicio de la transcripción rhaSR.

- 55 "CRP" significa "proteína reguladora del catabolito". "CRP" por lo general se refiere en la técnica a "proteína receptora de AMP cíclico", que tiene el significado sinónimo. CRP es una proteína reguladora controlada por AMP cíclico (AMPC) que media la activación de los operones catabólicos tales como el operón L-ramnosa.

Un "potenciador" es una secuencia de ácido nucleico que actúa para potenciar la transcripción de una unidad

transcripcional independiente de la identidad de la unidad transcripcional, la posición de la secuencia en relación con la unidad transcripcional o la orientación de la secuencia. Los vectores de la presente invención opcionalmente incluyen potenciadores.

- 5 “Unidad transcripcional” como se usa en el presente documento se refiere a una secuencia de ácido nucleico que normalmente se transcribe en una molécula de ARN sencilla. La unidad transcripcional puede contener un gen (monocistrónico) o dos (dicistrónico) o más genes (policistrónico) que codifican moléculas polipeptídicas funcionalmente relacionadas.

- 10 Una secuencia de ácido nucleico está “unida operablemente” cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor está operablemente unido a una secuencia de codificación si afecta a la transcripción de la secuencia; o una región de inicio de la transcripción, tal como un sitio de unión al ribosoma está operablemente unido a una secuencia de ácido nucleico que codifica, por ejemplo, un polipéptido, si se coloca de modo que se facilite la traducción del polipéptido. La unión se puede lograr mediante uniones en sitios de restricción convenientes. Si tales sitios no existen, los ligadores o adaptadores de oligonucleótidos sintéticos se usan de acuerdo con la práctica convencional

- 15 “Ácido nucleico” o “secuencia de ácido nucleico” o “ácidos nucleicos o secuencias de ácido nucleico aislados y purificados” como se refieren en la presente invención pueden ser ADN, ARN, o híbrido de ADN/ARN. En caso de que el ácido nucleico o la secuencia de ácido nucleico se localicen en un vector, normalmente es ADN. El ADN al que se hace referencia en el presente documento puede ser cualquier secuencia de polidesoxinucleótido, incluyendo, por ejemplo ADN bicatenario, ADN monocatenario, ADN bicatenario en el que una o ambas hebras se componen de dos o más fragmentos, ADN bicatenario en donde una o ambas hebras tienen un esqueleto fosfodiéster no interrumpido, ADN que contiene una o más porciones de cadena sencilla y una o más porciones de cadena doble, ADN bicatenario en donde las hebras de ADN son completamente complementarias, ADN bicatenario en donde las hebras de ADN son solo parcialmente complementarias, ADN circular, ADN covalentemente cerrado, ADN lineal, ADN covalentemente reticulado, ADNc, ADN sintetizado químicamente, ADN semi-sintético, ADN biosintético, ADN aislado naturalmente, ADN digerido de enzima, ADN de cizallamiento, ADN marcado, tal como ADN radiomarcado y ADN fluorocromo-marcado, ADN que contiene una o más especies de ácido nucleico que no ocurren naturalmente. Las secuencias de ADN se pueden sintetizar mediante técnicas estándar de química, por ejemplo, el procedimiento fosfotriéster o vía procedimientos de síntesis automatizados y procedimientos de PCR. La secuencia purificada y aislada de ADN también se puede producir por técnicas enzimáticas.

- 30 Una secuencia “sustancialmente idéntica” es una secuencia de nucleótidos que difiere de una secuencia de referencia únicamente en una o más sustituciones conservadoras, como se trata en el presente documento, o en una o más sustituciones, delecciones o inserciones no conservadoras, localizadas en las posiciones de la secuencia que no destruyen la función biológica de la molécula de ácido nucleico. Dicha secuencia puede ser cualquier valor de 80 % a 99% o, más en general, al menos 80%, 85%, 90% o 95%, o tanto como 96%, 97%, 98% o 99% idénticas cuando están alineadas óptimamente a nivel de aminoácido o nucleótido con la secuencia usada para comparación usando, por ejemplo, el programa Align Program 98 o FASTA. La longitud de las secuencias de comparación puede ser de al menos 5, 10, 15, 20, o 25 nucleótidos, o al menos 30, 40, o 50 nucleótidos. En realizaciones alternativas, la longitud de las secuencias de comparación puede ser de al menos 60, 70, 80 o 90 nucleótidos, o de más de 100, 200 o 500 nucleótidos.

- 40 La identidad de secuencia se puede medir fácilmente usando software de análisis de secuencia disponible para el público (p. ej., Sequenc Analysis Software Package of the Genetic Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis; 53705, o el software BLAST disponible en la National Library of Medicine, o como se describe en el presente documento). Ejemplos de software útil incluyen los programas Pile-up y PrettyBox. Dicho software aparea secuencias similares asignando grados de homología a varias sustituciones, delecciones, sustituciones y otras modificaciones. Secuencias sustancialmente idénticas incluyen secuencias homólogas, tales como secuencias relacionadas con COPI de especies no humanas, como se describe en el presente documento o se conoce en la técnica.

- 50 Como alternativa, o adicionalmente, dos secuencias de ácido nucleico pueden ser “sustancialmente idénticas” si hibridan en condiciones de rigurosidad alta. En algunas realizaciones, las condiciones de rigurosidad alta son, por ejemplo, condiciones que permiten la hibridación comparable con la hibridación que se produce usando una sonda de ADN de al menos 500 nucleótidos de longitud, en un tampón que contiene NaHPO₄, pH 7,2, SDS al 7%, EDTA 1 mM y BSA al 1% (fracción V), a una temperatura de 65°C, o un tampón que contiene formamida al 48%, 4,8x SSC, Tris-Cl 0,2 M, pH 7,6, 1x solución de Denhardt, sulfato de dextrano al 10% y SDS 0,1 %, a una temperatura de 42°C. (Estas son condiciones típicas de rigurosidad alta para hibridaciones Northern o Southern.) Las hibridaciones se pueden llevar a cabo en un periodo de aproximadamente de 20 a 30 minutos o de aproximadamente 2 a 6 horas, o de aproximadamente 10 a 15 horas, o de unas 24 horas o más. La hibridación de alta rigurosidad también depende del éxito de numerosas técnicas realizadas de forma rutinaria por los biólogos moleculares, tales como PCR de alta rigurosidad, secuenciación de ADN, análisis de polimorfismo conformacional de una hebra e hibridación *in situ*. En contraste con las hibridaciones Northern y Southern, estas técnicas normalmente se realizan con sondas relativamente cortas (p. ej., normalmente aproximadamente 16 nucleótidos o más para PCR o secuenciación y aproximadamente 40 nucleótidos o más para hibridación *in situ*). Las condiciones de rigurosidad alta usadas en

estas técnicas son bien conocidas para los expertos en la técnica de biología molecular y se pueden encontrar ejemplos en, por ejemplo, Ausubel y col., Current Protocols in Molecular Biology, John Wiley and Sons.

El ARN al que se hace referencia en el presente documento puede ser ARN monocatenario, ARNc, ARN bicatenario, ARN bicatenario en el que una o ambas hebras se componen de dos o más fragmentos, ARN bicatenario en el que una o ambas hebras tienen un esqueleto fosfodiéster no interrumpido, ARN que contiene una o más porciones de cadena sencilla y una o más porciones de cadena doble, ARN bicatenario en el que las hebras del ARN son completamente complementarias, ARN bicatenario en el que las hebras de ARN son solo parcialmente complementarias, ARN covalentemente reticulado, ARN digerido con enzimas, ARN cizallado, ARNm, ARN sintetizado químicamente, ARN semi-sintético, ARN biosintético, ARN aislado naturalmente, ARN marcado, tal como ARN radiomarcado y ARN marcado con fluorocromo, ARN que contiene una o más especies de ácido nucleico que no ocurren naturalmente.

Con "variantes" o "variantes de una secuencia" o "gen variante" se quiere decir una secuencia de ácido nucleico que varía de la secuencia referencia mediante sustituciones conservadoras de ácido nucleico, por lo cual uno o más ácidos nucleicos se sustituyen por otros con las mismas características. Las variantes abarcan tanto las secuencias degeneradas, secuencias con deleciones e inserciones, siempre y cuando tales secuencias modificadas muestren la misma función (funcionalmente equivalente) como la secuencia referencia.

Como se utiliza en el presente documento, los términos "polipéptido", "péptido", "proteína", "polipeptídico" y "peptídico" se usan de forma intercambiable para designar una serie de residuos de aminoácido conectados al otro por enlaces peptídicos entre los grupos alfa-amino y carboxi de residuos adyacentes.

El término "secuencia aislada y purificada del ácido nucleico" se refiere al estado en el cual la secuencia de ácido nucleico será, de acuerdo con la presente invención. La secuencia de ácido nucleico será libre o sustancialmente libre de material con el cual se asocia naturalmente, tal como otros ácidos nucleicos con los cuales se encuentran en su ambiente natural, o el ambiente en el cual se preparan (por ejemplo, cultivo celular) cuando dicha preparación es por tecnología recombinante realizada in vitro o in vivo.

Los términos "transformación", "transformado" o "introducción de un ácido nucleico en una célula huésped" indican cualquier procedimiento en el que un ácido nucleico extracelular como un vector, con o sin material de acompañamiento, entra a una célula huésped. El término "transformado celular" o "célula transformada" significa la célula o su progenie, en la cual el ácido nucleico extracelular ha sido introducido y de esta manera hospeda el ácido nucleico extracelular. El ácido nucleico puede ser introducido en la célula de tal manera que el ácido nucleico sea replicable, ya sea como un integrante cromosómico o como un elemento cromosómico adicional. La transformación de células huésped apropiadas con, por ejemplo, un vector de expresión se puede lograr por procedimientos bien conocidos tales como microinyección, electroporación, bombardeo de partículas o por procedimientos químicos tales como transformación mediada por fosfato de calcio, descritos en, por ejemplo, Maniatis y col. 1982, Molecular Cloning, A laboratory Manual, Cold Spring Harbor Laboratory o en Ausubel y col. 1994, Current protocols in molecular biology, John Wiley and Sons.

La "secuencia de ácido nucleico heteróloga" o "secuencia heteróloga de ácido nucleico a un huésped" significa una secuencia de ácido nucleico que codifica por ejemplo, un producto de expresión tal como un polipéptido que es extraño al huésped ("expresión heteróloga" o "producto heterólogo") es decir, una secuencia de ácido nucleico, que se origina de un donante diferente del huésped o una secuencia de ácido nucleico sintetizada químicamente que codifica por ejemplo, un producto de expresión tal como un polipéptido que es extraño al huésped. En caso de que el huésped sea una especie procariota particular, la secuencia heteróloga del ácido nucleico preferentemente se origina a partir de un género o familia diferente, más preferido de un orden o clase diferente, en particular a partir de un filo diferente (división) y más particularmente a partir de un dominio diferente (imperio) de organismos.

"Secuencia de ácido nucleico homóloga" o "secuencia de ácido nucleico homóloga a un huésped" significa una secuencia de ácido nucleico que codifica, por ejemplo, un producto de expresión, tal como un polipéptido que procede del huésped ("expresión homóloga" o "producto homólogo"), es decir una secuencia de ácido nucleico que se origina del huésped o una secuencia de ácido nucleico sintetizado químicamente que codifica, por ejemplo un producto de expresión tal como un polipéptido que procede del huésped.

La secuencia heteróloga del ácido nucleico que se origina a partir de un donante diferente a partir del huésped se puede modificar, antes de que se introduzca en una célula huésped, mediante mutaciones, inserciones, deleciones o sustituciones de ácidos nucleicos sencillos o una parte de la secuencia heteróloga del ácido nucleico, siempre y cuando tales secuencias modificadas muestren la misma función (funcionalmente equivalente) como la secuencia referencia. Una secuencia heteróloga del ácido nucleico como se hace referencia en el presente documento abarca tanto las secuencias nucleicas que se originan de un dominio diferente (imperio) de organismos, tales como a partir de eucariotas (de origen eucariota), tales como, por ejemplo, anticuerpos humanos que han sido utilizado en bibliotecas de expresión in fago y de las cuales ácidos nucleicos sencillos como una parte de la secuencias del ácido nucleico que se han modificado de acuerdo con el "uso del codón" de un huésped.

La "región de iniciación de la traducción" es una región señal que promueve el inicio de la traducción y que funciona como el sitio de unión al ribosoma tal como la secuencia Shine Dalgarno.

La "región de terminación de la transcripción" se refiere a una secuencia que causa que la ARN polimerasa termine la transcripción. La región de terminación de la transcripción, usualmente es parte de una unidad transcripcional y aumenta la estabilidad del ARNm.

El vector de acuerdo con la invención es preferiblemente un plásmido autónomo o de auto-replicación, un cósmido, un fago, un virus o un retrovirus. Una amplia variedad de combinaciones huésped/vector se pueden emplear en la expresión de las secuencias de ácido nucleico de la presente invención. Los vectores de expresión útiles, por ejemplo, pueden consistir en segmentos de secuencias de ácido nucleico cromosómico, no-cromosómico y/o sintético. Los vectores apropiados incluyen vectores con rango de huésped específico tales como vectores específicos para, por ejemplo, *E. coli*, así como vectores con un amplio rango de huésped tal como vectores útiles para bacterias tanto gramnegativas como grampositivas. Se pueden usar plásmidos de "baja copia", "media-copia" así como "alta copia".

Los vectores útiles para, por ejemplo, la expresión en *E. coli* son: pQE70, pQE60 y pQE-9 (QIAGEN, Inc.); pBluescript Vektoren, Phagescript Vektoren, pNH8A, pNH16a, pNH 18A, pNH46A (Stratagene Cloning Systems, Inc.); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia Bio-tech, Inc.); pLG338, pACYC184, pBR322, pUC18, pUC19, pKC30, pRep4, pACYC177, pACYC184, pRSF1010 y pBW22 (Wilms y col., 2001, Biotechnology y Bioengineering, 73 (2) 95-103) o derivados de estos. Otros plásmidos útiles son bien conocidos por alguien de habilidad en la técnica y se describen por ejemplo en "Cloning Vectors" (Eds. Pouwels P. H. y col. Elsevier, Amsterdam-New York-Oxford, 1985).

Los vectores preferidos de la presente invención son plásmidos autónomos o de auto-replicación, son más preferidos los vectores con rango de huésped específico, por ejemplo tal como vectores específicos para *E. coli*. Más preferidos son pBR322, pUC18, pACYC177, pACYC184, pRSF1010 y pBW22 o derivados de estos de los mismos.

En una realización preferida, la región del promotor rhaBAD del operón L-ramnosa, o una variante funcionalmente equivalente del mismo, es el promotor ajustable. La transcripción del promotor rhaBAD está extremadamente bien regulada y estrechamente controlada por la concentración de ramnosa del medio de cultivo. Ajustando la concentración de ramnosa del medio de cultivo se puede controlar muy bien la actividad del promotor rhaBAD. El promotor rhaBAD también cubre una amplia gama de intensidades de la expresión. En una realización concreta, el promotor rhaBAD consiste en la secuencia
gcccatcttctgtcagtaacgagaaggtcgcgaattcaggcgcttttagactgtcgtaatgaaattcag.

O una secuencia complementaria de esta y las variantes de estas. Una variante del promotor útil en la invención es una capaz de realizar esencialmente las mismas funciones que el promotor original, en particular la transcripción directa a modo de una ARN polimerasa. En otra realización preferida, el promotor ajustable es el promotor araBAD o una variante funcionalmente equivalente del mismo.

En otra realización preferida de la invención, el vector expresable en un huésped procariota comprende aparte de la región del promotor rhaBAD del operón L-ramnosa ligado operablemente a una unidad transcripcional, además las secuencias que codifican los activadores específicos de la L-ramnosa RhaS y RhaR, incluyendo sus respectivas secuencias del promotor nativo. Tras la expresión de las proteínas RhaS y RhaR controlan la actividad del promotor rhaBAD

Una "membrana biológica", "biomembrana" o simplemente "membrana" es una capa antipática que encierra o separa y que actúa como barrera dentro o alrededor de una célula. Casi siempre, es una bicapa lipídica compuesta por una doble capa de moléculas de tipo lípido, específicamente fosfolípidos, con proteínas ocasionales entrecruzadas.

Una "proteína transmembrana" o "proteína integral de membrana" es una proteína que atraviesa toda la membrana biológica. Las proteínas transmembrana se agregan y precipitan en agua. Requieren detergentes o disolventes no polares para extracción, aunque algunos de ellos (barriles beta) pueden también extraerse usando agentes desnaturizantes. Existen dos tipos principales de proteínas transmembrana. a) Alfa-hélice. Estas proteínas están presentes en todos los tipos de membranas biológicas, incluidas las membranas externas. Esta es la principal categoría de proteínas transmembrana. b) Barriles beta. Estas proteínas solo se encuentran en las membranas externas de bacterias gramnegativas, la pared celular de las bacterias grampositivas y las membranas externas de mitocondrias y cloroplastos.

Una "proteína de membrana interna" es una proteína que reside, al menos en parte, en la membrana interna de bacterias gramnegativas, tales como *E. coli*. Normalmente, una proteína de membrana interna integral es una alfa-hélice.

Una "proteína de membrana externa" es una proteína que reside, al menos en parte, en la membrana externa de bacterias gramnegativas, tales como *E. coli*. Normalmente, una proteína de membrana externa integral es un barril beta.

Una "proteína secretada" o "proteína secretora" se refiere a una proteína que se transloca a través de al menos una biomembrana.

La lisozima del bacteriófago T7 (lisozima de T7) es una proteína bifuncional que corta enlaces amida en la pared celular bacteriana y se une e inhibe la transcripción por la ARN polimerasa de T7. (Moffatt y Studier (1987) Cell 49: 221-7; Jeruzalmi y Steitz (1998) EMBO J 14: 4101-13; Cheng, y col. (1994) Proc Natl Acad Sci U S A 91:4034-8). Estas funciones no están separadas en dominios independientes y una molécula individual puede realizar únicamente una función a la vez (Jeruzalmi y Steitz (1998) EMBO J 14: 4101-13). La unión de la lisozima de T7 a la ARN polimerasa de T7 cierra la ARN polimerasa de T7 en un estado no procesador, de modo que se reduce el inicio de las cadenas de ADN de promotores de fago (Jeruzalmi and Steitz (1998) EMBO J 14: 4101-13). Hay mutantes de T7Lys que carecen de la actividad amidasa pero que todavía se pueden unir e inhibir la T7RNAP (Cheng, y col. (1994) Proc Natl Acad Sci U S A 91:4034-8). Ejemplos de mutantes deficientes en amidasa de la lisozima de T7 que conservan actividad inhibidora de ARN polimerasa de T7 son: Y46F, K128Q, K128W, K128Y, K128M, K128I (Cheng, y col. (1994) Proc Natl Acad Sci U S A 91:4034-8). En la presente invención se pueden usar "genes variantes de la lisozima de T7", es decir formas silvestre y mutantes de la T7Lys que conservan actividad de inhibición de la ARN polimerasa de T7. El mutante K128Y de la lisozima de T7 también se denomina LysY.

La ARN polimerasa de T7 es la ARN polimerasa codificada por el bacteriófago T7. También se pueden usar mutantes de la ARN polimerasa de T7 (Makarova, y col. (1995) Proc Natl Acad Sci U S A 92:12250-4).

Los promotores de T7 son los promotores reconocidos por la ARN polimerasa de T7 (Chamberlin y col., (1970) Nature 228, 227-231; Oakley y Coleman (1977) PNAS 74, 4266-4270; Rosa (1979) Cell 16, 815-825; Panayotatos y Wells (1979) Nucl. Acids Res. 13, 2227-2240; Dunn y Studier (1983) J. Mol. Biol. 166, 477-535/175, 111-112). El promotor $\phi 10$ es, con mayor frecuencia, usado en los sistemas de expresión de T7 y normalmente se denominan "promotor de T7". El operador lac se puede unir operablemente a un promotor de T7 (promotor de fusión T7/lac o T7lac), que ayuda a controlar la expresión basal en un sistema de expresión de T7 inducible bloqueando el promotor T7 diana con el represor lac (Dubendorff y Studier (1991) J Mol Biol 219:45-59). El promotor T7lac normalmente se denomina también "promotor de T7".

El sistema de T7 es el sistema más usado para producir proteínas recombinantes en *Escherichia coli*. Los genes diana se clonan en pET, u otros plásmidos del sistema T7 bajo el control de un promotor de T7; se induce la expresión proporcionando una fuente de ARN polimerasa de T7 en una célula huésped.

La expresión "ARNm recombinante" se refiere al ARNm que codifica un polipéptido que se va a producir.

La expresión "inductor" se refiere a una molécula pequeña que aumenta la transcripción de un operón. También se refiere a otras señales que pueden aumentar la transcripción de un operón, tal como la luz o la temperatura.

El origen de replicación (también denominado origen de replicación) es una secuencia concreta de ADN en la cual se inicia la replicación del ADN. La replicación del ADN puede progresar a partir de este punto bidireccional o unidireccionalmente.

Un "grupo de incompatibilidad (inc)" es un conjunto de plásmidos que interfieren en la replicación y/o reparto de cada uno y, por tanto, no se pueden mantener de forma estable juntos en el mismo cultivo. La expresión "número de copias de plásmido" se refiere al número medio de un plásmido concreto por célula. El número de copias de plásmido viene determinado por el origen de replicación.

La expresión "proteína diana" se refiere a una proteína recombinante que se sobreexpresa. La secuencia de ácido nucleico que codifica la proteína recombinante está unida operablemente a un promotor de T7 o a un promotor de fusión T7/lac u otros derivados de este tipo. La proteína diana puede ser: a) una proteína heteróloga u homóloga, b) una proteína citoplásmica soluble, secretora soluble o de membrana. En concreto, la proteína diana es cualquier proteína con requisitos postraduccionales complejos. Más específicamente, la proteína diana requiere a) dirigir a la membrana, inserción en la membrana o translocación a través de la membrana, o b) modificaciones postraduccionales, inserciones de grupos prostéticos o la formación de cromóforos. Más específicamente, la proteína diana es una proteína de membrana integral o una proteína secretora, una proteína modificada postraduccionalmente o una proteína que requiere la ayuda de chaperona(s) para su plegamiento.

"Sobreexpresión" y "producción" se van a usar de forma sinónima.

Un "marcador de resistencia a antibiótico" o "marcador de selección" es un fragmento de ADN que contiene un gen cuyo producto confiere resistencia a un antibiótico (p. ej., cloranfenicol, ampicilina, gentamicina, estreptomycin, tetraciclina, kanamicina, neomicina) o la capacidad para crecer en medios selectivos (p. ej., ura (uracilo), leu (leucina), trp (triptófano), his (histidina). Las cepas usadas son mutantes auxótrofos que no pueden formar las moléculas mencionadas anteriormente. El vector contiene el gen que puede convertir un auxótrofo en un no auxótrofo, es decir el vector (tras la transformación) permite que la célula crece en medio que carece de la(s) molécula(s) en el que no podía crecer antes de la transformación (Amberg y col. (2005) Methods in Yeast Genetics. Cold Spring Harbor Laboratory Press). Normalmente, los plásmidos que contienen un marcador de resistencia a antibiótico para forzar a la célula bacteriana a mantener el plásmido.

Una realización de la presente invención se muestra esquemáticamente en la Figura 2. Las células huésped se transforman con un vector que comprende un promotor ajustable, ejemplificados por el promotor rhaBAD, unido

operablemente a la lisozima de T7. En estas células huésped, los niveles de lisozima de T7 están regulados por la adición de L-ramnosa al medio de cultivo de un modo estrechamente correlacionad, en los que las concentraciones crecientes de L-ramnosa conducen a niveles crecientes de lisozima de T7. La lisozima de inhibe la ARN polimerasa de T7 (T7RNAP) hasta un grado en función de la concentración de L-ramnosa. La T7RNAP transcribe una proteína recombinante diana unida operablemente a un promotor de T7 de, por ejemplo, un vector del sistema pET. Los niveles de ARNm de dicha proteína recombinante dependen de la actividad de T7RNAP y, por tanto, de la concentración de L-ramnosa en el medio de cultivo. Los niveles de ARNm reducidos permite que *E. coli* se enfrente con la tensión por sobreexpresión, que continúe creciendo y que no se vea afectada por la sobreexpresión y, por último, que de más proteína producida. El sistema de acuerdo con la invención permite la inhibición de la ARN polimerasa de T7 por la lisozima T7 de un modo ajustable de forma continua para encontrar las condiciones de expresión óptimas para cada proteína diana.

La lisozima de T7 inhibe la ARN polimerasa de T7. La cantidad de lisozima de T7 respecto a la cantidad de ARN polimerasa de T7 determina el grado de inhibición. En un aspecto, la presente invención proporciona un vector expresable en un huésped, que comprende la región del rhaBAD del operón de L-ramnosa unido operablemente a una secuencia de ácido nucleico que codifica la lisozima T7. Dicho vector tiene que mantenerse de forma estable en el huésped mediante la ayuda de, por ejemplo, un marcador de selección.

Más específicamente, la lisozima de T7 se expresa en un plásmido. Dicho plásmido comprende un gen que codifica una lisozima de T7 unida operablemente al promotor ajustable, un marcador de resistencia a antibiótico y un origen de replicación que no está, preferentemente, en el mismo grupo de incompatibilidad que los plásmidos que contienen el origen de replicación ColEI. En una realización preferida, dicho promotor ajustable es en promotor rhaBAD, inducible por L-ramnosa, y dicho plásmido comprende, preferentemente, el operón rhaSR que codifica los dos activadores específicos de L-ramnosa RhaS y RhaR. En otra realización, el promotor ajustable es el promotor araBAD. Más específicamente, la lisozima de T7 se expresa, preferentemente, en los plásmidos pLEMO (SEC ID N° 1, también indicada como <400> 1), pTACO1-Lys (SEC ID N° 2, también indicada como <400> 2), pTACO1-LysY (SEC ID N° 3, también indicada como <400> 3), pTACO2-Lys (SEC ID N°4, también indicada como <400> 4), pTACO3-Lys o pTACO3-LysY (SEC ID N°5, también indicada como <400> 5). Dichos plásmidos tienen las secuencias proporcionadas más adelante. pTACO1 comprende un origen de replicación cloDF13 y un marcador de resistencia a espectinomicina; pTACO1-LysY comprende un origen de replicación cloDF13 y un marcador de resistencia a espectinomicina y codifica el mutante LysY; pTACO2 y pLEMO comprende un origen de replicación P15A y un marcador de resistencia a cloranfenicol y pTACO3 comprende un origen de replicación repA derivado de pSC101 y un marcador de resistencia a espectinomicina. Dichos plásmidos pueden fabricarse mediante técnicas sintéticas o recombinantes, que, junto con las demás tecnologías moleculares, pueden realizarse como se describe en Maniatis y col. 1982, Molecular Cloning, A laboratory Manual, Cold Spring Harbor Laboratory o en Ausubel y col. 1994, Current protocols in molecular biology, John Wiley and Sons.

Una proteína diana se tiene que expresar a partir de un vector de expresión. Dicho vector tiene que mantenerse de forma estable en el huésped mediante la ayuda de, por ejemplo, un marcador de selección. En particular, el proteína diana se tiene que expresar a partir de un sistema de plásmido de T7.

Una célula huésped se tiene que transformar con un a) vector que comprende un gen que codifica la lisozima de T7 unida operablemente a un promotor ajustable y b) un vector que comprende un gen que codifica una proteína diana unida operablemente a un promotor de T7, preferentemente la proteína diana está codificada en un plásmido del sistema T7.

En una realización comercialmente ventajosa de la presente invención, la célula huésped puede venderse solo portando solo el vector que comprende un gen que codifica la lisozima de T7. El vector se construye de un modo que sea compatible con los vectores de expresión disponibles comercialmente, tal como vectores pET, diseñado para portar un gen que codifica una proteína diana (gen de interés) unido operablemente a un promotor de T7. Dicho vector disponible comercialmente, que tiene el gen que codifica una proteína diana específica, puede estar disponible o construirse en el sitio del cliente y transformarse en la célula huésped de la presente invención en dicho centro.

El genoma de dicho huésped comprende un promotor unido operablemente a una secuencia de ácido nucleico que codifica una ARN polimerasa de T7. En particular, el genoma de dicho huésped comprende un promotor lac inducible por IPTG o uno de sus derivados, tales como el promotor mutante lacUV5 unido operablemente a una secuencia de ácido nucleico que codifica una ARN polimerasa de T7. En particular, el huésped es la cepa BL21 (DE3) de *Escherichia coli* o cualquier otro huésped capaz de expresar la ARN polimerasa de T7.

De acuerdo con la invención, otros promotores preferidos para inducir la lisozima de T7 incluyen el promotor araBAD, inducido por arabinosa, y tiene la secuencia

CCATAAGATTAGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCTACTGT
TTCTCCATACCCGTTTTTTTGGATGGAGT, SEC ID N° 14;

y el promotor rhaBAD, que es inducido por L-ramnosa, y tiene la secuencia

**GCCCATTTCCTGTCAGTAACGAGAAGGTCGCGAATTCAGGCGCTTTT TAGA
CTGGTCGTAATGAAATTCAGCAGGATCAC, SEC ID Nº 15.**

Un huésped que comprende a) un vector que comprende un promotor ajustable unido operablemente a una secuencia de ácido nucleico que codifica la lisozima de T7, b) un promotor unido operablemente a una secuencia de ácido nucleico que codifica la ARN polimerasa de T7 y c) un promotor de T7 unido operablemente a una secuencia de ácido nucleico que codifica una proteína diana que normalmente crece en condiciones estándar con respecto a la temperatura en medio de cultivo estándar.

Normalmente, las células huésped se cultivan en medio convencional, como se conoce en la técnica, tal como medio complejo como caldo de Luria o "medio de caldo nutriente para levaduras" o un medio que contiene glicerol, como han descrito Kortz y col., 1995, J. Biotechnol. 39, 59-65 o un medio con sales minerales, como han descrito Kulla y col., 1983, Arch. Microbiol, 135, 1. El medio podría modificarse según sea adecuado, por ejemplo añadiendo ingredientes adicionales, tales como tampones, sales, vitaminas o aminoácidos. También se pueden usar medios diferentes o combinaciones de medios durante el cultivo de las células. Preferentemente, el medio usado como medio básico no debería incluir L-ramnosa para conseguir una regulación estrecha de la región promotora de L-ramnosa.

Si el huésped es la cepa BL21 (DE3) de *Escherichia coli*, un cultivo durante la noche de dicho huésped se vuelve a diluir hasta una A₆₀₀ de 0,01 a 0,05. El medio de cultivo se suplementa con un inductor que induce la expresión de la lisozima de T7 del promotor ajustable. En concreto, el medio de cultivo se suplementa con 10 µM de L-ramnosa 10 mM para inducir la expresión de la lisozima de T7 del promotor rhaBAD. La concentración de L-ramnosa en el medio de cultivo determina la expresión de la lisozima de T7 que determina la inhibición de la ARN polimerasa de T7 y, por tanto, la expresión de dicha proteína diana y la concentración de ARNm recombinante en la célula huésped.

Cuando el cultivo de dicho huésped es a una A₆₀₀ de 0,4 a 0,5, la expresión de la proteína diana se induce por la adición del inductor al medio de cultivo, en concreto por la adición de IPTG 0,01 a 10,0 mM, más específicamente por la adición de IPTG 0,1-1 mM. El inductor induce la expresión de ARN polimerasa de T7 que se une al promotor de T7 unido operablemente a las secuencias de ácido nucleico que codifica la proteína diana y transcribe la proteína diana. La actividad de la ARN polimerasa de T7 depende de la concentración de la lisozima de T7 en la célula huésped. La cantidad de ARNm transcrito depende de la actividad de la ARN polimerasa de T7.

La expresión de la proteína diana se monitoriza tras la inducción de la expresión. La concentración del inductor, en particular L-ramnosa, que da la cantidad más elevada de proteína diana se escoge para la expresión de la proteína diana. Esto se puede monitorizar mediante, por ejemplo, SDS-PAGE combinada con tinción con azul de Coomassie/plata, transferencia Western o variantes de los mismos, como transferencia puntual (<http://www.cshprotocols.org/>). Si la proteína fluorescente verde (GFP) se condensa con la proteína sobreexpresada, también se puede usar fluorescencia de células enteras y en gel para monitorizar la expresión de la proteína diana. Cuando una proteína de membrana se condensa con la GFP, la GFP solo es fluorescente cuando la fusión se expresa en la membrana y no en los cuerpos de inclusión/agregados (Drew y col., Protein Sci. (2005) 8, 2011-7; Drew y col., Nat. Methods (2006) 4, 303-13).

La concentración de inductor para la inducción de la expresión de la lisozima de T7 se puede escoger de un modo tal que el índice de ARNm recombinante producido y, por tanto, el índice de proteína diana recombinante producida, sea óptimo para la biogénesis postraduccional de la proteína diana recombinante. Por ejemplo, con respecto a una proteína de membrana, sus efectos tóxicos de su sobreexpresión se pueden minimizar armonizando la traducción y la inserción en la membrana y, con respecto a una proteína que requiere la ayuda de una chaperona para el plegamiento adecuado, se puede establecer el índice de producción de la proteína diana recombinante para permitir que la(s) chaperona(s) mantengan la carga de trabajo del plegamiento (es decir, la capacidad de la chaperona es suficiente para ayudar al plegamiento de la proteína sobreexpresada).

Como sistemas de cultivo celular, cultivos continuos o discontinuos, tales como cultivos por lotes o cultivos de alimentación por lotes, se pueden aplicar en tubos de cultivos, matraces de agitación o fermentadores bacterianos.

Tras la expresión en la célula huésped, el producto expresado tal como el polipéptido de interés se puede recuperar del cultivo de células huésped. Con el fin de obtener un rendimiento máximo del producto expresado, las células normalmente se recogen al final del cultivo y se lisan, tal como lisado mediante tratamiento con lisozimas, ultrasonidos o prensa francesa. Por tanto, normalmente los polipéptidos se obtienen primero como lisado bruto de las células huésped. Después, se pueden purificar mediante procedimientos estándar de purificación de proteínas conocidos en la técnica, que pueden incluir precipitación diferencia, cromatografía de tamiz molecular, cromatografía de intercambio iónico, enfocado isoelectrico, electroforesis en gel, afinidad y cromatografía de inmunoafinidad. En, por ejemplo, Ausubel y col., citado anteriormente y Wu y col. (eds.), Academic Press Inc., N. Y.; Immunochemical Methods In Cell And Molecular Biology se describen estos procedimientos bien conocidos y de práctica rutinaria.

Las Figuras 1A-E muestran los mapas de plásmidos que expresan lisozima de T7 o variantes de la misma a partir de un promotor inducible por L-ramnosa. Un pTACO1-Lys comprende un origen de replicación cloDF13 y un marcador de resistencia a espectinomicina; B, pTACO1-LysY comprende un origen de replicación cloDF13 y un marcador de resistencia a espectinomicina y codifica el mutante LysY; C, pTACO2-Lys comprende un origen de replicación P15A y un marcador de resistencia a cloranfenicol; D, pLEMO comprende un origen de replicación P15A y un marcador de resistencia a cloranfenicol y codifica el mutante LysY; y E) pTACO3-LysY comprende un origen de replicación repA derivado de pSC101 y un marcador de resistencia a espectinomicina y codifica el mutante LysY. Los plásmidos de las Figuras 1A-E también incluyen los dos activadores específicos de L-ramnosa RhaS y RhaR.

Las secuencias de los plásmidos se proporcionan en el listado de secuencias al final de la descripción.

La Figura 2 muestra un modelo del control de la actividad de la ARN polimerasa de T7 mediante titulación de la expresión de la lisozima de T7. La expresión de la lisozima de T7 está bajo el control del promotor rhaBAD y está inducida por la adición de L-ramnosa. En este caso se usa el mutante LysY de la lisozima de T7 con deficiencia de la actividad amidasa.

La Figura 3 muestra los niveles de expresión de la lisozima de T7 en células Lemo21(DE3), cultivadas en presencia de concentraciones diferentes de L-ramnosa. La expresión de la lisozima de T7 se monitorizó mediante SDS-PAGE/transferencia Western usando un anticuerpo contra la lisozima de T7. Por carril se cargaron cantidades iguales, basándose en las lecturas de la DO₆₀₀, de material de células enteras. Los resultados muestran que el nivel de expresión de la lisozima de T7 aumenta con concentraciones crecientes de L-ramnosa en el medio de cultivo.

La Figura 4 muestra los niveles de ARNm de la proteína recombinante YidC-GFP, expresada en células Lemo21(DE3) (BL21 (DE3) que alojan pLemo) cultivadas en presencia de concentraciones diferentes de L-ramnosa (en el presente documento, rha). La figura muestra un rendimiento decreciente de ARNm, que indica un índice de transcripción menor, con concentraciones crecientes de L-ramnosa. Esto indica además una mayor inhibición de la ARN polimerasa de T7 por los niveles crecientes de la lisozima de T7 alcanzados al aumentar los niveles de L-ramnosa.

La Figura 5 muestra A, crecimiento y B, expresión proteica de las células Lemo21(DE3), que expresan la proteína de membrana recombinante YidC-GFP y cultivadas en presencia de concentraciones diferentes de L-ramnosa. La expresión de YidC-GFP se monitorizó mediante fluorescencia con GFP de células enteras. Los resultados de la Figura A muestran un nivel creciente de crecimiento con niveles crecientes de L-ramnosa (es decir, niveles de lisozima de T7), que indican un índice menor de expresión de la proteína de membrana, debido a la inhibición de la ARN polimerasa de T7, tiene un efecto positivo sobre la viabilidad y el crecimiento de las células. Los resultados de la figura B indican un nivel creciente de expresión de la proteína de membrana e inserción en la membrana con niveles crecientes de L-ramnosa, es decir con niveles crecientes de lisozima de T7 y un índice decreciente de transcripción.

La Figura 6 muestra la fracción de células que no expresan YidC-GFP 4 horas y 18 horas, respectivamente, después de la inducción con IPTG, en presencia de concentraciones diferentes de L-ramnosa. Los resultados muestran que, de un modo más evidente tras 18 horas, la fracción de células negativas para YidC-GFP, disminuye con niveles crecientes de L-ramnosa (es decir, niveles de lisozima de T7), lo que indica que concentraciones crecientes de L-ramnosa impiden el sobrecrecimiento de células que no expresan en el cultivo.

La Figura 7 muestra A, crecimiento y B, expresión proteica de las células Lemo21(DE3), que expresan la proteína recombinante YidC-GFP y cultivadas en presencia de concentraciones diferentes de L-ramnosa. Los resultados indican que, en comparación con las células control (que no expresan proteínas de membrana), la respiración de las células con actividad ARN polimerasa de T7 completa (ausencia de presencia de L-ramnosa y, por tanto, sin inhibición de la lisozima de T7) se ve afectada, mientras que la respiración de las células en presencia de una concentración óptimamente establecida de L-ramnosa (y, por tanto, actividad de ARN polimerasa de T7 óptima regulada por disminución) no se ve afectada.

La Figura 8 muestra la agregación de proteínas en células Lemo21(DE3), que expresan la proteína recombinante YidC-GFP durante 4 horas en presencia de concentraciones diferentes de L-ramnosa. A, agregados proteicos se aislaron y analizaron mediante SDS-PAGE. Los geles se tiñeron con azul de Coomassie coloidal. B, cuantificación de la fracción de proteínas agregadas en Lemo21(DE3) tras sobreexpresión de YidC-GFP. Los resultados muestran niveles decrecientes de agregados proteicos citoplásmicos con niveles crecientes de L-ramnosa, que indican que la actividad regulada por disminución de ARN polimerasa de T7, mediante niveles crecientes de lisozima de T7, conduce a niveles decrecientes de agregación proteica.

La Figura 9 muestra A, crecimiento y B, expresión proteica de la proteína recombinante YidC-GFP en un cultivo semicontinuo, en una comparación entre cepas diferentes (BL21 (DE3), BL21(DE3)pLysS and Lemo21(DE3)). Las cepas que sobreexpresan las fusiones YidC-GFP se retrodiluyeron a 1:10 cada 4 horas en medio con IPTG. Se cultivaron células Lemo21(DE3) en presencia de una concentración óptimamente establecida (1000 µM) de L-ramnosa. El crecimiento y la expresión proteica se monitorizaron midiendo la A₆₀₀ y la fluorescencia GFP, respectivamente. Los gráficos muestran valores acumulados. Los resultados muestran que Lemo21(DE3) que expresa YidC-GFP es completamente estable y funciona tan bien o mejor que las demás cepas.

La Figura 10 muestra una comparación de la expresión (fluorescencia GFP) de las proteínas de fusión de GFP en las diferentes cepas BL21 (DE3), BL21(DE3)pLysS y Lemo21(DE3). Las proteínas incluyen las proteínas de membrana de *E. coli* EnvZ GltP, PheP, PstA, RarD, UhpT, YfbF, YiaM, YidC, YijD, YqcE, YdaN; el transportador A de proteínas de membrana d *Shewanella oneidensis* (SOTA); el receptor de la neurotensina proteína de membrana de rata (NTR); las proteínas de membrana humanas tetraspanina TspA y TspB; y la proteína soluble eGFP. La concentración óptima de L-ramnosa se determinó y usó para cada proteína. Los niveles de expresión de eGFP se dividen por 50 para ajustarlos en el gráfico. La expresión de las proteínas de fusión de GFP y eGFP se monitorizó mediante fluorescencia con GFP de células enteras. La proteína sensora de osmolaridad de *E. coli* EnvZ y el transportador A de *S. oneidensis* (SOTA) solo se expresaron bien en in BL21 (DE3) y Lemo(DE3) sin nada de L-ramnosa. En la mayoría de los casos, Lemo21 (DE3) funcionó mejor que BL21(DE3) and BL21(DE3)pLysS, lo que indica que Lemo21 (DE3) permite la sobreexpresión de proteínas de membrana usando solo una cepa y una simple titulación de L-ramnosa en lugar de una multitud de cepas.

La Figura 11 muestra los niveles de expresión de la chaperona soluble que requiere luciferasa-His8, luciferasa marcada con una cola de His8 en el extremo C, en células Lemo21(DE3) cultivadas en presencia de concentraciones diferentes de L-ramnosa. La expresión de luciferasa-His8 se monitorizó mediante SDS-PAGE/transferencia Western usando un anticuerpo contra las colas de His. Por carril se cargaron cantidades iguales, basándose en las lecturas de la DO₆₀₀, de material de células enteras. En este caso, la concentración óptima para alcanzar los niveles de expresión más altos es 50 µM de L-ramnosa. Los resultados muestran que el sistema de expresión de la presente invención es adecuado y permite la optimización de la sobreexpresión también de proteínas solubles que requieren chaperonas.

La Figura 12 muestra A, niveles de expresión de la uracilo ADN glicosilasa de *E. coli*, una proteína citoplásmica soluble, de expresión en células Lemo21(DE3) cultivadas en presencia de concentraciones diferentes de L-ramnosa. Se cargaron cantidades iguales de células por carril de un gel de SDS-PAGE y el gel, tras la finalización de la carrera, se tiñó con azul de Coomassie, y B, expresión de uracil ADN glicosilasa de de *E. coli* en células Lemo21(DE3) cultivadas en presencia de L-ramnosa 0 µM, en comparación con la expresión de uracil ADN glicosilasa de de *E. coli* en BL21(DE3), BL21(DE) + pLysS, T7 Express, T7 Express + pLysY y T7 Express + pLysY/lq de *E. coli*. Los resultados de la Figura 12A son otro ejemplo que muestra que el sistema de expresión de la presente invención es adecuado para proteínas solubles. La expresión es mejor en Lemo21(DE3) en comparación con las otras cepas, aunque no se añadiera ramnosa. Es probable que el efecto se deba a la expresión de fondo de T7Lys..

Ejemplos

Todos los experimentos se pueden realizar siguiendo la descripción y con referencia a los manuales estándar conocidos en el campo, tales como Maniatis y col.(Sambrook, J., Fritsch, E.F., and Maniatis, T., in Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY, Vol. 1, 2, 3 (1989) or at <http://www.cshprotocols.org/>.)

Materiales usados: La amplificación del ADN se realizó usando el sistema de PCR Expand Long Template PCR de Roche. La extracción del ADN de los geles de agarosa se realizó con Qiaex II de Qiagen. La purificación de productos de reacciones enzimáticas se realizó con el kit de purificación QIAquick PCR Purification Kit de Qiagen. Las enzimas de restricción procedían de Fermentas. La ligasa de T4 procedía de Invitrogen. El kit para mutagénesis dirigida a sitio de Quickchange procedía de Stratagene. Los dNTP procedían de Fermentas. La secuenciación se realizó usando el kit de secuenciación en ciclos BigDye Terminator v1.1 de Applied biosystems.

Construcción de vectores que contienen la lisozima de T7 o variantes de la misma bajo el control del promotor rhaBAD.

Ejemplo 1

Construcción de pTACO1 y pTACO2: Una porción de pRha109 (Giacalone, y col. (2006) Biotechniques 40: 355-64) que contiene los genes rhaS, rhaR, el promotor rhaBAD y el sitio de clonación múltiple se amplificó mediante PCR usando los cebadores 'pRha_EheI_f, GCGCGCGCGCCCCGATAAGCTTAATTAATCTTTCTGCG' y 'pRha_XmaJI_r, CGCGCGCCTAGGGCATATGAATACGCCCTTTTCGGATG'. El producto de la PCR se pasó por un gel de agarosa al 1 % y se purificó usando Qiaex II de Qiagen. El producto de la PCR purificado se digirió con EheI y XmaJI y se ligó en la estructura de pACYC-DUET-1 y pCDF-DUET-1, respectivamente, (Novagene) (Tolia y Joshua-Tor (2006) Nat Methods 3:55-64) digeridos con EheI y XmaJI. El sitio de restricción XbaI dentro del origen p15A de pACYC-DUET-1 se eliminó mediante mutagénesis dirigida as tio Quickchange (TCTAGA a TCTAGG) siguiendo las instrucciones del fabricante (Stratagene). El plásmido resultante de la unión de dicho producto de PCR con pCDF-DUET-1 se denominó pTACO1, el plásmido resultante de la unión de dicho producto de PCR con pACYC-DUET-1 se denominó pTACO2.

Ejemplo 2

Construcción de pTACO1-Lys, pTACO1-LysY, pTACO2-Lys y pLEMO

pTAC01-Lys, pTAC01-LysY, pTACO2-Lys y pLemo se construyeron uniendo el producto amplificado por PCR de la lisozima de T7 o el mutante K128Y de la lisozima de T7 (Cheng, y col. (1994) Proc Natl Acad Sci U S A 91:4034-8) (LysY) en pTACO1 y pTACO2, respectivamente, se digirieron con Sall y BamHI. Los cebadores de PCR usados fueron 'T7Lys_Sall_f, GCGCGCGTCGACATGGCTCGTGTACAGTTTAA' y 'T7Lys_BamHI_r, CGCGCGGGATCCTTATCCACGGTCAGAAGTGA'.

Ejemplo 3

Construcción de pTACO3-LysY

Una porción de pCL1920 (Lerner e Inouye (1990) Nucleic Acids Res 18:4631) que contiene el gen aadA1 y el origen de replicación repA de pSC101 se amplificó mediante PCR usando los cebadores 'pCL1920_XmaJI_f, CGCGCGCCTAGGGACAGTAAGACGGGTAAGCC' y 'pCL1920_HindIII_r, GCGCGCAAGCTTCTAACGCTTGAGTTAAGCCG'. El producto purificado por PCR se digirió y clonó en el fragmento de pTACO1-LysY digerido con HindIII y XmaJI que contenía los genes rhaS, rhaR, el promotor rhaBAD y el gen LysY. El plásmido resultante se denominó pTACO3-LysY.

Las secuencias de nucleótidos de los plásmidos se confirmaron mediante secuenciación (BM Labbet AB, Furusund, Suecia).

Ejemplo 4

Construcción de células huésped portadoras de pLemo

Como células huésped se usaron células BL21 (DE3) de *E. coli*. BL21 (DE3) es una cepa bien conocida y disponible comercialmente portadora de DE3, un profago λ que comprende el gen de la ARN polimerasa de T7 y lacIq. El gen que codifica la ARN polimerasa de T7 está bajo el control del promotor *lacUV5*. Las células BL21 (DE3) químicas o electrocompetentes se transformaron con pLemo usando procedimientos estándar, lo que da como resultado Lemo21(DE3), es decir células BL21 (DE3) que alojan pLEMO.

Ejemplo 5

Expresión de proteína diana y titulación de la actividad de T7RNAP mediante expresión de la lisozima de T7: Las células de las cepas indicadas respectivamente se transformaron con un vector de expresión derivado de pET28(a+) que comprende el gen que codifica la proteína diana, bajo el control de un promotor de T7. Las proteínas diana de membrana se condensaron en el extremo C con GFP, mientras que las proteínas diana solubles luciferasa-His₈ y uracilo ADN glicosilasa de *E. coli* no se condensaron a GFP. Las células se cultivaron aeróbicamente en caldo Luria-Bertani estándar suplementado con canamizina (50 µg/ml) (y cloranfenicol (30 µg/ml) si la cepa de expresión contenía un plásmido derivado de pACYC). Los cultivos durante la noche se diluyeron a 1:50. El crecimiento se monitorizó midiendo la DO₆₀₀ con un espectrofotómetro Shimadzu UV-1601. La expresión de la lisozima de T7 se indujo mediante la adición de L-ramnosa 10-4000 µM al medio de cultivo ya al inicio. La expresión de la lisozima de T7 se monitorizó mediante SDS-PAGE/transferencia Western. Para todos los experimentos, la expresión de la proteína diana se indujo mediante la adición de isopropil-β-D-tiogalactopiranosido (IPTG) 0,4 mM (concentración final) a una DO₆₀₀ de 0,4-0,5 y las células se recogieron a varios puntos de tiempo tras la inducción para análisis posterior. La expresión de la proteína diana se monitorizó midiendo la fluorescencia de GFP como se ha descrito anteriormente (Drew, D., y col. (2005) Protein Sci 14:2011-7; Drew, D., y col. (2006) Nat Methods, 2006. 3:303-13). La fluorescencia de GFP es un buen indicador de la expresión de proteínas de membrana, ya que sólo en la membrana son fluorescentes las proteínas de membrana insertadas (Drew, D. E., y col. (2001) FEBS Letters 507:220-4). El análisis de las células que sobreexpresan proteínas de fusión-GFP y células control por medio de citometría de flujo se realizó como se ha descrito anteriormente usando un instrumento FACSCalibur (BD Biosciences) (Wagner, y col. (2007) Mol Cell Proteomics 6:1527-50). La expresión de la luciferasa-His₈ y la uracilo ADN glicosilasa de *E. coli* se monitorizó usando SDS-PAGE/transferencia Western y SDS-PAGE/tinción con azul de Coomassie, respectivamente.

LISTADO DE SECUENCIAS

<110> Xbrane Bioscience AB

<120> New vector

<130> P8207SE00

<160> 17

<170> PatentIn versión 3.4

<210> 1

ES 2 381 433 T3

<211> 5001

<212> ADN

<213> Artificial

5 <220>

<223> pLEMO

<400> 1

gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactggtcgt	60
aatgaaattc aggaggtggt cgacatggct cgtgtacagt ttaaacaacg tgaatctact	120
gacgcaatct ttgttcactg ctcggctacc aagccaagtc agaatgttgg tgtccgtgag	180
attcgccagt ggcacaaaga gcagggttgg ctcgatgtgg gataccactt tatcatcaag	240
cgagacggta ctgtggaggc aggacgagat gagatggctg taggctctca cgctaagggt	300
tacaaccaca actctatcgg cgtctgcctt gttggtggta tcgacgataa aggtaagttc	360
gacgctaact ttacgccagc ccaaagcaa tcccttcgct cactgcttgt cacactgctg	420
gctaagtacg aaggcgctgt gcttcgcgcc catcatgagg tggcgccgta cgcttgccct	480
tcgttcgacc ttaagcgttg gtgggagaag aacgaactgg tcacttctga ccgtggataa	540

ES 2 381 433 T3

ggatccccgc gccctcatcc gaaagggcgt attcatatgc cctaggctgc tgccaccgct	600
gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgctg	660
aaacctcagg catttgagaa gcacacggtc aactgcttc cggtagtcaa taaaccggtg	720
aaccagcaat agacataagc ggctatTTaa cgaccctgcc ctgaaccgac gaccgggtcg	780
aatttgcttt cgaatttctg ccattcatcc gcttattatc acttattcag gcgtagcacc	840
aggcgTTTaa gggcaccaat aactgcctta aaaaaattac gccccgccct gccactcatc	900
gcagtactgt tgtaattcat taagcattct gccgacatgg aagccatcac aaacggcatg	960
atgaacctga atcgccagcg gcacagcac cttgtgcct tgcgataat atttgcccat	1020
agtgaaaacg ggggcgaaga agttgtccat attggccacg tttaaataaa aactggtgaa	1080
actcaccag ggattggctg agacgaaaaa catattctca ataaaccctt tagggaaata	1140
ggccaggTtt tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa actgccggaa	1200
atcgctcgtgg tattcactcc agagcgatga aaacgtttca gtttgctcat ggaaaacggg	1260
gtaacaaggg tgaacactat cccatatcac cagctcaccg tctttcattg ccatacggaa	1320
ctccggatga gcattcatca ggcgggcaag aatgtgaata aaggccgat aaaacttgtg	1380
cttatttttc ttacgggtct ttaaaaaggc cgtaatatcc agctgaacgg tctggttata	1440
ggtacattga gcaactgact gaaatgcctc aaaatgttct ttacgatgcc attgggatat	1500
atcaacgggtg gtatatccag tgattttttt ctccatttta gcttccttag ctctgaaaa	1560
tctcgataac tcaaaaaata cgcccggtag tgatcttatt tcattatggt gaaagttgga	1620
acctcttacg tgccgatcaa cgtctcattt tcgccaaaag ttggcccagg gcttcccggg	1680
atcaacaggg acaccaggat ttatttatc tgccaagtga tcttcgtca caggatttta	1740

ES 2 381 433 T3

ttcggcgcaa agtgcgtcgg gtgatgctgc caacttactg atttagtgta tgatggtggt	1800
tttgagggtgc tccagtggct tctgtttcta tcagctgtcc ctectgttca gctactgacg	1860
gggtggtgcg taacggcaaa agcaccgccg gacatcagcg ctagcggagt gtatactggc	1920
ttactatggt ggactgatg aggggtgtag tgaagtgtt catgtggcag gagaaaaag	1980
gctgcaccgg tgcgtcagca gaatatgtga tacaggatat attccgcttc ctgctcact	2040
gactcgtac gctcggtcgt tcgactgcgg cgagcggaaa tggcttacga acggggcgga	2100
gatttcttgg aagatgccag gaagatactt aacagggaag tgagagggcc gcggcaaagc	2160
cgtttttcca taggctccgc cccctgaca agcatcacga aatctgacgc tcaaatacgt	2220
ggtggcgaaa cccgacagga ctataaagat accaggcggt tcccctggcg gtcctctgt	2280
gcgctctct gttctgcct ttcggtttac cgggtgcatt ccgctgttat ggccgcgttt	2340
gtctcattcc acgctgaca ctcagttccg ggtaggcagt tcgctccaag ctggactgta	2400
tgcacgaacc ccccgttcag tccgaccgct gcgccttacc cggtaactat cgtcttgagt	2460
ccaaccggga aagacatgca aaagcaccac tggcagcagc cactggtaat tgatttagag	2520
gagttagtct tgaagtcag cgccggttaa ggctaaactg aaaggacaag ttttggtgac	2580
tgcgctcctc caagccagtt acctcggttc aaagagttgg tagctcagag aaccttcgaa	2640
aaaccgccct gcaaggcgg ttttctgtt tcagagcaag agattacgcg cagaccaaaa	2700
cgatctcaag aagatcatct tattaatcag ataaaatatt tctaggtttc agtgcaattt	2760
atctcttcaa atgtagcacc tgaagtcagc ccatacgat ataagttgta attctcatgt	2820
tagtcatgcc ccgcgccac cggaaggagc tgactgggtt gaaggctctc aagggcacgc	2880

ES 2 381 433 T3

gtcgagatcc cgggtgcctaa tgagtgaagct aacttacatt aattgcgttg cgctcactgc 2940

ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg 3000

ggagaggcgg tttgcgtatt gggcgccga taagcttaat taatctttct gcgaattgag 3060

atgacgccac tggctgggcg tcatcccggg ttcccgggta aacaccaccg aaaaatagtt 3120

actatcttca aagccacatt cggtcgaaat atcactgatt aacaggcggc tatgctggag 3180

aagatattgc gcatgacaca ctctgacctg tcgcagatat tgattgatgg tcattccagt 3240

ctgctggcga aattgctgac gcaaaacgcg ctactgcac gatgcctcat cacaaaattt 3300

atccagcgca aagggaacttt tcaggctagc cgccagccgg gtaatcagct tatccagcaa 3360

cgtttcgctg gatgttggcg gcaacgaatc actggtgtaa cgatggcgat tcagcaacat 3420

caccaactgc ccgaacagca actcagccat ttcgtttagca aacggcacat gctgactact 3480

ttcatgctca agctgaccga taacctgccg cgccctgcgc atccccatgc tacctaagcg 3540

ccagtgtggt tgccctgcgc tggcgtaaa tcccggaatc gcccctgcc agtcaagatt 3600

cagcttcaga cgctccgggc aataaataat attctgcaa accagatcgt taacggaagc 3660

gtaggagtgt ttatcgtcag catgaatgta aaagagatcg ccacgggtaa tgcgataagg 3720

gcgatcggtg agtacatgca ggccattacc gcgccagaca atcaccagct cacaaaaatc 3780

atgtgtatgt tcagcaaaga catcttgccg ataacggtca gccacagcga ctgcctgctg 3840

gtcgctggca aaaaaatcat ctttgagaag ttttaactga tgcgccaccg tggctacctc 3900

ggccagagaa cgaagttgat tattcgcaat atggcgtaaa aatacgttga gaagattcgc 3960

gttattgcag aaagccatcc cgtccctggc gaatatcacg cggtgaccag ttaaactctc 4020

ggcgaaaaag cgtcgaaaag tggttactgt cgctgaatcc acagcgatag gcgatgtcag 4080

ES 2 381 433 T3

taacgctggc ctcgctgtgg cgtagcagat gtcgggcttt catcagtcgc aggcggttca 4140
 ggtatcgctg aggcgtcagt cccgtttgct gcttaagctg ccgatgtagc gtacgcagtg 4200
 aaagagaaaa ttgatccgcc acggcatccc aattcacctc atcggcaaaa tggtcctcca 4260
 gccaggccag aagcaagttg agacgtgatg cgctgttttc caggttctcc tgcaaactgc 4320
 ttttacgcag caagagcagt aattgcataa acaagatctc gcgactggcg gtcgagggtta 4380
 aatcattttc cccttcctgc tgttccatct gtgcaaccag ctgtcgcacc tgctgcaata 4440
 cgctgtgggtt aacgcgccag tgagacggat actgccatc cagctcttgt ggcagcaact 4500
 gattcagccc ggcgagaaac tgaaatcgat ccggcgagcg atacagcaca ttggtcagac 4560
 acagattatc ggtatgttca tacagatgcc gatcatgatc gcgtacgaaa cagaccgtgc 4620
 caccggtgat ggtatagggc tgccattaa acacatgaat acccgtgcca tgttcgacaa 4680
 tcacaatttc atgaaaatca tgatgatgtt caggaaaatc cgctgcggg agccgggggtt 4740
 ctatcgccac ggacgcgtta ccagacggaa aaaaatccac actatgtaat acggtcatac 4800
 tggcctcctg atgtcgtcaa cacggcgaaa tagtaatcac gaggtcaggt tcttacctta 4860
 aattttcgac ggaaaaccac gtaaaaaacg tcgatttttc aagatacagc gtgaattttc 4920
 aggaaatgcg gtgagcatca catcaccaca attcagcaaa ttgtgaacat catcacgttc 4980
 atctttccct gggtgccaat g 5001

<210> 2

<211> 4774

<212> ADN

5 <213> artificial

<220>

<223> pTACO1-Lys

10 <400> 2

ES 2 381 433 T3

gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactggctcgt	60
aatgaaattc aggaggttgt cgacatggct cgtgtacagt ttaaacaacg tgaatctact	120
gacgcaatct ttgttcactg ctcggctacc aagccaagtc agaatgttgg tgtccgtgag	180
attcgccagt ggcacaaaga gcaggggttg ctcgatgtgg gataccactt tatcatcaag	240
cgagacggta ctgtggaggc aggacgagat gagatggctg taggctctca cgctaagggt	300
tacaaccaca actctatcgg cgtctgcctt gttggtggta tcgacgataa aggtaagttc	360
gacgctaact ttacgccagc ccaaagtcaa tcccttcgct cactgcttgt cacactgctg	420
gctaagtacg aaggcgctgt gcttcgcgcc catcatgagg tggcgccgaa agcttgccct	480
tcgttcgacc ttaagcgttg gtgggagaag aacgaactgg tcacttctga ccgtggataa	540
ggatccccgc gccctcatcc gaaagggcgt attcatatgc cctaggctgc tgccaccgct	600
gagcaataac tagcataacc ccttgggggc tctaaacggg tcttgagggg ttttttgctg	660
aaacctcagg catttgagaa gcacacggtc aactgcttc cggtagtcaa taaaccggta	720
aaccagcaat agacataagc ggctatttaa cgaccctgcc ctgaaccgac gaccgggtca	780
tcgtggccgg atcttgcggc ccctcggctt gaacgaattg ttagacatta tttgccgact	840
accttggtga tctcgctttt cacgtagtgg acaaattctt ccaactgatc tgcgcgcgag	900
gccaagcgat cttcttcttg tccaagataa gcctgtctag cttcaagtat gacgggctga	960
tactgggccg gcaggcgctc cattgcccag tcggcagcga catccttcgg cgcgattttg	1020
ccggttactg cgctgtacca aatgcggggc aacgtaagca ctacatttcg ctcatcgcca	1080

ES 2 381 433 T3

gccacgtcgg gcggcgagtt ccatagcgtt aaggtttcat ttagcgctc aaatagatcc 1140
 tggtcaggaa ccgatcaaa gagttcctcc gccgctggac ctaccaaggc aacgctatgt 1200
 tctcttgctt ttgtcagcaa gatagccaga tcaatgtcga tcgtggctgg ctccaagata 1260
 cctgcaagaa tgtcattgcg ctgccattct ccaaattgca gttcgcgctt agctggataa 1320
 cgccacggaa tgatgtcgtc gtgcacaaca atggtgactt ctacagcgcg gagaatctcg 1380
 ctctctccag gggaagccga agtttccaaa aggtcgttga tcaaagctcg ccgcttggtt 1440
 tcatcaagcc ttacggtcac cgtaaccagc aaatcaatat cactgtgtgg cttcaggccg 1500
 ccatccactg cggagccgta caaatgtacg gccagcaacg tcggttcgag atggcgctcg 1560
 atgacgcaa ctacctctga tagttgagtc gatacttcgg cgatcaccgc ttccctcata 1620
 ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
 atatttgaat gtatttagaa aaataaaca atagctagct cactcggtcg ctacgtccg 1740
 ggcgtgagac tgcggcgggc gctgcggaca catacaaagt taccacaga ttccgtggat 1800
 aagcagggga ctaacatgtg aggcaaaaca gcagggccgc gccggtggcg tttttccata 1860
 ggctccgccc tctgccaga gttcacataa acagacgctt ttccggtgca tctgtgggag 1920
 ccgtgaggct caaccatgaa tctgacagta cgggcgaaac ccgacaggac ttaaagatcc 1980
 ccaccgtttc cggcggtcg ctccctcttg cgctctctg ttccgaccct gccgtttacc 2040
 ggatacctgt tccgcctttc tcccttacgg gaagtgtggc gctttctcat agctcacaca 2100
 ctggtatctc ggctcgggtg aggtcgttcg ctccaagctg ggctgtaagc aagaactccc 2160
 cgttcagccc gactgctgcg ctttatccgg taactgttca cttgagtcca acccgaaaaa 2220

ES 2 381 433 T3

gcacggtaaa acgccactgg cagcagccat tggtaactgg gagttcgag aggatttggt 2280

tagctaaaca cgcggttgct cttgaagtgt gcgccaaagt ccggctacac tggaaggaca 2340

gatttggttg ctgtgctctg cgaaagccag ttaccacggt taagcagttc cccaactgac 2400

ttaaccttcg atcaaaccac ctccccaggt ggtttttttcg tttacagggc aaaagattac 2460

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctactgaac cgctctagct 2520

ttcagtgcaa tttatctctt caaatgtagc acctgaagtc agccccatac gatataagtt 2580

gtaattctca tgtagtcat gccccgcgcc caccggaagg agctgactgg gttgaaggct 2640

ctcaagggca tcggtcgaga tcccggtgcc taatgagtga gctaacttac attaattgcg 2700

ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 2760

ggccaacgcg cggggagagg cggtttgct attgggcgcc cgataagctt aattaatctt 2820

tctgcgaatt gagatgacgc cactggctgg gcgtcatccc ggtttcccg gtaaaccacca 2880

ccgaaaaata gttactatct tcaaagccac attcggctga aatatcactg attaacaggc 2940

ggctatgctg gagaagatat tgcgcatgac acactctgac ctgtcgaga tattgattga 3000

tggtcattcc agtctgctgg cgaaattgct gacgcaaac gcgctcactg cacgatgcct 3060

catcacaaaa tttatccagc gcaaaggac ttttcaggct agccgccagc cgggtaatca 3120

gcttatccag caacgtttcg ctggatgttg gcggcaacga atcactggtg taacgatggc 3180

gattcagcaa catcaccaac tgcccgaaca gcaactcagc catttcgtta gcaaacggca 3240

catgctgact actttcatgc tcaagctgac cgataacctg ccgcgcctgc gccatcccca 3300

tgctacctaa gcgccagtgt ggttgccctg cgctggcggtt aaatcccga atcgccccct 3360

gccagtcaag attcagcttc agacgctccg ggcaataaat aatattctgc aaaaccagat 3420

ES 2 381 433 T3

cgттаасгга агсгtaggag tgттtатсгt cagcatgaat gтаааgаgа tсgссacggg	3480
таатсгgata agggcgatcg ttgagtacat gcaggccatt accgcgccag асаатсаса	3540
gctcacaaaa atcatgtgta tgttcagcaa agacatcttg cggataacgg tcagccacag	3600
cgactgcctg ctggtcgctg gcaaaaaaat catctttgag aagttttaac tgatgcgcca	3660
ccgtggctac ctcgccaga gaacgaagtt gattattcgc aatatggcgt асааатасgt	3720
tgagaagatt cgcgттattg cagaaagcca tcccgтссct ggcgaatатс acgcggtgac	3780
cagtтааact ctcggcgaaa aagcgтсgaa aagtggттac tgтсgтgaa тссacagсgа	3840
taggcgatgt cagтаасgt ggcтсgтg tggcgtagca gatgtсgggc тттcatcagt	3900
cgсaggсgt тсaggтatсg ctgaggсgtc agтссcgттt gтgгттaag ctгссgatgt	3960
agcgтacгca gtgaaagaga aaattgatcc gccacggcat cccaattcac ctcatcgгca	4020
aaatggтсct ccagccaggc cagaagcaag ttgagacgtg atгсgгтт ttccaggттc	4080
тсctгсааac tgctттtacg cagcaagagc agтаattгca тааасааgat ctгсгсgactg	4140
гсggтсgagg gтааатсatt ттсссctтсc tgгтгттсca тctгтгсаac cagгтгтсгс	4200
acгггггca atacггггг gtтаасгсгс cagтgagacg gataгггсс atссcagггт	4260
tgгггсgca actgattcag cссггсgаgа аactgaaатс gatсгггсgа гсgатасagс	4320
acattggгca gacacagatt atсгггatgt тсataсagat gccgatcatg atгсгггacg	4380
aaacagaccg тгссacсgt gatгггatag ggгггсссат тааасасатg аатасссггг	4440
ссатгттсga саатсасаат ттсатgаааа тсатгатгат гттсaggааа atссгссггс	4500
gggagссggg гттсатсгс сасггасгсг тtaccagacg gaaaaaatc сасactатgt	4560

ES 2 381 433 T3

aatacgggtca tactggcctc ctgatgtcgt caacacggcg aaatagtaat cacgaggtca 4620
 gggttcttacc ttaaattttc gacggaaaac cacgtaaaaa acgtcgattt ttcaagatac 4680
 agcgtgaatt ttcaggaaat gcggtgagca tcacatcacc acaattcagc aaattgtgaa 4740
 catcatcacg ttcattcttc cctggttgcc aatg 4774

<210> 3

<211> 4774

<212> ADN

5 <213> artificial

<220>

<223> pTACO1-LysY

10 <400> 3

gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactgggtcgt 60
 aatgaaattc aggagggtgt cgacatggct cgtgtacagt ttaaacaacg tgaatctact 120
 gacgcaatct ttgttctactg ctcggtctacc aagccaagtc agaattgttg tgcccgtag 180
 attcgccagt ggcacaaaga gcagggttg ctcgatgtgg gataccactt tatcatcaag 240
 cgagacggta ctgtggaggc aggacgagat gagatggctg taggctctca cgctaagggt 300
 tacaaccaca actctatcgg cgtctgcctt gttggtggta tcgacgataa aggtaagtgc 360
 gacgctaact ttacgccagc ccaaattgcaa tcccttcgct cactgcttgt cacactgctg 420
 gctaagtacg aaggcgctgt gcttcgcgcc catcatgagg tggcgccgta cgcttgccct 480
 tcgttcgacc ttaagcgttg gtgggagaag aacgaactgg tcacttctga ccgtggataa 540
 ggatccccgc gccctcatcc gaaagggcgt attcatatgc cctaggctgc tgccaccgct 600
 gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgctg 660

ES 2 381 433 T3

aaacctcagg catttgagaa gcacacggtc aactgcttc cggtagtcaa taaaccggta	720
aaccagcaat agacataagc ggctatttaa cgaccctgcc ctgaaccgac gaccgggtca	780
tcgtggccgg atcttgccgc ccctcggctt gaacgaattg ttagacatta tttgccgact	840
accttggtga tctcgccttt cacgtagtgg acaaattctt ccaactgatc tgcgcgcgag	900
gccaaagcat cttcttcttg tccaagataa gcctgtctag cttcaagtat gacgggctga	960
tactgggccg gcaggcgctc cattgcccag tcggcagcga catccttcgg cgcgattttg	1020
ccggttactg cgctgtacca aatgcgggac aacgtaagca ctacatttcg ctcatcgcca	1080
gccagtcgg gcggcgagtt ccatagcgtt aaggtttcat ttagcgcctc aaatagatcc	1140
tggtcaggaa ccggatcaaa gagttcctcc gccgctggac ctaccaaggc aacgctatgt	1200
tctcttgctt ttgtcagcaa gatagccaga tcaatgtcga tcgtggctgg ctcgaagata	1260
cctgcaagaa tgtcattgcg ctgccattct ccaaattgca gttcgcgctt agctggataa	1320
cgccacggaa tgatgtcgtc gtgcacaaca atggtgactt ctacagcgcg gagaatctcg	1380
ctctctccag gggaagccga agtttccaaa aggtcgttga tcaaagctcg ccgcgttggt	1440
tcataagcc ttacggtcac cgtaaccagc aaatcaatat cactgtgtgg cttcaggccg	1500
ccatccactg cggagccgta caaatgtacg gccagcaacg tcggttcgag atggcgctcg	1560
atgacgcaa ctacctctga tagttgagtc gatacttcgg cgatcacgcg ttccctcata	1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac	1680
atatttgaat gtatttagaa aaataaacia atagctagct cactcggctg ctacgctccg	1740
ggcgtgagac tgcggcgggc gctgcggaca catacaaagt taccacaga ttccgtggat	1800

ES 2 381 433 T3

aagcagggga ctaacatgtg aggcaaaaca gcagggccgc gccggtggcg tttttccata 1860
ggctccgccc tcttgccaga gttcacataa acagacgctt ttccggtgca tctgtgggag 1920
ccgtgaggct caaccatgaa tctgacagta cgggcgaaac ccgacaggac ttaaagatcc 1980
ccaccgtttc cggcgggtcg ctccctcttg cgctctcttg ttccgaccct gccgtttacc 2040
ggatacctgt tccgcctttc tcccttacgg gaagtgtggc gctttctcat agctcacaca 2100
ctggtatctc ggctcgggtgt aggtcgttcg ctccaagctg ggctgtaagc aagaactccc 2160
cggtcagccc gactgctgcg ccttatccgg taactgttca cttgagtcca acccggaaaa 2220
gcacggtaaa acgccactgg cagcagccat tggttaactgg gagttcgcag aggatttggt 2280
tagctaaaca cgcggttgct cttgaagtgt gcgccaaagt ccggctacac tggaaggaca 2340
gatttggttg ctgtgctctg cgaaagccag ttaccacggt taagcagttc cccaactgac 2400
ttaaccttcg atcaaaccac ctccccaggt ggttttttcg ttacagggc aaaagattac 2460
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctactgaac cgctctagct 2520
ttcagtgcaa tttatctctt caaatgtagc acctgaagtc agccccatac gatataagtt 2580
gtaattctca tgttagtcac gccccgcgcc caccggaagg agctgactgg gttgaaggct 2640
ctcaagggca tcggtcgaga tcccgggtgcc taatgagtga gctaacttac attaattgcg 2700
ttgcgctcac tgcccgtttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 2760
ggccaacgcg cgggggagagg cggtttgctt attgggcgcc cgataagctt aattaatctt 2820
tctgcgaatt gagatgacgc cactggctgg gcgtcatccc ggtttcccg gtaaacacca 2880
ccgaaaaata gttactatct tcaaagccac attcggtcga aatatcactg attaacaggc 2940
ggctatgctg gagaagatat tgcgcatgac acactctgac ctgtcgaga tattgattga 3000

ES 2 381 433 T3

tggtcattcc agtctgctgg cgaaattgct gacgcaaac gcgctcactg cacgatgcct	3060
catcacaaaa tttatccagc gcaaagggac ttttcaggct agccgccagc cgggtaatca	3120
gcttatccag caacgtttcg ctggatggtg gcggcaacga atcactggtg taacgatggc	3180
gattcagcaa catcaccaac tgcccgaaca gcaactcagc catttcgtta gcaaacggca	3240
catgctgact actttcatgc tcaagctgac cgataacctg ccgcgcctgc gccatcccca	3300
tgctacctaa gcgccagtgt ggttgccctg cgctggcggtt aaatcccga atcgccccct	3360
gccagtcaag attcagcttc agacgctccg ggcaataaat aatattctgc aaaaccagat	3420
cgtaacgga agcgtaggag tgtttatcgt cagcatgaat gtaaaagaga tcgccacggg	3480
taatgcgata agggcgatcg ttgagtacat gcaggccatt accgcgccag acaatcacca	3540
gctcacaaaa atcatgtgta tgttcagcaa agacatcttg cggataacgg tcagccacag	3600
cgactgcctg ctggtcgctg gcaaaaaaat catctttgag aagttttaac tgatgcgcca	3660
ccgtggctac ctcggccaga gaacgaagtt gattattcgc aatatggcgt acaaatacgt	3720
tgagaagatt cgcgttattg cagaaagcca tcccgtccct ggcgaatc acgcggtgac	3780
cagttaaaact ctcggcgaaa aagcgtcgaa aagtggttac tgcgctgaa tccacagcga	3840
taggcgatgt cagtaacgct ggccctcgctg tggcgtagca gatgtcgggc tttcatcagt	3900
cgcaggcgggt tcaggatatcg ctgaggcgtc agtcccgttt gctgcttaag ctgccgatgt	3960
agcgtacgca gtgaaagaga aaattgatcc gccacggcat cccaattcac ctcatcggca	4020
aaatggtcct ccagccaggc cagaagcaag ttgagacgtg atgcgctggt ttccaggttc	4080
tcctgcaaac tgcttttacg cagcaagagc agtaattgca taaacaagat ctcgcgactg	4140

ES 2 381 433 T3

gcggtcgagg gtaaattcatt ttccccttcc tgctgttcca tctgtgcaac cagctgtcgc 4200

acctgctgca atacgctgtg gttaacgcgc cagtgcgacg gatactgccc atccagctct 4260

tgtggcagca actgattcag cccggcgaga aactgaaatc gatccggcga gcgatacagc 4320

acattgggtca gacacagatt atcgggtatgt tcatacagat gccgatcatg atcgcgtacg 4380

aaacagaccg tgccaccggt gatggtatag ggctgccccat taaacacatg aatacccggtg 4440

ccatgttcga caatcacaat ttcatgaaaa tcatgatgat gttcaggaaa atccgcctgc 4500

gggagccggg gttctatcgc cacggacgcg ttaccagacg gaaaaaaatc cacactatgt 4560

aatacgggtca tactggcctc ctgatgtcgt caacacggcg aaatagtaat cacgaggtca 4620

ggttcttacc ttaaattttc gacggaaaac cacgtaaaaa acgtcgattt ttcaagatac 4680

agcgtgaatt ttcaggaaat gcggtgagca tcacatcacc acaattcagc aaattgtgaa 4740

catcatcacg ttcatctttc cctgggtgcc aatg 4774

<210> 4

<211> 5001

<212> ADN

5 <213> artificial

<220>

<223> pTACo2-Lys

10 <400> 4

gcccattttc ctgtcagtaa cgagaaggct gcgaattcag gcgcttttta gactggctcgt 60

aatgaaattc aggaggtggg cgacatggct cgtgtacagt ttaaacaacg tgaatctact 120

gacgcaatct ttgttactg ctcggtacc aagccaagtc agaattgttg tgtccgtgag 180

attcgccagt ggcacaaaga gcagggttg ctcgatgtgg gataccactt tatcatcaag 240

ES 2 381 433 T3

cgagacggta ctgtggaggc aggacgagat gagatggctg taggctctca cgctaagggt 300
 tacaaccaca actctatcgg cgtctgcctt gttggtggta tcgacgataa aggtaagttc 360
 gacgctaact ttacgccagc ccaaatacaa tcccttcgct cactgcttgt cacactgctg 420
 gctaagtacg aaggcgctgt gcttcgcgcc catcatgagg tggcgccgaa agcttgcctt 480
 tcgttcgacc ttaagcggtg gtgggagaag aacgaactgg tcacttctga ccgtggataa 540
 ggatccccgc gccctcatcc gaaagggcgt attcatatgc cctaggctgc tgccaccgct 600
 gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgctg 660
 aaacctcagg catttgagaa gcacacggtc aactgcttc cggtagtcaa taaaccggta 720
 aaccagcaat agacataagc ggctatttaa cgaccctgcc ctgaaccgac gaccgggtcg 780
 aatttgcttt cgaatttctg ccattcatcc gcttattatc acttattcag gcgtagcacc 840
 aggcgtttta gggcaccaat aactgcctta aaaaaattac gccccgccct gccactcatc 900
 gcagtactgt tgtaattcat taagcattct gccgacatgg aagccatcac aaacggcatg 960
 atgaacctga atcgccagcg gcatcagcac cttgtgcctt tgcgtataat atttgcccat 1020
 agtgaaaacg ggggcgaaga agttgtccat attggccacg tttaaataca aactggtgaa 1080
 actcaccacg ggattggctg agacgaaaaa catattctca ataaaccctt tagggaaata 1140
 ggccagggtt tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa actgccggaa 1200
 atcgtcgtgg tattcactcc agagcgatga aaacgtttca gtttgctcat ggaaaacggc 1260
 gtaacaaggg tgaacactat cccatataac cagctcaccg tctttcattg ccatacggaa 1320
 ctccggatga gcattcatca ggcgggcaag aatgtgaata aaggccgatg aaaacttgctg 1380

ES 2 381 433 T3

cttatTTTTtC tttacGGtCt ttaaaaagGc cGtaatatcc agctgaacGg tctggttata 1440
 ggtacattga gcaactgact gaaatgcctc aaaatgttct ttacgatgcc attgggatat 1500
 atcaacggtg gtatatccag tgatTTTTtt ctccatttta gcttccttag ctCctgaaaa 1560
 tctcgataac tcaaaaaata cgcccggtag tgatcttatt tcattatggt gaaagttgga 1620
 acctcttacg tgccgatcaa cgtctcattt tcgcaaaaag ttggcccagg gcttcccggT 1680
 atcaacaggg acaccaggat ttattttattc tgCGaagtga tcttcCGtca caggatttta 1740
 ttcggcgcaa agtgCGtCGg gtgatgctgc caacttactg atttagtgta tgatggtgtt 1800
 tttgaggtgc tccagtggct tctgtttcta tcagctgtcc ctCctgttca gctactgacg 1860
 gggTggtgCG taacggcaaa agcaccgCGg gacatcagCG ctagCGgagt gtatactggc 1920
 ttactatgtt ggcactgatg agggTgtcag tgaagtGctt catgtggcag gagaaaaaag 1980
 gctgcaccgg tgcgtcagca gaatatgtga tacaggatat attccGcttc ctCGctcact 2040
 gactCGctac gctCGgtCGt tcgactgCGg cGagCGgaaa tggcttacga acggggCGga 2100
 gatttcctgg aagatGCCag gaagatactt aacagggaaG tgagagggcc gCGgcaaagc 2160
 cgtttttcca taggctccgc cccctgaca agcatcacga aatctgacgc tcaaatcagt 2220
 ggtggCGaaa cccgacagga ctataaagat accagggcgtt tccctggCG gctccctCGt 2280
 gcgctctcct gttcctgcct ttCGgtttac cggTgtcatt cCGctgttat ggCGcGttt 2340
 gtctcattcc acgCctgaca ctCagttcCG ggtaggcagt tcGctccaag ctggactgta 2400
 tgcacgaacc cccCGttcag tccgaccGct gCGccttata cggtaactat cgtcttgagt 2460
 ccaaccCGga aagacatgca aaagcaccac tggcagcagc cactggtaat tgatttagag 2520
 gagttagtct tgaagtcatg cGccggttaa ggctaaactg aaaggacaag ttttggtgac 2580

ES 2 381 433 T3

tgcgctcctc caagccagtt acctcggttc aaagagttgg tagctcagag aaccttcgaa	2640
aaaccgccct gcaaggcggg tttttcgttt tcagagcaag agattacgcg cagaccaaaa	2700
cgatctcaag aagatcatct tattaatcag ataaaatatt tctaggtttc agtgcaattt	2760
atctcttcaa atgtagcacc tgaagtcagc ccatacgat ataagttgta attctcatgt	2820
tagtcatgcc ccgcgccac cggaaggagc tgactgggtt gaaggctctc aagggcacgc	2880
gtcgagatcc cgggtgcctaa tgagtgcgtt aacttacatt aattgcgttg cgctcactgc	2940
ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgcg	3000
ggagaggcgcg tttgcgtatt gggcgccga taagcttaat taatctttct gcgaattgag	3060
atgacgccac tggctgggcg tcatcccggt ttcccggtta aacaccaccg aaaaatagtt	3120
actatcttca aagccacatt cggtcgaaat atcactgatt aacaggcggc tatgctggag	3180
aagatattgc gcatgacaca ctctgacctg tcgcagatat tgattgatgg tcattccagt	3240
ctgctggcga aattgctgac gcaaaacgcg ctactgcac gatgcctcat cacaaaattt	3300
atccagcgca aagggaacttt tcaggctagc cgccagccgc gtaatcagct tatccagcaa	3360
cgtttcgctg gatgttggcg gcaacgaatc actggtgtaa cgatggcgat tcagcaacat	3420
caccaactgc ccgaacagca actcagccat ttcgttagca aacggcacat gctgactact	3480
ttcatgctca agctgaccga taacctgccg cgctgcgcc atcccatgc tacctaagcg	3540
ccagtgtggt tgccctgcgc tggcgtaaa tcccggaatc gcccctgcc agtcaagatt	3600
cagcttcaga cgctccgggc aataaataat attctgcaaa accagatcgt taacggaagc	3660
gtaggagtggt ttatcgtcag catgaatgta aaagagatcg ccacgggtaa tgcgataagg	3720

ES 2 381 433 T3

gcgatcggtg agtacatgca ggccattacc gcgccagaca atcaccagct cacaaaaatc 3780

atgtgtatgt tcagcaaaga catcttgcgg ataacgggtca gccacagcga ctgcctgctg 3840

gtcgctggca aaaaaatcat ctttgagaag ttttaactga tgcgccaccg tggctacctc 3900

ggccagagaa cgaagttgat tattcgcaat atggcgtaga aatacgttga gaagattcgc 3960

gttattgcag aaagccatcc cgtccctggc gaatatcacg cggtgaccag ttaaactctc 4020

ggcgaaaaag cgtcgaaaag tggttactgt cgctgaatcc acagcgatag gcgatgtcag 4080

taacgctggc ctcgctgtgg cgtagcagat gtcgggcttt catcagtcgc aggcggttca 4140

ggtatcgctg aggcgtcagt cccgtttgct gcttaagctg ccgatgtagc gtacgcagtg 4200

aaagagaaaa ttgatccgcc acggcatccc aattcacctc atcggaataa tggcctcca 4260

gccaggccag aagcaagttg agacgtgatg cgctgttttc caggttctcc tgcaaaactgc 4320

ttttacgcag caagagcagt aattgcataa acaagatctc gcgactggcg gtcgagggtg 4380

aatcattttc cccttctgc tgttccatct gtgcaaccag ctgtcgcacc tgctgcaata 4440

cgctgtgggt aacgcgccag tgagacggat actgcccatac cagctcttgt ggcagcaact 4500

gattcagccc ggcgagaaac tgaaatcgat ccggcgagcg atacagcaca ttggtcagac 4560

acagattatc ggtatgttca tacagatgcc gatcatgac gcgtacgaaa cagaccgtgc 4620

caccggtgat ggtatagggc tgcccattaa acacatgaat acccgtgcca tgttcgacaa 4680

tcacaatttc atgaaaatca tgatgatgtt caggaaaatc cgctgcggg agccgggggt 4740

ctatcgccac ggacgcgtta ccagacggaa aaaaatccac actatgtaat acggtcatac 4800

tggcctcctg atgtcgtcaa caggcgaaa tagtaatcac gaggtcaggt tcttacctta 4860

aattttcgac ggaaaaccac gtaaaaaacg tcgatttttc aagatacagc gtgaattttc 4920

ES 2 381 433 T3

aggaaatgcg gtgagcatca catcaccaca attcagcaaa ttgtgaacat catcacgttc 4980

atctttccct gggtgccaat g 5001

<210> 5

<211> 6358

<212> ADN

5 <213> artificial

<220>

<223> pTACO3-LysY

10 <400> 5

gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactggtcgt 60

aatgaaattc aggaggttgt cgacatggct cgtgtacagt ttaaacaacg tgaatctact 120

gacgcaatct ttgttactg ctcggtacc aagccaagtc agaatgttg tgtccgtgag 180

attcgccagt ggcacaaaga gcagggttg ctcgatgtg gataccactt tatcatcaag 240

cgagacggta ctgtggaggc aggacgagat gagatggctg taggctctca cgctaagggt 300

tacaaccaca actctatcgg cgtctgcctt gttggtggta tcgacgataa aggtaagttc 360

gacgctaact ttacgccagc ccaaagcaa tcccttcgct cactgcttgt cacactgctg 420

gctaagtacg aaggcgctgt gcttcgcgcc catcatgagg tggcgccgta cgcttgccct 480

tcgttcgacc ttaagcggtg gtgggagaag aacgaactgg tcacttctga ccgtggataa 540

ggatccccgc gccctcatcc gaaagggcgt attcatatgc cctagggaca gtaagacggg 600

taagcctggt gatgataccg ctgccttact ggggtgatta gccagtctga atgacctgtc 660

acgggataat ccgaagtggc cagactggaa aatcagaggg caggaaactgc tgaacagcaa 720

ES 2 381 433 T3

aaagtcagat agcaccacat agcagacccg ccataaaacg ccctgagaag cccgtgacgg	780
gcttttcttg tattatgggt agtttccttg catgaatcca taaaaggcgc ctgtagtgcc	840
atttaccccc attcactgcc agagccgtga gcgcagcgaa ctgaatgtca cgaaaaagac	900
agcgactcag gtgcctgatg gtcggagaca aaaggaatat tcagcgattt gcccgagctt	960
gcgaggggtgc tacttaagcc tttagggttt taaggctctgt tttgtagagg agcaaacagc	1020
gtttgcgaca tccttttgta atactgcgga actgactaaa gtagtgagtt atacacaggg	1080
ctgggatcta ttctttttat ctttttttat tctttcttta ttctataaat tataaccact	1140
tgaatataaa caaaaaaac acacaaaggt ctagcggaat ttacagaggg tctagcagaa	1200
tttacaagtt ttccagcaaa ggtctagcag aatttacaga taccacaac tcaaaggaaa	1260
aggactagta attatcattg actagcccat ctcaattggg atagtgatta aaatcaccta	1320
gaccaattga gatgtatgtc tgaattagtt gttttcaaag caaatgaact agcgattagt	1380
cgctatgact taacggagca tgaaccaag ctaattttat gctgtgtggc actactcaac	1440
cccacgattg aaaaccctac aaggaaagaa cggacggtat cgttcactta taaccaatac	1500
gctcagatga tgaacatcag tagggaaaat gcttatgggtg tattagctaa agcaaccaga	1560
gagctgatga cgagaactgt ggaaatcagg aatccttttg tttaaaggctt tgagattttc	1620
cagtggacaa actatgccaa gttctcaagc gaaaaattag aattagtttt tagtgaagag	1680
atattgcctt atcttttcca gttaaaaaaa ttcataaaat ataactctgga acatgttaag	1740
tcttttgaaa acaataactc tatgaggatt tatgagtggg tattaaaaga actaacacaa	1800
aagaaaactc acaaggcaaa tatagagatt agccttgatg aatttaagtt catgttaatg	1860
cttgaaaata actaccatga gtttaaaaagg ctttaaccaat gggttttgaa accaataagt	1920

ES 2 381 433 T3

aaagatttaa acacttacag caatatgaaa ttggtggttg ataagcgagg ccgcccgact	1980
gatacgttga ttttccaagt tgaactagat agacaaatgg atctcgtaac cgaacttgag	2040
aacaaccaga taaaaatgaa tggtgacaaa ataccaacaa ccattacatc agatttcctac	2100
ctacataacg gactaagaaa aacactacac gatgctttta ctgcaaaaat tcagctcacc	2160
agttttgagg caaaatTTTT gagtgacatg caaagtaagt atgatctcaa tggttcgttc	2220
tcaTggtca cgcaaaaaca acgaaccaca ctagagaaca tactggctaa atacggaagg	2280
atctgagggtt cttatggctc ttgtatctat cagtgaagca tcaagactaa caaacaaaag	2340
tagaacaact gttcacCGtt acatatcaaa gggaaaactg tccatatgca cagatgaaaa	2400
cggtgtaaaa aagatagata catcagagct tttacgagtt tttggtgcat tcaaagctgt	2460
tcaccatgaa cagatcgaca atgtaacaga tgaacagcat gtaacaccta atagaacagg	2520
tgaaaccagt aaaacaaagc aactagaaca tgaaattgaa cacctgagac aacttgttac	2580
agctcaacag tcacacatag acagcctgaa acaggcgatg ctgcttatcg aatcaaagct	2640
gccgacaaca cgggagccag tgacgcctcc cgtggggaaa aaatcatggc aattctggaa	2700
gaaatagcgc tttcagccgg caaaccggct gaagccggat ctgcgattct gataacaaac	2760
tagcaacacc agaacagccc gtttgCGggc agcaaaaccc gtgggaatta attcccctgc	2820
tcgcgcaggc tgggtgcaa gctctcgggt aacatcaagg cccgatcctt ggagcccttg	2880
ccctcccga cgatgatcgt gccgtgatcg aaatccagat ccttgaccCG cagttgcaaa	2940
ccctcactga tccgcatgcc cgttcatac agaagctggg cgaacaaacg atgctcgctt	3000
tccagaaaac cgaggatgcg aaccacttca tccggggTca gcaccaccgg caagcgccgc	3060

ES 2 381 433 T3

gacggccgag gtcttccgat ctctgaagc cagggcagat ccgtgcacag caccttgccg 3120
 tagaagaaca gcaaggccgc caatgcctga cgatgcgtgg agaccgaaac cttgcgctcg 3180
 ttcgccagcc aggacagaaa tgcctcgact tcgctgctgc ccaaggttgc cgggtgacgc 3240
 acaccgtgga aacggatgaa ggcacgaacc cagtggacat aagcctgttc ggttcgtaag 3300
 ctgtaatgca agtagcgtat gcgctcacgc aactggcca gaaccttgac cgaacgcagc 3360
 ggtggtaacg gcgcagtggc ggttttcatg gcttggtatg actgtttttt tggggtacag 3420
 tctatgcctc gggcatccaa gcagcaagcg cgttacgccg tgggtcgatg tttgatgtta 3480
 tggagcagca acgatgttac gcagcagggc agtcgcccta aaacaaagtt aaacatcatg 3540
 agggaagcgg tgatcgccga agtatcgact caactatcag aggtagttgg cgtcatcgag 3600
 cgccatctcg aaccgacgtt gctggccgta catttgtagc gtcgccagc ggatggcggc 3660
 ctgaagccac acagtgatat tgatttgctg gttacggtga ccgtaaggct tgatgaaaca 3720
 acgcggcgag ctttgatcaa cgaccttttg gaaacttcgg cttcccctgg agagagcgag 3780
 attctccgcg ctgtagaagt caccattgtt gtgcacgacg acatcattcc gtggcggttat 3840
 ccagctaagc gcgaactgca atttgagaaa tggcagcgca atgacattct tgcaggtatc 3900
 ttcgagccag ccacgatcga cattgatctg gctatcttgc tgacaaaagc aagagaacat 3960
 agcgttgcc tggtaggtcc agcggcggag gaactctttg atccggttcc tgaacaggat 4020
 ctatttgagg cgctaaatga aaccttaacg ctatggaact cgccgccga ctgggctggc 4080
 gatgagcgaa atgtagtgct tacgttgctc cgcatcttgt acagcgagc aaccggcaaa 4140
 atcgcgccga aggatgtcgc tgccgactgg gcaatggagc gcctgccggc ccagtatcag 4200
 cccgtcatac ttgaagctag acaggcttat cttggacaag aagaagatcg cttggcctcg 4260

ES 2 381 433 T3

cgcgagatc agttggaaga atttgtccac tacgtgaaag gcgagatcac caaggtagtc 4320
 ggcaaataat gtctaacaat tcgttcaagc cgacgccgct tcgcggcgcg gcttaactca 4380
 agcgtagaa gcttaattaa tctttctgcg aattgagatg acgccactgg ctgggcgtca 4440
 tcccggtttc cggggtaaac accaccgaaa aatagttact atcttcaaag ccacattcgg 4500
 tcgaaatata actgattaac aggcggctat gctggagaag atattgcgca tgacacactc 4560
 tgacctgtcg cagatattga ttgatggta ttccagtctg ctggcgaaat tgctgacgca 4620
 aaacgcgctc actgcacgat gcctcatcac aaaatttatc cagcgcaaag ggacttttca 4680
 ggctagccgc cagccgggta atcagcttat ccagcaacgt ttcgctggat gttggcggca 4740
 acgaatcact ggtgtaacga tggcgattca gcaacatcac caactgcccg aacagcaact 4800
 cagccatttc gttagcaaac ggcacatgct gactactttc atgctcaagc tgaccgataa 4860
 cctgccgcgc ctgcgccatc cccatgctac ctaagcgcca gtgtggttgc cctgcgctgg 4920
 cgttaaatcc cggaatcgcc ccctgccagt caagattcag cttcagacgc tccgggcaat 4980
 aaataatatt ctgcaaaacc agatcgtaa cggaagcgta ggagtgttta tcgtcagcat 5040
 gaatgtaaaa gagatcgcca cgggtaatgc gataagggcg atcgttgagt acatgcaggc 5100
 cattaccgcg ccagacaatc accagctcac aaaaatcatg tgtatgttca gcaaagacat 5160
 cttgcggata acggtcagcc acagcgactg cctgctggtc gctggcaaaa aaatcatctt 5220
 tgagaagttt taactgatgc gccaccgtgg ctacctcggc cagagaacga agttgattat 5280
 tcgcaatatg gcgtacaaat acgttgagaa gattcgcgtt attgcagaaa gccatcccg 5340
 ccctggcgaa tatcacgcgg tgaccagtta aactctcggc gaaaaagcgt cgaaaagtgg 5400

ES 2 381 433 T3

ttactgtcgc tgaatccaca gcgataggcg atgtcagtaa cgctggcctc gctgtggcgt 5460

agcagatgtc gggctttcat cagtcgcagg cggttcaggt atcgctgagg cgtcagtccc 5520

gtttgctgct taagctgccg atgtagcgta cgcagtgaaa gagaaaattg atccgccacg 5580

gcatcccaat tcacctcatc ggcaaaatgg tcctccagcc aggccagaag caagttgaga 5640

cgtgatgcgc tgttttccag gttctcctgc aaactgcttt tacgcagcaa gagcagtaat 5700

tgcataaaca agatctcgcg actggcggtc gagggtaaata cattttcccc ttctgtgtgt 5760

tccatctgtg caaccagctg tcgcacctgc tgcaatacgc tgtggttaac gcgccagtga 5820

gacggatact gcccatccag ctcttggtggc agcaactgat tcagcccggc gagaaactga 5880

aatcgatccg gcgagcgata cagcacattg gtcagacaca gattatcggg atgttcatac 5940

agatgccgat catgatcgcg tacgaaacag accgtgccac cggtgatggg atagggctgc 6000

ccattaaaca catgaatacc cgtgccatgt tcgacaatca caatttcatg aaaatcatga 6060

tgatgttcag gaaaatccgc ctgcgggagc cggggttcta tcgccacgga cgcgttacca 6120

gacggaaaaa aatccacact atgtaatacg gtcatactgg cctcctgatg tcgtcaacac 6180

ggcgaaatag taatcacgag gtcaggttct taccttaaata ttctgacgga aaaccacgta 6240

aaaaacgtcg atttttcaag atacagcgtg aattttcagg aaatgcggtg agcatcacat 6300

caccacaatt cagcaaattg tgaacatcat cacgttcac tttccctggg tgccaatg 6358

<210> 6

<211> 72

<212> ADN

5 <213> artificial

<220>

<223> Promotor rhaBAD

10 <400> 6

ES 2 381 433 T3

gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactggtcgt 60

aatgaaattc ag 72

<210> 7

<211> 38

<212> ADN

5 <213> artificial

<220>

<223> pRha_EheI_f

10 <400> 7

gcgcgcggcg cccgataagc ttaattaatc ttctgcg 38

<210> 8

<211> 37

15 <212> ADN

<213> artificial

<220>

<223> pRha_XmaJI_r

20 <400> 8

gcgcgccta gggcatatga atacgccctt tcggatg 37

<210> 9

25 <211> 32

<212> ADN

<213> artificial

<220>

30 <223> T7Lys_Sall_f

<400> 9

gcgcgcgtcg acatggctcg tgtacagtt aa 32

35 <210> 10

<211> 32

<212> ADN

<213> artificial

5 <220>

<223> T7Lys_BamHI_r

<400> 10

cgcgcgggat ccttatccac ggtcagaagt ga 32

10

<210> 11

<211> 32

<212> ADN

<213> artificial

15

<220>

<223> pCL1920_XmaJI_f

<400> 11

20 cgcgcgcccta gggacagtaa gacgggtaag cc 32

<210> 12

<211> 32

<212> ADN

25 <213> artificial

<220>

<223> pCL1920_HindIII_r

30 <400> 12

gcgcgcaagc ttctaacgct tgagttaagc cg 32

<210> 13

<211> 36

35 <212> ADN

<213> artificial

ES 2 381 433 T3

<220>

<223> Promotor Tet

<400> 13

5 ttgacactct atcattgata gagttatatt accact 36

<210> 14

<211> 81

<212> ADN

10 <213> artificial

<220>

<223> Promotor araBAD

15 <400> 14

ccataagatt agcggatcct acctgacgct ttttatcgca actctctact gtttctccat 60

acccgttttt ttggatggag t 81

<210> 15

<211> 81

<212> ADN

20 <213> artificial

<220>

<223> rhaBAD

25 <400> 15

gcccattttc ctgtcagtaa cgagaaggct gcgaattcag gcgcttttta gactggtcgt 60

aatgaaattc agcaggatca c 81

<210> 16

30 <211> 456

<212> ADN

<213> Bacteriófago T7

<400> 16

ES 2 381 433 T3

atggctcgtg tacagtttaa acaacgtgaa tctactgacg caatctttgt tcaactgctcg	60
gctaccaagc caagtcagaa tgttggtgtc cgtgagattc gccagtggca caaagagcag	120
ggttggctcg atgtgggata ccactttatc atcaagcgag acggtactgt ggaggcagga	180
cgagatgaga tggctgtagg ctctcacgct aagggttaca accacaactc tatcggcgtc	240
tgccttgttg gtggtatcga cgataaaggc aagttcgacg ctaactttac gccagcccaa	300
atgcaatccc ttcgctcact gcttgtcaca ctgctggcta agtacgaagg cgctgtgctt	360
cgcgcccatc atgaggtggc gccgaaagct tgcccttcgt tcgaccttaa gcgttggtgg	420
gagaagaacg aactggtcac ttctgaccgt ggataa	456

<210> 17

<211> 2820

5 <212> ADN

<213> Bacteriófago T7

<400> 17

tcgcgctgca ctggcgtaat gctgaccgga tggctatcgc taatggtctt acgctcaaca	60
ttgataagca acttgacgca atgttaatgg gctgatagtc ttatcttaca ggtcatctgc	120
gggtggcctg aataggtacg atttactaac tggaagaggc actaaatgaa cacgattaac	180
atcgctaaga acgacttctc tgacatcgaa ctggctgcta tcccggtcaa cactctggct	240

ES 2 381 433 T3

gaccattacg gtgagcgttt agctcgcgaa cagttggccc ttgagcatga gtcttacgag 300

atgggtgaag cacgcttccg caagatgttt gagcgtaac ttaaagctgg tgaggttgcg 360

gataacgctg ccgccaagcc tctcatcact accctactcc ctaagatgat tgcacgcac 420

aacgactggt ttgaggaagt gaaagctaag cgcggcaagc gcccgcacgc cttccagttc 480

ctgcaagaaa tcaagccgga agccgtagcg tacatcacca ttaagaccac tctggcttgc 540

ctaaccagtg ctgacaatac aaccgttcag gctgtagcaa ggcgaatcgg tcggggccatt 600

gaggacgagg ctgccttcgg tcgtatccgt gacctgaag ctaagcactt caagaaaaac 660

gttgaggaac aactcaaca ggcgtaggg cacgtctaca agaaagcatt tatgcaagtt 720

gtcgaggctg acatgctctc taagggtcta ctcggtggcg aggcgtggtc ttcgtggcat 780

aaggaagact ctattcatgt aggagtacgc tgcacgaga tgctcattga gtcaaccgga 840

atggttagct tacaccgcca aaatgctggc gtagtaggtc aagactctga gactatcgaa 900

ctcgcacctg aatacgtga ggctatcgca acccgtgcag gtgcgctggc tggcatctct 960

ccgatgttcc aacctgctg agttcctcct aagccgtgga ctggcattac tgggtggtggc 1020

tattgggcta acggtcgctg tctctggcg ctggtgcgta ctcacagtaa gaaagcactg 1080

atgcgctacg aagacgttta catgcctgag gtgtacaaag cgattaacat tgcgcaaaac 1140

accgcatgga aaatcaaca gaaagtccta gcggtcgcca acgtaatcac caagtggaag 1200

cattgtccgg tcgaggacat ccctgcgatt gagcgtgaag aactcccgat gaaaccggaa 1260

gacatcgaca tgaatcctga ggctctcacc gcgtggaaac gtgctgccgc tgctgtgtac 1320

cgcaaggaca gggctcgcaa gtctcgccgt atcagccttg agttcatgct tgagcaagcc 1380

ES 2 381 433 T3

aataagtttg ctaaccataa ggccatctgg ttccttaca acatggactg gcgcggtcgt 1440
gtttacgccg tgtcaatggt caaccgcaa ggtaacgata tgaccaaagg actgcttacg 1500
ctggcgaaag gtaaaccaat cggttaaggaa ggttactact ggctgaaaat ccacggtgca 1560
aactgtgcgg gtgtcgataa ggttccgttc cctgagcgca tcaagttcat tgaggaaaac 1620
cacgagaaca tcatggcttg cgctaagtct cactggaga acacttggtg ggctgagcaa 1680
gattctccgt tctgcttctt tgcgttctgc tttgagtacg ctggggtaca gcaccacggc 1740
ctgagctata actgctccct tccgctggcg tttgacgggt cttgctctgg catccagcac 1800
ttctccgcga tgctccgaga tgaggtaggt ggtcgcgcgg ttaacttgct tcctagttag 1860
accgttcagg acatctacgg gattgttgct aagaaagtca acgagattct acaagcagac 1920
gcaatcaatg ggaccgataa cgaagtagtt accgtgaccg atgagaacac tggtgaaatc 1980
tctgagaaag tcaagctggg cactaaggca ctggctggtc aatggctggc tcacggtggt 2040
actgcagtg tgactaagcg ttcagtcatg acgctggctt acgggtccaa agagttcggc 2100
ttccgtcaac aagtgtgga agataccatt cagccagcta ttgattccgg caaggggccg 2160
atgttcactc agccgaatca ggctgctgga tacatggcta agctgatttg ggaatctgtg 2220
agcgtgacgg tggtagctgc ggttgaagca atgaactggc ttaagtctgc tgctaagctg 2280
ctggctgctg aggtcaaaga taagaagact ggagagattc ttcgcaagcg ttgcgctgtg 2340
cattgggtaa ctctgatgg tttccctgtg tggcaggaat acaagaagcc tattcagacg 2400
cgcttgaacc tgatgttctt cggtcagttc cgcttacagc ctaccattaa caccaacaaa 2460
gatagcgaga ttgatgcaca caaacaggag tctggtatcg ctcttaactt tgtacacagc 2520
caagacggta gccaccttcg taagactgta gtgtgggcac acgagaagta cggaatcgaa 2580

ES 2 381 433 T3

tcttttgcac	tgattcacga	ctccttcggt	accattccgg	ctgacgctgc	gaacctgttc	2640
aaagcagtgc	gcgaaactat	ggttgacaca	tatgagtctt	gtgatgtact	ggctgatttc	2700
tacgaccagt	tcgctgacca	gttgacagag	tctcaattgg	acaaaatgcc	agcacttccg	2760
gctaaaggta	acttgaacct	ccgtgacatc	ttagagtcgg	acttcgcggt	cgcgtaacgc	2820

REIVINDICACIONES

1. Una célula huésped capaz de expresar ARN polimerasa de T7 tras inducción, comprendiendo la célula huésped:
un primer ácido nucleico que tiene un gen de la lisozima de T7 o un gen variante de la lisozima de T7 que codifica una variante de la lisozima de T7 que tiene la capacidad de inhibir la actividad de la ARN polimerasa de T7, bajo el control de un promotor, y
un segundo ácido nucleico que tiene un promotor de T7 unido operablemente a una secuencia de ácido nucleico que codifica un polipéptido diana,
que se caracteriza porque
el gen de la lisozima de T7 o el gen variante de la lisozima de T7 está bajo el control de un promotor ajustable;
- el promotor ajustable tiene un efecto sobre el índice de transcripción del gen de la lisozima de T7 que depende de la concentración o intensidad de un inductor del promotor, de forma que el índice de transcripción del gen de la lisozima de T7 aumenta con la concentración o intensidad creciente del inductor y que la expresión de la lisozima de T7 se puede ajustar de forma continua mediante dicho promotor ajustable;
- y porque
- el promotor ajustable se selecciona de un modo tal que la expresión del polipéptido diana se puede ajustar mediante el control de la expresión del gen de la lisozima de T7 o el gen variante de la lisozima de T7 cuando se induce la expresión de la ARN polimerasa de T7.
2. La célula huésped de la reivindicación 1, en la que dicho ácido nucleico comprende el promotor ajustable es un primer vector y dicho segundo ácido nucleico que comprende el promotor de T7 es un segundo vector y en el que el primero y el segundo vector son compatibles.
3. La célula huésped de cualquiera de las reivindicaciones 1 o 2, en la que dicho promotor ajustable es ajustable por ramnosa o arabinosa.
4. La célula huésped de cualquiera de las reivindicaciones 1 o 2, en la que dicho promotor ajustable es ajustable por luz.
5. La célula huésped de cualquiera de las reivindicaciones 1 o 2, en la que dicho promotor ajustable es ajustable por temperatura.
6. La célula huésped de cualquiera de las reivindicaciones 1 a 5, en la que dicho primer ácido nucleico comprende un marcador de selección.
7. La célula huésped de la reivindicación 6, en la que el marcador de selección es un marcador de selección antibiótico.
8. La célula huésped de cualquiera de las reivindicaciones 1 a 7, en la que dicha variante de la lisozima de T7 es LysY.
9. La célula huésped de cualquiera de las reivindicaciones 1 a 8, en la que dicho primer ácido nucleico es al menos un 80 % idéntico a un ácido nucleico escogido del grupo que comprende la SEC ID N° 1-5.
10. La célula huésped de la reivindicación 9, en la que dicho primer ácido nucleico es idéntico a un ácido nucleico escogido del grupo que comprende la SEC ID N° 1-5.
11. Una célula huésped de las reivindicaciones 1 a 10, seleccionada del grupo de *E. coli*, *Pseudomonas aeruginosa*, *Erwinia carotovora*, *Salmonella choleraesuis*, *Agrobacterium tumefaciens*, *Chromobacterium violaceum*, *Lactococcus lactis*, *Bacillus subtilis*, *Salmonella*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, CHO, NS0, HEK293, HeLa, Sf9, tabaco, arroz y *Leishmania tarentolae*.
12. Un procedimiento para producir un polipéptido diana, que comprende las etapas de:
 - a) proporcionar una célula huésped de acuerdo con una cualquiera de las reivindicaciones 1 a 11;
 - b) inducir la expresión del polipéptido diana
 - c) controlar la expresión del polipéptido diana ajustando dicho promotor ajustable y, por tanto, la expresión de dicha lisozima de T7;
 - y opcionalmente
 - d) aislar el polipéptido.

13. El procedimiento de la reivindicación 12, en el que la inducción de la expresión del polipéptido diana se realiza induciendo la expresión de la ARN polimerasa de T7 o una variante de la misma.

5 14. El procedimiento de cualquiera de las reivindicaciones 12 a 13, en el que el ajuste del promotor ajustable, o una variante del mismo, se consigue mediante cualquiera de: inducción ligera, ajuste de la temperatura, adición de un inductor químico, en concreto ramnosa o arabinosa.

15. Un ácido nucleico que comprende un gen de la lisozima de T7 o un gen variante de la lisozima de T7 que codifica una variante de la lisozima de T7 que tiene la capacidad de inhibir la actividad de la ARN polimerasa de T7, bajo el control de un promotor,

que se caracteriza porque

10 el gen de la lisozima de T7 o el gen variante de la lisozima de T7 está bajo el control de un promotor rhaBAD ajustable y que el promotor ajustable tiene un efecto sobre el índice de transcripción del gen de la lisozima de T7 que depende de la concentración de un inductor del promotor rhaBAD, de forma que el índice de transcripción del gen de la lisozima de T7 aumenta con la concentración creciente del inductor y que la expresión de la lisozima de T7 se puede ajustar de forma continua mediante dicho promotor rhaBAD ajustable.

15 16. Un ácido nucleico de acuerdo con la reivindicación 15, que es al menos un 80 % idéntico con un ácido nucleico escogido del grupo de las SEC ID N° 1-5.

17. Un ácido nucleico de acuerdo con la reivindicación 16 escogido del grupo de las SEC ID N° 1-5.

18. Uso del ácido nucleico de cualquiera de las reivindicaciones 15 a 17 para controlar la expresión de un polipéptido diana mediante la lisozima de T7.

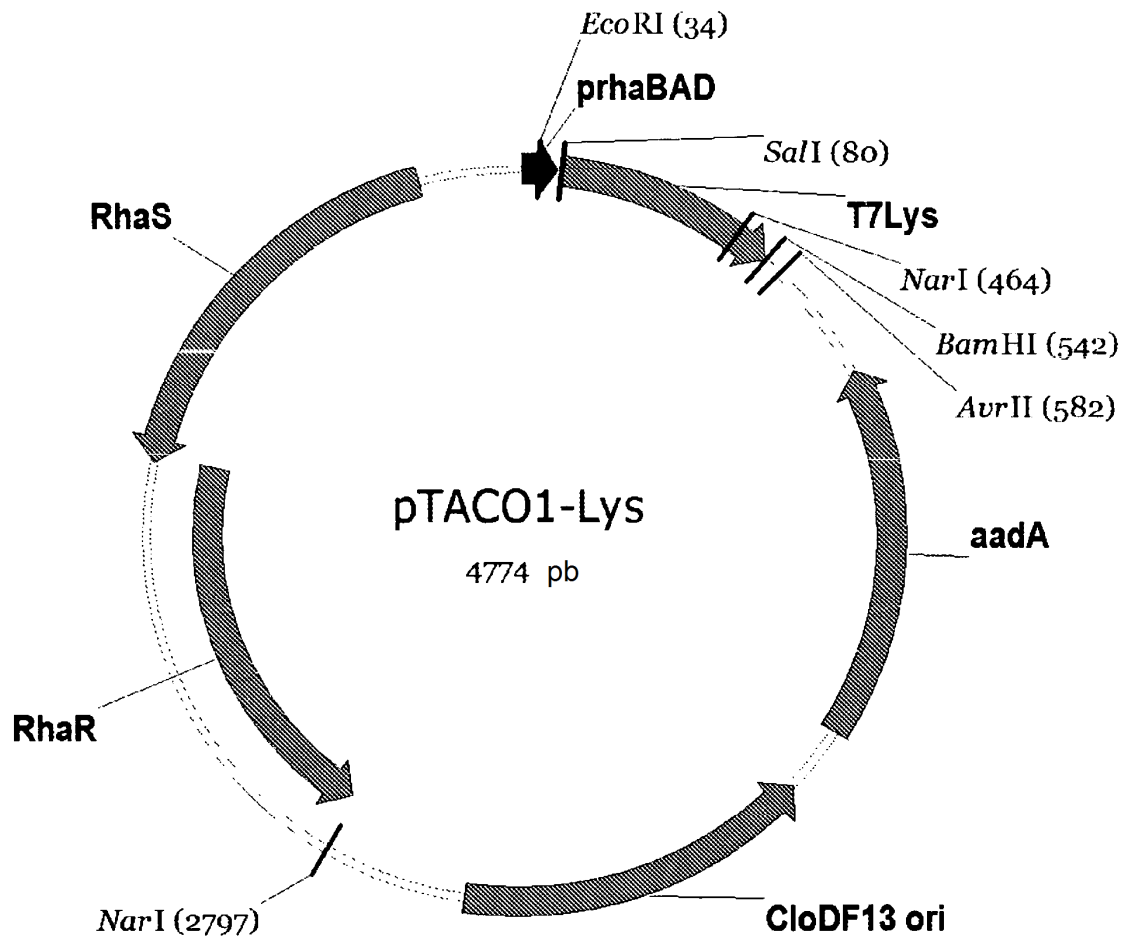


Figura 1 A

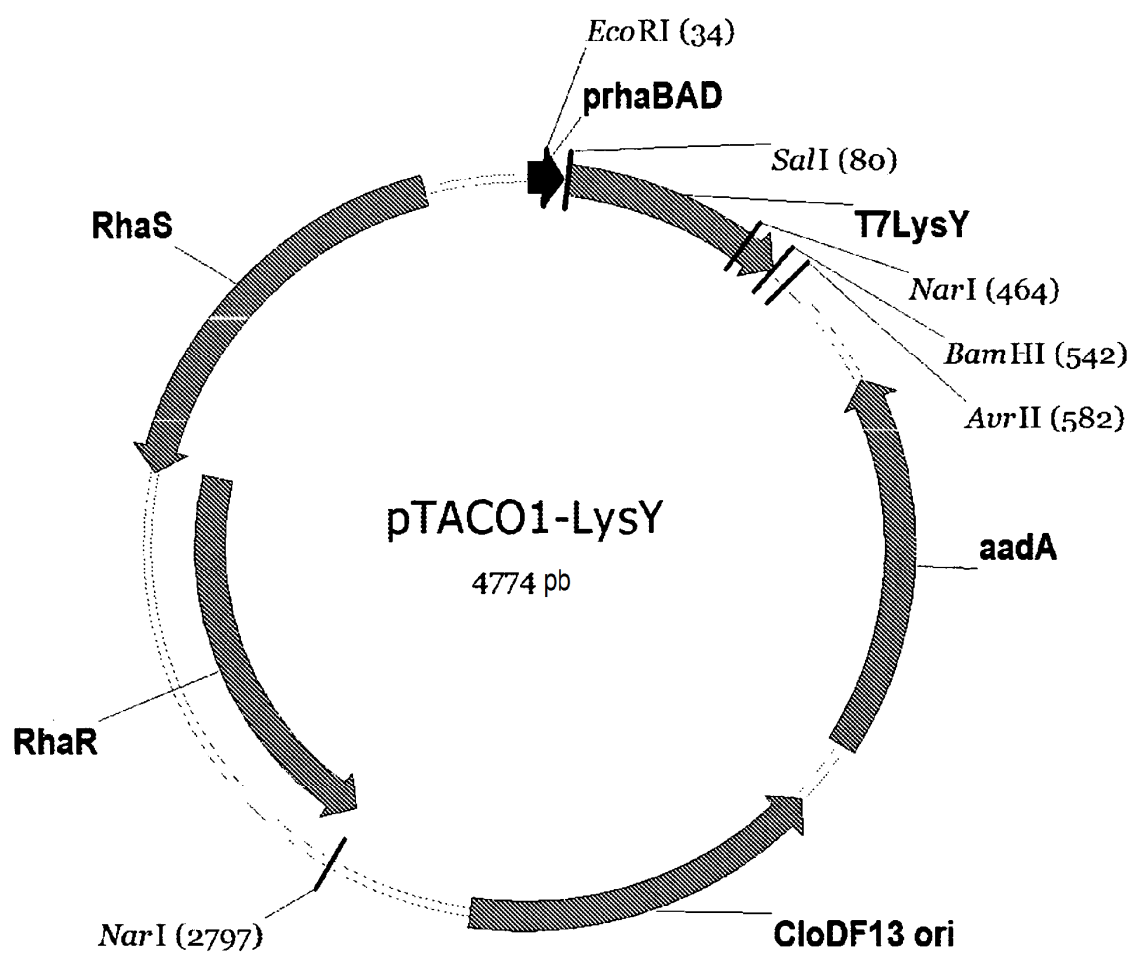


Figura 1 B

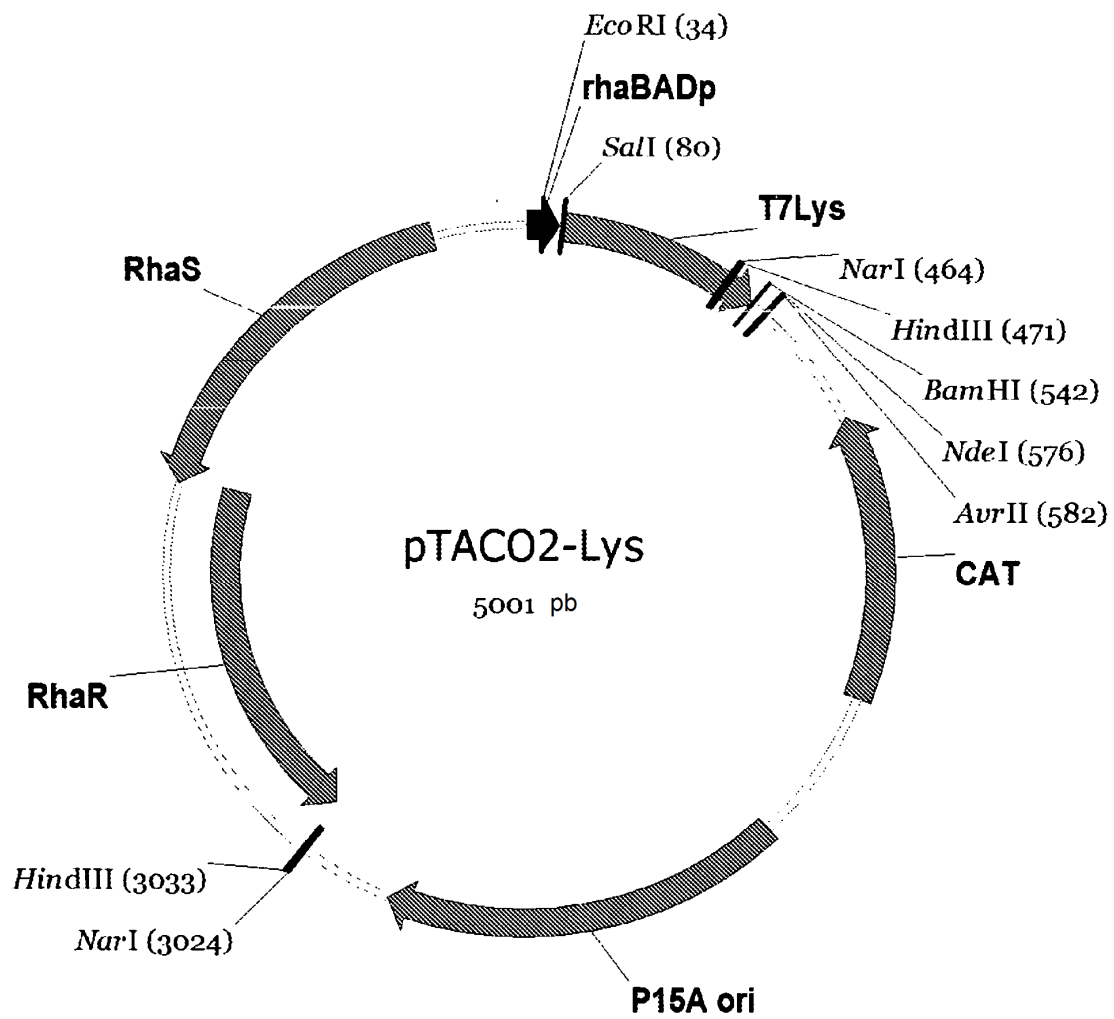


Figura 1 C

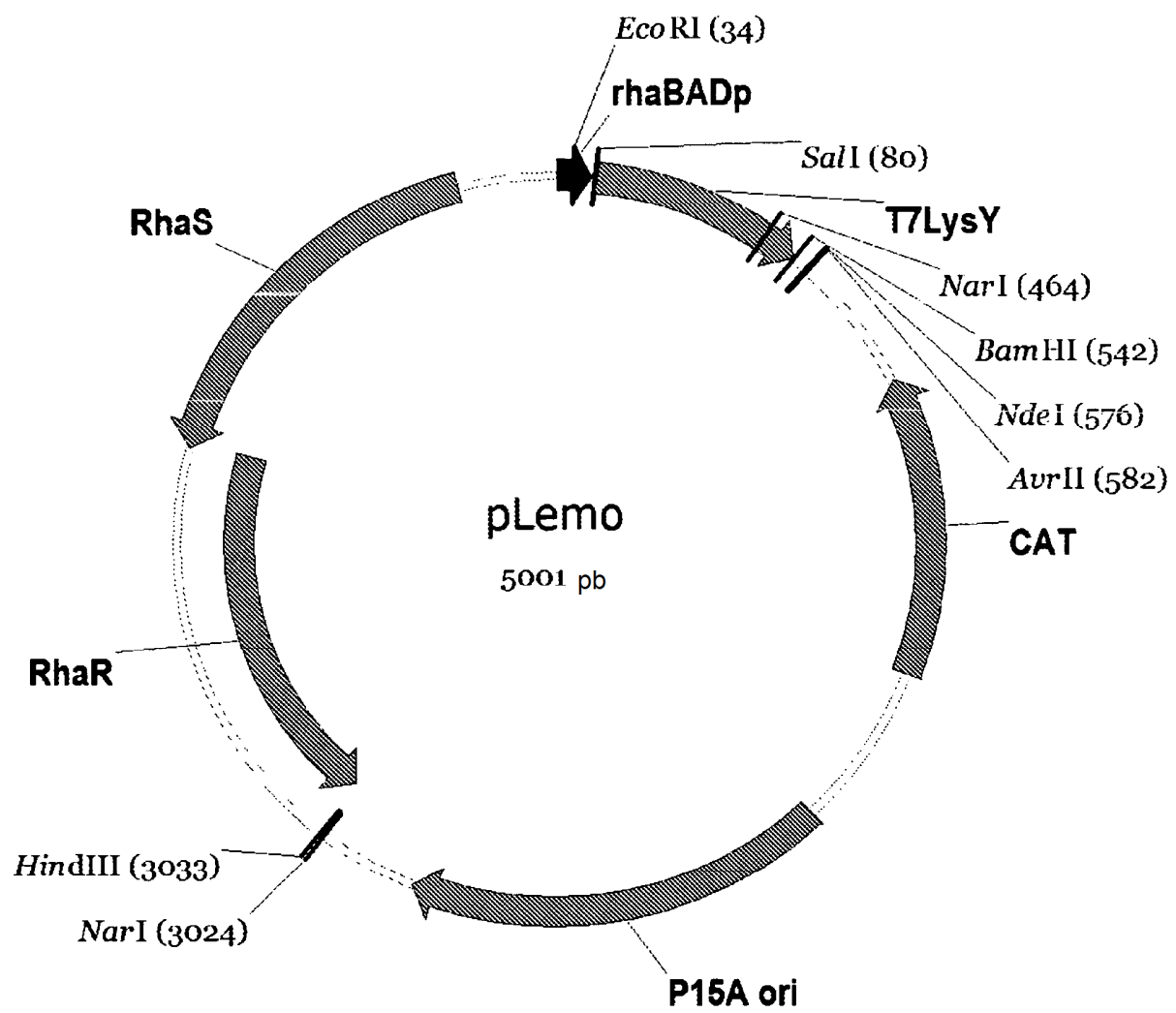


Figura 1 D

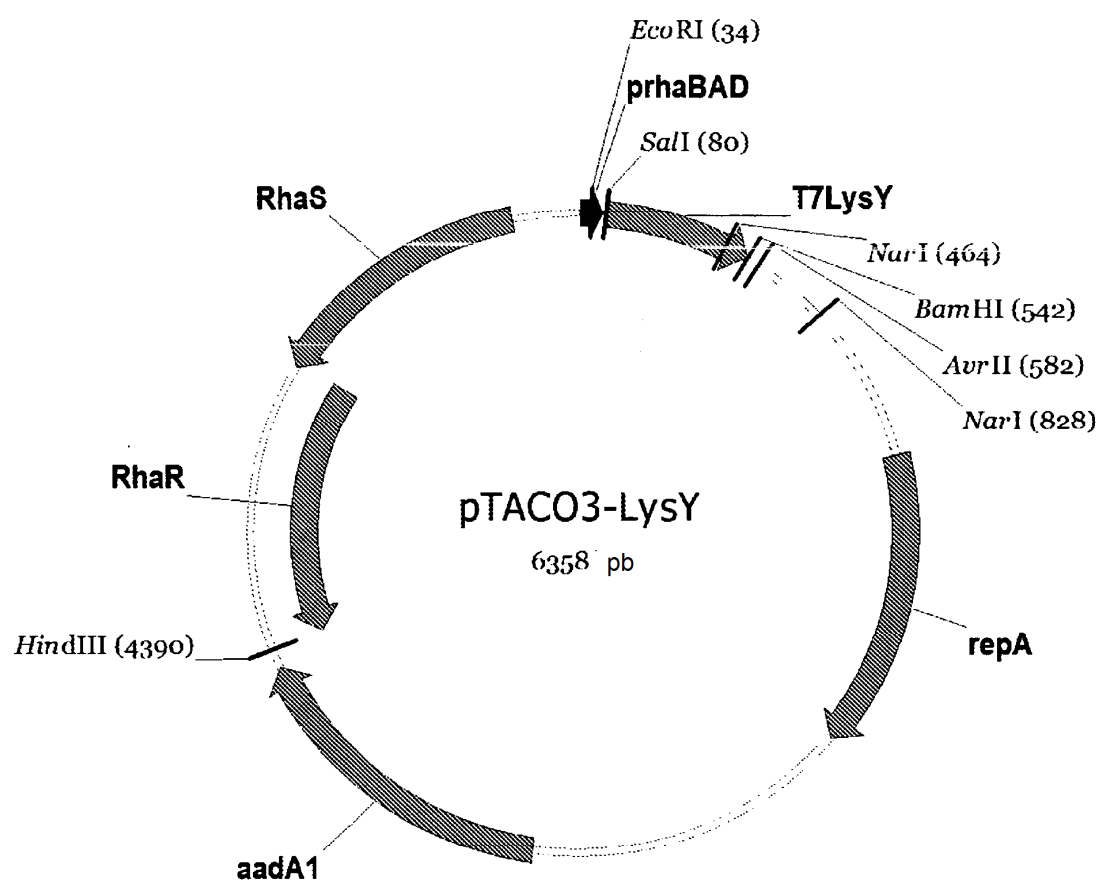


Figura 1 E

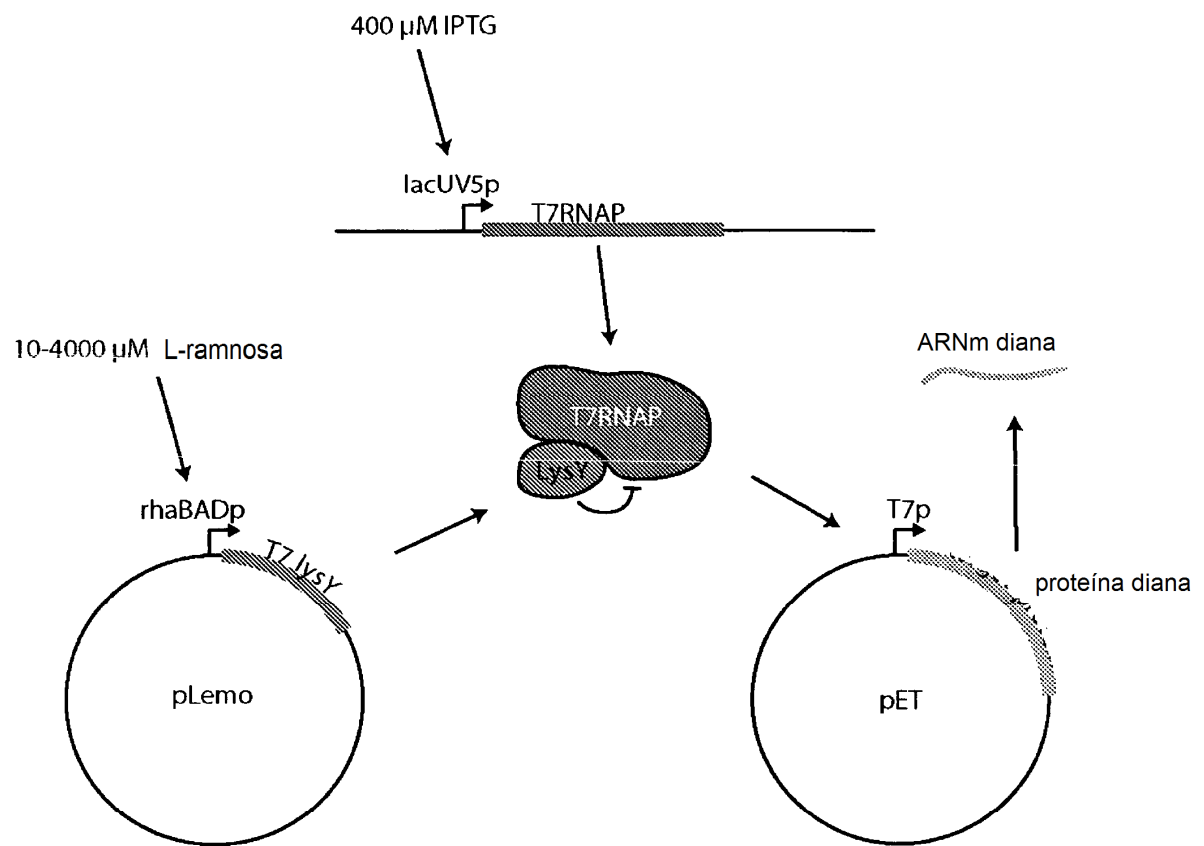


Figura 2

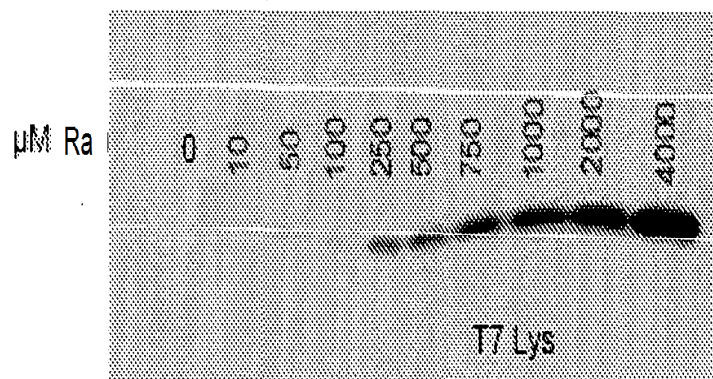


Figura 3

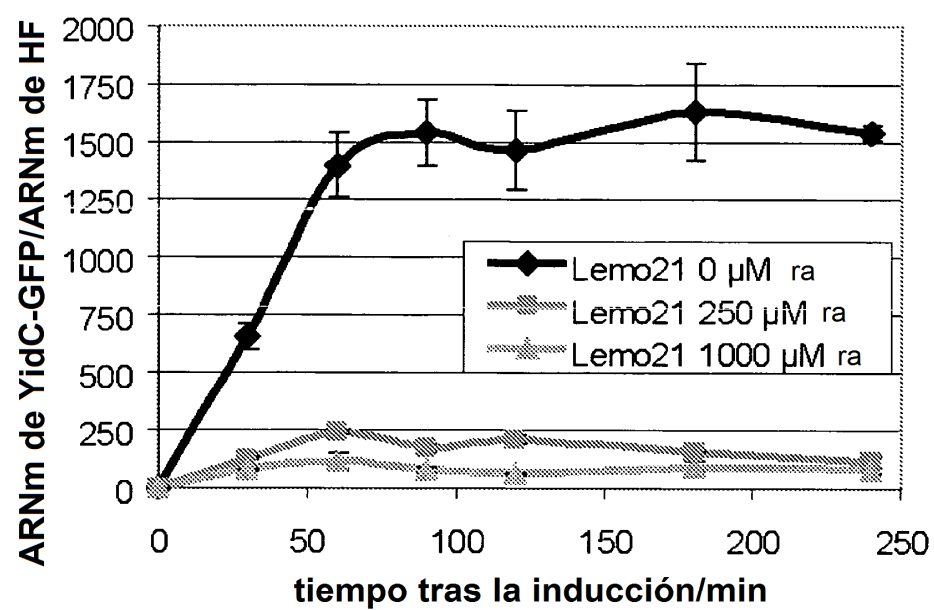
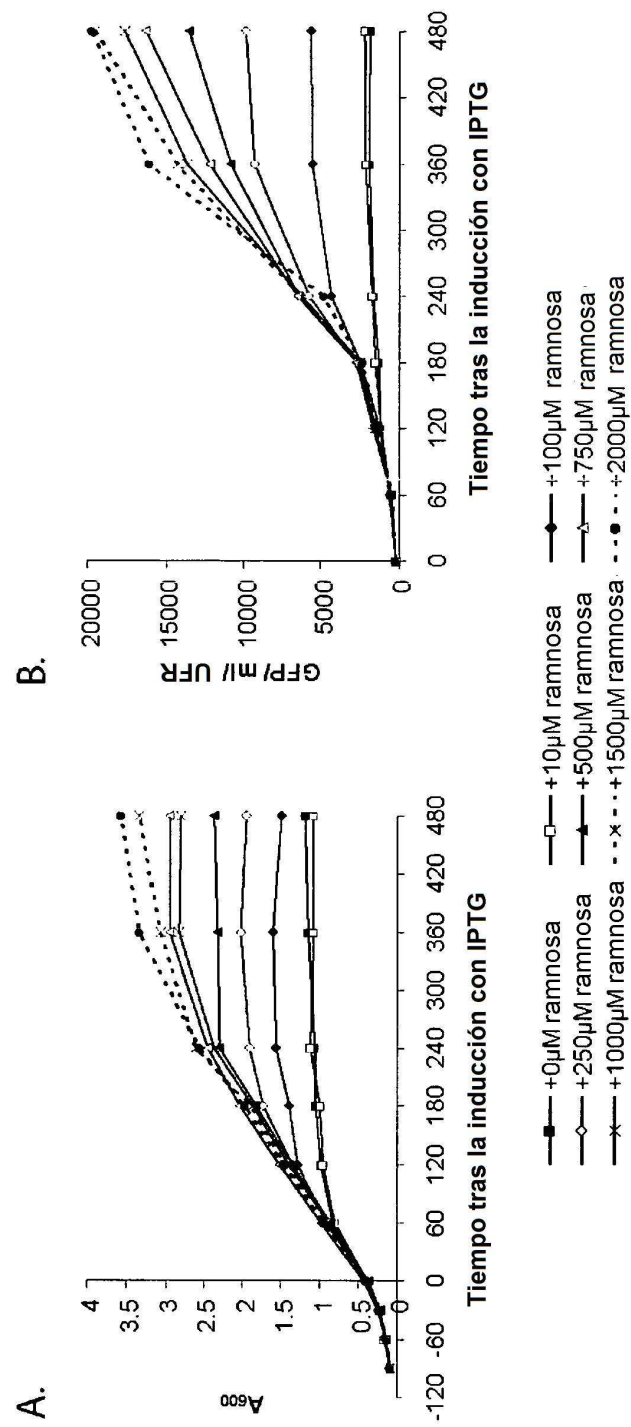


Figura 4



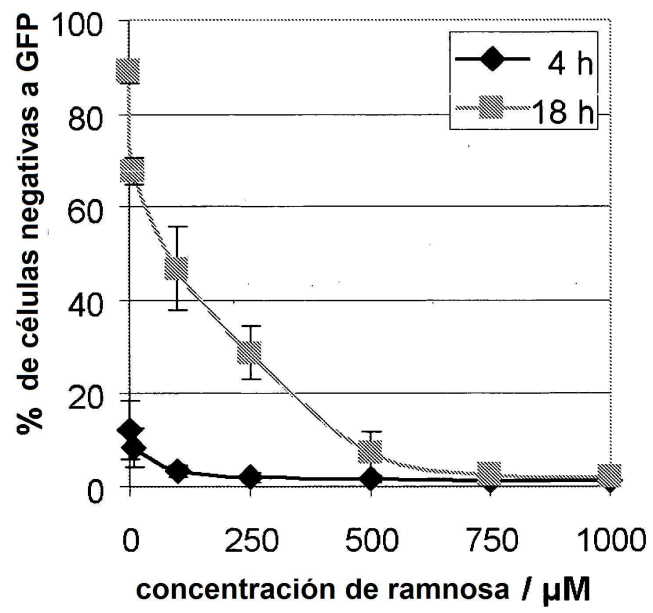


Figura 6

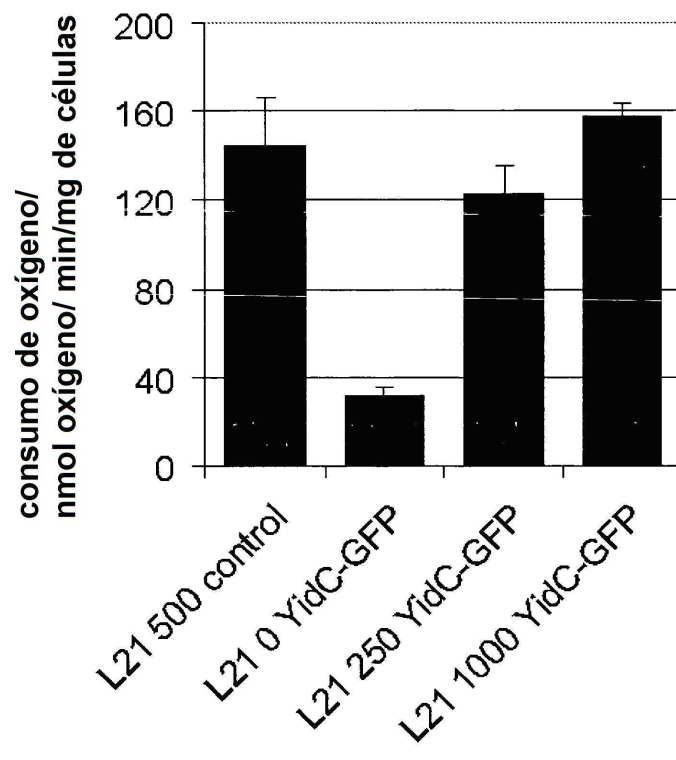


Figura 7

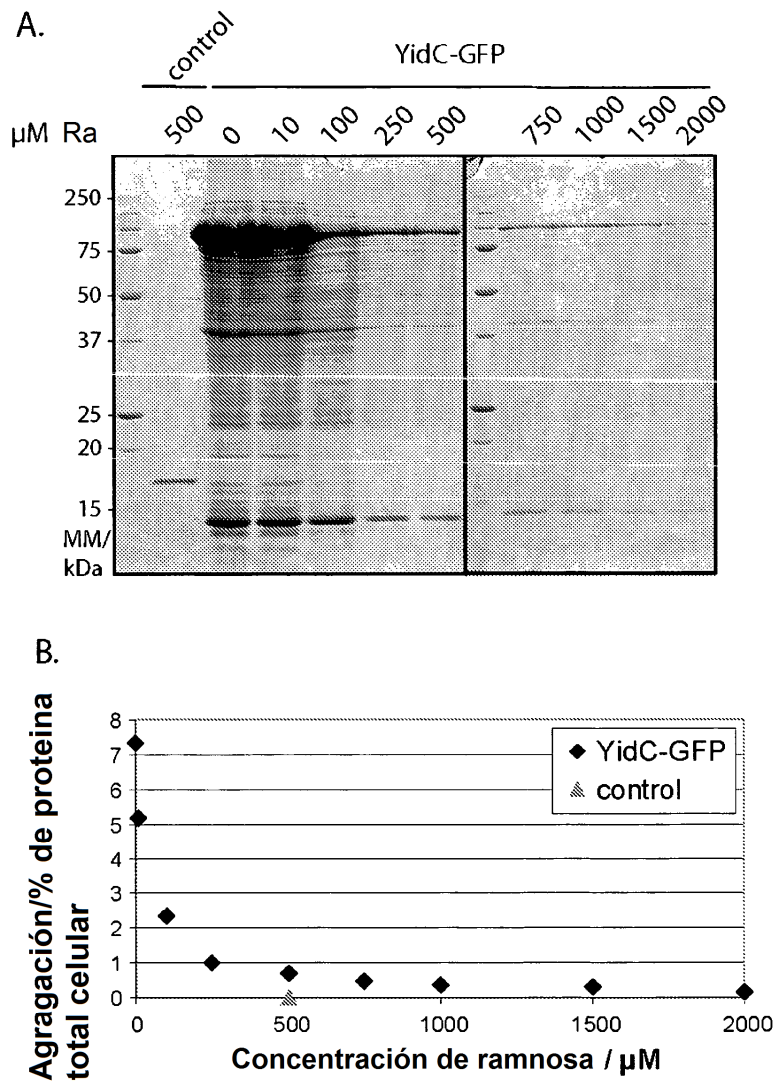


Figura 8

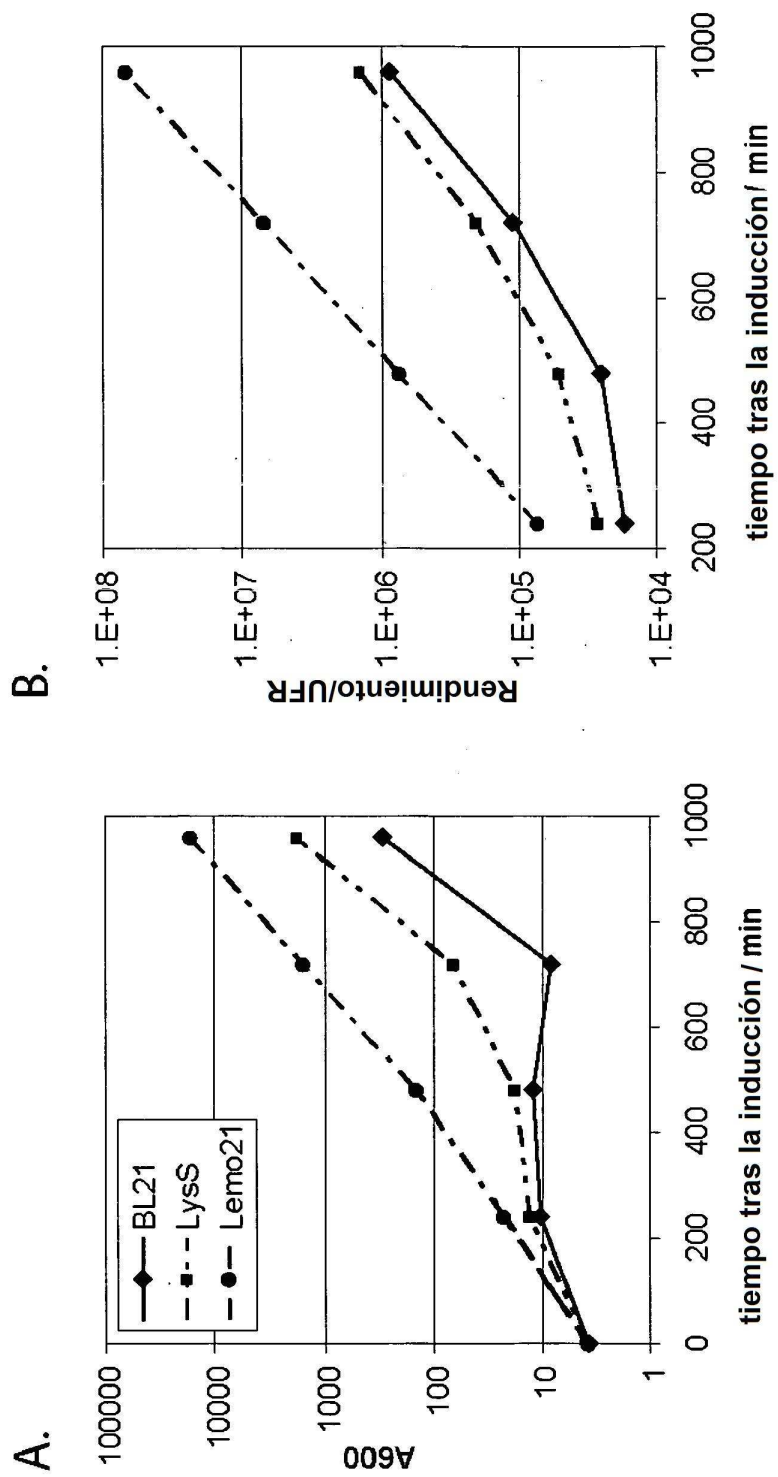


Figura 9

Figura 10 A

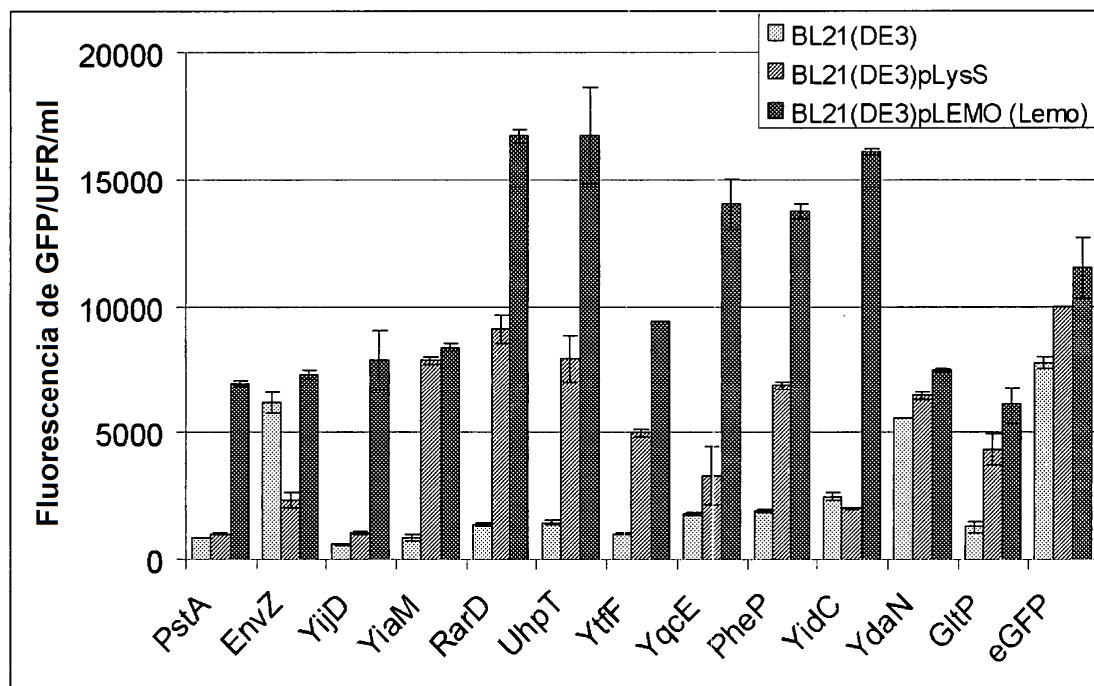
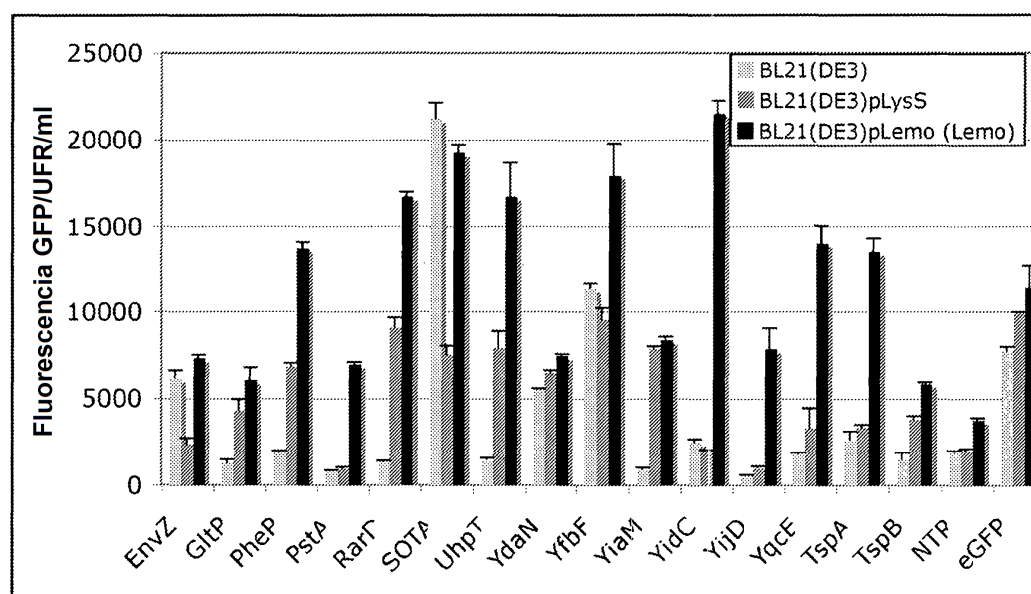


Figura 10 B



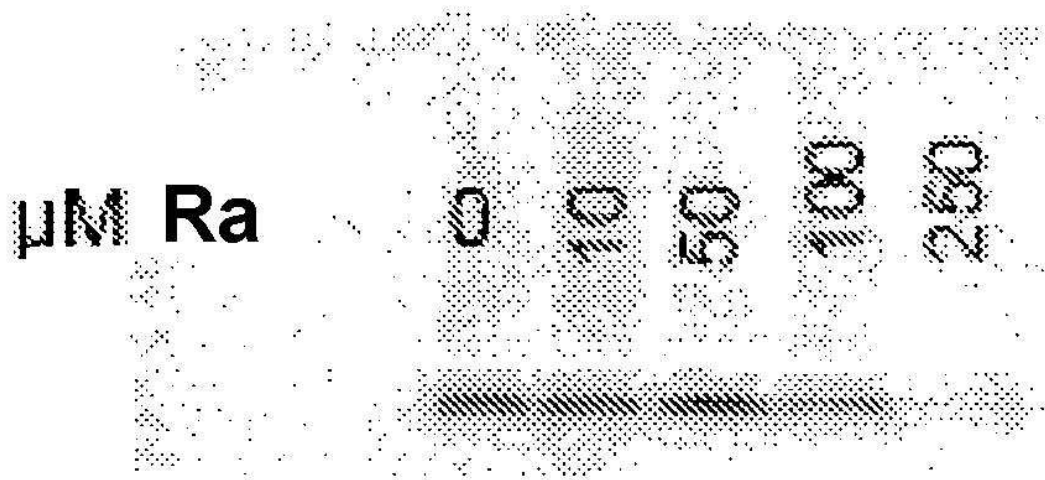


Figura 11

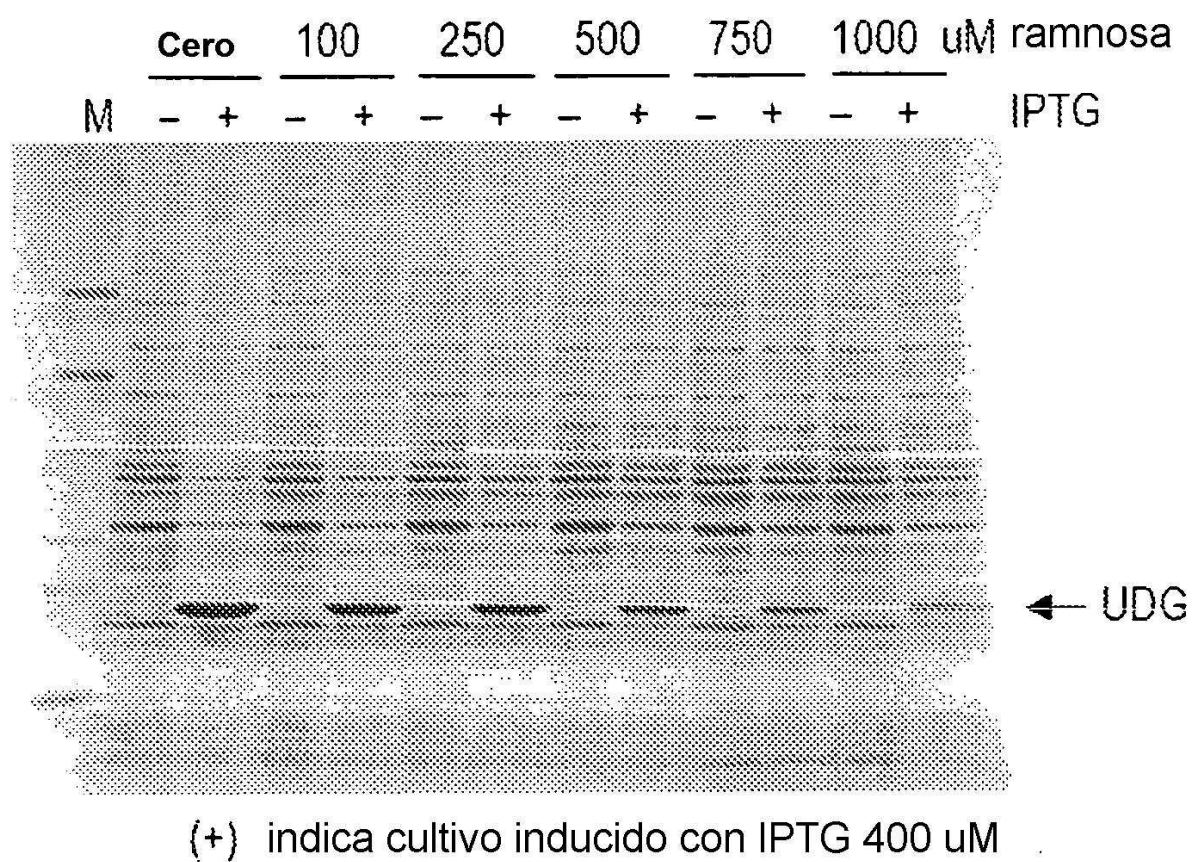
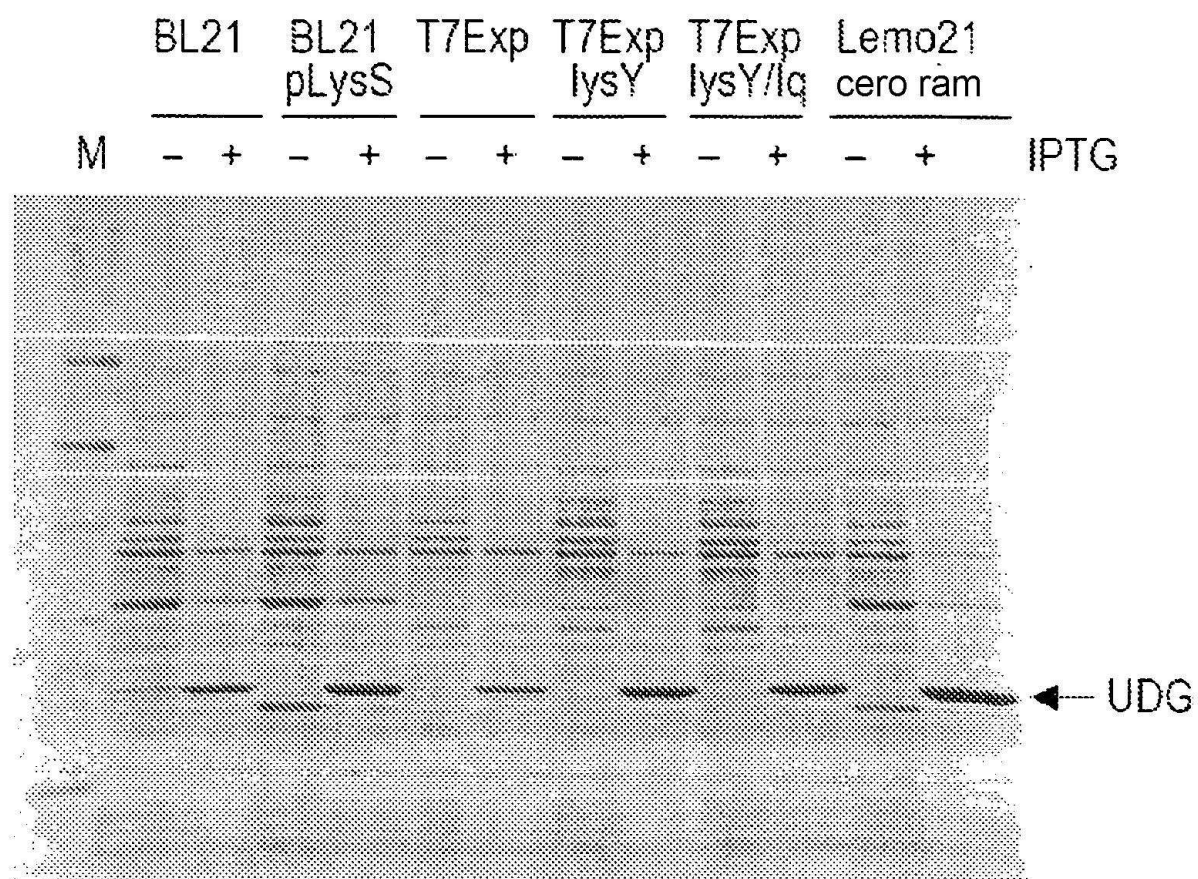


Figura 12A



(+) indica cultivo inducido con IPTG 400 μ M

Figura 12B