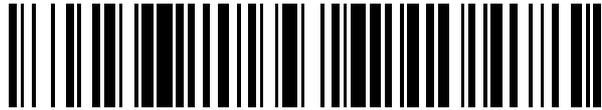


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 694 683**

51 Int. Cl.:

C07K 14/47 (2006.01)

C12N 9/50 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.10.2013 PCT/US2013/063804**

87 Fecha y número de publicación internacional: **17.04.2014 WO14058830**

96 Fecha de presentación y número de la solicitud europea: **08.10.2013 E 13779699 (1)**

97 Fecha y número de publicación de la concesión europea: **03.10.2018 EP 2906592**

54 Título: **Inactivación proteolítica de proteínas seleccionadas en extractos bacterianos para mejorar la expresión**

30 Prioridad:

12.10.2012 US 201261713245 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

26.12.2018

73 Titular/es:

**SUTRO BIOPHARMA, INC. (100.0%)
310 Utah Avenue, Suite 150
South San Francisco, California 94080, US**

72 Inventor/es:

**THANOS, CHRISTOPHER D.;
MURRAY, CHRISTOPHER J.;
YANG, JUNHAO y
STEPHENSON, HEATHER**

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 694 683 T3

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).

DESCRIPCIÓN

Inactivación proteolítica de proteínas seleccionadas en extractos bacterianos para mejorar la expresión

5 **Campo de la invención**

La invención se refiere a la tecnología de síntesis libre de células, y, en particular, a la inactivación proteolítica de proteínas seleccionadas en un extracto bacteriano. Un uso preferido es el aumento del rendimiento de los polipéptidos que tienen un aminoácido no nativo incorporado en un resto de aminoácido definido.

10

Antecedentes de la invención

El uso de extractos bacterianos libres de células para la síntesis de proteínas *in vitro* ofrece varias ventajas frente a los métodos convencionales de expresión de proteínas *in vivo*. Los sistemas libres de células pueden dirigir la mayoría, si no todos, los recursos metabólicos de la célula hacia la producción exclusiva de una proteína. Por otra parte, la falta de componentes de la pared y de la membrana celulares *in vitro* es ventajosa, ya que permite el control del entorno de la síntesis. Sin embargo, la eficacia de los extractos libres de células puede ser disminuida por las proteínas bacterianas que inhiben la síntesis de proteínas, ya sea directa o indirectamente. Por lo tanto, la inactivación de proteínas no deseadas que disminuyen la eficacia de la síntesis de proteínas debería aumentar el rendimiento de las proteínas deseadas en extractos libres de células. Por ejemplo, la inactivación de proteínas que disminuyen la eficacia de la síntesis de proteínas debería aumentar el rendimiento de los polipéptidos que tienen aminoácidos no nativos incorporados en un resto de aminoácido definido. La introducción de aminoácidos no nativos (aann) en polipéptidos es útil para aumentar la diversidad biológica y la función de las proteínas. Un enfoque para producir polipéptidos que tengan un aann incorporado en un resto de aminoácido definido es el uso de un aann, CUA ortogonal aminoacilado que contiene ARNt para la introducción del aann en el polipéptido naciente en un codón ámbar (parada) durante la traducción de la proteína. Sin embargo, la incorporación del aann en un codón ámbar puede ser inhibida por el complejo de terminación bacteriano nativo, que normalmente reconoce el codón de parada y termina la traducción. El factor de liberación 1 (RF1) es una proteína compleja de terminación que facilita la terminación de la traducción al reconocer el codón ámbar en una secuencia de ARNm. El reconocimiento de RF1 del codón de parada ámbar puede potenciar los productos de truncamiento prematuros en el sitio de incorporación de los aminoácidos no nativos y, por lo tanto, reducir el rendimiento proteico. Por lo tanto, la atenuación de la actividad de RF1 puede aumentar la incorporación de aann a proteínas recombinantes.

15

20

25

30

35

40

Previamente se ha demostrado que la incorporación de aann se puede aumentar mediante la atenuación de la actividad RF1 de 3 modos: 1) neutralizando la inactivación de anticuerpos de RF1; 2) desactivando genómicamente RF1 (en una cepa reforzada con RF2); y 3) retirando de manera específica del sitio la RF1 usando una cepa diseñada para expresar RF1 que contiene un marcador proteico para la eliminación mediante cromatografía de afinidad (Dominio de unión a quitina y marcador de His). La presente divulgación describe un nuevo método para inactivar a RF1 mediante la introducción de sitios de escisión proteolítica en la secuencia de aminoácidos de RF1. Los sitios de escisión no son accesibles a la proteasa durante el crecimiento de las células bacterianas, pero son escindidos por la proteasa cuando las células bacterianas se lisan para producir un extracto libre de células. Por lo tanto, el rendimiento de los polipéptidos de longitud completa que tienen un aann incorporado en un codón ámbar se aumenta en los extractos de células bacterianas que expresan variantes de RF1 modificadas descritas en el presente documento.

45 **Breve resumen de la invención**

La presente invención proporciona una proteína funcional de factor de liberación 1 (RF1) que puede ser escindida por una proteína de la membrana externa T1 (OmpT1) y tiene una secuencia de aminoácidos que tiene al menos un 90 % de identidad de secuencia con SEQ ID NO: 1, en la que hay un enlace peptídico escindible por OmpT1 situado dentro de la región del bucle de conmutación, y en la que la región del bucle de conmutación comprende los aminoácidos 287-304 de la SEQ ID NO: 1 modificada para incluir tres aminoácidos básicos adyacentes que están cargados positivamente a pH 7,0.

50

La invención también proporciona una proteína funcional de factor de liberación 1 (RF1) que es escindible por una proteína de la membrana externa T1 (OmpT1), en la que hay un enlace de péptido escindible por OmpT1 situado dentro de la región del bucle de conmutación, y en la que la región del bucle de conmutación comprende una secuencia de aminoácidos seleccionada del grupo de secuencias que consiste en:

55

QARRGSTRRNLLGSGDRS (SEQ ID NO:26);
 QQARRGTRRNLLGSGDRS (SEQ ID NO:27);
 QQAARRGRRNLLGSGDRS (SEQ ID NO:28);
 QQAEARRGRNLLGSGDRS (SEQ ID NO:29);
 QQAEAARRGNLLGSGDRS (SEQ ID NO:30);
 QQAEASARRGLLGSAGDRS (SEQ ID NO:31);
 QQAEASTARRGLGSGDRS (SEQ ID NO:32);
 QQAEASTRARRGGSGDRS (SEQ ID NO:33);

60

65

QQAEASTRRARRGSGDRS (SEQ ID NO:34);
 QQAEASTRRNARRGGDRS (SEQ ID NO:35);
 QQAEASTRRNLARRGDRS (SEQ ID NO:36);
 QQAEASTRRNLLARRGRS (SEQ ID NO:37);
 QQAEASTRRNLLGARRGS (SEQ ID NO:38);
 QQAEASTRRNLLGSARRG (SEQ ID NO:39);
 QQAEASTRRNLLGSGARR (SEQ ID NO:40);
 QQAWLAARRGRGGSGDRS (SEQ ID NO:41);
 QQAEWLAARRGRGGSGDRS (SEQ ID NO:42);
 QQAEWLAARRGRGGDRS (SEQ ID NO:43);
 QQWGGRWARKKGTIDRS (SEQ ID NO:44);
 QQAWGGRWARKKGTIDRS (SEQ ID NO:45); y
 QQAEWGGRWARKKGTIRS (SEQ ID NO:46).

15 La invención proporciona además una célula bacteriana que expresa tanto una proteína de la membrana externa T1 (OmpT1) como una proteína RF de la invención descrita en el presente documento.

En algunos casos, la proteína diana esencial se selecciona de RF1, RF2, RNAsa, tiorredoxina reductasa, glutarodoxina reductasa, glutatión reductasa, enzimas de degradación de aminoácidos, polifosfato quinasa o proteínas de choque en frío.

La invención también proporciona un método para reducir la actividad perjudicial de la proteína aRF1 de la invención en un sistema de síntesis libre de células *in vitro*, comprendiendo el método las etapas de:

- 25 i) cultivar una bacteria positiva en OmpT1 que exprese la proteína RF1 de la invención;
 ii) lisar las bacterias para crear un extracto de síntesis libre de células;
 iii) poner en contacto la proteína RF1 de la invención con OmpT1 en una cantidad suficiente para reducir la proteína intacta en un 50 %;
 iv) añadir un molde de ácido nucleico al extracto en el que el molde codifica una proteína de interés; y
 30 v) dejar que el sistema de síntesis libre de células produzca la proteína de interés.

La invención proporciona además un ácido nucleico que codifica la proteína funcional de factor de liberación 1 (RF1) de la invención.

35 La invención también proporciona un método para reducir la competencia de RF1 en un codón ámbar en un sistema de síntesis libre de células *in vitro*, comprendiendo el método las etapas de:

- i) cultivar una bacteria positiva en OmpT1 que exprese la proteína RF1 de la invención;
 ii) lisar las bacterias para crear un extracto de síntesis libre de células;
 40 iii) poner en contacto la proteína RF1 del extracto con OmpT1 en una cantidad suficiente para reducir la proteína RF1 intacta en un 50 %;
 iv) añadir un molde de ácido nucleico al extracto en el que el molde codifica una proteína de interés e incluye un codón de parada; y
 v) dejar que el sistema de síntesis libre de células produzca la proteína de interés.

45 La invención proporciona además un sistema de síntesis libre de células que comprende, en una sola mezcla de reacción:

- i) los componentes de un lisado bacteriano suficientes para traducir un molde de ácido nucleico que codifica una proteína;
 50 ii) un molde de ácido nucleico que codifica una proteína de interés y que tiene al menos un codón ámbar;
 iii) ARNt complementario al codón ámbar; e
 iv) la proteína RF1 de la invención.

55 La invención proporciona además un método para expresar proteína en un sistema de síntesis libre de células, que comprende:

- i. combinar un molde de ácido nucleico con un extracto bacteriano que comprende la proteína RF1 de la invención;
 e
 60 ii. expresar una proteína a partir del molde de ácido nucleico.

La invención también proporciona una célula bacteriana que comprende un ácido nucleico que codifica la proteína RF1 de la invención incorporado en el genoma de la célula bacteriana.

65 La invención también proporciona un extracto bacteriano que comprende una proteína RF1 de la invención.

Breve descripción de los dibujos

La **Figura 1** muestra que las variantes de proteína RF1 modificadas que contienen sitios de escisión de OmpT tienen actividad RF1 de tipo silvestre cuando se añaden a extractos libres de células que carecen de actividad OmpT. Se determinó la capacidad de variantes de RF-1 exógenas para reducir la supresión (aumentar la terminación de la cadena) cuando se añade un aminoácido no nativo en un codón ámbar introducido en la posición S378 de la proteína Fc (Fc_S378TAG). La Figura 1(A), Carril 1, es un extracto de control sin RF1 añadida exógenamente. El extracto contiene RF1 de tipo silvestre, lo que produce una reducción de la supresión ámbar, produciendo un Fc truncado (banda inferior del gel). El Carril 2 es un extracto de control que contiene RF1 de tipo silvestre (WT) añadida exógenamente. La adición de RF1 de tipo silvestre produce una reducción adicional en la supresión ámbar, lo que produce un Fc relativamente más truncado. Los Carriles 3-7 muestran extractos libres de células que contienen diferentes variantes de RF1 de la invención. Las cinco variantes poseen actividad RF1. La Figura 1(B) muestra la actividad relativa de las variantes de RF1, en la que la actividad RF1 WT añadida se establece en el 100 por ciento, y ninguna RF1 añadida se establece en cero por ciento.

La **Figura 2** muestra que los extractos libres de células de una cepa bacteriana diseñada para expresar una variante de RF1 modificada (A18) producen un aumento sustancial en el rendimiento de IgG de longitud completa al introducirse aann en diversas posiciones de la cadena pesada, en comparación con una cepa bacteriana que expresa una RF1 de tipo silvestre. 2(A) panel izquierdo: extracto de SBJY001, que contiene OmpT intacta y RF1 de tipo silvestre. 2(A) panel derecho: extracto de SBHS016 (cepa 16), que contiene OmpT intacta y variante A18 de RF1. La intensidad de la banda de IgG de longitud completa (intacta) es mayor en el extracto de la cepa 16, lo que indica que la incorporación de aann aumentó sustancialmente, con una disminución en las proteínas truncadas. 2 (B) muestra que el rendimiento soluble de IgG de las proteínas que contienen un aann en las posiciones indicadas de la cadena pesada se aumenta sustancialmente en el extracto de la cepa 16, lo que indica que la escisión por OmpT de la variante A18 de RF1 produce una cadena pesada menos truncada. El extracto de "RF1+" (RF1 positivo) se refiere a la RF1 de la cepa 001 que no es escindida por OmpT. El extracto "RF1-" (RF1 negativo) se refiere a la RF1 de la cepa 16 que se modifica al incluir sitios de escisión de OmpT, y por lo tanto, es degradada por la actividad OmpT.

La **Figura 3** muestra que los extractos libres de células de una cepa bacteriana diseñada para expresar una variante de RF1 modificada (A18) producen un aumento sustancial en la supresión ámbar correspondiente a una mayor incorporación de aann en las posiciones indicadas de la cadena pesada. 3(A) panel izquierdo: extracto de SBJY001, que contiene OmpT intacta y RF1 de tipo silvestre. 3(A) panel derecho: extracto de SBHS016 (cepa 16), que contiene OmpT intacta y variante A18 de RF1. La intensidad de la banda de la cadena pesada de longitud completa (intacta) (Hc) es mayor en el extracto de la cepa 16, lo que indica que la incorporación de aann aumentó sustancialmente, con una disminución en las proteínas truncadas. 3(B) muestra que la supresión ámbar de las proteínas que contienen un aann en las posiciones indicadas de la cadena pesada se aumenta sustancialmente en el extracto de la cepa 16, lo que indica que la escisión por OmpT de la variante A18 de RF1 produce un aumento de la supresión ámbar. El extracto de "RF1+" (RF1 positivo) se refiere a la RF1 de la cepa 001 que no es escindida por OmpT. El extracto "RF1-" (RF1 negativo) se refiere a la RF1 de la cepa 16 que se modifica al incluir sitios de escisión de OmpT, y por lo tanto, es degradada por la actividad OmpT.

Definiciones

La "aminoacilación" o el "aminoacilato" se refieren al proceso completo en el que un ARNt se "carga" con su aminoácido correcto, que es el resultado de añadir un grupo aminoacilo a un compuesto. En lo que respecta a la presente invención, un ARNt sometido a una aminoacilación o que ha sido aminoacilado es aquel que se ha cargado con un aminoácido, y un aminoácido sometido a aminoacilación o que ha sido aminoacilado es aquel que se ha cargado en una molécula de ARNt.

"Aminoacil-ARNt sintetasa" o "ARNt sintetasa" o "sintetasa" o "aaRs" o "RS" se refieren a una enzima que cataliza un enlace covalente entre un aminoácido y una molécula de ARNt. Esto da lugar a una molécula de ARNt "cargada", que es una molécula de ARNt que tiene sus respectivos aminoácidos unidos a través de un enlace éster.

"Codón" se refiere a un grupo de 3 nucleótidos consecutivos de un molde de ácido nucleico que especifica una determinada señal de aminoácido natural, aminoácido no nativo o señal de parada de la traducción (terminación de la cadena polipeptídica). Debido a la degeneración del código genético, un aminoácido puede ser especificado por más de un codón.

Un codón ámbar es una secuencia de terminación de cadena polipeptídica (UAG) en el ARN (TAG en el ADN), que actúa para terminar la traducción del polipéptido en la mayoría de los organismos. El codón ámbar también puede codificar el aminoácido proteogénico pirrolisina si el ARNt apropiado es cargado por su aminoacil-ARNt sintetasa afín.

"ARNt de codón ámbar" o "ARNt supresor de ámbar" o "ARNt anticodón de ámbar" se refieren a un ARNt que se une a un codón ámbar.

"Aminoácidos adyacentes" se refiere a los aminoácidos que preceden o siguen inmediatamente en la secuencia de aminoácidos de un polipéptido. Por ejemplo, el aminoácido de la posición dos de la secuencia de aminoácidos primaria

de un polipéptido es adyacente a los aminoácidos de las posiciones uno y tres. "Dos aminoácidos básicos adyacentes" se refiere a un aminoácido básico que precede o sigue inmediatamente a otro aminoácido básico en la secuencia de aminoácidos primaria de un polipéptido. Por ejemplo, el aminoácido básico Arg (R) de la posición 295 de la secuencia de aminoácidos de RF1 puede ser adyacente a un aminoácido básico tal como K o R introducido en la posición 296 de la secuencia de aminoácidos de RF1.

"Extracto libre de células derivado de bacterias" se refiere al preparado de mezclas de reacción *in vitro* capaz de traducir el ARNm en polipéptidos. Las mezclas incluyen ribosomas, ATP, aminoácidos y ARNt. Pueden derivarse directamente de bacterias lisadas, de componentes purificados o combinaciones de ambos.

Los "aminoácidos básicos" son polares y están cargados positivamente a valores de pH por debajo de sus pK_a , y son muy hidrófilos. Los ejemplos de aminoácidos básicos a pH neutro incluyen arginina (Arg = R), lisina (Lys = K) e histidina (His = H).

El "sistema de síntesis libre de células" se refiere a la síntesis *in vitro* de polipéptidos en una mezcla de reacción que comprende extractos biológicos y/o reactivos definidos. La mezcla de reacción comprenderá un molde para la producción de la macromolécula, por ejemplo, ADN, ARNm, etc.; monómeros para la macromolécula que se vaya a sintetizar, por ejemplo, aminoácidos, nucleótidos, etc.; y cofactores, enzimas y otros reactivos que son necesarios para la síntesis, por ejemplo, ribosomas, ARNt no cargados, ARNt cargados con aminoácidos no naturales, polimerasas, factores de transcripción, etc.

La expresión "la actividad de escisión es superior al 50 % de la RF1 de tipo silvestre" se refiere a la velocidad de escisión de una proteína modificada descrita en el presente documento que es superior al 50 % de la velocidad de escisión de la proteína de tipo silvestre en condiciones especificadas. Por ejemplo, la velocidad de escisión puede ser superior al 50 % de la velocidad de escisión de la proteína de tipo silvestre tras 30 minutos a 30 °C cuando las proteínas modificadas y de tipo silvestre están presentes a una concentración similar (por ejemplo, una concentración de 0,1-1,0 micromolar) en un extracto libre de células de bacterias que expresan OmpT1.

La expresión "actividad perjudicial" se refiere a una actividad que disminuye el rendimiento proteico durante la síntesis de proteínas en los extractos libres de células. Por ejemplo, la actividad perjudicial puede inhibir la transcripción y/o traducción celular en sistemas de síntesis libres de células. La actividad perjudicial puede incluir la reducción enzimática de GSSG que se requiere para el plegamiento eficaz de proteínas unidas por disulfuro producidas en una reacción de síntesis libre de células. La actividad perjudicial también puede incluir la terminación prematura de la cadena por un factor de liberación, de modo que la elongación del polipéptido se termina prematuramente en un codón de parada introducido.

Una "proteína diana" es una proteína que ha sido modificada para incluir un sitio de escisión de la proteasa, o una proteína que se modifica para incluir un aann. Una "proteína diana esencial" es una proteína que se requiere para el crecimiento y/o la supervivencia normales de una célula, tal como una célula bacteriana o una célula eucariota.

Una "proteína funcional de factor de liberación 1 (RF1)" se refiere a RF1 que conserva la misma o esencialmente la misma actividad que la proteína RF1 de tipo silvestre o sin modificar. La actividad RF1 funcional se puede ensayar, por ejemplo, midiendo la velocidad de crecimiento de las bacterias que expresan la proteína RF1 modificada, y comparando la velocidad de crecimiento con la de las bacterias que expresan RF1 de tipo silvestre o no modificada. La actividad funcional de otras proteínas modificadas para contener sitios de escisión de OmpT se puede determinar de manera similar, por ejemplo, midiendo la velocidad de crecimiento de las bacterias que expresan la proteína modificada, y comparando la velocidad de crecimiento con la de las bacterias que expresan la proteína de tipo silvestre o no modificada. La actividad RF1 funcional también se puede ensayar, por ejemplo, según la capacidad de la proteína RF1 modificada para reducir la incorporación de ARNn ortogonal de un aann en una posición específica de un ARNm que codifica una proteína diana, aumentando así la cantidad de terminación prematura de la cadena (es decir, aumentando la cantidad de proteína truncada).

La expresión "modificada para incluir", en el contexto de la presente invención, se refiere a sustituciones, adiciones o eliminaciones de aminoácidos en una secuencia de proteína o polipéptido, o a las sustituciones correspondientes, adiciones o eliminaciones en una secuencia de ácido nucleico que codifica la proteína modificada. La proteína modificada puede incluir sustituciones de aminoácidos que reemplazan un número igual de aminoácidos de la secuencia de proteínas de tipo silvestre, aminoácidos añadidos a la secuencia de tipo silvestre o aminoácidos eliminados de la secuencia de tipo silvestre. Por ejemplo, la proteína puede modificarse para incluir aminoácidos básicos que reemplacen al/a los aminoácido/s de tipo silvestre de la posición correspondiente de la secuencia de aminoácidos. La proteína también puede modificarse para incluir secuencias de aminoácidos que sean secuencias conocidas de escisión de proteasas, tales como secuencias de escisión de OmpT.

Aminoácido "no natural" o "no nativo" se refiere a los aminoácidos que no son uno de los veinte aminoácidos naturales que constituyen todas las proteínas que, sin embargo, son capaces de ser modificados biológicamente para incorporarse a las proteínas. Los aminoácidos no nativos pueden incluir enantiómeros D de péptidos o cualquier modificación posterior a la traducción de uno de los veinte aminoácidos naturales. Se puede usar una amplia variedad

de aminoácidos no nativos en los métodos de la invención. El aminoácido no nativo se puede escoger en función de las características deseadas del aminoácido no nativo, por ejemplo, la función del aminoácido no nativo, tal como la modificación de las propiedades biológicas de las proteínas tales como la toxicidad, la biodistribución o la semivida, propiedades estructurales, propiedades espectroscópicas, propiedades químicas y/o fotoquímicas, propiedades catalíticas, capacidad de reaccionar con otras moléculas (bien covalente o no covalentemente), o similares. Los aminoácidos no nativos que pueden usarse en los métodos de la invención pueden incluir, pero no se limitan a, un análogo no nativo de un aminoácido tirosina; un análogo no nativo de un aminoácido glutamina; un análogo no nativo de un aminoácido fenilalanina; un análogo no nativo de un aminoácido serina; un análogo no nativo de un aminoácido treonina; un grupo alquilo, arilo, acilo, azido, ciano, halo, hidrazina, hidrazida, hidroxilo, alqueno, alquino, éter, tiol, sulfonilo, seleno, éster, tioácido, borato, boronato, fosfo, fosfono, fosfina, heterocíclico, enona, imina, aldehído, hidroxilamina, ceto, o aminoácido sustituido con amino, o cualquier combinación de los mismos; un aminoácido con un agente de reticulación fotoactivable; un aminoácido marcado en el espín; un aminoácido fluorescente; un aminoácido con un nuevo grupo funcional; un aminoácido que interactúa covalente o no covalentemente con otra molécula; un aminoácido de unión a metales; un aminoácido que contiene metal; un aminoácido radiactivo; un aminoácido fotoactivado y/o fotoisomerizable; un aminoácido que contiene biotina o análogo de biotina; un aminoácido glicosilado o modificado con carbohidratos; un aminoácido que contiene ceto; aminoácidos que comprenden polietilenglicol o poliéter; un aminoácido sustituido con un átomo pesado; un aminoácido químicamente escindible o fotoescindible; un aminoácido con una cadena lateral elongada; un aminoácido que contiene un grupo tóxico; un aminoácido sustituido con azúcar, por ejemplo, una serina sustituida con azúcar o similares; un aminoácido que contiene azúcar ligado al carbono, por ejemplo, una serina sustituida con azúcar o similares; un aminoácido que contiene azúcar ligado al carbono; un aminoácido redox activo; un ácido que contiene alfa-hidroxi; un aminoácido que contiene ácido amino tio; un aminoácido alfa, alfa-disustituido; un beta-aminoácido; un aminoácido cíclico distinto de la pralina, etc.

"Molde de ácido nucleico" incluye un polinucleótido de ADN o ARN a partir del que se traducirá un polipéptido. Los expertos en la materia entenderán que un molde de ácido nucleico de ADN debe transcribirse primero en ARN, y que el ARN se traduce en un polipéptido. El ADN se puede transcribir en ARN *in vivo* o *in vitro*. Los métodos de transcripción *in vitro* de un molde de ADN son bien conocidos en la técnica. En algunas realizaciones, el molde de ADN se somete a la transcripción y a la traducción *in vitro* simultáneas.

La "proteína de la membrana externa T1" (OmpT1) es una serina proteasa de la membrana superficial y pertenece a la familia OmpTina de bacterias gram-negativas. Se sabe que OmpT1 escinde los péptidos antimicrobianos, activa el plasminógeno humano y degrada algunas proteínas recombinantes heterólogas. OmpT1 y sus homólogos escinden sustratos sintéticos entre restos dibásicos con alta eficacia catalítica. Se ha demostrado que la escisión de secuencias que contienen restos dibásicos es importante para la inactivación de los péptidos antibióticos y las colicinas, la proteólisis de las proteínas de la membrana bacteriana en *trans* y la degradación de las proteínas recombinantes expresadas en *E. coli*. En el contexto de la presente invención, OmpT1 tiene ciertas ventajas, ya que se encuentra en la membrana externa de la célula y, por lo tanto, no está en contacto con las posibles proteínas de sustrato que se encuentran dentro de la célula bacteriana intacta.

La expresión "sistema de fosforilación oxidativa de las bacterias permanece activo" se refiere a un lisado bacteriano que presenta la fosforilación oxidativa activa durante la síntesis de proteínas. Por ejemplo, el lisado bacteriano puede generar ATP usando enzimas ATP sintasa y reducción de oxígeno. Se entenderá que otros sistemas de traducción conocidos en la técnica también pueden usar una fosforilación oxidativa activa durante la síntesis de proteínas. La activación de la fosforilación oxidativa se puede demostrar mediante la inhibición de la vía usando inhibidores específicos, tales como los inhibidores de la cadena de transporte de electrones.

"Traducción" se refiere al proceso mediante el que un molde de ARN se convierte en un polipéptido que contiene aminoácidos naturales y/o no nativos, y es bien conocido en la técnica. La traducción implica la etapa de inicio, mediante la que un ribosoma se une al molde de ARN, en general, en el codón FMet (por ejemplo, AUG), y la etapa de elongación, mediante la que el anticodón de una molécula de ARNt cargada se empareja con un codón en el molde de ARN, repitiéndose esta etapa a medida que el ribosoma se desplaza en sentido descendente por el molde de ARN. Como cada anticodón de ARNt se empareja con su codón correspondiente, el grupo amino del aminoácido cargado en cada molécula de ARNt está unido covalentemente al grupo carboxilo del aminoácido anterior a través de enlaces peptídicos. En general, la traducción también implica la etapa de terminación, mediante la que el ribosoma se encuentra con un codón de parada de la traducción, terminando así la elongación de la cadena y la liberación del polipéptido del ribosoma. Sin embargo, en los métodos descritos en el presente documento, el molde de ARN puede comprender un ORF que tenga un codón de parada ámbar UAG, que es reconocido por el anticodón CUA de un ARNt cargado con un aann. Por lo tanto, en los métodos actuales, el codón ámbar no funciona necesariamente para terminar la traducción.

"Sistema de traducción" se refiere a una mezcla de componentes que es capaz de traducir el ARNm a polipéptidos *in vitro*. Un sistema de traducción puede ser un extracto libre de células, un lisado celular reconstituido o una mezcla purificada de componentes que sea capaz de traducir el ARNm a polipéptidos *in vitro*. Por ejemplo, el extracto libre de células o el lisado celular usados en los métodos descritos en el presente documento pueden derivarse de *Escherichia coli* (*E. coli*). El sistema de traducción también puede comprender un sistema de traducción *in vitro* purificado y

reconstituido. Los métodos descritos en el presente documento también pueden utilizar un sistema de traducción que presenta una fosforilación oxidativa activa durante la síntesis de proteínas. Por ejemplo, el sistema de traducción puede generar ATP usando enzimas ATP sintetasa. El sistema de traducción también puede comprender al menos un molde de ácido nucleico. El molde de ácido nucleico se puede transcribir y traducir simultáneamente a proteínas usando el sistema de traducción.

"Sistema de traducción ribosomal reconstituido" se refiere a una mezcla de componentes purificados que es capaz de traducir una molécula de ácido nucleico tal como ARNm *in vitro*, como se describe en Tan *et al.*, *Methods* 36, 279-290 (2005), y Forster *et al.*, patente de EE.UU. n.º 6.977.150.

Un "enlace de péptido escindible por OmpT1" es un enlace peptídico que puede ser escindido por OmpT1.

Un "motivo expuesto en la superficie" es un dominio que está presente en el exterior de una molécula de proteína. Los motivos expuestos en la superficie, en general, son más accesible a la escisión proteolítica, por ejemplo, por una serina proteasa tal como OmpT1. Un experto en la materia entendería que un motivo expuesto en la superficie se puede definir de varias maneras. Por ejemplo, un motivo expuesto en la superficie suele carecer de estructura secundaria, tal como hélice alfa o lámina beta. Además, los motivos expuestos en la superficie no suelen compartir una cantidad significativa de homología de la secuencia de aminoácidos, incluso entre secuencias de proteínas homólogas. Sin embargo, un experto en la materia puede calcular el área de superficie relativa accesible para las regiones de una proteína de interés usando el algoritmo GetArea incorporado en el software de modelización molecular pymol (véase en Internet, en pymolwiki.org/index.php/Get_Area).

Una "motivo expuesto en la superficie que tiene un factor B de al menos 50 Å²" se refiere a la movilidad de los átomos individuales en el motivo expuesto en la superficie. Los átomos de los motivos expuestos en la superficie pueden tener más movilidad que los átomos del núcleo de una proteína. En general, los factores B indican el movimiento vibratorio relativo de diferentes partes de una estructura, tal como una proteína. Los átomos con bajos factores B pertenecen a una parte de la proteína que está bien ordenada y, por lo tanto, tienen una movilidad relativamente baja. Los átomos con factores B altos, en general, pertenecen a una parte de la proteína que es relativamente flexible y, por lo tanto, tienen una movilidad relativamente alta. Cuanto mayor sea el factor B, más probable será la observación de errores de posición de un átomo individual en una estructura de proteína. Por ejemplo, un factor B de entre 40 y 60 sugiere que se pueden observar errores de posición de hasta 1,0 Angstrom. Un factor B superior a 60 sugiere que no es probable que el átomo esté dentro de 1,0 Angstrom de su posición observada. (Véase, por ejemplo, en Internet, en wiki.cmbi.ru.nl/index.php/B-factor). El factor B se calcula como:

$$B_i = 8\pi^2 U_i^2$$

en la que U_i^2 es la media del desplazamiento al cuadrado del átomo i . Esto produce un factor de ponderación sobre la contribución del átomo i a la transformada de Fourier de:

$$\exp\left(-B_i \frac{\sin^2 \Theta}{\lambda^2}\right)$$

A medida que aumenta U , aumenta B , y se reduce la contribución del átomo a la dispersión. Véase, por ejemplo, en Internet, en:

pldserver1.biochem.queensu.ca/~rlc/work/teaching/definitions.shtml. El factor B se mide en unidades de Å².

La expresión "diagrama de Ramachandran" se refiere a un método para la visualización de ángulos diedros de la cadena principal ψ (psi) frente a ϕ (Phi) para cada resto de aminoácido de una estructura de proteína. Los métodos para calcular los ángulos Phi y Psi se describen, por ejemplo, en Lovell, S. C. *et al.*, "Proteins: Structure, Function, and Genetics", 50:437-450 (2003), y Chen, V. B. *et al.*, "MolProbity: all-atom structure validation for macromolecular crystallography", *Acta Crystallographica* D66: 12-21 (2010). Es posible determinar los ángulos Phi y Psi en un diagrama de Ramachandran refiriéndose a las referencias anteriores o, por comodidad, se pueden usar programas de software disponibles gratuitamente en Internet. Por ejemplo, se puede cargar un archivo de coordenadas o un archivo PDB en el servidor MOLPROBITY ubicado en kinemage.biochem.duke.edu (véase Chen, V. B. *et al.*, anteriormente citados). Como alternativa, se puede usar Ramachandran Plot Explorer disponible en Internet, en boscoh.com/ramaplot.

La expresión "área de superficie total accesible a un disolvente" (SASA) se refiere al área de superficie de una proteína o un polipéptido que es accesible a un disolvente. El área de superficie accesible al disolvente se puede describir en unidades de angstrom al cuadrado, y se puede calcular usando el algoritmo de "bola rodante" desarrollado por Shrake

y Rupley (Shrake, A; Rupley, J. A. (1973). "Environment and exposure to solvent of protein atoms. Lysozyme and insulin". *J Mol Biol* 79 (2): 351-71.). El área de superficie accesible al disolvente también se puede calcular como se describe en Fraczekiewicz, R. y Braun, W., "Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules", *J. Comp. Chem.*, 19:319-333 (1998). Por comodidad, se puede calcular el área de superficie total accesible a un disolvente usando programas de software disponibles gratuitamente en Internet. Por ejemplo, se puede usar el conocido software GETAREA, que se encuentra en la página Web de la sección de Medicina de la Universidad de Texas curie.utmb.edu/getarea.html, para calcular el área de superficie accesible al disolvente (energía de solvatación) de una molécula de proteína, como se describe en Fraczekiewicz, R. y Braun, W. (1998), introduciendo las coordenadas atómicas en formato PDB (base de datos de proteínas) y especificando el radio deseado de la sonda de agua (parámetros: radio de la sonda de agua = 1,4).

La expresión "región del bucle de conmutación" se refiere a un motivo expuesto en la superficie que une los dominios tres y cuatro de RF1 y forma una conexión rígida que coloca el motivo GGQ del dominio 3 en contacto con el enlace éster de peptidil-ARNt en el centro de la peptidil transferasa (PTC) de la subunidad 50S. El reconocimiento del codón de parada por el complejo de terminación estabiliza una configuración reorganizada de la región del bucle de conmutación, que dirige el dominio 3 al PTC. La reorganización del bucle de conmutación produce la reorientación y la extensión de la hélice $\alpha 7$, de modo que los motivos GGQ se acoplen en el PTC. Véase, Korostelev, A. *et al.*, "Recognition of the amber UAG stop codon by release factor RF1", *EMBO Journal* (2010) 29, 2577-2585. La región del bucle de conmutación de RF1 corresponde a los aminoácidos 287 a 304 de SEQ ID NO: 1, y comprende la siguiente secuencia de aminoácidos: 287-QQAEASTRRNLLGSGDRS-304 (SEQ ID NO:4).

Una proteína "de tipo silvestre" es una proteína no modificada que tiene la secuencia de aminoácidos y/o actividad funcional de una proteína nativa o de origen natural. Por ejemplo, una proteína RF1 de tipo silvestre puede tener la secuencia de SEQ ID NO: 1. Una "proteína de control" puede ser una proteína de tipo silvestre o una proteína previamente modificada.

"Aminoácido nativo" se refiere a uno o más aminoácidos naturales codificados por el código genético. Un "aminoácido nativo endógeno" se refiere a un aminoácido nativo producido por las células hospedadoras usadas para generar el lisado.

"Aminoácidos no nativos" ("aann") se refiere a estructuras químicas que tienen la fórmula $\text{NH}_3\text{-(CR)-COOH}$, en la que R no es ninguno de los 20 sustituyentes más comunes que definen los aminoácidos naturales.

Los términos "idéntico" o porcentaje de "identidad", en el contexto de dos o más secuencias de ácidos nucleicos o polipéptidos, se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje específico de nucleótidos o restos de aminoácidos que son iguales (por ejemplo, el 60 % de identidad, opcionalmente, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % o 99 % de identidad en una región especificada, cuando se compara y alinea para obtenerse una correspondencia máxima en una ventana de comparación, o región designada, según lo medido usando los algoritmos BLAST y PSI-BLAST, que se describen en Altschul *et al.* (*J. Mol. Biol.* 215:403-10, 1990), y Altschul, *et al.* (*Nucleic Acids Res.*, 25:3389-3402, 1997), respectivamente. El programa informático para realizar los análisis BLAST está disponible al público a través del National Center for Biotechnology Information (véase, <http://www.ncbi.nlm.nih.gov>). Este algoritmo implica identificar en primer lugar pares de secuencias de alta valoración (HSP) identificando palabras cortas de longitud W en la secuencia de consulta, que coincide o satisface algún valor umbral de valor positivo T cuando se alinea con una palabra de la misma longitud en una secuencia de la base de datos. T se menciona como el umbral del valor de palabra adyacente (Altschul *et al.*, *supra*). Estos aciertos de palabra adyacente iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largos que los contengan. Los aciertos de palabra se prolongan en ambas direcciones a lo largo de cada secuencia todo el tiempo que pueda aumentarse el valor de alineación acumulativo. Los valores acumulativos se calculan usando, para secuencias nucleotídicas, los parámetros M (valor recompensa para un par de restos coincidentes; siempre > 0) y N (valor de penalización para restos no coincidentes; siempre < 0). Para secuencias de aminoácidos, se usa una matriz de valores para calcular el valor acumulativo. La prolongación de los aciertos de palabra en cada dirección se detiene cuando: el valor de alineación acumulativo queda fuera por una cantidad X de su valor conseguido máximo; el valor acumulativo llega a cero o por debajo, debido a la acumulación de una o más alineaciones de restos de valoración negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una expectativa (E) o 10, M = 5, N = -4 y una comparación de ambas hebras. Para las secuencias de aminoácidos, el programa BLASTP usa como valor predeterminado una longitud de palabra de 3, y una expectativa (E) de 10, y las alineaciones de la matriz de puntuación BLOSUM62 (Véase Henikoff y Henikoff, *Proc. Natl. Acad. Sci. EE.UU.* 89:10915, 1989) (B) de 50, expectativa (E) de 10, M = 5, N = -4 y una comparación de ambas hebras.

El "porcentaje de identidad de secuencia" se determina al comparar dos secuencias alineadas de manera óptima a lo largo de una ventana de comparación, en la que la parte de la secuencia polinucleotídica o polipeptídica en la ventana de comparación puede comprender adiciones o supresiones (es decir, huecos) en comparación con la secuencia de referencia (que no comprende adiciones o supresiones) para la alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las cuales se encuentra la base de ácido nucleico o el resto de aminoácido idénticos en ambas secuencias para producir el número de posiciones emparejadas, dividiendo el número

de posiciones coincidentes entre el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para producir el porcentaje de identidad de secuencia.

Una "ventana de comparación", como se usa en el presente documento, incluye la referencia a un segmento de una cualquiera del número de posiciones contiguas seleccionadas del grupo que consiste en de 20 a 600, por lo general, de aproximadamente 50 a aproximadamente 200, más habitualmente, de aproximadamente 100 a aproximadamente 150 en el que se puede comparar una secuencia con una secuencia de referencia del mismo número de posiciones contiguas una vez alineadas las dos secuencias de manera óptima. Los métodos de alineación de secuencias para la comparación son bien conocidos en la técnica.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, *Proc. Natl. Acad. Sci. EE.UU.* 90:5873-87, 1993). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña ($P(N)$), que proporciona una indicación de la probabilidad por la que sucedería una coincidencia entre dos secuencias de nucleótidos o aminoácidos por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es menor de aproximadamente 0,2, normalmente, menor de aproximadamente 0,01 y más normalmente menor de aproximadamente 0,001.

Cuando se usa el porcentaje de identidad de secuencia en referencia a un polipéptido, se reconoce que una o más posiciones de restos que no son idénticas pueden diferir en una sustitución conservativa de aminoácidos, en la que un primer resto de aminoácido es sustituido por otro resto de aminoácido que tiene propiedades químicas similares tales como una carga similar, o un carácter hidrófobo o hidrófilo similar, y, por lo tanto, no cambia las propiedades funcionales del polipéptido. Cuando las secuencias polipeptídicas difieren en sustituciones conservativas, el porcentaje de identidad de secuencia puede ajustarse hacia arriba para corregir la naturaleza conservativa de la sustitución. Dicho ajuste se puede hacer usando métodos bien conocidos, por ejemplo, la puntuación de una sustitución conservativa como un desapareamiento parcial más que completo, aumentando, por tanto, el porcentaje de identidad de secuencia. Por lo tanto, por ejemplo, cuando se asigna a un aminoácido idéntico una puntuación de 1 y se asigna a una sustitución no conservativa una puntuación de cero, se asigna a una sustitución conservativa una puntuación entre cero y 1. La puntuación de las sustituciones conservativas se puede calcular usando el algoritmo descrito en Pearson *et al.* (*Meth. Mol. Biol.* 24:307-331, 1994). La alineación también se puede realizar mediante el examen a simple vista y la alineación manual de las secuencias.

La expresión "variante modificada de forma conservativa", cuando se usa en referencia a una secuencia polinucleotídica en particular, se refiere a diferentes secuencias polinucleotídicas que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o cuando el polinucleótido no codifica una secuencia de aminoácidos, a secuencias prácticamente idénticas. A causa de la degeneración del código genético, un gran número de polinucleótidos funcionalmente idénticos codifican cualquier polipéptido dado. Por ejemplo, los codones CGU, CGC, CGA, CGG, AGA y AGG codifican, todos ellos, el aminoácido arginina. Por lo tanto, en cada posición en la que se especifica una arginina mediante un codón, el codón se puede alterar a cualquiera de los correspondientes codones descritos sin alterar el polipéptido codificado. Dichas variaciones de nucleótidos se denominan "variaciones silenciosas", que se pueden considerar un tipo de "variaciones modificadas de forma conservativa". Por tanto, se reconocerá que cada secuencia de polinucleótido desvelada en el presente documento como que codifica una variante de proteína también describe cada posible variación silenciosa. También se reconocerá que cada codón de un polinucleótido, excepto AUG, que es normalmente el único codón de la metionina, y UUG, que es normalmente el único codón para el triptófano, se puede modificar para producir una molécula funcionalmente idéntica mediante técnicas convencionales. Por consiguiente, cada variación silenciosa de un polinucleótido que no cambia la secuencia del polipéptido codificado se describe implícitamente en el presente documento.

Además, se reconocerá que las sustituciones, eliminaciones o adiciones individuales que alteran, añaden o eliminan un solo aminoácido o un pequeño porcentaje de aminoácidos (normalmente inferior al 10 %, y en general inferior al 1 %) en una secuencia codificada pueden considerarse variaciones modificadas de manera conservativa, siempre que la alteración resulte en la sustitución de un aminoácido con un aminoácido químicamente similar. Las sustituciones de aminoácidos conservativas de aminoácidos que proporcionan aminoácidos funcionalmente similares son bien conocidas en la técnica, incluidos los siguientes seis grupos, cada uno de los cuales contiene aminoácidos que se consideran sustitutos conservativos entre sí:

1) Alanina (Ala, A), Serina (Ser, S), Treonina (Thr, T);

2) Ácido aspártico (Asp, D), Ácido glutámico (Glu, E);

3) Asparagina (Asn, N), Glutamina (Gln, Q);

4) Arginina (Arg, R), Lisina (Lys, K)

5) Isoleucina (Ile, I), Leucina (Leu, L), Metionina (Met, M), Valina (Val, V); y

6) Fenilalanina (Phe, F), Tirosina (Tyr, Y), Triptófano (trp, W).

5 Dos o más secuencias de aminoácidos, o dos o más secuencias de nucleótidos se consideran "sustancialmente similares" si las secuencias de aminoácidos o las secuencias de nucleótidos comparten al menos el 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 99 % de identidad de secuencia entre sí, o con una secuencia de referencia en una ventana de comparación determinada. Dos o más proteínas también se consideran esencialmente similares si incorporan sustituciones conservativas de aminoácidos que proporcionan aminoácidos funcionalmente similares a la secuencia de aminoácidos.

10 "Promotor" se refiere a elementos de secuencia en el molde de ácido nucleico ubicados cadena arriba o cadena abajo desde el inicio de la transcripción. Las secuencias promotoras participan en el reconocimiento y en la unión de la ARN polimerasa y otras proteínas que inician la transcripción. Los ejemplos de promotores incluyen los promotores de bacteriófagos T7, SP6 y T3.

15 "La región no traducida 5'" se refiere a la secuencia de ácido nucleico situada cadena arriba o 5' del marco de lectura abierto de un molde de ácido nucleico. En el ARN, la región no traducida 5' precede al codón de inicio de la traducción presente en el molde de ácido nucleico. En el ADN, la región no traducida 5' se refiere a las secuencias de ácido nucleico que se transcriben al ARN y se ubican 5' en el codón de inicio de la traducción. En el ADN, el codón de inicio de la traducción normalmente es ATG, y en el ARN, el codón de inicio de la traducción normalmente es AUG.

20 La "secuencia primaria de aminoácidos" se refiere el orden de los restos de aminoácidos que están unidos entre sí por enlaces peptídicos en un polipéptido. El orden de los restos de aminoácidos, en general, se menciona comenzando en el extremo amino terminal del polipéptido y procediendo hacia el extremo carboxi terminal del polipéptido. La secuencia de aminoácidos primaria está determinada por la secuencia de nucleótidos de los codones de ARN en el molde de ácido nucleico.

25 "Marco de lectura abierto" se refiere a la secuencia de nucleótidos de un molde de ácido nucleico que se traduce a un polipéptido de interés. Como se usa en el presente documento, el marco de lectura abierto (ORF) puede incluir al menos un codón correspondiente a un resto de aminoácido definido que se une a un ARNt cargado con un aann. El al menos un codón puede ser un codón ámbar.

30 "ARNt sentido isoceptor" se refiere a diferentes especies de ARNt que se unen a codones alternativos para el mismo aminoácido.

35 "ARNt" o "ARN de transferencia" se refiere a una pequeña molécula de ARN que transfiere un aminoácido específico a una cadena polipeptídica en crecimiento en el sitio ribosomal de la síntesis de proteínas durante la traducción. A causa de la degeneración del código genético, un aminoácido puede asociarse con múltiples ARNt, mientras que cada tipo de molécula de ARNt solo puede asociarse con un tipo de aminoácido.

40 Los términos "polipéptido" y "proteína" se usan indistintamente, y se refieren a un compuesto que contiene dos o más aminoácidos unidos por enlaces peptídicos. Las proteínas pueden contener una o más cadenas polipeptídicas.

45 La expresión "cuando la proteína no está complejada" se refiere a la estructura de una proteína en solución, o una parte, una región o un dominio de la estructura de la proteína en solución, y no a la estructura de la proteína, o parte, región o dominio de la estructura de la proteína, cuando forma parte de un complejo con otras moléculas, tales como proteínas, ligandos, sustrato enzimático o complejos inhibidores, ribosomas, anticuerpos unidos o fragmentos de anticuerpos, ARN y ADN. El término también se refiere a la estructura de un dominio de proteína, tal como los aminoácidos que comprenden un motivo expuesto en la superficie, que está en solución.

50 Como se usa en el presente documento, El término "aproximadamente", cuando modifica cualquier cantidad, se refiere a la variación en la esa cantidad normalmente encontrada por un experto en la materia, por ejemplo, en la síntesis de proteínas o en los experimentos de cristalografía de rayos X. Por ejemplo, el término "aproximadamente" se refiere a la variación normal encontrada en las mediciones para una técnica analítica dada, tanto dentro como entre lotes o muestras. Por lo tanto, el término aproximadamente puede incluir una variación del 1-10 % del valor medido, tal como una variación del 5 % o 10 %. Las cantidades desveladas en el presente documento incluyen equivalentes a esas cantidades, Incluyendo las cantidades modificadas o no modificadas por el término "aproximadamente".

Descripción detallada de la invención

INTRODUCCIÓN

60 La presente divulgación describe proteínas diana que están modificadas de forma recombinante para incluir sitios de escisión de proteasa OmpT1 que pueden ser escindidos por OmpT1. El sitio de escisión por OmpT1 incluye un enlace peptídico escindible por OmpT1. En un caso, la proteína diana se modifica para incluir dos aminoácidos básicos adyacentes que se cargan positivamente a pH 7,0. En algunos casos, la proteína diana se modifica para introducir el sitio de escisión por OmpT1 en un motivo expuesto en la superficie de la proteína diana. En algunos casos, el enlace peptídico escindible por OmpT1 se sitúa en un motivo expuesto en la superficie que tiene un factor B de al menos

50 Å² cuando la proteína o el motivo expuesto en la superficie no están complejados (es decir, libres en solución y/o no forman parte de una estructura macromolecular que comprende otras moléculas). En algunos casos, el enlace peptídico escindible por OmpT1 está ubicado en un motivo expuesto en la superficie que tiene un área de superficie accesible al disolvente total (SASA) de entre aproximadamente 25 Å² y aproximadamente 225 Å². En algunos casos,

5 la proteína diana se modifica de manera que el enlace peptídico escindible por OmpT1 se ubique en una posición de la proteína que presente un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un diagrama de Ramachadran.

Como es entendido por un experto en la materia, lo anterior requiere información estructural para el/los dominio/s de proteína de interés. Sin embargo, los bucles de superficie suelen estar desordenados y, por lo tanto, no puede haber estructura cristalina disponible para la proteína diana de interés. Por lo tanto, en algunos casos, se puede crear un modelo de homología usando estructuras similares para predecir un SASA de entre aproximadamente 25 Å² y aproximadamente 225 Å², o un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un diagrama de Ramachadran, en ausencia de información estructural.

15 Se describe una célula bacteriana que expresa tanto OmpT1 como la proteína diana modificada de forma recombinante. En algunos casos, la proteína diana modificada de forma recombinante es una proteína esencial, por ejemplo, una proteína que se requiere para el crecimiento y/o la supervivencia celular normal.

20 Las proteínas diana descritas en el presente documento se seleccionan debido a que tienen actividad perjudicial en un sistema de síntesis libre de células *in vitro*. Por ejemplo, la proteína diana puede inhibir la transcripción y/o la traducción de moldes de ácido nucleico que codifican proteínas de interés. Por lo tanto, se describen métodos para reducir la actividad perjudicial de una proteína diana esencial modificada en un sistema de síntesis libre de células *in vitro* mediante la inactivación de la proteína diana con una proteasa OmpT1. En algunos casos, el método comprende cultivar una bacteria positiva en OmpT1 que expresa la proteína diana esencial modificada, en el que la proteína diana se modifica para incluir un sitio de escisión por OmpT1 que comprende un enlace peptídico escindible por OmpT1 en un motivo expuesto en la superficie; lisar las bacterias para crear un extracto de síntesis libre de células; poner en contacto la proteína diana esencial modificada con OmpT1 en una cantidad suficiente para reducir la cantidad de proteína diana intacta en un 50 %; añadir un molde de ácido nucleico al extracto, en el que el molde codifica una proteína de interés; y dejar que el sistema de síntesis libre de células produzca la proteína de interés. El método aprovecha la separación espacial de la proteína diana modificada y la proteasa OmpT1 durante el crecimiento celular, cuando la proteína diana puede ser importante para el crecimiento celular y/o la supervivencia, mientras que la proteína diana es escindida por OmpT1 cuando las células bacterianas se rompen y se lisan para producir un extracto libre de células.

35 En algunos casos del método, el enlace peptídico escindible por OmpT1 se encuentra dentro de un motivo expuesto en la superficie que tiene un factor B de al menos 50 Å² cuando la proteína o el motivo expuesto en la superficie no está complejado. Por ejemplo, el motivo expuesto en la superficie puede tener un factor B de 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190 o 200 Å² o superior, o puede estar en el intervalo de entre aproximadamente 50 y aproximadamente 200 Å², entre aproximadamente 50 y aproximadamente 150 Å², entre aproximadamente 50 y aproximadamente 100 Å², entre aproximadamente 60 y aproximadamente 200 Å², entre aproximadamente 60 y aproximadamente 150 Å², o entre aproximadamente 60 y aproximadamente 100 Å². En algunos casos del método, el enlace peptídico escindible por OmpT1 está ubicado dentro de un motivo expuesto en la superficie que tiene un área superficial accesible al disolvente total de entre aproximadamente 25 Å² y aproximadamente 225 Å². Por ejemplo, en algunos casos, el área de superficie accesible al disolvente total es de aproximadamente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220 o 225 Å² (cuando el término "aproximadamente" modifica cada uno de los valores precedentes). En un caso del método, el enlace peptídico escindible por OmpT1 se encuentra en una posición de la proteína que muestra un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un gráfico de Ramachadran. Como se ha descrito anteriormente, en ausencia de información estructural, se puede crear un modelo de homología usando estructuras similares para predecir un SASA de entre aproximadamente 25 Å² y aproximadamente 225 Å², o un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un diagrama de Ramachadran.

Las proteínas de interés mencionadas anteriormente incluyen proteínas que se han diseñado para incorporar aminoácidos no nativos (aann) en una ubicación definida de la secuencia de aminoácidos. La introducción de aann en proteínas puede dar lugar a proteínas con propiedades preferidas. Un método para introducir aann en proteínas o polipéptidos de interés usa ARNt ortogonales aminoacilados que reconocen un codón ámbar (parada) para introducir el aann en la cadena polipeptídica naciente durante la traducción de la proteína. Sin embargo, el rendimiento de las proteínas que tienen aann introducidos en las mismas puede ser reducido por las proteínas del complejo de terminación de la traducción que facilitan la terminación de la traducción al reconocer la terminación o los codones de parada en una secuencia de ARNm. El factor de liberación 1 (RF1) forma parte del complejo de terminación, y reconoce el codón de parada UAG (ámbar). El reconocimiento de RF1 del codón ámbar puede potenciar la terminación prematura de la cadena en el sitio de incorporación del aann, lo que reduce el rendimiento de las proteínas deseadas. Los métodos descritos en el presente documento resuelven este problema al disminuir la actividad funcional de RF1 en lisados de células bacterianas. Por lo tanto, en algunas realizaciones, la proteína diana esencial es RF1. La actividad funcional de RF1 disminuye al introducir sitios de escisión de la proteasa OmpT1 en RF1. OmpT1 es una enzima ubicada en la membrana celular externa de las bacterias intactas. Por lo tanto, la RF1 modificada no está

disponible como un sustrato para OmpT1 en las células intactas. Cuando la membrana bacteriana se rompe, por ejemplo, como en un lisado bacteriano, la enzima OmpT1 es capaz de entrar en contacto y escindir la RF1 modificada en el sitio de escisión introducido, disminuyendo así la actividad funcional de RF1. Los métodos son aplicables a otros factores de liberación encontrados en el complejo de terminación de la traducción, tales como RF2, que reconoce el codón de parada UGA. Los métodos también son aplicables a proteínas que degradan el ARNm durante la traducción *in vitro*, tales como las enzimas RNasa. Además, los métodos son en general aplicables a cualquier proteína de un extracto bacteriano que inhiba la transcripción y/o la traducción de una proteína de interés.

MÉTODOS GENERALES

Salvo que se defina de otro modo, todos los términos técnicos y científicos usados en este documento tienen el significado comprendido habitualmente por un experto en la materia a la que pertenece esta invención. Los expertos se dirigen en particular a Green, M. R. y Sambrook, J., eds., "Molecular Cloning: A Laboratory Manual", 4ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (2012), y Ausubel, F. M., *et al.*, "Current Protocols in Molecular Biology" (Suplemento 99), John Wiley & Sons, Nueva York (2012) para las definiciones y los términos de la técnica. También aparecen métodos convencionales en Bindereif, Schön, y Westhof (2005) "Handbook of RNA Biochemistry", Wiley-VCH, Weinheim, Alemania, que describe métodos detallados para la manipulación y el análisis del ARN. Los ejemplos de técnicas moleculares apropiadas para generar ácidos nucleicos recombinantes y las instrucciones suficientes para dirigir a los expertos a través de numerosos ejercicios de clonación se encuentran en Green, M. R. y Sambrook, J., (*Id.*); Ausubel, F. M., *et al.*, (*Id.*); Berger y Kimmel, "Guide to Molecular Cloning Techniques", *Methods in Enzymology* (Volumen 152 Academic Press, Inc., San Diego, Calif. 1987); y "PCR Protocols: A Guide to Methods and Applications" (Academic Press, San Diego, Calif. 1990).

Los métodos para la purificación de proteínas, cromatografía, electroforesis, centrifugación y cristalización se describen en Coligan *et al.* (2000) "Current Protocols in Protein Science", Vol. 1, John Wiley and Sons, Inc., Nueva York. Los métodos para la síntesis libre de células se describen en Spirin y Swartz (2008) "Cell-free Protein Synthesis", Wiley-VCH, Weinheim, Alemania. Los métodos para la incorporación de aminoácidos no nativos a proteínas usando síntesis libre de células se describen en Shimizu *et al.* (2006) *FEBS Journal*, 273, 4133-4140.

Los métodos de amplificación por PCR son bien conocidos en la técnica, y se describen, por ejemplo, en Innis *et al.*, "PCR Protocols: A Guide to Methods and Applications", Academic Press Inc. San Diego, Calif., 1990. Una reacción de amplificación normalmente incluye el ADN que se va a amplificar, una ADN polimerasa termoestable, dos cebadores oligonucleótidos, trifosfatos de desoxinucleótidos (dNTP), tampón de reacción y magnesio. Por lo general, un número deseable de ciclos térmicos está entre 1 y 25. Los métodos para el diseño de cebadores y la optimización de las condiciones de PCR son bien conocidos en la técnica y se pueden encontrar en textos de Biología molecular convencionales, tales como Ausubel *et al.*, "Short Protocols in Molecular Biology", 5ª edición, Wiley, 2002 e Innis *et al.*, "PCR Protocols", Academic Press, 1990. Los programas informáticos son útiles en el diseño de cebadores con la especificidad requerida y las propiedades de amplificación óptimas (por ejemplo, Oligo Versión 5.0 (National Biosciences)). En algunas realizaciones, los cebadores de PCR pueden contener además sitios de reconocimiento para endonucleasas de restricción, para facilitar la inserción del fragmento de ADN amplificado en sitios de enzimas de restricción específicos en un vector. Si se añaden sitios de restricción al extremo 5' de los cebadores de PCR, es preferible incluir unas cuantas bases 5' adicionales (por ejemplo, dos o tres) para permitir una escisión más eficaz por parte de la enzima. En algunas realizaciones, los cebadores de PCR también pueden contener un sitio promotor de ARN polimerasa, tal como T7 o SP6, para permitir la posterior transcripción *in vitro*. Los expertos en la materia conocen bien los métodos para la transcripción *in vitro* (véase, por ejemplo, Van Gelder *et al.*, *Proc. Natl. Acad. Sci. EEUU* 87:1663-1667, 1990; Eberwine *et al.*, *Proc. Natl. Acad. Sci. EEUU* 89:3010-3014, 1992).

PROTEÍNAS ESCINDIBLES POR OMP T1

La presente divulgación describe proteínas diana que pueden ser escindidas e inactivadas por enzimas proteolíticas tales como OmpT1. Para ser escindidas por la proteasa, las proteínas se modifican para introducir sitios de escisión proteolítica en la proteína. En algunos casos, el sitio de escisión introducido en la proteína diana modificada es un enlace peptídico escindible por OmpT1, y el enlace escindible por OmpT1 se introduce en un motivo expuesto en la superficie de la proteína diana. El sitio de escisión reconocido por OmpT1 comprende dos restos básicos adyacentes. Por lo tanto, las proteínas diana descritas en el presente documento se modifican para comprender dos restos básicos adyacentes que se carguen positivamente a pH 7,0.

Modificación de proteínas para introducir un sitio de escisión por OmpT1 en un motivo expuesto en la superficie que tiene un factor B de al menos 50 Å²

5 En algunos casos, la proteína diana se modifica para introducir el enlace peptídico escindible por OmpT1 en un motivo
 expuesto en la superficie que tenga un factor B de al menos 50 Å² cuando la secuencia de aminoácidos de la proteína
 diana o de la región del motivo no esta complejada. La ubicación de un motivo expuesto en la superficie con un factor
 B de al menos 50 Å² se puede determinar de varias maneras. Los factores B se dan para cada átomo en archivos
 10 cristalográficos del banco de datos de proteínas (PDB). El factor B de un motivo expuesto en la superficie se puede
 calcular usando el algoritmo GetArea incorporado en el software de modelización molecular pymol (véase en Internet,
 en pymolwiki.org/index.php/Get_Area). Como alternativa, si no hay disponible una estructura cristalográfica de rayos
 X de la proteína, pero hay disponible una estructura de RMN de la proteína, se puede usar el índice de
 superenrollamiento aleatorio (RCI) en lugar de un factor B alto (véase Berjanskii, M. V., *et al.*, "Application of the
 random coil index to studying protein flexibility", *J Biomol NMR*. enero de 2008; 40(1):31-48. Epub 6 de noviembre de
 2007).

15 Con el fin de modificar una proteína diana para introducir un sitio de escisión por OmpT1 en un motivo expuesto en la
 superficie que tenga un factor B de al menos 50 Å², se pueden cartografiar los factores B desde las estructuras
 cristalográficas de rayos X de la proteína hasta la secuencia de aminoácidos, y la secuencia modificada de forma
 recombinante para incluir un enlace peptídico escindible por OmpT1 en el/los sitio/s con factores B altos. A
 20 continuación, se puede analizar la proteína diana modificada de forma recombinante para determinar la degradación
 de la proteína en un extracto celular que contenga OmpT1. En algunos casos, el factor B es de al menos 50, 60, 70,
 80, 90 o 100 Å². En algunos casos, el factor B está entre aproximadamente 50 y aproximadamente 200 Å², o entre
 aproximadamente 50 y aproximadamente 150 Å², o entre aproximadamente 50 y 120 Å². Se entiende que los
 intervalos descritos en el presente documento incluyen todos los valores entre los puntos finales.

Modificación de proteínas para introducir un sitio de escisión por OmpT1 en un motivo expuesto en la superficie que tiene un área de superficie accesible al disolvente total de entre aproximadamente 25 Å² y aproximadamente 225 Å²

30 En algunos casos, la proteína diana se modifica para introducir el enlace peptídico escindible por OmpT1 en un motivo
 expuesto en la superficie que tenga un área superficial accesible al disolvente total de entre aproximadamente 25 Å²
 y aproximadamente 225 Å². La ubicación de un motivo expuesto en la superficie que tiene un área de superficie
 accesible al disolvente total de entre aproximadamente 25 Å² y aproximadamente 225 Å² se puede determinar usando
 métodos conocidos en la técnica. Por ejemplo, se puede usar el conocido software GETAREA para ubicar vértices
 35 expuestos a disolventes de átomos que interseccionan entre sí, como se describe en Fraczekiewicz, R. y Braun, W.,
 "Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules", *J.*
Comp. Chem., 19, 319-333 (1998). Por lo tanto, se puede calcular el área de superficie accesible al disolvente (energía
 de solvatación) de una molécula de proteína usando el software GETAREA, mediante la introducción de las
 coordenadas atómicas en formato PDB (base de datos de proteínas), especificando el radio deseado de la sonda de
 40 agua y especificando el nivel deseado de salida usando un formulario proporcionado en Internet, en el sitio web
curie.utmb.edu/getarea.html de la Universidad de Texas (parámetros: radio de la sonda de agua = 1,4). Otros métodos
 para determinar el área de superficie accesible al disolvente total se describen en Eisenberg, D. y McLachlan, A. D.
 (1986) *Nature*, 319, 199; Markley, J. L.; *et al.* (1998) *Pure & Appl. Chem.*, 70, 117; y Wesson, L. y Eisenberg, D. (1992)
Protein Sci., 1, 227.

Modificación de proteínas para introducir un sitio de escisión por OmpT1 una posición de la proteína que muestra un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un diagrama de Ramachandran

50 En algunos casos, la proteína diana se modifica para introducir el enlace peptídico escindible por OmpT1 en una
 posición de la proteína que presente un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un diagrama de
 Ramachandran. La ubicación de la posición en una proteína que muestra un ángulo Phi de 0 ° a -180 ° o un ángulo Psi
 de 0 ° a +180 ° en un diagrama de Ramachandran se puede determinar usando métodos conocidos en la técnica. Los
 métodos para calcular los ángulos Phi y Psi se describen, por ejemplo, en Lovell, S. C. *et al.*, "Proteins: Structure,
 Function, and Genetics", 50:437-450 (2003). Se pueden determinar los ángulos Phi y Psi en un diagrama de
 55 Ramachandran cargando un archivo de coordenadas o un archivo PDB en el servidor MOLPROBITY en Internet, en
kinemage.biochem.duke.edu (véase Chen, V. B., *et al.*, (2010) "MolProbity: all-atom structure validation for
 macromolecular crystallography". *Acta Crystallographica D*66: 12-21.). Como alternativa, se puede usar
 Ramachandran Plot Explorer disponible en Internet, en boscoh.com/ramaplot.

60 Las proteínas diana descritas en el presente documento se seleccionan debido a que pueden inhibir la producción de
 proteínas en un sistema de síntesis libre de células. Por lo tanto, en algunos casos, las proteínas modificadas descritas
 en el presente documento disminuyen la producción de proteínas de longitud completa en extractos bacterianos libres
 de células. Por ejemplo, las proteínas RF1 y RF2 forman parte del complejo de terminación que reconoce un codón
 65 de parada en el ARNm y termina la traducción de la cadena polipeptídica. La terminación prematura de la traducción
 es indeseable cuando se incorporan aminoácidos no nativos a una proteína usando sistemas de traducción libres de
 células como los descritos en el presente documento. La reducción del reconocimiento del codón de parada por el

complejo de terminación puede aumentar el rendimiento de las proteínas que incorporan aann. Por lo tanto, en algunos casos, la proteína RF1 y/o RF2, o un homólogo funcional de la misma, se modifica para contener sitios de escisión por OmpT1 en un motivo expuesto en la superficie que tenga un factor B de al menos 50 Å² cuando la proteína RF1 y/o RF2, u homólogo funcional de las mismas, está sin complejar. En algunos casos, la proteína RF1 y/o RF2, o un homólogo funcional de la misma, se modifican para contener sitios de escisión por OmpT1 en un motivo expuesto en la superficie que tiene un área de superficie accesible al disolvente total de entre aproximadamente 25 Å² y aproximadamente 225 Å². En algunos casos, la proteína RF1 y/o RF2, o un homólogo funcional de la misma, se modifican para contener sitios de escisión por OmpT1 en una posición de la proteína que muestra un ángulo Phi de 0 ° a -180 ° o un ángulo Psi de 0 ° a +180 ° en un diagrama de Ramachadran. En algunos casos, la proteína funcional modificada es esencialmente similar a RF1 (SEQ ID NO: 1) o RF2 (SEQ ID NO: 2).

La proteína RF1 incluye la proteína RF1 de tipo silvestre prototipo (SEQ ID NO: 1) de *E. coli*, así como las variaciones polimórficas y las muteínas creadas de forma recombinante. Las proteínas RF1 se definen como aquellas que tienen esencialmente la misma actividad biológica o capacidad funcional que el tipo silvestre (por ejemplo, al menos el 80 % de cualquiera de ellas), tienen al menos el 60 %, 70 %, 80 %, 90 % o 95 % de identidad de secuencia con la proteína RF1 prototipo y/o se unen a anticuerpos policlonales generados contra la proteína prototipo SEQ ID NO: 1.

Con respecto a la unión de una proteína RF1 a anticuerpos policlonales, la proteína RF1 se unirá en condiciones de inmunoensayo designadas a los anticuerpos especificados con una especificidad al menos del doble que la del fondo, donde los anticuerpos no se unen esencialmente en una cantidad significativa a otras proteínas presentes en la muestra. Por ejemplo, los anticuerpos policlonales generados contra RF1 (SEQ ID NO: 1), o isoformas o partes de la misma, se pueden seleccionar para obtener solamente aquellos anticuerpos que sean específicamente inmunorreactivos con RF1 y no con otras proteínas, excepto las variantes polimórficas de RF1. Se puede usar una variedad de formatos de inmunoensayo para seleccionar anticuerpos que son específicamente inmunorreactivos con una proteína en particular. Por ejemplo, los inmunoensayos de ELISA en fase sólida se utilizan de forma rutinaria para seleccionar anticuerpos que son específicamente inmunorreactivos con una proteína (véase, por ejemplo, Harlow y Lane, "Antibodies, A Laboratory Manual" (1988) para una descripción de formatos de inmunoensayo y de las condiciones que se pueden usar para determinar la inmunorreactividad específica). Normalmente, una reacción específica o selectiva será al menos el doble de la señal o ruido de fondo y, más normalmente, de más de 10 a 100 veces superior a la de fondo.

Como se describe en los ejemplos, RF1 se modificó satisfactoriamente introduciendo sitios de escisión por OmpT1 funcionales en un motivo expuesto en la superficie. Sin embargo, la introducción de sitios putativos de escisión por OmpT1 dibásicos en RF1 no siempre dio lugar a una proteína que pudiera ser escindida eficazmente por OmpT1. Por ejemplo, se introdujeron sustituciones de aminoácidos simples M74R, E76K, E84K, A85R, E87R, E108R, T293R y S304K en las regiones de bucle de RF1 junto a una Arg o Lys existente, creando así sitios de escisión dibásicos. Sin embargo, estas variantes no se escindieron de manera eficaz cuando las variantes de RF1 se expresaron en extractos de células positivos en OmpT. Por lo tanto, la presente invención proporciona el resultado inesperado de que proteínas tales como RF1 pueden modificarse de manera que la proteína pueda ser escindida por una proteasa tal como OmpT1.

Como entenderán los expertos en la materia, el sitio de escisión introducido puede ser escindido por una enzima con actividad enzimática de tipo OmpT1, de manera que la actividad enzimática proceda de un homólogo funcional o fragmento de OmpT1, o de una proteína modificada para tener actividad de tipo OmpT1.

45 **Escisión de proteínas no modificadas por OmpT1**

Las proteínas descritas en el presente documento también pueden contener sitios de escisión de la proteasa OmpT1 endógenos. Por ejemplo, en algunas realizaciones, la proteína no modificada o de tipo silvestre contiene una secuencia de aminoácidos dibásica que comprende un enlace peptídico escindible por OmpT1 que es escindible por OmpT1. La escisión de la proteína por OmpT1 se puede ensayar incubando la proteína purificada con extractos bacterianos libres de células que expresan OmpT1, y comparando la cantidad de escisión con la cantidad de escisión detectada en extractos libres de células que no expresan OmpT1. En algunos casos, las proteínas no modificadas que contienen sitios de escisión por OmpT1 endógenos son necesarias o se requieren para el crecimiento o la función normal de la célula, pero inhiben la traducción de las proteínas en extractos sin células. Por lo tanto, la divulgación describe métodos para inactivar selectivamente proteínas no modificadas mediante la escisión con OmpT1, en los que el momento de inactivación puede controlarse de manera que las proteínas sean funcionales durante el crecimiento celular, pero se inactiven en extractos celulares que expresen OmpT1.

60 **Molde**

Para producir las proteínas, se necesita un molde de ácido nucleico. Los moldes se usan para producir las proteínas modificadas para comprender un sitio de escisión escindible por OmpT1. Los moldes también se usan para producir proteínas de interés que se expresan en sistemas libres de células. Los moldes para la síntesis de proteínas libres de células pueden ser ARNm o ADN. El molde puede comprender secuencias para cualquier gen particular de interés, y puede codificar un polipéptido de longitud completa o un fragmento de cualquier longitud del mismo. Los ácidos nucleicos que sirven como moldes de síntesis de proteínas se derivan opcionalmente de una fuente natural, o pueden

ser sintéticos o recombinantes. Por ejemplo, los ADN pueden ser ADN recombinantes, por ejemplo, plásmidos, virus o similares.

En un caso, el molde de ADN comprende un ORF que codifica una proteína diana que comprende un sitio de escisión por OmpT1. Por ejemplo, el ORF puede codificar una proteína diana modificada que se requiere para el crecimiento y la supervivencia celulares normales, pero cuya actividad disminuye el rendimiento de las proteínas en extractos libres de células, ya sea directa o indirectamente. Por lo tanto, el ORF puede codificar un componente proteico del complejo de terminación, tal como RF1 o RF2, o variantes modificadas de las mismas. El ORF también puede codificar una RNasa que degrade los moldes de ARN.

En otro caso, el molde de ADN comprende un ORF que codifica una proteína de interés que se ha modificado para incorporar un aminoácido no nativo. Por lo tanto, el ORF puede codificar cualquier proteína que tenga importancia biológica que lleve incorporado un aann en una posición definida de la secuencia de aminoácidos. Los ejemplos no limitantes de dichas proteínas de interés incluyen anticuerpos, hormonas, citocinas y proteínas víricas. La divulgación usa codones sentido para la incorporación de aminoácidos no nativos, y elude el requisito de componentes ortogonales como se encuentra comúnmente en la técnica. En estos casos, el ORF comprende al menos un codón sentido isoaceptor. El ORF comprende además un codón correspondiente a un resto de aminoácido definido que reconoce un ARNt cargado con un aminoácido no nativo. En algunos casos, el ORF comprende un codón ámbar (UAG) que se une a un ARNt cargado con un aminoácido no nativo. En otros casos, el molde es capaz de traducir una proteína completa y funcional independientemente de si se escogen aminoácidos no nativos para incorporarlos a la proteína de interés.

Un molde de ADN que comprende el ORF de interés se unirá operativamente a al menos un promotor y a una o más otras secuencias reguladoras incluyendo, sin limitación, represores, activadores, potenciadores de la transcripción y la traducción, proteínas de unión al ADN, etc. Se pueden producir cantidades adecuadas del molde de ADN para su uso en el presente documento amplificando el ADN en vectores de clonación y huéspedes conocidos, o mediante la reacción en cadena de la polimerasa (PCR).

El molde de ADN puede comprender además el ORF de interés unido en fase a secuencias de ácidos nucleicos que codifican secuencias de aminoácidos que son útiles para aislar y purificar la proteína expresada, tales como, por ejemplo, marcadores de poli-aminoácido que se unen con alta afinidad a los medios cromatográficos. El marcador de poli-aminoácido se puede ubicar en el extremo 5' o en el extremo 3' del ORF, dando lugar a un marcador amino-terminal o carboxilo-terminal en la proteína expresada, respectivamente. En un caso, el ORF se une en fase a las secuencias que codifican un marcador de poli-histidina.

Se describen usos de un lisado bacteriano. Se puede construir un molde de ADN para la expresión bacteriana uniendo operativamente un ADN codificante de la proteína deseada tanto a una secuencia promotora como a un sitio de unión al ribosoma bacteriano (secuencia de Shine-Delgarno). Los promotores adecuados para su uso con el molde de ADN en los métodos de traducción-transcripción libres de células descritos incluyen cualquier secuencia de ADN capaz de potenciar la transcripción *in vivo* en las bacterias de las que se deriva el extracto bacteriano. Se prefieren los promotores que son capaces de iniciar eficazmente la transcripción dentro de la célula hospedadora. El ADN que codifica la proteína deseada y el ADN que contiene el promotor deseado y las secuencias de Shine-Dalgarno (SD) se pueden preparar mediante una variedad de métodos conocidos en la técnica. Como alternativa, las secuencias de ADN deseadas se pueden obtener a partir de clones existentes o, si no hay ninguno disponible, mediante la selección en genotecas de ADN y la construcción de las secuencias de ADN deseadas a partir de los clones de las genotecas.

Se pueden producir convenientemente moldes de ARN que codifiquen la proteína de interés a partir de una célula hospedadora recombinante transformada con un vector construido para expresar un ARNm con un sitio de unión al ribosoma bacteriano (secuencia SD) unido operativamente a la secuencia codificante del gen deseado, de manera que los ribosomas en la mezcla de reacción puedan unirse y traducir dicho ARNm. Por lo tanto, el vector porta cualquier promotor capaz de potenciar la transcripción del ADN en la célula hospedadora particular usada para la síntesis del molde de ARN.

Debido a que es difícil extraer el ARN no degradado de las bacterias, se prefiere el cultivo de células eucariotas superiores para la producción del molde de ARN. En principio, cualquier cultivo de células eucariotas superiores es viable, incluyendo los cultivos de células de vertebrados e invertebrados. El molde de ARN puede aislarse convenientemente en una fracción de ARN celular total extraída del cultivo de la célula hospedadora. El ARN celular total se puede aislar del cultivo de la célula hospedadora mediante cualquier método conocido en la técnica. El molde de ARN deseado puede aislarse junto con la mayoría del ARNm celular si el molde de ARN está diseñado para contener en su extremo 3' una señal de poliadenilación reconocida por la célula hospedadora eucariota. Por lo tanto, la célula hospedadora producirá el molde de ARN con una cola de poliadenilato (poli (A)). Los ARNm poliadenilados se pueden separar del grueso del ARN celular mediante cromatografía de afinidad en columnas de oligodesoxitimidilato (oligo(dT))-celulosa usando cualquier método conocido en la técnica. Si se conoce el tamaño del ARNm que codifica la proteína deseada, el preparado del ARNm puede purificarse más para moléculas de ARNm del tamaño particular mediante la electroforesis en gel de agarosa del ARN.

Expresión de proteínas modificadas que tienen sitios de escisión por OmpT1

Una vez producido el molde de ácido nucleico, el molde se usa para expresar la proteína diana recombinante que comprende un sitio de escisión por OmpT1 en una célula, o para sintetizar una proteína diana recombinante modificada en un sistema de traducción libre de células. Por ejemplo, el molde se puede añadir a un lisado celular en condiciones
 5 suficientes para traducir el molde en proteína. El lisado celular puede ser de células bacterianas o de células eucariotas. La proteína expresada se puede purificar luego usando métodos conocidos en la técnica, como se describe más adelante.

Purificación de proteínas para ensayar la actividad

10 Las proteínas que contienen sitios de escisión por OmpT1 se pueden purificar como es convencional en la técnica. Las proteínas se pueden recuperar y purificar mediante métodos que incluyen, pero sin limitación, precipitación con sulfato de amonio o etanol, extracción con ácido o base, cromatografía en columna, cromatografía en columna de afinidad, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de hidroxilapatita, cromatografía de lectina, electroforesis en gel, etc. Por ejemplo,
 15 las proteínas que comprenden marcadores de histidina N-terminal pueden purificarse usando medios de afinidad que contienen iones metálicos tales como el níquel y el cobalto. El medio de afinidad se lava luego para eliminar la proteína no unida, y las proteínas unidas se eluyen y se recuperan. En algunos casos, las proteínas marcadas con poli-histidina se purifican usando cromatografía de afinidad por metales inmovilizados (IMAC). Las proteínas modificadas también
 20 pueden purificarse usando cromatografía líquida de alto rendimiento (HPLC), u otros métodos adecuados cuando se desee una alta pureza. En el Ejemplo 1, se proporciona un método de purificación preferido.

Tras la purificación, las proteínas que contienen sitios de escisión por OmpT1 pueden poseer una configuración diferente de las configuraciones deseadas de los polipéptidos relevantes. Por lo tanto, las proteínas purificadas pueden someterse a condiciones que den lugar a la configuración de proteína preferida. Se conoce una diversidad de métodos
 25 de purificación/plegamiento de proteínas en la técnica, por ejemplo, Deutscher, *Methods in Enzymology* Vol. 182: "Guide to Protein Purification" (Academic Press, Inc. N.Y. 1990); Bollag *et al.*, "Protein Methods", 2.^a edición, (Wiley-Liss, N. Y. 1996). En general, es en ocasiones deseable desnaturalizar y reducir los polipéptidos expresados y después hacer que los polipéptidos se replieguen en su configuración preferida. Por ejemplo, guanidina, urea, DTT, DTE y/o
 30 una chaperona pueden añadirse a un producto de traducción de interés. Los métodos de reducción, desnaturalización y renaturalización de proteínas son bien conocidos por los expertos en la materia. Véase, por ejemplo, Debinski *et al.*, *J. Biol. Chem.* 268:14065-70 (1993); Buchner *et al.*, *Anal. Biochem.* 205:263-70 (1992).

Confirmación de la actividad biológica funcional de proteínas modificadas que tienen sitios de escisión por OmpT1

35 Dependiendo del uso deseado de las proteínas modificadas descritas en el presente documento, puede ser importante para las proteínas modificadas tener actividad biológica comparable a una proteína no modificada de tipo silvestre. Por ejemplo, si la proteína que se va a modificar es importante para el crecimiento normal de las bacterias, es deseable
 40 mantener los niveles de actividad normales de la proteína hasta que las células se lisen para producir el lisado libre de células para la traducción. Por lo tanto, los métodos de la divulgación proporcionan proteínas modificadas que contienen sitios de escisión por OmpT1 que tienen una actividad biológica comparable a la proteína nativa o de tipo silvestre. Las proteínas modificadas que conservan los niveles de actividad de tipo silvestre, o niveles de actividad que son comparables a los del tipo silvestre, se denominan en el presente documento proteínas funcionales. Se puede
 45 determinar la actividad específica de una proteína al determinar el nivel de actividad en un ensayo funcional. Como alternativa, se puede determinar la actividad específica de una proteína mediante la cuantificación de la cantidad de proteína presente en un ensayo no funcional, por ejemplo, inmunotinción, ELISA, cuantificación en gel teñido con coomasie o plata, etc., y determinando la proporción de proteína biológicamente activa con respecto a la proteína total. En general, una proteína modificada es comparable a una proteína de tipo silvestre si la actividad específica definida
 50 de este modo es al menos aproximadamente el 50 % de la proteína de tipo silvestre, o al menos aproximadamente el 60 %, aproximadamente 70 %, aproximadamente 80 %, aproximadamente el 90 % o superior de la de la proteína de tipo silvestre. Véase, por ejemplo, Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual" (Cold Spring Harbor Press, Cold Spring Harbor, N. Y. 1989).

55 La actividad funcional de proteínas modificadas se puede ensayar en una variedad de maneras. Por ejemplo, se puede comparar la velocidad de crecimiento de las bacterias que expresan la proteína modificada con la velocidad de crecimiento de las bacterias que expresan una proteína de tipo silvestre, de control o no modificada. La actividad funcional de las proteínas modificadas también se puede ensayar determinando la capacidad de la proteína modificada para terminar la traducción en un codón de parada en el molde de ARNm. La terminación de la traducción da lugar a
 60 proteínas truncadas y, por lo tanto, las actividades relativas de las proteínas modificadas pueden cuantificarse midiendo la cantidad de proteína truncada y de longitud completa, y comparando la proporción de proteína truncada con respecto a la de longitud completa con la resultante de la proteína de tipo silvestre, como se describe en los ejemplos y se muestra en la Figura 1. Una proteína modificada tiene una actividad comparable a una proteína de tipo silvestre si la proporción de proteína truncada con respecto a la de longitud completa es al menos aproximadamente
 65 un 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, o mayor de la proporción de proteína truncada con respecto a la de longitud completa usando una proteína de tipo silvestre.

Confirmación de que las proteínas modificadas son escindibles por OmpT1

Una vez determinado que las proteínas modificadas tienen una actividad biológica comparable a la de la proteína de tipo silvestre o no modificada, las proteínas modificadas se ensayan para determinar que son escindidas por OmpT1. Un método para ensayar que las proteínas modificadas son escindidas por OmpT1 es la adición de proteínas recombinantes que contienen sitios de escisión por OmpT1 a un extracto libre de células que contiene OmpT1 funcional. Si la proteína modificada es escindida por OmpT1, migrará a un peso molecular aparentemente más bajo que la proteína intacta durante la electroforesis en gel (por ejemplo, SDS-PAGE). La proteína modificada puede detectarse incluyendo un marcador radiactivo tal como ^{14}C en la reacción de traducción en condiciones adecuadas para la incorporación del marcador radiactivo a la proteína modificada. La migración de la proteína marcada radiactivamente se puede visualizar, por ejemplo, en una autorradiografía del gel. La escisión de la proteína modificada por OmpT1 también se puede detectar mediante análisis de transferencia Western, por ejemplo, transfiriendo las proteínas desde el gel a un soporte sólido, poniendo en contacto el soporte con anticuerpos que se unen a la proteína intacta y escindida, y visualizando los anticuerpos unidos usando un marcador detectable.

La actividad de escisión por OmpT1 se puede determinar mediante la comparación de la cantidad o velocidad de escisión de una proteína modificada por OmpT1 con la de una proteína no modificada. Por ejemplo, la velocidad de escisión de la proteína modificada por OmpT1 puede ser superior al 50 % de la velocidad de escisión de la proteína de tipo silvestre en condiciones específicas. En algunos casos, la velocidad de escisión de una proteína modificada descrita en el presente documento es superior al 50 % de la velocidad de escisión de la proteína de tipo silvestre tras 30 minutos a 30 °C cuando las proteínas modificadas y de tipo silvestre están presentes a una concentración similar (por ejemplo, una concentración de 0,1-1,0 micromolar) en un extracto libre de células de bacterias que expresan OmpT1. En algunos casos, más del 90 % de la proteína modificada es escindida por OmpT1 tras 60 minutos a 30 °C cuando la proteína modificada está presente a una concentración de 0,1-1,0 micromolar en un extracto libre de células de bacterias que expresan OmpT1.

Se entenderá que los sitios de escisión de OmpT introducidos en las proteínas modificadas descritas en el presente documento son escindibles por cualquier proteína o polipéptido que posea actividad enzimática de tipo OmpT. Todo lo que se requiere es que la enzima de tipo OmpT1 sea capaz de escindir una proteína modificada para contener los sitios de escisión por OmpT. Por ejemplo, la actividad proteolítica de tipo OmpT puede ser proporcionada por OmpT1 de tipo silvestre o nativa, o un homólogo funcional o fragmento de la misma. La actividad de tipo OmpT también puede ser proporcionada por una proteína que sea esencialmente idéntica o esencialmente similar a OmpT1. Por ejemplo, la proteína con actividad de tipo OmpT puede tener al menos el 60 %, 70 %, 80 %, 90 %, 95 % o 99 % de identidad de secuencia con OmpT1. En algunos casos, la proteína que tiene actividad de tipo OmpT es al menos un 60 %, 70 %, 80 %, 90 %, 95 % o 99 % idéntica a SEQ ID NO: 3. En algunos casos, la proteína que tiene actividad OmpT1 está unida específicamente por un anticuerpo policlonal que se une a OmpT1 y sus variantes. La selección de anticuerpos policlonales que se unen a OmpT1 y a sus variantes funcionales se puede realizar como se describe en el presente documento para seleccionar anticuerpos policlonales que se unen a RF1.

40

Transformación de las bacterias con las proteínas escindibles por OmpT1

Una vez determinado que las proteínas diana esenciales modificadas conservan la función de tipo silvestre y son susceptibles a la escisión por OmpT1, tal como se han descrito anteriormente, los ácidos nucleicos que codifican las proteínas modificadas se transforman en bacterias. Las bacterias pueden transformarse con el ácido nucleico en condiciones adecuadas para la incorporación del ácido nucleico al genoma de las bacterias. Por ejemplo, las bacterias pueden transformarse con oligonucleótidos que tengan secuencias que codifiquen los sitios de escisión por OmpT1 descritos en el presente documento usando el reemplazo alélico mediado por oligonucleótidos, como se describe en los ejemplos. La incorporación de las mutaciones deseadas al genoma bacteriano se puede determinar, por ejemplo, mediante la selección de colonias transformadas usando la PCR del ensayo de mutación de amplificación de desapareamientos (MAMA), como se describe en los ejemplos.

Uso de proteínas escindibles por OmpT1 para aumentar el rendimiento proteico en sistemas de traducción libres de células

55

Las proteínas diana modificadas que tienen sitios de escisión por OmpT1 se degradan eficazmente cuando se expresan en un extracto de células bacterianas que contiene la proteasa OmpT1 activa. La selección cuidadosa de proteínas inhibitoras que se modifican para ser escindibles por OmpT1 puede mejorar la productividad de los sistemas de síntesis libres de células. Por ejemplo, Por ejemplo, en un uso preferido, la escisión de proteínas diana seleccionadas por OmpT1 puede mejorar la eficacia de traducción de proteínas de longitud completa que tienen incorporados aminoácidos no nativos en extractos libres de células. En ciertos casos, el aminoácido no nativo se incorpora en un codón ámbar introducido en el molde de ARNm. Como se ha descrito anteriormente, la incorporación de aminoácidos no nativos en un codón ámbar puede ser inhibida por proteínas complejas de terminación tales como RF1 y RF2. Por lo tanto, en una realización particularmente deseable, la escisión e inactivación de RF1 y/o RF2 por OmpT1 puede aumentar el rendimiento de las proteínas diseñadas para incorporar un aminoácido no nativo en un codón ámbar.

65

Aminoácidos no nativos

- 5 Como se ha descrito anteriormente, Por ejemplo, en un uso preferido, la degradación de la proteína diana modificada por OmpT1 aumenta el rendimiento de proteínas que incorporan aminoácidos no nativos en los sistemas de síntesis libres de células. Los aminoácidos no nativos usados normalmente comprenden uno o más derivados o análogos de aminoácidos modificados químicamente, en los que las estructuras químicas tienen la fórmula $\text{NH}_3\text{-(CR)-COOH}$, en la que R no es ninguno de los 20 sustituyentes canónicos que definen los aminoácidos naturales. Los derivados de aminoácidos no nativos adecuados están disponibles en el mercado de proveedores tales como, por ejemplo, Bachem Inc., (Torrance, CA); Genzyme Pharmaceuticals (Cambridge, MA); Senn Chemicals (Dielsdorf, Suiza); Sigma-Aldrich (St. Louis, MO); Synthetec, Inc (Albany, OR). Preferentemente, los aminoácidos no nativos incluyen, pero sin limitación, derivados y/o análogos de glicina, tirosina, glutamina, fenilalanina, serina, treonina, prolina, triptófano, leucina, metionina, lisina, alanina, arginina, asparagina, valina, isoleucina, ácido aspártico, ácido glutámico, cisteína, histidina, así como beta-aminoácidos y homólogos, aminoácidos protegidos con BOC y aminoácidos protegidos con FMOC.
- 10
- 15 La generación de derivados de aminoácidos no nativos, análogos y miméticos que ya no están disponibles en el mercado se puede lograr de varias maneras. Por ejemplo, una forma es sintetizar un aminoácido no nativo de interés usando métodos de Química orgánica conocidos en la técnica, mientras que otra forma es usar métodos de síntesis quimioenzimática conocidos en la técnica. Véase, por ejemplo, Kamphuis *et al.*, *Ann. N. Y. Acad. Sci.*, 672:510-527, 1992; Ager D. J. y Fotheringham I. G., *Curr. Opin. Drug Discov. Devel.*, 4:800-807, 2001; y Weiner *et al.*, *Chem. Soc. Rev.*, 39:1656-1691, 2010; "Asymmetric Syntheses of Unnatural Amino Acids and Hydroxyethylene Peptide Isosteres", Wieslaw M. Kazmierski, ed., *Peptidomimetics Protocols*, Vol. 23, 1998; "Unnatural Amino Acids", Kumar G. Gadamasetti y Tamim Braish, ed., *Process Chemistry in the Pharmaceutical Industry*, Vol. 2, 2008.
- 20
- 25 Un experto en la materia reconocerá que hay muchos procedimientos y protocolos disponibles para la síntesis de aminoácidos no nativos, por ejemplo, según lo descrito en Wieslaw M. Kazmierski, ed., *Peptidomimetics Protocols*, Vol. 23, 1998; Wang L *et al.*, "Chemistry and Biology", 16:323-336, 2009; y Wang F., Robbins S., Guo J., Shen W. y Schultz P. G., *PLoS One*, 5:e9354, 2010.
- 30 Los aminoácidos no nativos pueden incluir aminoácidos L- y D-alfa no nativos. Los aminoácidos L-alfa pueden sintetizarse químicamente mediante métodos conocidos en la técnica, tales como, pero sin limitación, el acoplamiento reductor mediado por hidrógeno a través de la formación de enlaces C-C catalizada por rodio de conjugaciones hidrogenadas de alquinos con iminoacetatos de etilo (Kong *et al.*, *J. Am. Chem. Soc.*, 127:11269-11276, 2005). Como alternativa, se puede utilizar la producción semisintética por ingeniería metabólica. Por ejemplo, se pueden usar procedimientos de fermentación para sintetizar aminoácidos no nativos de *E. coli* que albergan una vía biosintética de cisteína rediseñada. (Véase, Maier T. H., *Nature*, 21:422-427, 2003). Se pueden producir mezclas racémicas de alfa-aminoácidos usando síntesis asimétrica de Strecker (como se describe en Zuend *et al.*, *Nature*, 461: 968-970 (2009)) o usando enzimas transaminasas para la síntesis a gran escala (tal como se encuentra en Taylor *et al.*, *Trends Biotechnol.*, 16:412-419, 1998. Los alfa-aminoácidos terciarios bicíclicos pueden producirse mediante la alquilación de bases de Schiff derivadas de glicina o nitroacetatos con electrófilos de éter cíclico, seguida de la apertura dl anillo inducida por ácido y la ciclación en NH_4OH (véase Strachan *et al.*, *J. Org. Chem.*, 71:9909-9911 (2006)).
- 35
- 40 Los aminoácidos no nativos pueden comprender además beta-aminoácidos, que son notablemente estables al metabolismo, presentan una degradación microbiana lenta, y son inherentemente estables a las proteasas y peptidasas. Un ejemplo de la síntesis de los beta-aminoácidos se describe en Tan C. Y. K. y Weaver D. F., *Tetrahedron*, 58:7449-7461, 2002.
- 45
- 50 En algunos casos, los aminoácidos no nativos comprenden aminoácidos modificados químicamente comúnmente usados en la síntesis de péptidos en fase sólida, incluyendo, pero sin limitación, aminoácidos protegidos con *tert*-butoxicarbonil-(Boc) o (9*H*-fluoren-9-ilmtox)carbonil (Fmoc). Por ejemplo, los derivados de Boc de leucina, metionina, treonina, triptófano y prolina pueden producirse por oxidación selectiva de la cadena lateral del 3,3-dimetildioxirano, como se describe en Saladino *et al.*, *J. Org. Chem.*, 64:8468-8474, 1999. Los derivados de Fmoc de alfa-aminoácidos se pueden sintetizar mediante la alquilación de nitroacetato de etilo y la transformación en derivados (véase Fu *et al.*, *J. Org. Chem.*, 66:7118-7124, 2001).
- 55
- 60 Los aminoácidos no nativos que pueden usarse pueden incluir, pero no se limitan a, análogos no nativos o derivados de los 20 sustituyentes de aminoácidos canónicos. Un experto en la materia apreciará que la síntesis de diversos aminoácidos no nativos puede implicar una serie de métodos químicos y quimioenzimáticos conocidos en la técnica. Los aminoácidos no nativos pueden sintetizarse de acuerdo con procedimientos conocidos en la técnica específicos para un derivado en particular de cada aminoácido no nativo. Sycheva *et al.*, *Microbiology*, 76: 712-718, 2007 describe un procedimiento para sintetizar los aminoácidos no nativos norvalina y norleucina. Los derivados de prolina dializados pueden producirse mediante síntesis estereoselectiva práctica (véase Belvisi *et al.*, *Tetrahedron*, 57:6463-6473, 2001). Por ejemplo, los derivados de triptófano se pueden sintetizar mediante la sustitución electrófila de indo catalizada con triflato de iterbio como se describe en Janczuk *et al.*, *Tetrahedron Lett.*, 43: 4271-4274, 2002, y la síntesis de derivados de 5-aril-triptófano se detalla en Wang *et al.*, *Tetrahedron*, 58:3101-3110, 2002. Los análogos de serina no nativos se pueden producir mediante la fragmentación beta de los radicales alcoxilo primarios (véase Boto *et al.*, *J. Org. Chem.*,
- 65

72:7260-7269, 2007). Como alternativa, se describe un procedimiento para la síntesis de fenilserina en Koskinen *et al.*, *Tetrahedron Lett.*, 36:5619-5622, 1995. El procedimiento para la síntesis de L-fenilglicina se describe en Cho *et al.*, *Biotechnol. Bioprocess. Eng.*, 11:299-305, 2006. Y se describe un método quimioenzimático para sintetizar D-4-hidroxifenilglicina en Yu *et al.*, *Folia Microbiol (Praga)*, 54:509-15; 2009. Un ejemplo no limitante de la producción de 2-naftilalanina o 2-naftilalanina protegida con Boc se detalla en Boaz *et al.*, *Org. Process Res. Dev.*, 9:472-478; 2005. La síntesis de yodo-L-tirosina y *p*-benzoil-L-fenilalanina se describe en Hino N., *Nat. Protoc.*, 1:2957-2962, 2007.

ARNt cargado

10 Para incorporar los aminoácidos no nativos descritos en el presente documento al polipéptido deseado, el aann descrito anteriormente necesita cargarse a ARNt de codones ámbar o sentido isoaceptores. La reacción de carga del ARNt, como se usa en el presente documento, se refiere a la reacción de aminoacilación del ARNt *in vitro*, en la que se aminoacilan los ARNt de codones sentido isoaceptores o codones ámbar deseados con sus respectivos aminoácidos de interés. La reacción de carga del ARNt comprende la mezcla de reacción de carga, un ARNt sentido isoaceptor, y como se usa en la presente invención, puede incluir aminoácidos naturales o no nativos. La reacción de carga del ARNt puede ocurrir *in situ* en la misma reacción que la reacción de traducción libre de células, o puede ocurrir en una reacción separada, en la que el ARNt cargado se añade luego a la reacción de traducción libre de células.

20 Las moléculas de ARNt para su uso en la reacción de carga de ARNt pueden sintetizarse a partir de un molde de ADN sintético para cualquier ARNt de elección tras la amplificación por PCR en presencia de los cebadores 5' y 3' apropiados. El molde de ADN bicatenario resultante, que contiene una secuencia de promotor T7, puede transcribirse luego *in vitro* usando la ARN polimerasa de T7 para producir la molécula de ARNt, que posteriormente se añade a la reacción de carga del ARNt.

25 La reacción de carga del ARNt puede ser cualquier reacción que aminoacile una molécula de ARNt de codón sentido o ámbar con un aminoácido deseado separado de la reacción de síntesis de proteínas. Esta reacción puede tener lugar en un extracto, una mezcla de reacción artificial o una combinación de ambos. Las condiciones de reacción de aminoacilación del ARNt adecuadas son bien conocidas por los expertos en la materia. Por lo general, la aminoacilación del ARNt se lleva a cabo en un tampón fisiológico con un valor de pH que varía de 6,5 a 8,5, fosfato de alta energía 0,5-10 mM (tal como ATP), MgCl₂ 5-200 mM, KCl 20-200 mM. Preferentemente, la reacción se realiza en presencia de un agente reductor (tal como ditiotreitól 0-10 mM). Cuando la aminoacil-ARNt sintetasa se añade exógenamente, la concentración de la sintetasa normalmente es de 1-100 nM. Un experto en la materia reconocerá fácilmente que estas condiciones se pueden variar para optimizar la aminoacilación del ARNt, tal como una alta especificidad para los aminoácidos preseleccionados, altos rendimientos y menor reactividad cruzada.

40 En otros casos, los ARNt isoaceptores o ámbar son cargados por aminoacil-ARNt sintetasa. Las reacciones de carga del ARNt pueden utilizar la aminoacil-ARNt sintetasa nativa específica para los ARNt sentido isoaceptores que se vayan a cargar, una aminoacil-ARNt sintetasa modificada por ingeniería genética o una aminoacil ARNt sintetasa "promiscua" capaz de cargar una molécula de ARNt con más de un tipo de aminoácidos. Las propias aminoacil-ARNt sintetasa promiscuas pueden ser modificadas por ingeniería genética, o pueden incluir aminoacil-ARNt sintetasa producidas de manera endógena que a veces se encuentran en la naturaleza. En el documento WO2010/081110, se describen métodos para cargar los ARNt isoaceptores con aminoácidos nativos y no nativos usando aminoacil-ARNt sintetasa.

Sistemas de traducción

50 Los ARNt sentido isoaceptores o ámbar cargados descritos anteriormente se combinan ahora con un sistema de traducción que puede comprender un extracto libre de células, un lisado de células o un sistema de traducción reconstituido, junto con el molde de ácido nucleico para la síntesis del polipéptido o la proteína que se desee que tenga aminoácidos no nativos en posiciones preseleccionadas (definidas). La mezcla de reacción comprenderá además monómeros para la macromolécula que se vaya a sintetizar, por ejemplo, aminoácidos, nucleótidos, etc., y dichos cofactores, enzimas y otros reactivos que son necesarios para la síntesis, por ejemplo, ribosomas, ARNt, polimerasas, factores de transcripción, etc. Además de los componentes anteriores, tales como un extracto libre de células, un molde de ácido nucleico y aminoácidos, se pueden añadir a la reacción materiales específicamente requeridos para la síntesis de proteínas. Los materiales incluyen regulares, ácido folínico, AMP cíclico, inhibidores de enzimas degradantes de proteínas o de ácido nucleico, inhibidores o reguladores de la síntesis de proteínas, ajustadores de los potenciales de oxidación/reducción, tensoactivos no desnaturalizantes, componentes tampón, espermina, espermidina, putrescina, etc. En la técnica, se conocen diversos sistemas de reacción de síntesis libres de células. Véase, por ejemplo, Kim, D. M. y Swartz, J. R. *Biotechnol. Bioeng.* 66:180-8 (1999); Kim, D. M. y Swartz, J. R. *Biotechnol. Prog.* 16:385-90 (2000); Kim, D. M. y Swartz, J. R. *Biotechnol. Bioeng.* 74:309-16 (2001); Swartz *et al.*, *Methods Mol Biol.* 267:169-82 (2004); Kim, D. M. y Swartz, J. R. *Biotechnol. Bioeng.* 85:122-29 (2004); Jewett, M. C. y Swartz, J. R., *Biotechnol. Bioeng.* 86:19-26 (2004); Yin, G. y Swartz, J. R., *Biotechnol. Bioeng.* 86:188-95 (2004); Jewett, M. C. y Swartz, J. R., *Biotechnol. Bioeng.* 87:465-72 (2004); Voloshin, A.M. y Swartz, J. R., *Biotechnol. Bioeng.* 91:516-21 (2005). Las condiciones adicionales para la síntesis libre de células de los polipéptidos deseados se describen en el documento WO2010/081110.

En algunos casos, se usa un molde de ADN para conducir la síntesis de proteínas *in vitro*, y se añade la ARN polimerasa a la mezcla de reacción para potenciar la transcripción del molde de ADN. Las ARN polimerasas adecuadas para su uso en el presente documento incluyen cualquier ARN polimerasa que funcione en las bacterias de las que se derive el extracto bacteriano. En otros casos, se usa un molde de ARN para dirigir la síntesis de proteínas *in vitro*, y los componentes de la mezcla de reacción se pueden mezclar entre sí en cualquier orden conveniente, pero preferentemente se mezclan en un orden en el que se añade el molde de ARN en último lugar, reduciendo así al mínimo la posible degradación del molde de ARN por parte de las nucleasas.

10 Sistemas de traducción libres de células

La síntesis de proteínas libres de células pueden aprovechar el poder catalítico de la maquinaria celular. La obtención de rendimientos máximos de proteínas *in vitro* requiere un suministro adecuado de sustratos, por ejemplo, trifosfatos de nucleósidos y aminoácidos, un entorno homeostático, la estabilidad del catalizador y la eliminación o evitación de subproductos inhibidores. La optimización de las reacciones sintéticas *in vitro* se beneficia de la recreación del estado *in vivo* de un organismo en rápido crecimiento. En algunos casos, la síntesis libre de células se realiza en una reacción en la que se activa la fosforilación oxidativa, es decir, el sistema CYTOMIM™. El sistema CYTOMIM™ se define usando una condición de reacción en ausencia de polietilenglicol con una concentración de magnesio optimizada. El sistema CYTOMIM™ no acumula fosfato, que se sabe que inhibe la síntesis de proteínas, mientras que las fuentes de energía secundarias convencionales producen una acumulación de fosfato. Otras diversas características del sistema CYTOMIM™ se describen en la patente de EE.UU. n.º 7.338.789.

La presencia de una vía de la fosforilación oxidativa activa se puede demostrar por la falta de necesidad de fuentes de energía secundarias, tales como fosfoenolpiruvato, fosfato de creatina, fosfato de acetilo o productos intermedios glicolíticos tales como glucosa, glucosa-6-fosfato y piruvato. La presencia de una vía de fosforilación oxidativa activa también se puede determinar por la sensibilidad de la vía hacia inhibidores, tales como los inhibidores de la cadena de transporte de electrones. Los ejemplos de inhibidores de la cadena de transporte de electrones incluyen 2-heptil-4-hidroxiquinolina-N-óxido (HQNO), 2,4-dinitrofenol, cianuro, azida, tenoiltrifluoroacetona y carbonil-cianuro-*m*-clorofenilhidrazona. Como alternativa, en una realización, el sistema de traducción libre de células no comprende una vía de fosforilación oxidativa activa.

La síntesis de proteínas *in vitro* o libre de células ofrece varias ventajas frente a los métodos convencionales de expresión de proteínas *in vivo*. Los sistemas libres de células pueden dirigir la mayoría, si no todos, los recursos metabólicos de la célula hacia la producción exclusiva de una proteína. Por otra parte, la falta de componentes de la pared y de la membrana celulares *in vitro* es ventajosa, ya que permite el control del entorno de la síntesis. Por ejemplo, los niveles de ARNt pueden cambiarse para reflejar el uso del codón de los genes que se expresan. El potencial redox, el pH o la fuerza iónica también se pueden modificar con mayor flexibilidad que con la síntesis de proteínas *in vivo* debido a que no existen problemas de crecimiento ni viabilidad celular. Además, puede lograrse fácilmente la recuperación directa de productos de proteínas purificadas y correctamente plegadas.

La productividad de los sistemas libres de células ha mejorado en 2 órdenes de magnitud en los últimos años, de aproximadamente 5 µg/ml-h a aproximadamente 500 µg/ml-h. Dichas mejoras han hecho de la síntesis de proteínas *in vitro* una técnica práctica para la investigación a escala de laboratorio y que proporciona una tecnología de plataforma para la expresión de proteínas de alto rendimiento. Indica además la viabilidad de usar tecnologías libres de células como un medio alternativo a la producción comercial de proteínas farmacéuticas a gran escala *in vivo*.

Generación de un lisado

Se usa y se describe un lisado celular para la traducción *in vitro* de una proteína diana. Por comodidad, el organismo usado como fuente del lisado puede denominarse organismo de origen o célula hospedadora. Las células hospedadoras pueden ser bacterias, levaduras, células de mamíferos o plantas, o cualquier otro tipo de célula capaz de sintetizar proteínas. Un lisado comprende componentes que son capaces de traducir ácido ribonucleico mensajero (ARNm) que codifica una proteína deseada, y opcionalmente, comprende componentes que son capaces de transcribir ADN que codifica una proteína deseada. Dichos componentes incluyen, por ejemplo, ARN polimerasa dirigida por ADN (ARN polimerasa), cualquier activador de la transcripción que sea necesario para el inicio de la transcripción del ADN que codifica la proteína deseada, ácidos ribonucleicos de transferencia (ARNt), aminoacil-ARNt sintetetas, ribosomas 70S, N¹⁰-formiltetrahidrofolato, formilmietionina-ARNt^{Met} sintetetas, peptidil transferasa, factores de inicio tales como IF-1, IF-2 e IF-3, factores de elongación tales como EF-Tu, EF-Ts y EF-G, factores de liberación tales como RF-1, RF-2 e RF-3, y similares.

Se describe el uso de una célula bacteriana de la que se deriva un lisado. En los métodos descritos, se puede usar un lisado bacteriano derivado de cualquier cepa de bacterias. El lisado bacteriano se puede obtener de la siguiente manera. Las bacterias de elección crecen durante la noche en cualquiera de una serie de medios de crecimiento y en condiciones de crecimiento que son bien conocidas en la técnica y que un profesional puede optimizar fácilmente para el crecimiento de las bacterias particulares. Por ejemplo, un entorno natural para la síntesis utiliza lisados celulares derivados de células bacterianas cultivadas en un medio que contiene glucosa y fosfato, en el que la glucosa está

presente a una concentración de al menos aproximadamente el 0,25 % (peso/volumen), más normalmente de al menos aproximadamente el 1 %; y habitualmente no más de aproximadamente el 4 %, más normalmente no más de aproximadamente el 2 %. Un ejemplo de dichos medios es el medio 2YTPG, sin embargo, un experto en la materia apreciará que muchos medios de cultivo pueden adaptarse para este fin, ya que existen muchos medios publicados adecuados para el crecimiento de bacterias tales como *E. coli*, usando tanto fuentes definidas como indefinidas de nutrientes. Las células que se han recogido durante la noche pueden lisarse suspendiendo el sedimento celular en un tampón de suspensión celular adecuado, y rompiendo las células suspendidas por ultrasonidos, rompiendo las células suspendidas en una prensa francesa, mediante homogenización de flujo continuo de alta presión o cualquier otro método conocido en la técnica es útil para una lisis celular eficaz. El lisado celular se centrifuga o se filtra para retirar los fragmentos de ADN grandes.

Ejemplos

Ejemplo 1

El presente ejemplo describe la modificación de la proteína RF1 para introducir secuencias de aminoácidos que son escindibles por la proteasa OmpT1.

Métodos:

Construcción de un molde para introducir sitios de escisión por OMPT1 en RF1

Se usó una estrategia basada en PCR para introducir los cambios de la secuencia de nucleótidos deseada en moldes de ácido nucleico codificantes de las proteínas RF1 modificadas descritas en el presente documento. La reacción de PCR se llevó a cabo usando la mezcla maestra Hot Start Flex 2X (NEB) de Phusion de acuerdo con los protocolos sugeridos por el fabricante. En general, se llevó a cabo una PCR solapante de dos etapas para introducir mutaciones de escisión por OmpT en el gen codificante de RF1, como se describe con mayor detalle a continuación. Las moldes de ADN generados por PCR se purificaron usando el kit de purificación de PCR QIAquick (QIAGEN) para la aplicación en la expresión libre de células. Tras la purificación por PCR, se secuenciaron las variantes mediante Mclab (South San Francisco, CA) para confirmar la presencia de las mutaciones esperadas.

Preparación de la proteína GamS para prevenir la degradación de los moldes de ADN en reacciones sin células

La eficacia de la transcripción a partir de un molde de ADN puede disminuir debido a la degradación del molde por exonucleasas bacterianas endógenas presentes en los extractos libres de células. Se sabe que la forma corta de la proteína Gam del fago λ (GamS) protege los moldes de ADN de la degradación al inhibir la actividad de RecBCD (Exonucleasa V) (véase Sitaraman, K., Esposito, D., Klarmann, G., Grice, S. F. L., Hartley, J. L. y Chatterjee, D. K. 2004, "A Novel Cell-free Protein Synthesis System". *J Biotechnol.* 110: 257-263. Por lo tanto, en este ejemplo, se usó la proteína GamS para estabilizar los moldes de PCR durante las reacciones de transcripción libres de células. Para producir la proteína GamS recombinante, se amplificó el gen GamS para incluir un marcador de poli-histidina C-terminal, GGSHHHHH (SEQ ID NO:50), mediante cebadores, 5'-ATATATCATATGAACGCTTATTACATTCAGGATCGTCTTGAG-3' (SEQ ID NO:51) y 5'-ATATATGTCGACTTAATGATGATGATGATGATGATGAGAACCCCTACCTCTGAATCA ATATCAACCTGGTGGTG-3' (SEQ ID NO:52) usando pKD46 (Datsenko, K. A. y Wanner, B. 2000, "One-step Inactivation of Chromosomal Genes in *Escherichia coli* K-12 Using PCR Products". *Proc. Natl. Acad. Sci. EE.UU.* 97: 6640-6645) como molde. Se subclonó el gen GamS en el plásmido de expresión libre de células pYD317 en los sitios de restricción NdeI/SalI. Se expresó GamS *in vitro* y se purificó mediante cromatografía de afinidad por metales inmovilizados (IMAC) con una pureza superior al 90 % (datos no mostrados). La proteína GamS se almacenó a -70 °C antes de la aplicación en un tampón de Tris-acetato 100 mM (pH 8,2), que también contenía acetato de potasio 160 mM, cloruro de sodio 200 mM y sacarosa al 10 %.

Cepas y plásmidos bacterianos

SBJY001 es un derivado de K12 de *E. coli* optimizado para la producción abierta de proteínas libres de células. Se transformó SBJY001 con el plásmido pKD46 (Coli Genetic Stock Center) que contenía genes de Red recombinasa del fago λ bajo un promotor de arabinosa inducible.

Construcción de SBHS002

SBHS002 es una cepa de *E. coli* con eliminación de ompT creada usando la transducción de P1. El lisado P1 se creó a partir de JW0554-1 (CGSC n.º 8680), una cepa de la colección Keio que contiene la mutación ompT::Kan^R flanqueada por sitios FRT. Luego se usó el lisado JW0554-1 P1 para introducir la mutación en SBJY001 mediante la transducción de P1. Se cultivaron colonias en LB con 30 μ g/ml de kanamicina para seleccionar la resistencia a la kanamicina. SBJY001ompT::Kan^R se transformó con plásmido de expresión 708-FLPe Cm^R (Gene Bridges). Se indujo la síntesis de FLP y se examinaron las colonias para determinar la pérdida de resistencia a la kanamicina. Las colonias resistentes a la kanamicina se secuenciaron para confirmar la eliminación de ompT. La cepa con eliminación de ompT

se conoce como SBHS002.

Clonación y expresión de RF1 marcada con His N-terminal

5 Se amplificó RF1 a partir de ADN genómico de la cepa A19 de *E. coli* usando los cebadores 5His-RF1: CATATGCATCACCATCACCATCACGGTGGTGGCTCTAAGCCTTCTATCGTTGCCA AACTGGAAGCC (SEQ ID NO: 138) y 3RF1: GTCGACTTATTCTGCTCGGACAACGCCGCCAG (SEQ ID NO: 139) que introdujeron un marcador de His N-terminal y los sitios de restricción NdeI/SalI. Se ligó la inserción en el vector de expresión pYD317 y se confirmó mediante secuenciación. RF1 se expresó en una reacción libre de células de 25 ml usando el plásmido, se purificó
10 mediante IMAC, y el tampón se intercambiaba a PBS.

Escisión de variantes de RF1 recombinantes usando extractos sin células

15 Para ensayar si las proteínas RF1 modificadas descritas en el presente documento eran escindibles por OmpT, se incubaron variantes de RF1 recombinantes con extractos celulares con o sin OmpT. Se cultivaron células SBJY001, que tenían una proteasa ompT intacta en la membrana externa, y SBHS002, de la que se eliminó ompT, en 5 ml de LB durante una noche a 37 °C. Se centrifugaron 50 µl de cada cultivo a 8.000 rpm durante 2 minutos y se lavaron dos veces con Tris 10 mM, cloruro de amonio 20 mM y cloruro de magnesio 10 mM. A continuación, se volvieron a suspender los sedimentos celulares en 50 µl del tampón, y se añadieron 10 µg de proteína RF1 de *E. coli* recombinante
20 purificada. Se incubaron las muestras a 37 °C. Se centrifugaron las muestras a 8.000 rpm durante 2 minutos. Se retiró el sobrenadante que contenía la proteína RF1, y se procesó en un gel SDS-PAGE.

Expresión libre de células de las variantes de RF1

25 Las reacciones de transcripción/traducción libres de células se llevaron a cabo en un volumen de 60 µl a 30 °C en placas de 24 pocillos profundos (n.º de cat. 95040470, Thermo Scientific) durante 4,5 horas. La concentración del molde de PCR para la expresión de la variante de RF1 era de 10 µg/ml. La composición de la reacción también incluía glutamato de magnesio 8 mM, glutamato de potasio 130 mM, piruvato sódico 35 mM, AMP 1,2 mM, 0,86 mM de cada uno de GMP, UMP y CMP, oxalato de sodio 4 mM, putrescina 1 mM, espermidina 1,5 mM, fosfato de potasio 15 mM,
30 tirosina 1 mM, 2 mM de cada uno de los otros 19 aminoácidos, ARN polimerasa T7 100 nM, extracto de células S30 al 30 % (v/v). Para facilitar la formación de disulfuro, se trató el extracto de células S30 con IAM 50 µM a temperatura ambiente durante 30 minutos antes de la reacción sin células. También se añadió una mezcla de glutatión oxidado (GSSG) 2 mM y glutatión reducido (GSH) 1 mM con disulfuro isomerasa DsbC de *E. coli* 4,3 µM. Para analizar las variantes de RF1 expresadas libres de células con SDS-PAGE y autorradiograma, se realizaron reacciones en presencia de cantidades traza de [¹⁴C]-Ieucina (300 µCi/mol; GE Life Sciences, NJ). Las variantes de RF1 se expresaron bien en un extracto celular OmpT positivo (OmpT +), que se preparó a partir de la cepa bacteriana SBJY001, o un extracto celular OmpT negativo (OmpT-), que se preparó a partir de la cepa bacteriana SBHS002 (SBJY001 Δ ompT). Para estabilizar los moldes de PCR en la reacción libre de células, también se añadió proteína GamS 1,4 µM para inhibir la actividad de RecBCD.
40

SDS-PAGE y autorradiografía

45 Para determinar si las proteínas RF1 modificadas eran escindibles por OmpT, se analizaron las proteínas RF1 traducidas en las reacciones libres de células descritas anteriormente mediante SDS-PAGE y autorradiografía. Las muestras de reacción libres de células marcadas con ¹⁴C se centrifugaron a la velocidad máxima en una centrífuga de sobremesa y se mezclaron 4 µl de sobrenadante con tampón de carga de muestras SDS-PAGE de Invitrogen y agua. Las muestras se cargaron en geles SDS-PAGE de Bis-Tris al 4~12 % y se procesaron con tampón de ejecución MES durante aproximadamente 45 minutos. Después, los geles se secaron y se expusieron a un tamiz de fósforo (63-0034-86, GE healthcare, EE. UU.) durante una noche, y luego se escanearon usando Storm 460 (GE healthcare, EE. UU.).
50

Introducción de los sitios de escisión por OmpT en RF1

Para identificar los posibles sitios de escisión por OmpT en la secuencia RF1, se introdujeron mutaciones únicas de Arg o Lys en diferentes regiones del bucle de RF1 junto a una Arg o Lys existente. Las regiones de bucle se predijeron
55 en función de la alineación de secuencias de los factores de liberación de clase I procariontes (véase Graille M., Heurgue-Hamard, V., Champ, S., Mora, L., Scrima, N., Ulryck, N., Tilbeurgh, H. y Buckingham, R. H. 2005, "Molecular Basis for Bacterial Class I Release Factor Methylation by PrmC". *Molecular Cell*. 20: 917-927.) Para generar las variantes deseadas, se diseñaron las secuencias de mutación en mitad de los cebadores oligo directos e inversos (Tabla 1). En la PCR de la primera etapa, el fragmento 5' se amplificó mediante PCR usando 5chiT2PT7, 5'-GCGTACTAGCGTACCACGTGGCTGGTGGCCGATTCAATGCAGCTGGCAGAC AGG-3' (SEQ ID NO:53) y los cebadores inversos (Tabla 1). El fragmento 5' incluía el promotor T7, la región constante de la secuencia N-terminal y el sitio de mutación. El fragmento 3' también se amplificó en la primera etapa usando cebadores directos (Tabla 1) y 3chiT2TT7, 5'-GCGTACTAGCGTACCACGTGGCTGGTGGCCGTTGAGTTTTCTCCTTCATTACAGAA ACGGC-3' (SEQ ID NO:54). El fragmento 3' incluía el sitio de mutación, la región C-terminal constante y las secuencias del
60 terminador T7. En la PCR de la segunda etapa, se ensamblaron los ADN del fragmento 5' y del fragmento 3' mediante una PCR solapada usando un cebador único 5chiT2,5'-GCGTACTAGCGTACCACGTGGCTGGTGG-3' (SEQ ID
65

NO:55).

5 En primer lugar, se introdujeron 9 mutaciones individuales, M74R, E76R, E84R, A85R, E87R, E108R, T293R, N296K y S304R, en las regiones de bucle junto a una Arg o Lys existente (**Tabla 2**). Estas variantes de RF1 y RF1 WT se expresaron usando moldes de PCR en extractos libres de células con o sin OmpT. Las proteínas RF1 se analizaron mediante SDS-PAGE y autorradiografía. Cuando se expresaron en un extracto de células OmpT negativo, las 10 variantes de RF1 mostraron una proteína entera de longitud completa en SDS-PAGE. Sin embargo, cuando se expresaron en extractos celulares que contenían OmpT, la variante N296R se digirió parcialmente, mientras que las otras variantes migraron como se esperaba para las proteínas no digeridas de longitud completa. N296 se encuentra en la región de bucle de conmutación de RF1, que es flexible y de fácil acceso mediante la digestión con proteasas.

15 En la primera serie de selección, se seleccionó N296K para la digestión parcial por OmpT en el extracto celular. En la segunda serie de selección, se generaron moldes de PCR de variantes mutantes dobles y mutantes triples usando el mismo método descrito anteriormente con los cebadores enumerados en la **Tabla 1**. Seis mutantes dobles, N296K/I297V, N296K/I297K, N296K/I297R, N296R/I297V, N296R/I297K, N296R/I297R, y 2 mutantes triples, N296K/I297R/I298K and N296K/I297R/I298R (**Tabla 2**), se ensayaron en expresión libre de células con o sin OmpT. Las ocho variantes mutantes de RF1 dobles y triples se escindieron con OmpT1 (datos no mostrados). Entre estas variantes, N296K/L297R/L298R fue la más sensible a la digestión por OmpT.

Tabla 1. Secuencias de cebadores directos e inversos usadas en la selección de sitios de escisión por OmpT

Variantes de RF1	Directos	SEQ ID NO:
WT	GCTCGATGATCCTGAAATGCGTGAGATGGCCGAGG	56
M74R	GCTCGATGATCCTGAACGCCGTGAGATGGCCGAGG	57
E76K	CGATGATCCTGAAATGCGTAAGATGGCCGAGGATGAAC	58
E84K	CAGGATGAAC TGCCCAAAGCTAAAGAAAAAAGCGAGCAAC	59
A85R	CAGGATGAAC TGCCGGAACGTAAGAAAAAAGCGAGCAAC	60
E87R	GGATGAAC TGCGGAAAGCTAAACGTAAGAAAGCGAGCAACTGGAAC	61
E108R	GCCAAAAGATCCTGATGACCCGTCGTAACGCCCTTCCTCG	62
T293R	CAACAGGCCGGAAGCGTCTCGCCGTCGTAACCTGC	63
N296K	GCGTCTACCCGTCGTAACCTGCTGGGGAGTGGCG	64
S304K	GGGAGTGGCGATCGCAAGGACCGTAACCCGTACTTAC	65
N296K/I297V	CGAAGCGTCTACCCGTCGTAAGTTCTGGGGAGTGGCGATCGCAGC	66
N296K/I297K	CGAAGCGTCTACCCGTCGTAAGTTCTGGGGAGTGGCGATCGCAGC	67
N296K/I297R	CGAAGCGTCTACCCGTCGTAACCGTCTGGGGAGTGGCGATCGCAGC	68
N296R/I297V	CGAAGCGTCTACCCGTCGTCGCGTTCTGGGGAGTGGCGATCGCAGC	69
N296R/I297K	CGAAGCGTCTACCCGTCGTCGCAAGCTGGGGAGTGGCGATCGCAGC	70
N296R/I297R	CGAAGCGTCTACCCGTCGTCGCCGTTCTGGGGAGTGGCGATCGCAGC	71
N296K/I297R/I298K	CAGGCCGAAGCGTCTACCCGTCGTAACCGTAAGGGAGTGGCGATCGCAGCAGC	72
N296K/I297R/I298R	CAGGCCGAAGCGTCTACCCGTCGTAACCGTGGGGAGTGGCGATCGCAGCAGC	73
	Inversos	
WT	CCTGGCCATCTCACGCATTTACAGGATCATCGAGC	74
M74R	CCTGGCCATCTCACGGGTTACAGGATCATCGAGC	75
E76K	GTTTCATCTCGCCATCTTACGCATTTACAGGATCATCG	76
E84K	GTTGCTCGCTTTTTTCTTTAGCTTTGGCGAGTTTCATCTG	77
A85R	GTTGCTCGC TTTTTTCTTTACGTTTCGGCGAGTTTCATCTG	78
E87R	GTTCCAGTTGCTCGCTTTTACGTTTAGCTTTGGCGAGTTTCATCTC	79

Variantes de RF1	Directos	SEQ ID NO:
E108R	CGAGGAAGCGGTACGACGGTCATCAGGATCTTTTGGC	80
T293R	GCAGGTTACGACGGCGAGACGCTTCGGCCTGTTG	81
N296K	CGCCACTCCCAGCAGTTTACGACGGGTAGAGCG	82
S304K	GTAAGTACGGTTACGGTCTTGGGATCGCCACTCCC	83
N296K/I297V	GCTGCGATCGCCACTCCCAGAACTTACGACGGGTAGACGCTTTCG	84
N296K/I297K	GCTGCGATCGCCACTCCCAGCTTTTACGACGGGTAGACGCTTTCG	85
N296K/I297R	GCTGCGATCGCCACTCCCAGACGTTTACGACGGGTAGACGCTTTCG	86
N296R/I297V	GCTGCGATCGCCACTCCCAGAACGGGACGCGGTAGACGCTTTCG	87
N296R/I297K	GCTGCGATCGCCACTCCCAGCTTTCGACGCGGGTAGACGCTTTCG	88
N296R/I297R	GCTGCGATCGCCACTCCCAGACGGGACGCGGTAGACGCTTTCG	89
N296K/I297R/I298K	GGTCGCTGCGATCGCCACTCCCCTTACGTTTACGACGGGTAGACGCTTCGGCCTG	90
N296K/I297R/I298R	GGTCGCTGCGATCGCCACTCCCAGACGTTTACGACGGGTAGACGCTTCGGCCTG	91

*Las secuencias de las mutaciones están subrayadas.

Tabla 2. Variantes de RF1 de sustituciones simples, dobles y triples

N.º	Variantes de sustitución simple
A-1	M74R
A-2	E76K
A-3	E84K
A-4	A85R
A-5	E87R
A-6	E108R
A-7	T293R
A-8	N296K
A-9	S304K
	Variantes de sustitución doble
A-11	N296K/I297V
A-12	N296K/I297K
A-13	N296K/I297R
A-14	N296R/I297V
A-15	N296R/I297K
A-16	N296R/I297R
	Variantes de sustitución triple
A-17	N296K/I297R/I298K
A-18	N296K/I297R/I298R

Inserción de péptidos de escisión por OmpT en la región de bucle de conmutación de RF1

5 Además de introducir las mutaciones de aminoácidos descritas anteriormente, RF1 también se modificó para reemplazar las secuencias de tipo silvestre de la región del bucle de conmutación por secuencias conocidas de péptidos susceptibles a la proteasa OmpT (véase Hwang B., Varadarajan, N., Li, H., Rodriguez, S., Iverson, B. L. y Georgiou, G. 2007, "Substrate Specificity of the Escherichia coli Outer Membrane Protease OmpP". *J. Bacteriol.* 189: 522-530; McCarter, J. D., Stephens, D., Shoemaker, K., Rosenberg, S., Kirsch, J. F. y Georgiou, G. 2004, "Substrate Specificity of the Escherichia coli Outer Membrane Protease OmpT". *J. Bacteriol.* 186: 5919-5925). Se construyeron 10 22 variantes de RF1, y se enumeran en la **Tabla 3**. Las variantes n.º 1 a 14 contenían ARR (SEQ ID NO: 47) para la digestión con OmpT. La variante n.º 15 contenía ARR en lugar de ARR (SEQ ID NO: 47), ya que se encuentra al final del bucle de conmutación. La variante n.º 16 contenía una mutación simple N296R. Las variantes n.º 17, 18 y 19 contenían un péptido de escisión por OmpT WLAARRGRG (SEQ ID NO:48). Las variantes n.º 20, 21 y 22 contenían 15 otro péptido de escisión por OmpT WGGRWARKKGTI (SEQ ID NO: 49).

Las secuencias de mutación fueron diseñadas en los cebadores oligo directos e inversos enumerados en la **Tabla 4**. En la PCR de la primera etapa, el fragmento 5' se amplificó mediante PCR usando el cebador 5chiT2PT7 y los cebadores inversos (**Tabla 4**). El fragmento 5' incluía el promotor T7, la región constante de la secuencia N-terminal y el sitio de mutación. El fragmento 3' también se amplificó en la primera etapa usando cebadores directos (Tabla 3) y 3chiT2TT7. El fragmento 3' incluía el sitio de mutación, la región C-terminal constante y las secuencias del terminador T7. En la PCR de la segunda etapa, se ensamblaron los ADN del fragmento 5' y del fragmento 3' mediante una PCR solapada usando un cebador único 5chiT2 según lo descrito anteriormente.

25 Las 22 variantes de RF1 (**Tabla 3**) se expresaron usando moldes de PCR en extractos celulares con o sin OmpT. Entre las variantes de inserción ARR (SEQ ID NO:47), la n.º 9 mostró la sensibilidad más alta hacia la digestión con OmpT. Las variantes n.º 17, 18 y 19, que contenían el péptido de escisión por OmpT WLAARRGRG (SEQ ID NO:48), fueron digeridas parcialmente por OmpT. Sin embargo, fueron mucho menos sensibles que las variantes n.º 20, 21 y 22, que contenían el péptido de escisión por OmpT WGGRWARKKGTI (SEQ ID NO:49). Las variantes n.º 9, 20, 21 y 30 22 se seleccionaron como las variantes de RF1 más sensibles a la digestión con OmpT de estas 22 construcciones.

Tabla 3. Secuencias de péptidos de escisión por OmpT insertadas en el bucle de conmutación de RF1

ES 2 694 683 T3

N.º	Péptido**	SEQ ID NO:
*0	QQAEASTRRNLLGSGDRS	4
1	QARRGSTRRNLLGSGDRS	26
2	QQARRGSTRRNLLGSGDRS	27
3	QQAARRGRRNLLGSGDRS	28
4	QQAEARRGRNLLGSGDRS	29
5	QQAEAARRGNLLGSGDRS	30
6	QQAEASARRGLLGSGDRS	31
7	QQAEASTARRGLGSGDRS	32
8	QQAEASTRARRGGSGDRS	33
9	QQAEASTRRARRGSGDRS	34
10	QQAEASTRRNARRGGDRS	35
11	QQAEASTRRNLARRGDRS	36
12	QQAEASTRRNLLARRGRS	37
13	QQAEASTRRNLLGARRGS	38
14	QQAEASTRRNLLGSARRG	39
15	QQAEASTRRNLLGSGARR	40
16	QQAEASTRRLLGSGDRS	6
17	QQAWLAARRGRGGSGDRS	41
18	QQAEWLAARRGRGGSGDRS	42
19	QQAEAWLAARRGRGGDRS	43
20	QQWGGRWARKKGTIGDRS	44
21	QQAWGGRWARKKGTIDRS	45
22	QQAEWGGRWARKKGTIRS	46

*Secuencia del péptido del bucle de conmutación WT (Q287 a S304)
 **Las secuencias de aminoácidos y peptídicas insertadas están subrayadas.

Tabla 4. Secuencias de cebadores directos e inversos para la inserción de péptidos de escisión por OmpT

N.º	Directos	SEQ ID NO:
0	GCCGAAAGCGTACCCGTCGTAACCTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTA CTTACAACCTCCCG	92
1	<u>CGCCGTGGT</u> TCTACCCGTCGTAACCTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	93
2	<u>GCACGCCGTGGT</u> ACCCGTCGTAACCTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	94
3	<u>GCCGCACGCCGTGGI</u> CGTCGTAACCTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	95
4	GCCGAAAGCGCGCGTGGTGCCTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA CTTCCCG	96
5	GCCGAAAGCGCGCGTGGTAACTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA CTTCCCG	97
6	GCCGAAAGCGTGCACGCCGTGGTCTGCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	98
7	GCCGAAAGCGTACCGCACGCCGTGGTCTGGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	99
8	GCCGAAAGCGTACCCGTGCACGCCGTGGTGGGAGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	100
9	GCCGAAAGCGTACCCGTCGTGCACGCCGTGGTGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	101
10	GCCGAAAGCGTACCCGTCGTAAACGCACGCCGTGGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	102
11	GCCGAAAGCGTACCCGTCGTAACCTGGCAGCCCGTGGTGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	103
12	GCCGAAAGCGTACCCGTCGTAACCTGCTGGCAGCCCGTGGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	104
13	GCCGAAAGCGTACCCGTCGTAACCTGCTGGGGCAGCCCGTGGTGGCGATCGCAGCGACCGTAACCCGTAACCCGTA TTCCCG	105

N.º	Directos	SEQ ID NO:
14	GCCGAAAGCGTCTACCCGTCGTAACCTGCTGGGGAGTGCAGCCGGTGGIGACCCGGTAACCGTACTTACAAC TTCCCC	106
15	GCCGAAAGCGTCTACCCGTCGTAACCTGCTGGGGAGTGGCGACCCGGIGACCCGTAACCGTACTTACAAC TTCCCC	107
16	GCCGAAAGCGTCTACCCGTCGTCGTCGCTGGGGAGTGGCGATCGCAGCGACCGTAACCGTACTTACAAC TTCCCC	108
17	GCTGGCTGGCAGCGCGTCCGCGTCTGGCGGGAGTGGCGATCGCAGCGACCGTAACCGTACTTACAA CTTCCCC	109
18	GCCGAAAGCGTGGCAGCGCGTCCGCGTCTGGCGAGTGGCGATCGCAGCGACCGTAACCGTACTTACAA CTTCCCC	110
19	GCCGAAAGCGTGGCTGGCAGCGCGTCCGCGTCTGGCGGGGATCGCAGCGACCGTAACCGTACTTACAA CTTCCCC	111
20	IGGGGTGGCCGTTGGGCTCGCAAGAAAGGTACTATTGGCGATCGCAGCGACCGTAACCGTACTTACAAC TTCCCC	112
21	GCTGGGGTGGCCGTTGGGCTCGCAAGAAAGGTACTATTGATCGCAGCGACCGTAACCGTACTTACAAC TTCCCC	113
22	GCCGAAAGCGTGGCGTGGGCTCGCAAGAAAGGTACTATTGCGCAGCGACCGTAACCGTACTTACAAC TTCCCC	114
	Inversos	
0	GCTGCGATCGCCACTCCCCAGCAGGTTACGACGGGTAGACGCTTCGGCCTGTTGGCGTTTGCCATTTC GCAGC	115
1	GCTGCGATCGCCACTCCCCAGCAGGTTACGACGGGTAGACGCTTCGGCCTGTTGGCGTTTGCCATTTC GCAGC	116
2	GCTGCGATCGCCACTCCCCAGCAGGTTACGACGGGTACCGCGGCTGTTGGCGTTTGCCATTTC GCAGC	117
3	GCTGCGATCGCCACTCCCCAGCAGGTTACGACGGGTGCGGCTGTTGGCGTTTGCCATTTC GCAGC	118
4	GCTGCGATCGCCACTCCCCAGCAGGTTACGACGGGTGCGGCTGTTGGCGTTTGCCATTTC GCAGC	119

N.º	Directos	SEQ ID NO:
5	GCTGGATCGCCACTCCCACGAGGTTACCACGGCGTTCGGCTCTGGCGTTTG CCATTTACGAGC	120
6	GCTGGATCGCCACTCCCACGAGACACGGCGTGCAGCGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	121
7	GCTGGATCGCCACTCCCACGAGACACGGCGTGCAGCGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	122
8	GCTGGATCGCCACTCCCACGGCGTGCAGCGGTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	123
9	GCTGGATCGCCACTACCACGGCGTGCAGCGGTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	124
10	GCTGGATCGCCACACGGCGTGCAGCGGTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	125
11	GCTGGATCACACGGCGTGCAGGTTACGAGGTTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	126
12	GCTGGACACGGCGTGCAGGTTACGAGGTTAGACGCTTCGGCTGTGGCGTTTGCCATTTC AGCAGC	127
13	GCTACCACGGCGTGCAGGTTACGAGGTTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	128
14	ACCACGGCGTGCAGGTTACGAGGTTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	129
15	ACGGCGTGCAGGTTACGAGGTTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	130
16	GCTGGATCGCCACTCCCACGAGACGAGCGGTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	131
17	GCTGGATCGCCACTCCCACGAGACGAGCGGTAGACGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	132
18	GCTGGATCGCCACTGCCACGAGCGGCGTGCAGCGCTTCGGCTGTGGCGTTTGCCATTCA GCAGC	133

N.º	Directos	SEQ ID NO:
19	GCTGCGATCGCCGCGCCACGACCGCGAGCGGCTGCCAGCCACGGCTTCGGCCTGTTGGCGTTTGCCATTTCAGCAGC	134
20	GCTGCGATCGCCAATAGTACCTTTCTTGGAGCCCAACGGCCACCCCACTGTTGGCGTTTGCCATTTCAGCAGC	135
21	GCTGCGATCAATAGTACCTTTCTTGGAGCCCAACGGCCACCCCAAGGCCTGTTGGCGTTTGCCATTTCAGCAGC	136
22	GCTGCGAATAGTACCTTTCTTGGAGCCCAACGGCCACCCCAATCGGCCTGTTGGCGTTTGCCATTTCAGCAGC	137
*Las secuencias de las mutaciones están subrayadas.		

Ejemplo 2

Habiendo demostrado que la RF1 se modificó con éxito para ser escindible por la proteasa OmpT1, este ejemplo describe la construcción de cepas bacterianas recombinantes que expresan tanto la RF1 modificada como la OmpT1 intacta. El fin de este ejemplo es mostrar que las variantes de RF1 expresadas por las cepas recombinantes se escinden en extractos libres de células de las cepas recombinantes que expresan OmpT1.

Reemplazo alélico mediado por oligonucleótidos

Para generar cepas bacterianas que expresaran las proteínas RF1 modificadas descritas en el presente documento, se usó la sustitución alélica mediada por oligonucleótidos (OMAR) para insertar las mutaciones de RF1 en el genoma bacteriano. El protocolo de OMAR se adaptó de un protocolo presentado anteriormente (Wang y Church, *Methods in Enzymology*, 2011, 498, 409-426). En resumen, se cultivó SBJY001 que contenía el plásmido pKD46 en 3 ml de LB y 50 µg/ml de ampicilina a 30 °C a DO₆₀₀ de 0,3. A continuación, se indujeron las células con L-arabinosa 1 mM a 37 °C durante 45 min. El sedimento celular se lavó dos veces con glicerol frío al 10 % y se volvió a suspender en 30 µl de glicerol frío al 10 %. Se añadieron 5 µM de cada oligo a las células resuspendidas. Los oligos sintéticos (Integrated DNA Technologies) tenían una longitud de 90 pares de bases y se diseñaron para hibridarse con la hebra retrasada durante la replicación del ADN (véase **Tabla 5**). Las células se sometieron a electroporación a 1.800 V durante 5 ms en una cubeta de 1 mm. Luego, se recuperaron en 3 ml de LB y 50 µg/ml de Amp. Este proceso se repitió durante 13 ciclos. Las células se diluyeron y se sembraron en placas de agar LB y se cultivaron a 37 °C durante la noche.

PCR de MAMA para identificar cepas bacterianas con las mutaciones deseadas

Se seleccionaron colonias bacterianas usando una adaptación de la PCR del ensayo de mutación de amplificación de desapareamientos (MAMA) para identificar las cepas con las mutaciones deseadas en RF1 (Cha *et. al.*, "PCR Methods and Applications", 1992, 2, 14-20). En resumen, se usó un cebador 5' universal junto con un cebador 3' que era específico de cada mutación para diferenciar entre una colonia mutante y una colonia WT (véase la **Tabla 5**). Los oligos se ordenaron desde el operón MWG Eurofins. Se usó la supermezcla para PCR Platinum® Blue (Invitrogen) para ejecutar la PCR de MAMA. La PCR se realizó a 95 °C 3 min, x 30 (95 °C 15 s, 58 °C 20 s, 72 °C 1 min) y 72 °C 5 min. Los productos de PCR se ejecutaron en un E-gel de 96 pocillos (Invitrogen) para visualizar las bandas.

Preparación de extracto y transferencia Western

Las cepas SBHS015, SBHS016 y SBHS017, que fueron diseñadas para contener variantes de RF1 modificadas, se cultivaron en 500 ml de TB a 37 °C agitando durante la noche en matraces de agitación Tunair. Las células se sedimentaron a 6.000 xg durante 15 minutos. El sedimento celular se lavó 2 veces con 6 ml de tampón S30 (Tris 10 mM, acetato de magnesio 14 mM y acetato de potasio 60 mM): 1 g de sedimento celular. Las células se volvieron a suspender luego en 2 ml de tampón S30: 1 g de sedimento celular. Las células resuspendidas se lisaron usando un homogenizador. Luego se clarificó el extracto 2 veces a 15.000 xg durante 30 minutos. El extracto se activó durante 1, 2 o 3 horas en un baño de agua a 30 °C. Se fabricó un anticuerpo anti-RF1 inoculando conejos con proteína RF1 de *E. coli* recombinante purificada que luego se purificó usando una matriz de afinidad (YenZym Antibodies LLC). La especificidad del anticuerpo se confirmó mediante ELISA y transferencias Western de la proteína recombinante. Las muestras de sedimento celular, lisado y extracto se procesaron en un gel SDS-PAGE y se transfirieron a una membrana de PVDF usando el sistema iBlot® (Invitrogen). Se usó el anticuerpo anti-RF1 primario seguido de un anticuerpo conjugado con fosfatasa alcalina anti-conejo secundario (Invitrogen). Las bandas se visualizaron usando una solución de sustrato cromogénico de fosfatasa alcalina que contenía 5-bromo-4-cloro-3-indolil-1-fosfato y nitroazul de tetrazolio (Invitrogen).

Tabla 5. Secuencias de oligonucleótidos para la PCR de OMAR y MAMA

Nombre del oligo	Secuencia del oligo (5' a 3')	SED ID NO:
1opRF1KR (OMAR)	GGGAAGTTGTAAGTACGGTTACGGTTCGCTGCGATCcCCTgaaCCaAGacGcTTtCGACGGGTAGACGCTTCGGCCTGTTGGCGTTTTGCC	140
1opRF1KRR (OMAR)	GGGAAGTTGTAAGTACGGTTACGGTTCGCTGCGATCcCCTgaaCCacgacGcTTtCGACGGGTAGACGCTTCGGCCTGTTGGCGTTTTGCC	141
1opRF1KRK (OMAR)	GGGAAGTTGTAAGTACGGTTACGGTTCGCTGCGATCcCCTgaaCCcttacGcTTtCGACGGGTAGACGCTTCGGCCTGTTGGCGTTTTGCC	142

Nombre del oligo	Secuencia del oligo (5' a 3')	SED ID NO:
3KR op-PCR (MAMA)	GCG ATC CCC TGA ACC AAG ACG C	143
3 KRR op-PCR (MAMA)	CGATCcCCtgaaCCacgacGc	144
3KRK op-PCR (MAMA)	TGCGATCcCCtgaaCCcttacGc	145
5 RF1 op-PCR (MAMA)	CGTGACGGGGATAACGAACGCC	146

Se identificaron tres cepas bacterianas recombinantes que habían incorporado satisfactoriamente mutaciones de RF1 en sus genomas. La cepa SBHS015 contiene la variante de RF1 de mutación doble N296K/I297R, denominada variante A13. La cepa SBHS016 contiene la variante de RF1 de mutación triple N296K/L297R/L298R, denominada variante A18. La cepa SBHS017 contiene la variante de RF1 de mutación triple N296K/I297R/I298K, denominada variante A17. Las tres cepas también expresan OmpT1 intacta, pues derivan de la cepa parental SBJY001. Cuando las cepas recombinantes se lisaron e incubaron durante varios períodos de tiempo (0, 1, 2 y 3 horas), las variantes de la proteína RF1 modificadas se escindieron eficazmente en todos los puntos de tiempo ensayados por el extracto libre de células, según lo determinó mediante el análisis de transferencia Western (datos no mostrados). Por el contrario, la escisión de RF1 de tipo silvestre no modificada por la cepa SBJY001 no se detectó durante los mismos períodos de tiempo, lo que indica que la RF1 de tipo silvestre no se escinde de manera eficaz en extractos libres de células que contienen OmpT1 intacta.

Este ejemplo demuestra que las cepas bacterianas recombinantes se diseñaron para expresar variantes modificadas de RF1, y que las variantes de RF1 se escindieron mediante extractos libres de células de cepas positivas en OmpT1.

Ejemplo 3

El Ejemplo 3 demuestra que las proteínas RF1 intactas modificadas para incluir sitios de escisión por OmpT1 en la región del bucle de conmutación tienen función RF1 de tipo silvestre.

Ensayo funcional de variantes de RF1 recombinantes

Para ensayar la función de las variantes de RF1 recombinantes descritas en el presente documento, se determinó la capacidad de las variantes de RF1 para terminar la traducción en un codón ámbar. Se expresó una proteína Fc con una mutación TAG en presencia de 500 nM de RF1 mutante o de tipo silvestre recombinante purificada de *E. coli* y 2 µM de aminoácido no natural en el extracto SBHS002 (una cepa con eliminación de OmpT). Se ejecutaron reacciones libres de células de 60 µl a 30 °C durante 5 horas en presencia de ¹⁴C-Leu. Las reacciones finales se centrifugaron, obteniéndose la fracción soluble, se procesaron en un gel reductor de SDS-PAGE, se transfirieron a una membrana de PVDF (Invitrogen) y se expusieron a una pantalla de fósforo durante una noche. La pantalla de fósforo se visualizó usando un Storm Imager, y se usó ImageQuant para determinar las intensidades de banda relativas. Las actividades relativas de los mutantes se determinaron comparando la cantidad de proteína Fc truncada con el control negativo (sin la adición de RF1 exógena) y el control positivo (con la adición de RF1 WT). El porcentaje de proteína truncada se determinó mediante la ecuación: (recuentos de proteínas truncadas/recuentos de proteínas totales) x 100 %. La actividad relativa de RF1 se determinó usando la ecuación: [(variante de proteína truncada-proteína truncada del control negativo)/(proteína RF1 WT truncada-proteína truncada del control negativo)] x 100 %.

Como se muestra en la **Figura 1**, las variantes de RF1 descritas en el presente documento tienen actividad RF1, según lo demostrado por el truncamiento de la traducción cuando se incorpora un aminoácido no nativo (pAzF) a un codón ámbar introducido en la posición S378 de la proteína Fc. En particular, la variante A13 de RF1 (que tiene sustituciones N296K/L297R) poseía niveles de actividad similares a los de la RF1 de tipo silvestre.

Este ejemplo demuestra que las proteínas RF1 intactas modificadas para incluir sitios de escisión por OmpT en la región del bucle de conmutación tienen actividad RF1 funcional (por ejemplo, supresión de ámbar reducida).

Ejemplo 4

El Ejemplo 4 demuestra el aumento de la incorporación de aminoácidos no naturales a la cadena pesada de IgG de la proteína Herceptina usando extractos libres de células que comprenden variantes de RF1 que tienen sitios de escisión por OmpT1 en la región del bucle de conmutación. Los extractos libres de células proceden de las cepas bacterianas descritas en el Ejemplo 2.

Métodos

Mutagénesis dirigida de la cadena pesada de Herceptina

Para introducir un aann en la cadena pesada de Herceptina, se mutó el molde de ADN codificante de Herceptina para introducir codones ámbar en diferentes posiciones de la secuencia de codificación. La mutagénesis dirigida se realizó usando un plásmido pYD que contenía la región codificante de Herceptin6xHis en el extremo C como el molde de ADN y oligonucleótidos sintéticos (Operón) que contengan codones ámbar en ambas direcciones, sentido y antisentido (Tabla 6). Los oligonucleótidos de cada mutación se mezclaron con el molde de ADN y la polimerasa Phusion® (Thermo, n.º de cat. F531s) hasta un volumen final de 20 µl. La concentración final de cada componente fue de 0,16 µM de cada oligonucleótido, 0,5 ng/µl de ADN molde, 0,02 U/µl de polimerasa Phusion® en tampón HF (Thermo) que contenía MgCl₂ 1,5 mM y dNTP 200 µM. La mezcla se incubó a 98 °C 5 min, 18 ciclos de PCR [98 °C 30 s, 55 °C 1 min, 72 °C 4 min], 10 min a 72 °C, y se almacenó a 4 °C durante hasta 16 h. Se añadió DpnI (NEB) a la mezcla hasta una concentración final de 0,6 U/µl y se incubó a 37 °C durante 1 h. Se transformaron 5 µl de cada mezcla en 50 µl de células de químicamente competentes según el procedimiento del fabricante (Invitrogen, placa de 96 pocillos MultiShot™ TOP10). Las células transformadas se recuperaron en 200 µl de SOC (Invitrogen) a 37 °C durante 1 h, y se sembraron en placas sobre agar Luria-Bertani (LB) complementado con 50 µg/ml de kanamicina (Teknova). Tras 24 h a 37 °C, se seleccionaron colonias usando Qpix2 (Genetix) en 200 µl de LB con glicerol al 7,5 % y 50 µg/ml de kanamicina, y se cultivaron a 37 °C durante 24 h, se usaron 20 µl de cultivo para la amplificación en círculo rodante y se secuenciaron mediante la extensión del cebador usando los cebadores T7 (5'-TAATACGACTCAC- TATAGG-3'; SEQ ID NO:147) y T7 term (5'-GCTAGTTATTGCTCAGCG-3'; SEQ ID NO:148) (Sequetech). La secuencia se analizó mediante el secuenciador (Gene Codes).

Tabla 6. Cebadores para introducir codones ámbar en las posiciones indicadas de Herceptina

Variante	Oligo sentido (SEQ ID NO:)	Oligo antisentido (SEQ ID NO:)
SP-00067_V422	CGTTGGCAGCAGGGTAATTAGTT CAGCTGCAGCGTTATG (149)	CATAACGCTGCAGCTGAACTAA TTACCCTGCTGCCAACG (150)
SP-00127_S415	GCAAGCTGACCGTCGATAAATA GCGTTGGCAGCAGGGTAATG (151)	CATTACCCTGCTGCCAACGCTA TTTATCGACGGTCAGCTTGC (152)
SP-00128_Q418	CGATAAAAGCCGTTGGTAGCAG GGTAATGTGTTTCAG (153)	CTGAACACATTACCCTGCTACC AACGGCTTTTATCG (154)
SP-00114_P343	GCAAAGCGAAAGGCCAATAGCG TGAACCGCAGGTC (155)	GACCTGCGGTTACGCTATTGG CCTTCGCTTTGC (156)
SP-00112_G341	GACGATCAGCAAAGCGAAATAG CAACCGCGTGAACCGCAG (157)	CTGCGGTTACGCGGTTGCTAT TTCGCTTTGCTGATCGTC (158)
SP-00102_K320	GCTGAATGGTAAAGAATACTAG TGCAAAGTGAGCAACAAGG (159)	CCTTGTTGCTCACTTTGCACTAG TATTCTTTACCATTTCAGC (160)
SP-00066_F404	CTGGACAGCGACGGTAGCTAGT TTCTGTATAGCAAGCTG (161)	CAGCTTGCTATACAGAACTAG CTACCGTCGCTGTCCAG (162)
SP-00113_Q342	GCAAAGCGAAAGGCTAGCCGCG TGAACCGCAG (163)	CTGCGGTTACGCGGCTAGCCT TTCGCTTTGC (164)
SP-00096_T299	GTGAGGAACAATAACAATAGCTA GTATCGCGTAGTGAGCGTGC (165)	GCACGCTCACTACGCGATACTA GCTATTGTATTGTTCTCAC (166)
SP-00120_Y373	GGTGAAGGGCTTTTAGCCGAGC GACATCGC (167)	GCGATGTCGCTCGGCTAAAAGC CCTTCACC (168)
SP-00094_N297	CGCGTGAGGAACAATACTAGAG CACGTATCGCGTAGTG (169)	CACTACGCGATACGTGCTCTAG TATTGTTCTCACGCG (170)

Variante	Oligo sentido (SEQ ID NO:)	Oligo antisentido (SEQ ID NO:)
SP-00125_F405	GACAGCGACGGTAGCTTCTAGCT GTATAGCAAGCTGAC (171)	GTCAGCTTGCTATACAGCTAGA AGCTACCGTCGCTGTC (172)

Preparación de extracto

5 Se cosechó la cepa SBHS016 de *E. coli* con la variante A18 de RF1 modificada a una densidad final de DO40-55, y se centrifugó a 14.000 g en una centrifuga Sharples Modelo AS14 durante 10 minutos para eliminar el medio agotado. Se volvió a suspender la pasta celular hasta la homogeneidad con una proporción de 6 ml/g de células de tampón S30 (MgAcO 14 mM, KAcO 60 mM, Tris 10 mM) y se centrifugó nuevamente a 14.000 g durante 10 minutos usando Sharples AS14 para la eliminación adicional del medio agotado. Se volvió a suspender la pasta de células clarificada resultante en S30 con una proporción de 2 ml/g de células, y las células se lisaron mediante una sola pasada a través de un homogenizador Avestin Emulsiflex C-55 a 117,21 MPa (17.000 psi). El homogenizado se clarificó por centrifugación a 14.000 g dos veces durante 30 minutos cada vez, y los sedimentos resultantes se desecharon. La solución de extracto celular resultante se incubó a 30 °C durante 2 horas y luego se centrifugó nuevamente a 14.000 g usando Sharples AS14 para la eliminación de partículas. Esta solución final se congeló en LN₂ y se almacenó a -80 °C hasta que fue necesaria para la síntesis de proteínas sin células.

15 Se descongelaron los extractos libres de células a temperatura ambiente y se incubaron con 50 uM yodoacetamida durante 30 min. Se realizaron reacciones libres de células a 30 °C durante hasta 10 h que contenían extracto tratado con yodoacetamida al 30 % (v/v) con glutamato de magnesio 8 mM, glutamato de amonio 10 mM, glutamato de potasio 130 mM, piruvato sódico 35 mM, AMP 1,2 mM, 0,86 mM de cada GMP, UMP y CMP, aminoácidos 2 mM (1 mM para la tirosina), oxalato de sodio 4 mM, putrescina 1 mM, espermidina 1,5 mM, fosfato de potasio 15 mM, RNAP T7 100 nM, DsbC de *E. coli* 2,5 uM, Levadura PDI 5 uM, glutatión oxidado (GSSG) 2 mM y aminoacil ARNt pN3F de ARNt de levadura 15 uM. Para marcar la proteína sintetizada con ¹⁴C, también se añadió 1-[U-¹⁴C]-leucina al 3,33 % (v/v) (300 mCi/mmol; GE Life Sciences, Piscataway, NJ) a la reacción. Las concentraciones del plásmido de variante TAG de cadena pesada y del plásmido de cadena ligera de tipo silvestre fueron de 7,5 ug/ml y 2,5 ug/ml, respectivamente. Como control, se realizó la expresión libres de células de la cadena ligera de tipo silvestre en paralelo con las variantes TAG.

30 Se seleccionaron 12 sitios difíciles de suprimir de la cadena pesada basándose en el estudio interno de los presentes inventores, que incluían N297, T299, K320, G341, Q342, P343, Y373, F404, F405, S415, Q418 e V422. Se escogió S136 como control positivo, que tiene una supresión relativamente alta. Se usaron dos extractos celulares, de las cepas SBJY001 (que expresa RF1 de tipo silvestre) y SBHS016 (que expresa RF1 modificada), para comparar la capacidad de los extractos para incorporar aann en estos sitios difíciles de suprimir.

35 Se llevaron a cabo 60 ul de reacciones sin células en placas de 24 pocillos. Una vez completadas las reacciones sin células, se realizó la precipitación con TCA para medir las proteínas totales y solubles sintetizadas. En paralelo, los geles no reductores y reductores fueron para el ensayo de autorradiografía. Para el gel no reductor, se mezclaron 4 ul de muestra, 8 ul de H₂O DI y 4 ul de 4 x tampón LDS (Invitrogen, Carlsbad, CA) antes de cargarlos en el gel. Para el gel reductor, se mezclaron 4 ul de muestra, 1 ul de DTT 1 M, 7 ul de H₂O DI y 4 ul de 4 x tampón LDS (Invitrogen, Carlsbad, CA) y se calentaron en una transferencia en caliente a 70 °C durante 5 minutos. Las muestras se analizaron con geles SDS-PAGE de Bis-Tris al 4~12 % (Invitrogen, Carlsbad, CA) de acuerdo con las recomendaciones del fabricante. Los geles se secaron y se analizaron mediante autorradiografía usando un Storm 840 PhosphorImager después de aproximadamente 16 horas de exposición.

$$45 \text{ Rendimiento de IgG} = \frac{\text{Intensidad de la banda de longitud completa de IgG desde el gel no reductor}}{\text{La suma de la intensidad de todas las bandas desde el gel reductor}} \times [\text{proteína soluble}]$$

donde la intensidad de la banda se determina con el software ImageQuant™ y la [proteína soluble] se estimó mediante el método de precipitación en TCA.

50 Se determinaron las supresiones del codón ámbar en diferentes sitios de la cadena pesada por [¹⁴C]-autorradiografía de geles reductores SDS-PAGE. Las variantes de TAG de cadena pesada de tipo silvestre de longitud completa y de cadena pesada suprimida se ejecutan a 49 Kd en SDS-PAGE. La cadena pesada no suprimida (truncada) de las variantes de TAG se ejecuta en un peso molecular más bajo.

55 La supresión de TAG en la cadena pesada se predice según:

$$\text{Supresión} = \frac{\text{Intensidad de la banda de la variante de TAG de cadena pesada suprimida}}{\text{Intensidad de la banda de la cadena pesada de tipo silvestre}}$$

donde la intensidad de la banda se determina con el software ImageQuant™.

60

Para generar los materiales para la purificación, las reacciones se escalaron hasta 1 ml en placas de Petri de 10 cm bajo la misma condición.

Producción de ARNt

5 Las transcripciones de todos los transcritos de ARNt tRNA^{phe}_{CUA} se realizaron en las siguientes condiciones: 20-50 ng/μl de pYD318-tRNA^{phe} AAA-HDV-T7trm, NaCl 40 mM, MgCl₂ 10 mM, DTT 10 mM, dNTP 4 mM, espermidina 2,5 mM, PPIasa 1 U/ml, 2,5 mg/ml de ARN polimerasa T7 y Tris 40 mM (pH 7.9). Las moléculas de ARNt se separaron del ARN parental y del producto de ARN de ribozima HDV mediante cromatografía de filtración en gel usando una resina tándem Sephacryl 100 o 300 en columnas XK50/100. Las columnas de calibrado se desarrollaron en Tris 50 mM (pH 6,5) y NaCl 250 mM. Las fracciones que contenían ARNt se combinaron, se mezclaron con 1/10 de volumen de acetato de sodio 3 M (pH 5,2) y se añadió un volumen igual de isopropanol para precipitar el ARN. El ARNt se almacenó como un sedimento o se volvió a suspender en Tris 10 mM (pH 6,5) y EDTA 0,1 mM.

15 Los ARNt para su uso en reacciones de aminoacilación se tratan con polinucleótido quinasa T4 en MES 100 mM, pH 5,5, MgCl₂ 10 mM y 2-mercaptoetanol 10 mM durante 1 h para eliminar el fosfato 2',3'-cíclico que abandona los grupos 2',3'-OH en el extremo 3' del ARNt. El ARNt tratado con T4 PNK se extrajo en fenol:cloroformo: alcohol isoamílico e se intercambió con tampón usando una columna G25 que eliminó el fosfato inorgánico y el exceso de fenol. El ARNt se precipitó con isopropanol o etanol, y se volvió a suspender en Tris 10 mM (pH 6,5) y EDTA 0,1 mM. El ARNt se replegó calentando el ARNt a 70 °C durante 20 minutos. Luego, se añadió MgCl₂ 10 mM y la mezcla se equilibró lentamente a temperatura ambiente.

25 La escisión de ribozimas HDV de los transcritos de ARNt, mientras se producen extremos 3' homogéneos, deja un resto de fosfato 2'-3' cíclico que interfiere con la aminoacilación posterior. Se ha encontrado que esto puede eliminarse usando polinucleótido quinasa T4 (PNK). El ARNt 40 mM se incubó a 37 °C con 0,050 mg/ml de PNK en MES 50 mM (pH 5,5), MgCl₂ 10 mM, NaCl 300 mM y EDTA 0,1 mM. La desfosforilación se analizó mediante dos métodos diferentes. La desfosforilación se confirmó mediante electroforesis en gel desnaturalizante. Como se ha informado, el ARNt desfosforilado tiene una movilidad reducida en la electroforesis en geles de ácido/urea. Se diluyeron partes alícuotas que contenían 3 μg de ARNt fosforilado 2 veces en tampón de carga (acetato de sodio 100 mM (pH 5,2), urea 7 M, 1 mg/ml de colorante azul de bromofenol) y se cargaron en una acrilamida 19:1 al 6,5 %, acetato de sodio 100 mM (pH 5,2), gel de urea 7 M (40 cm x 34 cm) y se sometió a electroforesis durante la noche a 40 W. Los geles se tiñeron con azul de metileno al 0,06 %, acetato de sodio 0,5 M (pH 5,2) durante 30 minutos y se destiñeron con agua desionizada. Ambos ensayos indicaron una desfosforilación significativa después de solo 5 minutos. La desfosforilación se completó esencialmente después de 1 hora, el ARNt se replegó calentando a 70 °C, la adición de MgCl₂ 10 mM y luego se enfrió lentamente a temperatura ambiente. La concentración de ARN se midió usando un espectrofotómetro Nano-Drop 1000 (Thermo Scientific) y se confirmó mediante electroforesis en gel.

Aminoacilación del ARNt supresor del ámbar

40 Las condiciones para la aminoacilación no natural son HEPES 50 mM pH 8,1, KCl 40 mM, MgCl₂ 75 mM, ATP 5 mM, tRNA^{phe}_{CUA} 8-40 μM, DTT 10 mM, aminoácido 2 mM (pN₃F) y PheRS T415A D243A 40 μM. La determinación del porcentaje de aminoacilación de tRNA^{phe}_{CUA} se lleva a cabo mediante resolución de HPLC HIC de los restos aminoacilados y no aminoacilados de ARNt. Este método permite controlar el grado de aminoacilación del presente ARNt una vez procesado y ya listo para su uso para su incorporación a proteínas. Las reacciones se incuban a 37 °C durante 15 minutos y se inactivan con 2,5 volúmenes de acetato de sodio 300 mM pH 5,5. La muestra inactivada se extrae con fenol: cloroformo:alcohol isoamílico 25:24:1 pH 5,2 (ambion) y se agita con formación de vórtice durante 2 min. Luego se centrifuga a 14.000 rcf durante 10-30 min a 4 °C para separar las fases acuosa (ARNt) y orgánica (proteína). La fase acuosa (que contiene ARNt) se retira y se añade a una columna de exclusión por tamaño de resina G25 sephadex pre-equilibrada (NaOAc 300 mM) que se separa según el tamaño de la molécula. El eluyente se mezcla con 2,5 volúmenes de etanol al 100 % y se incuba a -80 °C durante 15-30 minutos, y se centrifuga a 12.000-14.000 rcf durante 30-45 minutos. El ARNt aminoacilado ahora se encuentra en un sedimento que puede almacenarse a -80 °C o volverse a suspender en un tampón ligeramente ácido para inyectarse en la HPLC y/o usar en reacciones de OCFS.

55 La columna de HPLC C5 HIC se equilibra en tampón A (fosfato de potasio 50 mM y sulfato de amonio 1,5 M, pH 5,7) hasta que la traza UV no fluctúa de cero. Se mezclan 1-10 μg de ARNt con 100 μl de tampón A x2 (fosfato de potasio 100 mM y sulfato de amonio 3 M). La muestra se inyecta y se ejecuta en un gradiente de tampón A a tampón B (fosfato de potasio 50 mM e isopropanol al 5 %) durante 50 minutos.

Incorporación de pN₃F a mutantes turboGFP TAG:

60 Para controlar la fluorescencia de la proteína verde fluorescente en una construcción en la que hay un codón ámbar (codón de parada), se clonó el ADN que codifica turboGFP (Evrogen, Rusia) en el presente vector de expresión de OCFS pYD317. Se insertó un codón de parada (TAG) mediante mutagénesis por PCR solapada en los nucleótidos correspondientes al aminoácido Lisina 37, Tirosina 50 y Glutamato 205 (y combinaciones) de acuerdo con la estructura cristalina de turboGFP (pdb 2G6X). Por lo tanto, cualquier supresión del codón de parada con un ARNt cargado resultará en fluorescencia. Las reacciones se incubaron a 30 °C en un espectrofotómetro (Molecular Devices,

SpectraMaxM5) durante cinco horas con una cubierta adhesiva (VWR, 9503130) y la intensidad de fluorescencia se midió a intervalos de 10 minutos, $\lambda_{Ex} = 476 \text{ nm}$ y $\lambda_{Em} = 510$. Se añadió la mezcla de reacción OCFS inmediatamente a la microplaca con inhibidor para un volumen de reacción final de 25 μl que contenía 30 % de extracto de S30, 24 $\mu\text{g/ml}$ de ARN polimerasa de T7, L-tirosina 1 mM (Sigma, T8566), premezcla*, pN₃F-tRNA^{Phe}_{CUA} 10-60 μM o tRNA^{Phe}_{CUA} sin cargar, y plásmido turboGFP 3 nM en agua tratada con DEPC (G Biosciences, 786-109). Se usó una reacción de control positivo usando turboGFP sin el codón de parada para asegurar que las reacciones procedieran con velocidades similares a las observadas previamente, mientras que las reacciones que contenían turboGFP Y50TAG también se realizaron sin ARNt para garantizar que no se detectara fluorescencia (control negativo). Las eficacias de supresión se calcularon mediante la comparación de la fluorescencia del control positivo con la fluorescencia del molde que contiene el codón ámbar.

Como se muestra en la **Figura 2**, la expresión de las variantes TAG de cadena pesada de IgG en el extracto libre de células SBJY001, que expresa OmpT1 y RF-1 de tipo silvestre, dio lugar a un rendimiento relativamente bajo de IgG soluble de longitud completa. Por el contrario, la expresión de las variantes TAG de cadena pesada de IgG en el extracto libre de células SBHS016, que expresa OmpT1 y una proteína RF-1 modificada (variante A18) que contiene la sustitución triple N296K/L297R/L298R, dio lugar a un rendimiento relativamente alto de IgG de longitud completa. Además, como se muestra en la **Figura 3**, la expresión de las variantes TAG de cadena pesada de IgG en el extracto libre de células SBJY001 dio lugar a una supresión ámbar relativamente pobre (es decir, la proteína de cadena pesada se truncó). Por el contrario, la expresión de las variantes TAG de la cadena pesada de IgG en el extracto libre de células SBHS016 dio lugar a una supresión ámbar relativamente alta (es decir, se truncó mucho menos de la proteína de cadena pesada), lo que corresponde a los rendimientos sustancialmente más altos de IgG de longitud completa observados en **Figura 2**.

Este ejemplo demuestra que la escisión por OmpT1 de RF1 proporciona una mejora espectacular en el rendimiento de las cadenas pesadas que incorporan el aann deseado en un codón ámbar en la secuencia de codificación, en comparación con la RF1 intacta que no es escindible por OmpT1.

LISTADO DE SECUENCIAS

<110> Thanos, Christopher D.
Murray, Christopher J.
Yang, Junhao Stephenson,
Heather Sutro Biopharma, Inc.

<120> Inactivación proteolítica de proteínas seleccionadas en extractos bacterianos para mejorar la expresión

<130> 91200-889617

<140> WO todavía sin asignar
<141> Todavía sin asignar

<150> US 61/713.245
<151> 12/10/2012

<160> 172

<170> FastSEQ para Windows Versión 4.0

<210> 1
<211> 360
<212> PRT
<213> *Escherichia coli*

<220>
<223> Factor de liberación 1 (RF1)

<400> 1

ES 2 694 683 T3

Met Lys Pro Ser Ile Val Ala Lys Leu Glu Ala Leu His Glu Arg His
 1 5 10 15
 Glu Glu Val Gln Ala Leu Leu Gly Asp Ala Gln Thr Ile Ala Asp Gln
 20 25 30
 Glu Arg Phe Arg Ala Leu Ser Arg Glu Tyr Ala Gln Leu Ser Asp Val
 35 40 45
 Ser Arg Cys Phe Thr Asp Trp Gln Gln Val Gln Glu Asp Ile Glu Thr
 50 55 60
 Ala Gln Met Met Leu Asp Asp Pro Glu Met Arg Glu Met Ala Gln Asp
 65 70 75 80
 Glu Leu Arg Glu Ala Lys Glu Lys Ser Glu Gln Leu Glu Gln Gln Leu
 85 90 95
 Gln Val Leu Leu Leu Pro Lys Asp Pro Asp Asp Glu Arg Asn Ala Phe
 100 105 110
 Leu Glu Val Arg Ala Gly Thr Gly Gly Asp Glu Ala Ala Leu Phe Ala
 115 120 125
 Gly Asp Leu Phe Arg Met Tyr Ser Arg Tyr Ala Glu Ala Arg Arg Trp
 130 135 140
 Arg Val Glu Ile Met Ser Ala Ser Glu Gly Glu His Gly Gly Tyr Lys
 145 150 155 160
 Glu Ile Ile Ala Lys Ile Ser Gly Asp Gly Val Tyr Gly Arg Leu Lys
 165 170 175
 Phe Glu Ser Gly Gly His Arg Val Gln Arg Val Pro Ala Thr Glu Ser
 180 185 190
 Gln Gly Arg Ile His Thr Ser Ala Cys Thr Val Ala Val Met Pro Glu
 195 200 205
 Leu Pro Asp Ala Glu Leu Pro Asp Ile Asn Pro Ala Asp Leu Arg Ile
 210 215 220
 Asp Thr Phe Arg Ser Ser Gly Ala Gly Gly Gln His Val Asn Thr Thr
 225 230 235 240

Asp Ser Ala Ile Arg Ile Thr His Leu Pro Thr Gly Ile Val Val Glu
 245 250 255
 Cys Gln Asp Glu Arg Ser Gln His Lys Asn Lys Ala Lys Ala Leu Ser
 260 265 270
 Val Leu Gly Ala Arg Ile His Ala Ala Glu Met Ala Lys Arg Gln Gln
 275 280 285
 Ala Glu Ala Ser Thr Arg Arg Asn Leu Leu Gly Ser Gly Asp Arg Ser
 290 295 300
 Asp Arg Asn Arg Thr Tyr Asn Phe Pro Gln Gly Arg Val Thr Asp His
 305 310 315 320
 Arg Ile Asn Leu Thr Leu Tyr Arg Leu Asp Glu Val Met Glu Gly Lys
 325 330 335
 Leu Asp Met Leu Ile Glu Pro Ile Ile Gln Glu His Gln Ala Asp Gln
 340 345 350
 Leu Ala Ala Leu Ser Glu Gln Glu
 355 360

<210> 2

5 <211> 365

<212> PRT

<213> *Escherichia coli*

10 <220>

<223> Factor de liberación 2 (RF2)

<400> 2

ES 2 694 683 T3

```

Met Phe Glu Ile Asn Pro Val Asn Asn Arg Ile Gln Asp Leu Thr Glu
 1      5      10
Arg Ser Asp Val Leu Arg Gly Tyr Leu Asp Tyr Asp Ala Lys Lys Glu
 20
Arg Leu Glu Glu Val Asn Ala Glu Leu Glu Gln Pro Asp Val Trp Asn
 35      40      45
Glu Pro Glu Arg Ala Gln Ala Leu Gly Lys Glu Arg Ser Ser Leu Glu
 50      55      60
Ala Val Val Asp Thr Leu Asp Gln Met Lys Gln Gly Leu Glu Asp Val
 65      70      75      80
Ser Gly Leu Leu Glu Leu Ala Val Glu Ala Asp Asp Glu Glu Thr Phe
 85      90      95
Asn Glu Ala Val Ala Glu Leu Asp Ala Leu Glu Glu Lys Leu Ala Gln
 100      105      110
Leu Glu Phe Arg Arg Met Phe Ser Gly Glu Tyr Asp Ser Ala Asp Cys
 115      120      125
Tyr Leu Asp Ile Gln Ala Gly Ser Gly Gly Thr Glu Ala Gln Asp Trp
 130      135      140
Ala Ser Met Leu Glu Arg Met Tyr Leu Arg Trp Ala Glu Ser Arg Gly
 145      150      155      160
Phe Lys Thr Glu Ile Ile Glu Glu Ser Glu Gly Glu Val Ala Gly Ile
 165      170      175
Lys Ser Val Thr Ile Lys Ile Ser Gly Asp Tyr Ala Tyr Gly Trp Leu
 180      185      190
Arg Thr Glu Thr Gly Val His Arg Leu Val Arg Lys Ser Pro Phe Asp
 195      200      205
Ser Gly Gly Arg Arg His Thr Ser Phe Ser Ser Ala Phe Val Tyr Pro
 210      215      220
Glu Val Asp Asp Asp Ile Asp Ile Glu Ile Asn Pro Ala Asp Leu Arg
 225      230      235      240
Ile Asp Val Tyr Arg Thr Ser Gly Ala Gly Gly Gln His Val Asn Arg
 245      250      255
Thr Glu Ser Ala Val Arg Ile Thr His Ile Pro Thr Gly Ile Val Thr
 260      265      270
Gln Cys Gln Asn Asp Arg Ser Gln His Lys Asn Lys Asp Gln Ala Met
 275      280      285

Lys Gln Met Lys Ala Lys Leu Tyr Glu Leu Glu Met Gln Lys Lys Asn
 290      295      300
Ala Glu Lys Gln Ala Met Glu Asp Asn Lys Ser Asp Ile Gly Trp Gly
 305      310      315      320
Ser Gln Ile Arg Ser Tyr Val Leu Asp Asp Ser Arg Ile Lys Asp Leu
 325      330      335
Arg Thr Gly Val Glu Thr Arg Asn Thr Gln Ala Val Leu Asp Gly Ser
 340      345      350
Leu Asp Gln Phe Ile Glu Ala Ser Leu Lys Ala Gly Leu
 355      360      365

```

<210> 3
 <211> 317
 <212> PRT
 <213> *Escherichia coli*

<220>
 <223> Proteína de la membrana externa T1 (OmpT1)

<220>
 <221> SEÑAL
 <222> (1)...(20)
 <223> péptido señal

<400> 3

```

Met Arg Ala Lys Leu Leu Gly Ile Val Leu Thr Thr Pro Ile Ala Ile
-20          -15          -10          -5
Ser Ser Phe Ala Ser Thr Glu Thr Leu Ser Phe Thr Pro Asp Asn Ile
          1          5          10
Asn Ala Asp Ile Ser Leu Gly Thr Leu Ser Gly Lys Thr Lys Glu Arg
          15          20          25
Val Tyr Leu Ala Glu Glu Gly Gly Arg Lys Val Ser Gln Leu Asp Trp
          30          35          40
Lys Phe Asn Asn Ala Ala Ile Ile Lys Gly Ala Ile Asn Trp Asp Leu
          45          50          55          60
Met Pro Gln Ile Ser Ile Gly Ala Ala Gly Trp Thr Thr Leu Gly Ser
          65          70          75
Arg Gly Gly Asn Met Val Asp Gln Asp Trp Met Asp Ser Ser Asn Pro
          80          85          90
Gly Thr Trp Thr Asp Glu Ser Arg His Pro Asp Thr Gln Leu Asn Tyr
          95          100          105
Ala Asn Glu Phe Asp Leu Asn Ile Lys Gly Trp Leu Leu Asn Glu Pro
          110          115          120
Asn Tyr Arg Leu Gly Leu Met Ala Gly Tyr Gln Glu Ser Arg Tyr Ser
          125          130          135          140
Phe Thr Ala Arg Gly Gly Ser Tyr Ile Tyr Ser Ser Glu Glu Gly Phe
          145          150          155
Arg Asp Asp Ile Gly Ser Phe Pro Asn Gly Glu Arg Ala Ile Gly Tyr
          160          165          170
Lys Gln Arg Phe Lys Met Pro Tyr Ile Gly Leu Thr Gly Ser Tyr Arg
          175          180          185
Tyr Glu Asp Phe Glu Leu Gly Gly Thr Phe Lys Tyr Ser Gly Trp Val
          190          195          200
Glu Ser Ser Asp Asn Asp Glu His Tyr Asp Pro Gly Lys Arg Ile Thr
          205          210          215          220
Tyr Arg Ser Lys Val Lys Asp Gln Asn Tyr Tyr Ser Val Ala Val Asn
          225          230          235
Ala Gly Tyr Tyr Val Thr Pro Asn Ala Lys Val Tyr Val Glu Gly Ala
          240          245          250
Trp Asn Arg Val Thr Asn Lys Lys Gly Asn Thr Ser Leu Tyr Asp His
          255          260          265
Asn Asn Asn Thr Ser Asp Tyr Ser Lys Asn Gly Ala Gly Ile Glu Asn
          270          275          280
          Tyr Asn Phe Ile Thr Thr Ala Gly Leu Lys Tyr Thr Phe
          285          290          295

```

- 5 <210> 4
- <211> 18
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> región sintética del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
- <400> 4

```

          Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Leu Gly Ser Gly Asp
          1          5          10          15
15 Arg Ser

```

- <210> 5
- <211> 18
- <212> PRT

ES 2 694 683 T3

<213> Secuencia artificial

<220>

5 <223> mutación N296K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 5

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Leu Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

10

<210> 6

<211> 18

<212> PRT

<213> Secuencia artificial

15

<220>

<223> mutación N296R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

20

<400> 6

Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Leu Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

25

<210> 7

<211> 18

<212> PRT

<213> Secuencia artificial

30

<220>

<223> mutación L297K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 7

Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Lys Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

35

<210> 8

<211> 18

<212> PRT

<213> Secuencia artificial

40

<220>

<223> mutación L297R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

45

<400> 8

Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Arg Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

50

<210> 9

<211> 18

ES 2 694 683 T3

<212> PRT
<213> Secuencia artificial

5 <220>
<223> mutación L297V sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 9

10 Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Val Leu Gly Ser Gly Asp
1 5 10 15
Arg Ser

15 <210> 10
<211> 18
<212> PRT
<213> Secuencia artificial

20 <220>
<223> mutación L298K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 10

Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Lys Gly Ser Gly Asp
1 5 10 15
Arg Ser

25 <210> 11
<211> 18
<212> PRT
<213> Secuencia artificial

30 <220>
<223> mutación L298R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 11

35 Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Arg Gly Ser Gly Asp
1 5 10 15
Arg Ser

40 <210> 12
<211> 18
<212> PRT
<213> Secuencia artificial

45 <220>
<223> mutación N296K,L297K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 12

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Lys Leu Gly Ser Gly Asp
1 5 10 15
Arg Ser

50 <210> 13
<211> 18

ES 2 694 683 T3

<212> PRT
 <213> Secuencia artificial

5 <220>
 <223> mutación N296K,L297R = A13 sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 13

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Arg Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

10

<210> 14
 <211> 18
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <223> mutación N296K,L297V sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

20

<400> 14

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Val Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

25

<210> 15
 <211> 18
 <212> PRT
 <213> Secuencia artificial

30

<220>
 <223> mutación N296K,L297V sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

35

Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Lys Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

40

<210> 16
 <211> 18
 <212> PRT
 <213> Secuencia artificial

45

<220>
 <223> mutación N296R,L297R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 16

Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Arg Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

50

<210> 17

ES 2 694 683 T3

<211> 18
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> mutación N296R,L297V sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 17

10

Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Val Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

<210> 18
 <211> 18
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <223> mutación N296K,L297K,L298K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

20

<400> 18

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Lys Lys Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

25

<210> 19
 <211> 18
 <212> PRT
 <213> Secuencia artificial

30

<220>
 <223> mutación N296K,L297K,L298R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

35

<400> 19

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Lys Arg Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

40

<210> 20
 <211> 18
 <212> PRT
 <213> Secuencia artificial

45

<220>
 <223> mutación N296K,L297R,L298K = A17 sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 20

Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Arg Lys Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

50

ES 2 694 683 T3

<210> 21
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> mutación N296K,L297R,L298R = A18 sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 10
 <400> 21

 Gln Gln Ala Glu Ala Ser Thr Arg Arg Lys Arg Arg Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

 <210> 22
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 15
 <220>
 <223> mutación N296R,L297R,L298R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 20
 <400> 22

 Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Arg Arg Gly Ser Gly Asp
 1 5 10 15
 Arg Ser
 25

 <210> 23
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 30
 <220>
 <223> mutación N296R,L297K,L298R sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 35
 <400> 23

 Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Lys Arg Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

 <210> 24
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 40
 <220>
 <223> mutación N296R,L297R,L298K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 45
 <400> 24
 50

 Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Arg Lys Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

ES 2 694 683 T3

<210> 25
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> mutación N296R,L297K,L298K sintética de la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 10
 <400> 25

 Gln Gln Ala Glu Ala Ser Thr Arg Arg Arg Lys Lys Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

 <210> 26
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 15
 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 20
 <400> 26

 Gln Ala Arg Arg Gly Ser Thr Arg Arg Asn Leu Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser
 25
 <210> 27
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 30
 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 35
 <400> 27

 Gln Gln Ala Arg Arg Gly Thr Arg Arg Asn Leu Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser
 40
 <210> 28
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 45
 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304
 50
 <400> 28

 Gln Gln Ala Ala Arg Arg Gly Arg Arg Asn Leu Leu Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

ES 2 694 683 T3

<211> 18
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 33

10

Gln Gln Ala Glu Ala Ser Thr Arg Ala Arg Arg Gly Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

<210> 34
 <211> 18
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

20

<400> 34

Gln Gln Ala Glu Ala Ser Thr Arg Arg Ala Arg Arg Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

25

<210> 35
 <211> 18
 <212> PRT
 <213> Secuencia artificial

30

<220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

35

<400> 35

Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Ala Arg Arg Gly Gly Asp
 1 5 10 15
 Arg Ser

40

<210> 36
 <211> 18
 <212> PRT
 <213> Secuencia artificial

45

<220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 36

Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Ala Arg Arg Gly Asp
 1 5 10 15
 Arg Ser

50

<210> 37

ES 2 694 683 T3

<211> 18
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 37

10
 Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Leu Ala Arg Arg Gly
 1 5 10 15
 Arg Ser

<210> 38
 <211> 18
 <212> PRT
 <213> Secuencia artificial

15 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 38

20
 Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Leu Gly Ala Arg Arg
 1 5 10 15
 Gly Ser

25 <210> 39
 <211> 18
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 39

35
 Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Leu Gly Ser Ala Arg
 1 5 10 15
 Arg Gly

40 <210> 40
 <211> 18
 <212> PRT
 <213> Secuencia artificial

45 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 40

50
 Gln Gln Ala Glu Ala Ser Thr Arg Arg Asn Leu Leu Gly Ser Gly Ala
 1 5 10 15
 Arg Arg

<210> 41

ES 2 694 683 T3

<211> 18
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

<400> 41

10

Gln Gln Ala Trp Leu Ala Ala Arg Arg Gly Arg Gly Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

<210> 42
 <211> 18
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

20

<400> 42

Gln Gln Ala Glu Trp Leu Ala Ala Arg Arg Gly Arg Gly Ser Gly Asp
 1 5 10 15
 Arg Ser

25

<210> 43
 <211> 18
 <212> PRT
 <213> Secuencia artificial

30

<220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

35

<400> 43

Gln Gln Ala Glu Ala Trp Leu Ala Ala Arg Arg Gly Arg Gly Gly Asp
 1 5 10 15
 Arg Ser

40

<210> 44
 <211> 18
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

45

<400> 44

Gln Gln Trp Gly Gly Arg Trp Ala Arg Lys Lys Gly Thr Ile Gly Asp
 1 5 10 15
 Arg Ser

50

<210> 45

ES 2 694 683 T3

<211> 18
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

10 <400> 45

```

  Gln Gln Ala Trp Gly Gly Arg Trp Ala Arg Lys Lys Gly Thr Ile Asp
  1      5      10      15
  Arg Ser
  
```

<210> 46
 <211> 18
 <212> PRT
 <213> Secuencia artificial

15 <220>
 <223> secuencia del péptido de escisión por OmpT sintética insertada en la región del bucle de conmutación del Factor de liberación 1 (RF1), aminoácidos 287 a 304

20 <400> 46

```

  Gln Gln Ala Glu Trp Gly Gly Arg Trp Ala Arg Lys Lys Gly Thr Ile
  1      5      10      15
  Arg Ser
  
```

25 <210> 47
 <211> 4
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> péptido de escisión por OmpT sintético

35 <400> 47

```

  Ala Arg Arg Gly
  1
  
```

40 <210> 48
 <211> 9
 <212> PRT
 <213> Secuencia artificial

45 <220>
 <223> péptido de escisión por OmpT sintético

<400> 48

```

  Trp Leu Ala Ala Arg Arg Gly Arg Gly
  1      5
  
```

50 <210> 49
 <211> 12
 <212> PRT
 <213> Secuencia artificial

55 <220>

ES 2 694 683 T3

<223> péptido de escisión por OmpT sintético

<400> 49

Trp Gly Gly Arg Trp Ala Arg Lys Lys Gly Thr Ile
 1 5 10

5

<210> 50

<211> 9

<212> PRT

10 <213> Secuencia artificial

<220>

<223> secuencia C-terminal sintética de la forma corta de la proteína Gam del fago lambda (GamS)

15 <400> 50

Gly Gly Ser His His His His His His
 1 5

20

<210> 51

<211> 42

<212> ADN

<213> Secuencia artificial

<220>

25 <223> forma corta sintética del cebador de amplificación del gen de la proteína Gam del fago lambda (GamS)

<400> 51

atatatcata tgaacgctta ttacattcag gatcgtcttg ag 42

30

<210> 52

<211> 72

<212> ADN

<213> Secuencia artificial

35

<220>

<223> forma corta sintética del cebador de amplificación del gen de la proteína Gam del fago lambda (GamS)

<400> 52

atatatgtcg acttaatgat gatgatgatg atgagaaccc cctacctctg aatcaatatc 60
 aacctggtgg tg 72

40

<210> 53

<211> 58

<212> ADN

45 <213> Secuencia artificial

<220>

<223> fragmento 5' sintético que incluye el promotor T7, región constante de la secuencia N-terminal y el cebador oligo de amplificación por PCR de primera etapa del sitio de mutación 5chiT2PT7

50

<400> 53

gctgactagc gtaccacgtg gctggtggcc gattcattaa tgcagctggc acgacagg 58

55

<210> 54

<211> 60

<212> ADN

<213> Secuencia artificial

<220>

ES 2 694 683 T3

<223> fragmento 3' sintético que incluye sitio de mutación, región C-terminal constante y cebador oligo de amplificación por PCR de primera etapa de las secuencias del terminador T7 3chiT2TT7

5 <400> 54
gcgtactagc gtaccacgtg gctggtggcg gtgagtttct tccttcatta cagaaacggc 60

10 <210> 55
<211> 28
<212> ADN
<213> Secuencia artificial

15 <220>
<223> ensamblaje sintético del fragmento 5' y del fragmento 3' solapado al cebador individual de PCR de segunda etapa 5chiT2

20 <400> 55
gcgtactagc gtaccacgtg gctggtgg 28

25 <210> 56
<211> 35
<212> ADN
<213> Secuencia artificial

30 <220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 WT

35 <400> 56
gctcgatgat cctgaaatgc gtgagatggc gcagg 35

40 <210> 57
<211> 35
<212> ADN
<213> Secuencia artificial

45 <220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 M74R

50 <400> 57
gctcgatgat cctgaacgcc gtgagatggc gcagg 35

55 <210> 58
<211> 38
<212> ADN
<213> Secuencia artificial

60 <220>
<223> cebador directo de selección del sitio de escisión por OmpT de la variante de RF1 E76K sintético

65 <400> 58
cgatgatcct gaaatcgta agatggcgca ggatgaac 38

70 <210> 59
<211> 40
<212> ADN
<213> Secuencia artificial

75 <220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 E84K

80 <400> 59
caggatgaac tgcgcaaagc taaagaaaa agcgagcaac 40

85 <210> 60
<211> 40
<212> ADN
<213> Secuencia artificial

ES 2 694 683 T3

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 A85R

5 <400> 60
caggatgaac tgcgcaacg taaagaaaa agcgagcaac 40

<210> 61
<211> 44
10 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 E87R

15 <400> 61
ggatgaactg cgcgaagcta aacgtaaaag cgagcaactg gaac 44

<210> 62
<211> 38
20 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 E108R

25 <400> 62
gccaaaagat cctgatgacc gtcgtaacgc cttctctg 38

<210> 63
<211> 34
30 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 T293R

35 <400> 63
caacaggccg aagcgctctg ccgctgtaac ctgc 34

40 <210> 64
<211> 34
<212> ADN
<213> Secuencia artificial

45 <220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K

50 <400> 64
gctgtaccg gtcgtaaaact gctggggagt ggcg 34

<210> 65
<211> 36
55 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 S304K

60 <400> 65
gggagtggcg atcgcaagga ccgtaaccgt acttac 36

<210> 66
<211> 46
65 <212> ADN
<213> Secuencia artificial

ES 2 694 683 T3

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297V

5 <400> 66
cgaagcgtct acccgtcgta aagttctggg gagtggcgat cgcagc 46

<210> 67
<211> 46
10 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297K sintético

15 <400> 67
cgaagcgtct acccgtcgta aaaagctggg gagtggcgat cgcagc 46

<210> 68
<211> 46
20 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297R sintético

25 <400> 68
cgaagcgtct acccgtcgta aacgtctggg gagtggcgat cgcagc 46

<210> 69
<211> 46
30 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296R/I297V

35 <400> 69
cgaagcgtct acccgtcgtc gcgttctggg gagtggcgat cgcagc 46

<210> 70
<211> 46
40 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296R/I297K

45 <400> 70
cgaagcgtct acccgtcgtc gcaagctggg gagtggcgat cgcagc 46

<210> 71
<211> 46
50 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296R/I297R

55 <400> 71
cgaagcgtct acccgtcgtc gccgtctggg gagtggcgat cgcagc 46

<210> 72
<211> 55
60 <212> ADN
<213> Secuencia artificial

65

ES 2 694 683 T3

<220>
 <223> cebador directo sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297R/I298K
 5
 <400> 72
 caggccgaag cgtctaccg tcgtaaact aaggggagtg gcgatcgag cgacc 55
 <210> 73
 <211> 55
 <212> ADN
 <213> Secuencia artificial
 10
 <220>
 <223> cebador directo de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297R/I298R sintético
 15
 <400> 73
 caggccgaag cgtctaccg tcgtaaact cgggggagtg gcgatcgag cgacc 55
 20
 <210> 74
 <211> 35
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 WT
 30
 <400> 74
 cctgcgcat ctcacgcatt tcaggatcat cgagc 35
 <210> 75
 <211> 35
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 M74R
 40
 <400> 75
 cctgcgcat ctcacggcgt tcaggatcat cgagc 35
 <210> 76
 <211> 38
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> cebador inverso de selección del sitio de escisión por OmpT de la variante de RF1 E76K sintético
 50
 <400> 76
 gttcatcctg cgccatctta cgcatttcag gatcatcg 38
 <210> 77
 <211> 40
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> cebador inverso de selección del sitio de escisión por OmpT de la variante de RF1 E84K sintético
 60
 <400> 77
 gttgctcgt ttttcttta gcttgcgca gttcatcctg 40
 65
 <210> 78
 <211> 40

ES 2 694 683 T3

<212> ADN
<213> Secuencia artificial

5 <220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 A85R

<400> 78
gttgctcgct ttttcttta cgttcgcgca gttcatcctg 40

10 <210> 79
<211> 44
<212> ADN
<213> Secuencia artificial

15 <220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 E87R

<400> 79
gtccagttg ctcgcttta cgttagctt cgcgcagttc atcc 44

20 <210> 80
<211> 38
<212> ADN
<213> Secuencia artificial

25 <220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 E108R

30 <400> 80
cgaggaaggc gttacgacgg tcatcaggat ctttggc 38

35 <210> 81
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 T293R

40 <400> 81
gcaggttacg acggcgagac gcttcggcct gttg 34

45 <210> 82
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K

50 <400> 82
cgccactccc cagcagttta cgacgggtag acgc 34

55 <210> 83
<211> 36
<212> ADN
<213> Secuencia artificial

60 <220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 S304K

<400> 83
gtaagtacgg ttacggtcct tgcatcgcg actccc 36

65 <210> 84
<211> 46

ES 2 694 683 T3

<212> ADN
<213> Secuencia artificial

<220>
5 <223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297V

<400> 84
gctgcatcg cactcccca gaactttacg acgggtagac gcttcg 46

10 <210> 85
<211> 46
<212> ADN
<213> Secuencia artificial

15 <220>
<223> cebador inverso de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297K sintético

<400> 85
20 gctgcatcg cactcccca gcttttacg acgggtagac gcttcg 46

<210> 86
<211> 46
<212> ADN
<213> Secuencia artificial

25 <220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297R

<400> 86
30 gctgcatcg cactcccca gacgtttacg acgggtagac gcttcg 46

<210> 87
<211> 46
<212> ADN
35 <213> Secuencia artificial

<220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296R/I297V

40 <400> 87
gctgcatcg cactcccca gaacgcgacg acgggtagac gcttcg 46

<210> 88
<211> 46
45 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296R/I297K

50 <400> 88
gctgcatcg cactcccca gctgcatcg acgggtagac gcttcg 46

<210> 89
<211> 46
55 <212> ADN
<213> Secuencia artificial

<220>
60 <223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296R/I297R

<400> 89
gctgcatcg cactcccca gacggcgacg acgggtagac gcttcg 46

65 <210> 90
<211> 55

ES 2 694 683 T3

<212> ADN
<213> Secuencia artificial

<220>
5 <223> cebador inverso sintético de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297R/I298K

<400> 90
10 ggtcgtcgcg atgccactc cccttacgtt tacgacgggt agacgcttcg gcctg 55

<210> 91
<211> 55
<212> ADN
<213> Secuencia artificial

15 <220>
<223> cebador inverso de selección del sitio de escisión por OmpT de la variante de RF1 N296K/I297R/I298R sintético

20 <400> 91
ggtcgtcgcg atgccactc ccgacgacgtt tacgacgggt agacgcttcg gcctg 55

<210> 92
<211> 75
25 <212> ADN
<213> Secuencia artificial

<220>
30 <223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 0

<400> 92

gcccgaagcgt ctaccctgcg taacctgctg gggagtggcg atcgcagcga ccgtaaccgt 60
acttacaact tcccg 75

35 <210> 93
<211> 75
<212> ADN
<213> Secuencia artificial

40 <220>
<223> cebador oligo directo de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 1 sintético

<400> 93

cgccgtgggt ctaccctgcg taacctgctg gggagtggcg atcgcagcga ccgtaaccgt 60
acttacaact tcccg 75

45 <210> 94
<211> 75
<212> ADN
50 <213> Secuencia artificial

<220>
<223> cebador oligo directo de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 2 sintético

55 <400> 94

gcacgcctgcg gtaccctgcg taacctgctg gggagtggcg atcgcagcga ccgtaaccgt 60
acttacaact tcccg 75

<210> 95

ES 2 694 683 T3

<211> 75
<212> ADN
<213> Secuencia artificial

5 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 3

<400> 95

10 **gccgcacgcc gtggtcgtcg taacctgctg gggagtggcg atcgcagcga ccgtaaccgt 60**
acttacaact tcccg 75

<210> 96
<211> 75
<212> ADN
15 <213> Secuencia artificial

<220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 4

20 <400> 96

gccgaagcac gccgtggtcg taacctgctg gggagtggcg atcgcagcga ccgtaaccgt 60
acttacaact tcccg 75

<210> 97
25 <211> 75
<212> ADN
<213> Secuencia artificial

<220>
30 <223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 5

<400> 97

gccgaagcgg cagcgcgtgg taacctgctg gggagtggcg atcgcagcga ccgtaaccgt 60
acttacaact tcccg 75

35 <210> 98
<211> 75
<212> ADN
<213> Secuencia artificial

40 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 6

<400> 98

45 **gccgaagcgt ctgcacgccg tggctctgctg gggagtggcg atcgcagcga ccgtaaccgt 60**
acttacaact tcccg 75

<210> 99
50 <211> 75
<212> ADN
<213> Secuencia artificial

<220>
55 <223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 7

<400> 99

ES 2 694 683 T3

gccgaagcgt ctaccgcacg ccgtggtctg gggagtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75

5 <210> 100
<211> 75
<212> ADN
<213> Secuencia artificial

10 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 8
<400> 100

gccgaagcgt ctaccctgctc acgccgtggt gggagtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75

15 <210> 101
<211> 75
<212> ADN
<213> Secuencia artificial

20 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 9
<400> 101

25 gccgaagcgt ctaccctgctg tgcacgccgt ggtagtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75

30 <210> 102
<211> 75
<212> ADN
<213> Secuencia artificial

35 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 10
<400> 102

gccgaagcgt ctaccctgctg taacgcacgc cgtggtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75

40 <210> 103
<211> 75
<212> ADN
<213> Secuencia artificial

45 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 11
<400> 103

50 gccgaagcgt ctaccctgctg taacctggca cgccgtggtg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75

55 <210> 104
<211> 75
<212> ADN
<213> Secuencia artificial

ES 2 694 683 T3

<220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 12

5 <400> 104

```
gccgaagcgt ctaccgctcg taacctgctg gcacgccgtg gtcgcagcga ccgtaaccgt 60
attacaact tcccg 75
```

<210> 105
<211> 75
10 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 13

15 <400> 105

```
gccgaagcgt ctaccgctcg taacctgctg ggggcacgcc gtggtagcga ccgtaaccgt 60
attacaact tcccg 75
```

20 <210> 106
<211> 75
<212> ADN
<213> Secuencia artificial

25 <220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 14

<400> 106

```
gccgaagcgt ctaccgctcg taacctgctg gggagtgcac gccgtggtga ccgtaaccgt 60
attacaact tcccg 75
```

30 <210> 107
<211> 75
<212> ADN
35 <213> Secuencia artificial

<220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 15

40 <400> 107

```
gccgaagcgt ctaccgctcg taacctgctg gggagtggcg cacgccgtga ccgtaaccgt 60
attacaact tcccg 75
```

45 <210> 108
<211> 75
<212> ADN
<213> Secuencia artificial

<220>
50 <223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 16

<400> 108

```
gccgaagcgt ctaccgctcg tcgtctgctg gggagtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75
```

ES 2 694 683 T3

<210> 109
<211> 75
<212> ADN
5 <213> Secuencia artificial

<220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 17

10 <400> 109

```
gcctggctgg cagcgcgtcg cggtcgtggc gggagtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75
```

<210> 110
15 <211> 75
<212> ADN
<213> Secuencia artificial

<220>
20 <223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 18

<400> 110

```
gccgaatggc tggcagcgcg tcgctgctgt ggcagtggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75
```

25

<210> 111
30 <211> 75
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 19

35 <400> 111

```
gccgaagcgt ggctggcagc gcgtcgcggt cgtggcggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75
```

<210> 112
40 <211> 75
<212> ADN
<213> Secuencia artificial

<220>
45 <223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 20

<400> 112

```
tgggggtggcc gttgggctcg caagaaaggt actattggcg atcgcagcga ccgtaaccgt 60
attacaact tcccg 75
```

50 <210> 113
<211> 75
<212> ADN
<213> Secuencia artificial

55 <220>

ES 2 694 683 T3

<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 21

<400> 113

```
5      gcctggggtg gccgttgggc tcgcaagaaa ggtactattg atcgcagcga ccgtaaccgt 60  
      attacaact tcccg                                     75
```

<210> 114

<211> 75

<212> ADN

10 <213> Secuencia artificial

<220>

<223> cebador oligo directo sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 22

15 <400> 114

```
      gccgaatggg gtggccgttg ggctcgcaag aaaggtaact ttcgcagcga ccgtaaccgt 60  
      attacaact tcccg                                     75
```

<210> 115

20 <211> 75

<212> ADN

<213> Secuencia artificial

<220>

25 <223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 0

<400> 115

```
      gctgcatcg ccaactcccca gcaggttacg acgggtagac gttcggcct gttggcgttt 60  
      tgccatttca gcagc                                     75
```

30 <210> 116

<211> 75

<212> ADN

<213> Secuencia artificial

35 <220>

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 1

<400> 116

40 <400> 116

```
      gctgcatcg ccaactcccca gcaggttacg acgggtagaa ccacggcgtg cttggcgttt 60  
      tgccatttca gcagc                                     75
```

<210> 117

45 <211> 75

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 2

50 <400> 117

```
      gctgcatcg ccaactcccca gcaggttacg acgggtacca cggcgtgcct gttggcgttt 60  
      tgccatttca gcagc                                     75
```

55 <210> 118

<211> 75

ES 2 694 683 T3

<212> ADN
<213> Secuencia artificial

5 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 3
<400> 118

```
gctgcatcg ccaactcccca gcaggttacg acgaccacgg cgtgcggcct gttggcgttt 60  
tgccatttca gcagc 75
```

10 <210> 119
<211> 75
<212> ADN
<213> Secuencia artificial

15 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 4
<400> 119

```
gctgcatcg ccaactcccca gcaggttacg accacggcgt gcttcggcct gttggcgttt 60  
tgccatttca gcagc 75
```

25 <210> 120
<211> 75
<212> ADN
<213> Secuencia artificial

30 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 5
<400> 120

```
gctgcatcg ccaactcccca gcaggttacc acggcgtgcc gcttcggcct gttggcgttt 60  
tgccatttca gcagc 75
```

35 <210> 121
<211> 75
<212> ADN
<213> Secuencia artificial

40 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 6
<400> 121

```
gctgcatcg ccaactcccca gcagaccacg gcgtgcagac gcttcggcct gttggcgttt 60  
tgccatttca gcagc 75
```

45 <210> 122
<211> 75
<212> ADN
<213> Secuencia artificial

50 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 7
55 <400> 122

ES 2 694 683 T3

gctgcatcg ccaactcccca gaccacggcg tgcggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

5 <210> 123
<211> 75
<212> ADN
<213> Secuencia artificial

10 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 8
<400> 123

gctgcatcg ccaactcccaac cacggcgtgc acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

15 <210> 124
<211> 75
<212> ADN
<213> Secuencia artificial

20 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 9
<400> 124

gctgcatcg ccactaccac ggcgtgcacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

25 <210> 125
<211> 75
<212> ADN
<213> Secuencia artificial

30 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 10
35 <400> 125

gctgcatcg ccaccacggc gtgcgttacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

40 <210> 126
<211> 75
<212> ADN
<213> Secuencia artificial

45 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 11
<400> 126

gctgcatca ccacggcgtg ccagggttacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

50 <210> 127
<211> 75
<212> ADN

ES 2 694 683 T3

<213> Secuencia artificial

<220>

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 12

5

<400> 127

```
gctgcgacca cggcgtgccca gcaggttacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75
```

10

<210> 128

<211> 75

<212> ADN

<213> Secuencia artificial

15

<220>

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 13

<400> 128

```
gctaccacgg cgtgccccca gcaggttacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75
```

20

<210> 129

<211> 75

<212> ADN

25

<213> Secuencia artificial

<220>

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 14

30

<400> 129

```
accacggcgt gcaactccccca gcaggttacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75
```

35

<210> 130

<211> 75

<212> ADN

<213> Secuencia artificial

<220>

40

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 15

<400> 130

```
acggcgtgcg ccaactccccca gcaggttacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75
```

45

<210> 131

<211> 75

<212> ADN

50

<213> Secuencia artificial

<220>

<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 16

55

<400> 131

ES 2 694 683 T3

gctgcatcg ccaactcccca gcagacgacg acgggtagac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

5 <210> 132
<211> 75
<212> ADN
<213> Secuencia artificial

10 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 17
<400> 132

gctgcatcg ccaactcccg cagcaccgag acgcgctgcc agccaggcct gttggcgttt 60
tgccatttca gcagc 75

15 <210> 133
<211> 75
<212> ADN
<213> Secuencia artificial

20 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 18
<400> 133

gctgcatcg ccaactgccac gaccgagacg cgtgccagc cattcggcct gttggcgttt 60

25 tgccatttca gcagc 75

30 <210> 134
<211> 75
<212> ADN
<213> Secuencia artificial

35 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 19
<400> 134

gctgcatcg cgcaccagc cgcgacgcgc tgccagccac gcttcggcct gttggcgttt 60
tgccatttca gcagc 75

40 <210> 135
<211> 75
<212> ADN
<213> Secuencia artificial

45 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 20
<400> 135

50 gctgcatcg ccaatagtag ctttcttgag agcccaacgg ccaccccact gttggcgttt 60
tgccatttca gcagc 75

<210> 136
<211> 75

ES 2 694 683 T3

<212> ADN
<213> Secuencia artificial

5 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 21
<400> 136

```
gctgcatca atagtacctt tcttgcgagc ccaacggcca cccaggcct gttggcgttt 60  
tgccatttca gcagc 75
```

10 <210> 137
<211> 75
<212> ADN
<213> Secuencia artificial

15 <220>
<223> cebador oligo inverso sintético de la inserción del péptido de escisión por OmpT de la variante de RF1 n.º 22
20 <400> 137

```
gctgcaata gtacctttct tgcgagccca acggccaccc cattcggcct gttggcgttt 60  
tgccatttca gcagc 75
```

25 <210> 138
<211> 66
<212> ADN
<213> Secuencia artificial

30 <220>
<223> cebador de amplificación de RF1 sintético 5His-RF1
<400> 138

```
catatgcatc accatcacca tcacgggtgt ggccttaagc cttctatcgt tgccaaactg 60  
gaagcc 66
```

35 <210> 139
<211> 33
<212> ADN
<213> Secuencia artificial

40 <220>
<223> cebador de amplificación de RF1 sintético 3RF1

45 <400> 139
gtcgacttat tctgctcgg acaacgccgc cag 33 <210> 140

<211> 90
<212> ADN
<213> Secuencia artificial

50 <220>
<223> oligo de PCR sintético de reemplazo alélico mediado por oligonucleótido (OMAR) 1opRF1 KR
<400> 140

```
gggaagttgt aagtacggtt acggtcgctg cgatcccctg aaccaagacg ctttcgacgg 60  
gtagacgctt cggcctgttg gcgttttgcc 90
```

55

ES 2 694 683 T3

<211> 22
<212> ADN
<213> Secuencia artificial

5 <220>
<223> oligo de PCR sintético de ensayo de mutación de amplificación de desapareamientos (MAMA) 5 RF1 op-PCR

10 <400> 146
cgtgacgggg ataacgaacg cc 22

<210> 147
<211> 19
<212> ADN
15 <213> Secuencia artificial

<220>
<223> cebador sintético de secuenciación de la extensión de cebadores y amplificación en círculo rodante T7

20 <400> 147
taatacgact cactatagg 19

<210> 148
<211> 18
25 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador sintético de secuenciación de la extensión de cebadores y amplificación en círculo rodante T7 term

30 <400> 148
gctagttatt gctcagcg 18

<210> 149
<211> 39
35 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00067_V422 para la introducción del codón ámbar en herceptina

40 <400> 149
cgttggcagc aggtaatta gttcagctgc agcgttatg 39

45 <210> 150
<211> 39
<212> ADN
<213> Secuencia artificial

50 <220>
<223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00067_V422 para la introducción del codón ámbar en herceptina

55 <400> 150
cataacgctg cagctgaact aattaccctg ctgccaacg 39

<210> 151
<211> 42
60 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00127_S415 para la introducción del codón ámbar en herceptina

65

ES 2 694 683 T3

<400> 151
 gcaagctgac cgtcgataaa tagcgttggc agcagggtaa tg 42

5 <210> 152
 <211> 42
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00127_S415 para la introducción del codón ámbar en herceptina

15 <400> 152
 cattaccctg ctgccaacgc tatttatcga cggtcagctt gc 42

20 <210> 153
 <211> 36
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00128_Q418 para la introducción del codón ámbar en herceptina

30 <400> 153
 cgataaaagc cgttgtagc aggtaatgt gttcag 36

35 <210> 154
 <211> 36
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00128_Q418 para la introducción del codón ámbar en herceptina

45 <400> 154
 ctgaacacat taccctgcta ccaacggctt ttatcg 36

50 <210> 155
 <211> 35
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00114_P343 para la introducción del codón ámbar en herceptina

60 <400> 155
 gcaaagcgaa aggccaatag cgtgaaccgc aggtc 35

65 <210> 156
 <211> 35
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00114_P343 para la introducción del codón ámbar en herceptina

70 <400> 156
 gacctgcggt tcacgctatt ggccttgc tttgc 35

75 <210> 157
 <211> 40

ES 2 694 683 T3

<212> ADN
 <213> Secuencia artificial

5 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00112_G341 para la introducción del codón ámbar en herceptina

10 <400> 157
 gacgatcagc aaagcgaaat agcaaccgcg tgaaccgcag 40

15 <210> 158
 <211> 40
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00112_G341 para la introducción del codón ámbar en herceptina

25 <400> 158
 ctgcggttca cgcggttgct attcgctt gctgatcgtc 40

30 <210> 159
 <211> 41
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00102_K320 para la introducción del codón ámbar en herceptina

40 <400> 159
 gctgaatggt aaagaatact agtgcaaagt gagcaacaag g 41

45 <210> 160
 <211> 41
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00102_K320 para la introducción del codón ámbar en herceptina

55 <400> 160
 ccttggtgct cacttgcac tagtattctt taccattcag c 41

60 <210> 161
 <211> 39
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00066_F404 para la introducción del codón ámbar en herceptina

70 <400> 161
 ctggacagcg acggtagcta gtttctgtat agcaagctg 39

75 <210> 162
 <211> 39
 <212> ADN
 <213> Secuencia artificial

80 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00066_F404 para la introducción del codón ámbar en herceptina

ES 2 694 683 T3

<400> 162
 cagcttgcta tacagaaact agctaccgtc gctgtccag 39

5 <210> 163
 <211> 32
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00113_Q342 para la introducción del codón ámbar en herceptina

15 <400> 163
 gcaaagcgaa aggctagccg cggaaccgc ag 32

20 <210> 164
 <211> 32
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00113_Q342 para la introducción del codón ámbar en herceptina

30 <400> 164
 ctgcggttca cgcggttagc ctttcgcttt gc 32

35 <210> 165
 <211> 42
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00096_T299 para la introducción del codón ámbar en herceptina

45 <400> 165
 gtgaggaaca atacaatagc tagtatcgcg tagtgagcgt gc 42

50 <210> 166
 <211> 42
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00096_T299 para la introducción del codón ámbar en herceptina

60 <400> 166
 gcacgctcac tacgcatagc tagctattgt attgtcctc ac 42

65 <210> 167
 <211> 30
 <212> ADN
 <213> Secuencia artificial

70 <220>
 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00120_Y373 para la introducción del codón ámbar en herceptina

75 <400> 167
 ggtgaagggc ttttagccga gcgacatcgc 30

80 <210> 168
 <211> 30

ES 2 694 683 T3

<212> ADN
<213> Secuencia artificial

<220>
5 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00120_Y373 para la introducción del codón ámbar en herceptina

<400> 168
10 gcgatgtcgc tcggctaaaa gcccttcacc 30

<210> 169

<211> 38
<212> ADN
15 <213> Secuencia artificial

<220>
20 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00094_N297 para la introducción del codón ámbar en herceptina

<400> 169
20 cgcgtgagga acaatactag agcacgtatc gcgtagtg 38

<210> 170
<211> 38
25 <212> ADN
<213> Secuencia artificial

<220>
30 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00094_N297 para la introducción del codón ámbar en herceptina

<400> 170
35 cactacgcga tacgtgctct agtattgttc ctcacgcg 38

<210> 171
<211> 38
<212> ADN
40 <213> Secuencia artificial

<220>
40 <223> cebador oligo sentido sintético de la variante de mutagénesis dirigida SP-00125_F405 para la introducción del codón ámbar en herceptina

<400> 171
45 gacagcgacg gtagcttcta gctgtatagc aagctgac 38

<210> 172
<211> 38
<212> ADN
50 <213> Secuencia artificial

<220>
55 <223> cebador oligo antisentido sintético de la variante de mutagénesis dirigida SP-00125_F405 para la introducción del codón ámbar en herceptina

<400> 172
55 gtcagcttgc tatacagcta gaagctaccg tcgctgtc 38

REIVINDICACIONES

1. Una proteína de factor de liberación 1 (RF1) que puede ser escindida por una proteína de la membrana externa T1 (OmpT1) y tiene una secuencia de aminoácidos que tiene al menos un 90 % de identidad de secuencia con SEQ ID NO: 1, en la que hay un enlace peptídico escindible por OmpT1 situado dentro de la región del bucle de conmutación, y en la que la región del bucle de conmutación comprende los aminoácidos 287-304 de la SEQ ID NO: 1 modificada para incluir tres aminoácidos básicos adyacentes que están cargados positivamente a pH 7,0.
2. La proteína RF1 de la reivindicación 1, en la que:
- Los tres aminoácidos básicos adyacentes se seleccionan independientemente del grupo que consiste en arginina y lisina;
 - la región del bucle de conmutación se modifica para tener cuatro o cinco aminoácidos básicos adyacentes;
 - La asparagina nativa en la posición 296 se sustituye por uno de los tres aminoácidos básicos adyacentes; o
 - la región del bucle de conmutación comprende una secuencia de aminoácidos seleccionada del grupo de secuencias que consiste en:
 - QQAEASTRRKLLGSGDRS (SEQ ID NO:5),
 - QQAEASTRRLLGSGDRS (SEQ ID NO:6),
 - QQAEASTRRKLLGSGDRS (SEQ ID NO:12),
 - QQAEASTRRKRLGSGDRS (SEQ ID NO:13),
 - QQAEASTRRKVLGSGDRS (SEQ ID NO:14),
 - QQAEASTRRRLGSGDRS (SEQ ID NO:15),
 - QQAEASTRRRLGSGDRS (SEQ ID NO:16),
 - QQAEASTRRRVLGSGDRS (SEQ ID NO:17),
 - QQAEASTRRKKKGSGDRS (SEQ ID NO:18),
 - QQAEASTRRKKRGS GDRS (SEQ ID NO:19),
 - QQAEASTRRKKRGS GDRS (SEQ ID NO:20),
 - QQAEASTRRKRRGS GDRS (SEQ ID NO:21),
 - QQAEASTRRRRRGS GDRS (SEQ ID NO:22),
 - QQAEASTRRRKRGS GDRS (SEQ ID NO:23),
 - QQAEASTRRRRKGS GDRS (SEQ ID NO:24), and
 - QQAEASTRRRKKGS GDRS (SEQ ID NO:25).
3. Una proteína de factor de liberación 1 (RF1) que es escindible por una proteína de la membrana externa T1 (OmpT1), en la que hay un enlace de péptido escindible por OmpT1 situado dentro de la región del bucle de conmutación, y en la que la región del bucle de conmutación comprende una secuencia de aminoácidos seleccionada del grupo de secuencias que consiste en:
- QARRGSTRRNLLGSGDRS (SEQ ID NO:26);
 - QQARRGTRRNLLGSGDRS (SEQ ID NO:27);
 - QQARRGRRNLLGSGDRS (SEQ ID NO:28);
 - QQAEARRGRNLLGSGDRS (SEQ ID NO:29);
 - QQAEARRGNLLGSGDRS (SEQ ID NO:30);
 - QQAEASARRGLLGS GDRS (SEQ ID NO:31);
 - QQAEASTARRGLGS GDRS (SEQ ID NO:32);
 - QQAEASTRARRGGSGDRS (SEQ ID NO:33);
 - QQAEASTRRARRGS GDRS (SEQ ID NO:34);
 - QQAEASTRRNARRGGDRS (SEQ ID NO:35);
 - QQAEASTRRNLARRGDRS (SEQ ID NO:36);
 - QQAEASTRRNLLARRGRS (SEQ ID NO:37);
 - QQAEASTRRNLLGARRGS (SEQ ID NO:38);
 - QQAEASTRRNLLGSARRG (SEQ ID NO:39) QQAEASTRRNLLGSGARR (SEQ ID NO:40);
 - QQAWLAARRGRGGSGDRS (SEQ ID NO:41);
 - QQAEWLAARRGRGS GDRS (SEQ ID NO:42);
 - QQAEAWLAARRGRGS GDRS (SEQ ID NO:43);
 - QQWGGRWARKKGTIGDRS (SEQ ID NO:44);
 - QQAWGGRWARKKGTIDRS (SEQ ID NO:45); y
 - QQAEWGGRWARKKGTIRS (SEQ ID NO:46).
4. Una célula bacteriana que expresa tanto una proteína de la membrana externa T1 (OmpT1) como una proteína RF1 de la reivindicación 1, 2 o 3.
5. La célula bacteriana de la reivindicación 4, en la que: la célula bacteriana es una *E. coli*.
6. Un método para reducir la actividad perjudicial de una proteína RF1 de las reivindicaciones 1, 2 o 3 en un sistema

de síntesis libre de células *in vitro*, comprendiendo el método las etapas de:

- 5 i) cultivar una bacteria positiva en OmpT1 que expresa la proteína RF1 de la reivindicación 1, 2 o 3;
- ii) lisar las bacterias para crear un extracto de síntesis libre de células;
- iii) poner en contacto la proteína RF1 de la reivindicación 1, 2 o 3 con OmpT1 en una cantidad suficiente para reducir la proteína intacta en un 50 %;
- iv) añadir un molde de ácido nucleico al extracto en el que el molde codifica una proteína de interés; y
- v) dejar que el sistema de síntesis libre de células produzca la proteína de interés.

10 7. El método de la reivindicación 6, en el que:

- (a) la bacteria positiva en OmpT1 es *E. coli*;
- (b) el sistema de fosforilación oxidativa de las bacterias permanece activo después de la lisis celular y durante la síntesis de una proteína de interés; o
- 15 (c) el extracto de síntesis libre de células coloca un aminoácido no nativo en un codón ámbar de una proteína de interés.

8. Un ácido nucleico que codifica la proteína RF1 de la reivindicación 1, 2 o 3.

20 9. Un método para reducir la competencia de RF1 en un codón ámbar en un sistema de síntesis libre de células *in vitro*, comprendiendo el método las etapas de:

- i) cultivar una bacteria positiva en OmpT1 que expresa la proteína RF1 de la reivindicación 1, 2 o 3; y
- 25 ii) lisar las bacterias para crear un extracto de síntesis libre de células;
- iii) poner en contacto la proteína RF1 del extracto con OmpT1 en una cantidad suficiente para reducir la proteína RF1 intacta en un 50 %;
- iv) añadir un molde de ácido nucleico al extracto en el que el molde codifica una proteína de interés e incluye un codón de parada; y
- v) dejar que el sistema de síntesis libre de células produzca la proteína de interés.

30 10. El método de la reivindicación 9, en el que:

- (a) la bacteria positiva en OmpT1 es de *E. coli*;
- (b) el sistema de fosforilación oxidativa de las bacterias permanece activo después de la lisis celular y durante la síntesis de una proteína de interés; o
- 35 (c) el extracto de síntesis libre de células coloca un aminoácido no nativo en el codón ámbar de una proteína de interés.

40 11. Un sistema de síntesis libre de células que comprende en una sola mezcla de reacción:

- i) los componentes de un lisado bacteriano suficientes para traducir un molde de ácido nucleico que codifica una proteína;
- ii) un molde de ácido nucleico que codifica una proteína de interés y que tiene al menos un codón ámbar;
- iii) ARNt complementario al codón ámbar; e
- 45 iv) la proteína RF1 de la reivindicación 1, 2 o 3.

12. El sistema de síntesis libre de células de la reivindicación 11, en el que:

- (a) la mezcla de reacción comprende además un aminoácido no natural y el aminoácido sintetasa correspondiente, la sintetasa capaz de cargar el ARNt complementario al codón ámbar con el aminoácido no natural; o
- 50 (b) el sistema genera ATP a través de un sistema activo de fosforilación oxidativa.

13. Un método para expresar proteína en un sistema de síntesis libre de células, que comprende:

- 55 i) combinar un molde de ácido nucleico con un extracto bacteriano que comprende la proteína RF1 de la reivindicación 1, 2 o 3 en condiciones que permiten la síntesis de proteínas; y
- ii) expresar una proteína a partir del molde de ácido nucleico.

60 14. Una célula bacteriana que comprende ácido nucleico que codifica la RF1 funcional de la reivindicación 1, 2 o 3 incorporada en el genoma de la célula bacteriana.

15. Un extracto bacteriano que comprende una proteína RF1 de la reivindicación 1 2 o 3.

Figura 1

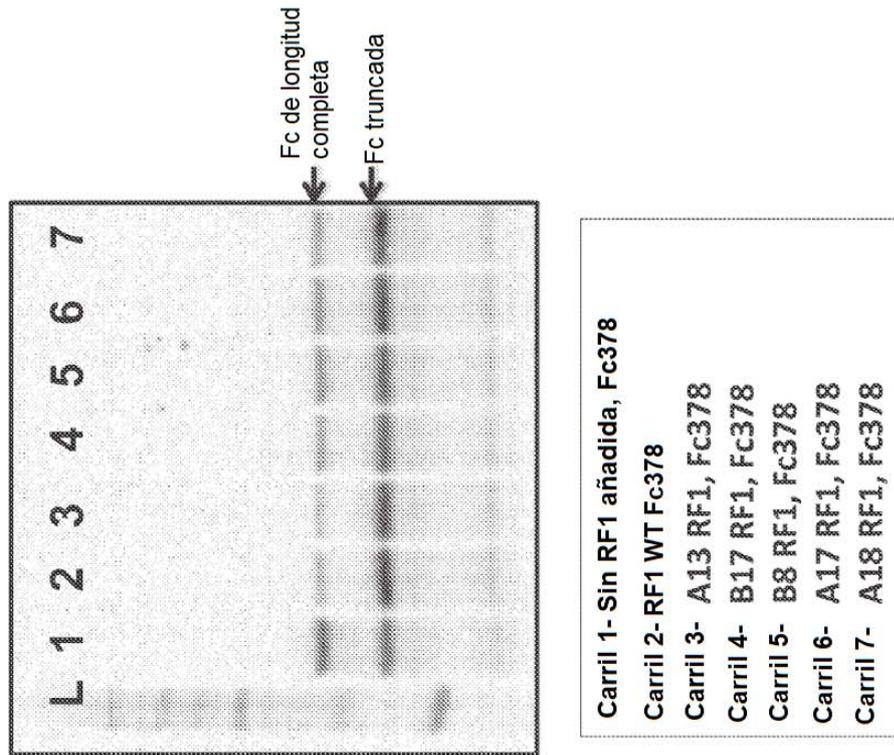
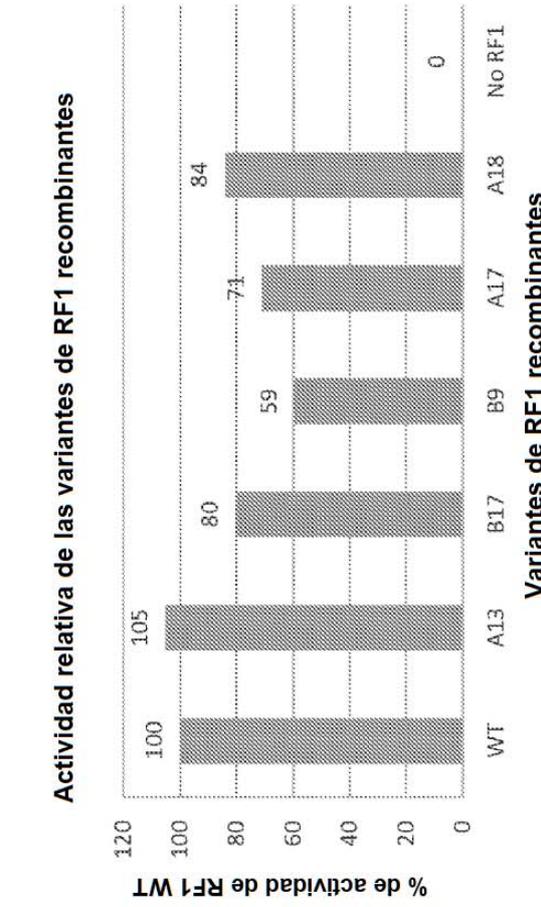


Figura 1B

Figura 1A

Figura 2

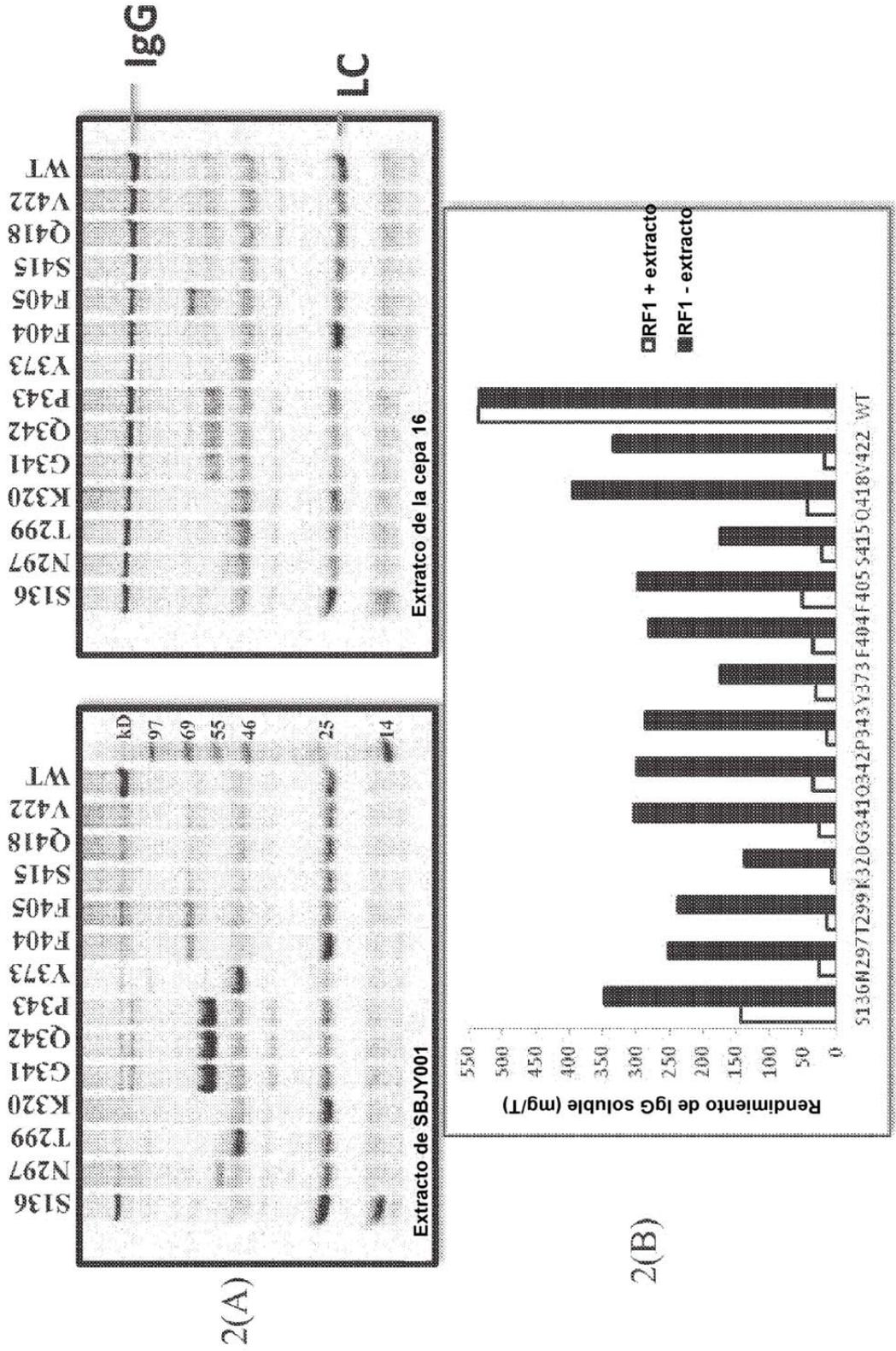


Figura 3

