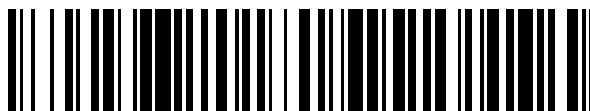


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 539 280**

51 Int. Cl.:

C12N 15/77 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.10.2007 E 07821516 (7)**

97 Fecha y número de publicación de la concesión europea: **18.03.2015 EP 2082045**

54 Título: **Procedimiento de reducción de la expresión génica mediante el uso de codones modificado**

30 Prioridad:

24.10.2006 EP 06122882

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

29.06.2015

73 Titular/es:

**BASF SE (100.0%)
67056 Ludwigshafen, DE**

72 Inventor/es:

**HEROLD, ANDREA;
KLOPPROGGE, CORINNA;
SCHRÖDER, HARTWIG;
ZELDER, OSKAR y
JEONG, WEOL KYU**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 539 280 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

Procedimiento de reducción de la expresión génica mediante el uso de codones modificado

Objeto de la invención

5 La presente invención está dirigida a un procedimiento de reducir la cantidad de al menos un polipéptido en una célula de *Corynebacterium glutamicum*, que comprende la etapa de expresar en una célula de *C. glutamicum* una secuencia de nucleótidos modificada en lugar de una secuencia de nucleótidos no modificada que codifica dicho polipéptido que tiene la misma secuencia de aminoácidos y/o función, en la que dicha secuencia de nucleótidos deriva de la secuencia de nucleótidos no modificada, de forma que al menos un codón de la secuencia de nucleótidos no modificada está sustituida por un codón de uso menos frecuente de acuerdo con el uso de codones de *C. glutamicum*, en el que dicho al menos un codón es el codón de iniciación que está sustituido con un codón de iniciación que se usa con menos frecuencia en *C. glutamicum*, en el que dicho codón de iniciación usado con menor frecuencia es GTG o TTG.

También se describen secuencias de nucleótidos que codifican un polipéptido con un uso de codón que se ha ajustado para usar codones que solo se usan rara vez de acuerdo con el uso de codones del organismo huésped.

15 También se describen el uso de dichas secuencias y procedimientos para producir sustancias químicas finas, tales como aminoácidos, azúcares, lípidos, aceites, hidratos de carbono, vitaminas, cofactores etc.

Antecedentes

20 En un lote de procesos biotecnológicos es necesario modular la expresión génica. Por tanto, para algunas aplicaciones es necesario aumentar la expresión de un determinado producto génico y para aumentar de este modo la cantidad y/o actividad de, por ejemplo, una proteína en la célula huésped en la que se (sobre) expresa el gen de interés. De forma similar, puede ser necesario reducir la cantidad de expresión de un gen endógeno en una célula huésped. Adicionalmente, puede ser deseable ajustar el nivel de expresión de genes endógenos o heterólogos.

25 La producción fermentativa de las denominadas sustancias químicas finas normalmente se lleva a cabo en microorganismos tales como *Corynebacterium glutamicum* (*C. glutamicum*), *Escherichia coli* (*E. coli*), *Saccharomyces cerevisiae* (*S. cerevisiae*), *Schizosaccharomyces pombe* (*S. pombe*), *Pichia pastoris* (*P. pastoris*), *Aspergillus niger*, *Bacillus subtilis*, *Ashbya gossypii* o *Gluconobacter oxydans*.

30 Normalmente se usan y se necesitan sustancias químicas finas que incluyen, por ejemplo, ácidos orgánicos tales como ácido láctico, aminoácidos proteogénicos o no proteogénicos, bases púricas y pirimidínicas, hidratos de carbono, compuestos aromáticos, vitaminas y cofactores, lípidos, ácidos grasos saturados e insaturados, en la industria farmacéutica, agrícola, cosmética, así como alimentaria y de piensos.

35 Como respecto a, por ejemplo, el aminoácido metionina, actualmente la producción anual mundial alcanza aproximadamente 500.000 toneladas. El proceso de producción industrial actual no es mediante fermentación sino un proceso químico de varias etapas. La metionina es el primer aminoácido limitante en ganado de aves alimentadas y, por ello, se ha aplicado principalmente como suplemento de alimentos. Se han publicado varios intentos en la técnica anterior de producir metionina, por ejemplo usando microorganismos tales como *E. coli*.

40 Otros aminoácidos, tales como glutamato, lisina, treonina y treonina, se producen mediante, por ejemplo, procedimientos de fermentación. Para estos fines, determinados microorganismos tales como *C. glutamicum* han mostrado que son particularmente adecuados. La producción de aminoácidos mediante fermentación tiene la ventaja concreta de que solo se producen L-aminoácidos y que se evitan las sustancias químicas problemáticas para el medio ambiente, tales como los disolventes tal como se usan normalmente en la síntesis química.

45 Algunos de los intentos en la técnica anterior para producir sustancias químicas finas tales como aminoácidos, lípidos, vitaminas o hidratos de carbono en microorganismos tales como *E. coli* y *C. glutamicum* han intentado alcanzar este objetivo mediante, por ejemplo, incremento de la expresión de los genes implicados en las rutas de biosíntesis de las respectivas sustancias químicas finas. Si, por ejemplo, se sabe que una determinada etapa de la ruta biosintética de un aminoácido tal como metionina o lisina es limitante de la velocidad, la sobreexpresión de la respectiva enzima puede permitir la obtención de un microorganismo que da más producto de la reacción catalizada y, por tanto, en última instancia conducirá a una producción aumentada del aminoácido respectivo. De un modo similar, si se sabe que una determinada etapa enzimática en la ruta biosintética de, por ejemplo, un aminoácido deseado no es deseable porque canaliza gran cantidad de energía metabólica a la formación de subproductos no deseados, se puede contemplar regular por disminución la expresión de la actividad enzimática respectiva con el fin de favorecer únicamente las reacciones metabólicas que en última instancia conducen a la formación del aminoácido en cuestión.

55 Los intentos para aumentar la producción de, por ejemplo, metionina y lisina mediante regulación por aumento y/o disminución de la expresión de los genes implicados en la ruta biosintética de la producción de metionina o lisina se describen en, por ejemplo, los documentos WO 02/10209, WO 2006008097 o WO2005059093.

Normalmente, la sobreexpresión de un gen determinado en un microorganismo tal como *E. coli* o *C. glutamicum* u otras células huésped tal como *P. pastoris*, *A. niger* o incluso sistemas de cultivo celular de mamíferos se puede conseguir mediante transformación de la respectiva célula con un vector que comprende una secuencia de nucleótidos que codifica la proteína deseada y que además comprende elementos que permiten al vector dirigir la expresión de la secuencia de nucleótidos que codifica, por ejemplo, una enzima determinada. Usando este abordaje, las proteínas extrañas, es decir proteínas que están codificadas por secuencias que no se encuentran de forma natural en la célula huésped que se usa para expresión, además de proteínas endógenas específicas de células huésped, se pueden sobreexpresar. Otros procedimientos típicos incluyen incrementar el número de copias de los respectivos genes en el cromosoma, insertando promotores fuertes para regular la transcripción de la copia en el cromosoma de los respectivos genes y potenciar el inicio de la traducción mediante optimización del sitio de unión al ribosoma (RBS).

Kinnaird y Bums, 1991, Journal of Molecular Biology, 221(3), 733-736 divulgan un raro efecto de codón sobre la velocidad de traducción de un gen de *Neurospora*. El documento describe que, como resultado de dos mutaciones de desplazamiento de marco que se compensan mutuamente, se generaron tres codones sucesivos con la tercera posición A en el gen de (glutamato deshidrogenasa específica de NADP: CDH) de *Neurospora crassa*.

Malumbres et al., 1993, Gene, 134, 15-24 divulga preferencias de codones en *Corynebacteria*. En particular, el documento describe el análisis del uso de codones (CU) de 34 genes de las especies íntimamente relacionadas *Brevibacterium lactofermentum* y *Corynebacterium glutamicum* (BLCG) y la comparación con los de 23 genes de otras especies de *Brevibacterium* y *Corynebacterium*. El contenido de G+C de los genes de BLCG se encontró que variaba de 50 a 62 %.

Romero y Garcia, 1991, FEMS Microbiology Letters, 84, 325-330 divulga el inicio de la traducción en los codones AUC, AUA y AUU en *Escherichia coli*.

Charlet et al., 2005, Molecular Microbiology, 56(5), 1302-1313 divulga una reducción de la expresión de proteínas antigénicas MPB70 y MPB83 en cepas BCG de *Mycobacterium bovis* debido a una mutación en el codón de iniciación en sigK.

Para la regulación por disminución de la expresión de determinados factores en, por ejemplo, microorganismos, se dispone de una multitud de tecnologías tales como abordajes de genes defectivos, tecnología antisentido, tecnología de ARNi etc. Algunas de las tecnologías para la regulación por disminución de genes conducen a una completa pérdida de la función para los respectivos factores debido a la ausencia de, por ejemplo, cualquier enzima producida. Esto puede suponer un problema cuando el objetivo global es únicamente un determinado grado de reducción en la cantidad de la proteína. En estos casos, se puede eliminar la copia de tipo salvaje del respectivo gen y reemplazarlo con una versión mutante que muestre una menor actividad o que lo exprese a partir de un promotor débil. No obstante, dichos abordajes son bastante molestos y otros medios de reducción parcial de la expresión de, por ejemplo, una proteína sin cambiar su secuencia de aminoácidos son altamente deseables, no solo para producir sustancias químicas finas en, por ejemplo, microorganismos, sino también para otros fines.

En vista de esta situación, es un objeto de la presente invención proporcionar procedimientos and secuencias de nucleótidos que codifican polipéptidos que se pueden usar para reducir la expresión de polipéptidos en *C. glutamicum*. Es un objeto adicional proporcionar dichos procedimientos y secuencias para producir sustancias químicas finas.

Estos y otros objetivos como sean evidentes para la descripción de la invención siguiente se resuelven mediante la presente invención como se describe en las reivindicaciones independientes. Las reivindicaciones dependientes se refieren a realizaciones preferidas.

Sumario de la invención

En una realización, la invención se refiere a un procedimiento de reducción de la cantidad de al menos un polipéptido en una célula de *Corynebacterium glutamicum*. Este procedimiento de reducir la cantidad de al menos un polipéptido comprende la etapa de expresar en una célula de *C. glutamicum* una secuencia de nucleótidos modificada en lugar de una secuencia de nucleótidos no modificada que codifica dicho polipéptido que tiene la misma secuencia de aminoácidos y/o función, en la que dicha secuencia de nucleótidos deriva de la secuencia de nucleótidos no modificada de forma que al menos un codón de la secuencia de nucleótidos no modificada está sustituida por un codón de uso menos frecuente de acuerdo con el uso de codones de *C. glutamicum*, en el que dicho al menos un codón es el codón de iniciación que está sustituido con un codón de iniciación que se usa con menos frecuencia en *C. glutamicum*, en el que dicho codón de iniciación usado con menor frecuencia es GTG o TTG. En una realización preferida, el uso de codones de referencia se determinará en base al grupo de proteínas abundantes de la célula de *C. glutamicum*. La secuencia de nucleótidos que se habrá optimizado de este modo para la expresión reducida de un polipéptido, también se puede designar como secuencia de nucleótidos modificada mientras que la secuencia endógena se puede designar como de partida o no modificada.

En una realización preferida, los codones muy frecuentes o extremadamente frecuentes se intercambian por codones raros, preferentemente muy raros y, lo más preferentemente, extremadamente raros. El uso de codones de

referencia se basará en el uso de codones de *C. glutamicum* y, preferentemente, en el uso de codones de proteínas abundantes de *C. glutamicum*.

Este procedimiento de reducción de la cantidad de un polipéptido comprende, por tanto, la etapa de expresar en *C. glutamicum* una secuencia de nucleótidos modificada que codifica dicho polipéptido en lugar de la secuencia de nucleótidos endógena que codifica un polipéptido de sustancialmente la misma secuencia de aminoácidos y/o función. La secuencia de nucleótidos modificada deriva de la secuencia de nucleótidos endógena (no modificada) de forma que los codones de la secuencia de nucleótidos no modificada se intercambia en la secuencia de nucleótidos modificada con codones menos usados y, preferentemente, con los codones menos usados. El uso de codones de referencia se basará en el uso de codones de *C. glutamicum* y, preferentemente, en el uso de codones de proteínas abundantes de *C. glutamicum*. En una realización preferida, de este modo, los codones frecuentes, muy frecuentes o extremadamente frecuentes se intercambian por codones raros, muy raros o extremadamente raros. En una realización particularmente preferida, uno, algunos o todos los codones se reemplazan por los codones menos usados.

En una realización al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de las secuencias de nucleótidos no modificadas pueden sustituirse en la secuencia de nucleótidos modificada por los codones menos usados para el respectivo aminoácido. En una realización incluso más preferida, el número mencionado anteriormente de codones a reemplazar se refiere a codones frecuentes, muy frecuentes o extremadamente frecuentes. En una realización particularmente preferida, estos codones se reemplazan por codones raros, preferentemente muy raros y lo más preferentemente extremadamente raros. En otra realización particularmente preferida, el número codones anterior se reemplaza por los codones menos usados, respectivamente. En todos estos casos, el uso de codones de referencia se basa en el uso de codones del organismo huésped y, preferentemente en el uso de codones de proteínas abundantes de la célula de *C. glutamicum*.

En una realización preferida, se usará una secuencia de nucleótidos modificada que solo usa codones raros, muy raros, extremadamente raros o, preferentemente, los codones menos usados para cada aminoácido reemplazado, como se ha determinado para el grupo de proteína abundantes de la célula de *C. glutamicum*.

Los polipéptidos que se van a expresar mediante este procedimiento de reducción de la cantidad de un polipéptido en una célula de *C. glutamicum* cell son polipéptidos endógenos, con la condición de que la secuencia de nucleótidos modificada no debe ser idéntica a las secuencias de nucleótidos de partida de los respectivos polipéptidos de la célula huésped.

Usando este procedimiento, la expresión del respectivo polipéptido se puede reducir en la célula huésped en al menos un 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o 95 %. la extensión de la reducción de la expresión se determina en comparación con el nivel de expresión del polipéptido endógeno que se expresa a partir de la secuencia de nucleótidos no modificada endógena en condiciones comparables.

Una realización preferida of este procedimiento de reducir la cantidad de un polipéptido en una célula huésped se refiere a procedimientos en los que la cantidad de un polipéptido que se está expresando en *C. glutamicum* se reduce.

Este procedimiento de comprende la etapa de expresar en *C. glutamicum* una secuencia de nucleótidos modificada que codifica dicho polipéptido en lugar de la secuencia de nucleótidos endógena que codifica el respectivo polipéptido de sustancialmente la misma secuencia de aminoácidos y/o función. En una realización al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de las secuencias de nucleótidos no modificadas pueden sustituirse en la secuencia de nucleótidos modificada por los codones menos usados para el respectivo aminoácido. En una realización incluso más preferida, el número mencionado anteriormente de codones a reemplazar se refiere a codones frecuentes, muy frecuentes o extremadamente frecuentes. En una realización particularmente preferida, estos codones se reemplazan por codones raros, preferentemente muy raros y lo más preferentemente extremadamente raros. En otra realización particularmente preferida, el número codones anterior se reemplaza por los codones menos usados, respectivamente. En todos estos casos, el uso de codones de referencia se basará en el uso de codones de *C. glutamicum*. Preferentemente, el uso de codones de referencia se basará en el uso de codones de proteínas abundantes de *C. glutamicum*.

Para reducir la expresión de polipéptidos del género de *C. glutamicum* basada en el uso de codones alterado, en otra realización la secuencia de nucleótidos no modificada que codifica el polipéptido puede estar modificada de modo que al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de las secuencias de nucleótidos no modificadas pueden sustituirse en la secuencia de nucleótidos modificada por los codones menos usados para el respectivo aminoácido. En una realización incluso más preferida, el número mencionado anteriormente de codones a reemplazar se refiere a codones frecuentes, muy frecuentes o extremadamente frecuentes. En una realización particularmente preferida, estos codones se reemplazan por codones raros, preferentemente muy raros y lo más preferentemente extremadamente raros. En otra realización particularmente preferida, el número codones anterior se reemplaza por los codones menos usados, respectivamente. En todos estos casos, el uso de codones de referencia se basará en el uso de codones de *C. glutamicum*. Preferentemente, el uso de codones de referencia se basará en el uso de codones de proteínas abundantes de *C. glutamicum*.

al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de la secuencia de nucleótidos no modificada está reemplazada por codones de uso menos frecuente para el respectivo aminoácido de acuerdo con la Tabla 1 y, preferentemente, de acuerdo con la Tabla 2 . En una
 5 realización incluso más preferida, el número mencionado anteriormente de codones a reemplazar se refiere a codones frecuentes, muy frecuentes y extremadamente frecuentes. En otra realización particularmente preferida, los números de codones anteriores se reemplazan por uno de los dos codones menos frecuentes que codifican el o los aminoácidos respectivos como se expone en la Tabla 1 y, preferentemente, en la Tabla 2.

En otra realización de la invención que se refiere a un procedimiento de disminuir la cantidad del polipéptido en *C. glutamicum*, dicha secuencia de nucleótidos modificada usa los codones ACG para treonina, GAT para ácido aspártico, GAA para ácido glutámico, AGA y AGG para arginina y/o TTG para el codón de iniciación que codifica metionina.
 10

En otra realización más de la invención que se refiere a un procedimiento de disminuir la cantidad del polipéptido en *C. glutamicum*, se puede seleccionar al menos un codón de las secuencias de nucleótidos modificadas del uso de
 15 codones de la Tabla 3 y, preferentemente de la Tabla 4.

Por lo tanto, otra forma de realización particularmente preferida de la invención se refiere a procedimientos de disminuir la cantidad de un polipéptido de en *C. glutamicum* en el que al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez,
 20 preferentemente al menos 1 %, al menos 2 %, al 4 % menos, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de la secuencia de nucleótidos no modificada se sustituyen en la secuencia de nucleótidos modificada resultante por los codones para el respectivo aminoácido de acuerdo con la tabla 3 y preferentemente según la tabla 4. en una realización incluso más preferida el número de codones mencionado anteriormente a reemplazar se refiere a codones frecuentes, muy frecuentes,
 25 extremadamente frecuentes o los más frecuentes.

Otra forma de realización particularmente preferida de la invención se refiere a procedimientos de disminuir la cantidad de un polipéptido en *C. glutamicum* en la que todos los codones de la secuencia de nucleótidos no modificada se sustituyen en la secuencia de nucleótidos modificada resultante por los codones para el aminoácido respectivo de acuerdo con la tabla 3 y preferentemente de acuerdo con la tabla 4.

Las secuencias de nucleótidos modificadas, como pueden usarse en los procedimientos descritos anteriormente de reducir la cantidad de un polipéptido en el organismo huésped *C. glutamicum*.
 30

Dichas secuencias de nucleótidos modificadas derivan de las secuencias de nucleótidos endógenas que codifican un polipéptido de sustancialmente la misma secuencia de aminoácidos y/o función, ajustándose el uso de codones de la secuencia de nucleótidos modificada e tal manera que los codones de las secuencias no modificadas (de tipo salvaje) están reemplazadas por codones de uso menos frecuente. El uso de codones de referencia se basará en el uso del codón del organismo huésped y Preferentemente en el uso de codones de proteínas abundantes del organismo huésped *C. glutamicum*.
 35

La solicitud describe además secuencias de nucleótidos modificadas que se han derivado para un polipéptido específico mediante la sustitución de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos de diez, al menos del 1 %, al menos del 2 %, al menos del 4 %, al menos 6 %, al menos del 8 %, al menos 10 %, al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, al menos 90 % o al menos el 95 % o todos los codones de las secuencias de nucleótidos no modificadas están reemplazados en la secuencia de nucleótidos modificada por codones utilizados con menos frecuencia para el respectivo aminoácido. El número de codones anteriormente mencionados a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes. Los números de codones anteriores pueden sustituirse por los codones de uso menos frecuente. En todos los casos, el uso de codones de referencia se basará en el uso del codón del organismo huésped *C. glutamicum*, y preferentemente en el uso de codones de proteínas abundantes del organismo huésped *C. glutamicum*.
 40
 45

Se describe además que las secuencias de nucleótidos modificadas usarán para cada aminoácido reemplazado el codón menos frecuentemente utilizado de acuerdo con el uso de codones de la célula huésped y de acuerdo con el uso de codones de las proteínas abundantes de la célula huésped.
 50

En el caso de secuencias de nucleótidos modificadas que se van a expresar en *C. glutamicum* para reducir la cantidad del respectivo polipéptido codificado, de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos de diez, preferentemente al menos el 1 %, al menos el 2 %, al menos el 4 %, al menos el 6 %, al menos el 8 %, al menos el 10 %, más preferentemente al menos el 20 %, al menos el 40 %, al menos el 60 %, al menos el 80 %, incluso más preferentemente al menos el 90 % o al menos el 95 % y lo más preferentemente todos los codones de las secuencias de nucleótidos no modificadas pueden reemplazarse en la secuencia de nucleótidos modificada por
 55

- 5 codones utilizados con menos frecuencia para el respectivo aminoácido. En una realización incluso más preferida, el número de codones anteriormente mencionados a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes. En otra realización particularmente preferida, el número de codones anterior se reemplaza por los codones menos usados. En todos estos casos, el uso de codones de referencia se basará en el uso del codón de *C. glutamicum*, y preferentemente en el uso de codones de las proteínas abundantes de *C. glutamicum*.
- 10 En el caso de las secuencias de nucleótidos modificadas que se van a expresar en *C. glutamicum* para reducir la cantidad del respectivo polipéptido modificado, al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al 4 % menos, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de la secuencia de nucleótidos no modificada se sustituyen en la secuencia de nucleótidos modificada resultante por los codones para el respectivo aminoácido de acuerdo con la tabla 1 y preferentemente de acuerdo con la tabla 2. En una realización incluso más preferida el número de codones mencionado anteriormente a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes. En otra realización particularmente preferida, los números de codones anteriores se reemplazan por uno de los dos codones menos frecuentes que codifican el o los aminoácidos respectivos como se expone en la Tabla 1 y, preferentemente, en la Tabla 2.
- 15 En otra realización, la secuencia de nucleótidos MODIFICADA que se utiliza para reducir la expresión de un polipéptido en *C. glutamicum* utiliza los codones ACG para treonina, GAT para ácido aspártico, GAA para ácido glutámico, AGA y AGG para arginina y / o TTG para el codón de iniciación que codifica metionina.
- 20 En aún otra realización de la invención, la secuencia de nucleótidos modificada que se utiliza para reducir la expresión de un polipéptido en *C. glutamicum* puede comprender al menos un codón siendo seleccionado de la Tabla 3 y preferentemente de la Tabla 4.
- 25 La solicitud describe además secuencias de nucleótidos modificadas para reducir la expresión de un polipéptido en *Corynebacterium* y particularmente preferido en *C. glutamicum* en el que al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o menos 95 % y lo más preferentemente todos los codones de la secuencia de nucleótidos no modificada se sustituyen en la secuencia de nucleótidos modificada resultante por los codones para el aminoácido respectivo de acuerdo con la Tabla 3 y, preferentemente, de acuerdo con la Tabla 4. El número de codones mencionado anteriormente a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes y los más frecuentes.
- 30 La solicitud describe además secuencias de nucleótidos modificadas para reducir la expresión de un polipéptido en *C. glutamicum* en la que todos los codones de la secuencia de nucleótidos no modificada se sustituyen en la secuencia de nucleótidos modificada resultante por los codones para el aminoácido respectivo de acuerdo con la Tabla 3 y, preferentemente, de acuerdo con la Tabla 4.
- 35 En la situación en la que se utilizan para reducir la expresión de polipéptidos en *C. glutamicum* estas secuencias de nucleótidos modificadas se seleccionan preferentemente del grupo que comprende secuencias de nucleótidos que codifican genes de las rutas biosintéticas de sustancias químicas finas tales como aminoácidos, azúcares, hidratos de carbono, lípidos, aceites, vitaminas, cofactores, etc., si se sabe que la regulación por disminución de dichos genes favorece la producción de sustancias químicas finas. Particularmente pueden seleccionarse del grupo que comprende las secuencias que codifican genes de las rutas biosintéticas que están implicadas en la síntesis de aminoácidos, tales como glicina, lisina, cisteína, triptófano o metionina.
- 40 Una realización particularmente preferida de la invención se refiere a un procedimiento de disminución de la expresión de la isocitrato deshidrogenasa en un microorganismo adaptando el uso de codones como se describe en el presente documento. El microorganismo puede ser *C. glutamicum*. Estos procedimientos se pueden usar para mejorar la síntesis de aminoácidos y, en particular, de metionina y/o lisina.
- 45 En el presente documento también se describe una célula huésped que comprende una secuencia de nucleótidos modificada que conducirá a una expresión reducida de un polipéptido en la célula huésped como se ha descrito anteriormente. Dicha célula de *C. glutamicum* puede comprender una secuencia de nucleótidos modificada que conducirá a una menor expresión de isocitrato deshidrogenasa.
- 50 Además se describe en el presente documento el uso de las secuencias de nucleótidos modificadas mencionadas anteriormente y / o células huésped para la producción de sustancias químicas finas tales como aminoácidos, azúcares, lípidos, aceites, vitaminas, cofactores, hidratos de carbono, etc. Se pueden utilizar en particular para la producción de aminoácidos tales como glicina, lisina, treonina, cisteína o metionina.
- 55

Figuras

La Figura 1 muestra la secuencia de tipo salvaje de la isocitrato deshidrogenasa (SEC ID N° 1)

La Figura 2 a) muestra la isocitrato deshidrogenasa (ICD) modificada por el uso de codones que lleva una mutación de ATG-GTG (SEC ID N° 2). La mutación está sombreada en gris. La Figura 2 b) muestra el inserto de vector que se usó para reemplazar el gen endógeno (SEC ID N° 3). La mutación está sombreada en gris. Los sitios de restricción están subrayados.

La Figura 3 a) muestra la isocitrato deshidrogenasa (ICD) modificada por el uso de codones CA2 (SEC ID N° 4). La mutación está sombreada en gris. La Figura 3 b) muestra el inserto de vector que se usó para reemplazar el gen endógeno (SEC ID N° 5). La mutación está sombreada en gris. Los sitios de restricción están subrayados.

Descripción detallada de la invención

10 Como se ha expuesto en la parte introductoria, puede ser deseable en algunos casos sobreexpresar, por ejemplo, genes extraños, en una determinada célula huésped, ya que este enfoque permite conferir características novedosas y únicas a una célula huésped, por ejemplo si un gen que codifica para una cierta actividad enzimática que se introduce de forma natural no se encuentra en la célula huésped.

15 Sin embargo, la sobreexpresión de genes extraños que no tienen equivalente en el huésped usando, por ejemplo, vectores de expresión tales como plásmidos se ha encontrado con problemas. Lo mismo se ha observado para la sobreexpresión de genes que tienen un homólogo en el organismo huésped en cuanto a su función, pero que utilizan una secuencia de nucleótidos que normalmente no se encuentra dentro del organismo huésped. El fracaso de las células huésped tales como *E. coli* o *C. glutamicum* para expresar ciertas secuencias extrañas (heterólogas) puede deberse al uso de codones alterado (véase, por ejemplo WO 2004/042059).

20 El código genético es degenerado, lo que significa que un determinado aminoácido puede estar codificado por un número de diferentes tripletes de bases. El uso de codones se refiere a la observación de que un determinado organismo típicamente no va a usar cada posible codón para un determinado aminoácido con la misma frecuencia. En su lugar, un organismo normalmente mostrará ciertas preferencias, es decir, un sesgo de codones específicos, lo que significa que estos codones se encuentran con mayor frecuencia en los genes transcritos de un organismo.

25 Sin embargo, la manipulación del uso de codones se ha utilizado en el pasado para aumentar la expresión de genes extraños. Para reducir la expresión de los genes, más habitualmente una parte o la región de codificación completa se ha eliminado del genoma (el denominado enfoque "knock out"). Del mismo modo, los mutantes con actividad reducida o la modulación de la actividad transcripcional, por ejemplo, por promotores débiles se han aplicado en este contexto.

30 La presente Invención se basa en parte en el descubrimiento sorprendente de que, si se intercambian los codones de una secuencia de nucleótidos que codifica un polipéptido endógeno contra codones que se encuentran con menor frecuencia y se que utilizan en la célula huésped, se observará un nivel de expresión reducido de este polipéptido debido al uso del codón modificado.

35 El uso de codones que se utilizan con menos frecuencia en el organismo huésped para la reducción de la expresión de un polipéptido tiene numerosas ventajas. Dado que la secuencia de nucleótidos MODIFICADA puede integrarse en el locus genómico nativo, la integridad genómica del organismo estará muy conservada. Por otra parte, dependiendo del número de codones que se intercambian y la frecuencia de los codones que se introducen puede ser posible ajustar con precisión la reducción del factor específico.

40 La presente invención se basa además en el sorprendente hallazgo de que la determinación del uso de codones de un organismo puede dar resultados diferentes dependiendo de si el uso de codones se determina para abundantes proteínas de la célula huésped o para el organismo huésped como un todo.

45 Típicamente, las tablas de uso de codones en la técnica anterior para organismos tales como *E. coli* etc. se han basado en un análisis del genoma completo. Los inventores de la presente invención han encontrado para el caso de *C. glutamicum* que el análisis del uso de codones de proteínas abundantes dará resultados bastante diferentes en comparación con las frecuencias de uso de codones como se determina para el organismo completo de *C. glutamicum*. Sin desear quedar ligada a teoría alguna, se supone que la frecuencia de uso de codones específicos de proteínas abundantes en un organismo, tal como *C. glutamicum* refleja ciertos requisitos en cuanto a la composición de codones de una secuencia de nucleótidos altamente expresada.

50 La distribución de uso de codones específicos de genes altamente expresados puede, por ejemplo, reflejar las preferencias de codones que son reconocidos por ARNt que también están frecuente t abundantemente disponibles en las células de los organismos huésped. Del mismo modo estos codones pueden reflejar las estructuras de ARN de transcripción que por su disposición espacial se pueden traducir de manera más eficiente.

55 La identificación de frecuencias de uso de codones no sobre la base de todo el organismo, sino para las proteínas abundantes abre sólo así la intrigante posibilidad de definir los codones que probablemente tendrán una tendencia a dirigir una fuerte expresión de un polipéptido. Si se seleccionan específicamente dichos codones frecuentes y se

reemplazan por un codón menos frecuentemente usado como se ha descrito anteriormente, esto, además, debería ayudar a reducir la expresión del polipéptido respectivo. Por otra parte, si un codón sólo se utiliza muy raramente en proteínas abundantes, esto puede indicar que dicho codón no se traduce eficazmente en el organismo huésped. Esta supuesta incapacidad puede usarse para una realización preferida adicional de la presente invención en la que los codones que sean elegidos deliberadamente para la secuencia de nucleótidos modificada que sólo rara vez se utilizan en las proteínas abundantes de la célula huésped.

Parece razonable suponer que el hallazgo de que las proteínas de expresión alta en una célula huésped tienen un uso de codones diferente en comparación con la situación en la que se determina el uso de codones para todos los genes de un organismo no se limitará a *C. glutamicum*, sino que también se observa para otros organismos tales como *E. coli*, células de levadura, células vegetales, células de insecto o células de cultivo celular de mamífero.

Por tanto, la invención se refiere a reducir la expresión de polipéptidos en una célula de *C. glutamicum* mediante el uso de secuencias de nucleótidos modificada en lugar de las secuencias endógenas en las que los codones de las secuencias endógenas se intercambian para los codones de uso menos frecuente del organismo huésped. Ese uso de codones de referencia se determinará preferentemente. En una realización preferida, los codones se intercambian por los codones menos usados del organismo huésped. En una realización preferida adicional, se sustituirán particularmente los codones de las secuencias endógenas que usan codones que se sabe que se encuentran con mayor frecuencia en las proteínas abundantes del organismo huésped.

La expresión "célula huésped" u "organismo" para los fines de la presente invención se refiere a *C. glutamicum*.

En realizaciones preferidas de la invención, las células huésped pueden seleccionarse del grupo que comprende *Corynebacterium glutamicum* ATCC13032, *Corynebacterium glutamicum* KFCC10065 y *Corynebacterium glutamicum* ATCC21608, así como cepas que se derivan las mismas mediante, por ejemplo, mutagénesis y selección clásica o por mutagénesis dirigida.

Otras cepas particularmente preferidas de *C. glutamicum* se pueden seleccionar del grupo que comprende ATCC13058, ATCC13059, ATCC13060, ATCC21492, ATCC21513, ATCC21526, ATCC21543, ATCC13287, ATCC21851, ATCC21253, ATCC21514, ATCC21516, ATCC21299, ATCC21300, ATCC39684, ATCC21488, ATCC21649, ATCC21650, ATCC19223, ATCC13869, ATCC21157, ATCC21158, ATCC21159, ATCC21355, ATCC31808, ATCC21674, ATCC21562, ATCC21563, ATCC21564, ATCC21565, ATCC21566, ATCC21567, ATCC21568, ATCC21569, ATCC21570, ATCC21571, ATCC21572, ATCC21573, ATCC21579, ATCC19049, ATCC19050, ATCC19051, ATCC19052, ATCC19053, ATCC19054, ATCC19055, ATCC19056, ATCC19057, ATCC19058, ATCC19059, ATCC19060, ATCC19185, ATCC 13286, ATCC21515, ATCC21527, ATCC21544, ATCC21492, NRRL B8183, NRRL W8182, B12NRRLB12416, NRRLB12417, NRRLB12418 y NRRLB11476.

La abreviatura KFCC significa Federación Coreana de Colección de Cultivos, ATCC significa Colección Americana de Cultivos Tipo y la abreviatura DSM significa Deutsche Sammlung von Mikroorganismen. La abreviatura NRRL significa colección de cultivos ARS Northern Regional Research Laboratory, Peoria, IL, EE.UU.

Particularmente preferidos son los microorganismos de *Corynebacterium glutamicum* que ya son capaces de producir sustancias químicas finas tales como L-lisina, L-metionina y / o L-treonina. Por lo tanto se prefieren particularmente la cepa *Corynebacterium glutamicum* ATCC13032 y derivados de esta cepa.

La expresión "reducción de la cantidad de al menos un polipéptido en una célula huésped" se refiere a la situación de que si uno reemplaza una secuencia de nucleótidos endógena que codifica un polipéptido con una secuencia de nucleótidos modificada de acuerdo con la invención que codifica un polipéptido de sustancialmente la misma secuencia de aminoácidos y / o función, una cantidad reducida del polipéptido codificado se expresará dentro de las células huésped. Esto, por supuesto, supone que la comparación se hace para los tipos de células huésped comparables, situaciones basales genética comparables etc.

La expresión "secuencia de nucleótidos" para los fines de la presente invención se refiere a cualquier molécula de ácido nucleico que codifica los polipéptidos tales como péptidos, proteínas, etc. Estas moléculas de ácido nucleico pueden estar hechas de ADN, ARN o análogos de los mismos. Sin embargo se prefieren las moléculas de ácido nucleico hechas de ADN.

Las expresiones "secuencia de nucleótidos no modificada" o "secuencia de nucleótidos de partida" para los fines de la presente invención típicamente se refieren a una secuencia de nucleótidos endógena que codifica un polipéptido cuya expresión se quiere reducir. Estas secuencias de nucleótidos no modificadas o de partida no se han modificado con respecto a su uso de codones, es decir, ningún codón se ha reemplazado por codones de uso menos frecuente.

Las expresiones "secuencia de nucleótidos no modificada" y "secuencia de nucleótidos de partida" no necesariamente tienen que ser equivalentes a una secuencia de nucleótidos endógena. Se puede, por ejemplo prever la situación de que un gen endógeno que codifica el factor X se ha delecionado y reemplazado por una versión mutada del factor X portador de una mutación puntual que conduce a una reducción de la actividad de este factor. Sin embargo, aparte de la mutación puntual, la secuencia de codificación del factor X puede no haberse modificado. A partir de una secuencia de nucleótidos, se puede reducir aún más la expresión del factor X mutado

mediante la sustitución de codones de la secuencia de codificación por codones de uso menos frecuente. Por lo tanto, la secuencia de nucleótidos de partida no será idéntica a una secuencia endógena. Sin embargo, la secuencia de nucleótidos de partida se caracteriza por que no se ha modificado en base a la información del uso de codones para dar lugar a una expresión reducida.

5 Por lo tanto, las expresiones "no modificada" o "secuencia de nucleótidos de partida" se refieren a una secuencia nucleotídica que codifica una proteína endógena o versiones mutadas cuyo uso de codones no se ha modificado para reducir la expresión del polipéptido codificado mediante la sustitución de codones con menos frecuencia de uso según lo determinado por el respectivo organismo huésped y preferentemente para el grupo de proteínas abundantes de los mismos.

10 La expresión "secuencia de nucleótidos modificada" para los fines de la presente invención se refiere a una secuencia que se ha modificado con la intención de reducir la expresión del polipéptido respectivo codificado en una célula huésped mediante el ajuste de la secuencia de la secuencia de nucleótidos no modificada/de partida originalmente diferente. Por lo tanto, los codones de la secuencia no modificada / de partida son reemplazados por codones de uso menos frecuente. El uso de codones de referencia se basa en el uso de codones del organismo huésped.

La persona experta en la técnica es claramente consciente de que la modificación de la secuencia de nucleótidos de partida describe el proceso de optimización con respecto al uso de codones.

20 Si, por ejemplo, la secuencia de codificación de una enzima endógena se ajusta con el fin de reducir la expresión de este factor, los cambios introducidos se pueden identificar fácilmente mediante la comparación de la secuencia modificada y la secuencia de partida que en tal caso es la secuencia de tipo salvaje. Por otra parte, ambas secuencias codificarán en este caso la misma secuencia de aminoácidos.

25 Si, sin embargo, la secuencia de codificación de, por ejemplo, una enzima de tipo salvaje endógena se ajusta como se ha descrito y si la secuencia resultante se modificó adicionalmente simultánea o posteriormente mediante, por ejemplo, delección de aminoácidos, inserción de aminoácidos adicionales o introducción de mutaciones puntuales con el fin de transmitir por ejemplo, nuevas propiedades a la enzima (tales como inhibición por retroalimentación reducida), la secuencia de nucleótidos modificada resultante y la secuencia de nucleótidos modificada puede no codificar secuencias de aminoácidos idénticas. En tal situación, no hay una secuencia de partida en el sentido de que la secuencia de partida y la secuencia modificada codifican la misma secuencia de aminoácidos simplemente porque la mutación que se ha introducido no se había descrito antes. Por otra parte, la función de la codificada puede o no puede verse afectada por las mutaciones introducidas. Si por ejemplo se inserta una mutación puntual en una enzima regulada por retroalimentación, la enzima aún puede catalizar la reacción respectiva, pero puede ser resistente a la inhibición por retroalimentación. Si por ejemplo, se introducen aminoácidos en el extremo N- o C-terminal, esto puede no tener ningún impacto sobre la función. Una persona experta se dará cuenta de que a pesar de las diferencias entre las secuencias de nucleótidos modificada y no modificada, el procedimiento de la invención se ha usado porque la secuencia de partida sin la mutación introducida será conocida en forma de secuencia de tipo salvaje y las diferencias de las secuencias modificada y de partida para dichos codones que no codifican la mutación introducida indicarán claramente que la optimización de uso de codones como se ha descrito anteriormente se ha llevado a cabo. Por lo tanto, la optimización del uso de codones será clara a partir de una comparación de la secuencia de partida y modificada para los codones que codifican los aminoácidos en las mismas posiciones o equivalentes.

40 Esto quiere decir, cuando se afirma en el contexto de la presente invención, que las secuencias de nucleótidos modificada y de partida codifican proteínas de secuencia de aminoácidos y/o función sustancialmente idéntica. La secuencia de nucleótidos modificada y de partida será típicamente al menos 60 %, 65 %, preferentemente al menos 70 %, 75 %, 80 %, 85 % y más preferentemente al menos 90 %, 95 o al menos 98 % idéntica en cuanto a la secuencia de aminoácidos.

45 Se ha indicado anteriormente que en una realización preferida, sólo se reemplazarán codones con codones de uso menos frecuente de acuerdo con el uso de codones del organismo huésped, pero también se seleccionarán los codones a reemplazar de acuerdo con determinados criterios. La sustitución de codones que se utiliza con frecuencia en el organismo huésped puede tener un efecto positivo adicional debido a que el aumento de la frecuencia de uso indica la selección de tales codones por un organismo si la expresión se desea aumentar. La sustitución de estos codones frecuentes por los codones usados con menos frecuencia y preferentemente por los codones menos frecuentes para reducir la expresión, de este modo no sólo establece "frenos" en la secuencia, sino que también elimina los elementos que generalmente impulsan expresión. La sustitución de los codones que se utilizan con frecuencia en proteínas abundantes o ubicuas es particularmente preferida.

55 La expresión "proteínas abundantes" para los fines de la presente invención se refiere al grupo de proteínas altamente expresadas dentro de una célula u organismo huésped.

La persona experta en la técnica está familiarizado con la identificación del grupo de proteínas abundantes en una célula u organismo huésped. Esto puede conseguirse, mediante, por ejemplo, electroforesis en gel 2D. En la

electroforesis en gel 2D, una mezcla de proteínas, tal como un extracto celular en bruto se separó en geles de proteínas por, por ejemplo, tamaño y punto isoeléctrico. Posteriormente estos geles se tiñen y la intensidad de las diversas manchas es una indicación de la cantidad total de proteína presente en la célula.

5 El uso de paquetes de software estándar se seleccionará un grupo de proteínas cuyas intensidades de señal están por encima de un cierto nivel de fondo umbral y definirá este grupo de proteínas como proteínas abundantes. Los paquetes de software típicos usados para este propósito incluyen por ejemplo, Melanie3 (Geneva Bioinformatics SA).

10 La persona experta en la técnica es consciente de que diferentes células huésped tales como microorganismos, células vegetales, células de insectos, etc., varían con respecto al número y tipo de proteínas abundantes en una célula. Incluso dentro del mismo organismo, diferentes cepas pueden mostrar un perfil de expresión un tanto heterogéneo a nivel de proteína. Por lo tanto, se analizarán típicamente diferentes cepas y se considerará que las proteínas que se encuentran en todas las cepas son abundantes.

15 Un buen parámetro de selección para la definición de un grupo de proteínas abundantes para los fines de la presente invención es considerar solo las 10 a 300 y preferentemente las 10 a 30 proteínas más abundantes como se detecta en el procedimiento de electroforesis en gel 2D anteriormente descrito. Preferentemente solo se considerarán las proteínas citosólicas para el grupo de proteínas abundantes, solamente. Por lo tanto, en una realización preferida, la expresión "proteínas abundantes" se refiere al grupo de las aproximadamente 13, 14 o 15 proteínas abundantes en extractos citosólicos de células enteras de organismos huésped identificadas por electroforesis en gel de 2D.

20 Una vez que se han identificado las proteínas abundantes, se pueden utilizar herramientas de software, tales como la función "Cusp" de la caja de herramientas de la versión 2.2.0 de EMBOSS que se puede descargar en [HTTP://EMBOSS.sourceforge.net/download/](http://EMBOSS.sourceforge.net/download/). Otros paquetes de software que se pueden utilizar están disponibles en www.entelechon.com (por ejemplo, Leto IO).

25 Como se ha explicado anteriormente, el principio de la invención es reducir la cantidad de al menos un polipéptido en una célula de *Corynebacterium glutamicum*, que comprende la etapa de expresar en una célula de *C. glutamicum* una secuencia de nucleótidos modificada en lugar de una secuencia de nucleótidos no modificada que codifica dicho polipéptido que tiene la misma secuencia de aminoácidos y/o función, en la que dicha secuencia de nucleótidos deriva de la secuencia de nucleótidos no modificada de forma que al menos un codón de la secuencia de nucleótidos no modificada está sustituida por un codón de uso menos frecuente de acuerdo con el uso de codones de *C. glutamicum*, en el que dicho al menos un codón es el codón de iniciación que está sustituido con un codón de iniciación que se usa con menos frecuencia en *C. glutamicum*, en el que dicho codón de iniciación usado con menor frecuencia es GTG o TTG.

35 A menos que se indique lo contrario, la expresión "codones frecuentes" se refiere a la frecuencia relativa por la cual un cierto codón de todos los codones posibles que codifican un aminoácido específico es utilizado por las proteínas de la célula huésped y, preferentemente, por las abundantes proteínas de la célula huésped.

Un codón se considera que es "frecuente" si se utiliza a una frecuencia relativa de más de 40 %. Es "muy frecuente" si se utiliza con una frecuencia relativa de más de 60 % y una frecuencia relativa de más de 80 % es indicativa de un codón de "extremadamente frecuente". Una vez más las frecuencias relativas se basan en el uso de codones de las proteínas de la célula huésped a menos que se indique lo contrario.

40 A menos que se indique lo contrario, la expresión "codones frecuentes" se refiere a la frecuencia relativa por la cual un cierto codón de todos los codones posibles que codifican un aminoácido específico es utilizado por las proteínas de la célula huésped.

45 Un codón se considera que es "raro" si se utiliza menos de 20 % para el aminoácido específico. Un codón "muy raro" se utilizará a una frecuencia de menos de 10 % y un codón "extremadamente raro" se utilizará a una frecuencia de menos de 5 %.

50 Como los aminoácidos metionina y triptófano están codificados por un codón solamente, la respectiva frecuencia de codones es siempre 100 %. Sin embargo, el aminoácido treonina está codificado por cuatro codones, a saber, ACU, ACC, ACA y ACG. Para todo el organismo de *C. glutamicum*, estos codones se utilizan a una frecuencia relativa de 20,4 %, 52,9 %, 12,5 % y 14,3 % (véase la tabla 1, experimento 1). En vista de las explicaciones anteriores, el codón ACC es, por lo tanto, un codón frecuente y los codones ACA y ACG y son codones raros.

55 El aminoácido alanina está codificada por cuatro codones, a saber, GCU, GCC, GCA, y GCG. Para todo el organismo de *C. glutamicum*, estos codones se utilizan a una frecuencia relativa de 23,7 %, 25,4 %, 29,3 % y 21,6 % (véase la tabla 1, experimento 1). Sin embargo, en el grupo de las proteínas abundantes, estos codones se utilizan a las frecuencias relativas de 46,8 %, 9,9 %, 35,9 % y 7,4 % (véase la tabla 2, el experimento 1). Así, GCU para el grupo de proteínas abundantes resulta ser un codón frecuente. Por otro lado, GCG es particularmente evitado en las proteínas abundantes (véase la tabla 2, experimento 1). Así, se puede considerar reemplazar específicamente el codón GCU en la secuencia de partida por GCG en la secuencia modificada si la reducción de la expresión en *C.*

glutamicum es el objetivo.

En algunas realizaciones, puede ser suficiente y preferido reemplazar codones con codones de uso menos frecuente. En una realización preferida, la secuencia de nucleótidos modificada puede utilizar para cada uno de los codones más frecuentes el codón menos frecuentemente utilizado.

5 Los términos "expresar", "que expresa", "expresado" y "expresión" se refieren a la expresión de un producto génico (por ejemplo, una enzima biosintética de un gen de una ruta) en un organismo huésped. La expresión se puede hacer mediante alteración genética del microorganismo que se utiliza como un organismo de partida. En algunas realizaciones, un microorganismo puede estar alterado genéticamente (por ejemplo, ingeniería genética) para expresar un producto génico en un mayor nivel relativo al producido por el microorganismo de partida o en un microorganismo equiparable que no se ha alterado. La alteración genética incluye, entre otros, alterar o modificar secuencias reguladoras o sitios asociados con la expresión de un gen particular (por ejemplo, mediante la adición de promotores fuertes, promotores inducibles o promotores múltiples o retirando las secuencias reguladoras de tal manera que la expresión sea constitutiva), modificando la localización cromosómica de un gen particular, alterando secuencias de ácido nucleico adyacentes a un gen particular tal como un sitio de unión al ribosoma o un terminador de la transcripción, incrementando el número de copias de un gen particular, modificando proteínas (por ejemplo, proteínas reguladoras, supresores, potenciadores, activadores transcripcionales y similares) implicadas en la transcripción de un gen particular y / o traducción de un producto génico particular, o cualquier otro medio convencional de alterar la regulación de la expresión de un gen particular mediante la rutina en la técnica (incluyendo, entre otros, el uso de moléculas de ácido nucleico antisentido, por ejemplo, para bloquear la expresión de proteínas represoras).

El procedimiento de reducir la cantidad de un polipéptido de acuerdo con la invención comprende la etapa de expresar una secuencia de nucleótidos modificada que codifica un polipéptido en *Corynebacterium glutamicum* en lugar de una secuencia de nucleótidos de partida que codifica para un polipéptido de sustancialmente la misma secuencia de aminoácidos y / o función. La secuencia de nucleótidos modificada deriva de la secuencia de nucleótidos de partida de tal manera que los codones de la secuencia de nucleótidos no modificada se intercambian en la secuencia de nucleótidos modificada con codones usados con menos frecuencia y, preferentemente, los codones menos frecuentemente utilizados, como se ha definido anteriormente en el presente documento. En una realización preferida, los codones frecuentes y preferentemente muy o extremadamente frecuentes se intercambian de este modo por codones raros y, preferentemente, codones muy raros y extremadamente raros. En una realización preferida adicional, los codones más frecuentemente utilizados son reemplazados por los codones menos utilizados. En todos estos casos, la referencia es el uso de codones basado en el uso de codones de *Corynebacterium glutamicum* y preferentemente en el uso de codones de proteínas abundantes de *Corynebacterium glutamicum*.

En una realización preferida se va a utilizar una secuencia de nucleótidos modificada que utiliza solamente codones raros y preferentemente muy raros y extremadamente raros para cada aminoácido sustituido según lo determinado por las proteínas de la célula huésped. En una realización preferida adicional de este último aspecto de la invención, la secuencia de nucleótidos modificada utiliza para cada aminoácido el codón de uso menos frecuente.

En todos estos casos, el codón de iniciación puede ser un "punto caliente" preferido para la introducción de codones de uso menos frecuente, es decir, GTG en lugar de ATG y aún más preferentemente TTG.

40 Los polipéptidos que se van a expresar por este procedimiento de reducir la cantidad de un polipéptido en una célula huésped pueden ser polipéptidos endógenos con la condición de que la secuencia de nucleótidos modificada no deba ser idéntica a las secuencias de nucleótidos de tipo salvaje no modificadas de los respectivos polipéptidos endógenos de *Corynebacterium glutamicum*.

El uso de estos procedimientos y los organismos antes mencionados, la expresión del respectivo polipéptido puede reducirse en la célula huésped en al menos 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o 95 %. La extensión de la reducción de la expresión se determina en comparación con el nivel de expresión de la secuencia de nucleótidos de partida en condiciones comparables.

Se entiende que no es siempre deseable para reducir la expresión tanto como sea posible. En ciertos casos, por ejemplo, una represión de 25 % puede ser suficiente y deseable. La presente invención ofrece la posibilidad de ajustar la represión, por ejemplo, no sustituyendo todos los codones por los codones menos frecuente, pero por ejemplo, introduciendo sólo dos o tres codones raros en posiciones seleccionadas.

Además se describen secuencias de nucleótidos modificadas que pueden utilizarse para reducir la expresión de un polipéptido en una célula huésped.

55 Tales secuencias de nucleótidos modificadas derivan de las secuencias de nucleótidos de partida que codifican un polipéptido de sustancialmente la misma secuencia y / o función con el uso de codones de la secuencia de nucleótidos modificada de aminoácidos que se ajusta de tal manera que los codones de las secuencias no modificadas (de tipo salvaje) son reemplazados por codones de uso menos frecuente. El uso de codones de referencia se basa en el uso de codones del organismo huésped y preferentemente en el uso de codones de las

proteínas abundantes del organismo huésped.

Además se describen tales secuencias de nucleótidos modificadas que se han obtenido para un polipéptido específico mediante la sustitución de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, en menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o menos 95 % y más preferentemente todos los codones de las secuencias de nucleótidos no modificadas se sustituyen en la secuencia de nucleótidos modificada por codones utilizados con menos frecuencia para el respectivo aminoácido. El número de codones anteriormente mencionados a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes que están sustituidos por condiciones raras, muy raras o extremadamente raras. Los números de codones anteriores pueden sustituirse por los codones de uso menos frecuente. En todos los casos, el uso de codones de referencia se basa en el uso de codones del organismo huésped y preferentemente en el uso de codones de las proteínas abundantes del organismo huésped.

Las secuencias de nucleótidos modificados pueden utilizar para cada aminoácido sustituido el codón menos frecuentemente usado como se determina para las proteínas de la célula huésped.

Además se describen vectores que comprenden tales secuencias de nucleótidos modificadas. Estos vectores pueden utilizarse para expresar dichas secuencias si por ejemplo las secuencias endógenas se han eliminado a nivel cromosómico. Los vectores también pueden usarse para reemplazar las secuencias endógenas con las secuencias de nucleótidos modificadas.

También se describen células huésped que comprenden las secuencias de nucleótidos modificadas anteriormente mencionadas y / o los vectores antes mencionados.

Un microorganismo particularmente preferido para llevar a cabo el procedimiento anterior de reducir la expresión es *Corynebacterium glutamicum*.

Se describe adicionalmente el uso de células huésped que se seleccionan a partir de bacterias corineiformes tales como bacterias del género *Corynebacterium*. Particularmente preferidas son las especies *Corynebacterium glutamicum*, *Corynebacterium acetoglutamicum*, *Corynebacterium acetoacidophilum*, *Corynebacterium thermoaminogenes*, *Corynebacterium callunae*, *Corynebacterium ammoniagenes*, *Corynebacterium melassecola* y *Corynebacterium efficiens*, *Brevibacteria* o las especies *Brevibacterium flavum*, *Brevibacterium lactofermentum* y *Brevibacterium divarecatum*.

También se describen células huésped seleccionadas del grupo que comprende *Corynebacterium glutamicum* ATCC13032, *C. acetoglutamicum* ATCC15806, *C. acetoacidophilum* ATCC13870, *Corynebacterium thermoaminogenes* FERMBP-1539, *Corynebacterium melassecola* ATCC 17965, *Corynebacterium efficiens* DSM 44547, *Corynebacterium efficiens* DSM 44549, *Brevibacterium flavum* ATCC14067, *Brevibacterium lactofermentum* ATCC13869, *Brevibacterium divarecatum* ATCC 14020, *Corynebacterium glutamicum* KFCC10065 and *Corynebacterium glutamicum* ATCC21608 así como cepas que derivan de las mismas por ejemplo, mutagénesis y selección clásica o por mutagénesis dirigida.

También se describe el uso de los procedimientos antes mencionados, secuencias de nucleótidos modificadas, vectores y / o células huésped para producir sustancias químicas finas. Se pueden utilizar secuencias de nucleótidos modificadas que se seleccionan del grupo que comprende secuencias de nucleótidos que codifican genes de las rutas de biosíntesis de las sustancias químicas finas para los que se sabe que la represión potencia la producción de sustancias químicas finas.

La expresión "sustancia química fina" es bien conocida por el experto en la técnica y comprende compuestos que se pueden usar en diferentes partes de la industria farmacéutica, la industria agrícola, así como en la industria cosmética, alimentaria y de piensos. Las sustancias químicas finas pueden ser los productos finales o intermedios que se necesitan para otras etapas de síntesis. Las sustancias químicas finas también incluyen monómeros para la síntesis de polímeros.

Las sustancias químicas finas e definen como todas las moléculas que contienen al menos dos átomos de carbono y, además, al menos un heteroátomo que no es un átomo de carbono o hidrógeno. Preferentemente sustancias químicas finas se refieren a moléculas que comprenden al menos dos átomos de carbono y, además, al menos un grupo funcional, tal como hidroxil, amino, tiol, carbonilo, carboxil, metoxil, éter, éster-, amido, fosfoéster-, tioéster- o tioéster-grupo

Sustancias químicas finas comprenden por lo tanto preferentemente ácidos orgánicos tales como ácido láctico, ácido succínico, ácido tartárico, ácido itacónico etc. Las sustancias químicas finas comprenden, además, aminoácidos, bases de púricas y pirimidínicas, nucleótidos, lípidos, ácidos grasos saturados e insaturados, tales como ácido araquidónico, alcoholes, por ejemplo, dioles tales como propanodiol y butanodiol, hidratos de carbono tales como ácido hialurónico y trehalosa, compuestos aromáticos tales como vainillina, vitaminas y cofactores, etc.

Un grupo particularmente preferido de sustancias químicas finas para los fines de la presente invención son productos biosintéticos seleccionados del grupo que comprende ácidos orgánicos, proteínas, aminoácidos, lípidos etc. Otras sustancias químicas finas particularmente preferidas se seleccionan del grupo de compuestos que contienen azufre tales como tionina, cisteína, homocisteína, cistationina, glutatión, biotina, tiamina y/o ácido lipoico.

- 5 El grupo de las sustancias químicas finas más preferido incluye aminoácidos entre los cuales se prefieren particularmente glicina, lisina, metionina, cisteína y treonina.

Un procedimiento preferido de acuerdo con la presente invención se refiere a un procedimiento para reducir la cantidad de al menos un polipéptido en *C. glutamicum* en la que se utilizan los principios anteriores. Por lo tanto, se expresará una secuencia de nucleótidos modificada en lugar de la secuencia de nucleótidos de partida ambos de las cuales codifican sustancialmente la misma secuencia de aminoácidos y / o función en la que al menos un codón de la secuencia de nucleótidos de partida se sustituye en la secuencia de nucleótidos modificada por un codón menos usado. La referencia uso de codones se determina para *C. glutamicum*. Preferentemente, el uso de codones de referencia se determina para las proteínas abundantes de *C. glutamicum*.

15 Por supuesto, las definiciones como se han proporcionado anteriormente para el significado de las expresiones "secuencias de nucleótidos modificadas", "secuencias de nucleótidos no modificadas", "codones raros", "codones frecuentes" etc. se aplican igual para estas realizaciones preferidas de la invención.

Una realización particularmente preferida de la invención se refiere a un procedimiento de disminución de la expresión de la isocitrato deshidrogenasa en un microorganismo adaptando el uso de codones como se describe en el presente documento. El microorganismo puede ser *C. glutamicum*. Estos procedimientos se pueden usar para mejorar la síntesis de aminoácidos y, en particular, de metionina y/o lisina.

20 También se describe una célula huésped, que puede ser *C. glutamicum*, que comprende una secuencia de nucleótidos modificada que dará lugar a una expresión reducida de isocitrato deshidrogenasa

Como se ha mencionado anteriormente, puede ser preferible sustituir los codones que se encuentran con frecuencia en el grupo de proteínas abundantes. Los proteínas abundantes de, por ejemplo., *C. glutamicum* se pueden determinar como se ha descrito anteriormente por electroforesis en gel de proteína 2D. Para este propósito, las cepas de *C. glutamicum* pueden cultivarse en condiciones estándar. Después, los extractos de células se pueden preparar usando los protocolos de lisis comunes. Después de la lisis, los extractos celulares se centrifugan y aproximadamente 25 a 50 mg se analizan mediante 2D-PAGE estándar. Un ejemplo del enfoque se puede encontrar a continuación en el ejemplo 1, así como en la parte materiales y Procedimientos de HansMeier et al. (Proteomics 2006, 6, 233-250)

25 Siguiendo este enfoque se pueden identificar las proteínas abundantes en *C. glutamicum* seleccionando las 10 a 300 proteínas citosólicas más abundantes o identificando de 10 a 30 proteínas citosólicas que se observa que están presente en cantidades elevadas en diversas cepas. Estos resultados se supone que son también representativos para el grupo de proteínas abundantes en otras especies de *Corynebacterium*.

35 Para los fines de la presente invención, la expresión "proteínas abundantes de *C. glutamicum*" puede hacer referencia al grupo que comprende los siguientes factores proteicos (el número de acceso de la secuencia de nucleótidos se muestra entre paréntesis):

Factor de elongación Tu (nº de acceso en Genbank: X77034)

Glicerín-aldehído-3-fosfato-deshidrogenasa (nº de acceso en Genbank: BX927152, 6, nt. 289401-288397)

40 Fructosa bisfosfato aldolasa (nº de acceso en Genbank: BX927156, 6, nt. 134992-133958)

Factor de elongación Ts (nº de acceso en Genbank: BX927154, 6, nt. 14902-14075)

proteína hipotética (nº de acceso en Genbank: BX927155, 6, nt. 213489-214325)

Enolasa (nº de acceso en Genbank: BX927150, nt. 338561-339838)

Peptidil-prolil-Cis-trans isomerasa (nº de acceso en Genbank: BX927148, nt. 34330-34902)

45 Superóxido dismutasa (nº de acceso en Genbank: AB055218)

Fosfoglicerato deshidrogenasa (nº de acceso en Genbank: BX927151, nt. 306039-307631)

proteína SSU Rib S1P (nº de acceso en Genbank: BX927152, 6, nt. 26874-28334)

Triosa-fosfato-isomerasa (nº de acceso en Genbank: BX927155, 6, nt. 286884-286105)

Isopropilmalat-sintasa (nº de acceso en Genbank: X70959)

Butano-2,3-dioldeshidrogenasa (nº de acceso en Genbank: BX927156, nt. 20798-21574)

Fumarato-hidratasa (nº de acceso en Genbank: BX927151, 6, nt. 18803-17394)

5 En base a estas catorce proteínas mencionadas anteriormente se puede crear una tabla de uso de codones usando la función "CUSP" de la herramienta EMBOSS mencionada anteriormente

El grupo descrito anteriormente de catorce proteínas puede usarse particularmente para determinar o definir el grupo de proteínas abundantes en la cepa de *C. glutamicum* ATCC 13032 y/o los derivados (obtenidos mediante, por ejemplo, mutagénesis clásica y selección o ingeniería genética) se usan en el análisis de electroforesis en gel 2D.

10 Utilizando la función CUSP de la caja de herramientas EMBOSS, versión uno, se puede crear una tabla de uso de codones que refleje el uso de codones de las proteínas abundantes de *Corynebacterium* en general y de preferentemente de *C. glutamicum*.

15 Sorprendentemente, el uso de codones de estas proteínas abundantes difiere significativamente del uso de codones tal como se determina para todo el genoma de *C. glutamicum* como queda claro a partir de una comparación de las tablas 1 y 2 (véase el Experimento 1 a continuación). El uso de codones de todo el genoma de *C. glutamicum* puede determinarse, por ejemplo, a partir de cepas que están completamente secuenciadas tal como la cepa ATCC13032 cepa y las Tablas de Uso de Codones uso puede por ejemplo, generarse mediante la función CUSP de la caja de herramientas EMBOSS antes mencionada o están disponibles en, por ejemplo, [HTTP://www.kazusa.or.jp](http://www.kazusa.or.jp). Se obtienen resultados altamente comparables si uno utiliza las más abundantes proteínas citosólicas como se menciona en la Tabla 4 de HansMeier et al. (Véase anteriormente).

20 Una forma de realización preferida del aspecto de la invención que se refiere a procedimientos de reducir la cantidad de polipéptidos en una célula huésped derivada de la especie *Corynebacterium glutamicum* comprende la etapa de expresar en *C. glutamicum* una secuencia de nucleótidos modificada que codifica al menos un polipéptido en lugar de una secuencia nucleotídica no modificada endógena que codifica un polipéptido de sustancialmente la misma secuencia y / o la función de aminoácidos, en el que dicha secuencia de nucleótidos modificada deriva de la
25 secuencia de nucleótidos de partida de tal manera que uno, algunos o preferentemente todos los codones de las secuencias de nucleótidos de partida están sustituidos en la secuencia de nucleótidos modificada por codones utilizados con menos frecuencia para el aminoácido respectivo. En una realización aún más preferida, los codones a reemplazar son codones frecuentes, muy frecuente o extremadamente frecuentes. En una realización particularmente preferida, estos codones se reemplazan por codones raros, muy raros o extremadamente raros. En
30 otra realización particularmente preferida, los números anteriores de codones se sustituyen por los codones menos frecuentemente utilizados, respectivamente. En todos estos casos. el uso de codones de referencia se basa en el uso de codones de *C. glutamicum*. Preferentemente, el uso de codones de referencia se determina para las proteínas abundantes de *C. glutamicum*.

35 Para reducir la expresión de polipéptidos del género de *C. glutamicum* basada en el uso de codones alterado, en otra realización la secuencia de nucleótidos no modificada que codifica el polipéptido puede estar modificada de modo que al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los
40 codones de la secuencia de nucleótidos no modificada está reemplazada por codones de uso menos frecuente para el respectivo aminoácido de acuerdo con la Tabla 1 y, preferentemente, de acuerdo con la Tabla 2 .

En otra realización más de la invención, en el que el procedimiento se usa para reducir la cantidad de polipéptidos en *C. glutamicum*, los codones de la secuencia de nucleótidos modificada se seleccionan de uno de los dos codones de uso menos frecuente que codifican el o los respectivos aminoácido(s) como se expone en la Tabla 1 y, preferentemente en la Tabla 2.

En otra realización más de la invención que se refiere a un procedimiento de disminuir la cantidad de polipéptidos en *C. glutamicum*, al menos uno, algunos o todos los codones de las secuencias de nucleótidos modificadas se pueden seleccionar del uso de codones de la Tabla 3.

50 En otra realización de la invención, el procedimiento puede depender de secuencias de nucleótidos modificadas que usan los codones ACG para treonina, GAT para ácido aspártico, GAA para ácido glutámico, AGA y AGG para arginina y/o TTT para el codón de iniciación.

55 Otra realización particularmente preferida de la invención se refiere a procedimientos de disminuir la cantidad de un polipéptido en *C. glutamicum*, en el que al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más

preferentemente todos los codones de la secuencia de nucleótidos no modificada están reemplazados en la secuencia de nucleótidos modificada resultante por los codones para el respectivo aminoácido de acuerdo con la Tabla 3 y, preferentemente, de acuerdo con la Tabla 4. En una realización incluso más preferida, el número mencionado anteriormente de codones a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes.

Otra realización particularmente preferida de la invención se refiere a procedimientos de disminuir la cantidad de un polipéptido en *C. glutamicum* en el que todos los codones de la secuencia de nucleótidos no modificada están reemplazados en la secuencia de nucleótidos modificada resultante por los codones para el respectivo aminoácido de acuerdo con la Tabla 3 y, preferentemente, de acuerdo con la Tabla 4.

El codón de iniciación puede ser un "punto caliente" preferido para introducir los codones menos usados, es decir GTG en lugar de ATG e incluso más preferentemente TTG.

El organismo huésped se puede seleccionar también de *Corynebacterium glutamicum*. También es preferida la cepa de *C. glutamicum* mencionada anteriormente y es particularmente preferida la cepa de *Corynebacterium glutamicum* ATCC13032 y todas sus derivadas. Las cepas ATCC 13286, ATCC 13287, ATCC 21086, ATCC 21127, ATCC 21128, ATCC 21129, ATCC 21253, ATCC 21299, ATCC 21300, ATCC 21474, ATCC 21475, ATCC 21488, ATCC 21492, ATCC 21513, ATCC 21514, ATCC 21515, ATCC 21516, ATCC 21517, ATCC 21518, ATCC 21528, ATCC 21543, ATCC 21544, ATCC 21649, ATCC 21650, ATCC 21792, ATCC 21793, ATCC 21798, ATCC 21799, ATCC 21800, ATCC 21801, ATCC 700239, ATCC 21529, ATCC 21527, ATCC 31269 y ATCC 21526, que se sabe que producen lisina, también se pueden usar preferentemente. Las otras cepas mencionadas anteriormente también se pueden usar.

La extensión de la reducción de la expresión de proteínas puede ser la misma que se ha mencionado anteriormente.

También se describen secuencias de nucleótidos (modificadas) como se pueden usar en los procedimientos descritos anteriormente de reducción de la cantidad de un polipéptido en *C. glutamicum*.

En el caso de secuencias de nucleótidos modificadas que se van a expresar en *C. glutamicum* para reducir la cantidad del respectivo polipéptido codificado, al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de las secuencias de nucleótidos no modificadas pueden sustituirse en la secuencia de nucleótidos modificada por los codones menos usados para el respectivo aminoácido. El número de codones mencionado anteriormente que se va a reemplazar también puede hacer referencia a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes que se reemplazan por codones raros, muy raros o extremadamente raros. El número anterior de codones también se puede reemplazar por los codones menos usados. En todos estos casos, el uso de codones de referencia se basará en el uso de codones de *C. glutamicum*. Preferentemente, el uso de codones de referencia se basará en el uso de codones de proteínas abundantes de *C. glutamicum*.

En el caso de secuencias de nucleótidos modificadas que se van a expresar en *C. glutamicum* para reducir la cantidad del respectivo polipéptido codificado, al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de la secuencia de nucleótidos no modificada están reemplazados en la secuencia de nucleótidos modificada resultante por los codones menos usados para el respectivo aminoácido de acuerdo con la Tabla 1 y, preferentemente, de acuerdo con la Tabla 2. El número de codones mencionado anteriormente a reemplazar también se puede referir a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes. El número de codones mencionado anteriormente también se puede reemplazar por los codones menos usados de la Tabla 1 y, preferentemente, de la Tabla 2.

En lo que se refiere a las secuencias de nucleótidos para reducir la expresión en *C. glutamicum*, al menos una, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más preferentemente todos los codones de la secuencia de nucleótidos no modificada están reemplazados en la secuencia de nucleótidos modificada resultante por uno de los dos codones menos usados para el respectivo aminoácido de acuerdo con la Tabla 1 y, preferentemente, de acuerdo con la Tabla 2. El número de codones mencionado anteriormente a reemplazar se puede referir a codones frecuentes, muy frecuentes, extremadamente frecuentes o los más frecuentes. La estimación de su un codón es frecuente, muy frecuente o extremadamente frecuente se basarán, p preferentemente, en el uso de codones de proteínas abundantes de *C. glutamicum* como se representa en la Tabla 2.

También se describe que las secuencias de nucleótidos modificadas usan los codones ACG para treonina, GAT para ácido aspártico, GAA para ácido glutámico, AGA y AGG para arginina y/o TTG para el codón de iniciación.

La secuencia de nucleótidos modificada que se usa para reducir la expresión de un polipéptido en *C. glutamicum* puede comprender al menos uno, algunos o todos seleccionados de la Tabla 3 y, preferentemente, de la Tabla 4.

5 También se describen secuencia de nucleótidos modificadas para reducir la expresión de un polipéptido en *C. glutamicum*, en la que al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, preferentemente al menos 1 %, al menos 2 %, al menos 4 %, al menos 6 %, al menos 8 %, al menos 10 %, más preferentemente al menos 20 %, al menos 40 %, al menos 60 %, al menos 80 %, incluso más preferentemente al menos 90 % o al menos 95 % y lo más
10 preferentemente todos los codones de la secuencia de nucleótidos no modificada están reemplazados en la secuencia de nucleótidos modificada resultante por los codones para el respectivo aminoácido de acuerdo con la Tabla 3 y, preferentemente, de acuerdo con la Tabla 4 . El número de codones mencionado anteriormente a reemplazar se refiere a codones frecuentes, muy frecuentes, extremadamente frecuentes y los más frecuentes.

15 También se describen secuencias de nucleótidos modificadas para reducir la expresión de un polipéptido en *C. glutamicum* en el que todos los codones de la secuencia de nucleótidos no modificada están reemplazados en la secuencia de nucleótidos modificada resultante por los codones para el respectivo aminoácido de acuerdo con la Tabla 3 y, preferentemente, de acuerdo con la Tabla 4.

20 En todos estos casos, el codón de iniciación puede ser un "punto caliente" preferido para reemplazar los codones menos usados, es decir GTG en lugar de ATG e incluso más preferentemente TTG. Este aspecto se refiere particularmente a disminuir la cantidad y/o la actividad de la isocitrato deshidrogenasa, por ejemplo, para mejorar la síntesis de metionina.

También se describen vectores que comprenden las secuencias de nucleótidos modificadas mencionadas anteriormente que se pueden usar para disminuir la expresión de proteínas en *Corynebacterium* y, preferentemente, en *C. glutamicum*.

25 Un vector que comprende las secuencias de nucleótidos mencionadas anteriormente se usa para dirigir la expresión de una secuencia de nucleótidos modificada in *C. glutamicum* para aumentar la cantidad de un polipéptido en estas células huésped. Dichos vectores pueden ser, por ejemplo, vectores plasmídicos que se replican de forma autónoma en bacterias corineiformes. Ejemplos son pZ1 (Menkel et al. (1989), Applied and Environmental Microbiology 64: 549-554), pEKEx1 (Eikmanns et al.(1991), Gene 102: 93-98), pHS2-1 (Sonnen et al. (1991), Gene 107: 69-74) Estos
30 vectores se basan en los plásmidos crípticos pHM1519, pBL1 oder pGA1. Otros vectores son pCLiK5MCS (documento WO2005059093), o vectores basados en pCG4 (documento US-A 4,489,160) o pNG2 (Serwold- Davis et al. (1990), FEMS Microbiology Letters 66, 119-124) o pAG1 (documento US-A 5,158,891). Ejemplos de otros vectores adecuados se pueden encontrar en el Handbook of Corynebacterium (editado por Eggeling y Bott, ISBN 0-8493-1821-1, 2005).

35 Al optimizar el uso de codones también se deben considerar otros factores de influencia, como la estructura del ARNm resultante.

También se describen células huésped que comprenden los vectores mencionados anteriormente y las secuencias de nucleótidos modificadas que se pueden usar para disminuir la expresión de proteínas en *C. glutamicum*.

40 Un experto en la técnica está familiarizado con cómo reemplazar, por ejemplo, un gen o secuencia de nucleótidos endógena que codifican un polipéptido determinado con una secuencia de nucleótidos modificada. Esto se puede conseguir mediante, por ejemplo, introducción de una construcción adecuada (plásmido sin origen de replicación, fragmento de ADN lineal sin origen de replicación), mediante electroporación, transformación química, conjugación u otros procedimientos de transformación adecuados. A esto le sigue, por ejemplo, recombinación homóloga usando
45 marcadores seleccionables que garanticen que solo se identifican las células portadoras de la secuencia de nucleótidos modificada en lugar de la secuencia endógena de origen natural. Otros procedimientos incluyen alteración génica del locus cromosómico endógeno y la expresión de las secuencias modificadas de, por ejemplo, plásmidos. Otros procedimientos más incluyen, por ejemplo, transposición. A continuación se proporcionará información adicional sobre los vectores y células huésped que se pueden usar.

50 Los procedimientos descritos anteriormente de reducción de la cantidad de un polipéptido en una célula huésped y, particularmente, para reducir la cantidad de un polipéptido en *C. glutamicum* puede realizarse con la secuencia de nucleótidos modificada seleccionada del grupo que comprende secuencias de nucleótidos que codifican genes de las rutas biosintéticas de sustancias químicas finas si se sabe que la represión de la expresión de estos genes potencia la producción de las sustancias químicas finas. Esto es especialmente útil cuando el polipéptido es responsable de la síntesis de subproductos/productos secundarios indeseados. Reducir la cantidad de un
55 polipéptido tal como de una enzima que forma parte de una ruta biosintética de las sustancias químicas finas mencionadas anteriormente puede permitir el incremento de la síntesis de las sustancias químicas finas mediante, por ejemplo, apagado de la producción de subproductos y mediante canalización del flujo metabólico hacia una dirección preferida.

De forma similar, las secuencias de nucleótidos modificadas, los vectores y las células huésped mencionadas anteriormente se pueden usar para producir sustancias químicas finas in *C. glutamicum*.

5 También se describe un vector que se puede usar para reemplazar una secuencia endógena por la secuencia modificada. Por supuesto, se puede también usar un vector que sea adecuado para la expresión de dichas secuencias de nucleótidos en la respectiva célula huésped. Esto puede realizarse, por ejemplo, su la secuencia modificada se expresa desde un vector y la secuencia endógena se ha silenciado antes mediante, por ejemplo, alteración génica.

10 En este contexto se prefieren las cepas de *Corynebacterium glutamicum* que ya son capaces de producir sustancias químicas finas tales como L-lisina, L-metionina y/o L-treonina. Dicha cepa es, por ejemplo, *Corynebacterium glutamicum* ATCC13032 y derivadas de la misma. Las cepas ATCC 13286, ATCC 13287, ATCC 21086, ATCC 21127, ATCC 21128, ATCC 21129, ATCC 21253, ATCC 21299, ATCC 21300, ATCC 21474, ATCC 21475, ATCC 21488, ATCC 21492, ATCC 21513, ATCC 21514, ATCC 21515, ATCC 21516, ATCC 21517, ATCC 21518, ATCC 21528, ATCC 21543, ATCC 21544, ATCC 21649, ATCC 21650, ATCC 21792, ATCC 21793, ATCC 21798, ATCC 21799, ATCC 21800, ATCC 21801, ATCC 700239, ATCC 21529, ATCC 21527, ATCC 31269 y ATCC 21526, que se sabe que producen lisina, también se pueden usar preferentemente. Las otras cepas mencionadas anteriormente también se pueden usar.

20 Estas secuencias de nucleótidos modificadas de nuevo pueden seleccionarse, preferentemente, del grupo que comprende secuencias de nucleótidos que codifican genes de las rutas biosintéticas de sustancias químicas finas. Las definiciones y preferencias en cuanto al significado y deseo de las sustancias químicas finas proporcionadas anteriormente se aplican igualmente. Por tanto, el grupo de los productos químicos finos más preferidos incluye aminoácidos entre los cuales se prefieren particularmente glicina, lisina, metionina, cisteína y treonina.

25 En una realización preferida de los procedimientos de reducir la cantidad de un polipéptido en *C. glutamicum*, las secuencias de nucleótidos modificadas puede, por tanto, seleccionarse del grupo que comprende secuencias que codifican treonina-deshidratasa, homoserina-O-acetiltransferasa, O-acetilhomoserina-sulfhidrilasa, fosfoenolpiruvato-carboxilasa (pepCK), piruvato-oxidasa (poxB), homoserina-cinasa, homoserina-deshidrogenasa, exportador de treonina, salida de treonina, asparaginasa, aspartato-decarboxilasa y treonina-sintasa, citrato sintasa, aconitasa, isocitrato-deshidrogenasa, alfa-cetoglutarato deshidrogenasa, succinil-CoA-sintasa, succinato-deshidrogenasa, fumarasa, malato-quinona oxidoreductasa, malato deshidrogenasa, piruvato cinasa, enzima málica, alr (alanina racemasa), atr43 (ABC trans- porter), ccpA1 (proteína de control de catabolitos a), ccpA2 (proteína de control de catabolitos), chrA (regulador de respuesta de dos componentes), chrS (histidina cinasa), citB (regulador de la transcripción), citE (citrato liasa E), citE (citrato liasa E), clpC (proteasa), csp1, ctaF (4. subunidad de la citocromo aa3 oxidasa), dctA (proteína transportadora de C4-dicarboxilato), dctQ sodit (proteína transportadora de C4-dicarboxilato), dead (ADN/ARN helicasa), def (péptido deformilasa), dep33 (proteína B de resistencia a múltiples fármacos), dep34 (proteína de flujo), fda (fructosa bisfosfato idolasa), gorA (glutatio reductasa), gpi/pgi (glucosa-6-P-isomerasa), hisC2 (his- tidinol fosfato aminotransferasa), hom (homoserín deshidrogenasa), lipA (lipoato sintasa), lipB (lipoproteína-ligasa B), lrp (regulador de la respuesta de leucina), luxR (regulador de la transcripción), luxS (proteína cinasa de la transducción sensorial), lysR1 (regulador de la transcripción), lysR2 (regulador de la transcripción), lysR3 (regulador de la transcripción), mdhA (malato deshidrogenasa), menE (ácido O-succinilbenzoico CoA ligasa), mikE17 (facto de transcripción), mqq (malato-quinon oxidoreductasa), mtrA mtrB (proteína sensora cpxA, componente regulador sensorial), nadA (quinolinato sintasa A), nadC (nicotinato nucleótido pirofosfatasa, otsA (trehalosa-6-P-sintasa), otsB, treY, treZ (trehalosa fosfatasa, maltooligosil-trehalosa sintasa maltooligosil-trehalosa trehalohidrolasa), pepC (aminopeptidasa I), , pfKA pfbK (1 und 6-fosfofructoquinasa), poxB (piruvato oxidasa), pstC2 (proteína transportadora de fosfato unida a la membrana), rplK (PS1-proteína), sucC sucD (succinil CoA sintetasa), sugA (proteína transportadora de azúcar), tmk (timidilato cinasa), zwa2, metK metZ, glyA (serínhidroximetiltransferasa), sdhC sdhA sdhB (succinato DH), smtB (regulador de la transcripción), cgl1 (regulador de la transcripción), hspR (regulador de la transcripción), cgl2 (regulador de la transcripción), cebR (regulador de la transcripción), cgl3 (regulador de la transcripción), gatR (regulador de la transcripción), glcR (regulador de la transcripción), tcmR (regulador de la transcripción), smtB2 (regulador de la transcripción), dtxR (regulador de la transcripción), degA (regulador de la transcripción), galR (regulador de la transcripción), tipA2 (regulador de la transcripción), mall (regulador de la transcripción), cgl4 (regulador de la transcripción), arsR (regulador de la transcripción), merR (regulador de la transcripción), hrcA (regulador de la transcripción), glpR2 (regulador de la transcripción), lexA (regulador de la transcripción), ccpA3 (regulador de la transcripción), degA2 (regulador de la transcripción) en el caso de lisina. En una realización preferida adicional de los procedimientos de reducción de la cantidad de un polipéptido en *C. glutamicum*, las secuencias de nucleótidos modificadas Pueden seleccionarse de este modo del grupo que comprende secuencias que codifican homoserina-cinasa, treonina-deshidratasa, treonina-sintasa, meso-diaminopimelat D-deshidrogenasa, fosfoenolpiruvato-carboxilasa, piruvato-oxidasa, dihidrodipicolinato-sintasa, dihidrodipicolinato-reductasa y diaminopicolinato- descarboxilasa en el caso de metionina. La isocitrato deshidrogenasa puede preferirse particularmente cuando se hace referencia a la producción de metionina.

60 En una realización preferida de los procedimientos de reducir la cantidad de un polipéptido en *C. glutamicum*, las secuencias de nucleótidos modificadas pueden seleccionarse de este modo del grupo que comprende secuencias que codifican treonina-deshidratasa, homoserina O-acetiltransferasa, serina-hidroximetiltransferasa, O-

acetilhomoserina-sulfidrilasa, meso- diaminopimelato D-deshidrogenasa, fosfoenolpiruvato-carboxicinasas, piruvato-oxidasa, dihidrodipicolinato-sintetasa, dihidrodipicolinato-reductasa, asparaginasa, aspartato-descarboxilasa, exportador de lisina, acetolactato-sintasa, cetol-aid-reductoisomerasa, aminotransferasa de cadena ramificada, metionina sintasa dependiente de coenzima B12, metionina sintasa independiente de coenzima B12, dihidroxi ácido deshidratasa y diaminopicolinato-descarboxilasa en el caso de treonina.

En general, el experto en la técnica está familiarizado con el diseño de construcciones tales como vectores para dirigir la expresión de un polipéptido en microorganismos tales como *E. coli* y *C. glutamicum*. El experto en la técnica también conoce bien las condiciones de cultivo de microorganismos tales como *C. glutamicum* y *E. coli* así como procedimientos para recoger y purificar sustancias químicas finas tales como aminoácidos y, particularmente, lisina, metionina y treonina, a partir de los microorganismos mencionados anteriormente. Algunos de estos aspectos se expondrán con detalle más adelante.

El experto en la técnica también está familiarizado con técnicas que permiten modificar la secuencia de nucleótidos no modificada original en una secuencia de nucleótidos modificada que codifica los polinucleótidos de aminoácidos idénticos pero con diferente uso de codones. Esto se puede conseguir mediante, por ejemplo, mutagénesis basada en la reacción en cadena de la polimerasa, mediante procedimientos de clonación conocidas habitualmente, mediante síntesis química etc. Algunos de estos procedimientos se exponen en los ejemplos. A continuación se describirá y expondrá con detalle cómo se pueden realizar manipulaciones genéticas en microorganismos tales como *E. coli* y particularmente *Corynebacterium glutamicum*.

Vectores y células huésped

En el presente documento se describen vectores, preferentemente vectores de expresión, que contienen una secuencia de nucleótidos modificada como se ha mencionado anteriormente. Como se usa en el presente documento, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido.

Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN circular bicatenario en el que se pueden ligar segmentos adicionales de ADN. Otro tipo de vector es un vector viral, en el que se pueden ligar en el genoma viral segmentos adicionales de ADN.

Ciertos vectores son capaces de una replicación autónoma en una célula huésped en la que se introducen (p. ej., vectores bacterianos que tienen un origen de replicación bacteriano y vectores episomales de mamífero). Otros vectores (por ejemplo, vectores no episomales de mamífero) se integran en el genoma de una célula huésped tras la introducción en la célula huésped y, de este modo, se replican junto con el genoma del huésped. Además, ciertos vectores son capaces de dirigir la expresión de los genes a los que están unidos operativamente.

Dichos vectores se denominan en el presente documento "vectores de expresión".

En general, los vectores de expresión de utilidad en técnicas de ADN recombinante suelen estar en forma de plásmidos. En la presente especificación, "plásmido" y "vector" se pueden usar de forma intercambiable, ya que el plásmido es la forma de vector usada con más frecuencia. También se incluyen otras formas de vectores de expresión, tales como vectores virales (por ejemplo, retrovirus defectuosos en la replicación, adenovirus y virus adenoasociados), que sirven funciones equivalentes.

Los vectores de expresión recombinantes descritos en el presente documento pueden comprender un ácido nucleico modificado como se ha mencionado anteriormente en una forma adecuada para la expresión del ácido nucleico respectivo en una célula huésped, que significa que los vectores de expresión incluyen una o más secuencias reguladoras seleccionadas en base a las células huésped que se van a usar para la expresión, que está operativamente unida a la secuencia de ácido nucleico que se va a expresar.

Dentro de un vector de expresión recombinante, con "operablemente unido", se quiere decir que la secuencia de nucleótidos de interés está relacionada con la(s) secuencia(s) reguladora(s) de un modo que permita la expresión de la secuencia de nucleótidos (p. ej., en un sistema de transcripción/traducción in vitro o en una célula huésped cuando el vector se introduce en la célula huésped). Con la expresión "secuencia reguladora" se pretende incluir promotores, sitios de unión al represor, sitios de unión al activador, potenciadores y otros elementos de control de la expresión (por ejemplo, terminadores, señales de poliadenilación u otros elementos de estructura secundaria del ARNm). Dichas secuencias reguladoras se describen en, por ejemplo, Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Las secuencias reguladoras incluyen aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células huésped y aquellas que dirigen la expresión directa de la secuencia de nucleótidos solo en determinadas células huésped. Secuencias reguladoras preferidas son, por ejemplo, promotores tales como *cos-*, *tac-*, *trp-*, *tet-*, *trp-*, *tet-*, *lpp-*, *lac-*, *lpp-lac-*, *lacIq-*, *T7-*, *T5-*, *T3-*, *gal-*, *trc-*, *ara-*, *SP6-*, *amy*, *SP02*, *e-Pp-* *ore PL*, *SOD*, *EFTu*, *EFT*, *GroEL*, *MetZ* (últimos 5 de *C. glutamicum*), que se usan preferentemente en bacterias. Secuencias reguladoras adicionales son, por ejemplo, promotores de levaduras y hongos, tales como *ADC 1*, *MFa*, *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH*, promotores de plantas tales como *CaMV/35S*, *SSU*, *OCS*, *lib4*, *usp*, *STLS1*, *B33*, promotores *nos* o *ubiquitina* o *faseolina*. También es posible usar promotores artificiales. Un experto en la técnica apreciará que el diseño del

vector de expresión puede depender de factores tales como la elección de la células huésped que se va a transformar, el nivel de expresión de proteínas deseado etc. Los vectores de expresión descritos en el presente documento pueden introducirse en las células huésped para producir de este modo proteínas o péptidos, incluyendo proteínas de fusión o péptidos, codificados por las secuencias de nucleótidos modificadas mencionadas anteriormente.

Los vectores de expresión recombinantes descritos en el presente documento se pueden diseñar para la expresión de las secuencias de nucleótidos modificadas como se ha mencionado anteriormente en células procariotas o eucariotas. Por ejemplo, las secuencias de nucleótidos modificadas como se ha mencionado anteriormente se pueden expresar en células bacterianas tales como *C. glutamicum* y *E. coli*, células de insecto (usando vectores de expresión de baculovirus), levaduras y otras células fúngicas (véase Romanos, M. A. et al. (1992), *Yeast* 8: 423-488; van den Hondel, C. A. M.J. J. et al.(1991) en: *More Gene Manipulations in Fungi*, J. W.Bennet & L. L. Lasure, eds., p. 396-428: Academic Press: San Diego; y van den Hondel, C. A. M. J. J. & Punt, P. J.(1991) en: *Applied Molecular Genetics of Fungi*, Peberdy, J. F. et al., eds., p. 1-28, Cambridge University Press: Cambridge), algas y células vegetales multicelulares (véase Schmidt, R. y Willmitzer, L. (1988) *Plant Cell Rep.*: 583-586). Las células huésped adecuadas se tratan adicionalmente en Goeddel, *Gene Expression Technology : Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternativamente, el vector de expresión recombinante puede también transcribirse y traducirse in vitro, por ejemplo usando las secuencias reguladoras del promotor de T7 y la polimerasa de T7.

La expresión de proteínas en procariotas se lleva a cabo con mayor frecuencia con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o que no son de fusión.

Los vectores de fusión añaden una serie de aminoácidos a una proteína codificada en ellos, normalmente al extremo amino de la proteína recombinante, pero también en el extremo C o fusionados dentro de regiones adecuadas en las proteínas. Dichos vectores de fusión normalmente sirven para cuatro fines: 1) aumentar la expresión de proteína recombinante; 2) aumentar la solubilidad de la proteína recombinante y 3) ayudar en la purificación de la proteína recombinante actuando como ligando en la purificación de afinidad 4) para proporcionar una "marca" para la detección posterior de la proteína. A menudo, en los vectores de expresión de fusión se introduce un sitio de escisión proteolítica en la unión del resto de fusión y la proteína recombinante para permitir la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión. Dichas enzimas, y sus secuencias de reconocimiento afines, incluyen el Factor Xa, la trombina y la enteroquinasa.

Los vectores de expresión de fusión incluyen pGEX (Pharmacia Biotech Inc; Smith, D. B. y Johnson, K. S. (1988) *Gene* 67: 31-40], pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Pharmacia, Piscataway, NJ) que contienen glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente.

Ejemplos de vectores de expresión de no fusión de *E. coli* incluyen pTrc (Amann et al., (1988) *Gene* 69: 301-315), pLG338, pACYC184, pBR322,pUC18, pUC19, pKC30, pRep4,pHSI, pHS2, pPLc236, pMBL24, pLG200, pUR290,pIN-III113-BI, egtII, pBdCl, y pET Ild (Studier et al., *Gene Expression Technology : Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89; and Pouwels et al., eds. (1985) *Cloning Vectors*. Elsevier: New York ISBN 0 444 904018). La expresión del gen diana desde el vector pTrc depende de la transcripción con la ARN polimerasa del huésped desde un promotor de fusión híbrido trp-lac. La expresión del gen diana del vector pET Ild depende de la transcripción a partir de un promotor de fusión gnlO-lac de T7 mediada por una ARN polimerasa viral coexpresada (T7gnl). Esta polimerasa viral la suministra las cepas huésped BL21 (DE3) o HMS174 (DE3) a partir de un profago X residente que aloja un gen de T7gnl bajo el control transcripcional del promotor de lacUV 5. Para la transformación de otras variedades de bacterias se pueden seleccionar vectores adecuados. Por ejemplo, se sabe que los plásmidos pIJ101, pIJ364, pIJ702 y pIJ361 son útiles en la transformación de *Streptomyces*, mientras que los plásmidos pUB110, pC194 o pBD214 son adecuados para la transformación de especies de *Bacillus*. Varios plásmidos de uso en la transferencia de información genética en *Corynebacterium* incluyen pHM1519, pBL1, pSA77 o pAJ667 (Pouwels et al., eds. (1985) *Cloning Vectors*. Elsevier: New York ISBN 0 444 904018).

Ejemplos de vectores lanzadera adecuados de *C. glutamicum* y *E. coli* son, por ejemplo, pClik5aMCS (documento WO2005059093) o se pueden encontrar en Eikmanns et al (*Gene*. (1991) 102, 93-8).

Ejemplos de vectores adecuados para manipular corinebacterias se pueden encontrar en el *Handbook of Corynebacterium* (editado por Eggeling y Bott, ISBN 0-8493-1821-1, 2005). Se puede encontrar una lista de vectores lanzadera de *E. coli* - *C. glutamicum* (tabla 23.1), una lista de vectores de expresión lanzadera de *E. coli* - *C. glutamicum* (tabla 23.2), una lista de vectores que se pueden usar para la integración de ADN en el cromosoma de *C. glutamicum* (tabla 23.3), una lista de vectores de expresión para la integración en el cromosoma de *C. glutamicum* (tabla 23.4.) así como una lista de vectores para la integración específica de sitio en el cromosoma de *C. glutamicum* (tabla 23.6).

El vector de expresión de proteínas descrito en el presente documento puede ser un vector de expresión en levaduras. Ejemplos de vectores para la expresión en la levadura *S. cerevisiae* incluyen pYepSecl (Baldari, et al., (1987) *Embo J.* 6: 229-234)" 2i, pAG-1, Yep6, Yep13, pEMBLYe23, pMFa (Kurjan y Herskowitz, (1982) *Cell* 30: 933-943), pJRY88 (Schultz et al., (1987) *Gene* 54: 113-123), y pYES2 (Invitrogen Corporation, San Diego, CA). Vectores

y procedimientos para la construcción de vectores adecuados para usar en otros hongos, tales como hongos filamentosos, incluyen los detallados en : van den Hondel, C. A. M. J. J. & Punt. P. J. (1991) en: Applied Molecular Genetics of Fungi, J. F. Peberdy, et al., eds., p. 1-28, Cambridge University Press: Cambridge, y Pouwels et al., eds. (1985) Cloning Vectors. Elsevier: New York (ISBN 0 444 904018).

- 5 Para los fines de la presente invención, se entiende que un vínculo operativo es la disposición secuencial del promotor, la secuencia de codificación, el terminador y, opcionalmente, otros elementos reguladores de un modo tal que cada uno de los elementos reguladores puede cumplir su función, de acuerdo con su determinación, cuando expresan la secuencia de codificación.

10 Las secuencias de nucleótidos modificadas como se ha mencionado anteriormente se pueden expresar en células vegetales unicelulares (tal como algas) o en células vegetales de plantas superiores (por ejemplo, los espermatofitos, tales como plantas de cosecha). Ejemplos de vectores de expresión en plantas incluyen los detallados en: Becker, D., Kemper, E., Schell, J. y Masterson, R. (1992) Plant Mol. Biol. 20: 1195-1197; y Bevan, M. W. (1984) Nucl. Acid. Res. 12: 8711-8721, e incluyen pLGV23, pGHlac+, pBIN19, pAK2004, y pDH51 (Pouwels et al., eds. (1985) Cloning Vectors. Elsevier: New York ISBN 0 444 904018).

- 15 Para otros sistemas de expresión para células procariotas y eucariotas véanse los capítulos 16 y 17 de Sambrook, J. et al. Molecular Cloning: A Laboratory Manual. 3ª, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2003.

20 El vector de expresión recombinante en mamíferos puede ser capaz de dirigir la expresión del ácido nucleico, preferentemente en un tipo de célula concreto, por ejemplo en células vegetales (por ejemplo, los elementos reguladores específicos de tejido se usan para expresar el ácido nucleico). En la técnica se conocen elementos reguladores específicos de tejido.

25 También se describen organismos o células huésped en las que se ha introducido un vector de expresión recombinante como se ha mencionado en el presente documento. Las expresiones "célula huésped" y "célula huésped recombinante" se usan de forma intercambiable en el presente documento. Debe entenderse que dichos términos no sólo se refieren a la célula sujeto concreta, sino también a la progenie, o potencial progenie, de dicha célula. Dado que se pueden producir ciertas modificaciones en las generaciones posteriores debido a mutaciones o influencias ambientales, dicha progenie puede, de hecho, no ser idéntica a la célula parental, pero todavía está incluida en el alcance del término tal como se usa en la presente memoria descriptiva.

30 El ADN del vector se puede introducir en células procariotas y eucariotas a través de técnicas convencionales de transformación o transfección. Como se usa en el presente documento, con los términos "transformación" y "transfección", "conjugación" y "transducción" se pretende hacer referencia a diversas técnicas reconocidas en la materia para introducir ácido nucleico extraño (por ejemplo, ADN o ARN lineal (por ejemplo, un vector lializado o una construcción génica sola sin un vector) o ácido nucleico en forma de un vector (por ejemplo, un plásmido, un fago, fagemido, transposón u otro ADN) en una célula huésped, incluyendo coprecipitación en fosfato cálcico o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección, competencia natural, transferencia mediada por conjugación química o electroporación. Procedimientos adecuados para transformar o transfectar células huésped se pueden encontrar en Sambrook y col., Molecular Cloning: A Laboratory Manual. 3ª, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2003) y otros manuales de laboratorio.

40 Para identificar y seleccionar estos integrantes, un gen que codifica un marcador seleccionable (por ejemplo, resistencia a antibióticos) se introduce generalmente en las células huésped junto con el gen de interés. Marcadores seleccionables preferidos incluyen los que confieren resistencia a fármacos, tales como G418, higromicina, kanamicina, tetraciclina, ampicilina y metotrexato. El ácido nucleico que codifica un marcador seleccionable se puede introducir en una célula huésped en el mismo vector que el que codifica las secuencias de nucleótidos modificadas mencionadas anteriormente o se puede introducir en un vector distinto. Las células que se han transfectado establemente con el ácido nucleico introducido pueden identificarse mediante selección farmacológica (por ejemplo, las células que han incorporado el gen del marcador seleccionable sobrevivirán mientras que las otras células mueren).

50 Cuando se usan plásmidos sin un origen de replicación y dos genes marcadores diferentes (por ejemplo, pClik int sacB), también es posible generar cepas sin marcadores que tienen parte del inserto insertado en el genoma. Esto se consigue mediante dos acontecimientos consecutivos de recombinación homóloga (véase también Becker et al., APPLIED AND ENVIRONMENTAL MICROBIOLOGY, 71 (12), p. 8587-8596). La secuencia del plásmido pClik int sacB se puede encontrar en el documento WO2005059093; SEC ID N° 24; el plásmido se denomina pCIS en este documento).

55 Se pueden producir microorganismos recombinantes que contienen sistemas seleccionados que permitan la expresión regulada del gen introducido. Por ejemplo, la inclusión de una de las secuencias de nucleótidos optimizadas mencionadas anteriormente en un vector colocándolas bajo el control del operón lac permite la expresión del gen únicamente en presencia de IPTG. Dichos sistemas reguladores son bien conocidos en la técnica.

En una realización, el procedimiento comprende cultivar los organismos como se describe en el presente documento (en los que un vector de expresión recombinante o en cuyo genoma se ha introducido un gen que comprende las secuencias de nucleótidos modificadas como se ha mencionado anteriormente) en un medio adecuado para la producción de sustancias químicas fina. En otra realización, el procedimiento comprende además aislar la sustancia química fina del medio o de la célula huésped.

Crecimiento de *Escherichia coli* y *Corynebacterium glutamicum*- Medios y condiciones de cultivo

El experto en la técnica está familiarizado con el cultivo de microorganismos comunes tales como *C. glutamicum* y *E. coli*. Por tanto, una enseñanza general se proporcionará más adelante en cuanto al cultivo de *C. glutamicum*. Se puede recuperar información correspondiente de libros de texto convencionales para el cultivo de *E. coli*.

Las cepas de *E. coli* crecen de forma rutinaria en caldo MB y LB, respectivamente (Follettie et al. (1993) J. Bacteriol. 175, 4096-4103). El medio mínimo para *E. coli* es M9 y MCGC modificado (Yoshihama et al. (1985) J. Bacteriol. 162, 591-507), respectivamente. Se puede añadir glucosa a una concentración final del 1 %. Se pueden añadir antibióticos en las cantidades siguientes (microgramos por mililitro): ampicilina, 50; kanamicina, 25; ácido nalidíxico, 25. Se pueden añadir aminoácidos, vitaminas, y otros suplementos en las cantidades siguientes: metionina, 9,3 mM; arginina, 9,3 mM; histidina, 9,3 mM; tiamina, 0,05 mM. Las células de *E. coli* se cultivan de forma rutinaria a 37 °C, respectivamente.

Las Corinebacterias modificadas genéticamente normalmente se cultivan en medios de crecimiento sintéticos o naturales. Una serie de medios de crecimiento diferentes para *Corynebacteria* se conocen bien y están disponibles fácilmente (Lieb et al. (1989) Appl. Microbiol. Biotechnol., 32: 205-210; von der Osten et al. (1998) Biotechnology Letters, 11: 11-16; Patente DE 4,120,867; Liebl(1992) "The Genus *Corynebacterium*, in: The Prokaryotes, Volume II, Balows, A. et al., eds. Springer-Verlag). También se pueden encontrar instrucciones en el Handbook of *Corynebacterium* (editado por Eggeling y Bott, ISBN 0-8493-1821-1, 2005).

Estos medios consisten en una o más fuentes de carbono, fuentes de nitrógeno, sales inorgánicas, vitaminas y oligoelementos. Las fuentes de carbono preferidas son azúcares, tales como mono, di o polisacáridos. Por ejemplo, glucosa, fructosa, manosa, galactosa, ribosa, sorbosa, ribosa, lactosa, maltosa, sacarosa, glicerol, rafinosa, almidón o celulosa sirven como fuentes de carbono muy buenas.

También es posible suministrar azúcar a los medios a través de compuestos complejos tales como melazas u otros subproductos del refinado del azúcar. También puede ser ventajoso suministrar mezclas de diferentes fuentes de carbono. Otras posibles fuentes de carbono son alcoholes y ácidos orgánicos, tales como metanol, etanol, ácido acético o ácido láctico. Las fuentes de nitrógeno normalmente son compuestos de nitrógeno orgánicos o inorgánicos o materiales que contienen estos compuestos. Fuentes de nitrógeno de ejemplo incluyen gas amonio o sales de amonio, tales como NH₄Cl o (NH₄)₂SO₄, NH₄OH, nitratos, urea, aminoácidos o fuentes de nitrógeno complejas como licor de maceración del maíz, harina de soja, proteína de soja, extracto de levaduras, extracto de carne y otras.

La sobreproducción de metionina es posible usando diferentes fuentes de azufre. Se pueden usar sulfatos, tiosulfatos, sulfitos y también fuentes de azufre reducido como H₂S y sulfuros y derivados. Asimismo se pueden usar fuentes de azufre orgánico como metilo, mercaptano, tioglicolatos, tiocianatos y tiourea, aminoácidos que contienen azufre como cisteína y otros compuestos que contienen azufre, para conseguir una producción eficiente de metionina. También puede ser posible el formiato como suplemento, como también otras fuentes de C1 tales como metanol o formaldehído.

Compuestos de sales inorgánicas que se pueden incluir en los medios incluyen las sales cloruro, fósforo o sulfato de calcio, magnesio, sodio, cobalto, molibdeno, potasio, manganeso, cinc, cobre y hierro. Se pueden añadir compuestos quelantes al medio para mantener los iones metálicos en solución. Compuestos quelantes particularmente útiles incluyen dihidroxifenoles, como catecol or protocatechuato, o ácidos orgánicos, tales como ácido cítrico. Es típico que los medios contengan también otros factores de crecimiento, tales como vitaminas o promotores del crecimiento, ejemplos de los cuales incluyen biotina, riboflavina, tiamina, ácido fólico, ácido nicotínico, pantotenato y piridoxina. Los factores de crecimiento y las sales con frecuencia proceden de componentes de medios complejos, tales como extracto de levadura, melazas, licor de maceración del maíz y otros. La composición exacta de los compuestos de los medios depende fuertemente del experimento inmediato y se decide de forma individual para cada caso específico. La información sobre la optimización de medios está disponible en el libro de texto "Applied Microbiol. Physiology, A Practical Approach (Eds. P. M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19 963577 3). También es posible seleccionar medios de crecimiento de proveedores comerciales, como estándar 1 (Merck) o BHI (infusión de cerebro-corazón, DIFCO) u otros.

Todos los componentes del medio se deben esterilizar con calor (20 minutos a 1,5 bares y 121 °C) o mediante filtración en esterilidad. Los componentes se pueden esterilizar juntos o, en caso necesario, por separado.

Todos los componentes de los medios pueden estar presentes al comienzo del crecimiento u, opcionalmente, se pueden añadir de forma continua o discontinua. Las condiciones de cultivo se definen por separado para cada experimento.

- La temperatura debería estar en un intervalo entre 15 °C y 45 °C. La temperatura se puede mantener constante y se puede alterar durante el experimento. El pH del medio puede estar en el intervalo de 5 a 8,5, preferentemente aproximadamente 7,0, y puede mantenerse mediante la adición de tampones al medio. Un tampón de ejemplo para este fin es un tampón de fosfato potásico. Los tampones sintéticos tales como MOPS, HEPES, ACES y otros se pueden usar de forma alternativa o simultáneamente. También es posible mantener un pH del cultivo constante a través de la adición de NaOH o NH₄OH durante el crecimiento. Si se usan componentes del medio complejo, tal como extracto de levadura, la necesidad de tampones adicionales se puede reducir debido al hecho de que muchos compuestos complejos tienen altas capacidades tampón. Si se usa un fermentador para cultivar los microorganismos, el pH también se puede controlar usando amoníaco gaseoso.
- El tiempo de incubación normalmente está en el intervalo de horas a varios días. Este tiempo se selecciona con el fin de permitir que se acumule la cantidad máxima de producto en el caldo. Los experimentos de crecimiento divulgados se pueden llevar a cabo en diversos recipientes, tales como placas de microtitulación, tubos de vidrio, matraces de vidrio o fermentadores de vidrio o metal de diferentes tamaños. Para seleccionar un gran número de clones, los microorganismos se deben cultivar en placas de microtitulación, tubos de vidrio o matraces de agitación, con o sin deflectores. Preferentemente se usan matraces de agitación de 100 ml, cargados con 10 % (en volumen) del medio de crecimiento requerido. Los matraces deberán agitarse en un agitador rotatorio (amplitud de 25 mm) usando un intervalo de velocidad de 100-300' rpm. Las pérdidas por evaporación se pueden disminuir mediante el mantenimiento de una atmósfera húmeda; como alternativa, deberá realizarse una corrección matemática para las pérdidas por evaporación.
- Si se analizan clones modificados genéricamente, también se deberán analizar un clon control no modificado o un clon control que contiene el plásmido básico sin ningún inserto. El medio se inocula a una DO600 de 0,5-1,5 usando células cultivadas en placas de agar, tales como placas de CM (10 g/l de glucosa, 2,5g/l de NaCl, 2g/l de urea, 10g/l de polipeptona, 5 g/l de extracto de levadura, 5g/l de extracto de carne, 22 g/l de NaCl, 2 g/l de urea, 10 g/l de polipeptona, 5 g/l de extracto de levadura, 5 g/l de extracto de carne, 22 g/l de agar, pH 6,8 con NaOH 2M) que se había incubado a 30 °C.

La inoculación del medio se consigue mediante introducción de una suspensión salina de células de *C. glutamicum* de placas CM o adición de un líquido de precultivo de esta bacteria.

Cuantificación de aminoácidos e intermedios de metionina

- El análisis se realiza mediante HPLC (Agilent 1100, Agilent, Waldbronn, Alemania) con un cartucho de guarda y una columna Synergi de 4 mm (MAX-RP 80 Å, 150 * 4,6 mm) (Phenomenex, Aschaffenburg, Alemania). Antes de la inyección, los analitos se obtienen usando o-ftaldialdehído (OPA) y mercaptoetanol como agente reductor (2-MCE). Adicionalmente, los grupos sulfhidrilo se bloquean con ácido yodoacético. La separación se lleva a cabo a un caudal de 1 ml/min usando NaH₂PO₄ 40 mM (eluyente A, pH=7.8, ajustado con NaOH) como fase polar y una mezcla de agua y metanol (100 / 1) como fase no polar (eluyente B). Se aplica el gradiente siguiente: Inicio 0 % de B; 39 min 39 % de B; 70 min 64 % de B; 100 % de B durante 3,5 min; 2 min 0 % de B para el equilibrado. La derivación a temperatura ambiente se es automática como se describe a continuación. Inicialmente se mezclan 0,5 µl de 0,5 % de 2- MCE en bicina (0,5M, pH 8,5) con 0,5 µl de extracto celular. Después se añaden 1,5 µl de 50 mg/ml de ácido yodoacético en bicina (0,5M, pH 8,5), seguido de la adición de 2,5 µl de tampón de bicina (0,5M, pH 8,5). La derivatización se realiza añadiendo 0,5 µl de 10 mg/ml de reactivo OPA disuelto en 1/45/54 v/v/v de 2-MCE/MeOH/bicina (0,5M, pH 8,5). Por último, la mezcla se diluye con 32 µl de H₂O. Entre cada una de las etapas de pipeteado anteriores existe un tiempo de espera de 1 minuto. Después, se inyecta un volumen total de 37,5 µl en la columna. Obsérvese que los resultados analíticos se pueden mejorar significativamente, si la aguja automuestreadora se limpia periódicamente durante (por ejemplo en el tiempo de espera) y después de la preparación de la muestra. La detección se realiza con un detector de fluorescencia (340 nm de excitación, emisión a 450 nm, Agilent, Waldbronn, Alemania). Para la cuantificación se usa ácido α-amino butírico (ABA) como patrón interno.

Definición del protocolo recombinante

- A continuación se describirá cómo se puede construir una cepa de *C. glutamicum* con una eficiencia mayor de producción de metionina implementando los hallazgos de las predicciones anteriores. Antes de describir la construcción de la cepa se proporciona una definición de un acontecimiento/protocolo de recombinación que se usará en lo siguiente.

- "Campbell in" como se usa en el presente documento, hace referencia a una células huésped original en la que una molécula completa de ADN bicatenario circular (por ejemplo, un plásmico basado en pCLIK int sacB) se ha integrado en un cromosoma mediante un único acontecimiento de recombinación homóloga (un acontecimiento de cruzamiento) y que tiene como resultado una inserción eficaz de una versión linealizada de dicha molécula de ADN circular en una primera secuencia de ADN del cromosoma que es homóloga de una primera secuencia de ADN de dicha molécula de ADN circular. "Campbelizado in" hace referencia a la secuencia de ADN linealizada que se ha integrado en el cromosoma de un transformante "Campbell in". Un "Campbell in" contiene una duplicación de la primera secuencia de ADN homóloga, cada copia de la cual incluye y rodea una copia del punto de cruzamiento de

recombinación homóloga. El nombre procede del Profesor Alan Campbell, quien propuso por primera vez este tipo de recombinación.

"Campbell out" como se usa en el presente documento, hace referencia a una célula descendiente de un transformante "Campbell in" en la que se ha producido un segundo acontecimiento de recombinación homóloga (salida) entre una segunda secuencia de ADN contenida en el ADN insertado linealizado del ADN "Campbelizado in" y una segunda secuencia de ADN de origen cromosómico, que es homóloga de la segunda secuencia de ADN de dicho inserto linealizado, de modo que el segundo acontecimiento de recombinación en la delección (descarga) de una porción de la secuencia de ADN integrada pero, cabe destacar, que también tiene como resultado una porción (esta puede ser tan pequeña como una sola base) del ADN *Campbelizado in* que queda en el cromosoma, deforma que en comparación con la célula huésped original, la célula "Campbell out" contiene uno o más cambios intencionados en el cromosoma (por ejemplo, una sustitución de una sola base, sustituciones de varias bases, inserción de un gen o secuencia de ADN heteróloga, inserción de una copia adicional o copias de un gen homólogo o un gen homólogo modificado, o inserción de una secuencia de ADN que comprende más de uno de estos ejemplos mencionados anteriormente indicados en lo que antecede).

Una célula o cepa "Campbell out" normalmente se obtiene, pero no necesariamente, mediante contraselección en una porción (la porción que se desea descargar) de la secuencia de ADN "Campbelizado in", por ejemplo el gen *sacB* de *Bacillus subtilis*, que es letal cuando se expresa en una célula que se cultiva en presencia de aproximadamente 5 % a 10 % de sacarosa. Con o sin contraselección se puede obtener o identificar una célula "Campbell out" mediante detección selectiva de la célula deseada usando cualquier fenotipo detectable, tal como, entre otros, la morfología de las colonias, el color de las colonias, la presencia o ausencia de resistencia a antibióticos de una secuencia de ADN dada mediante reacción en cadena de la polimerasa, presencia o ausencia de auxotrofia, hibridación de ácido nucleico de las colonias, detección selectiva de anticuerpos etc. Las expresiones "Campbell in" y "Campbell out" también se pueden usar como verbos en varios tiempos para hacer referencia al procedimiento o proceso descrito anteriormente.

Se entiende que los acontecimientos de recombinación homóloga que conduce a un "Campbell in" o "Campbell out" se pueden producir en una serie de bases de ADN dentro de la secuencia de ADN homólogo y dado que las secuencias homólogas serán idénticas entre sí para al menos parte de este intervalo, normalmente no es posible especificar exactamente cuando se ha producido el acontecimiento de cruzamiento. En otras palabras, no es posible especifica con precisión qué secuencia procedía originalmente del ADN insertado y cuál era originalmente del ADN cromosómico. Además, la primera secuencia de ADN homóloga y la segunda secuencia de ADN homóloga normalmente están separadas por una región de no homología parcial y es esta región de no homología que permanece depositada en un cromosoma de la célula "Campbell out" cell.

A efectos prácticos, en *C. glutamicum*, la primera y la secuencia de ADN homóloga típicas tienen al menos una longitud de aproximadamente 200 pares de bases, y pueden ser de hasta miles de pares de bases de longitud, aunque el procedimiento se puede hacer funcionar con secuencias más cortas o más largas. Por ejemplo, una longitud para la primera y la segunda secuencias homólogas puede variar desde aproximadamente 500 a 2.000 bases, y la obtención de un "Campbell out" a partir de un "Campbell in" se facilita disponiendo la primera y la segunda secuencias homólogas para que tengan aproximadamente la misma longitud, preferentemente con una diferencia de menos de 200 pares de bases y, lo más preferentemente, teniendo la más corta de las dos al menos un 70 % de la longitud de la más larga en pares de bases. El "procedimiento de *Campbell in* y -Out" se describe en el documento WO2007012078

A continuación se ilustrará la invención mediante varios ejemplos.

Ejemplos

A continuación se mostrará cómo se identificó el uso de codones de proteínas abundantes en *C. glutamicum*. Adicionalmente se presentan ejemplos que muestran que el uso de secuencias de nucleótidos modificadas que se han optimizado con respecto al uso de codones de proteínas abundantes o al organismo de *C. glutamicum* se pueden usar para aumentar o reducir la cantidad de una proteína en *C. glutamicum*. Esto se muestra para genes extraños así como para genes endógenos.

1. Identificación de proteínas abundantes en C. glutamicum

Se prepararon extractos celulares a partir de la cepa de *C. glutamicum* ATCC 13032 y de algunos derivados. Para este fin, 250 mg de cultivo celular en condiciones estándar se sedimentaron y suspendieron en 750 µl de tampón de lisis (TRIS 20 mM, EDTA 5 mM, pH 7,5) que contiene una mezcla del inhibidor de proteasa (Complete, Roche). Las células se rompieron a 4 °C en un molino mezclador (Retsch, MM 2000) usando esferas de vidrio de 0,25 - 0,5 mm. Los restos celulares se eliminaron mediante centrifugación a 22.000 rpm durante 1 hora a 4 °C. Las concentraciones proteicas se determinaron mediante el procedimiento de Popov (Popov et al. (1975) Acta. Biol. Med. Germ, 34, 1441-1446). Los extractos celulares se usaron inmediatamente o se congelaron en alícuotas a -80 °C.

Para electroforesis en gel de poliacrilamida 2D de las proteínas, 30 mg del extracto proteico bruto se resuspendieron en 450 µl de tampón de rehidratación (urea 8M, tiourea 2M, 1 % de CHAPS, DTT 20 mM, 1 % de Amfolinas 3,5 - 10)

5 y unos pocos granos de azul de bromofenol. Para enfocado isoeléctrico se usaron tiras de 24 cm de IPG precoladas con un gradiente de pH lineal de 4,5 a 5,5 en una unidad de enfocado isoeléctrico Multiphor II (Amersham Biosciences). Las proteínas se centraron usando un programa de gradiente de hasta 3500 V resultante en 65.000 Vh en total. Los geles de IPG enfocados se equilibraron durante 15 minutos en un tampón que contenía Tris-HCl 1,5 M (pH 8,8), urea 6M, 30 % (vol/vol) de glicerol, 2 % (p/vol) de dodecilsulfato sódico y 1 % (p/vol) de DTT. Para la segunda etapa de equilibrado, el DTT se reemplazó con 5 % (p/vol) de yodoacetamida, y se añadieron unos pocos granos de azul de bromofenol. La segunda dimensión se pasó en geles de dodecilsulfato sódico -12,5 % de poliacrilamida en un aparato de Ettan Dalt (Amersham Biosciences) según las recomendaciones del fabricante y después los geles se tiñeron con plata (Blum et al. (1987), Electrophoresis, 8, 93-99) en una tinción automática doméstica.

10 Las manchas de las proteínas se escindieron de los geles preparativos teñidos con azul de Coomassie (300 mg de carga de proteína total cada uno) y se digirieron con tripsina modificada (Roche, Mannheim) como describen Hermann et al. (Electrophoresis (2001), 22, 1712-1723). Las identificaciones con espectrometría de masas se realizaron en una LCQ advantage (Thermo Electron) tras la separación en nano- HPLC de los péptidos (LC Packings, columna RP18, longitud de 15 cm, i.d. 75 µm), usando el software MASCOT (David et al. (1999) Electrophoresis, 20,3551-3567).

20 En base a los resultados de la electroforesis en gel 2D se identificaron 14 proteínas como abundantes en *C. glutamicum*, ya que estas proteínas se pudieron observar en cantidades elevadas en todos los geles. Estas proteínas son: Factor de elongación Tu (nº de acceso en Genbank: X77034), glicerina-aldehído-3-fosfato-deshidrogenasa (nº de acceso en Genbank: BX927152, 6, nt. 289401-288397), fructosa bisfosfato aldolasa (nº de acceso en Genbank: BX927156, 6, nt. 134992-133958), Factor de elongación Ts (nº de acceso en Genbank: BX927154, 6, nt. 14902-14075), proteína hipotética (nº de acceso en Genbank: BX927155, 6, nt. 213489-214325), enolasa (nº de acceso en Genbank: BX927150, nt. 338561-339838) peptidil-prolil cis-trans isomerasa (nº de acceso en Genbank: BX927148, nt. 34330-34902), superóxido dismutasa (nº de acceso en Genbank: AB055218) fosfo-
25 glicerato deshidrogenasa (nº de acceso en Genbank: BX927151, nt. 306039-307631) proteína SSU Rib S1P(nº de acceso en Genbank: BX927152, 6, nt. 26874-28334) triosefosfato- isomerasa (nº de acceso en Genbank: BX927152, 6, nt. 286884-286105) isopropil malato sintasa (nº de acceso en Genbank: X70959) butan-2,3-dioldeshidrogenasa (nº de acceso en Genbank: BX927156, nt. 20798-21574) y fumarato hidratasa (nº de acceso en Genbank: BX927151, 6, nt. 18803-17394).

30 Las secuencia de codificación de estos genes se introdujeron en la función "Cusp" de la caja de herramientas EMBOSS usando parámetros estándar en un abordaje independiente de la secuencia genómica de la cepa completa de *C. glutamicum* ATCC 13032 para generar una tabla de uso de codones para el organismo como un todo.

35 Las frecuencias de uso de los codones determinadas para las 14 proteínas abundantes mencionadas anteriormente se usaron para calcular las frecuencias de uso de los codones para las proteínas abundantes en *C. glutamicum*. Las frecuencias relativas de codones del uso de los codones de las proteínas abundantes en *C. glutamicum* se encuentran en la tabla 2, mientras que las frecuencias de uso relativo de los codones del organismo como un todo se encuentran en la tabla 1.

Tabla 1 - Frecuencias relativas de uso de codones de *Corynebacterium glutamicum* ATCC 13032.

UUU 37,1	UCU 17,3	UAU 33,8	UGU 36,5
UUC 62,9	UCC 33,6	UAC 66,2	UGC 63,5
UUA 5,3	UCA 13,0	UAA 53,1	UGA 16,7
UUG 20,3	UCG 11,9	UAG 30,2	UGG 100
CUU 17,2	CCU 23,3	CAU 32,1	CGU 24,5
CUC 22,5	CCC 20,2	CAC 67,9	CGC 44,7
CUA 6,1	CCA 34,9	CAA 38,5	CGA 11,8
CUG 28,6	CCG 21,6	CAG 61,5	CGG 8,8
AUU 37,7	ACU 20,4	AAU 33,4	AGU 7,8

ES 2 539 280 T3

AUC 59,2	ACC 52,9	AAC 66,4	AGC 16,4
AUA 3,1	ACA 12,5	AAA 39,9	AGA 4,1
AUG 100	ACG 14,2	AAG 60,1	AGG 6,1
GUU 26,0	GCU 23,7	GAU 55,6	GGU 30,3
GUC 27,7	GCC 25,4	GAC 44,4	GGC 42,4
GUA 10,1	GCA 29,3	GAA 56,3	GGA 18,9
GUG 36,2	GCG 21,6	GAG 43,7	GGG 8,4
ATG* 72,5			
GTG* 20,5			
TTG* 7,0			
*Indica los codones de iniciación; las frecuencias relativas se expresan en porcentaje,			

Tabla 2 - Frecuencias relativas de uso de codones de 14 proteínas abundantes en *Corynebacterium glutamicum* ATCC 13032.

UUU 10,6	UCU 20,2	UAU 3,6	UGU 25,0
UUC 89,4	UCC 65,4	UAC 96,4	UGC 75,0
UUA 0,8	UCA 3,3	UAA 92,9	UGA 0,0
UUG 7,5	UCG 2,2	UAG 7,1	UGG 100
CUU 20,8	CCU 37,6	CAU 5,8	UCU 39,6
CUC 25,4	ACC 4,4	GAC 94,2	UGC 57,6
CUA 2,6	CCA 51,4	CAA 7,7	CGA 2,2
CUG 42,9	CCG 6,6	CAG 92,3	CGG 0,6
AUU 17,1	ACU 18,9	AAU 9,7	AGU 0,8
AUC 82,6	ACC 78,9	AAC 90,3	AGC 8,1
AUA 0,3	ACA 1,4	AAA 7,1	AGA 0,0
AUG 100	ACG 0,8	AAG 92,9	AGG 0,0
GUU 47,9	GCU 46,8	GAU 34,9	GGU 32,3
GUC 34,0	GCC 9,9	GAC 65,1	GGC 59,0
GUA 6,5	GCA 35,9	GAA 32,5	GGC 8,2

ES 2 539 280 T3

GUG 11,6	GCG 7,4	GAG 67,5	GGG 0,5
ATG* 78,6			
GTG* 21,4			
TTG* 0,0			
*Indica los codones de iniciación; las frecuencias relativas se expresan en porcentaje			

A continuación se usó la Tabla 1 para determinar los codones que son menos frecuentes para cada aminoácido en *C. glutamicum*. Esta información se muestra en la tabla 3, dada a continuación.

5

Tabla 3 - Los codones menos usados en *Corynebacterium glutamicum* ATCC 13032.

UUU	F
AGU	S
UAU	Y
UGU	C
UGA	Terminación
AUG	M
UGG	W
UUA	L
CCC	P
CAU	H
AGA	R
CAA	Q
AUA	I
ACA	T
AAU	N
TTG (si es codón de iniciación)	M
AAA	K
GUA	V
GCG	A
GAC	D
GGG	G
GAG	E

A continuación se usó la Tabla 2 para determinar los codones que son menos frecuentes para cada aminoácido en las proteínas abundantes de *C. glutamicum*. Esta información se muestra en la tabla 4, dada a continuación.

10

Tabla 4 - Los codones menos usados en las proteínas abundantes de *Corynebacterium glutamicum* ATCC 13032.

UUU	F
AGU	S
UAU	Y
UGU	C
UGA	Terminación
AUG	M
UGG	W
UUA	L
CCC	P
CAU	H
AGA, AGG	R
CAA	Q
AUA	I
ACG	T
AAU	N
TTG (si es codón de iniciación)	M
AAA	K
GUA	V
GCG	A
GAU	D
GGG	G
GAA	E

- 5 La Tabla 5 muestra las frecuencias de los codones que no se calculan en base a los codones que codifican un aminoácido específico sino en base a todos los codones para todos los aminoácidos. Los valores entre paréntesis indican el número absoluto del respectivo codón. Las frecuencias relativas de la tabla 1 se calcularon en base a estos números absolutos. Los valores hacen referencia al genoma de *C. glutamicum*.

Tabla 5 - Uso de codones de *Corynebacterium glutamicum* ATCC 13032,

UUU	13,4(25821)	UCU	11,0(21227)	UAU	7,5(14384)	UGU	2,4(4605)
UUC	22,8(43837)	UCC	21,4(41118)	UAC	14,7(28214)	UGC	4,2(8015)
UUA	5,1(9795)	UCA	8,3(15898)	UAA	1,7(3272)	UGA	0,5(1032)
UUG	19,6(37762)	UCG	7,6(14639)	UAG	1,0(1859)	UGG	14,1(27072)
CUU	16,7(32074)	CCU	11,3(21668)	CAU	6,8(12991)	CGU	13,7(26310)
CUC	21,8(41988)	CCC	9,7(18716)	CAC	14,3(27445)	CGC	24,9(47939)
CUA	5,9(11320)	CCA	16,9(32429)	CAA	13,0(24975)	CGA	6,6(12698)

ES 2 539 280 T3

CUG	27,7(53261)	CCG	10,4(20070)	CAG	20,7(39864)	CGG	4,9(9466)
AUU	21,7(41804)	ACU	12,6(24184)	AAU	10,9(21056)	AGU	4,9(9515)
ABC	34,1(65557)	ACC	32,5(62592)	AAC	21,8(42037)	AGC	10,4(20019)
AUA	1,8(3483)	ACA	7,7(14747)	AAA	13,9(26703)	AGA	2,3(4445)
AUG	22,1(42484)	ACG	8,8(16879)	AAG	20,9(40213)	AGG	3,3(6398)
GUU	20,8(40069)	GCU	25,4(48864)	GAU	33,0(63429)	GGU	24,3(46678)
GUC	22,2(42696)	GCC	27,2(52264)	GAC	26,4(50716)	GGC	34,0(65427)
GUA	8,1(15628)	GCA	31,3(60329)	GAA	35,7(68737)	GGA	15,2(29219)
GUG	28,9(55708)	GCG	23,2(44613)	GAG	27,7(53381)	GGG	6,7(12923)

Las frecuencias se indican a continuación de los codones en/1000.

La Tabla 6 muestra las frecuencias de los codones que no se calcularon en base a los codones que codifican un aminoácido específico sino en base a todos los codones para todos los aminoácidos. Los valores entre paréntesis indican el número absoluto de valores en referencia al grupo de proteínas abundantes en *C. glutamicum*.

5

Tabla 6 - Uso de codones de 14 proteínas abundantes en *Corynebacterium glutamicum* ATCC 13032,

UUU	3,6(18)	UCU	10,9(55)	UAU	0,8(4)	UAU	1,2(6)
UUC	30,0(152)	UCC	35,2(178)	UAC	21,1(107)	UGC	3,6(18)
UUA	0,6 (3)	UCA	1,8(9)	UAA	2,6(13)	UGA	0,0(0)
UUG	5,7(29)	UCG	1,2(6)	UAG	0,2(1)	UGG	8,3(42)
CUU	16,0(81)	CCU	13,4(68)	CAU	1,2(6)	CGU	17,4(88)
CUC	19,6(99)	CCC	1,6(8)	CAC	19,4(98)	CGC	25,3(128)
CUA	2,0(10)	CCA	18,4(93)	CAA	2,6(13)	CGA	1,0(5)
CUG	33,0(167)	CCG	2,4(12)	CAG	30,6(155)	CGG	0,2(1)
AUU	9,9(50)	ACU	10,5(53)	AAU	4,0(20)	AGU	0,4(2)
ABC	47,8(242)	ACC	43,7(221)	AAC	36,8(186)	AGC	4,3(22)
AUA	0,2(1)	ACA	0,8(4)	AAA	3,6 (18)	AGA	0,0(0)
AUG	17,2(87)	ACG	0,4(2)	AAG	46,4(235)	AGG	0,0(0)
GUU	43,5(220)	GCU	54,9(278)	GAU	21,9(111)	GGU	28,1(142)
GUC	30,8(156)	GCC	11,7(59)	GAC	40,9(207)	GGC	51,2(259)
GUA	5,9(30)	GCA	42,1(213)	GAA	27,9(141)	GGA	7,1(36)

GUG 10,5(53) GCG 8,7(44) GAG 57,9(293) GGG 0,4(2)

Las frecuencias se indican a continuación de los codones en /1000.

Sorprendentemente, hay muchas diferencias significativas en el uso de codones entre las tablas generadas usando todas las proteínas (genoma completo, tabla 1) en comparación con la situación en la que solo se consideran los genes abundantes especificados anteriormente (tabla 2). En la tabla 7 siguiente se muestran algunos de los ejemplos.

5

Tabla 7 - Frecuencia relativa de codones usados

Aminoácido	Codón	Genoma entero	Proteínas abundantes
q	caa	38,5 %	7,7
q	cag	61,5 %	91,3 %
y	tac	66,0 %	96,4 %
y	tac	33,8 %	3,6

3. Reducción de la expresión de una proteína endógena

Anteriormente en el presente documento se ha indicado que la modificación del uso de codones se puede usar para disminuir la expresión de proteínas. Esto se ilustrará a continuación para las proteínas isocitrato deshidrogenasa (icd) y diaminopimelato sintasa (dapA).

10

3.1 Reducción de la expresión de isocitrato deshidrogenasa (icd)

Clonación

Para reducir la actividad de la isocitrato deshidrogenasa (código de acceso en Genbank X71489), se realizaron dos cambios diferentes en el uso de codones. En todos los casos, los codones de la secuencia de codificación se cambiaron sin modificar la secuencia de aminoácidos de la proteína codificada. Las manipulaciones se realizaron todas en la única copia del cromosoma del gen *icd* de *Corynebacterium glutamicum*. La posterior medición de la actividad de la ICD permite directamente una lectura del efecto, ya que se puede asumir que refleja que el nivel de expresión proporcionado por la propia enzima no se ha modificado. Las modificaciones se muestran en la tabla 8.

15

20

Tabla 8 - Resumen de los intercambios de codones en ICD

	Nombre	Descripción	posiciones de aminoácidos afectados
1.	ICD ATG → GTG	Cambio del codón de iniciación de ATG a GTG	1 (Met)
2	ICD CA2	Cambio de del codón de glicina y de isoleucina de GGC ATT a GGG ATA	32 (Gly), 33 (Ile)

La secuencia de ICD ATG-GTG se representa en la figura 2 a). La secuencia de ICD CA se representa en la figura 3 a). Para introducir estas mutaciones en la copia del cromosoma de la región de codificación de *icd* se construyeron 2 plásmidos diferentes que permiten la manipulación sin marcadores con 2 acontecimientos de recombinación homóloga consecutivos .

25

Con este fin, las secuencias de ICD ATG-GTG y de ICD CA2 se clonaron en el vector pClik int *sacB* (Becker et al (2005), Applied and Environmental Microbiology, 71 (12), p.8587-8596) siendo un plásmido que contiene los elementos siguientes

- gen de resistencia a kanamicina
- gen de *SacB*, que se puede usar como marcador de selección positiva, ya que las células portadoras de este

30

- gen no pueden crecer en medio que contiene sacarosa.
- origen de replicación para *E. coli*
- sitio de clonación múltiple (MCS)

Este plásmido permite la integración de secuencias en el locus genómico de *C. glutamicum*.

5 **Construcción de los plásmidos**

Todos los insertos se amplificaron mediante PCR usando ADN genómico de ATCC 13032 como molde. La modificación de la región de codificación se consiguió mediante PCR de fusión usando los oligonucleótidos siguientes: La tabla muestra los cebadores usados, así como el ADN molde

Tabla 9 - Resumen de los cebadores para la clonación de las construcciones de *idh*

	PCR A	PCR B	PCR fusión	
ICD ATG → GTG	Old 441	Old 443	Old 441	Cebador 1
	Old 444	Old 442	Old 442	Cebador 2
	ADN genom. de ATCC 13032	ADN genom. de ATCC 13032	PCR A+B	Molde
ICD CA2	Old 441	Old 447	Old 441	Cebador 1
	Old 448	Old 442	Old 442	Cebador 2
	ADN genom. de ATCC 13032	ADN genom. de ATCC 13032	PCR A+B	Molde
Old 441 GAGTACCTCGAGCGAAGACCTCGCAGATTCCG (SEC ID N° 6)				
Old 442 CATGAGACGCGTGAATCTGCAGACCACTCGC (SEC ID N° 7)				
Old 443 GAGACTCGTGGCTAAGATCATCTG (SEC ID N° 8)				
Old 444 CAGATGATCTTAGCCACGAGTCTC (SEC ID N° 9)				

10

En todos los casos, el producto de la PCR de fusión se purificó, se digirió con XhoI y MluI, se purificó de nuevo y se ligó en pClik int sacB que se había linealizado con las mismas enzimas de restricción. Mediante secuenciación se confirmó la integridad de la construcción.

15 La secuencia de codificación de la secuencia optimizada ICD ATG → GTG se muestra en la Figura 2 (SEC ID N° 2). La secuencia de codificación de la secuencia optimizadas ICAD CA2 se muestra en la Figura 3 (SEC ID N° 4).

Construcción de las cepas con los niveles de expresión de ICD modificados

A continuación, los plásmidos se usaron para reemplazar la región de codificación nativa de estos genes por regiones de codificación con el uso de codificación modificado. La cepa usada fue ATCC 13032 lysC fbr.

20 Son necesarios dos acontecimientos de recombinación consecutivos uno en cada región, aguas arriba y aguas abajo, respectivamente, para cambiar la secuencia de codificación completa. El procedimiento de reemplazar los genes endógenos por los genes optimizados se describe, en principio, en la publicación de Becker et al. (véase en lo que antecede). Las etapas más importantes son:

- Introducción de los plásmidos en la cepa mediante electroporación. La etapa se describe en, por ejemplo, el documento DE 10046870.
- 25 – La selección de clones que han integrado con éxito el plásmido tras un primer acontecimiento de recombinación homóloga en el genoma. Esta selección se consigue mediante el cultivo en placas de agar con kanamicina. Además de la etapa de selección, el éxito de la recombinación se puede comprobar mediante PCR de las colonias.
- 30 – Los cebadores usados para confirmar la presencia del plásmido en el genoma fueron: BK1776 (AACGGCAGGTATATGTGATG) (SEC ID N° 12) y OLD 450 (CGAGTAGGTCGCGAGCAG) (SEC ID N° 13). Los clones positivos dan una banda de aproximadamente 600 pb)
- Incubando un clon positivo en un medio sin kanamicina se permite un segundo acontecimiento de recombinación
- Los clones en los que la estructura del vector se ha eliminado con éxito mediante un segundo acontecimiento de recombinación se identifican mediante crecimiento en medio que contiene sacarosa. Solo sobrevivirán los clones

que hayan perdido la estructura del vector que comprende el gen SacB.

- Después, los clones en los que dos acontecimientos de recombinación han conducido a un reemplazo con éxito de la región de codificación de *idh* nativa se identificaron mediante secuenciación de un producto de PCR que abarca la región relevante. El producto de PCR se generó usando ADN genómico de clones individuales como molde y los cebadores OLD 441 y OLD 442. El producto de la PCR se purificó y secuenció con Old 471 (GAATCCAACCCACGTTTCAGGC) (SEC ID N° 14)

Se pueden usar diferentes cepas de *C. glutamicum* para reemplazar la copia endógena de ICD. No obstante, se prefiere usar una cepa de producción de lisina de *C. glutamicum* tal como, por ejemplo, ATCC 13032 *lysC^{fbr}* u otros derivados de ATCC13032 o ATCC 13286

ATCC13032 *lysC^{fbr}* se puede producir a partir de ATCC13032. Con el fin de generar una cepa productora de lisina se realizó un intercambio alélico del gen salvaje de *lysC* en *C. glutamicum* ATCC13032. Con este fin se introdujo un intercambio e nucleótidos en el gen de *lys C* de forma que la proteína resultante porta una isoleucina en la posición 311 en lugar de treonina. La construcción detallada de esta cepa se describe en la solicitud de patente WO2005059093. El n° de acceso del gen *lysC* es P26512.

Para analizar el efecto del uso de codones modificado IDH ATG-GTG y IDH CA2, las cepas optimizadas se comparan con la productividad de lisina de la cepa parental.

Determinación de la actividad de ICD

En uno de dos clones de cada cepa mutantes se analizó la actividad de ICD. Las células se cultivaron en cultivo líquido durante la noche a 30 °C, se recogieron en fase de crecimiento exponencial mediante centrifugación. Las células se lavaron dos veces con Tris 50 mM, pH 7,0. 200 mg de células se resuspendieron en 800 µl de tampón de lisis (Tris-HCl 50 mM, pH 7,0, MgCl₂ 10 mM, DTT 1 mM, 10 % de glicerol) y se rompieron mediante batido con esferas (Ribolyser, 2x 30s, intensidad 6). Los restos celulares se sedimentaron mediante centrifugación (centrífuga de mesa, 30 min, 13 K). El sobrenadante resultante es un extracto de proteínas solubles que se usó como el siguiente ensayo enzimático.

La actividad de ICD se monitorizó mediante el incremento de la absorción a 340 nm debido a la reducción de NADP en un volumen total de of 1 ml en las siguientes condiciones:

Trietanolamina-cloruro 30 mM, pH 7,4, NADP 0,4 mM , DL-Isocitrato 8 mM, MnSO₄ 2 mM, lisado celular correspondiente a 0,1-0,2 mg de proteínas

Las actividades de ICD se calcularon usando el coeficiente de extinción molar de 6,22/mM*cm para NADPH.

Resultados

Las actividades de ICD medidas fueron como sigue:

Tabla 10- Actividad de ICD

Cepa	Clon	Actividad específica del extracto celular U*µmol/ml*min*mg de proteína
ATCC <i>lysC fbr</i>		0,29
ATCC <i>lysC fbr</i>		0,26
ICD ATG → GTG		0,04
ICD CA2	1	0,10
ICD CA2	2	0,10

Efecto sobre la productividad de lisina

Para analizar el efecto de la expresión modificada de ICD sobre la productividad de lisina, las cepas optimizadas se comparan con la productividad de lisina de las cepas parentales.

Con este fin, las cepas se cultivaron en placas CM (10 % de sacarosa, 10 g/l de glucosa, 2,5 g/l de NaCl, 2 g/l de urea, 10 g/l de Bacto Peptona, 10 g/l de extracto de levadura, 22 g/l de agar) durante 2 días a 30 °C. Después, se rasparon las células de las placas y se resuspendieron en solución salina. Para el cultivo principal se incubaron 10

ml del medio 1 y 0,5 g de CaCo3 en autoclave en un matraz Erlenmeyer de 100 ml junto con la suspensión celular hasta una DO 600 de 1,5. Después, las células se cultivaron durante 72 horas en un agitador del tipo Infors AJ118 (Infors, Bottmingen, Suiza) a 220 rpm. El medio 1 tenía la misma concentración como se ha mencionado en el experimento 3.3:

5 Después, se determinó la concentración de lisina que se segrega al medio. Esto se realizó usando HPLC en un sistema Agilent 1100 Series LC HPLC. Una precolumna de derivación con orto-ftalaldehído permitió cuantificar el aminoácido formado. La separación de la mezcla de aminoácido se puede realizar en una columna Hypersil AA- (Agilent).

10 Los valores de concentración de lisina determinados mostrados son los datos promedio de 2 cultivos independientes. Las desviaciones del promedio siempre fueron inferiores al 4 %

Tabla 11-Productividad de lisina

Cepa	Clon	Cantidad relativa de lisina [%]	DO relativa [%]
ATCC lysC fbr		100,00	100,00 %
ATCC lysC fbr		99,81	101,22 %
ICD ATG → GTG		102,34	92,77
ICD CA2	1	101,44	99,80
ICD CA2	2	104,85	96,23

15 Se puede ver fácilmente que las cepas con un actividad ICD menor tienen productividades mayores de lisina. Como toda la fuente de carbono se ha usado tras 72 horas, se puede ver directamente que el rendimiento de carbono (cantidad de producto formado por azúcar consumida) es mayor en estas cepas.

La cantidad de otros aminoácidos en el medio después de otro cultivo también se analizó . Es interesante le hecho de que la cantidad de glicina también se ve influida por la expresión de ICD.

En la tabla se muestra la cantidad de glicina obtenida tras 48 horas de cultivo. Cada clon se cultivó en tres matraces independientes.

20

Tabla 12-Productividad de glicina

		Glicina [µmol/l]
Medio		5,72
ATCC lysC fbr		78,56
		81,33
		79,40
ATCC lysC fbr		86,29
		78,51
		81,35
ICD ATG → GTG		110,44
		108,51
		117,36
ICD CA2	1	103,29
		104,56
		96,26

ICD CA2	2	98,83
		98,72
		99,39

Se puede ver que las cepas con una actividad de ICD menor acumulan cantidades mayores de glicina.

5 En un experimento adicional, la isocitrato deshidrogenasa portadora de la mutación ATG-GTG mencionada anteriormente en el codón de iniciación se clonó en pClik como se ha descrito anteriormente, lo que conduce a pClik int sacB ICD (ATG-GTG) (SEC ID N° 15). Después, la cepa M2620 se construyó mediante *campbelización in* y *campbelización out* del plásmido pClik int sacB ICDH (ATG-GTG) (SEC ID N° 15) en el genoma de la cepa OM469. La cepa OM469 se ha descrito en el documento WO2007012078.

10 La cepa se cultivó como se describe en el documento WO2007020295. Tras 48 horas de incubación a 30°C, las muestras se analizaron para determinar el consumo de azúcar. Se descubrió que las cepas habían usado todo el azúcar añadido, lo que significa que todas las cepas habían usado la misma cantidad de fuente de carbono. La metionina sintetizada se determinó mediante HPLC como se ha descrito anteriormente y en el documento WO2007020295.

Tabla 13-Producción de metionina

Cepa	Metionina (mM)
OM 469	10,2
M2620	23,7

15 A partir de los datos de la tabla 13 se puede ver que la cepa M2620 con un codón de iniciación alterado del gen de ICDH y, por tanto, una actividad ICDH alterada, tiene una productividad de metionina más alta. Como toda la fuente de carbono se ha usado tras 48 horas, se puede ver directamente que el rendimiento de carbono (cantidad de producto formado por azúcar consumida) para la metionina producida es mayor en esta cepa.

3.2. Reducción de la expresión de proteínas de *dapA* en *C. glutamicum*

20 A continuación se describe cómo disminuir la cantidad de *dapA* adaptando el uso de codones como se ha mencionado antes.

La enzima dihidropicolinato sintasa (*dapA*) es importante para la biosíntesis de lisina. La secuencia salvaje de *dapA* de *C. glutamicum* se proporciona como la SEC ID N° 16. El uso de codones de la secuencia de codificación de *dapA* se determinó usando la función Cusp del paquete de software EMBOSS .

25 Después, las bases correspondientes a las posiciones 12 (T a G), 27 (C a A), 30 (A a G), 45 (C a G) y 54 (A a G) se alteraron usando procedimientos establecidos de mutagénesis (Quikchange kit, Stratagene La Jolla USA), lo que tiene como resultado codones alterados que seguían codificando los mismos aminoácidos (SEC ID N° 17).

La secuencia del gen mutado se clonó en un vector para integración cromosómica como se ha mencionado en el documento WO2007020295.

30 En resumen, la cepa M2059 se construyó mediante *campbelización in* y *out* del plásmido pK19 PSOD ask (SEC ID N° 18) en la cepa OM99 para sobreexpresar el gen de la aspartoquinasa. La cepa M2121 se construyó mediante *campbelización in* y *out* del plásmido pCLIK Int Psod Glucosa-6-PDH (SEC ID N° 19) en la cepa M2059 para sobreexpresar el gen de la glucosa-6-fosfato deshidrogenasa (gen *zwf*). La cepa M2199 se construyó mediante *campbelización in* y *out* del plásmido pK19 sacB Glu-6-P-DH (SEC ID N° 20) reemplazo en la cepa M2121 para
35 expresar el gen mutante de la glucosa-6-fosfato deshidrogenasa (gen *zwf*). La cepa M2271 se construyó mediante *campbelización in* y *out* del plásmido pK19 sacB Glu-6-P-DH (SEC ID N° 21) en la cepa M2199 para expresar un gen *pycA* mutado. La cepa M2376 se construyó mediante *campbelización in* y *out* del plásmido ppCLIK int sacB, codón *dapA* bajo (SEC ID N° 22) en la cepa M2271 para expresar un gen *dapA* mutado.

40 La cepa resultante se denominó M2376 y se cultivó como se describe en 2007020295. Los aminoácidos se determinaron tras 48 horas de incubación y se describen como la concentración en mM (véase la tabla 14).

Tabla 14 - Producción de lisina y homoserina

Cepa	homoserina (mM)
M2271	78
M2376 (= M2271 + dapA mutada)	103

Se encontró que las cepas que expresan el gen *dapA* mutado mostraron un incremento de la acumulación de homoserina en comparación con la cepa isogénica que muestra que la actividad de *dapA* disminuía en la versión sin mutar de *dapA*

5 LISTADO DE SECUENCIAS

<110> BASF AG

<120> Procedimiento de reducción de la expresión génica mediante el uso de codones modificado

<130> B 8441 / DB

<150> EP 06122882.1

10 <151> 2006-10-24

<160> 22

<170> PatentIn version 3.3

<210> 1

<211> 2217

15 <212> ADN

<213> *Corynebacterium glutamicum*

<400> 1

ES 2 539 280 T3

atggctaaga	tcactctggac	ccgcaccgac	gaagcaccgc	tgctcgcgac	ctactcgcctg	60
aagccggtcg	tcgaggcatt	tgctgctacc	gcgggcattg	aggtcgagac	ccgggacatt	120
tcactcgcctg	gacgcatcct	cgcccagttc	ccagagcgcc	tcaccgaaga	tcagaaggta	180
ggcaacgcac	tcgcagaact	cggcgagctt	gctaagactc	ctgaagcaaa	catcattaag	240
cttccaaaca	tctccgcttc	tgttccacag	ctcaaggctg	ctattaagga	actgcaggac	300
cagggctacg	acatcccaga	actgcctgat	aacgccacca	ccgacgagga	aaaagacatc	360
ctcgcacgct	acaacgctgt	taagggttcc	gctgtgaacc	cagtgcctgcg	tgaaggcaac	420
tctgaccgcc	gcgcaccaat	cgctgtcaag	aactttgtta	agaagttccc	acaccgcatg	480
ggcgagtgg	ctgcagattc	caagaccaac	gttgcaacca	tggatgcaaa	cgacttccgc	540
cacaacgaga	agtccatcat	cctcgacgct	gctgatgaag	ttcagatcaa	gcacatcgca	600
gctgacggca	ccgagaccat	cctcaaggac	agcctcaagc	ttcttgaagg	cgaagttcta	660
gacggaaccg	ttctgtccgc	aaaggcactg	gacgcattcc	ttctcgagca	ggtcgctcgc	720
gcaaaggcag	aaggtatcct	cttctccgca	cacctgaagg	ccaccatgat	gaaggtctcc	780
gaccaatca	tcttcggcca	cgttgtgcgc	gcttacttcg	cagacgtttt	cgcacagtac	840
ggtgagcagc	tgctcgcagc	tggcctcaac	ggcgaaaacg	gcctcgcctgc	aatcctctcc	900
ggcttggagt	ccctggacaa	cggcgaagaa	atcaaggctg	cattcgagaa	gggcttggaa	960
gacggcccag	acctggccat	ggttaactcc	gctcgcggca	tcaccaacct	gcatgtccct	1020
tccgatgtca	tcgtggacgc	ttccatgcc	gcaatgattc	gtacctccgg	ccacatgtgg	1080
aacaaagacg	accaggagca	ggacaccctg	gcaatcatcc	cagactcctc	ctacgctggc	1140
gtctaccaga	ccgttatcga	agactgccgc	aagaacggcg	cattcgatcc	aaccaccatg	1200
ggtaccgtcc	ctaacgttgg	tctgatggct	cagaaggctg	aagagtacgg	ctcccatgac	1260
aagaccttcc	gcatcgaagc	agacgggtgtg	gttcagggtg	tttctccaa	cggcgacggt	1320
ctcatcgagc	acgacgttga	ggcaaatgac	atctggcgtg	catgccaggt	caaggatgcc	1380
ccaatccagg	attgggtaaa	gcttgctgtc	acccgctccc	gtctctccgg	aatgcctgca	1440

ES 2 539 280 T3

gtgttctggt tggatccaga gcgcgcacac gaccgcaacc tggcttcctt cgttgagaag 1500
 tacctggctg accacgacac cgagggcctg gacatccaga tcctctccc tggtgaggca 1560
 acccagctct ccatcgaccg catccgccgt ggcgaggaca ccatctctgt caccggtaac 1620
 gttctgcgtg actacaacac cgaccttctt ccaatcctgg agctgggcac ctctgcaaag 1680
 atgctgtctg tcgttccttt gatggctggc ggcggactgt tcgagaccgg tgctggtgga 1740
 tctgctccta agcacgtcca gcaggttcag gaagaaaacc acctgcgttg ggattccctc 1800
 ggtgagttcc tcgactggc tgagtccttc cgccacgagc tcaacaaca cggcaacacc 1860
 aaggccggcg ttctggctga cgctctggac aaggcaactg agaagctgct gaacgaagag 1920
 aagtcccat cccgcaaggt tggcgagatc gacaaccgtg gctcccactt ctggctgacc 1980
 aagttctggg ctgacgagct cgctgctcag accgaggacg cagatctggc tgctaccttc 2040
 gcaccagtcg cagaagcact gaacacaggc gctgcagaca tcgatgctgc actgctcgca 2100
 gttcagggtg gagcaactga ccttggtggc tactactccc ctaacgagga gaagctcacc 2160
 aacatcatgc gcccagtcgc acagttcaac gagatcgttg acgactgaa gaagtaa 2217

<210> 2

<211> 2217

5 <212> ADN

<213> Artificial

<220>

<223> isocitrato deshidrogenasa, *Corynebacterium glutamicum*, mutación ATG-GTG en el primer codón

<400> 2

gtggctaaga tcatctggac ccgcaccgac gaagcaccgc tgctcgcgac ctactcgctg 60
 aagccggctg tcgaggcatt tgctgctacc gcgggcattg aggtcgagac ccgggacatt 120
 tctctcgctg gacgcatcct cgcccagttc ccagagcgcc tcaccgaaga tcagaaggta 180
 ggcaacgcac tcgcagaact cggcgagctt gctaagactc ctgaagcaaa catcattaag 240
 cttccaaaca tctccgcttc tgttccacag ctcaaggctg ctattaagga actgcaggac 300
 cagggctacg acatcccaga actgcctgat aacgccacca ccgacgagga aaaagacatc 360
 ctgcacgct acaacgctgt taagggttcc gctgtgaacc cagtgctgcg tgaaggcaac 420
 tctgaccgcc gcgcaccaat cgctgtcaag aactttgtta agaagttccc acaccgcatg 480
 ggcgagtggc ctgcagattc caagaccaac gttgcaacca tggatgcaaa cgacttccgc 540
 cacaacgaga agtccatcat cctcgacgct gctgatgaag ttcagatcaa gcacatcgca 600
 gctgacggca ccgagaccat cctcaaggac agcctcaagc ttcttgaagg cgaagtctta 660
 gacggaaccg ttctgtccgc aaaggcactg gacgcattcc ttctcgagca ggctcgctcg 720
 gcaaaggcag aaggtatcct cttctccgca cacctgaagg ccacatgat gaaggtctcc 780
 gacccaatca tcttcggcca cgttgtgctg gcttacttcg cagacgtttt cgcacagtac 840
 ggtgagcagc tgctcgcagc tggcctcaac ggcgaaaacg gcctcgctgc aatcctctcc 900

10

ES 2 539 280 T3

ggcttggagt ccctggacaa cggcgaagaa atcaaggctg cattcgagaa gggcttggaa 960
 gacggcccag acctggccat ggttaactcc gctcgcggca tcaccaacct gcatgtccct 1020
 tccgatgtca tcgtggacgc ttccatgcca gcaatgattc gtacctccgg ccacatgtgg 1080
 aacaaagacg accaggagca ggacaccctg gcaatcatcc cagactcctc ctacgctggc 1140
 gtctaccaga ccgttatcga agactgccgc aagaacggcg cattcgatcc aaccaccatg 1200
 ggtaccgtcc ctaacgttgg tctgatggct cagaaggctg aagagtacgg ctcccatgac 1260
 aagaccttcc gcatcgaagc agacgggtgtg gttcagggtg tttcctcaa cggcgacgtt 1320
 ctcatcgagc acgacgttga ggcaaatgac atctggcgtg catgccaggt caaggatgcc 1380
 ccaatccagg attgggtaaa gcttgctgtc acccgctccc gtctctccgg aatgcctgca 1440
 gtgttctggg tggatccaga gcgcgcacac gaccgcaacc tggcttccct cgttgagaag 1500
 tacctggctg accacgacac cgagggcctg gacatccaga tcctctcccc tgttgaggca 1560
 acccagctct ccatcgaccg catccgccgt ggcgaggaca ccatctctgt caccggtaac 1620
 gttctgcgtg actacaacac cgacctctc ccaatcctgg agctgggcac ctctgcaaag 1680
 atgtgtctg tcgttcctt gatggctggc ggcggactgt tcgagaccgg tgctggtgga 1740
 tctgctccta agcacgtcca gcaggttcag gaagaaaacc acctgcgttg ggattccctc 1800
 ggtgagttcc tcgcactggc tgagtcctc cgccacgagc tcaacaacaa cggcaacacc 1860
 aaggccggcg ttctggctga cgctctggac aaggcaactg agaagctgct gaacgaagag 1920
 aagtcccat cccgcaaggc tggcgagatc gacaaccgtg gctcccactt ctggctgacc 1980
 aagttctggg ctgacgagct cgctgctcag accgaggacg cagatctggc tgctacctc 2040
 gcaccagtcg cagaagcact gaacacaggc gctgcagaca tcgatgctgc actgctcgca 2100
 gttcaggggtg gagcaactga ccttggtggc tactactccc ctaacgagga gaagctcacc 2160
 aacatcatgc gccagtcgc acagttcaac gagatcgttg acgcaactgaa gaagtaa 2217

- 5 <210> 3
- <211> 1002
- <212> ADN
- <213> Artificial

- 10 <220>
- <223> Vectorinsert con isocitrato deshidrogenasa, *Corynebacterium glutamicum*, mutación ATG-GTG

<400> 3

ES 2 539 280 T3

```

ctcgagcgaa gacctcgag attccgatat tccaggaacc gccatgatcg aaatcccctc    60
agatgacgat gcacttgcca tcgagggacc ttctccatc gatgtgaaat ggctgccccg    120
caacggccgc aagcacggtg aattgttgat ggaaaccctg gccctccacc atgaagaaac    180
agaagctgca gccacctccg aaggcgaact tgtgtgggag actcctgtgt tctccgccac    240
tggcgaacag atcacagaat ccaaccacg ttcaggcgac tactactgga ttgctggcga    300
aagtgggtgc gtgaccagca ttcgtcgatc tctagtgaag gagaaaggcc tcgaccgttc    360
ccaagtggca ttcattgggt attggaaca cggcgtttcc atgcggggct gaaactgcca    420

ccataggcgc cagcaattag tagaactctg tattctaggt agctgaacaa aagagcccat    480
caaccaagga gactcgtggc taagatcatc tggaccgcga ccgacgaagc accgctgctc    540
gcgacctact cgctgaagcc ggtcgtcgag gcatttgctg ctaccgcggg cattgaggtc    600
gagaccggg acatttcaact cgctggacgc atcctcgcgc agttcccaga gcgcctcacc    660
gaagatcaga aggtaggcaa cgcactcgca gaactcggcg agcttgctaa gactcctgaa    720
gcaaacaatca ttaagcttcc aaacatctcc gcttctgttc cacagctcaa ggctgctatt    780
aaggaaactgc aggaccaggg ctacgacatc ccagaactgc ctgataacgc caccaccgac    840
gaggaaaaag acatcctcgc acgctacaac gctgttaagg gttccgctgt gaaccagtg    900
ctgctgtaag gcaactctga ccgccgcgca ccaatcgtg tcaagaactt tgtaagaag    960
tccccacacc gcatgggcca gtggtctgca gattccacgc gt                                1002

```

<210> 4

5 <211> 2217

<212> ADN

<213> Artificial

<220>

10 <223> isocitrato deshidrogenasa, *Corynebacterium glutamicum*, mutación GGC ATT--->GGG ATA en las posiciones 32 y 33

<400> 4

ES 2 539 280 T3

atggctaaga	tcctctggac	ccgcaccgac	gaagcaccgc	tgctcgcgac	ctactcgcctg	60
aagccggctc	tcgaggcatt	tgctgctacc	gcgggatag	aggtcgagac	ccgggacatt	120
tcactcgcctg	gacgcctcct	cgcccagttc	ccagagcgcc	tcaccgaaga	tcagaaggta	180
ggcaacgcac	tcgcagaact	cggcgagctt	gctaagactc	ctgaagcaaa	catcattaag	240
cttccaaaca	tctccgcttc	tgttccacag	ctcaaggctg	ctattaagga	actgcaggac	300
cagggctacg	acatcccaga	actgcctgat	aacgccacca	ccgacgagga	aaaagacatc	360
ctcgcacgct	acaacgctgt	taagggttcc	gctgtgaacc	cagtgcctgc	tgaaggcaac	420
tctgaccgcc	gcgcaccaat	cgctgtcaag	aactttgtta	agaagtccc	acaccgcatg	480
ggcgagtgg	ctgcagattc	caagaccaac	gttgcaacca	tggatgcaaa	cgacttccgc	540
cacaacgaga	agtccatcat	cctcgacgct	gctgatgaag	ttcagatcaa	gcacatcgca	600
gctgacggca	ccgagaccat	cctcaaggac	agcctcaagc	ttcttgaagg	cgaagttcta	660
gacggaaccg	ttctgtccgc	aaaggcactg	gacgcattcc	ttctcgagca	ggtcgctcgc	720
gcaaaggcag	aaggtatcct	cttctccgca	cacctgaagg	ccaccatgat	gaaggctccc	780
gaccaatca	tcttcggcca	cgttgtgcgc	gcttacttcg	cagacgtttt	cgcacagtac	840
ggtgagcagc	tgctcgcagc	tggcctcaac	ggcgaaaacg	gcctcgcctg	aatcctctcc	900
ggcttggagt	ccctggacaa	cggcgaagaa	atcaaggctg	cattcgagaa	gggcttggaa	960
gacggcccag	acctggccat	ggttaactcc	gctcgcggca	tcaccaacct	gcatgtccct	1020
tccgatgtca	tcgtggacgc	ttccatgcca	gcaatgattc	gtacctccgg	ccacatgtgg	1080

ES 2 539 280 T3

aacaaagacg accaggagca ggacaccctg gcaatcatcc cagactcctc ctacgctggc 1140
 gtctaccaga ccggtatcga agactgccgc aagaacggcg cattcgatcc aaccaccatg 1200
 ggtaccgtcc ctaacgttgg tctgatggct cagaaggctg aagagtacgg ctcccatgac 1260
 aagaccttcc gcatcgaagc agacgggtg gttcagggtg tttcctcaa cggcgacggt 1320
 ctcatcgagc acgacgttga ggcaaagac atctggcgtg catgccaggt caaggatgcc 1380
 ccaatccagg attgggtaaa gcttgctgtc acccgctccc gtctctccgg aatgcctgca 1440
 gtgttctggt tggatccaga gcgcgcacac gaccgcaacc tggcttccct cgttgagaag 1500
 tacctggctg accacgacac cgagggcctg gacatccaga tcctctcccc tgttgaggca 1560
 acccagctct ccatcgaccg catccgccgt ggcgaggaca ccatctctgt caccggtaac 1620
 gttctgcgtg actacaacac cgacctctc ccaatcctgg agctgggac ctctgcaaag 1680
 atgctgtctg tcgttccttt gatggctggc ggcggactgt tcgagaccgg tgctggtgga 1740
 tctgctccta agcacgtcca gcaggttcag gaagaaaacc acctgcgctg ggattccctc 1800
 ggtgagttcc tcgactggc tgagtccttc cgccacgagc tcaacaacaa cggcaacacc 1860
 aaggccggcg ttctggctga cgctctggac aaggcaactg agaagctgct gaacgaagag 1920
 aagtcccat cccgcaagg tggcgagatc gacaaccgtg gctcccactt ctggctgacc 1980
 aagttctggg ctgacgagct cgctgctcag accgaggacg cagatctggc tgctacctc 2040
 gcaccagtcg cagaagcact gaacacaggc gctgcagaca tcgatgctgc actgctcgca 2100
 gttcaggggtg gagcaactga ccttgggtggc tactactccc ctaacgagga gaagctcacc 2160
 aacatcatgc gccagtcgc acagttcaac gagatcgctg acgactgaa gaagtaa 2217

<210> 5

5 <211> 1002

<212> ADN

<213> Artificial

<220>

10 <223> Vectorinsert con isocitrato deshidrogenasa, *Corynebacterium glutamicum*, mutación GGC ATT--->GGG ATA en las posiciones 32 y 33

<400> 5

ES 2 539 280 T3

ctcgagcgaa gacctcgag attccgatat tccaggaacc gccatgatcg aaatccccctc 60
 agatgacgat gcacttgcca tcgagggacc ttcctccatc gatgtgaaat ggctgccccg 120
 caacggccgc aagcacggtg aattgttgat ggaaaccctg gccctccacc atgaagaaac 180
 agaagctgca gccacctccg aaggcgaact tgtgtgggag actcctgtgt tctccgccac 240
 tggcgaacag atcacagaat ccaaccacg ttcaggcgac tactactgga ttgctggcga 300
 aagtgggtgc gtgaccagca ttcgctgac tctagtgaaa gagaaaggcc tcgaccgttc 360
 ccaagtggca ttcatggggg attggaaaca cggcgtttcc atgcggggct gaaactgcca 420
 ccatagggcg cagcaattag tagaacctg tattctaggt agctgaacaa aagagccccat 480
 caaccaagga gactcatggc taagatcadc tggacccgca ccgacgaagc accgctgctc 540
 gcgacctact cgctgaagcc ggctcgtcag gcatttgctg ctaccgcggg gatagaggtc 600

gagacccggg acatttact cgctggacgc atcctcgccc agttcccaga ggcctcacc 660
 gaagatcaga aggtaggcaa cgcactcgca gaactcggcg agcttgctaa gactcctgaa 720
 gcaaacatca ttaagcttcc aaacatctcc gcttctgttc cacagctcaa ggctgctatt 780
 aaggaactgc aggaccaggg ctacgacatc ccagaactgc ctgataacgc caccaccgac 840
 gaggaaaaag acatcctcgc acgctacaac gctgttaagg gttccgctgt gaaccagtg 900
 ctgctgtaag gcaactctga ccgcccgcga ccaatcgtg tcaagaactt tgtaagaag 960
 ttccacacc gcattggcga gtggtctgca gattccacgc gt 1002

<210> 6

<211> 32

5 <212> ADN

<213> Artificial

<220>

<223> oligonucleótido; Cebador old 441

10

<400> 6

gagtacctcg agcgaagacc tcgagattc cg 32

<210> 7

15 <211> 32

45 <212> ADN

<213> Artificial

<220>

<223> Oligonucleótido; Cebador old 442

50

<400> 7

catgagacgc gtggaatctg cagaccactc gc 32

5

<210> 8

<211> 24

<212> ADN

<213> Artificial

10

<220>

<223> Oligonucleótido; Cebador old 443

<400> 8

15 gagactcgtg gctaagatca tctg 24

<210> 9

<211> 24

<212> ADN

20 <213> Artificial

<220>

<223> Oligonucleótido; Cebador old 444

25 <400> 9

cagatgatct tagccacgag tctc 24

<210> 10

<211> 18

30 <212> ADN

<213> Artificial

<220>

<223> Oligonucleótido; Cebador old 447

35

<400> 10

ctaccgcggg gatagagg 18

<210> 11

<211> 18

<212> ADN

5 <213> Artificial

<220>

<223> oligonucleótido; Cebador old 448

10 <400> 11

cctctatccc cgcggtag 18

<210> 12

<211> 20

15 <212> ADN

<213> Artificial

<220>

<223> Oligonucleótido; cebador BK1776

20

<400> 12

aacggcaggt atatgtgatg 20

<210> 13

25 <211> 18

<212> ADN

<213> Artificial

<220>

30 <223> oligonucleótido; Cebador old 450

<400> 13

cgagtaggtc gcgagcag 18

35

<210> 14

<211> 21

ES 2 539 280 T3

<212> ADN

<213> Artificial

<220>

5 <223> oligonucleótido; Cebador old 471

<400> 14

gaatccaacc cacgttcagg c 21

10 <210> 15

<211> 5293

<212> ADN

<213> Artificial

15 <220>

<223> plásmido pClik int sacB ICDH (ATG-GTG)

<400> 15

20 **ctcgagcga gacctcgag attccgat tccaggaacc gccatgatc aaatcccctc 60**
 agatgacgat gcacttgcca tcgagggacc ttctccatc gatgtgaaat ggctgccccg 120

caacggccgc	aagcacggtg	aattggtgat	ggaaaccctg	gccctccacc	atgaagaaac	180
agaagctgca	gccacctccg	aaggcgaact	tgtgtgggag	actcctgtgt	tctccgccac	240
tggcgaacag	atcacagaat	ccaacccacg	ttcaggcgac	tactactgga	ttgctggcga	300
aagtggtgtc	gtgaccagca	ttcgtcgatc	tctagtgaaa	gagaaaggcc	tcgaccgttc	360
ccaagtggca	ttcatggggg	attggaaaca	cggcgtttcc	atgcggggct	gaaactgcc	420
ccataggcgc	cagcaattag	tagaacactg	tattctaggt	agctgaacaa	aagagcccat	480
caaccaagga	gactcgtggc	taagatcatc	tggaccgcga	ccgacgaagc	accgctgctc	540
gcgacctact	cgctgaagcc	ggcgtcgag	gcatttgctg	ctaccgctgg	cattgaggtc	600
gagaccggg	acatttcact	cgctggacgc	atcctcgccc	agttcccaga	gcgctcacc	660
gaagatcaga	aggtaggcaa	cgcactcgca	gaactcggcg	agcttgctaa	gactcctgaa	720
gcaaacatca	ttaagcttcc	aaacatctcc	gcttctgttc	cacagctcaa	ggctgctatt	780
aaggaactgc	aggaccaggg	ctacgacatc	ccagaactgc	ctgataacgc	caccaccgac	840
gaggaaaaag	acatcctcgc	acgctacaac	gctgttaagg	gttccgctgt	gaaccagtg	900
ctgcgtgaag	gcaactctga	ccgccgcgca	ccaatcgctg	tcaagaactt	tgtaagaag	960
ttccacacc	gcatgggcga	gtggtctgca	gattccacgc	gtcatatgac	tagttcggac	1020
ctagggatat	cgtcgacatc	gatgctcttc	tgcgttaatt	aacaattggg	atcctctaga	1080
cccgggattt	aatgatccg	ctagcgggct	gctaaaggaa	gcggaacacg	tagaaagcca	1140
gtccgcagaa	acggtgctga	ccccggatga	atgtcagcta	ctgggctatc	tggacaaggg	1200
aaaacgcaag	cgcaaagaga	aagcaggtag	cttgacgtgg	gcttacatgg	cgatagctag	1260
actgggcggt	tttatggaca	gcaagcgaac	cggaattgcc	agctggggcg	ccctctggta	1320
aggttgggaa	gccctgcaaa	gtaaactgga	tggctttctt	gccgccaagg	atctgatggc	1380
gcaggggatc	aagatctgat	caagagacag	gatgaggatc	gtttcgcatg	attgaacaag	1440
atggattgca	cgcaggttct	ccggccgctt	gggtggagag	gctattcggc	tatgactggg	1500
cacaacagac	aatcggctgc	tctgatgccg	ccgtgttccg	gctgtcagcg	caggggcgcc	1560
cggttctttt	tgtcaagacc	gacctgtccg	gtgccctgaa	tgaactgcag	gacgaggcag	1620
cgcggctatc	gtggctggcc	acgacgggcg	ttccttgccg	agctgtgctc	gacgttgcca	1680
ctgaagcggg	aagggactgg	ctgctattgg	gcgaagtgcc	ggggcaggat	ctcctgtcat	1740
ctcaccttgc	tcctgccgag	aaagtatcca	tcatggctga	tgcaatgcgg	cggctgcata	1800
cgcttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	acatcgcatc	gagcgagcac	1860
gtactcggat	ggaagccggt	cttgtcgatc	aggatgatct	ggacgaagag	catcaggggc	1920
tcgcccagc	cgaactgttc	gccaggctca	aggcgcgcat	gcccgacggc	gaggatctcg	1980
tcgtgacca	tggcgatgcc	tgcttgccga	atatcatggt	ggaaaatggc	cgcttttctg	2040
gattcatcga	ctgtggccgg	ctgggtgtgg	cggaccgcta	tcaggacata	gcgttggtca	2100
cccgtgatat	tgctgaagag	cttggcggcg	aatgggctga	ccgcttcctc	gtgctttacg	2160

gtatcgccgc	tcccgattcg	cagcgcacg	ccttctatcg	ccttcttgac	gagttcttct	2220
gagcgggact	ctggggttcg	aatgaccga	ccaagcgacg	cccaacctgc	catcacgaga	2280
tttcgattcc	accgccgcct	tctatgaaag	gttgggcttc	ggaatcgttt	tccgggacgc	2340
cggctggatg	atcctccagc	gcggggatct	catgctggag	ttcttcgccc	acgctagcgg	2400
cgcgccggcc	ggccccggtg	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	2460
ggcgtcttc	cgcttcctcg	ctcactgact	cgctgcgctc	ggtcgttcgg	ctgcggcgag	2520
cggtatcagc	tactcaaag	gcggtaatac	ggttatccac	agaatcaggg	gataacgcag	2580
gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcggtgc	2640
tggcgttttt	ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	2700
agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	2760
tcgtgcgctc	tctgtttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	2820
cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	2880
ttcgtccaa	gctgggctgt	gtgcacgaac	cccccgttca	gcccgaccgc	tgcgccttat	2940
ccggtaaacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	3000
ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	3060
ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	3120
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaacaacc	accgctggta	3180
gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	3240
atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgtaagggga	3300
ttttggtcat	gagattatca	aaaaggatct	tcacctagat	ccttttaaag	gccggccgcg	3360
gccgccatcg	gcattttctt	ttgcgttttt	atttgttaac	tgtaattgt	ccttgttcaa	3420
ggatgctgtc	tttgacaaca	gatgttttct	tgcttttgat	gttcagcagg	aagctcggcg	3480
caaacgttga	ttgtttgtct	gcgtagaatc	ctctgtttgt	catatagctt	gtaatcacga	3540
cattgtttcc	tttcgcttga	ggtacagcga	agtgtgagta	agtaaagggt	acatcgttag	3600
gatcaagatc	catttttaac	acaaggccag	ttttgttcag	cggcttgat	gggccagtta	3660
aagaattaga	aacataacca	agcatgtaaa	tatcgttaga	cgtaatgccg	tcaatcgta	3720
ttttgatcc	gcgggagtca	gtgaacaggt	accatttgcc	gttcatttta	aagacgttcg	3780
cggttcaat	ttcatctggt	actgtgtag	atgcaatcag	cggtttcatc	acttttttca	3840
gtgtgtaatc	atcgtttagc	tcaatcatac	cgagagcgcc	gtttgctaac	tcagccgtgc	3900
gttttttatc	gctttgcaga	agtttttgac	tttcttgacg	gaagaatgat	gtgcttttgc	3960
catagtatgc	tttgttaaat	aaagattctt	cgccttggtg	gccatcttca	gttccagttg	4020
ttgcttcaaa	tactaagtat	ttgtggcctt	tatcttctac	gtagtgagga	tctctcagcg	4080
tatggttgtc	gcctgagctg	tagttgcctt	catcgatgaa	ctgctgtaca	ttttgatacg	4140
ttttccgctc	accgtcaaag	attgatttat	aatcctctac	accgttgatg	ttcaaagagc	4200

ES 2 539 280 T3

tgtctgatgc	tgatacgtta	acttgtgcag	ttgtcagtg	ttgtttgccg	taatgtttac	4260
cggagaaatc	agtgtagaat	aaacggattt	ttccgtcaga	tgtaaagtgt	gctgaacctg	4320
accattcttg	tgtttggctt	tttaggatag	aatcatttgc	atcgaatttg	tcgctgtctt	4380
taaagacgcg	gccagcgttt	ttccagctgt	caatagaagt	ttcgccgact	ttttgataga	4440
acatgtaa	cgatgtgtca	tccgcatttt	taggatctcc	ggctaatagca	aagacgatgt	4500
ggtagccgtg	atagtttgcg	acagtgccgt	cagcgttttg	taatggccag	ctgtcccaaa	4560
cgtccaggcc	ttttgcagaa	gagatatttt	taattgtgga	cgaatcaaat	tcagaaaactt	4620
gatatttttc	atftttttgc	tgttcagggg	tttgcagcat	atcatggcgt	gtaatatggg	4680
aaatgccgta	tgtttcctta	tatggctttt	ggttcgtttc	tttcgcaaac	gcttgagttg	4740
cgctcctgc	cagcagtgcg	gtagtaaagg	ttaatactgt	tgcttgtttt	gcaaactttt	4800
tgatgttcat	cgttcatgtc	tcctttttta	tgtactgtgt	tagcggctctg	cttcttccag	4860
ccctcctggt	tgaagatggc	aagttagtta	cgcacaataa	aaaaagacct	aaaatatgta	4920
aggggtgacg	ccaaagtata	cactttgccc	tttacacatt	ttaggtcttg	cctgctttat	4980
cagtaacaaa	cccgcgcgat	ttacttttcg	acctcattct	attagactct	cgtttgatt	5040
gcaactggtc	tattttcctc	ttttgtttga	tagaaaatca	taaaaggatt	tgcagactac	5100
gggcctaaag	aactaaaaaa	tctatctggt	tcttttcatt	ctctgtattt	tttatagttt	5160
ctgttgcatg	ggcataaagt	tgctttttta	atcaccaattc	agaaaatatc	ataatatctc	5220
atctcactaa	ataatagtga	acggcaggta	tatgtgatgg	gttaaaaagg	atcggcggcc	5280
gctcgattta	aat					5293

<210> 16

<211> 906

5 <212> ADN

<213> *Corynebacterium glutamicum*

<400> 16

ES 2 539 280 T3

atgagcacag	gtttaacagc	taagaccgga	gtagagcact	tcggcaccgt	tggagtagca	60
atggttactc	cattcacgga	atccggagac	atcgatatcg	ctgctggccg	cgaagtcgcg	120
gcttatttgg	ttgataaggg	cttggattct	ttggttctcg	cgggcaccac	tggatgaatcc	180
ccaacgacaa	ccgccgctga	aaaactagaa	ctgctcaagg	ccgttcgtga	ggaagttggg	240
gatcgggcca	agctcatcgc	cggtgtcggg	accaacaaca	cgcggacatc	tgtggaactt	300
gcggaagctg	ctgcttctgc	tggcgcagac	ggccttttag	ttgtaactcc	ttattactcc	360
aagccgagcc	aagagggatt	gctggcgcac	ttcggtgcaa	ttgctgcagc	aacagagggt	420
ccaatttgtc	tctatgacat	tcttggtcgg	tcaggtattc	caattgagtc	tgataccatg	480
agacgcctga	gtgaattacc	tacgattttg	gcggtcaagg	acgccaaggg	tgacctcggt	540
gcagccacgt	cattgatcaa	agaaacggga	cttgcctggt	attcaggcga	tgaccacta	600
aaccttgttt	ggcttgcttt	gggcggatca	ggtttcattt	ccgtaattgg	acatgcagcc	660
cccacagcat	tacgtgagtt	gtacacaagc	ttcgaggaag	gcgacctcgt	ccgtgcgcgg	720
gaaatcaacg	ccaaactatc	accgctggta	gctgcccaag	gtcgcctggg	tggagtcagc	780
ttggcaaaag	ctgctctgcg	tctgcagggc	atcaacgtag	gagatcctcg	acttccaatt	840
atggctccaa	atgagcagga	acttgaggct	ctccgagaag	acatgaaaaa	agctggagtt	900
ctataa						906

<210> 17

5 <211> 906

<212> ADN

<213> Artificial

<220>

10 <223> dapA optimizado de *Corynebacterium glutamicum*

<400> 17

ES 2 539 280 T3

atgagcacag ggttaacagc taagacaggg gtagagcact tcgggaccgt tggggtagca 60
atggttactc cattcacgga atccggagac atcgatatcg ctgctggccg cgaagtcgcg 120
gcttatttgg ttgataaggg cttggattct ttggttctcg cgggcaccac tggatgaatcc 180
ccaacgacaa ccgccgctga aaaactagaa ctgctcaagg ccgttcgtga ggaagttggg 240
gatcgggcca agctcatcgc cgggtgctgga accaacaaca cgcggacatc tgtggaactt 300
gcggaagctg ctgcttctgc tggcgcagac ggccttttag ttgtaactcc ttattactcc 360
aagccgagcc aagagggatt gctggcgcac ttcggtgcaa ttgctgcagc aacagaggtt 420
ccaatttgtc tctatgacat tcctggtcgg tcaggtattc caattgagtc tgataccatg 480
agacgcctga gtgaattacc tacgattttg gcggtcaagg acgccaaggg tgacctcgtt 540
gcagccacgt cattgatcaa agaaacggga cttgcctggt attcaggcga tgaccacta 600
aaccttgttt ggcttgcttt gggcggatca ggtttcattt ccgtaattgg acatgcagcc 660
cccacagcat tacgtgagtt gtacacaagc ttcgaggaag gcgacctcgt ccgtgcgcgg 720
gaaatcaacg ccaaactatc accgctggta gctgcccaag gtcgcttggg tggagtcagc 780
ttggcaaaag ctgctctgcg tctgcagggc atcaacgtag gagatcctcg acttccaatt 840
atggctccaa atgagcagga acttgaggct ctccgagaag acatgaaaaa agctggagtt 900
ctataa 906

<210> 18

5 <211> 8438

<212> ADN

<213> Artificial

<220>

10 <223> plásmido pK19 PSOD ask

<400> 18

aattcgcct tcaccgcggc tttggacatc actgctacgt agccaaacaa tgcaccgctc 60
acaagaccaa ggatgagggc tttgtccttc ttttaatacgt attccgcaag cagccacatt 120
ccaccatta ctgcaacgcc gactaaaagt actggaatcc atcgatcgag tgggggtggg 180

15

ggtttccggg	aagggggcgt	cccaaaacga	tcatgatgcc	cacggctacg	gtgaggaggg	240
tagcccagaa	gatttcagtt	cggcgtagtc	ggtagccatt	gaatcgtgct	gagagcggca	300
gcgtgaacat	cagcgacagg	acaagcactg	gttgactac	caagaggggtg	ccgaaaccaa	360
gtgctactgt	ttgtaagaaa	tatgccagca	tcgcggtact	catgcctgcc	caccacatcg	420
gtgtcatcag	agcattgagt	aaaggtgagc	tccttaggga	gccatctttt	gggggtgcgga	480
gcgcatccg	gtgtctgacc	acggtgcccc	atgcgattgt	taatgccgat	gctagggcga	540
aaagcacggc	gagcagattg	ctttgcactt	gattcagggt	agttgactaa	agagttgctc	600
gcgaagtagc	acctgtcact	tttgtctcaa	atattaaatc	gaatatcaat	atatggtctg	660
tttattggaa	cgcgctcccag	tggctgagac	gcatccgcta	aagccccagg	aagggcgaat	720
tctgcagata	tccatcacac	tggcgccgc	tcgagcatgc	atctagctta	tcgccattcg	780
ccattcaggc	tgcgcaactg	ttgggaaggg	cgatcgggtc	gggcctcttc	gctattacgc	840
cagctggcga	aagggggatg	tgctgcaagg	cgattaagtt	gggtaacgcc	agggttttcc	900
cagtcacgac	gttgtaaaac	gacggccagt	gaattccgtg	gcacggaaat	cgaggtagaa	960
gacattactc	aggcaaccga	aagggcgaat	tccgtggcac	ggaaatcgag	gtagaagaca	1020
ttactcaggc	aaccgaaagg	gcgaattcgc	ccttgcgggc	caggattttc	taaaacagga	1080
tgctgcccac	ctttaagcgc	ctcatcagcg	gtaacatca	cgggttcggg	tgcgaaaaac	1140
catgccataa	caggaatggt	cctttcgaaa	attgaggaag	ccttatgccc	ttcaacccta	1200
cttagctgcc	aattattccg	ggcttgtgac	ccgctaccgc	ataaataggt	cggctgaaaa	1260
atctcgttgc	aatataaaca	aaaaggccta	tcattgggag	gtgtcgcacc	aagtactttt	1320
gcgaagcgc	atctgacgga	ttttcaaaag	atgtatatgc	tcggtgcgga	aacctacgaa	1380
aggatTTTTT	accggtggcc	ctggctgtac	agaaatatgg	cggttcctcg	cttgagagtg	1440
cggaacgcat	tagaaacgtc	gctgaacgga	tcggtgccac	caagaaggct	ggaaatgatg	1500
tcgtggttgt	ctgctccgca	atgggagaca	ccacggatga	acttctagaa	cttgcagcgg	1560
cagtgaatcc	cgttccgcca	gctcgtgaaa	tggatatgct	cctgactgct	ggtgagcgta	1620
tttctaacgc	tctcgtcgcc	atggctattg	agtccttgc	gcagaagccc	aatctttcac	1680
gggctctcag	gctggtgtgc	tcaccaccga	gcgccacgga	aacgcacgca	ttgttgatgt	1740
cactccagggt	cgtgtgcgtg	aagcactcga	tgagggcaag	atctgcattg	ttgctggttt	1800
ccagggtggt	aataaagaaa	cccgcgatgt	caccacgttg	ggcgtgggtg	gttctgacac	1860
cactgcagtt	gcgttggcag	ctgctttgaa	cgctgatgtg	tgtgagattt	actcggacgt	1920
tgacgggtgtg	tataccgctg	accgcgcat	cgttccta	gcacagaagc	tgaaaaagct	1980
cagcttcgaa	gaaatgctgg	aacttgctgc	tgttggctcc	aagatTTTgg	tgctgcgcag	2040
tgttgaatac	gctcgtgcat	tcaatgtgcc	acttcgcgta	cgctcgtctt	atagtaatga	2100
tcccggcact	ttgattgccg	gctctatgga	ggatattcct	gtggaagaag	cagtccttac	2160
cgggtgctgca	accgacaagt	ccgaagccaa	agtaaccggt	ctgggtattt	ccgataagcc	2220

aggcgaggct gcgaaggttt tccgtgctgt ggctgatgca gaaatcaaca ttgacatggt 2280
 tctgcagaac gtctcttctg tagaagacgg caccaccgac atcatcttca cctgccctcg 2340
 ttccgacggc cgccgcgcga tggagatctt gaagaagctt caggttcagg gcaactggac 2400
 caatgtgctt tacgacgacc aggtcggcaa agtctccctc gtgggtgctg gcatgaagtc 2460
 tcaccaggt gttaccgcag agttcatgga agctctgctc gatgtcaacg tgaacatcga 2520
 attgatttcc acctctgaga ttcgtatttc cgtgctgatc cgtgaagatg atctggatgc 2580
 tgctgcacgt gcattgcatg agcagttcca gctgggctgc gaagacgaag ccgtcgttta 2640
 tgcaggcacc ggacgctaaa gttttaaagg agtagtttta caatgaccac catcgcagtt 2700
 gttggtgcaa ccggccaggt cggccaggtt atgctcacc ttttgaaga gcgcaatttc 2760
 ccagctgaca ctgttcgttt ctttgcttcc ccacgttccg caggccgtaa gattgaattc 2820
 gccctttcgg ttgcctgagt aatgtcttct acctcgattt ccgtgccacg gaattcgagc 2880
 tcggtaccgc gggatcctct agagtcgacc tgcaggcatg caagcttggc gtaatcatgg 2940
 tcatagctgt ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc 3000
 ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcy 3060
 ttgcgctcac tgcccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 3120
 ggccaacgcg cggggagagg cggtttgcgt attgggcgct cttccgcttc ctcgctcact 3180
 gactcgtgc gctcggctgt tcggctgctg cgagcggat cagctcactc aaaggcggta 3240
 atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 3300
 caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tttccatag gctccgcccc 3360
 cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaacc gacaggacta 3420
 taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg 3480
 ccgcttaccg gatacctgtc cgcctttctc cttcgggaa gcgtggcgtt ttctcatagc 3540
 tcacgctgta ggtatctcag ttcgggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 3600
 gaacccccg ttcagcccga ccgctgctcc ttatccggt actatcgtct tgagtccaac 3660
 ccggtaaagc acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg 3720
 aggtatgtag gcggtgctac agagttcttg aagtgggtgg ctaactacgg ctacactaga 3780
 agaacagtat ttggtatctg cgctctgctg aagccagtta cttcggaaa aagagttggt 3840
 agctcttgat ccggcaaaaa aaccaccgct ggtagcggtg gttttttgt ttgcaagcag 3900
 cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 3960
 gacgctcagt ggaacgaaaa ctacggttaa gggattttgg tcatgagatt atcaaaaagg 4020
 atcttcacct agatcctttt ggggtgggcg aagaactcca gcatgagatc cccgcgctgg 4080
 aggatcatcc agccctgata gaaacagaag cactggagc acctcaaaaa caccatcata 4140
 cactaaatca gtaagtggc agcatcacc gacgcacttt gcgccgaata aatacctgtg 4200
 acggaagatc acttcgcaga ataaataaat cctggtgtcc ctgttgatac cgggaagccc 4260

tgggccaact	tttggcgaaa	atgagacggt	gatcggcagc	taagaggttc	caactttcac	4320
cataatgaaa	taagatcact	accgggcgta	ttttttgagt	tatcgagatt	ttcaggagct	4380
gatagaaaca	gaagccactg	gagcacctca	aaaacacccat	catacactaa	atcagtaagt	4440
tggcagcatc	acccgacgca	ctttgcgccg	aataaatacc	tgtgacggaa	gatcacttcg	4500
cagaataaat	aaatcctggt	gtccctgttg	ataccgggaa	gccctgggcc	aacttttggc	4560
gaaaatgaga	cgttgatcgg	cacgtaagag	gttccaactt	tcaccataat	gaaataagat	4620
cactaccggg	cgtatTTTTT	gagttatcga	gattttcagg	agctctttgg	catcgtctct	4680
cgctgtccc	ctcagttcag	taatttcctg	catttgcctg	tttccagtcg	gtagatattc	4740
cacaaaacag	caggggaagca	gcgcttttcc	gctgcataac	cctgcttcgg	ggtcattata	4800
gcgatTTTTT	cggtatatcc	atcctTTTTT	gcacgatata	caggatTTTg	ccaaaggggt	4860
cggtgtagact	ttccttggtg	tatccaacgg	cgtcagccgg	gcaggatagg	tgaagtaggc	4920
ccaccgcga	gcggggtgtc	cttcttcaact	gtcccttatt	cgcacctggc	ggtgctcaac	4980
gggaatcctg	ctctgcgagg	ctggccggct	accgccggcg	taacagatga	gggcaagcgg	5040
atggctgatg	aaaccaagcc	aaccaggaag	ggcagcccac	ctatcaaggt	gtactgcctt	5100
ccagacgaac	gaagagcgat	tgaggaaaag	gcggcggcgg	ccggcatgag	cctgtcggcc	5160
tacctgctgg	ccgtcggcca	gggctacaaa	atcacgggcg	tcgtggacta	tgagcacgtc	5220
cgcgagggcg	tcccggaaaa	cgattccgaa	gcccaacctt	tcatagaagg	cggcggtgga	5280
atcgaaatct	cgtgatggca	ggttgggcgt	cgcttggtcg	gtcatttcgc	tcggtacca	5340
tcggcatttt	cttttgcggt	taactgttaa	ttgtccttgt	tcaaggatgc	tgtctttgac	5400
aacagatggt	ttcttgcctt	tgatgttcag	caggaagctc	ggcgcaaacg	ttgattgttt	5460
gtctgcgtag	aatcctctgt	ttgtcatata	gcttgtaatc	acgacattgt	ttcctttcgc	5520
ttgaggtaca	gcgaagtgtg	agtaagtaaa	ggttacatcg	ttaggatcaa	gatccatttt	5580
taacacaagg	ccagttttgt	tcagcggctt	gtatgggcca	gttaaagaat	tagaaacata	5640
accaagcatg	taaataatcgt	tagacgtaat	gccgtcaatc	gtcatttttg	atccgcggga	5700
gtcagtgaac	aggtaccatt	tgccgttcat	tttaaagacg	ttcgcgcggt	caatttcac	5760
tgttactgtg	ttagatgcaa	tcagcggttt	catcactttt	ttcagtgtgt	aatcatcgtt	5820
tagctcaatc	ataccgagag	cgccgtttgc	taactcagcc	gtgcgTTTTT	tatcgctttg	5880
cagaagtttt	tgactttctt	gacggaagaa	tgatgtgctt	ttgccatagt	atgctttggt	5940
aaataaagat	tcttcgcctt	ggtagccatc	ttcagttcca	gtgtttgctt	caaatactaa	6000
gtattttgtg	cctttatctt	ctacgtagtg	aggatctctc	agcgtatgg	tgtcgcctga	6060
gctgtagttg	ccttcacgca	tgaactgctg	tacattttga	tacgtttttc	cgtcaccgtc	6120
aaagattgat	ttataatcct	ctacaccggt	gatgttcaaa	gagctgtctg	atgctgatac	6180
gttaacttgt	gcagttgtca	gtgtttgttt	gccgtaatgt	ttaccggaga	aatcagtgta	6240
gaataaacgg	atTTTTccgt	cagatgtaaa	tgtggctgaa	cctgaccatt	cttgtgtttg	6300

gtcttttagg	atagaatcat	ttgcatcgaa	tttgtcgcctg	tctttaaaga	cgcgccagc	6360
gtttttccag	ctgtcaatag	aagtttcgcc	gactttttga	tagaacatgt	aatcgcgatg	6420
gtcatccgca	tttttaggat	ctccggctaa	tgcaaagacg	atgtggtagc	cgatgatagt	6480
tgcgacagtg	ccgtcagcgt	tttgtaatgg	ccagctgtcc	caaacgtcca	ggccttttgc	6540
agaagagata	tttttaattg	tggacgaatc	aaattcagga	acttgatatt	tttcattttt	6600
ttgctgttca	gggatttgca	gcataatcatg	gcgtgtaata	tgggaaatgc	cgatggtttc	6660
cttatatggc	ttttggttcg	tttctttcgc	aaacgcttga	gttgcgcctc	ctgccagcag	6720
tgcggtagta	aaggttaata	ctgttgcttg	ttttgcaaac	tttttgatgt	tcacgtttca	6780
tgtctccttt	tttatgtact	gtgtagcgg	tctgcttctt	ccagccctcc	tgtttgaaga	6840
tggcaagtta	gttacgcaca	ataaaaaaag	acctaaaata	tgtaaggggt	gacgccaag	6900
tatacacttt	gccctttaca	catttttaggt	cttgcctgct	ttatcagtaa	caaaccgcg	6960
cgatttactt	ttcgacctca	ttctattaga	ctctcgtttg	gattgcaact	ggtctatttt	7020
cctcttttgt	ttgatagaaa	atcataaaaag	gatttgcaga	ctacgggcct	aaagaactaa	7080
aaaatctatc	tgtttctttt	cattctctgt	attttttata	gtttctgttg	catgggcata	7140
aagttgcctt	tttaatcaca	attcagaaaa	tatcataata	tctcatttca	ctaaataata	7200
gtgaacggca	ggtatatgtg	atgggttaaa	aaggatcgat	cctctagcga	accccagagt	7260
cccgcctcaga	agaactcgtc	aagaaggcga	tagaaggcga	tgcgctgcga	atcgggagcg	7320
gcgataccgt	aaagcacgag	gaagcggcca	gcccattcgc	cgccaagctc	ttcagcaata	7380
tcacgggtag	ccaacgctat	gtcctgatag	cggtccgcca	caccagccg	gccacagtcg	7440
atgaatccag	aaaagcggcc	attttccacc	atgatattcg	gcaagcaggc	atcgccatgg	7500
gtcacgacga	gatcctcgcc	gtcgggcattc	cgcgcttga	gcctggcgaa	cagttcggct	7560
ggcgcgagcc	cctgatgctc	ttcgtccaga	tcacctgat	cgacaagacc	ggcttccatc	7620
cgagtacgtg	ctcgtctgat	gcgatgtttc	gcttgggtgg	cgaatgggca	ggtagccgga	7680
tcaagcgtat	gcagccgccc	cattgcatca	gccatgatgg	atactttctc	ggcaggagca	7740
aggtgagatg	acaggagatc	ctgccccggc	acttcgccc	atagcagcca	gtcccttccc	7800
gcttcagtga	caacgtcgag	cacagctgcg	caaggaacgc	ccgtcgtggc	cagccacgat	7860
agccgcgctg	cctcgtcttg	gagttcattc	agggcaccgg	acaggtcggg	cttgacaaaa	7920
agaaccgggc	gcccctgccc	tgacagccgg	aacacggcgg	catcagagca	gccgattgtc	7980
tgttgtgccc	agtcatagcc	gaatagcctc	tccaccaag	cggccggaga	acctgcgtgc	8040
aatccatctt	gttcaatcat	gcgaaacgat	cctcatcctg	tctcttgatc	agatcttgat	8100
cccctgccc	atcagatcct	tggcggcaag	aaagccatcc	agtttacttt	gcagggcttc	8160
ccaaccttac	cagagggcgc	cccagctggc	aattccggtt	cgcttgctgt	ccataaaaacc	8220
gcccagtcta	gctatcgcca	tgtaagccca	ctgcaagcta	cctgctttct	ctttgcgctt	8280
gcgttttccc	ttgtccagat	agcccagtag	ctgacattca	tccgggggtca	gcaccgtttc	8340

ES 2 539 280 T3

tgccgactgg ctttctacgt gttccgcttc ctttagcagc ctttgcgccc tgagtgcttg 8400
cggcagcgtg aagctagtaa cggccgcccag tgtgctgg 8438

<210> 19

<211> 5685

5 <212> ADN

<213> Artificial

<220>

<223> plásmido pCLIK Int Psod Glucosa-6-PDH

10

<400> 19

ES 2 539 280 T3

cgcgtttgcc caaatagtgg tcgatgcgga acacagaaga ttctgggaag actgcgttga 60
 ccagctgggt gagctcgtgt gcggattcga ggttgtggcc gaaaggcttc tcgatgatca 120
 cgcggcgcca tgcttcttcg gtggattcag ccatgccgga acgctccagc tgggtggcaga 180
 ccgctgtgaa ggaatctggt ggaatggaca ggtagtaagc ccagttgccg gcggtgccgc 240
 gggttttgtc gatgcgcttg agtgttgcag cgaggttgtc gaaagctgca tcatcatcaa 300
 agttgccgcg aacaaattcc ataccctcgg cgaggcgtc ccaaacttt tcacggaatt 360
 ccgtacgagc accagcactt gcggcatcgc gtacgtatth ttcaaagtct tctttggacc 420
 attcgcggcg gccgtaacct accaacgaga atcctggggg cagcaatccg cggtttgcta 480
 gatcataaat ggcgggggagc agcttctttc gagccaagtc gccagtgaca ccgaagatca 540
 ccatgccgga agggccagcg atgcggggga gtcgtttatc ctgcgggtcg cgcagtgggg 600
 ttgtccagct ggaggggggtc gtgtttgtgc tcacgggtaa aaaatccttt cgtaggtttc 660
 cgcaccgagc atatacatct ttgaaaatc cgtcagatgg cgcttcgcaa aagtacttgg 720
 tgcgacacct cccaatgata ggcctttttg ttgatattgc aacgaaattt ttcagccgac 780
 ctatttatcg ggtagcgggt cacaagcccg gaataattgg cagctagatg gtagtgtcac 840
 gatcctttct ttaatgaaag atgtgtaacg gccacataag atcgaactaa ttcgatttca 900
 tgtcgcggtt actgatgagc cgtgctgatt ctacttcaga cgagcttcca tggactcaag 960
 cagttcgctc caagaagcaa cgaacttgtc cacaccctcg gtctccagga cctggaagac 1020
 atctgccaag tcaacgcca gagcctcaag ctggggagaac acagcgtcag cttctgccgc 1080
 ggagttggac aggggtgtcac cgtgcagggt gccctgtccc agaaccgctc cgatggtgcc 1140
 ttctggcatg gtgttgacgg tgtttgacc agccagctcg gaaacgtaa gagttgcagc 1200
 gtacgcaggg ttcttcacgc cgggtgatgc ccacagtggt cgctgagtgt tggcaccttc 1260
 aggcagctcg gcggcgtcga aaagctcctt gtacacagcg taagcgcgct gagcgttggc 1320
 aacgcctgcc ttgccgcgca gagccaaagc ctcatcggat ccgattgcct cgaggcgtt 1380
 gtcgatctca acgtcgacat cgatgctctt ctgcgttaat taacaattgg gatcctctag 1440
 acccgggatt taaatgatcc gctagcgggc tgctaaagga agcggaacac gtagaaagcc 1500
 agtccgcaga aacggtgctg accccggatg aatgtcagct actgggctat ctggacaagg 1560
 gaaaacgcaa gcgcaaagag aaagcaggta gcttgcaagt ggcttacatg gcgatagcta 1620

gactggg	cggtttat	gacagca	agcaagc	gaaacg	ccggaatt	gcagctg	ggggc	gccctct	gggt	1680	
aaggtt	gggaag	ccctg	caagtaa	actg	atggctt	tctt	tgccg	ccaag	gatctg	atgg	1740
cgaggg	gatca	agatct	gata	tcaag	agaca	ggatg	aggat	cgtttc	gcat	gattg	1800
gatgg	attgc	acgcag	gttc	tccgg	ccgct	tgggt	ggaga	ggctatt	ccg	ctatg	1860
gcaca	acaga	caatcg	gctg	ctctg	atgcc	gccgt	gttc	ggctg	tcagc	gcaggg	1920
ccggtt	cttt	ttgtca	agac	cgacct	gtcc	ggtgc	ccctg	atgaact	gca	ggacg	1980
gcg	cggtat	cgtgg	ctgg	cacgac	ggc	gttc	cttgc	cagctg	tgct	cgacgt	2040
actga	agcgg	gaaggg	actg	gctg	ctattg	ggcga	agtgc	cgggg	cagga	tctc	2100
tctca	cttg	ctcctg	ccga	gaaagt	atcc	atcat	ggctg	atgca	atgcg	gcgg	2160
acgctt	gatc	cggtac	ctg	cccatt	cgac	cacca	agcga	aacat	cgcat	cgagc	2220
cgta	ctc	ggaag	ccgg	tcttg	tcgat	caggat	gatc	tggac	gaaga	gcatc	2280
ctcgc	ccag	ccgaact	gtt	cgccag	gctc	aaggc	gcga	tgccc	gacgg	cgagg	2340
gtcgt	gacc	atggc	gatgc	ctgctt	gccg	aatat	catgg	tggaa	aatgg	ccgctt	2400
ggatt	catc	actgt	ggccg	gctgg	gtgtg	gcgg	accg	ct	cagg	acat	2460
acc	gtg	ata	ttgt	gaaga	gcttg	gcgg	c	gaat	ggctg	accg	2520
gg	tatc	gccg	ctcc	gattc	gcagc	gcac	tc	gcctt	cttga	cgagt	2580
tgag	cg	ggg	gac	tctgg	gggtt	gaaat	gaccg	acca	agc	gac	2640
attt	cg	attc	cacc	gccc	ttct	atg	aaa	ggtt	ggc	tt	2700
ccg	gt	ggat	gatc	ctcc	ag	cg	ggg	gatc	tcat	gctg	2760
gcgc	g	ccgg	c	gg	ctg	tg	aaat	accg	cac	agat	2820
agg	cg	ctct	ccg	ctt	ctc	gctc	actg	tcg	ctg	cgct	2880
gcg	gt	atc	ag	ggc	g	taata	cggt	tat	cca	caga	2940
ggaa	aga	aaca	tgtg	ag	ca	aaa	agg	ccag	ga	accg	3000
ctgg	cg	tttt	tccat	agg	ct	cccc	ct	gacg	atc	acaaa	3060
cag	agg	tg	gac	cccg	ac	agg	act	ataa	agat	acc	3120
ctc	gt	gc	gct	ctt	cc	gacct	g	ctt	acc	ggat	3180
tcg	ga	ag	cg	ctt	t	tcat	ag	ctca	cgct	gtag	3240
gtt	cg	ct	cca	ag	ctg	ggg	ctt	ag	ccg	accg	3300
tcc	g	ta	act	at	ct	ttg	ga	gtcc	a	accg	3360
gcc	act	gg	ta	ac	ag	gatt	ag	cag	ag	gagg	3420
tggt	gg	cc	ta	act	ac	gg	ct	ta	ag	g	3480
ccag	tt	acct	tcg	g	aaa	ag	ag	ttg	g	tag	3540
agc	g	gtg	gtt	tttt	gtt	caag	cag	ca	g	gcga	3600
gatc	ctt	g	ta	tctt	ctac	ggg	gt	ctg	ga	ac	3660

attttggTca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ggccggccgc 3720
 ggccgccatc ggcattttct tttgcgtttt tatttgtaa ctgtaattg tccttgttca 3780
 aggatgctgt ctttgacaac agatgttttc ttgcctttga tgttcagcag gaagctcggc 3840
 gcaaacgTtg attgTttgtc tgcgtagaat cctctgtttg tcatatagct tgtaatcacg 3900
 acattgtttc ctttcgcttg aggtacagcg aagtgtgagt aagtaaaggT tacatcgTta 3960
 ggatcaagat ccatttttaa cacaaggcca gttttgttca gcggtttgta tgggccagtt 4020
 aaagaattag aaacataacc aagcatgtaa atatcgTtag acgtaatgcc gtcaatcgTc 4080
 atttttgatc cgcgggagtc agtgaacagg taccatttgc cgTtcatttt aaagacgTtc 4140
 gcgcgTtcaa tttcatctgt tactgtgtta gatgcaatca gcggtttcat cacttttttc 4200
 agtgtgtaat catcgTttag ctcaatcata ccgagagcgc cgTttgctaa ctCagccgtg 4260
 cgTtttttat cgctttgcag aagTttttga ctttcttgac ggaagaatga tgtgcttttg 4320
 ccatagtatg ctttgTtaaa taaagattct tcgccttggT agccatcttc agTtccagtg 4380
 tttgcttcaa atactaagta tttgtggcct ttatcttcta cgtagtgagg atctctcagc 4440
 gtatggtTgt cgctgagct gtagttgcct tcatcgatga actgctgtac attttgatac 4500
 gtttttccgt caccgtcaaa gattgattta taatcttcta caccgttgat gttcaaagag 4560
 ctgtctgatg ctgatacgTt aactgtgca gttgtcagtg tttgtttgcc gtaatgttta 4620
 ccggagaaat cagtgtagaa taaacggatt tttccgTcag atgtaaagtG ggctgaacct 4680
 gaccattctt gtgtttggTc ttttaggata gaatcattg catcgaattt gTcgtgtct 4740
 ttaaagacgc ggccagcgTt tttccagctg tcaatagaag tttcgccgac tttttgatag 4800
 aacatgtaaa tcgatgtgtc atccgcattt ttaggatctc cggctaatgc aaagacgatg 4860
 tggtagccgt gatagTttgc gacagtgccg tcagcgTttt gtaatggcca gctgtcccaa 4920
 acgtccaggc cttttgcaga agagatattt ttaattgtgg acgaatcaaa ttcagaaact 4980
 tgatattttt catttttttg ctgttcaggg atttgcagca tatcatggcg tgtaatatgg 5040
 gaaatgccgt atgtttcctt atatggcttt tggTtcgttt ctttcgcaaa cgcttgagtt 5100
 gcgcctcctg ccagcagTgc ggtagtaaag gTtaatactg ttgcttgttt tgcaaaactt 5160
 ttgatgttca tcgttcatgt ctCctttttt atgtactgtg ttagcggTct gcttcttcca 5220
 gccctcctgt ttgaagatgg caagttagtt acgcacaata aaaaagacc taaaatatgt 5280
 aaggggtgac gccaaagtat acactttgcc ctttacacat tttaggTctt gcctgcttta 5340
 tcagtaacaa acccgcgcga tttactttc gacctattc tattagactc tcgtttggat 5400
 tgcaactggt ctattttcct cttttgtttg atagaaaatc ataaaaggat ttgcagacta 5460
 cgggcctaaa gaactaaaaa atctatctgt ttcttttcat tctctgtatt ttttatagtt 5520
 tctgttgcac gggcataaag ttgcctttt aatcacaatt cagaaaatat cataatatct 5580
 catttcacta aataatagtg aacggcaggt atatgtgatg ggttaaaaag gatcggcggc 5640
 cgctcgattt aaatctcgag aggcctgacg tcgggcccgg tacca 5685

<210> 20

<211> 7632

<212> ADN

<213> Artificial

5

<220>

<223> plásmido pK19 sacB Glu-6-P-DH

<400> 20

10

cgcccttccc aacagttgcg cagcctgaat ggcgaatggc gataagctag cttcacgctg 60
 ccgcaagcac tcagggcgca agggctgcta aaggaagcgg aacacgtaga aagccagtcc 120
 gcagaaacgg tgctgacccc ggatgaatgt cagctactgg gctatctgga caagggaaaa 180
 cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat agctagactg 240
 ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgccct ctggtaaggt 300
 tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct gatggcgag 360
 gggatcaaga tctgatcaag agacaggatg aggatcgttt cgcatgattg aacaagatgg 420
 attgcacgca ggttctccgg ccgcttgggt ggagaggcta ttcggctatg actgggcaca 480
 acagacaatc ggctgctctg atgccgccgt gttccggctg tcagcgcagg ggcgcccggt 540
 tctttttgtc aagaccgacc tgtccggtgc cctgaatgaa ctccaagacg aggcagcgcg 600
 gctatctggt ctggccacga cgggcgttcc ttgvcagct gtgctcgacg ttgtcactga 660
 agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc tgtcatctca 720
 ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcggc tgcatacgtt 780
 tgatccggct acctgcccac tcgaccacca agcgaaacat cgcacgcagc gagcacgtac 840
 tcggatggaa gccgggtctt tcgatcagga tgatctggac gaagagcatc aggggctcgc 900
 gccagccgaa ctgttcgcca ggctcaaggc gcggatgccc gacggcgagg atctcgtcgt 960
 gacctatggc gatgcctgct tgccgaatat catggtggaa aatggccgct tttctggatt 1020
 catcgactgt ggccggctgg gtgtggcggg ccgctatcag gacatagcgt tggctacccg 1080
 tgatattgct gaagagcttg gcggcgaatg ggctgaccgc ttcctcgtgc tttacggtat 1140
 cgccgctccc gattcgcagc gcacgcctt ctatgcctt cttgacgagt tcttctgagc 1200
 gggactctgg ggttcgctag aggatcgatc ctttttaacc catcacatat acctgccggt 1260
 cactattatt tagtgaaatg agatattatg atattttctg aattgtgatt aaaaaggcaa 1320
 ctttatgccc atgcaacaga aactataaaa aatacagaga atgaaaagaa acagatagat 1380
 ttttagttc tttaggcccg tagtctgcaa atccttttat gattttctat caaacaaaag 1440
 aggaaaatag accagttgca atccaaacga gagtctaata gaatgaggtc gaaaagtaaa 1500
 tcgvcgggt ttgttactga taaagcaggc aagacctaaa atgtgtaaag ggcaaagtgt 1560
 atactttggc gtcaccctt acatatttta ggtctttttt tattgtgcgt aactaacttg 1620
 ccatcttcaa acaggagggc tggaagaagc agaccgctaa cacagtacat aaaaaggag 1680
 acatgaacga tgaacatcaa aaagtttgca aaacaagcaa cagtattaac ctttactacc 1740

gcactgctgg caggaggcgc aactcaagcg tttgcgaaag aaacgaacca aaagccatat 1800
aaggaaacat acggcatttc ccatattaca cgccatgata tgctgcaaat ccctgaacag 1860
caaaaaaatg aaaaatatca agtttctgaa tttgattcgt ccacaattaa aaatatctct 1920
tctgcaaaag gcctggacgt ttgggacagc tggccattac aaaacgctga cggcactgtc 1980
gcaaactatc acggctacca catcgtcttt gcattagccg gagatcctaa aaatgcggat 2040
gacacatcga tttacatggt ctatcaaaaa gtcggcgaaa cttctattga cagctggaaa 2100
aacgctggcc gcgtctttaa agacagcgc aaattcgatg caaatgattc taccctaaaa 2160
gaccaaacc acgaatggtc aggttcagcc acatttacat ctgacggaaa aatccgttta 2220
ttctacactg atttctccg taaacattac ggcaaaaa cactgacaac tgcacaagtt 2280
aacgtatcag catcagacag ctctttgaac atcaacgggtg tagaggatta taaatcaatc 2340
tttgacgggtg acggaaaaac gtatcaaaat gtacagcagt tcatcgatga aggcaactac 2400
agctcaggcg acaaccatac gctgagagat cctcactacg tagaagataa aggccacaaa 2460
tacttagtat ttgaagcaaa cactggaact gaagatggct accaaggcga agaattctta 2520
tttaacaaag catactatgg caaaagcaca tcattcttcc gtcaagaaag tcaaaaactt 2580
ctgcaaagcg ataaaaaacg cacggctgag ttagcaaacg gcgctctcgg tatgattgag 2640
ctaaacgatg attacacact gaaaaaagtg atgaaaccgc tgattgcatc taacacagta 2700
acagatgaaa ttgaacgcgc gaacgtcttt aaaatgaacg gcaaatggta cctgttctact 2760
gactcccgcg gatcaaaaat gacgattgac ggcattacgt ctaacgatat ttacatgctt 2820
ggttatgttt ctaattcttt aactggccca tacaagccgc tgaacaaaac tggccttggtg 2880
ttaaaaatgg atcttgatcc taacgatgta acctttactt actcacactt cgctgtacct 2940
caagcgaaag gaaacaatgt cgtgattaca agctatatga caaacagagg attctacgca 3000
gacaaacaat caacgtttgc gccgagcttc ctgctgaaca tcaaaggcaa gaaaacatct 3060
gttgctcaag acagcatcct tgaacaagga caattaacag ttaacaaata aaaacgcaaa 3120
agaaaatgcc gatgggtacc gagcgaaatg accgaccaag cgacgcccaa cctgccatca 3180
cgagatttcg attccaccgc cgcttctat gaaaggttg gcttcggaat cgttttccgg 3240
gacgccctcg cggacgtgct catagtccac gacgcccgtg atttttagc cctggccgac 3300
ggccagcagg taggccgaca ggctcatgcc ggccgcccgc gccttttct caatcgctct 3360
tcgttcgtct ggaaggcagt acaccttgat aggtgggctg cccttcctgg ttggcttggt 3420
ttcatcagcc atccgcttg cctcatctgt tacgccggcg gtagccggcc agcctcgag 3480
agcaggattc ccgttgagca ccgccaggtg cgaataaggg acagtgaaga aggaacaccc 3540
gctcgcgggt gggcctactt cacctatcct gcccggctga cgccgttgga tacaccaagg 3600
aaagtctaca cgaacccttt ggcaaaatcc tgtatatcgt gcgaaaaagg atggatatac 3660
cgaaaaaatc gctataatga ccccgaagca gggttatgca gcggaaaagc gctgcttccc 3720
tgctgttttg tgaatatct accgactgga aacaggcaaa tgcaggaaat tactgaactg 3780

aggggacagg	cgagagacga	tgccaaagag	ctcctgaaaa	tctcgataac	tcaaaaaata	3840
cgcccggtag	tgatcttatt	tcattatggt	gaaagttgga	acctcttacg	tgccgatcaa	3900
cgtctcattt	tcgcaaaaag	ttggcccagg	gcttcccggg	atcaacaggg	acaccaggat	3960
ttattttatt	tgcaagtga	tcttccgtca	caggatatta	ttcggcgcaa	agtgcgtcgg	4020
gtgatgctgc	caacttactg	atthtagtga	tgatgggtgt	tttgagggtgc	tccagtggct	4080
tctgtttcta	tcagctcctg	aaaatctcga	taactcaaaa	aatacgcccc	gtagtgatct	4140
tatttcatta	tggtgaaagt	tggaacctct	tacgtgccga	tcaacgtctc	atthtcgcca	4200
aaagttggcc	cagggcttcc	cggtatcaac	agggacacca	ggatttattt	attctgcgaa	4260
gtgatcttcc	gtcacaggta	tttattcggc	gcaaagtgcg	tcgggtgatg	ctgccaactt	4320
actgatttag	tgtatgatgg	tgthtttgag	gtgctccagt	ggcttctggt	tctatcaggg	4380
ctggatgatc	ctccagcgcg	gggatctcat	gctggagttc	ttcggcccacc	ccaaaaggat	4440
ctaggtgaag	atcctthttg	ataatctcat	gaccaaatac	ccttaacgtg	agthttcggt	4500
ccactgagcg	tcagaccccc	tagaaaagat	caaaggatct	tcttgagatc	ctthttttct	4560
gcgcgtaatc	tgctgcttgc	aaacaaaaaa	accaccgcta	ccagcgggtgg	ttgttttgcc	4620
ggatcaagag	ctaccaactc	thtttccgaa	ggtaactggc	ttcagcagag	cgcagatacc	4680
aaatactgth	cttctagtgt	agccgtagtt	agccaccac	ttcaagaact	ctgtagcacc	4740
gcctacatac	ctcgtctctg	taatctggt	accagtggtc	gctgccagtg	gcgataagtc	4800
gtgtcttacc	gggttgact	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	4860
aacggggggt	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	aactgagata	4920
cctacagcgt	gagctatgag	aaagcggcc	gcttcccga	gggagaaagg	cggacaggta	4980
tccggtaagc	ggcagggctg	gaacaggaga	gcgcacgagg	gagcttccag	gggaaacgc	5040
ctggatctt	tatagctctg	tcgggtttcg	ccacctctga	cttgagcgtc	gattthttgtg	5100
atgctcgtca	ggggggcgga	gcctatggaa	aaacgccagc	aacgcggcct	thttacgggt	5160
cctggcctth	tgctggcctt	ttgctcacat	gttctthcct	gcgttatccc	ctgattctgt	5220
ggataaccgt	attaccgcct	ttgagtgagc	tgataccgct	cgccgcagcc	gaacgaccga	5280
gcgcagcag	tcagtgagcg	aggaagcgg	agagcgccca	atacgcaaac	cgctctccc	5340
cgcgcgttgg	ccgattcatt	aatgcagctg	gcacgacagg	thtcccgact	ggaaagcggg	5400
cagtgagcgc	aacgcaatta	atgtgagtta	gctcactcat	taggcacccc	aggctthtaca	5460
ctthtatgctt	ccggctcgta	tgthgtgtgg	aattgtgagc	ggataacaat	thcacacagg	5520
aaacagctat	gacctgatt	acgccaagct	tgaaatttgc	tggtgtgtgg	tatccggaag	5580
thcaaagatc	atthtttgcc	cctaaattat	ggcctgcgcc	aggtgtgacc	gttgcgggaa	5640
agcatttcat	cagcgtctt	tggacccac	gtaccgctg	ggtaatcctc	tggttctcca	5700
tcggcatccc	atgcttcaag	aattggatcc	agaatcttcc	agctcagttc	cacttctctg	5760
thggtaggg	agaggctgga	thcatctaac	agcgcattca	aaatgaggcg	ctcgtatgct	5820

<210> 21

```

tcaggtgatt cttcagtgaa ggattctgag taggagaagt ccatgttgac gtcacggact 5880
tccatggcag aacctggaac cttggaaccg aagcggatga gcacaccttc atcaggctgc 5940
acgcgaatca cgatggcggt ttggccaagg gatacagtca tgtcgccgtc gaaaggctgg 6000
tgtggtgctg ctttaaacac cacggcaatc tcagtaacac ggcgaccaag acgcttaccg 6060
gtgcgcaggt agaacggcac accagcccag cgacgagacg tgatctctaa ggtacaagcc 6120
gcaaaagtct cagtggtgga ctcagggttg aagccatctt cttcgcgaag tcccttgact 6180
aactcagagc cctgccaaacc ggcagcgtac tgaccacgag cggagggttt atccaatggg 6240
tagcacggct ttgtcgcaga gaggaccttg atcttttctg cctgcagctg cgctggcacg 6300
aaagaaattg gttcttccat ggcaaccaga gccaagagct ggatcaggtg gttctggatg 6360
acgtcgcggg ctgcccgat gccgtcgtag taaccagcac gtccaccaa gccaatatct 6420
tcagtcatgg tgatctggac gtggtcaacg tagttggagt tccacagtgg ctcaaacagc 6480
tggttagcaa aacgcagagc caggatgttt tgaactgttt ccttgcccaa atagtggctg 6540
atgcggaaca cagaagattc tgggaagact gcggtgacca gctggttgag ctctgtgctg 6600
gattcaggtt tgtggccgaa aggcttctcg atgatcacgc ggcgccatgc ttcttcggtg 6660
gattcagcca tgccggaacg ctccagctgg tggcagaccg ctgtgaagga atctggtgga 6720
atggacaggt agtaagccca gttgccggcg gtgccgcggg ttttgtcgat gcgcttgagt 6780
gttgcagcga ggttgtcgaa agctgcatca tcatcaaagt tgccgcgaac aaattccata 6840
ccctcggcga ggcgctccca aacattttca cgggaattccg tacgagcacc agcacttgcg 6900
gcatcgcgta cgtatTTTTT aaagtcttct ttggaccatt cgcggcggcc gtaacctacc 6960
aacgagaatc ctgggggag caatccgcgg tttgctagat cataaatggc ggggagcagc 7020
ttcttctcag ccaagtcgcc agtgacaccg aagatcacca tgccggaagg gccagcgatg 7080
cgggggagtc gtttatcctg cgggtcgcgc agtgggtttg tccagctgga gggggtcgtg 7140
tttgtgctca cgatggtagt gtcacgatcc tttctttaat gaaagatgtg taacggccac 7200
ataagatcga actaattcga tttcatgtcg ccgttactga tgcagcgtgc tgattctact 7260
tcagacgagc ttccatggac tcaagcagtt cgctccaaga agcaacgaac ttgtccacac 7320
cctcggctct caggacctgg aagacatctg ccaagtcaac gccagagacc tcaagctggg 7380
agaacacagc gtcagcttct gccgcggagt tggacagggg gtcaccgtgc aggttgccct 7440
gctccagaac cgcgtcgatg gtgccttctg gcatggtggt gacggtgttt ggaccagcca 7500
gctcggaaac gtaaagagtt gcagcgtacg cagggttctt cacgccggtg gatgccaca 7560
gtgggcgctg agtgttggca cttcaggca gctcggcggc gtcgaaaagc tccttgata 7620
caggaattcc gg 7632

```

<210> 21

<211> 7438

5 <212> ADN

<213> Artificial

<220>

<223> plásmido pK19 PGro PycA

5

<400> 21

ttatttgta	actgtaatt	gtccttgttc	aaggatgctg	tctttgacaa	cagatgtttt	60
cttgcccttg	atgttcagca	ggaagctcgg	cgcaaacggt	gattgtttgt	ctgcgtagaa	120
tcctctgttt	gtcatatagc	ttgtaatcac	gacattgttt	cctttcgctt	gaggtacagc	180
gaagtgtgag	taagtaaagg	ttacatcggt	aggatcaaga	tccattttta	acacaaggcc	240
agttttgttc	agcggcttgt	atgggccagt	taaagaatta	gaaacataac	caagcatgta	300
aatatcgta	gacgtaatgc	cgtaaatcgt	catttttgat	ccgcgggagt	cagtgaacag	360
gtaccatttg	ccgttcattt	taaagacggt	cgcgcgttca	atttcatctg	ttactgtgtt	420
agatgcaatc	agcggtttca	tcactttttt	cagtgtgtaa	tcatcgttta	gctcaatcat	480
accgagagcg	ccgtttgcta	actcagccgt	gcgtttttta	tcgctttgca	gaagtttttg	540
actttcttga	cggaagaatg	atgtgctttt	gccatagtat	gctttgttaa	ataaagattc	600
ttcgcccttg	tagccatctt	cagttccagt	gtttgcttca	aataactaagt	atttgtggcc	660
tttatcttct	acgtagtgag	gatctctcag	cgatagggtg	tcgcctgagc	tgtagttgcc	720
ttcatcgatg	aactgctgta	cattttgata	cgtttttccg	tcaccgtcaa	agattgattt	780
ataatcctct	acaccgttga	tgttcaaaga	gctgtctgat	gctgatacgt	taacttgtgc	840
agttgtcagt	gtttgtttgc	cgtaatgttt	accggagaaa	tcagtgtaga	ataaacggat	900
ttttccgtca	gatgtaaata	tggctgaacc	tgaccattct	tgtgtttggt	cttttaggat	960
agaatcattt	gcatcgaatt	tgtcgtgttc	tttaaagacg	cgccagcgt	ttttccagct	1020
gtcaatagaa	gtttcgccga	ctttttgata	gaacatgtaa	atcgatgtgt	catccgcatt	1080
tttaggatct	ccggctaata	caaagacgat	gtggtagccg	tgatagtttg	cgacagtgcc	1140
gtcagcgttt	tgtaatggcc	agctgtccca	aacgtccagg	ccttttgtag	aagagatatt	1200
tttaattgtg	gacgaatcaa	attcaggaac	ttgatatttt	tcattttttt	gctgttcagg	1260
gatttgcagc	atatcatggc	gtgtaatatg	ggaaatgccg	tatgtttcct	tatatggctt	1320
ttggttcggt	tctttcgcaa	acgcttgagt	tcgcctctct	gccagcagtg	cggtagtaaa	1380
ggtaataact	gttgcttgtt	ttgcaaactt	ttgatgttcc	atcgttcatg	tctccttttt	1440
tatgtactgt	gttagcggtc	tgcttcttcc	agccctctct	tttgaagatg	gcaagttagt	1500
tacgcacaat	aaaaaaagac	ctaaaatatg	taaggggtga	cgccaaagta	tacactttgc	1560
cctttacaca	ttttaggtct	tgcttcttct	atcagtaaca	aacccgcgcg	atttactttt	1620
cgacctcatt	ctattagact	ctcgtttgga	ttgcaactgg	tctattttcc	tcttttgttt	1680
gatagaaaat	cataaaagga	tttgtagact	acgggcctaa	agaactaaaa	aatctatctg	1740
tttcttttca	ttctctgtat	ttttatagtt	ttctgttgca	tgggcataaa	gttgccctttt	1800
taatcacaat	tcagaaaata	tcataatatc	tcatttctct	aaataatagt	gaacggcagg	1860
tatatgtgat	gggttaaaaa	ggatcgatcc	tctagcgaac	cccagagtcc	cgctcagaag	1920

aactcgtcaa	gaaggcgata	gaaggcgatg	cgctgcgaat	cgggagcggc	gataccgtaa	1980
agcacgagga	agcggtcagc	ccattcgccg	ccaagctctt	cagcaatatc	acgggtagcc	2040
aacgctatgt	cctgatagcg	gtccgccaca	cccagccggc	cacagtcgat	gaatccagaa	2100
aagcggccat	ttccaccat	gatattcggc	aagcaggcat	cgccatgggt	cacgacgaga	2160
tcctcgccgt	cgggcatccg	cgcttgagc	ctggcgaaca	gttcggctgg	cgcgagcccc	2220
tgatgctctt	cgtccagatc	atcctgatcg	acaagaccgg	cttccatccg	agtacgtgct	2280
cgctcgatgc	gatgtttcgc	ttggtggtcg	aatgggcagg	tagccggatc	aagcgtatgc	2340
agccgccgca	ttgcatcagc	catgatggat	actttctcgg	caggagcaag	gtgagatgac	2400
aggagatcct	gccccggcac	ttcgcccaat	agcagccagt	cccttcccgc	ttcagtgaca	2460
acgtcgagca	cagctgcgca	aggaacgccc	gtcgtggcca	gccacgatag	ccgcgctgcc	2520
tcgtcttggg	gttcattcag	ggcaccggac	aggtcggctc	tgacaaaaag	aaccgggcgc	2580
ccctgcgctg	acagccggaa	cacggcggca	tcagagcagc	cgattgtctg	ttgtgcccag	2640
tcatagccga	atagcctctc	cacccaagcg	gccggagaac	ctgcgtgcaa	tccatcttgt	2700
tcaatcatgc	gaaacgatcc	tcacctgtc	tcttgatcag	atcttgatcc	cctgcgccat	2760
cagatccttg	gcggaagaa	agccatccag	tttactttgc	agggcttccc	aaccttacca	2820
gagggcgccc	cagctggcaa	ttccggttcg	cttgctgtcc	ataaaaccgc	ccagtctagc	2880
tatcgccatg	taagcccact	gcaagctacc	tgctttctct	ttgcgcttgc	gttttcctt	2940
gtccagatag	cccagtagct	gacattcatc	cggggtcagc	accgtttctg	cggactggct	3000
ttctacgtgt	tccgcttctc	ttagcagccc	ttgcgccctg	agtgettgcg	gcagcgtgaa	3060
gctagatgca	tgctcgagcg	gccgccagtg	tgatggatat	ctgcagaatt	cgcccttccg	3120
gcgaagtgtc	tgctcgcgtg	attgtgcttc	ctttggctac	taaccacgc	gccaagatgc	3180
gttccttgcg	ccacggtttt	gtgaagctgt	tctgccgccg	taactctggc	ctgatcatcg	3240
gtggtgtcgt	ggtggcaccg	accgcgtctg	agctgatcct	accgatcgct	gtggcagtga	3300
ccaaccgtct	gacagttgct	gatctggctg	ataccttcgc	ggtgtacca	tcattgtcag	3360
gttcgattac	tgaagcagca	cgtcagctgg	ttcaacatga	tgatctaggc	taatttttct	3420
gagtcttaga	ttttgagaaa	accaggatt	gctttgtgca	ctcctgggtt	ttactttgt	3480
taagcagttt	tggggaaaag	tgcaaagttt	gcaaagtta	gaaatatttt	aagaggtaa	3540
atgtctgcag	gtggaagcgt	ttaaattcgt	taaacttggc	caaatgtggc	aacctttgca	3600
aggtgaaaaa	ctggggcggg	gtaagggcga	attccagcac	actggcggcc	gttactagct	3660
tatcgccatt	cgccattcag	gctgcgcaac	tgttgggaag	ggcgatcggt	gcgggcctct	3720
tcgctattac	gccagctggc	gaaaggggga	tgtgctgcaa	ggcgattaag	ttgggtaacg	3780
ccagggtttt	cccagtcacg	acgttgtaaa	acgacggcca	gtgaattcaa	cctgtggcgc	3840
aacgctgtat	ataacctgcg	tacggcttaa	agtttggtcg	ccatgtgaat	ttttagcacc	3900
ctcaacagtt	gagtgtcggc	actctcgggg	gtagagtgcc	aaatagggtg	tttgacacac	3960

agttgttcac ccgcgacgac ggctgtgctg gaaaccaca accggcacac acaaaatfff 4020
 tctcatggag ggattcatcg tgtcgactca cacatcttca acgcttccag cattcaaaaa 4080
 gatcttggtg gcaaaccgcg gcgaaatcgc ggtccgtgct tccgtgagc cactcgaaac 4140
 cgggtgcagcc acggtagcta tttacccccg tgaagatcgg ggatcattcc accgctcttt 4200
 tgctttctgaa gctgtccgca ttggtaccga aggtcacca gtcaaggcgt acctggacat 4260
 cgatgaaatt atcgggtgcag ctaaaaaagt taaagcagat gccatttacc cgggatacgg 4320
 cttcctgtct gaaaatgccc agcttgcccc cgagtgtgcg gaaaacggca ttacttttat 4380
 tggcccaacc ccagaggttc ttgatctcac cggtgataag tctcgcgagg taaccgccc 4440
 gaagaaggct ggtctgccag ttttggcggg atccaccccc agcaaaaaca tcgatgagat 4500
 cgttaaaagc gctgaaggcc agacttacc catctttgtg aaggcagttg ccggtggtgg 4560
 cggacgcggt atgcgttttg ttgcttcacc tgatgagctt cgcaaattag caacagaagc 4620
 atctcgtgaa gctgaagcgg ctttcggcga tggcgcggta tatgtcgaac gtgctgtgat 4680
 taaccctcag catattgaag tgcagatcct tggcgatcac actggagaag ttgtacacct 4740
 ttatgaacgt gactgctcac tgcagcgtcg tcaccaaaaa gttgtcgaag ttgcgccagc 4800
 acagcatttg gatccagaac tgcgtgatcg catttgtgcg gatgcagtaa agttctgccg 4860
 ctccattggt taccagggcg cgggaaccaa gggcgaattc ctctggataa tcatcgcggt 4920
 agttacgagc ggcgcgaatg caagggcgaa ttcgagctcg gtacccgggg atcctctaga 4980
 gtcgacctgc aggcattgca gcttggcgtg atcatggtca tagctgtttc ctgtgtgaaa 5040
 ttgttatccg ctcaaatc cacacaacat acgagccgga agcataaagt gtaaagcctg 5100
 gggtgccata tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca 5160
 gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgagg ggagagggcg 5220
 tttgcgtatt gggcgtctt cgccttcctc gctcactgac tcgctgcgct cggctgctcg 5280
 gctgcggcga gcggtatcag ctcaactcaa ggcggaata cggttatcca cagaatcagg 5340
 ggataacgca ggaagaaca tgtgagcaaa aggcagcaa aaggccagga accgtaaaaa 5400
 ggccgcgttg ctggcgttt tccataggct cgccttcctc gacgagcatc acaaaaatcg 5460
 acgctcaagt cagaggtggc gaaaccgac aggactataa agataccagg cgtttcccc 5520
 tggaagctcc ctctgctgct ctctgttcc gaccctgccg cttaccggat acctgtccgc 5580
 cttctcctc tcgggaagcg tggcgtttc tcatagctca cgctgtaggt atctcagttc 5640
 ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa cccccgctc agcccgaccg 5700
 ctgcgcctta tccggttaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 5760
 actggcagca gccactggta acaggattag cagagcgagg tatgtaggag gtgctacaga 5820
 gttcttgaag tgggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 5880
 tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 5940
 caccgctggt agcgggtggt tttttgtttg caagcagcag attacgcgca gaaaaaagg 6000

```

atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 6060
acgttaaggg attttgggtca tgagattatc aaaaaggatc ttcacctaga tccttttggg 6120
gtgggcgaag aactccagca tgagatcccc gcgctggagg atcatccagc cctgatagaa 6180
acagaagcca ctggagcacc tcaaaaacac catcatacac taaatcagta agttggcagc 6240
atcacccgac gcactttgcg ccgaataaat acctgtgacg gaagatcact tcgcagaata 6300
aataaatcct ggtgtccctg ttgataccgg gaagccctgg gccaaacttt ggcgaaaatg 6360
agacgttgat cggcacgtaa gaggttccaa ctttcaccat aatgaaataa gatcactacc 6420
gggcgtattt tttgagttat cgagattttc aggagctgat agaaacagaa gccactggag 6480
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cgacgcactt 6540
tgcgccgaat aaatacctgt gacggaagat cacttcgcag aataaataaa tcctggtgtc 6600
cctgttgata ccgggaagcc ctgggccaac ttttggcgaa aatgagacgt tgatcggcac 6660
gtaagagggt ccaactttca ccataatgaa ataagatcac taccgggcgt attttttgag 6720
ttatcgagat tttcaggagc tctttggcat cgtctctcgc ctgtccccctc agttcagtaa 6780
tttctgcat ttgcctgttt ccagtcggtg gatattccac aaaacagcag ggaagcagcg 6840
cttttccgct gcataaccct gcttcggggt cattatagcg attttttcgg tatatccatc 6900
ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc cttggtgtat 6960
ccaacggcgt cagccgggca ggataggtga agtaggcccc cccgcgagcg ggtgttcctt 7020
cttactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc tgcgaggctg 7080
gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa ccaagccaac 7140
caggaagggc agcccaccta tcaaggtgta ctgccttcca gacgaacgaa gagcgattga 7200
ggaaaaggcg gcggcgggcg gcatgagcct gtcggcctac ctgctggccg tcggccaggg 7260
ctacaaaatc acgggcgctg tggactatga gcacgtccgc gagggcgtcc cggaaaacga 7320
ttccgaagcc caacctttca tagaaggcgg cggtggaatc gaaatctcgt gatggcaggt 7380
tgggcgtcgc ttggtcggtc atttcgctcg gtacccatcg gcattttctt ttgcgttt 7438

```

<210> 22

<211> 4980

5 <212> ADN

<213> Artificial

<220>

<223> plásmido pCLIK int sacB dapA codón bajo

10

<400> 22

ES 2 539 280 T3

acgcgtgcgg tagaatgttt aagaaagctg caggtagcag cgccaactgt tttcggatgat	60
tttgagattg aaactttggc agacggatcg caaatggcaa caagcccgta tgtcatggac	120
ttttaacgca aagctcacac ccacgagcta aaaattcata tagttaagac aacatTTTTG	180
gctgtaaaag acagccgtaa aaacctcttg ctcgtgtcaa ttgttcttat cggaatgtgg	240
cttgggcgat tgttatgcaa aagttgtag gttttttgcg gggttgttta acccccaaat	300

gaggaagaa ggtaacctg aactctatga gcacagggtt aacagctaag acaggggtag 360
agcacttcgg gaccgttggg gtagcaatgg ttactccatt cacggaatcc ggagacatcg 420
atatcgctgc tggccgcgaa gtcgcggctt atttggttga taagggcttg gattctttgg 480
ttctcgcggg caccactggt gaatcccaa cgacaaccgc cgctgaaaaa ctagaactgc 540
tcaaggccgt tcgtgaggaa gttggggatc gggcgaagct catcgccggg gtcggaacca 600
acaacacgcg gacatctgtg gaacttgcgg aagctgctgc ttctgctggc gcagacggcc 660
ttttagttgt aactccttat tactccaagc cggctcgacat cgatgctctt ctgcttaat 720
taacaattgg gatcctctag acccgggatt taaatcgcta gcgggctgct aaaggaagcg 780
gaacacgtag aaagccagtc cgcagaaacg gtgctgacct cggatgaatg tcagctactg 840
ggctatctgg acaagggaaa acgcaagcgc aaagagaaaag caggtagctt gcagtgggct 900
tacatggcga tagctagact gggcggtttt atggacagca agcgaaccgg aattgccagc 960
tggggcgccc tctggtaagg ttgggaagcc ctgcaaagta aactggatgg ctttcttgcc 1020
gccaaaggatc tgatggcgcga ggggatcaag atctgatcaa gagacaggat gaggatcgtt 1080
tcgcatgatt gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct 1140
attcggctat gactgggac aacagacaat cggctgctct gatgccgccg tgttccggct 1200
gtcagcgcag gggcgcggg ttctttttgt caagaccgac ctgtccggtg ccctgaatga 1260
actgcaggac gaggcagcgc ggctatctgt gctggccacg acgggcggtc cttgcccagc 1320
tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg 1380
gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc 1440
aatgcggcgg ctgcatacgc ttgatccggc tacctgcccc ttcgaccacc aagcgaaca 1500
tcgcatcgag cgagcacgta ctcggatgga agccggctct gtcgatcagg atgatctgga 1560
cgaagagcat caggggctcg cgccagccga actgttcgcc aggtcaagg cgcgatgcc 1620
cgacggcgag gatctcgtcg tgaccatgg cgatgcctgc ttgccgaata tcatggtgga 1680
aaatggccc ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca 1740
ggacatagcg ttggctacct gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 1800
cttctcgtg ctttacggta tcgccgctcc cgattcgcag cgcacgcct tctatgcct 1860
tcttgacgag ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgcgccc 1920
aacctgccat cacgagattt cgattccacc gccgccttct atgaaagggt gggcttcgga 1980
atcgttttcc gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc 2040
ttcggccacg ctagcggcgc gccggccggc ccgggtgtgaa ataccgcaca gatgcgtaag 2100
gagaaaatac cgcatcaggc gctcttccgc ttctctgctc actgactcgc tgcgctcggg 2160
cgttcggctg cggcgcggc tatcagctca ctcaaaggcg gtaatacggg tatccacaga 2220
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 2280
taaaaaggcc gcgttgctgg cgttttcca taggctccgc cccctgacg agcatcacia 2340

aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgcagga	ctataaagat	accaggcggt	2400
tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggatacct	2460
gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	2520
cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	2580
cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	2640
atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtagt	taggcgggtgc	2700
tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	agaaggacag	tatttggtat	2760
ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	2820
acaaaccacc	gctggtagcg	gtgggttttt	tgtttgcaag	cagcagatta	cgcgcagaaa	2880
aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	agtggaacga	2940
aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	aggatcttca	cctagatcct	3000
tttaaaggcc	ggccgcggcc	gccatcggca	ttttcttttg	cgtttttatt	tgtaaactgt	3060
taattgtcct	tgttcaagga	tgctgtcttt	gacaacagat	gttttcttgc	ctttgatggt	3120
cagcaggaag	ctcggcgcaa	acgttgattg	tttgtctgcg	tagaatcctc	tgtttgtcat	3180
atagcttgta	atcacgacat	tgtttccttt	cgcttgaggt	acagcgaagt	gtgagtaagt	3240
aaaggttaca	tcgttaggat	caagatccat	ttttaacaca	aggccagttt	tgttcagcgg	3300
cttgtatggg	ccagttaaag	aattagaaac	ataaccaagc	atgtaaatat	cgttagacgt	3360
aatgccgtca	atcgtcattt	ttgatccgcg	ggagtcagtg	aacaggtacc	atgtgccggt	3420
cattttaaag	acgttcgcgc	gttcaatttc	atctgttact	gtgttagatg	caatcagcgg	3480
tttcatcact	tttttcagtg	tgtaatcatc	gtttagctca	atcataccga	gagcgcctgt	3540
tgctaactca	gccgtgcggt	ttttatcgct	ttgcagaagt	ttttgacttt	cttgacggaa	3600
gaatgatgtg	cttttgccat	agtatgcttt	gttaaataaa	gattcttcgc	cttggtagcc	3660
atcttcagtt	ccagtgtttg	cttcaaatac	taagtatttg	tggcctttat	cttctacgta	3720
gtgaggatct	ctcagcgtat	ggttgtcgcc	tgagctgtag	ttgccttcat	cgatgaactg	3780
ctgtacattt	tgatacgttt	ttccgtcacc	gtcaaagatt	gatttataat	cctctacacc	3840
gttgatgttc	aaagagctgt	ctgatgctga	tacgttaact	tgtgcagttg	tcagtgtttg	3900
tttgccgtaa	tgtttaccgg	agaaatcagt	gtagaataaa	cggatttttc	cgtcagatgt	3960
aaatgtggct	gaacctgacc	attcttgtgt	ttggcttttt	aggatagaat	catttgcatac	4020
gaatttgtcg	ctgtctttta	agacgcggcc	agcgtttttc	cagctgtcaa	tagaagtttc	4080
gccgactttt	tgatagaaca	tgtaaactga	tgtgtcatcc	gcatttttag	gatctccggc	4140
taatgcaaag	acgatgtggg	agccgtgata	gtttgcgaca	gtgccgtcag	cgttttgtaa	4200
tggccagctg	tcccaaactg	ccaggccttt	tgcagaagag	atatttttaa	ttgtggacga	4260
atcaaattca	gaaacttgat	atttttcatt	tttttgctgt	tcagggattt	gcagcatatc	4320
atggcgtgta	atatgggaaa	tgccgtatgt	ttccttatat	ggcttttggg	tcgtttcttt	4380

ES 2 539 280 T3

cgcaaacgct	tgagttgcg	ctcctgccag	cagtgcggt	gtaaaggta	atactgttgc	4440
ttgttttgca	aactttttga	tgttcatcgt	tcatgtctcc	ttttttatgt	actgtgttag	4500
cggtctgctt	cttcagccc	tcctgtttga	agatggcaag	ttagttacgc	acaataaaaa	4560
aagacctaaa	atatgtaagg	ggtgacgcca	aagtatacac	tttgccttt	acacatttta	4620
ggtcttgcc	gctttatcag	taacaaacc	gcgcgattta	cttttcgacc	tcattctatt	4680
agactctcgt	ttggattgca	actggtctat	tttcctcttt	tgtttgatag	aaaatcataa	4740
aaggatttgc	agactacggg	cctaaagaac	taaaaaatct	atctgtttct	tttcattctc	4800
tgtatTTTTT	atagtttctg	ttgcatgggc	ataaagttgc	ctttttaatc	acaattcaga	4860
aaatatcata	atatctcatt	tcactaaata	atagtgaacg	gcaggtatat	gtgatggggt	4920
aaaaggatc	ggcggccgct	cgatttaa	ctcgagaggc	ctgacgtcgg	gcccgtacc	4980

REIVINDICACIONES

1. Procedimiento de reducción de la cantidad de al menos un polipéptido en una célula de *Corynebacterium glutamicum*, que comprende la etapa de expresar en una célula de *C. glutamicum* una secuencia de nucleótidos modificada en lugar de una secuencia de nucleótidos no modificada que codifica dicho polipéptido que tiene la misma secuencia de aminoácidos y/o función, en el que dicha secuencia de nucleótidos deriva de la secuencia de nucleótidos no modificada, de forma que al menos un codón de la secuencia de nucleótidos no modificada esté sustituido por un codón de uso menos frecuente de acuerdo con el uso de codones de *C. glutamicum*, en el que dicho al menos un codón es el codón de iniciación que está sustituido con un codón de iniciación que se usa con menos frecuencia en *C. glutamicum*, en el que dicho codón de iniciación usado con menor frecuencia es GTG o TTG.
2. Procedimiento de acuerdo con la reivindicación 1, en el que la secuencia de nucleótidos modificada deriva de la secuencia de nucleótidos no modificada, de forma que al menos un codón adicional de la secuencia de nucleótidos no modificada esté reemplazado en la secuencia de nucleótidos modificada por un codón menos usado de acuerdo con el uso de codones de proteínas abundantes de *C. glutamicum*.
3. Procedimiento de acuerdo con la reivindicación 1 o 2, en el que al menos un codón de la secuencia de nucleótidos no modificada está reemplazado en la secuencia de nucleótidos modificada por uno de los dos codones menos usados como se expone en la tabla 1.
4. Procedimiento de acuerdo con la reivindicación 3, en el que todos los codones de dicha secuencia de nucleótidos modificada para cada aminoácido son seleccionados del uso de codones de la tabla 3.
5. Procedimiento de acuerdo con la reivindicación 1 o 2, en el que al menos un codón de la secuencia de nucleótidos no modificada está reemplazado en la secuencia de nucleótidos modificada por uno de los dos codones menos usados como se expone en la tabla 2.
6. Procedimiento de acuerdo con la reivindicación 5, en el que dicha secuencia de nucleótidos modificada usa los codones ACG para treonina; o GAT para ácido aspártico; o GAA para ácido glutámico; o AGA y AGG para arginina y/o TTG para el codón de iniciación.
7. Procedimiento de acuerdo con la reivindicación 5 o 6, en el que todos los codones de dicha secuencia de nucleótidos modificada para cada aminoácido son seleccionados del uso de codones de la tabla 4.
8. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en el que se usa la cepa ATCC 13032 de *C. glutamicum* o derivados de la misma.
9. Procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 8, en el que el procedimiento es usado para disminuir la expresión de dicho polipéptido en la célula de *C. glutamicum* en al menos 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o 95 %, determinándose la extensión de la reducción de la expresión en comparación con el nivel de expresión del polipéptido que se expresa a partir de la secuencia de nucleótidos no modificada en condiciones comparables.

Figura 1

Secuencia génica de tipo salvaje de la isocitrato deshidrogenasa (SEC ID N°1):

atggctaagatcactcggaccgcaccgacgaagcaccgctgctcggacctaactcgcctgaagccggctcgcgaggca
 ttgtgctaccgcgggcatcagagtcgagaccgggacattcaactcgcctggacgcacatcctcgcccagttcccagag
 cgcctcaccgaagatcagaaggtaggcaacgcactcgcagaactcggcgagcttgctaagactcctgaagcaaacatc
 attaagcttccaaacatctccgctctcgttccacagctcaaggctgctattaaggaactgcaggaccagggtacgac
 atcccagaactgcctgataacgccaccaccgacgaggaaaaaagacatcctcgcacgctacaacgctgttaagggttcc
 gctgtgaaccagtgctgctgtaaggcaactctgaccgcccgcaccaatcgcctgtcaagaactttgttaagaagttc
 ccacaccgcatggcgagtggtctgcagattccaagaccaacttgcaaccaatggaatgcaaacgacttccgccacaac
 gagaagtccatcactcogacgctgctgataagttcagatcaagcacaatcgcagctgacggcaccgagaccatcctc
 aaggacagcctcaagctctctgaaggcgaagttctagacggaaaccgtctgtccgcaaaggcaactggacgcaatcctt
 ctgcagcaggctcgcctcgcgcaaaggcagaaggtatcctctctcgcacacctgaaggccaatgataaggtctcc
 gacccaatcactctcggccacgttgctgctgcttactctgcagacgttttcgcacagtacggtagcagctgctcgc
 gctggcctcaacggcgaaaacggcctcgcctgcaatcctctcggcttgaggtccctggacaacggcgaaagaaatcaag
 gctgcatcogagaagggttggaagacggcccagacctggccaatggttaactccgctcgcggcaatccaacctgcat
 gtcctctccgatgctcogtggaagcttccaatgcagcaatgattcgtacctccggccacatgtaggaacaaagacgac
 caggagcaggacaccctggcaatcactcccagactcctctacgctggcgtctaccagaccgttatcgaagactgccc
 aagaacggcgcatcgaatccaaccaccaatgggtaccgtccctaacgttggtctgataggctcagaaggctgaagagtac
 ggctcccatgacaagaccttccgcatcgaagcagaagggttggttcagggtggttctctccaacggcgacgttctc
 gagcagcagcttgaggcaaatgacatctggcgtgcaatgccaggtcaaggatgcccacatccaggattgggtaaagctt
 gctgtcaccogctccgctctctccggaatgcctgcagtggttctgggttggaatccagagcgcgcacacgaccgcaac
 ctgcttccctcgttgagaagtacctggctgaccacgacaccgaggcctggacatccagatcctctccccgttgaggca
 acccagctctccatcgaaccgcatcgcgctggcgaggacaccatctctgtcaccggtaacgttctgctgactacaac
 accgacctcttcccaatcctggagctgggacacctcgcgaagatgctgtctgtcgttctcttgaatggctggcggg
 cggatctgcagaccggctgctgggtggaatctgctcctaaagcagctccagcaggttcaggaagaaaaccacctgcgt
 tgggat
 tccctcgggtgagttctctgcactggctgagtccttccgcccagagctcaacaaacacggcaacaccaaggccggcgtt
 ctggtgacgctctggacaaggcaactgagaagctgctgaacgaagagaagtccccatcccgaagggtggcgagatc
 gacaaccgtggctcccacttctggctgaccaagtctctgggctgacgagctcgcctgctcagaccgaggaacgagatc
 gctgctaccttgcaccagctcgcagaagcactgaaacagggcctgcagacatcgaatgctgcaactgctcagttcag
 ggtggagcaactgacctgggtggctactactccccatacagaggagaagctcaccacatcagtcgcccagtcgcacag
 ttcaacgagatcgttgacgcactgaagaagtaa

Figura 2

a) Secuencia de la isocitrato deshidrogenasa ATG-GTG (SEC ID N° 2):

gtggetaagatcatctggaccgcaccgaagaagcaacgctgctcgcgacctactcgctgaagccggtogtgcaggca
 ttgtgctaccgcgggcatgaggctcgagaccgggacatttcaactcgctggagcgatcctcgcccagttcccagag
 cgcctcaccgaagatcagaaggtaggcaacgcactcgcaagaactcggcgagcttgctaagactcctgaagcaaacatc
 attaagcttccaaacatctccgcttctgttccacagctcaaggctgctattaaggaactgcaggaccagggtacgac
 atcccagaactgctgataaacgccaccaccgacgaggaaaaagacatcctcgcaagctacaacgctgttaagggttcc
 gctgtgaaccagtgctgctgaaggcaactctgaccgcgcgaccaatcgctgtcaagaactttgttaagaagtcc
 ccacaccgcatgggaggtggtctgcagattccaagaccaacgttgcaaccatggatgcaaacgacttccgccacaac
 gagaagtccatcctcgcagctgctgatgaagtccagatcaagcacaatcgagctgacggcaccgagaccatcctc
 aaggacagcctcaagcttctgaaggcgaagtcttagacggaacggttctgtccgcaaggcactggagcattcctt
 ctcgagcaggtcgctcgcaaaaggcagaaggtatcctcttctcgcacacctgaaggccaccatgataaggtctcc
 gaccatcctctcggccacgttgtgctgcttactctcgacagcttctcgcacagctacggtagcagctgctcgca
 gctggcctcaacggcgaaaacggcctcgctgcaatcctctccggctggagtcctggacaacggcgagaagaatcaag
 gctgcatcgcagaagggcttgggaagcggcccagacctggccatgggtaactcgcctcgcgccaaccacaacctgcat
 gtccctccgagtctcagctgctgacgcttccatgcccagcaatgatctgctacctcggccacaatggaacaagaagcagc
 caggagcaggacacctggcaatcctcccagactcctcctacgctggcgtctaccagaccgttaccgaagactgcgcg
 aagaaaggcgcatcagatcccaaccaccaagggtaccgtccctaacgttggctgatggctcagaaggctgaagagtag
 ggctcccatgacaagaccttccgcatcgaagcagacgggtgtggttcagggtgttctcctcaacggcgagcttctcatc
 gagcagcagcttgaggcaaaatgacatctggcgtgcatgcccaggtcaaggatgcccacaaccaggattgggtaaaagt
 gctgtcaccgctcccgctctctccggaaatgcctgcagtgcttctgggttggatccagagcgcgcaacgaccgcaacctg
 gcttccctcgttgagaagtacctggctgaccacgacaccgagggcctggacatccagatcctctccctggttgaggca
 acccagctctccatcgaccgcatccgctggtggcgaggacacctctctgtcaccggtaacgttctgcgtgactcaac
 accgacctctcccaatcctggagctgggcaacctctgcaagaatgctgtctgtcgttctttagtggtggcgcgga
 ctgctcgagaccggtgctggtggatctgctcctaaagcagctccagcaggttcaggaagaaaaccacctgcttgggga
 tccctcgggtgagttcctcgcactggctgagctcctcgcgcaagcagctcaacaacaacggcaacaaccaggccgct
 ctggctgacgctctggacaaggcaactgagaagctgctgaacgaagagaagtcaccaaccgcaaggtggcgagatc
 gacaacgggtggctcccactctggctgaccaagtctgggctgacgagctcgtgctcagaccgaggacgcagatctg
 gctgtacctctgcaccagctcgagaagcactgaacacagggcctgcagacatcagatgctgcaactgctcagctcag
 ggtggagcaactgacctgggtggctactactccctaacgaggagaagctcaacaacatcagcggccagctgcacag
 ttcaacgagatcgttgacgcactgaagaagtaa

b) Secuencia del inserto usado para pCiik int sac ICD ATG_GTG (SEC ID N° 3):

ctcgagcgaagacctcgagattccgatatccaggaacggccatgacgaaa tccctcagatgacgatgcaactgac
 catcgaggacaccttccatcagatgtgaaatggctgccccgcaacggccgcaagcaggtgaaattggtgagtgaaac
 cctggccctccaccaatgaagaacacagaagctgcagccacctcgaaggcgaacttggtgaggagactcctgtgtctc
 cgccactggcgaacagatcacagaatccaacccacgttcaggcgactactactggaattgctggcgaaagtgggtgctgt
 gaccagcatctgctgatactctagtgaaagagaaggcctcgaccgttcccaagtggcattcatggggatattggaaaca
 cgggcttccaatgccccgctgaaactgccaccaataggcgcagcaattagtagaacactgtaattctaggtagctgaac
 aaaagagcccaatcaaccaaggagactcgtggctaaagatcatctggaccgcaccgacgaagcaaccgctcgcgacc
 tactcgtgaagccggtcgtcgaggcatctgtgctaccgcgggcatgaggctcgagaccgggacatttcaactcgcct
 ggagcgaatcctcgcccagttcccagagcgcctcaaccgaagatcagaaggtaggcaacgcactcgcagaactcggcgag
 ctgtgctaagactcctgaagcaaacatcattaagcttccaacatctccgcttctgttccacagctcaaggctgctatt
 aaggaaactgcaggaccagggtacgacatcccagaactgcctgataacggccaccaccgacgaggaaaaagacatcctc
 gacgctacaacgctgttaagggttccgctgtgaacccagtgctgctgtaaggcaactctgacccgcgcaaccaatc
 gctgtcaagaacttgttaagaagttcccacaccgcatggggcagtggtctgcagattccacgct

Figura 3

a) Secuencia de la isocitrato deshidrogenasa CA2 (SEC ID N° 4):

atggctaagatca tctggaccgcaccgacgaagcaccgctgctcgcgacctactcgctgaagccggctcgtcgaggca
 ttgtgcttaccgcggggatagaggctcgagaccgggacatttcactcgcctggacgcacacctcgcaccagttcccagag
 cgcctcaccogaagatcagaaggttaggcaacgcactcgcgagaactcggcgagcttgctaagactcctgaagcaaacatc
 attaagcttccaaacatctccgcttctgttccacagctcaaggctgctattaaggaactgcaggaccagggttacgac
 atcccagaactgcotgataacgcccaccaccgacgaggaaaaagacatcctcgcacgctacaacgctgttaagggtcc
 gctgtgaaccagtgctgctggaaggcaactctgaccgcccgcaccaatcgtgtcaagaactttgttaagaagttc
 ccacaccgcatggggcagtggtctgcagatccaagaccaacgcttgcaacca tggatgcaaacgacttccgccacaac
 gagaagtccatcactcgcagcctgctgataagttcagatcaagcacatcgcagctgacggcaccgagacca tccctc
 aaggacagcctcaagcttcttgaaggcgaagtcttagacggaaccgcttotgtccgcaaaaggcaactggacgcatctct
 ctcgagcaggtcctcgcgcaaaaggcagaaggtatcctcttctcgcacacctgaaggccacca tga tgaaggctctc
 gacccaatcactctcggccacgttggtgctgcttactctgcagacgtttctgcacagta cggtgagcagctgctcgc
 gctggcctcaacggcgaaaaacggcctcgcctgcaatcctctcgcgctggagctccctggacaa cggcgaagaaa tcaag
 gctgcatctcgagaagggtctggaagacggccagacctggccatggtaactcgcctcgcgca tcaaccaacctgcat
 gtcccttccgatgtca tctgtggaagcttccatgccagcaatgatctgtacctcggccacatgtggaacaaagacgac
 caggagcaggacacctggcaatcactccagactcctctcgcctggcgtctaccagacctta tcaagactgcccgc
 aagaacggcgcatctcga tccaaccacca tgggtaccgtccctaacgttggctctga tggctcagaaggctgaaggtac
 ggctcccatgacaagaccttccgcatcgaagcagacggtgtgggtcagggtgtttctccaa cggcgacgttctcactc
 gagcacgacgttgaggcaaa tgcactctggcgtgcatgcccaggctcaaggatgccccaatccaggat tgggtaaagctt
 gctgtcaccgctcccgctctctcgcgaa tgcctgcagtgctctgggtggatccagagcgcgcaca cgcaccgcaacctg
 gcttccctcgttgagaagtacctggctgacca cgcaccgagggcctggacatccagatcctctcccctgttgaggca
 accagctctctcga tgcaccgcatccgcccgtggcgaggacacctctctgtcaccggtaacgttctcgtgactacaac
 accgacctctcccaatcctggagctgggcaacctctgcaaaagatgctgtctgtctccttga tggctggcggcgga
 ctgtctcagagaccggtgctgggtgga tctgctcctaaagcagctccagcaggttcaggaagaaaaccacctgctgtgggat
 tccctcgggtgagttcctcgcactggctgagctcctcgcaccagagctcaacaa caacggcaacccaaggccggcgtt
 ctggctgacgctctggacaaggcaactgagaagctgctgaa cgaagagaag tcccca tcccgcaagggttggcgagatc
 gacaaccgtggctcccactctggctgaccaagttctgggctgacgagctcgc tgc tcaagaccgaggacgcagatctg
 gctgtacctctcgcactcgcagaagcactgaa cacagggcgtgcagacatcga tgetgactgctcgcagttcag
 ggtgggcaactccttgggtggctactactccccctaa cggaggagaagctcaccaacatcactgcgccagctcgcacag
 tcaacgagatcgttgacgactgaagaagtaa

b) Secuencia del inserto usado para pCiik int sac ICD CA2 ATG_GTG (SEC ID N° 5):

ctcgagcgaagacctcgcagattccgata tccaggaaccggccatgatacgaatccccctcagatgacgatgcacttgc
catcgagggaacctcctccatcga tgtgaaa tggctgccccgcaacggccgcaagcacggtgaattggtgatggaaac
 cctggccctccacca tgaagaaa cagaagctgcagccacctccgaaggcgaacttg tgtgggagactcctgtgttctc
 cggcactggcgaaacagatcacaagaatccaaccacgcttcaggcgactactactgga ttgctggcgaaagtgggtcgt
 gaccagca tctcgtcga tctctagtgaagagaaaggcctcgaaccttcccaagtggtattca tgggggtattggaaca
 cggcgttcca tgcggggctgaaactgccacca tagggcgcagcaat tagtagaacactgta tctaggtagctgaa
 aaaagagccca tcaaccaaggagactca tggctaaagatca tctggaccgcaaccgacgaagcaccgctgctcgcgacc
 tactcgtgaagccggtcgtcgaaggcatttgc tgc taccgcccggatagaggctcagagaccgggacatttcaactcgt
 ggacgcatcctcgcaccagttcccagagcgcctcaccgaaga tcaagaaggtaggcaacgcactcgcagaaactcggcgag
 cttgctaagactcctgaagcaaacatcattaa gcttccaaacatctccgcttctgttccacagctcaaggctgctatt
 aaggaaactgcaggaccagggtacgacatcccagaactgctgataacgccaccaccgacgaggaaaaagacatcctc
 gcacgcta caacgctgttaagggttccgctgtgaaccagtgctgctggaaggcaactctgaccgcccgcgcaaccaatc
 gctgtcaagaactttgttaagaagttcccaca ccgcatgggcgagtggtctgcagattccacgct