

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 435 990**

51 Int. Cl.:

**C12N 9/78**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.07.2005 E 05762814 (1)**

97 Fecha y número de publicación de la concesión europea: **21.08.2013 EP 1814983**

54 Título: **Enzimas mejoradas**

30 Prioridad:

**07.07.2004 EP 04015584**

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

**26.12.2013**

73 Titular/es:

**DSM IP ASSETS B.V. (100.0%)  
HET OVERLOON 1  
6411 TE HEERLEN, NL**

72 Inventor/es:

**EBERT, SYBILLE;  
HOHMANN, HANS-PETER;  
LEHMANN, MARTIN;  
MOUNCEY, NIGEL, JOHN y  
WYSS, MARKUS**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 435 990 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

## Enzimas mejoradas

La presente invención proporciona enzimas modificadas con actividad de GTP ciclohidrolasa II mayor que las enzimas de tipo salvaje respectivas. Las enzimas modificadas y polinucleótidos que las codifican se pueden usar para la producción de riboflavina, precursores de riboflavina, mononucleótido de flavina (FMN), dinucleótido de flavina y adenina (FAD), y sus derivados.

La riboflavina (vitamina B<sub>2</sub>) es sintetizada por todas las plantas y muchos microorganismos, pero no es producida por animales superiores. Debido a que es un precursor de coenzimas tales como el dinucleótido de flavina y adenina y el mononucleótido de flavina, que son necesarios en la oxidación enzimática de hidratos de carbono, la riboflavina es esencial para el metabolismo básico. En animales superiores, una insuficiencia de riboflavina puede provocar pérdida del cabello, inflamación de la piel, deterioro de la visión, e insuficiencia del crecimiento.

La manipulación mediante ingeniería de cepas de producción de riboflavina con mayores velocidades y rendimientos de producción de riboflavina se ha logrado en el pasado de muchas maneras diferentes. Por ejemplo, (1) se usó mutagénesis clásica para generar variantes con mutaciones aleatorias en el genoma del organismo de elección, seguido de la selección por mayor resistencia a análogos de purina y/o identificando una mayor producción de riboflavina. (2) Como alternativa, las enzimas terminales de la biosíntesis de riboflavina, es decir, las enzimas que catalizan la conversión de trifosfato de guanosa (GTP) y ribulosa-5-fosfato en riboflavina, estaban sobreexpresadas, dando como resultado también un mayor flujo hacia el producto diana. Sin embargo, en este último enfoque, la fuerte sobreexpresión de las proteínas de la biosíntesis de riboflavina impone una carga metabólica adicional en las células hospedantes, lo que, a su vez, puede inducir reacciones de respuesta de estrés y otros efectos negativos indeseables en la fisiología de las células.

Las enzimas requeridas que catalizan la biosíntesis de riboflavina a partir de trifosfato de guanosa (GTP) y ribulosa-5-fosfato están codificadas por cuatro genes (*ribG*, *ribB*, *ribA*, y *ribH*) en *B. subtilis*. Estos genes están situados en un operón, cuyo orden génico difiere del orden de las reacciones enzimáticas catalizadas por las enzimas. Por ejemplo, GTP ciclohidrolasa II, que cataliza la primera etapa en la biosíntesis de riboflavina, es codificada por el tercer gen en el operón, *ribA*. El gen *ribA* también codifica una segunda actividad enzimática, es decir, 3,4-dihidroxi-2-butanona 4-fosfato sintasa (DHBPS), que cataliza la conversión de ribulosa-5-fosfato a la unidad de cuatro carbonos 3,4-dihidroxi-2-butanona 4-fosfato (DHBP). La desaminasa y la reductasa son codificadas por el primer gen del operón, *ribG*. La penúltima etapa en la biosíntesis de la riboflavina está catalizada por lumazina sintasa, el producto del último gen *rib*, *ribH*. La riboflavina sintasa, que controla la última etapa de la ruta, es codificada por el segundo gen del operón, *ribB*. La función del gen situado en el extremo 3' del operón *rib* es actualmente incierta; sin embargo, su producto génico no es necesario para la síntesis de riboflavina.

La transcripción del operón de riboflavina a partir del promotor *ribP1* está controlada por un mecanismo de atenuación que implica una región líder reguladora situada entre *ribP1* y *ribG*. Las mutaciones *ribO* en esta región líder dan como resultado la expresión desregulada del operón de riboflavina. La expresión desregulada también se observa en cepas que contienen mutaciones con cambio de sentido en el gen *ribC*. Se ha demostrado que el gen *ribC* codifica la flavina cinasa/FAD sintasa de *B. subtilis* (Mack, M., et al., J. Bacteriol., 180:950-955, 1998). Las mutaciones desregulantes reducen la actividad de flavocinasa del producto del gen *ribC* dando como resultado concentraciones intracelulares reducidas de mononucleótido de flavina (FMN), la molécula efectora del sistema regulador de riboflavina.

Recientemente, se modificó genéticamente mediante ingeniería a *Bacillus subtilis* para producir niveles elevados de riboflavina durante un ciclo de fermentación corto (patente U.S. n° 5.837.528). Este enfoque combinó selección clásica de mutantes genéticos y mejora de la fermentación con ingeniería genética de los genes biosintéticos de riboflavina desregulando e incrementando el nivel de expresión génica. En este sistema, la expresión de los genes *rib* se incrementó al mutar el gen *ribC* que codifica flavocinasa, enlazando los genes *rib* a promotores constitutivos fuertes, e incrementando el número de copias de los genes *rib*.

Como ya se ha discutido anteriormente, la sobreexpresión de los genes *rib* plantea una carga adicional en las cepas de producción, lo cual puede tener, potencialmente, un impacto negativo sobre la producción de precursores de riboflavina, fiboflavina, FMN, FAD, o sus derivados. A fin de eludir este defecto, es un objeto de la presente invención describir mutantes de GTP ciclohidrolasa II con actividad específica incrementada. El uso de tales enzimas mutantes en cepas de producción, ya sea solas o combinadas con mutantes mejorados de las otras proteínas Rib, permitirá mayores velocidades de flujo con menos carga o ninguna carga adicional sobre el metabolismo de las células.

Como se usa aquí, la expresión "GTP ciclohidrolasa II" puede incluir cualquier enzima que sea capaz de catalizar la conversión de GTP en 2,5-diamino-6-ribosilamino-4 (3H)-pirimidinona-5'-fosfato DRAPP. Es irrelevante si esta enzima es capaz de catalizar otras reacciones, como por ejemplo la conversión de ribulosa-5-fosfato en DHBP. Una "GTP ciclohidrolasa II" puede ser homóloga a una o más de las enzimas cuyas secuencias de aminoácidos se muestran en la Figura 1 o en la Tabla 4. "Homóloga" se refiere a una GTP ciclohidrolasa II que es al menos

alrededor de 50% idéntica, preferiblemente al menos alrededor de 60% idéntica, más preferiblemente al menos alrededor de 70%, incluso más preferiblemente al menos alrededor de 80%, incluso más preferiblemente al menos alrededor de 85%, incluso más preferiblemente al menos alrededor de 90% o 95% idéntica, y lo más preferible al menos alrededor de 98% idéntica a una o más de las secuencias de aminoácidos como se muestran en la Figura 1 o en la Tabla 4.

La expresión "% de identidad", como se conoce en la técnica, significa el grado de relación entre secuencias polipeptídicas o polinucleotídicas, según sea el caso, como se determina mediante el emparejamiento entre cadenas de tales secuencias. La "identidad" se puede determinar fácilmente mediante métodos conocidos, por ejemplo con el programa BESTFIT (GCG Wisconsin Package, versión 10.2, Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752, USA) usando los siguientes parámetros: penalización de creación de salto 8, penalización de extensión de salto 2 (parámetros por defecto).

"Enzima de tipo salvaje" o "GTP ciclohidrolasa II de tipo salvaje" puede incluir cualquier GTP ciclohidrolasa II homóloga a una cualquiera de las enzimas mostradas en la Figura 1 o en la Tabla 4 que se usa como punto de partida para diseñar mutantes con mayor actividad según la presente invención. "Tipo salvaje", en el contexto de la presente invención, puede incluir tanto secuencias de GTP ciclohidrolasa II derivables de la naturaleza así como variantes de enzimas GTP ciclohidrolasa II sintéticas (en tanto que sean homólogas a una cualquiera de las secuencias mostradas en la Figura 1 o en la Tabla 4), si se pueden hacer más activas mediante cualquiera de las enseñanzas de la presente invención. Las expresiones "GTP ciclohidrolasa II de tipo salvaje" y "GTP ciclohidrolasa II no modificada" se usan aquí de forma intercambiable.

Un "mutante", "enzima mutante", o "GTP ciclohidrolasa II mutante", puede incluir cualquier variante derivable de una enzima/ GTP ciclohidrolasa II de tipo salvaje dada (según la definición anterior) de acuerdo con las enseñanzas de la presente invención, y que es más activa que la enzima de tipo salvaje respectiva. Para el alcance de la presente invención, no es relevante cómo se obtiene el mutante o mutantes; tales mutantes se pueden obtener, por ejemplo, mediante mutagénesis dirigida al sitio, mutagénesis de saturación, mutagénesis al azar/evolución dirigida, mutagénesis química o por UV de todas las células/organismos, y otros métodos que son conocidos en la técnica. Estos mutantes también se pueden generar, por ejemplo, diseñando genes sintéticos, y/o se pueden producir mediante traducción in vitro (libre de células). Para ensayar la actividad específica, los mutantes se pueden (sobre)expresar mediante métodos conocidos por los expertos en la técnica. Las expresiones "GTP ciclohidrolasa II mutante" y "GTP ciclohidrolasa II modificada" se usan aquí de forma intercambiable. Esto también se aplica a las expresiones "enzima mutante" y "enzima modificada".

"Precursor de riboflavina" y "derivados de riboflavina, FMN o FAD", en el contexto de esta invención de esta solicitud de patente, puede incluir cualquiera y todos los metabolitos que requieran como enzima intermedia GTP ciclohidrolasa II en su (bio)síntesis. En el contexto de esta solicitud de patente, es irrelevante si tales rutas (bio)sintéticas son naturales o no naturales (es decir, rutas que no se producen en la naturaleza, sino que están manipuladas biotecnológicamente mediante ingeniería). Preferiblemente, las rutas sintéticas son de naturaleza bioquímica. Los precursores de riboflavina y derivados de riboflavina, FMN o FAD, incluyen pero no se limitan a: DRAPP; 5-amino-6-ribosilamino-2,4(1H,3H)-pirimidindiona-5'-fosfato; 2,5-diamino-6-ribitilamino-4(3H)-pirimidinona-5'-fosfato; 5-amino-6-ribitilamino-2,4(1H,3H)-pirimidindiona-5'-fosfato; 5-amino-6-ribitilamino-2,4(1H,3H)-pirimidindiona; 6,7-dimetil-8-ribitilumazina (DMRL); y flavoproteínas. El término "riboflavina" también incluye derivados de riboflavina, tales como, *por ejemplo*, riboflavin-5-fosfato, y sus sales, tales como, *por ejemplo*, riboflavin-5-fosfato de sodio.

En general, es un objeto de la presente invención proporcionar una enzima que tiene actividad de GTP ciclohidrolasa II, estando dicha enzima modificada de tal manera que las propiedades catalíticas son más favorables (es decir, que muestran mayor actividad específica) que aquellas de la enzima GTP ciclohidrolasa II no modificada.

De este modo, la presente invención se refiere a una GTP ciclohidrolasa II modificada seleccionada de *Bacillus*, en la que

(i) la actividad específica de la enzima modificada aumenta en al menos alrededor de 10 % en comparación con la enzima no modificada correspondiente, y

(ii) la secuencia de aminoácidos de la enzima modificada comprende una o más sustituciones que incluyen 1, 2, 3, 4, 5, o 6 sustituciones en la posición o posiciones de aminoácidos que corresponden a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEC ID NO:2,

en la que dichas sustituciones conducen a restos de aminoácidos seleccionados del grupo que consiste en:

(a) alanina en una posición que corresponde a la posición 261 de SEQ ID NO:2,

(b) alanina en una posición que corresponde a la posición 270 de SEQ ID NO:2,

(c) treonina en una posición que corresponde a la posición 276 de SEQ ID NO:2,

- (d) arginina en una posición que corresponde a la posición 279 de SEQ ID NO:2,
- (e) arginina en una posición que corresponde a la posición 308 de SEQ ID NO:2, y
- (f) isoleucina en una posición que corresponde a la posición 347 de SEQ ID NO:2.

De este modo, la presente invención se refiere a una GTP ciclohidrolasa II modificada seleccionada de *Bacillus* que muestra un incremento en la actividad específica de al menos alrededor de 10 % en comparación con la enzima no modificada correspondiente, en la que se han sustituido restos de aminoácidos en las posiciones que corresponden a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEQ ID NO:2, conduciendo dichas sustituciones a restos de aminoácidos seleccionados del grupo que consiste en:

- (a) alanina en una posición que corresponde a la posición 261 de SEQ ID NO:2,
- (b) alanina en una posición que corresponde a la posición 270 de SEQ ID NO:2,
- (c) treonina en una posición que corresponde a la posición 276 de SEQ ID NO:2,
- (d) arginina en una posición que corresponde a la posición 279 de SEQ ID NO:2,
- (e) arginina en una posición que corresponde a la posición 308 de SEQ ID NO:2, y
- (f) isoleucina en una posición que corresponde a la posición 347 de SEQ ID NO:2.

La expresión "al menos una sustitución" significa una o más mutaciones en una posición como se define anteriormente que conduce a una GTP ciclohidrolasa II modificada que tiene mayor actividad específica en comparación con la enzima no modificada. Una enzima modificada como se describe anteriormente puede consistir en solamente 1, 2, 3, 4, 5 ó 6 sustituciones en una posición como se define anteriormente que conduce a una mayor actividad específica en comparación con la enzima no modificada, pero también puede incluir otras mutaciones de aminoácidos en otras posiciones, en tanto que la enzima modificada resultante tenga actividad específica aumentada. De este modo, la enzima modificada comprende una o más sustituciones, incluyendo 1, 2, 3, 4, 5 ó 6 mutaciones, en una posición o posiciones de aminoácidos correspondientes a las posiciones 261, 270, 276, 279, 308 y/o 347 de la secuencia de aminoácidos de GTP ciclohidrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO: 2. Los ejemplos de tales mutaciones en posiciones distintas de las definidas anteriormente son mutaciones de aminoácidos en una posición correspondiente a la posición de aminoácidos 196, 282, y/o 325 de SEC ID NO:2.

Como se usa aquí, la expresión "actividad específica" representa la velocidad de reacción de las enzimas GTP ciclohidrolasa II de tipo salvaje y mutantes en condiciones de reacción apropiadamente definidas como se describe, *por ejemplo*, en Ritz et al. (J. Biol. Chem. 276, 22273-22277, 2001), Koh et al. (Mol. Gen. Genet. 251, 591-598, 1996), o Schramek et al. (J. Biol. Chem. 276, 44157-44162, 2001), o como se describe en detalle en el Ejemplo 2. La "actividad específica" define la cantidad de un sustrato consumido y/o producto producido en un período de tiempo dado y por cantidad definida de proteína a una temperatura definida. Típicamente, la "actividad específica" se expresa en  $\mu$ moles de sustrato consumido o producto formado por min. por mg de proteína. Típicamente,  $\mu$ mol/min. se abrevia mediante U (= unidad). Por lo tanto, las definiciones de unidad para actividad específica de  $\mu$ mol/min./ (mg de proteína) o U/(mg de proteína) se usan de forma intercambiable a lo largo de este documento. Se entiende que en el contexto de la presente invención, la actividad específica se debe de comparar en base a una longitud similar, o preferiblemente idéntica, de la cadena polipeptídica. La invención no se debe eludir incrementando el tamaño de una enzima de tipo salvaje dada a través de la formación, *por ejemplo*, de una proteína de fusión, reduciendo de ese modo la actividad específica aparente de la enzima global.

Según la presente invención, la GTP ciclohidrolasa II modificada muestra una actividad específica que es mayor en al menos alrededor de 10% que la de la enzima no modificada correspondiente. Preferiblemente, la actividad específica de la GTP ciclohidrolasa II modificada de la invención está incrementada en al menos alrededor de 25, 40, 60, 70, 80, 85, 90%, más preferiblemente al menos alrededor de 70%, en comparación con la GTP ciclohidrolasa II no modificada correspondiente (para la medición de la actividad específica, véase más abajo). Preferiblemente, incremento en la actividad específica se refiere a las condiciones experimentales descritas en el Ejemplo 1 de esta solicitud. Aproximadamente 0,004-0,02 U/ml (que corresponde a aproximadamente 40  $\mu$ g/ml de GTP ciclohidrolasa II de *Bacillus subtilis*, o 20  $\mu$ g/ml para los mejores mutantes descritos aquí), preferiblemente de forma aproximada 0,004 U/ml de actividad de GTP ciclohidrolasa II, estuvieron presentes en la mezcla de ensayo, y la reacción se llevó a cabo a 37°C.

La secuencia de aminoácidos de la GTP ciclohidrolasa II modificada de la invención contiene al menos una sustitución como se define anteriormente, cuando se compara con la secuencia de aminoácidos de la GTP ciclohidrolasa II no modificada correspondiente.

La secuencia de aminoácidos de la GTP ciclohidrolasa II modificada contiene al menos una sustitución de aminoácido cuando se compara con la secuencia de aminoácidos de la GTP ciclohidrolasa II no modificada correspondiente, es decir, comprende una o más, incluyendo 1, 2, 3, 4, 5, o 6, sustituciones de aminoácidos en las

posiciones de aminoácidos que corresponden a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEC ID NO:2, preferiblemente 2, 3, 4 ó 5 sustituciones de aminoácidos. De este modo, la enzima modificada contiene preferiblemente al menos 2, al menos 3, al menos 4 o al menos 5 sustituciones cuando se compara con la secuencia de aminoácidos de la GTP ciclohrolasa II no modificada correspondiente.

- 5 En una realización, se proporciona una GTP ciclohrolasa II modificada procedente de *Bacillus*, preferiblemente *Bacillus subtilis*, que muestra un incremento en la actividad específica de al menos alrededor de 10% en comparación con la enzima no modificada correspondiente, en la que se han sustituido los restos de aminoácidos en las posiciones correspondientes a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEC ID NO:2, conduciendo dichas sustituciones a restos de aminoácidos seleccionados del grupo que consiste en:
- 10 (a) alanina en una posición correspondiente a la posición 261 de SEC ID NO:2,  
 (b) alanina en una posición correspondiente a la posición 270 de SEC ID NO:2,  
 (c) treonina en una posición correspondiente a la posición 276 de SEC ID NO:2,  
 (d) arginina en una posición correspondiente a la posición 279 de SEC ID NO:2,  
 (e) arginina en una posición correspondiente a la posición 308 de SEC ID NO:2, y  
 15 (f) isoleucina en una posición correspondiente a la posición 347 de SEC ID NO:2.

En una realización, la enzima no modificada corresponde a GTP ciclohrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO:2. De este modo, la enzima modificada que tiene una mayor actividad específica en comparación con la enzima de tipo salvaje comprende una o más sustituciones, incluyendo 1, 2, 3, 4, 5, o 6, en las posiciones de aminoácidos que corresponden a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEC ID NO:2. En una realización adicional, la enzima modificada que tiene mayor actividad específica como se define anteriormente contiene una mutación o mutaciones de aminoácidos junto a las posiciones de aminoácidos como anteriormente, estando dicha mutación o mutaciones adicionales en una posición seleccionada del grupo que consiste en la posición 196, 282, 235 y cualquier combinación de las mismas, preferiblemente sustituciones de aminoácidos, más preferiblemente las sustituciones son Y196C (sustitución de tirosina por cisteína), A282T (sustitución de alanina por treonina) o F325Y (sustitución de fenilalanina por tirosina).

20

25

Una GTP ciclohrolasa II no modificada puede ser cualquier GTP ciclohrolasa II para la cual sea deseable incrementar la actividad específica. Las enzimas de GTP ciclohrolasa II modificadas incluyen, pero no se limitan a, enzimas de GTP ciclohrolasa II derivables de la naturaleza, tales como enzimas de origen eucariota o procariota, preferiblemente fúngico o bacteriano. Más preferiblemente, la enzima no modificada se selecciona de aquellas mostradas en la Figura 1 o en la Tabla 4, o que es homóloga a cualquiera de las secuencias de aminoácidos como se muestra en la Figura 1 o en la Tabla 4, en particular seleccionadas del grupo que consiste en *Ashbya*, *Saccharomyces*, *Eremothecium*, *Candida*, *Neurospora*, *Schizosaccharomyces*, *Archeoglobus*, *Streptomyces*, *Helicobacter*, *Escherichia*, *Corynebacterium*, *Thermotoga*, *Arabidopsis*, *Lycopersicum*, *Oryza*, *Alcaligenes*, *Pseudomonas*, *Dinococcus*, *Lactobacillus*, *Photobacterium* y *Bacillus* y preferiblemente seleccionadas del grupo que consiste en *Candida guilliermondii*, *Ashbya gossypii* (*Eremothecium ashbyii*) (SEC ID NO:33), *Saccharomyces cerevisiae*, *Neurospora crassa*, *Schizosaccharomyces pombe*, *Archeoglobus fulgidus*, *Streptomyces coelicolor*, *Helicobacter pylori* J99, *Escherichia coli* (SEC ID NO:35), *Corynebacterium glutamicum* (SEC ID NO:37), *Bacillus amyloliquefaciens* (SEC ID NO:39), *Bacillus cereus* (SEC ID NO:41), *Bacillus halodurans* (SEC ID NO:43), *Thermotoga maritima*, *Arabidopsis thaliana*, *Lycopersicum exculentum*, *Oryza sativum*, *Alcaligenes eutrophus*, cepa KT2440 de *Pseudomonas putida*, *Corynebacterium efficiens*, *Deinococcus radiodurans*, *Lactobacillus plantarum*, *Photobacterium phosphoreum*, cepa KT2440 de *Pseudomonas putida* (segundo gen) y *Bacillus subtilis* (SEC ID NO:2). Muy preferiblemente, la enzima no modificada se obtiene de *Bacillus subtilis*.

30

35

40

La GTP ciclohrolasa II modificada de la invención se puede obtener mutando la GTP ciclohrolasa II no modificada correspondiente. En una realización, la enzima no modificada corresponde a la GTP ciclohrolasa II de *B. subtilis* mostrada en SEC ID NO:2, y la enzima modificada comprende una o más sustituciones de aminoácidos, incluyendo 1, 2, 3, 4, 5, o 6, en las posiciones de los aminoácidos 261, 270, 276, 279, 308 y/o 347 de SEC ID NO:2, en la que la actividad específica de dicha enzima modificada está incrementada en comparación con la enzima no modificada, como se define aquí.

45

Preferiblemente, la al menos una sustitución es en una o más posiciones de aminoácidos seleccionadas del grupo que consiste en posiciones de aminoácidos que corresponden a las posiciones 261, 279, 308 y 347 de la secuencia de aminoácidos de GTP ciclohrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO:2. De este modo, en una realización, la GTP ciclohrolasa II modificada comprende una o más sustituciones, incluyendo 1, 2, 3 ó 4, en las posiciones de aminoácidos que corresponden a las posiciones 261, 279, 308 y/o 347 de SEC ID NO:2. En una realización preferida, la enzima modificada se obtiene de *B. subtilis* y comprende las posiciones de aminoácidos sustituidas 261, 279, 308 y/o 347 como se muestra en SEC ID NO:2, correspondientes a los aminoácidos V261, Q279, K308, y M374, respectivamente.

50

55

En otra realización preferida, la al menos una sustitución es en una o más posiciones de aminoácidos seleccionadas del grupo que consiste en posiciones de aminoácidos que corresponden a las posiciones 270, 279, 308 y 347 de la secuencia de aminoácidos de GTP ciclohidrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO:2. De este modo, en una realización, la GTP ciclohidrolasa II modificada comprende una o más sustituciones, incluyendo 1, 2, 3 ó 4, en las posiciones de aminoácidos correspondientes a las posiciones 270, 279, 308 y/o 347 de SEC ID NO:2. Preferiblemente, la enzima modificada se obtiene de *B. subtilis*, y comprende las posiciones de aminoácidos sustituidas 270, 279, 308 y/o 347 como se muestra en SEC ID NO:2, correspondientes a los aminoácidos G270, Q279, K308, y M374, respectivamente.

En una realización adicional preferida, la al menos una sustitución es en una o más posiciones de aminoácidos seleccionadas del grupo que consiste en las posiciones de aminoácidos correspondientes a las posiciones 276, 279, 308 y 347 de la secuencia de aminoácidos de GTP ciclohidrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO:2. De este modo, en una realización adicional, la GTP ciclohidrolasa II modificada comprende una o más sustituciones, incluyendo 1, 2, 3 ó 4, en las posiciones de aminoácidos que corresponden a las posiciones 276, 279, 308 y/o 347 de la secuencia de aminoácidos de GTP ciclohidrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO:2. Preferiblemente, la enzima modificada es obtenible de *B. subtilis*, y comprende las posiciones de aminoácidos sustituidas 276, 279, 308 y/o 347 como se muestra en SEC ID NO:2, correspondientes a los aminoácidos A276, Q279, K308, y M374, respectivamente.

Una GTP ciclohidrolasa II modificada puede comprender una o más sustituciones, incluyendo sólo una, en una posición de aminoácido como se define anteriormente, tal sustitución de aminoácido, puede incluir una sustitución en una posición de aminoácidos que corresponde a la posición 261, 270, 276, 279, 308, o 347 de la secuencia de aminoácidos de GTP ciclohidrolasa II de *Bacillus subtilis* como se muestra en SEC ID NO:2. El aminoácido presente en la GTP ciclohidrolasa II no modificada correspondiente a la posición 261 puede ser valina, el aminoácido presente en la GTP ciclohidrolasa II no modificada correspondiente a la posición 270 puede ser glicina, el aminoácido presente en la GTP ciclohidrolasa II no modificada correspondiente a la posición 276 puede ser alanina, el aminoácido presente en la GTP ciclohidrolasa II no modificada correspondiente a la posición 279 puede ser glutamina, el aminoácido presente en la GTP ciclohidrolasa II no modificada correspondiente a la posición 308 puede ser lisina, y el aminoácido presente en la GTP ciclohidrolasa II no modificada correspondiente a la posición 347 puede ser metionina.

El aminoácido en la secuencia de la GTP ciclohidrolasa II no modificada se cambia de manera que el aminoácido correspondiente a la posición 261 se cambia a alanina (*por ejemplo* V261A), el aminoácido correspondiente a la posición 270 se cambia a alanina (*por ejemplo* G270A), el aminoácido correspondiente a la posición 276 se cambia a treonina (*por ejemplo* A276T), el aminoácido correspondiente a la posición 279 se cambia a arginina (*por ejemplo* Q279A), el aminoácido correspondiente a la posición 308 se cambia a arginina (*por ejemplo* K308R), y el aminoácido correspondiente a la posición 347 se cambia a isoleucina (*por ejemplo* M374I). En una realización, la enzima modificada se obtiene de *B. subtilis* que comprende una sustitución de aminoácido en una posición de SEC ID NO:2 que se selecciona del grupo que consiste en la posición 261, 270, 276, 279, 308, y 347. Preferiblemente, la sustitución se selecciona del grupo que consiste en V261A, G270A, A276T, Q279R, K308R y M347I.

Una GTP hidrolasa II modificada puede comprender una o más sustituciones, incluyendo dos, en posiciones de aminoácidos como se definen anteriormente, tales sustituciones de aminoácidos, pueden incluir sustituciones en las posiciones de aminoácidos que corresponden a dos de las posiciones como se definen anteriormente, por ejemplo combinaciones de posiciones que corresponden a las posiciones 261/270, 261/276, 261/279, 261/308, 261/347, 270/276, 270/279, 270/308, 270/347, 276/279, 276/308, 276/347, 279/308, 279/347, o 308/347 como se muestra en SEC ID NO:2. Se prefieren las sustituciones de aminoácidos tales como V261A/A276T, V261A/Q279R, V261A/K308R, V261A/M347I, G270A/Q279R, G270A/K308R, G270A/M347I, A276T/Q279R, A276T/K308R, o A276T/M347I, en las que las posiciones corresponden a las posiciones de aminoácidos de SEC ID NO:2. En una realización, tales sustituciones preferidas están comprendidas en una GTP ciclohidrolasa II modificada obtenible de *B. subtilis* en la que la enzima no modificada corresponde a SEC ID NO:2. Preferiblemente, la GTP hidrolasa II de *B. subtilis* modificada como en SEC ID NO:2 comprende las sustituciones V261A/A276T o A276T/M347I.

Una GTP hidrolasa II modificada puede comprender una o más sustituciones, incluyendo tres, en posiciones de aminoácidos como se definen anteriormente, tales sustituciones de aminoácidos, pueden incluir sustituciones en las posiciones de aminoácidos que corresponden a tres de las posiciones como se definen anteriormente, en particular combinaciones de posiciones que corresponden a las posiciones 261/279/308, 261/279/347, 261/308/347, 270/279/308, 270/279/347, 270/308/347, 276/279/308, 276/308/347, o 276/279/347 como se muestra en SEC ID NO:2. Se prefieren las sustituciones de aminoácidos tales como V261A/Q279R/K308R, V261A/K308R/M347I, V261A/Q279R/M347I, G270A/Q279R/K308R, G270A/K308R/M347I, G270A/Q279R/M347I, A276T/Q279R/K308R, A276T/K308R/M347I, o A276T/Q279R/M347I, en las que las posiciones corresponden a las posiciones de aminoácidos de SEC ID NO:2. En una realización, tales sustituciones preferidas están comprendidas en una GTP ciclohidrolasa II modificada obtenible de *B. subtilis* en la que la enzima no modificada corresponde a SEC ID NO:2. Preferiblemente, la GTP hidrolasa II de *B. subtilis* modificada como en SEC ID NO:2 comprende las sustituciones A276T/Q279R/M347I.

Una GTP hidrolasa II modificada puede comprender una o más sustituciones, incluyendo cuatro, en posiciones de aminoácidos como se definen anteriormente, tales sustituciones de aminoácidos, pueden incluir sustituciones en las posiciones de aminoácidos que corresponden a cuatro de las posiciones como se definen anteriormente, en particular combinaciones de posiciones que corresponden a las posiciones 261/279/308/347, 270/279/308/347, or 276/279/308/347 como se muestra en SEC ID NO:2. Se prefieren las sustituciones de aminoácidos tales como V261A/Q279R/K308R/M347I, G270A/Q279R/K308R/M347I o A276T/Q279R/K308R/M347I, en las que las posiciones corresponden a las posiciones de aminoácidos de SEC ID NO:2. En una realización, tales sustituciones preferidas están comprendidas en una GTP ciclohidrolasa II modificada obtenible de *B. subtilis* en la que la enzima no modificada corresponde a SEC ID NO:2. Preferiblemente, la GTP hidrolasa II de *B. subtilis* modificada como en SEC ID NO:2 comprende las sustituciones A276T/Q279R/K308R/M347I.

Las más preferidas son las combinaciones de sustituciones descritas en la Tabla 1 ó 2 (véase más arriba). Las posiciones de aminoácidos identificadas en estos ejemplos se pueden transferir a enzimas de GTP ciclohidrolasa II de diferente origen, como por ejemplo se muestra en la Figura 1 o en la Tabla 4.

La GTP ciclohidrolasa II modificada de la invención puede comprender aminoácidos extraños, preferiblemente en su término N o C. "Aminoácidos extraños" significa aminoácidos que no están presentes en una GTP ciclohidrolasa II nativa (que aparece en la naturaleza), preferiblemente un tramo de al menos alrededor de 3, al menos alrededor de 5 o al menos alrededor de 7 aminoácidos contiguos que no están presentes en una GTP ciclohidrolasa II nativa. Los tramos preferidos de aminoácidos extraños incluyen, pero no se limitan a, "etiquetas" que facilitan la purificación de la GTP ciclohidrolasa II modificada producida recombinantemente. Los ejemplos de tales etiquetas incluyen, pero no se limitan a, una etiqueta "His<sub>6</sub>", una etiqueta FLAG, una etiqueta myc, y similar. Para el cálculo de la actividad específica, es necesario que los valores sean corregidos para estos aminoácidos adicionales (véase también anteriormente).

En otra realización, la GTP ciclohidrolasa II modificada puede contener una o más, por ejemplo dos, supresiones cuando se compara con la secuencia de aminoácidos de la GTP ciclohidrolasa II no modificada correspondiente. Preferiblemente, las supresiones afectan a los aminoácidos N o C-terminales de la GTP ciclohidrolasa II no modificada correspondiente, y no reducen significativamente las propiedades funcionales, por ejemplo la actividad específica, de la enzima.

Los polipéptidos y polinucleótidos de la presente invención, incluyendo las enzimas GTP ciclohidrolasa II modificadas, se pueden proporcionar en forma aislada, y preferiblemente están purificadas hasta homogeneidad. Como se usa aquí, el término "aislada" significa que el material se retira de su entorno original (*por ejemplo*, el entorno natural si es de origen natural). Por ejemplo, un polinucleótido o polipéptido de origen natural presente en un microorganismo vivo no está aislado, pero el mismo polinucleótido o polipéptido, separado de parte o de todos los materiales coexistentes en el sistema natural, está aislado. Tales polinucleótidos podrían ser parte de un vector, y/o tales polinucleótidos o polipéptidos podrían ser parte de una composición y todavía estar aislados por cuanto tal vector o composición no es parte de su entorno natural. Un polipéptido aislado es preferiblemente mayor que 80% puro, más preferiblemente mayor que 90% puro, incluso más preferiblemente mayor que 95% puro, lo más preferible mayor que 99% puro. La pureza se puede determinar según métodos conocidos en la técnica, *por ejemplo* mediante SDS-PAGE y tinción proteica subsiguiente. Las bandas de proteínas se pueden cuantificar entonces mediante densitometría. Otros métodos para determinar la pureza están dentro del nivel de la pericia normal.

La invención se refiere además a un polinucleótido que comprende una secuencia nucleotídica que codifica una GTP ciclohidrolasa II modificada según la invención. "Polinucleótido", como se usa aquí, se refiere a un polirribonucleótido o polidesoxirribonucleótido que puede ser ARN o ADN sin modificar, o ARN o ADN modificado. Los polinucleótidos incluyen, pero no se limitan a, ADN mono- o bicatenario, ADN que es una mezcla de regiones mono- y bicatenarias, ARN mono- y bicatenario, y ARN que es una mezcla de regiones mono- y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser monocatenarios o, más típicamente, bicatenarios, o una mezcla de regiones mono- y bicatenarias. El término "polinucleótido" incluye ADN o ARN que comprende una o más bases inusuales, *por ejemplo* inosina, o una o más bases modificadas, *por ejemplo* bases tritiladas.

El polinucleótido de la invención puede obtenerse fácilmente modificando una secuencia polinucleotídica que codifica una GTP ciclohidrolasa II no modificada. Los ejemplos de tales secuencias polinucleotídicas que codifican enzimas de GTP ciclohidrolasa II no modificadas incluyen, pero no se limitan a, las secuencias de aminoácidos de la Figura 1 o en la Tabla 4, en particular las SEC ID NOs:2, 39, 41, y 43. Los ejemplos no limitantes de polinucleótidos que codifican enzimas de GTP ciclohidrolasa II modificadas según la invención se muestran en SEC ID NOs:6, 8, 10, 12, 14, 16, 18, 20, 22, 24 y 26.

Los métodos para introducir mutaciones, *por ejemplo* adiciones, supresiones y/o sustituciones, en la secuencia nucleotídica que codifica la GTP ciclohidrolasa II no modificada incluyen, pero no se limitan a, mutagénesis dirigida al sitio y métodos a base de PCR.

Las secuencias de ADN de la presente invención se pueden construir partiendo de secuencias de ADN genómico o de ADNc que codifican enzimas GTP ciclohidrolasa II conocidas en el estado de la técnica, como están disponibles a partir de, *por ejemplo*, Genbank (Intelligenetics, California, USA), European Bioinformatics Institute (Hinton Hall,

Cambridge, GB), NBRF (Georgetown University, Medical Centre, Washington DC, USA) y Vecbase (University of Wisconsin, Biotechnology Centre, Madison, Wisconsin, USA), o a partir de una información de secuencia descrita en la Figura 1 o en la Tabla 4 mediante métodos de mutagénesis *in vitro* [véase, *por ejemplo*, Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory Press, Nueva York]. Otra posibilidad de mutar una secuencia de ADN dada que también puede ser adecuada para la práctica de la presente invención es la mutagénesis usando la reacción en cadena de la polimerasa (PCR). El ADN como material de partida se puede aislar mediante métodos conocidos en la técnica y descritos, *por ejemplo*, en Sambrook et al. (Molecular Cloning) a partir de las cepas/organismos respectivos. Sin embargo, se entiende que el ADN que codifica una GTP ciclohidrolasa II a construir/mutar según la presente invención también se puede preparar en base a una secuencia de ADN producida, *por ejemplo* mediante construcción de un gen sintético mediante métodos conocidos en la técnica (como se describe, *por ejemplo*, en el documento EP 747483).

El polinucleótido de la invención puede ser un polinucleótido aislado, es decir, un polinucleótido que está sustancialmente libre de otras secuencias de ácidos nucleicos, tales como, pero sin limitarse a, otro ADN y ARN cromosómico y extracromosómico. Para obtener polinucleótidos aislados, se pueden usar métodos de purificación de ácidos nucleicos convencionales, conocidos por la persona experta en la técnica. El término también abarca polinucleótidos recombinantes y polinucleótidos sintetizados químicamente.

En todavía otras realizaciones, la invención se refiere a un polinucleótido funcional en el que un promotor, un sitio de unión al ribosoma, si es necesario como en el caso de células bacterianas, y un terminador están operablemente enlazados con un polinucleótido según la invención. En todavía una realización adicional, la invención se refiere a un vector o plásmido que comprende tal polinucleótido. El vector o plásmido comprende preferiblemente al menos un gen marcador. La expresión "operablemente enlazado", como se usa aquí, se refiere a la asociación de secuencias de ácidos nucleicos en un único fragmento de ácido nucleico de manera que la función de una se ve afectada por la otra. Por ejemplo, un promotor está operablemente enlazado con una secuencia codificante cuando es capaz de afectar a la expresión de esa secuencia codificante, es decir, la secuencia codificante está bajo el control transcripcional del promotor. Las secuencias codificantes pueden estar operablemente enlazadas a secuencias reguladoras en orientación sentido o antisentido. El término "expresión" se refiere a la transcripción de una secuencia de ADN en ARNm y/o a la traducción de ARNm en una secuencia de aminoácidos. El término "sobrexpresión" significa la producción de un producto génico en un organismo modificado (*por ejemplo*, modificado mediante transformación o transfección) que excede los niveles de producción en el organismo no modificado correspondiente desregulando la expresión del gen y/o multiplicando el propio gen dentro del organismo.

Una vez que se han obtenido secuencias de ADN completas de la presente invención, se pueden integrar en vectores o se pueden introducir directamente en el genoma de un organismo hospedante mediante métodos conocidos en la técnica y descritos en, *por ejemplo*, Sambrook et al. (s.a.) para (sobre)expresar en sistemas hospedantes apropiados el polipéptido codificado. Sin embargo, un experto en la técnica sabe que también las propias secuencias de ADN se pueden usar para transformar los sistemas hospedantes adecuados de la invención para obtener la (sobre)expresión del polipéptido codificado.

Las células hospedantes adecuadas pueden ser células eucariotas o procariotas. Los ejemplos de células hospedantes adecuadas incluyen, pero no se limitan a, células bacterianas tales como células de cianobacterias, estreptococos, estafilococos, enterococos, *por ejemplo* bacilos como, *por ejemplo*, *Bacillus subtilis*, o *Streptomyces*, como *por ejemplo* *Streptomyces lividans* o *Streptococcus pneumoniae*, *E. coli* como, *por ejemplo* cepas de *E. coli* K12, *por ejemplo* M15 o HB 101. Las células hospedantes pueden ser una célula fúngica, incluyendo células de levadura, tales como células de *Aspergilli*, *por ejemplo* *Aspergillus niger* o *Aspergillus oryzae*, *Trichoderma*, *por ejemplo* *Trichoderma reesei*, *Ashbya*, *por ejemplo* *Ashbya gossypii*, *Eremothecium*, *por ejemplo* *Eremothecium ashbyii*, *Saccharomyces*, *por ejemplo* *Saccharomyces cerevisiae*, *Candida*, *por ejemplo* *Candida flareri*, *Pichia*, *por ejemplo* *Pichia pastoris*, *Hansenula polymorpha*, *por ejemplo* *H. polymorpha* (DSM 5215), y *Kluyveromyces*. Una célula hospedante adecuada se puede seleccionar además de células animales, incluyendo células de mamíferos, tales como *por ejemplo* CHO, COS, HeLa, 3T3, BHK, 293, CV-1 y células de insectos como células de *Drosophila* S2 y *Spodoptera* Sf9; y células vegetales tales como células de una gimnosperma o angiosperma.

Los vectores que se pueden usar para la expresión en hongos son conocidos en la técnica y se describen, *por ejemplo*, en el documento EP 420358, o por Cullen et al. (Bio/Technology 5, 369-376, 1987), Ward (en Molecular Industrial Mycology, Systems and Applications for Filamentous Fungi, Marcel Dekker, Nueva York, 1991), Upshall et al. (Bio/Technology 5, 1301-1304, 1987), Gwynne et al. (Bio/Technology 5, 71-79, 1987), o Punt et al. (J. Biotechnol. 17, 19-34, 1991), y para levadura, por Sreekrishna et al. (J. Basic Microbiol. 28, 265-278, 1988; Biochemistry 28, 4117-4125, 1989), Hitzemann et al. (Nature 293, 717-722, 1981) o en los documentos EP 183070, EP 183071, EP 248227, o EP 263311. Los vectores adecuados que se pueden usar para la expresión en *E. coli* son conocidos en la técnica, como se describe por Sambrook et al. (s.a.). Los vectores que se pueden usar para la expresión en Bacilli son conocidos en la técnica y se describen, *por ejemplo*, en los documentos EP 207459 o EP 405370, por Yansura y Henner en Proc. Natl. Acad. Sci. USA 81, 439-443 (1984), o por Henner, Le Grice y Nagarajan en Meth. Enzymol. 185, 199-228, 1990. Los vectores que se pueden usar para la expresión en *H. polymorpha* son conocidos en la técnica, como se describe, *por ejemplo*, en Gellissen et al., Biotechnology 9, 291-295, 1991.



5 Cualquiera de tales vectores ya poseen elementos reguladores, *por ejemplo* promotores, o los polinucleótidos de la presente invención se pueden manipular mediante ingeniería para que contengan tales elementos. Los elementos promotores adecuados que se pueden usar son conocidos en la técnica, y son, *por ejemplo*, para *Trichoderma reesei*, el promotor cbh1 o pki1, para *Aspergillus oryzae* el promotor amy, y para *Aspergillus niger*, el promotor glaA, alcA, aphA, tpiA, gpdA y pkiA. Los elementos promotores adecuados que se pueden usar para la expresión en levadura son conocidos en la técnica, y son, *por ejemplo*, el promotor pho5 o gap para la expresión en *Saccharomyces cerevisiae*, y *por ejemplo* el promotor aox1 para *Pichia pastoris*, o el promotor FMD o MOX para *H. polymorpha*.

10 Los promotores y vectores adecuados para la expresión bacteriana incluyen, *por ejemplo*, un promotor sintético descrito por Giacomini et al. (Gene 144, 17-24, 1994), el promotor vegI de *Bacillus subtilis*, o el promotor T5 de bacteriófago fuerte. Las enseñanzas apropiadas para la expresión de las enzimas GTP ciclohidrolasa II (mutantes) reivindicadas en bacterias, ya sea mediante plásmidos apropiados o a través de la integración de secuencias de ADN que codifican GTP ciclohidrolasa II en el ADN cromosómico, se pueden encontrar en muchos lugares, *por ejemplo* la patente US nº 6.322.995.

15 En consecuencia, los vectores que comprenden un polinucleótido de la presente invención, preferiblemente para la expresión de dichos polinucleótidos en hospedantes bacterianos, fúngicos, animales o vegetales, y tales hospedantes bacterianos o fúngicos, animales o vegetales transformados también son un objeto de la presente invención.

20 La invención se refiere además a un método para producir riboflavina, un precursor de riboflavina, FMN, FAD, o uno o más de sus derivados, que comprende:

(a) cultivar la célula hospedante de la invención en un medio adecuado en condiciones que permitan la expresión en dicha célula hospedante de la GTP ciclohidrolasa II modificada; y

(b) opcionalmente separar el producto (riboflavina, un precursor de riboflavina, FMN, FAD, o uno o más de sus derivados) del medio.

25 Tal método se puede usar para la producción biotecnológica de uno o más de los siguientes productos: riboflavina, un precursor de riboflavina, FMN, FAD, o uno o más de sus derivados. Tales derivados pueden incluir flavoproteínas.

30 Los métodos de manipulación genética y metabólica mediante ingeniería de células hospedantes adecuadas según la presente invención son conocidos por la persona experta en la técnica. De forma similar, los métodos de purificación (potencialmente) adecuados para riboflavina, un precursor de riboflavina, FMN, FAD, o uno o más de sus derivados, son bien conocidos en el campo de la biosíntesis química fina y de la producción.

35 Se entiende que los métodos para la producción biotecnológica de un producto de riboflavina, un precursor de riboflavina, FMN, FAD, o uno o más de sus derivados según la presente invención no están limitados a procedimientos de fermentación de células completas como se describe anteriormente, sino que también pueden usar, *por ejemplo*, células hospedantes permeabilizadas, extractos celulares brutos, extractos celulares aclarados de restos celulares mediante, *por ejemplo*, centrifugación o filtración, o incluso rutas de reacción reconstituidas con enzimas aisladas. También las combinaciones de tales procedimientos están dentro del alcance de la presente invención. En el caso de la biosíntesis libre de células (tal como con rutas de reacción reconstituidas), es irrelevante si las enzimas aisladas se han preparado mediante y aislado de una célula hospedante, mediante transcripción/traducción in vitro, o todavía por otros medios.

40 La invención se refiere además a un método para producir una GTP ciclohidrolasa II modificada de la invención, que comprende:

(a) cultivar una célula hospedante de la invención en condiciones que permitan la expresión de la GTP ciclohidrolasa II modificada de la invención; y

45 (b) recuperar la GTP ciclohidrolasa II modificada a partir de células o a partir del medio.

Las enzimas de GTP ciclohidrolasa II modificadas de la invención se pueden preparar a partir de células hospedantes manipuladas genéticamente mediante ingeniería que comprenden sistemas de expresión apropiados.

50 Para la producción recombinante de los polipéptidos de la invención, las células hospedantes se pueden manipular genéticamente mediante ingeniería para incorporar polinucleótidos o vectores o plásmidos de la invención. La introducción de un polinucleótido o vector en la célula hospedante se puede efectuar mediante métodos estándar conocidos en la técnica, tales como transfección con fosfato de calcio, transfección mediada por DEAE-dextrano, microinyección, transfección mediada por lípidos catiónicos, electroporación, transducción, introducción balística e infección.

Para producir las enzimas de GTP ciclohidrolasa II modificadas de la invención, se puede usar una gran variedad de sistemas de expresión. Tales vectores incluyen, entre otros, los descritos más arriba. Generalmente, a este respecto, para la expresión, se puede usar cualquier sistema o vector adecuado para mantener, propagar o expresar polinucleótidos y/o para expresar un polipéptido en un hospedante.

5 En sistemas de expresión recombinantes en eucariotas, para la secreción de una proteína traducida a la luz del retículo endoplásmico, al espacio periplásmico o al medio extracelular, se pueden incorporar en el polipéptido expresado señales de secreción apropiadas. Estas señales pueden ser endógenas al polipéptido, o pueden ser señales heterólogas.

10 Los polipéptidos de la invención se pueden recuperar y purificar a partir de cultivos de células recombinantes mediante métodos bien conocidos que incluyen precipitación con sulfato de amonio o etanol, extracción ácida, cromatografía de intercambio aniónico o catiónico, cromatografía con fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía con hidroxapatita, y cromatografía de líquidos de altas prestaciones. Cuando el polipéptido se desnaturaliza durante el aislamiento y/o purificación, para regenerar la conformación activa se pueden emplear técnicas bien conocidas para el repliegamiento proteico.

15 Las enzimas de GTP ciclohidrolasa II de la presente invención también se pueden expresar en plantas según métodos como se describen, *por ejemplo*, por Pen et al. en *Bio/Technology* 11, 811-814, 1994 o en el documento EP 449375, preferiblemente en semillas, como se describe, *por ejemplo*, en el documento EP 449376. Algunos ejemplos adecuados de promotores y terminadores incluyen aquellos de los genes de nopalina sintasa (*nos*), de octopina sintasa (*ocs*) y del virus del mosaico de la coliflor (CaMV). Un tipo de promotor vegetal eficiente que se puede usar es un promotor vegetal de alto nivel. Tales promotores, en enlace operable con las secuencias genéticas de la presente invención, deberían ser capaces de promover la expresión de un producto génico de la presente invención. Los promotores vegetales de alto nivel que se pueden usar en esta invención incluyen el promotor de la subunidad pequeña (ss) de la ribulosa-1,5-bisfosfato carboxilasa, por ejemplo de haba de soja, y el promotor de la proteína de unión a clorofila *a/b*.

25 Cuando se desea una producción comercial de las actuales proteínas, se puede aplicar una variedad de metodologías de cultivo. Por ejemplo, la producción a gran escala de un producto génico específico, sobreexpresado a partir de un hospedante microbiano recombinante, se puede lograr tanto mediante metodologías de cultivo discontinuas o continuas. Los métodos de cultivo discontinuos o alimentado discontinuo son habituales y bien conocidos en la técnica, y los ejemplos se han descrito por Thomas D. Brock en *Biotechnology: A Textbook of Industrial Microbiology*, Segunda Edición (1989), Sinauer Associates, Inc., Sunderland, Mass., o Deshpande, *Appl. Biochem. Biotechnol.* 36, 227-234, 1992. Los métodos para modular nutrientes y factores de crecimiento para procesos de cultivo continuos, así como las técnicas para maximizar la velocidad de formación de productos, son bien conocidos en la técnica de la microbiología industrial, y en Brock, más arriba, se detalla una variedad de métodos.

35 Los medios de fermentación pueden contener además sustratos de carbono adecuados. Los sustratos adecuados pueden incluir, pero no se limitan a, monosacáridos tales como glucosa o fructosa, oligosacáridos tales como lactosa o sacarosa, polisacáridos tales como almidón o celulosa, o sus mezclas, y mezclas sin purificar a partir de materias primas renovables. Se contempla que la fuente de carbono utilizada en la presente invención puede englobar una amplia variedad de sustratos que contienen carbono, y sólo estará limitada por la elección del organismo.

40 La invención se refiere además a un método para la preparación de una GTP ciclohidrolasa II que tiene una mayor actividad específica, que comprende las siguientes etapas:

(a) proporcionar un polinucleótido que codifica una primera GTP ciclohidrolasa II con una actividad específica que se va a incrementar en al menos alrededor de 10%;

45 (b) introducir una o más sustituciones en la secuencia polinucleotídica de manera que la secuencia polinucleotídica modificada codifica una GTP ciclohidrolasa II modificada seleccionada de *Bacillus* que muestra un incremento en la actividad específica de al menos alrededor de 10% en comparación con la enzima no modificada correspondiente, en la que se han sustituido restos de aminoácidos en las posiciones que corresponden a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEC ID NO:2, conduciendo por ejemplo a sustituciones a restos de aminoácidos seleccionados del grupo que consiste en:

- 50
- (a) alanina en una posición correspondiente a la posición 261 de SEC ID NO:2,
  - (b) alanina en una posición correspondiente a la posición 270 de SEC ID NO:2,
  - (c) treonina en una posición correspondiente a la posición 276 de SEC ID NO:2,
  - (d) arginina en una posición correspondiente a la posición 279 de SEC ID NO:2,
  - (e) arginina en una posición correspondiente a la posición 308 de SEC ID NO:2, y

(f) isoleucina en una posición correspondiente a la posición 347 de SEC ID NO:2

(c) insertar opcionalmente el polinucleótido modificado en un vector o plásmido;

(d) introducir el polinucleótido o el vector o plásmido en una célula hospedante adecuada; y

5 (e) cultivar la célula hospedante en condiciones que permiten la expresión de la GTP ciclohidrolasa II modificada.

En una realización, la etapa (c) o el método descrito anteriormente se lleva a cabo vía mutagénesis saturada. Sin embargo, se entiende que éste puede no ser el único camino para definir el aminoácido que debería sustituir un aminoácido en una posición dada de la GTP ciclohidrolasa II de tipo salvaje a fin de obtener una GTP ciclohidrolasa II modificada con mayor actividad específica, como se define aquí.

10 La preparación de una GTP ciclohidrolasa II modificada que tiene mayor actividad específica a partir de una GTP ciclohidrolasa II no modificada como se describe anteriormente, por ejemplo vía mutagénesis saturada, incluye, pero no se limita a, la preparación de proteínas de GTP ciclohidrolasa II mutadas a partir de proteínas no modificadas como la Figura 1 o en la Tabla 4, en particular aquellas identificadas mediante SEC ID NOs:2, 39, 41, y 43, tales como por ejemplo las proteínas de GTP ciclohidrolasa II no modificadas de *Bacillus subtilis*. Los cebadores usados para la reacción de PCR son tales que un cebador, *por ejemplo* el cebador sentido, puede contener una secuencia nucleotídica mutada, y el otro cebador, *por ejemplo* el cebador antisentido, puede contener la secuencia nucleotídica de tipo salvaje. La PCR con estos pares de cebadores y ADN genómico del *ribA* de tipo salvaje puede dar como resultado un producto de PCR que posee la mutación particular en una posición dada, dependiendo de la secuencia nucleotídica mutada del cebador usado. Tras la purificación de los productos de la PCR resultantes usando métodos estándar como, *por ejemplo*, el kit de purificación mediante PCR QIAquick (Qiagen), el ADN se puede cortar con enzimas de restricción tales como *Bam*HI y *Eco*RI, se puede ligar en un vector adecuado, *por ejemplo* pQE60, y se puede transformar en una cepa que es negativa para GTP ciclohidrolasa II. Un ejemplo de tal cepa es la cepa Rib7 de *E. coli* (Richter et al., J. Bacteriol. 175, 4045-4051, 1993) que contiene el plásmido pREP4. Tras la continuación de la secuencia correcta mediante secuenciación de ADN, la RibA mutada se puede purificar y caracterizar como se describe anteriormente.

Las realizaciones preferidas de este método corresponden a las realizaciones preferidas de la GTP ciclohidrolasa II modificada, los polinucleótidos que la codifican, los vectores y plásmidos, las células hospedantes, y los métodos descritos aquí. La GTP ciclohidrolasa II primera y segunda corresponden a la GTP ciclohidrolasa II no modificada y modificada, respectivamente (véase más arriba).

30 Es un objeto de la presente invención proporcionar un polinucleótido que comprende una secuencia de ácido nucleico que codifica una GTP ciclohidrolasa II modificada como se describe anteriormente, un vector, preferiblemente un vector de expresión, que comprende tal polinucleótido, una célula hospedante que se ha transformado mediante tal polinucleótido o vector, un procedimiento para la preparación de una GTP ciclohidrolasa II de la presente invención, en el que la célula hospedante como se describe anteriormente se cultiva en condiciones de cultivo adecuadas y la GTP ciclohidrolasa II se aísla de tal célula hospedante o del medio de cultivo mediante métodos conocidos en la técnica, y un procedimiento para la producción biotecnológico de riboflavina, un precursor de riboflavina, FMN, FAD, o uno o más derivados de los mismos, basado en una célula hospedante que se ha transformado mediante tal polinucleótido o vector, y/o que puede tener integrado de forma estable tal polinucleótido en su cromosoma o cromosomas.

40 Además, es un objeto de la presente invención proporcionar una secuencia de ADN que se puede obtener mediante el método denominado de reacción en cadena de la polimerasa ("PCR") mediante cebadores de PCR diseñados en base a las secuencias de ADN descritas específicamente de la presente invención. Se entiende que las secuencias de ADN así obtenidas codifican enzimas de GTP ciclohidrolasa II con al menos la misma distribución que aquellas a partir de las que se diseñan, y muestran propiedades de actividad comparables.

45 Las diversas realizaciones de la invención descritas aquí se pueden combinar de forma cruzada.

Figura 1: Alineamiento de múltiples secuencias calculado mediante el programa PILEUP del paquete de programa GCG de 92 secuencias de GTP ciclohidrolasa II encontradas por el programa BLASTN usando bases de datos estándar como SWISS-PROT y TrEMBL (*Candida guilliermondii*, *Ashbya gossypii*, *Saccharomyces cerevisiae*, *Neurospora crassa*, *Schizosaccharomyces pombe*, *Archaeoglobus fulgidus*, *Streptomyces coelicolor*, *Helicobacter pylori* J99, *Helicobacter pylori*, *Pyrococcus furiosus*, *Thermotoga maritima*, *Chlamydia muridarum*, *Chlamydia trachomatis*, *Chlamydia caviae* GPIC, *Arabidopsis thaliana*, *Lycopersicon esculentum*, *Oryza sativa*, *Alcaligenes eutrophus*, *Neisseria meningitidis* (serogrupo A), *Neisseria meningitidis* (serogrupo B, dos enzimas de GTP ciclohidrolasa II), *Pseudomonas putida* (dos enzimas de GTP ciclohidrolasa II), *Pseudomonas syringae* (dos enzimas de GTP ciclohidrolasa II), *Actinobacillus actinomycetemcomitans* (*Haemophilus actinomycetemcomitans*), *Haemophilus influenzae*, *Pasteurella multocida*, *Escherichia coli*, *Escherichia coli* O6, *Salmonella typhimurium*, *Yersinia pestis*, *Buchnera aphidicola* (subsp. *Acyrtosiphon pisum*) (*Acyrtosiphon pisum* bacteria simbiótica), *Buchnera aphidicola* (subsp. *Schizaphis graminum*), *Wigglesworthia glossinidia brevipalpis*, *Buchnera aphidicola* (subsp. *Baizongia pistaciae*), *Mycobacterium leprae*, *Mycobacterium tuberculosis*, *Corynebacterium efficiens*,

5 *Corynebacterium glutamicum*, *Corynebacterium ammoniagenes* (*Brevibacterium ammoniagenes*), *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Actinobacillus pleuropneumoniae*, *Lactococcus lactis* (*Streptococcus lactis*), *Streptococcus agalactiae*, *Streptococcus pneumoniae*, *Clostridium acetobutylicum*, *Fusobacterium nucleatum*,  
 10 *Anabaena spec.*, *Synechocystis spec.*, *Synechococcus elongatus* (*Thermosynechococcus elongatus*), *Bacillus amyloliquefaciens*, *Bacillus subtilis*, *Bacillus cereus*, *Bacillus halodurans*, *Clostridium perfringens*, *Clostridium tetani*, *Chlorobium tepidum*, *Aquifex aeolicus*, *Leptospira interrogans*, *Deinococcus radiodurans*, *Bacteroides thetaiotaomicron*, *Caulobacter crescentus*, *Coxiella burnetii*, *Rhizobium etli*, *Lactobacillus plantarum*, *Pseudomonas glumae*, *Streptomyces avermitilis*, *Photobacterium phosphoreum*, *Azospirillum brasilense*, *Agrobacterium tumefaciens*, *Rhizobium meliloti* (*Sinorhizobium meliloti*), *Brucella melitensis*, *Brucella suis*, *Rhizobium loti* (*Mesorhizobium loti*), *Nitrosomonas europaea*, *Ralstonia solanacearum* (*Pseudomonas solanacearum*),  
 15 *Xanthomonas axonopodis*, *Xanthomonas campestris*, *Vibrio parahaemolyticus*, *Vibrio vulnificus*, *Vibrio cholerae*, *Vibrio fischeri*, *Shewanella oneidensis*, *Photobacterium phosphoreum*, *Photobacterium leiognathi*, *Pseudomonas aeruginosa*, *Dehalospirillum multivorans*, *Xylella fastidiosa*). La numeración se refiere al alineamiento realizado. Algunas de las secuencias de aminoácidos codifican una enzima que tiene justo la actividad de GTP ciclohidrolasa II como las enzimas de *Ashbya gossypii*, *Streptomyces coelicolor*, *Helicobacter pylori* J99, *Helicobacter pylori*, *Arabidopsis thaliana*, *Alcaligenes eutrophus*, *Neisseria meningitidis* (serogrupo A), *Neisseria meningitidis* (serogrupo B), *Pseudomonas putida*, *Pseudomonas syringae*, *Actinobacillus actinomycetemcomitans* (*Haemophilus actinomycetemcomitans*), *Haemophilus influenzae*, *Pasteurella multocida*, *Escherichia coli*, *Escherichia coli* 06, *Salmonella typhimurium*, *Yersinia pestis*, *Buchnera aphidicola* (subsp. *Acyrtosiphon pisum*) (*Acyrtosiphon pisum* bacteria simbiótica), *Buchnera aphidicola* (subsp. *Schizaphis graminum*), *Wigglesworthia glossinidia brevipalpis*, *Buchnera aphidicola* (subsp. *Baizongia pistaciae*), *Pseudomonas glumae*, *Streptomyces avermitilis*, o *Photobacterium phosphoreum*. Otras enzimas como la enzima RibA de *B. subtilis* contienen, además, un dominio que tiene actividad de DHBP sintasa. La secuencia de aminoácidos de RibA procedente de *B. subtilis* está subrayada. Las posiciones que son homólogas/equivalentes a los restos de aminoácidos encontrados que tienen un efecto positivo sobre la actividad específica (restos de aminoácidos 261, 270, 276, 279, 308, 347) y sobre la sensibilidad a proteasas (196) de RibA de *B. subtilis*, y que se discuten en uno de los siguientes ejemplos, están en letras negritas. La numeración usada para esas posiciones se realiza según la secuencia de aminoácidos de tipo salvaje de *B. subtilis*. La figura comienza con el nombre de las secuencias usadas, el número de acceso de la base de datos, y entre paréntesis el organismo fuente de la secuencia.

30 Los siguientes ejemplos no limitantes ilustran adicionalmente la invención.

### Ejemplo 1: Medida de la actividad de GTP ciclohidrolasa II y determinación de la actividad específica

35 La actividad enzimática usada para medir la actividad de GTP ciclohidrolasa II se adaptó de (J. Biol. Chem. 276, 22273-22277, 2001). El tampón de ensayo final contenía 50 mM de Tris/HCl, pH 8,5, 10 mM de MgCl<sub>2</sub>, 7,5 mM de mercaptoetanol, 2,5 mM de GTP y 0,1 mg/ml de seroalbúmina bovina. Tras la purificación (véase el Ejemplo 5), la enzima se mantuvo en un tampón que contiene 50 mM de Tris/HCl, pH 8,5, 10 mM de MgCl<sub>2</sub>, 7,5 mM de mercaptoetanol, y 10% de glicerol. El sustrato se añadió a la enzima, y se siguió durante 20-30 min. la absorción a 310 nm, a la que GTP no muestra absorción. La mezcla de reacción final contenía entre 0,02 y 0,04 mg/ml de GTP ciclohidrolasa II procedente de *B. subtilis* o uno de los mutantes como se muestra en la Tabla 1 ó 2. Para el cálculo de la actividad, se usó un coeficiente de absorción de 6,28 [mM<sup>-1</sup> cm<sup>1</sup>] para DRAPP. La determinación de proteínas se realizó con el ensayo de proteínas de Bio-Rad (número de Catálogo 500-0002, Bio-Rad Laboratories AG, Nenzlingerweg 2, CH-4153 Reinach, Suiza).

45 Según la definición de "actividad específica" dada anteriormente, una unidad es la cantidad de RibA que cataliza la formación de 1 μmol de DRAPP por minuto en las condiciones como se describe anteriormente. La actividad específica es la actividad de DRAPP que se forma por 1 mg de RibA por minuto en las condiciones como se describe anteriormente. Usando las definiciones mencionadas anteriormente, la actividad específica de GTP ciclohidrolasa II de la proteína RibA etiquetada con His<sub>6</sub> de *B. subtilis* fue 0,115 U/mg.

### Ejemplo 2: Ensayo de la calidad del ensayo enzimático

50 Un ensayo óptimo debería cumplir un número de requisitos, tal como linealidad con la concentración enzimática y linealidad con el tiempo. Usando las condiciones descritas en el Ejemplo 1 y 22 μg de enzima, se siguió durante 25 min. el incremento en la absorción a 310 nm. Para ensayar en qué intervalo el ensayo es lineal con la concentración enzimática, se ensayó la dependencia del ensayo con la concentración enzimática creciente (0-40 μg de RibA etiquetada con His<sub>6</sub>). El ensayo demostró ser lineal a lo largo de 25 min. y entre 0 y 40 μg de RibA etiquetada con His<sub>6</sub> de *B. subtilis*.

55 Tras esto, se ensayó la dependencia de la actividad de GTP ciclohidrolasa II de RibA etiquetada con His<sub>6</sub> de *B. subtilis* sobre la concentración de GTP. Se usaron las condiciones como se describe en el Ejemplo 1. Sin embargo, la concentración de GTP varió de 0,05 a 2,5 mM de concentración final. Los datos indican un valor *K<sub>m</sub>* para GTP de 0,07 mM, y una actividad específica de alrededor de 115 mU/mg de proteína a 37°C para la actividad de GTP ciclohidrolasa II de la enzima RibA etiquetada con His<sub>6</sub> de *B. subtilis*. Los experimentos de este ejemplo mostraron que el ensayo de actividad de GTP ciclohidrolasa II es, de hecho, lineal con el tiempo y con la concentración de enzima (GTP ciclohidrolasa II), y que, en las condiciones dadas para GTP ciclohidrolasa II de *Bacillus subtilis*, una

concentración de GTP de 2,5 mM puede ser óptima para permitir medidas fiables de la actividad específica de la enzima.

### Ejemplo 3: Aislamiento de ADN genómico a partir de *Bacillus subtilis*

5 Se hizo crecer *B. subtilis* a 30°C en caldo de infusión de ternera lechal (Becton Dickinson, Sparks, MD 21152, USA) toda la noche. Se transfirieron 1,5 ml de cultivo a un tubo de 1,5 ml y se centrifugó. El pelete celular se resuspendió en 0,5 ml de tampón de suspensión (50 mM de Tris/HCl, pH 7,5, 50 mM de Na<sub>2</sub>EDTA, 15% de sacarosa y 1 mg/ml de lisozima recientemente añadida). Después de 10 minutos de incubación a temperatura ambiente, se añadió 1 µl de dietiloxidiformiato. Después, se añadieron 10 µl de disolución de SDS al 10% y el tubo se invirtió varias veces. El tubo se incubó durante 5 min. a 70°C para liberar el ADN bacteriano. Se añadieron 50 µl de acetato de potasio 5 M, el tubo se enfrió en hielo y se dejó allí durante 45 min. Después de esto, la muestra se centrifugó durante 30 min. a 4°C. El sobrenadante se transfirió a un nuevo tubo de 1,5 ml, que se llenó (hasta 1,5 ml) con etanol a temperatura ambiente. Después de 5 min. de centrifugación, el sobrenadante se desechó y el pelete de ADN se secó. Después, el ADN se lavó con etanol al 70% y al 96%, y se disolvió en 10 mM de Tris/HCl, pH 7,5, 1 mM de EDTA, y 10 µg/ml de RNase A.

### 15 Ejemplo 4: Construcción de los plásmidos de expresión para expresar *ribA* que codifica GTP ciclohidrolasa II y DHBP sintasa a partir de *B. subtilis* y sus mutantes

La clonación del gen *ribA* (SEC ID NO:1) de *B. subtilis* que codifica la GTP ciclohidrolasa II y la DHBP sintasa se realizó mediante PCR. El ADN genómico de *B. subtilis* se aisló según el Ejemplo 3. 100 ng de este ADN o de un molde que codifica una forma mutada del gen *ribA* se usaron para una PCR usando los cebadores RibA 1S (SEC ID NO:27) y RibA 1AS (SEC ID NO:28). Se usaron las siguientes condiciones de PCR: 2 µM de cada cebador, 0,2 mM de cada nucleótido, 2,5 U de una ADN polimerasa de corrección (Stratagene, Gebouw California, 1101 CB Amsterdam Zuidoost, Países Bajos), y 100 ng de ADN genómico en el tampón apropiado según se suministra junto con la ADN polimerasa.

La regulación de la temperatura fue la siguiente:

25 Etapa 1: 3 min. a 95°C

Etapa 2: 30 s a 95°C

Etapa 3: 30 s a 52°C

Etapa 4: 60 s a 72°C

Las etapas 2 a 4 se repitieron 30 veces.

30 El producto de la PCR de 1,3 kb se usó como molde para PCR 2, en la que el cebador RibA 1S se sustituyó por el cebador RibA 2S (SEC ID NO:29). El producto PCR de esta reacción (SEC ID NO:3), que codifica una versión N-terminalmente etiquetada con His<sub>6</sub> de RibA de *B. subtilis* (SEC ID NO:4), se separó mediante electroforesis en gel de agarosa, se eluyó del gel, se digirió con *EcoRI* y *BamHI*, y se ligó en el vector pQE60 digerido con *EcoRI* y *BamHI* (Qiagen AG, Hilden, Alemania). El plásmido se denominó pQE60ribANhis.

### 35 Ejemplo 5: Caracterización de las enzimas de tipo salvaje y mutantes

La generación de enzimas mutadas se llevó a cabo usando métodos descritos anteriormente y que son conocidos por la persona experta en la técnica. Los mutantes de RibA de *B. subtilis* que se investigaron adicionalmente se representan en la Tabla 1. Todos los genes mutantes se clonaron en un vector pQE60 como se describe en el Ejemplo 4. Todos los constructos finales contenían una etiqueta de His<sub>6</sub> N-terminal.

40 Tabla 1: Mutantes de Rib A como se definen mediante los intercambios de aminoácidos en comparación con la proteína RibA de tipo salvaje de *B. subtilis* (el número define las posiciones de aminoácidos respectivas en SEC ID NO:2)

Mutante	SEC ID NO (ADN)	SEC ID NO (proteína)
RibA M347I	9	10
RibA G270R	23	24
RibA K220E,G270A	19	20
RibA Y196C,A276T,A282T (PCR III)	11	12
RibA Y196C,A276T	5	6

Mutante	SEC ID NO (ADN)	SEC ID NO (proteína)
RibA Y196C,V261A	7	8
RibA Y196C,V261A,A276T	25	26
RibA Y196C,A276T,A282T,M347I	13	14
RibA Y196C,A276T,Q279R,A282T, M347I	15	16
RibA Y196C,A276T,Q279R,A282T, K308R,M347I (constructo C)	17	18
RibA Y196C,A276T,Q279R,A282T, K308R,F325Y,M347I (constructo E)	21	22

5 Las enzimas mutantes de RibA se expresaron a partir de los plásmidos del Ejemplo 4 y se purificaron como se describe en "The QiaExpressionist", Qiagen, Hilden, Alemania, Marzo de 2001, 5ª edición. Las propiedades enzimáticas de las enzimas purificadas (mutantes de RibA) se analizaron como se describe en los Ejemplos 1 y 2. La Tabla 2 compara las actividades específicas de GTP ciclohidrolasa II de los mutantes de RibA (véase la Tabla 1) con la de la GTP ciclohidrolasa II de la RibA de tipo salvaje de *B. subtilis*. La actividad se midió usando las versiones enzimáticas etiquetadas con His<sub>6</sub> N-terminalmente de RibA como se describe en el Ejemplo 4. Los números definen las posiciones de aminoácidos respectivas en SEC ID NO:2.

10 Tabla 2: Comparación de las actividades específicas de GTP ciclohidrolasa II de RibAs de *B. subtilis* mutadas y de tipo salvaje (WT) (todas etiquetadas con His<sub>6</sub> N-terminalmente)

Mutaciones	Actividad específica (en % con respecto a RibA de tipo salvaje)
WT	100
M347I	120
G270R	140
G270A,(K220E)	160
(Y196C),A276T,(A282T)	160
(Y196C),A276T	160
(Y196C),V261A	160
(Y196C),V261A,A276T	160
(Y196C),A276T,(A282T),M347I	180
(Y196C),A276T,Q279R, (A282T),M347I	185
(Y196C),A276T,Q279R,(A282T),K308R, M347I	201
(Y196C),A276T,Q279R,(A282T),K308R,(F325Y),M347I	200

15 Las sustituciones de aminoácidos entre paréntesis no tienen muy probablemente efecto sobre la actividad de GTP ciclohidrolasa II de los mutantes. El intercambio de aminoácidos Y196C reduce la sensibilidad de RibA a las proteasas.

**Ejemplo 6: Construcción de cepas de *B. subtilis* recombinantes que sobreexpresan mutantes de RibA que muestran una mayor actividad específica de GTP ciclohidrolasa II**

En el siguiente ejemplo, las secuencias polinucleotídicas de RibA mutado de Y196C,A276T,A282T (PCR III), RibA Y196C,A276T,Q279R,A282T,K308R,M347I

(constructo C), y RibA Y196C,A276T,Q279R,A282T,K308R,F325Y,M347I (constructo E) se introdujeron en primer lugar en un vector que contiene el promotor constitutivo fuerte  $P_{vegl}$ , y después se manipularon posteriormente en *E. coli*. La transformación de un microorganismo de *B. subtilis* competente natural con la secuencia polinucleotídica y las secuencias del vector de flaqueo da como resultado una cepa de *B. subtilis* que sobreexpresa el *ribA* mutado.

5 Para la construcción de la secuencia polinucleotídica y de las cepas de *B. subtilis* se usaron técnicas estándar de ADN recombinante. Véanse, por ejemplo, Sambrook et al., Molecular Cloning. A Laboratory Manual (2ª Ed.), Cold Spring Harbor Laboratory Press (1989), y Harwood y Cutting, Molecular Biology Methods for Bacillus, John Wiley and Sons (1990).

10 Para amplificar el *ribA* mutado, se amplificó mediante PCR, usando ADN procedente de un plásmido que contiene los mutantes PCR III, constructo C o constructo E, y como cebadores RibANde+1 (SEC ID NO:30) y RibA4AS (SEC ID NO:31), un fragmento de ADN de 1,2 kb que contiene toda la secuencia codificante de *ribA*.

15 Las condiciones de reacción para la reacción de PCR consistieron en 30 ciclos de desnaturalización a 95°C durante 1 min., recombinación a 52°C durante 1 min., y extensión a 72°C durante 2 min. Para minimizar los errores generados por la PCR, se usó la *Pfu* Turbo DNA polymerase (Stratagene, Gebouw California, 1101 CB Amsterdam Zuidoost, Países Bajos). Los productos de la PCR se purificaron usando el kit de purificación de PCR QIAquick (Qiagen), y se digirieron doblemente usando *NdeI* y *BamHI*. Los productos de la PCR digeridos se clonaron en el vector pXI16 (Huembelin et al., J. Ind. Microbiol. Biotechnol. 22, 1-7, 1999), que consiste en enzimas de restricción adecuadas para la clonación de secuencias polinucleotídicas inmediatamente en dirección 3' del promotor constitutivo fuerte  $P_{vegl}$  de *B. subtilis*. El vector pXI16 también contiene el terminador transcripcional *cryT* procedente de *B. thuringiensis*, las secuencias de flaqueo de *sacB* para la recombinación homóloga en el genoma de *B. subtilis* mediante un suceso de doble intercambio genético, y un marcador de resistencia a eritromicina. El hecho de que cada plásmido contuviese el *ribA* mutado se confirmó mediante secuenciación de ADN.

25 Cada plásmido se digirió con *Apal* para eliminar la región espaciadora del promotor  $P_{vegl}$ , se religó y se digirió nuevamente con *FspI*, y se transformó en células de *B. subtilis* 1012 competentes naturales. Los transformantes se seleccionaron en placas TBAB (triptosa sangre agar base, Becton Dickinson, Sparks, MD 21152, USA) que contiene eritromicina a una concentración final de 2 µg/ml. La secuenciación de ADN verificó que las secuencias polinucleotídicas de *ribA* mutadas fueron correctas en estas cepas. La sobreproducción de riboflavina se ensayó según el Ejemplo 7.

30 Las secuencias polinucleotídicas de *ribA* mutadas conducidas por el promotor  $P_{vegl}$  se introdujeron en la cepa RB50::(pRF69)<sub>n</sub>::(pRF93)<sub>m</sub> de *B. subtilis* que sobreproduce riboflavina, que se ha descrito en Perkins et al., J. Ind. Microbiol. Biotechnol. 22:8-18 (1999), mediante transducción generalizada. Las técnicas estándar usando el PBS 1 bacteriófago se emplearon según Harwood y Cutting, Molecular Biology Methods for Bacillus, John Wiley and Sons (1990). Los transductantes se seleccionaron en placas de TBAB que contienen eritromicina hasta una concentración final de 2 µg/ml. Los transformantes se comprobaron mediante análisis de PCR y secuenciación de ADN, para verificar la inserción correcta de la secuencia polinucleotídica de *ribA* mutada.

#### **Ejemplo 7: Producción mejorada de riboflavina usando una GTP ciclohrolasa II con actividad específica incrementada**

40 Para ensayar el efecto *in vivo* de mutaciones que afectan a la actividad específica de GTP ciclohrolasa II, se introdujeron los mutantes de GTP ciclohrolasa II (RibA) de *Bacillus subtilis* PCR III, constructo C, o constructo E en cepas de *B. subtilis* que sobreproducen riboflavina, tales como la cepa RB50::(pRF69)<sub>n</sub>::(pRF93)<sub>m</sub> (Perkins et al., J. Ind. Microbiol. Biotechnol. 22, 8-18, 1999), por ejemplo en el locus de *sacB*. La producción de riboflavina se comparó directamente en dos cepas recombinantes de *B. subtilis* que difieren solamente por la presencia o ausencia de las mutaciones en el gen *ribA*. El cultivo de las cepas de *Bacillus* se realizó como se describe en el Ejemplo 8.

#### **Ejemplo 8: Condiciones de cultivo para evaluar la producción de riboflavina**

45 La producción de riboflavina se ensayó en cultivos alimentados-discontinuos de la cepa RB50::(pRF69)<sub>n</sub>::(pRF93)<sub>m</sub> de *B. subtilis* que sobreproduce riboflavina, en la que los mutantes de GTP ciclohrolasa II PCR III, constructo C, o constructo E conducidos por el promotor  $P_{vegl}$  se integraron en el locus de *sacB* (véase el Ejemplo 6). La fermentación de las cepas se realizó como se describe en el documento EP 405370.

#### **Ejemplo 9: Métodos analíticos para la determinación de riboflavina**

50 Para la determinación de riboflavina, se puede usar el siguiente método analítico (Bretzel et al., J. Ind. Microbiol. Biotechnol. 22, 19-26, 1999).

55 El sistema cromatográfico fue un sistema Hewlett-Packard 1100 equipado con una bomba binaria, un termostato de columna y un automuestreador enfriado. Se usó en línea tanto un detector de conjunto de diodos como un detector de fluorescencia. Se registraron dos señales, UV a 280 nm y traza de fluorescencia a excitación de 446 nm, emisión 520 nm.

Se usó una columna de acero inoxidable Supercosil LC-8-DB (150 x 4,6 mm, tamaño de partículas 3 µm) junto con un cartucho de guarda. Las fases móviles fueron 100 mM de ácido acético (A) y metanol (B). Se usó una elución en gradiente según el siguiente esquema:

Tiempo [min.]	%A	%B
0	98	2
6	98	2
15	50	50
25	50	50

5 La temperatura de la columna se ajustó a 20°C, y el caudal fue 1,0 ml/min. El tiempo del experimento fue 25 min. Las muestras de fermentación se diluyeron, se filtraron y se analizaron sin tratamiento adicional. La riboflavina se cuantificó mediante comparación con un patrón externo. Los cálculos se basan en la señal de UV a 280 nm. Como material estándar, se usó riboflavina adquirida de Fluka (9471 Buchs, Suiza) (pureza ≥ 99,0%).

**Ejemplo 10: Identificación de los restos correspondientes en las enzimas de GTP ciclohrolasa II que son homólogos a GTP ciclohrolasa II de *Bacillus subtilis***

10 Un alineamiento de múltiples secuencias de aminoácidos de 92 enzimas diferentes de GTP ciclohrolasa II encontrado por el programa BLASTN usando bases de datos estándar tales como SWISS-PROT y TrEMBL (véase la Figura 1) se calculó con el programa "PILEUP" (GCG Wisconsin Package, versión 10.2, Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752, USA) usando los siguientes parámetros: penalización de creación de salto 8, penalización de extensión de salto 2, y matriz blosum62.cmp (parámetros por defecto).

15 Una GTP ciclohrolasa II homóloga en el contexto de la presente invención puede mostrar similitud de secuencia con cualquiera de las secuencias de aminoácidos de GTP ciclohrolasa II mostradas en la Figura 1. La Figura 1 proporciona un ejemplo de un alineamiento de múltiples secuencias de 92 secuencias de aminoácidos de GTP ciclohrolasa II, estando subrayada la secuencia de la GTP ciclohrolasa II procedente de *Bacillus subtilis*. La Figura 1 sirve justamente como ejemplo y no pretende ser una colección completa de todas las enzimas de GTP  
20 ciclohrolasa II conocidas. Se espera que los restos homólogos, *es decir*, restos de las diferentes enzimas de GTP ciclohrolasa II que están localizadas en la misma posición en el alineamiento de las secuencias de aminoácidos (*es decir*, están localizadas en la misma columna en, *por ejemplo*, la Figura 1), estén similarmente colocados en la estructura tridimensional de cada proteína y satisfagan en cada proteína una función comparable por estructura y por función. Los restos de aminoácidos homólogos a los restos de aminoácidos de la GTP ciclohrolasa II  
25 procedente de *B. subtilis* que se discuten en los Ejemplos están subrayados en negrita en la Figura 1, y la posición respectiva en la secuencia de aminoácidos de *B. subtilis* está añadida encima de la columna respectiva del alineamiento.

Los restos de aminoácidos de 92 organismos diferentes que corresponden a posiciones de aminoácidos específicas, *es decir*, posiciones que son homólogas/equivalentes a los restos de aminoácidos que se encuentra que tienen un  
30 efecto positivo sobre la actividad específica (restos de aminoácidos 261, 270, 276, 279, 308, 347) de la secuencia de aminoácidos de GTP ciclohrolasa II de *Bacillus subtilis* (SEC ID NO:2), se resumen en la Tabla 4, en la que el número en la columna de la izquierda define el organismo (según la Fig. 1), comenzando con el nombre de la secuencia usada, el número de acceso de la base de datos, y entre paréntesis el organismo fuente de la secuencia:

- 35 (1) gch2\_bacsu: SWISS-PROT: gch2\_bacsu (*Bacillus subtilis*)  
 (2) gch2\_cangu: geneseq:aay69776 (*Candida guilliermondii*)  
 (3) gch2\_ashgo: TrEMBL: CAA02912: (*Ashbya gossypii* (*Eremothecium gossypii*))  
 (4) gch2\_yeast: SWISS-PROT: gch2\_yeast (*Saccharomyces cerevisiae*)  
 (5) gch2\_neucr: TrEMBL: Q871B3 (*Neurospora crassa*)  
 (6) gch2\_schpo: TrEMBL: Q9P7M9 (*Schizosaccharomyces pombe*)  
 40 (7) gch2\_arcfu: SWISS-PROT: gch2\_arcfu (*Archaeoglobus fulgidus*)  
 (8) gch2\_strco: SWISS-PROT: gch2\_strco (*Streptomyces coelicolor*)  
 (9) gch2\_helpj: SWISS-PROT: gch2\_helpj (*Helicobacter pylori* J99)  
 (10) gch2\_helpy: SWISS-PROT: gch2\_helpy (*Helicobacter pylori*)  
 (11) gch2\_pyrfu: TrEMBL: Q8U4L7 (*Pyrococcus furiosus*)  
 45 (12) gch2\_thema: SWISS-PROT: gch2\_thema (*Thermotoga maritima*)  
 (13) gch2\_chlmu: SWISS-PROT: gch2\_chlmu (*Chlamydia muridarum*)  
 (14) gch2\_chltr: SWISS-PROT: gch2\_chltr (*Chlamydia trachomatis*)  
 (15) gch2\_chlca: TrEMBL: AAP05635 (*Chlamydia caviae* GPIC)  
 (16) gch2\_chlpn: SWISS-PROT: gch2\_chlpn (*Chlamydia pneumoniae*)  
 50 (17) gch2\_arath: SWISS-PROT: gch2\_arath (*Arabidopsis thaliana*)  
 (18) gch2\_lyces: TrEMBL: CAC09119 (*Lycopersicon esculentum*)  
 (19) gch2\_orysa: TrEMBL: AAO72560 (*Oryza sativum*)



## ES 2 435 990 T3

- (20) gch2\_alceu: TrEMBL: Q9F184 (*Alcaligenes eutrophus*)
- (21) gch2\_neima: SWISS-PROT: gch2\_neima (*Neisseria meningitidis* (serogrupo A))
- (22) gch2\_neimb: SWISS-PROT: gch2\_neimb (*Neisseria meningitidis* (serogrupo B))
- 5 (23) gch2\_psepk: SWISS-PROT: gch2\_psepk (*Pseudomonas putida* (cepa KT2440))
- (24) gch2\_psesm: SWISS-PROT: gch2\_psesm (*Pseudomonas syringae* (pv. tomate ))
- (25) gch2\_actac: TrEMBL: Q9JRR0 (*Actinobacillus actinomycetemcomitans* (*Haemophilus actinomycetemcomitans*))
- (26) gch2\_haein: SWISS-PROT: gch2\_pasmu gch2\_haein (*Haemophilus influenzae*)
- (27) gch2\_pasmu: SWISS-PROT: (*Pasteurella multocida*)
- 10 (28) gch2\_ec06: TrEMBL: Q8FHU5 (*Escherichia coli* 06)
- (29) gch2\_ecoli: SWISS-PROT: gch2\_ecoli (*Escherichia coli*)
- (30) gch2\_salty: TrEMBL: Q8XFY7 (*Salmonella typhimurium*)
- (31) gch2\_yerpe: TrEMBL: Q8ZEF0 (*Yersinia pestis*)
- (32) gch2\_buca: SWISS-PROT: gch2\_buca (*Buchnera aphidicola* (subsp. *Acyrtosiphon pisum*) (*Acyrtosiphon pisum* bacteria simbiótica))
- 15 (33) gch2\_bucap: SWISS-PROT: gch2\_bucap (*Buchnera aphidicola* (subsp. *Schizaphis graminum*))
- (34) gch2\_wigbr: SWISS-PROT: gch2\_wigbr (*Wigglesworthia glossinidia brevipalpis*)
- (35) gch2\_bucbp: SWISS-PROT: gch2\_wigbr (*Buchnera aphidicola* (subsp. *Baizongia pistaciae*))
- (36) gch2\_mycte: TrEMBL: Q9CCP4 (*Mycobacterium leprae*)
- 20 (37) gch2\_myctu: SWISS-PROT: gch2\_myctu (*Mycobacterium tuberculosis*)
- (38) gch2\_coref: TrEMBL: Q8FT57 (*Corynebacterium efficiens*)
- (39) gch2\_corgl: GENESEQP: AAB79913 (*Corynebacterium glutamicum*)
- (40) gch2\_coram: SWISS-PROT: gch2\_coram (*Corynebacterium ammoniagenes* (*Brevibacterium ammoniagenes*))
- 25 (41) gch2\_staau: TrEMBL: Q8NW14 (*Staphylococcus aureus* (cepa MW2))
- (42) gch2\_staep: GENESEQP: ABP40248 (*Staphylococcus epidermidis*)
- (43) gch2\_actpl: SWISS-PROT: gch2\_actpl (*Actinobacillus pleuropneumoniae*)
- (44) gch2\_lacla: TrEMBL: Q9CGU7 (*Lactococcus lactis* (subsp. *lactis*) (*Streptococcus lactis*))
- (45) gch2\_stcag: TrEMBL: Q8E658 (*Streptococcus agalactiae* (serotipo III))
- 30 (46) gch2\_stcpn: TrEMBL: Q8DRF1 (*Streptococcus pneumoniae* (cepa ATCC BAA-255/R6))
- (47) gch2\_cloac: TrEMBL: Q97LG9 (*Clostridium acetobutylicum*)
- (48) gch2\_fusnu: TrEMBL: Q8RIR1 (*Fusobacterium nucleatum* (subsp. *nucleatum*))
- (49) gch2\_anasp: TrEMBL: Q8RIR1 (*Anabaena* sp. (cepa PCC 7120))
- (50) gch2\_syny3: SWISS-PROT: gch2\_syny3 (*Synechocystis* sp. (cepa PCC 6803))
- 35 (51) gch2\_synel: TrEMBL: Q8DI64 (*Synechococcus elongatus* (*Thermosynechococcus elongatus*))
- (52) gch2\_bacam: SWISS-PROT: gch2\_bacam (*Bacillus amyloliquefaciens*)
- (53) gch2\_bacce: TrEMBL: AAP11030 (*Bacillus cereus* ATCC 14579)
- (54) gch2\_bacha: TrEMBL: Q9KCL5 (*Bacillus halodurans*)
- (55) gch2\_clope: TrEMBL: Q8XMX0 (*Clostridium perfringens*)
- 40 (56) gch2\_clote: TrEMBL: Q897Q8 (*Clostridium tetani*)
- (57) gch2\_chite: TrEMBL: Q8KC35 (*Chlorobium tepidum*)
- (58) gch2\_aquae: SWISS-PROT: gch2\_aquae (*Aquifex aeolicus*)
- (59) gch2\_lepin: TrEMBL: Q8F701 (*Leptospira interrogans*)
- (60) gch2\_deira: TrEMBL: Q9RXZ9 (*Deinococcus radiodurans*)
- 45 (61) gch2\_bacth: TrEMBL: Q8A528 (*Bacteroides thetaiotaomicron*)
- (62) gch2\_caucr: TrEMBL: Q9A9S5 (*Caulobacter crescentus*)
- (63) gch2\_coxbu: TrEMBL: AAO90191 (*Coxiella burnetii* RSA 493)
- (64) gch2\_rhiet: TrEMBL: Q8KL38 (*Rhizobium etli*)
- (65) gch2\_lacpl: TrEMBL: Q88X17 (*Lactobacillus plantarum*)
- 50 (66) gch2\_psegl: TrEMBL: Q8RS38 (*Pseudomonas glumae*)
- (67) gch2\_strav: TrEMBL: BAC71833 (*Streptomyces avermitilis*)
- (68) gch2\_phopo: SWISS-PROT: gch2\_phopo (*Photobacterium phosphoreum*)
- (69) gch2\_azobr: SWISS-PROT: gch2\_azobr (*Azospirillum brasilense*)
- (70) gch2\_agrtu: TrEMBL: Q8UHC9 (*Agrobacterium tumefaciens* (cepa C58 / ATCC 33970))
- 55 (71) gch2\_rhime: TrEMBL: Q92RH2 (*Rhizobium meliloti* (*Sinorhizobium meliloti*))
- (72) gch2\_brume: TrEMBL: Q8YFL5 (*Brucella melitensis*)
- (73) gch2\_brusu: TrEMBL: Q8G298 (*Brucella suis*)
- (74) gch2\_rhilo: TrEMBL: Q985Z3 (*Rhizobium loti* (*Mesorhizobium loti*))
- (75) gch2\_braja: TrEMBL: Q89RZ7 (*Bradyrhizobium japonicum*)
- 60 (76) gch2\_niteu: TrEMBL: CAD86468 (*Nitrosomonas europaea* ATCC 19718)
- (77) gch2\_ralso: TrEMBL: Q8Y1H7 (*Ralstonia solanacearum* (*Pseudomonas solanacearum*))
- (78) gch2\_neime: TrEMBL: Q9JZ77 (*Neisseria meningitidis* (serogrupo B, segunda enzima encontrada))
- (79) gch2\_xanax: TrEMBL: Q8PPD7 (*Xanthomonas axonopodis* (pv. *citri*))
- (80) gch2\_xanca: TrEMBL: Q8PCM8 (*Xanthomonas campestris* (pv. *campestris*))
- 65 (81) gch2\_vibpa: TrEMBL: Q87RU5 (*Vibrio parahaemolyticus*)
- (82) gch2\_vibvu: TrEMBL: Q8DF98 (*Vibrio vulnificus*)

## ES 2 435 990 T3

- 5  
10
- (83) gch2\_vibch: TrEMBL: Q9KPU3 (*Vibrio cholerae*)
  - (84) gch2\_vibfi: TrEMBL: Q8G9G5 (*Vibrio fischeri*)
  - (85) gch2\_sheon: TrEMBL: Q8EBP2 (*Shewanella oneidensis*)
  - (86) gch2\_phoph: TrEMBL: Q8G9H7 (*Photobacterium phosphoreum*)
  - (87) ribb\_phole: TrEMBL: Q93E93 (*Photobacterium leiognathi*)
  - (88) gch2\_psepu: TrEMBL: Q88GB1 (*Pseudomonas putida* (cepa KT2440, segunda enzima encontrada))
  - (89) gch2\_psesy: TrEMBL: Q882G0 (*Pseudomonas syringae* (pv. Tomate, segunda enzima encontrada))
  - (90) gch2\_pseae: TrEMBL: Q9HWX4 (*Pseudomonas aeruginosa*)
  - (91) ribb\_dehmu: SWISS-PROT: ribb\_dehmu (*Dehalospirillum multivorans*)
  - (92) gch2\_xylfa: TrEMBL: Q87D69 (*Xylella fastidiosa* (cepa Temecula1 / ATCC 700964))

Tabla 4: Posiciones/restos de aminoácidos que corresponden a las posiciones V261, G270, A276, Q279, K308, y M347 de RibA de *B. subtilis* como en SEC ID NO:2. Los números en la columna de la izquierda se refieren a los diferentes organismos (véase anteriormente).

							SEC ID NO ADN; proteína
(1)	V261	G270	A276	Q279	K308	M347	1; 2
(2)	T181	G190	A196	L199	N228	I267	
(3)	T126	G135	A141	L144	N182	I221	32; 33
(4)	N153	G162	A168	L171	N211	V250	
(5)	T255	G264	A270	L273	N305	I344	
(6)	T176	G185	A191	L194	N223	I262	
(7)	V139	G248	A254	M257	K287	I326	
(8)	A73	G82	A88	Q91	A121	V160	
(9)	A60	G69	A75	R78	A106	M145	
(10)	A60	G69	A75	R78	A106	M145	
(11)	T151	G160	A166	T169	E297	I336	
(12)	V245	G254	F260	Y263	S291	V330	
(13)	V267	G276	A282	Y285	A315	V354	
(14)	I267	G276	A282	Y285	A315	V354	
(15)	I272	G281	A287	Y290	A320	I359	
(16)	I272	G281	A287	Y290	A320	I359	
(17)	I91	G100	S106	Q109	N139	M178	
(18)	I299	G308	A314	Q317	N347	M386	
(19)	I295	G304	A310	L313	N343	M382	
(20)	V61	G70	A76	K79	R109	V148	
(21)	A60	G69	A75	A78	H107	V146	
(22)	A60	G69	A75	A78	H107	V146	
(23)	A59	G68	A74	A77	E106	L145	
(24)	A59	G68	A74	A77	E106	L145	
(25)	A61	G70	A76	A79	S108	I147	
(26)	A61	G70	A76	A79	S108	V147	
(27)	A61	G70	A76	A79	S108	V147	
(28)	A80	G89	A95	Q98	A127	V166	
(29)	A59	G68	A74	Q77	A106	V145	34; 35

ES 2 435 990 T3

							SEC ID NO ADN; proteína
(30)	A59	G68	A74	H77	A106	V145	
(31)	A59	G68	A74	R77	A106	V145	
(32)	A57	G66	S72	R75	A104	I143	
(33)	A59	G68	A74	R77	S106	I145	
(34)	A59	G68	A74	H77	S106	I145	
(35)	A60	G69	A75	E78	A107	I146	
(36)	V279	G288	A294	M297	Q327	M366	
(37)	V269	G278	A284	M287	Q317	M356	
(38)	V281	G290	S296	M299	Q329	I368	
(39)	V273	G282	S288	L291	Q321	L360	36; 37
(40)	V274	G283	S289	I292	S322	A361	
(41)	I259	G268	S274	Y277	E305	I344	
(42)	I263	G272	S276	Y279	E309	I348	
(43)	A264	G273	A279	Q282	E311	I350	
(44)	A262	G271	A277	K280	S309	L348	
(45)	V261	G270	A276	Q279	H308	I347	
(46)	V261	G270	A276	M279	H308	L347	
(47)	V264	G273	A279	A282	M311	V350	
(48)	I261	G270	A276	R279	N308	I347	
(49)	A301	R310	A316	M319	S348	I387	
(50)	A265	R274	A280	M283	S312	I351	
(51)	A267	R276	A282	M285	S315	I353	
(52)	V261	G270	A276	Q279	R308	M347	38; 39
(53)	V260	G269	A275	Q278	K307	L346	40; 41
(54)	V264	G273	A279	Q282	K311	M350	42; 43
(55)	V244	G253	A259	K262	K291	I330	
(56)	A265	G274	A280	A283	N312	I351	
(57)	T267	G276	A282	M285	N314	M353	
(58)	V270	R279	A285	M288	E317	M356	
(59)	I263	G272	A278	M281	N310	M349	
(60)	A269	G278	A284	A287	A316	L355	
(61)	I265	G274	A281	M283	K314	M351	
(62)	A260	G269	S276	Q278	A307	V346	
(63)	V265	G274	A280	E283	A311	I350	
(64)	A283	G292	A298	A301	S330	V369	
(65)	V262	G271	A277	K280	A309	V348	
(66)	V70	G79	S85	L88	R117	V156	
(67)	V66	G75	A81	R84	A112	L151	

ES 2 435 990 T3

							SEC ID NO ADN; proteína
(68)	G61	G70	T76	I79	K108	I147	
(69)	L250	G259	A265	E268	R297	V336	
(70)	V264	-	Y274	K277	K304	I340	
(71)	V264	-	I274	R277	K304	I340	
(72)	V264	-	I275	R278	R305	I345	
(73)	V264	-	I275	R278	R305	I345	
(74)	V264	-	I274	A277	-	I340	
(75)	I261	-	V271	H274	-	I330	
(76)	L262	S271	A277	V280	D303	M338	
(77)	L285	S294	A300	A303	Q327	M362	
(78)	F263	S271	A277	H280	D303	L336	
(79)	L262	G271	A277	A280	R304	L339	
(80)	L262	G271	A277	A280	R304	L339	
(81)	V261	S271	A277	R280	R303	M343	
(82)	L261	S271	A277	R280	R303	M343	
(83)	L261	S271	A277	R280	R303	M343	
(84)	I261	S271	A277	R280	K303	M343	
(85)	L260	S270	A276	R279	K302	M343	
(86)	I262	-	A275	R278	K301	M341	
(87)	I262	-	A275	R278	Q301	I339	
(88)	L262	N272	A278	K281	R305	L342	
(89)	L262	N272	A278	R281	R305	L342	
(90)	L262	R270	A276	K279	H303	M339	
(91)	L269	Y276	A282	Y285	-	I324	
(92)	L230	G239	S245	T248	G277	I316	

Los ejemplos mostrados en la Tabla 4 sirven como ilustración del principio. Los restos correspondientes se pueden determinar para todas las otras secuencias de aminoácidos de GTP ciclohidrolasa II que son homólogas a una cualquiera de las secuencias mostradas en la Figura 1 o en la Tabla 4.

5 LISTADO DE SECUENCIAS

<110> DSM IP Assets B.V.

<120> Enzimas mejoradas

<130> 22257

<160> 43

10 <170> PatentIn version 3.2

<210> 1

<211> 1197

<212> ADN

ES 2 435 990 T3

<213> *Bacillus subtilis*

<400> 1

atgtttcatc cgatagaaga agcactggac gctttaaaaa aaggcgaagt catcatcgtt	60
gtagatgatg aagacagaga aatgaagga gactttgtgg ctcttgccga gcatgcaacg	120
ccggaagtca ttaactttat ggcgacacat gggagaggac tgatctgcac gccgctcagt	180
gaggaaatcg cagacaggct tgatcttcac cctatggttg agcataatac agactctcac	240
cacactgcat ttaccgtaag catagacatc cgtgaaacga agacaggatc cagcgtcaa	300
gaaagatctt ttaccgttca agcattgctg gacagcaaat ccgtgccatc tgattttcag	360
cgcccggggc acatttttcc actgattgcg aaaaaaggag gtgtcctgaa aagagcgggc	420
catacagaag ctgctgttga tcttgctgaa gcttggtgat ctccaggagc cggcgtcatt	480
tgtgaaatta tgaatgaaga cggaacgatg gcgagagtgc ctgagctcat tgaattgcg	540
aaaagcatc aattaaat gatcaccatt aaggatttga ttcaataccg ttacaatctg	600
acaacacttg tcgagcgtga agttgacatt acgctgccta ctgattttgg gacatttaag	660
gtttatggat acacaaatga ggtagatgga aaagagcatg tcgcatttgt gatgggagat	720
gtgccgttcg gagaagaacc ggtattggtc cgggtgcatt cagaatgtct cacaggtgac	780
gtgtttggct ctcatcgctg tgattgcgga ccgcagctgc acgccgcgct gaaccaaat	840
gccgcagaag gccgtggagt gctcctgtac ttgcgccaag aaggacgagg catcggttta	900
atcaataaat taaaagctta taagcttcag gaacaaggct atgacaccgt agaagccaat	960
gaggcgcttg gattcttgcc ggatcttcgc aactatggca tcggagcaca aattttacgc	1020
gacctcggtg tccggaatat gaagcttttg acgaataatc ccgcaaaaat cgcaggcctt	1080
gaaggctacg gactcagtat ttcagaaaga gtgccgcttc aaatggaggc gaaagaacac	1140
aataaaaaat atttgcaaac caaatgaac aagctaggtc atttacttca tttctaa	1197

<210> 2

5 <211> 398

<212> PRT

<213> *Bacillus subtilis*

<400> 2

ES 2 435 990 T3

Met Phe His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu  
 1 5 10 15

Val Ile Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe  
 20 25 30

Val Ala Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala  
 35 40 45

Thr His Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala  
 50 55 60

Asp Arg Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His  
 65 70 75 80

His Thr Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly  
 85 90 95

Ile Ser Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser  
 100 105 110

Lys Ser Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu  
 115 120 125

Ile Ala Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala  
 130 135 140

Ala Val Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile  
 145 150 155 160

Cys Glu Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu  
 165 170 175

Ile Glu Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp  
 180 185 190

Leu Ile Gln Tyr Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val  
 195 200 205

ES 2 435 990 T3

Asp Ile Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr  
 210 215 220

Thr Asn Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp  
 225 230 235 240

Val Pro Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys  
 245 250 255

Leu Thr Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln  
 260 265 270

Leu His Ala Ala Leu Asn Gln Ile Ala Ala Glu Gly Arg Gly Val Leu  
 275 280 285

Leu Tyr Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu  
 290 295 300

Lys Ala Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn  
 305 310 315 320

Glu Ala Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala  
 325 330 335

Gln Ile Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn  
 340 345 350

Asn Pro Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser  
 355 360 365

Glu Arg Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr  
 370 375 380

Leu Gln Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 385 390 395

<210> 3

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 3

ES 2 435 990 T3

atgagaggat ctcaccatca ccacccat gggatcgatc atatgtttca tccgatagaa 60  
 gaagcactgg acgctttaa aaaaggcgaa gtcacatcg ttgtagatga tgaagacaga 120  
 gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
  
 atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg 240  
 cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta 300  
 agcatagacc atcgtgaaac gaagacaggt atcagcgctc aagaaagatc ttttaccggt 360  
 caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt 420  
 ccaactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
 gatcttgctg aagcttgtgg atctccagga gccggcgctca tttgtgaaat tatgaatgaa 540  
 gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
 atgatcacca ttaaggattt gattcaatac cgttacaatc tgacaacact tgtcgagcgt 660  
 gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
 gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa 780  
 ccggtattgg tccgggtgca ttcagaatgt ctcacagggtg acgtgtttgg ctctcatcgc 840  
 tgtgattgcg gaccgcagct gcacgccgcg ctgaaccaa ttgccgcaga aggccgtgga 900  
 gtgctcctgt acttgcgcca agaaggacga ggcacgggt taatcaataa attaaaagct 960  
 tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
 ccgatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat 1080  
 atgaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
 atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
 accaaaatga acaagctagg tcatttactt catttctaa 1239

<210> 4

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 4



ES 2 435 990 T3

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
1 5 10 15

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
20 25 30

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
50 55 60

ES 2 435 990 T3

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

ES 2 435 990 T3

290		295		300											
Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	Ile	Asn	Lys	Leu	Lys	Ala
305						310				315					320
Tyr	Lys	Leu	Gln	Glu	Gln	Gly	Tyr	Asp	Thr	Val	Glu	Ala	Asn	Glu	Ala
				325					330					335	
Leu	Gly	Phe	Leu	Pro	Asp	Leu	Arg	Asn	Tyr	Gly	Ile	Gly	Ala	Gln	Ile
			340					345					350		
Leu	Arg	Asp	Leu	Gly	Val	Arg	Asn	Met	Lys	Leu	Leu	Thr	Asn	Asn	Pro
		355					360					365			
Arg	Lys	Ile	Ala	Gly	Leu	Glu	Gly	Tyr	Gly	Leu	Ser	Ile	Ser	Glu	Arg
	370					375					380				
Val	Pro	Leu	Gln	Met	Glu	Ala	Lys	Glu	His	Asn	Lys	Lys	Tyr	Leu	Gln
385					390					395					400
Thr	Lys	Met	Asn	Lys	Leu	Gly	His	Leu	Leu	His	Phe				
				405					410						

<210> 5

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 5

```

atgagaggat ctcaccatca ccatcacat gggatcgatc atatgtttca tccgatagaa      60
gaagcactgg acgctttaa aaaaggcgaa gtcatcatcg ttgtagatga tgaagacaga      120
gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt      180
atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg      240
cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta      300
agcatagacc atcgtgaaac gaagacaggt atcagcgctc aagaaagatc ttttaccgtt      360
caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt      420
ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt      480
gatcttgctg aagcttgtgg atctccagga gccggcgtca tttgtgaaat tatgaatgaa      540
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa      600
atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtecgagcgt      660

```

ES 2 435 990 T3

gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa 780  
ccggtattgg tccgggtgca ttcagaatgt ctcacagggtg acgtgtttgg ctctcatcgc 840  
tgtgattgcg gaccgcagct gcacgccacg ctgaacccaaa ttgccgcaga aggccgtgga 900  
gtgctcctgt acttgcgcca agaaggacga ggcacggtt taatcaataa attaaaagct 960  
tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
ccgatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat 1080  
atgaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
accaaaatga acaagctagg tcatttactt catttctaa 1239

<210> 6

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 6

Met	Arg	Gly	Ser	His	His	His	His	His	His	Gly	Ile	Asp	His	Met	Phe
1				5					10					15	
His	Pro	Ile	Glu	Glu	Ala	Leu	Asp	Ala	Leu	Lys	Lys	Gly	Glu	Val	Ile
			20					25					30		
Ile	Val	Val	Asp	Asp	Glu	Asp	Arg	Glu	Asn	Glu	Gly	Asp	Phe	Val	Ala
		35					40					45			
Leu	Ala	Glu	His	Ala	Thr	Pro	Glu	Val	Ile	Asn	Phe	Met	Ala	Thr	His
	50					55					60				
Gly	Arg	Gly	Leu	Ile	Cys	Thr	Pro	Leu	Ser	Glu	Glu	Ile	Ala	Asp	Arg
65					70					75					80
Leu	Asp	Leu	His	Pro	Met	Val	Glu	His	Asn	Thr	Asp	Ser	His	His	Thr
				85					90					95	
Ala	Phe	Thr	Val	Ser	Ile	Asp	His	Arg	Glu	Thr	Lys	Thr	Gly	Ile	Ser
			100					105					110		
Ala	Gln	Glu	Arg	Ser	Phe	Thr	Val	Gln	Ala	Leu	Leu	Asp	Ser	Lys	Ser
		115					120					125			

ES 2 435 990 T3

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
195 200 205

Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
260 265 270

Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His  
275 280 285

Ala Thr Leu Asn Gln Ile Ala Ala Glu Gly Arg Gly Val Leu Leu Tyr  
290 295 300

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
305 310 315 320

Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
325 330 335

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn Asn Pro  
355 360 365

ES 2 435 990 T3

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 7

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 7

atgagaggat ctcacatca ccatcacat gggatcgatc atatgtttca tccgatagaa 60  
 gaagcactgg acgctttaa aaaaggcgaa gtcacatcg ttgtagatga tgaagacaga 120  
 gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
 atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgagacagg 240  
 cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta 300  
 agcatagacc atcgtgaaac gaagacaggt atcagcgctc aagaaagatc ttttaccggt 360  
 caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt 420  
 cactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
 gatcttgctg aagcttggtg atctccagga gccggcgctca tttgtgaaat tatgaatgaa 540  
 gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
 atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtcgagcgt 660  
 gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
 gaggtagatg gaaaagagca tgtcgcatth gtgatgggag atgtgccggt cggagaagaa 780  
 ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgcgtttg ctctcatcgc 840  
 tgtgattgcg gaccgcagct gcacgcccg ctgaaccaa ttgcgcgaga aggccgtgga 900  
 gtgctcctgt acttgcgcca agaaggacga ggcacgggt taatcaataa attaaaagct 960  
 tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
 ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat 1080  
 atgaagcttt tgacgaataa tccgcgaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
 atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200

accaaaatga acaagctagg tcatttactt catttctaa

1239

<210> 8

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 8

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
1 5 10 15

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
20 25 30

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
50 55 60

Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
65 70 75 80

Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
85 90 95

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

ES 2 435 990 T3

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
 195 200 205

Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
 210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
 225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
 245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
 260 265 270

Gly Asp Ala Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His  
 275 280 285

Ala Ala Leu Asn Gln Ile Ala Ala Glu Gly Arg Gly Val Leu Leu Tyr  
 290 295 300

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
 305 310 315 320

Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
 325 330 335

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
 340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn Asn Pro  
 355 360 365

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 9

<211> 1239

<212> ADN



ES 2 435 990 T3

<213> *Bacillus subtilis*

<400> 9

atgagaggat ctcaccatca ccatcaccat gggatcgatc atatgtttca tccgatagaa	60
gaagcactgg acgctttaaa aaaaggcgaa gtcatcatcg ttgtagatga tgaagacaga	120
gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt	180
atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgagacagg	240
cttgatcttc accctatggg tgagcataat acagactctc accacactgc atttaccgta	300
agcatagacc atcgtgaaac gaagacaggt atcagcgctc aagaaagatc ttttaccggt	360
caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt	420
ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt	480
gatcttgctg aagcttgtgg atctccagga gccggcgtea tttgtgaaat tatgaatgaa	540
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa	600
atgatcacca ttaaggattt gattcaatac cgttacaatc tgacaacact tgtcgcgct	660
gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat	720
gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa	780
ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgtgtttgg ctctcatcgc	840
tgtgattgcg gaccgcagct gcacgccgcg ctgaaccaa ttgccgcaga aggccgtgga	900
gtgctcctgt acttgcgcca agaaggacga ggcatcggtt taatcaataa attaaaagct	960
tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg	1020
ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat	1080
ataaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt	1140
atctcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa	1200
accaaaatga acaagctagg tcatttactt catttctaa	1239

<210> 10

5 <211> 412

<212> PRT

<213> *Bacillus subtilis*

<400> 10

ES 2 435 990 T3

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
1                   5                                   10                                   15

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
                  20                                   25                                   30

ES 2 435 990 T3

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
50 55 60

Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
65 70 75 80

Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
85 90 95

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
195 200 205

Gln Tyr Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr

ES 2 435 990 T3

		260						265						270			
Gly	Asp	Val	Phe	Gly	Ser	His	Arg	Cys	Asp	Cys	Gly	Pro	Gln	Leu	His		
		275					280					285					
Ala	Ala	Leu	Asn	Gln	Ile	Ala	Ala	Glu	Gly	Arg	Gly	Val	Leu	Leu	Tyr		
	290					295					300						
Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	Ile	Asn	Lys	Leu	Lys	Ala		
305					310					315					320		
Tyr	Lys	Leu	Gln	Glu	Gln	Gly	Tyr	Asp	Thr	Val	Glu	Ala	Asn	Glu	Ala		
				325					330					335			
Leu	Gly	Phe	Leu	Pro	Asp	Leu	Arg	Asn	Tyr	Gly	Ile	Gly	Ala	Gln	Ile		
			340					345					350				
Leu	Arg	Asp	Leu	Gly	Val	Arg	Asn	Ile	Lys	Leu	Leu	Thr	Asn	Asn	Pro		
		355					360						365				
Arg	Lys	Ile	Ala	Gly	Leu	Glu	Gly	Tyr	Gly	Leu	Ser	Ile	Ser	Glu	Arg		
	370					375					380						
Val	Pro	Leu	Gln	Met	Glu	Ala	Lys	Glu	His	Asn	Lys	Lys	Tyr	Leu	Gln		
385					390					395					400		
Thr	Lys	Met	Asn	Lys	Leu	Gly	His	Leu	Leu	His	Phe						
				405					410								

<210> 11

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 11

atgagaggat ctcacatca ccatcacat gggatcgatc atatgtttca tccgatagaa	60
gaagcactgg acgctttaa aaaaggcgaa gtcatcatcg ttgtagatga tgaagacaga	120
gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt	180
atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg	240
cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta	300
agcatagacc atcgtgaaac gaagacaggt atcagcgctc aagaaagatc ttttaccggt	360
caagcattgc tggacagcaa atccgtgcca totgattttc agcgtccggg gcacattttt	420

ES 2 435 990 T3

ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
gatcttgctg aagcttgtgg atctccagga gccggcgctca tttgtgaaat tatgaatgaa 540  
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgctgagcgt 660  
gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaaat 720  
gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa 780  
ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgtgtttgg ctctcatcgc 840  
tgtgattgcg gaccgcagct gcacgccacg ctgaaccaa ttgccacaga aggccgtgga 900  
gtgctcctgt acttgcgcca agaaggacga ggcacgggtt taatcaataa attaaaagct 960  
tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat 1080  
atgaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
atctcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
accaaaatga acaagctagg tcatttactt catttctaa 1239

<210> 12

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 12

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
1 5 10 15  
His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
20 25 30  
Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
35 40 45  
Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
50 55 60  
Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
65 70 75 80  
Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
85 90 95

ES 2 435 990 T3

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
195 200 205

Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
260 265 270

Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His  
275 280 285

Ala Thr Leu Asn Gln Ile Ala Thr Glu Gly Arg Gly Val Leu Leu Tyr  
290 295 300

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
305 310 315 320

Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
325 330 335

ES 2 435 990 T3

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
 340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn Asn Pro  
 355 360 365

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 13

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 13

atgagaggat ctcaccatca ccacacatc gggatcgatc atatgtttca tccgatagaa 60  
 gaagcactgg acgctttaa aaaaggcgaa gtcacatcgt ttgtagatga tgaagacaga 120  
 gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
 atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg 240  
 cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta 300  
 agcatagacc atcgtgaaac gaagacaggt atcagcgcctc aagaaagatc ttttaccggt 360  
 caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt 420  
 ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
 gatcttgctg aagcttgtgg atctccagga gccggcgtca tttgtgaaat tatgaatgaa 540  
 gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaa 600  
 atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtcgagcgt 660  
 gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
 gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa 780  
 ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgtgtttgg ctctcatcgc 840  
 tgtgattgeg gaccgcagct gcacgccacg ctgaaccaa ttgccacaga aggccgtgga 900  
 gtgctcctgt acttgcgcca agaaggacga ggcacggtt taatcaataa attaaaagct 960

ES 2 435 990 T3

tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
 ccggatcttc gcaactatgg catcggagca caaatTTTtac gcgacctcgg tgtccggaat 1080  
 ataaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
 atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
 accaaaatga acaagctagg tcatttactt catttctaa 1239

<210> 14

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 14

Met	Arg	Gly	Ser	His	His	His	His	His	His	His	Gly	Ile	Asp	His	Met	Phe
1			5							10					15	
His	Pro	Ile	Glu	Glu	Ala	Leu	Asp	Ala	Leu	Lys	Lys	Gly	Glu	Val	Ile	
			20					25					30			
Ile	Val	Val	Asp	Asp	Glu	Asp	Arg	Glu	Asn	Glu	Gly	Asp	Phe	Val	Ala	
		35					40					45				
Leu	Ala	Glu	His	Ala	Thr	Pro	Glu	Val	Ile	Asn	Phe	Met	Ala	Thr	His	
	50					55					60					
Gly	Arg	Gly	Leu	Ile	Cys	Thr	Pro	Leu	Ser	Glu	Glu	Ile	Ala	Asp	Arg	
65					70					75					80	
Leu	Asp	Leu	His	Pro	Met	Val	Glu	His	Asn	Thr	Asp	Ser	His	His	Thr	
			85						90					95		
Ala	Phe	Thr	Val	Ser	Ile	Asp	His	Arg	Glu	Thr	Lys	Thr	Gly	Ile	Ser	
			100					105					110			
Ala	Gln	Glu	Arg	Ser	Phe	Thr	Val	Gln	Ala	Leu	Leu	Asp	Ser	Lys	Ser	
		115					120					125				
Val	Pro	Ser	Asp	Phe	Gln	Arg	Pro	Gly	His	Ile	Phe	Pro	Leu	Ile	Ala	
	130					135					140					
Lys	Lys	Gly	Gly	Val	Leu	Lys	Arg	Ala	Gly	His	Thr	Glu	Ala	Ala	Val	
145					150					155					160	



ES 2 435 990 T3

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
 165 170 175  
 Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
 180 185 190  
 Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
 195 200 205  
 Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
 210 215 220  
 Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
 225 230 235 240  
 Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
 245 250 255  
 Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
 260 265 270  
 Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His  
 275 280 285  
 Ala Thr Leu Asn Gln Ile Ala Thr Glu Gly Arg Gly Val Leu Leu Tyr  
 290 295 300  
 Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
 305 310 315 320  
 Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
 325 330 335  
 Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
 340 345 350  
 Leu Arg Asp Leu Gly Val Arg Asn Ile Lys Leu Leu Thr Asn Asn Pro  
 355 360 365  
 Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380  
 Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

ES 2 435 990 T3

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 15

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 15

```

atgagaggat ctcaccatca ccatcaccat gggatcgcgc atatgtttca tccgatagaa      60
gaagcactgg acgctttaaa aaaaggcgaa gtcacatcgc ttgtagatga tgaagacaga      120
gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt      180
atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg      240
cttgatcttc accctatggg tgagcataat acagactctc accacactgc atttaccgta      300
agcatagacc atcgtgaaac gaagacaggt atcagcgcgc aagaaagatc ttttaccggt      360
caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt      420
ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt      480
gatcttgctg aagcttgtgg atctccagga gccggcgcgc tttgtgaaat tatgaatgaa      540
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa      600
atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtcgagcgt      660
gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat      720
gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa      780
ccggtattgg tccgggtgca ttcagaatgt ctcacagggtg acgtgttttg ctctcatcgc      840
tgtgattgcg gaccgcagct gcacgccacg ctgaaccgaa ttgccacaga aggccgtgga      900
gtgctcctgt acttgcgcca agaaggacga ggcacgcggt taatcaataa attaaaagct      960
tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgcct tggattcttg     1020
ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat     1080
ataaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt     1140
atctcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa     1200
accaaatga acaagctagg tcatttactt catttctaa                                1239
    
```

<210> 16

<211> 412

10 <212> PRT

<213> *Bacillus subtilis*

ES 2 435 990 T3

<400> 16

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
 1 5 10 15

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
 20 25 30

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
 35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
 50 55 60

Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
 65 70 75 80

Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
 85 90 95

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
 100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
 115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
 130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
 145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
 165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
 180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
 195 200 205

Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
 210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn

ES 2 435 990 T3

225						230											235												240							
Glu	Val	Asp	Gly	Lys	Glu	His	Val	Ala	Phe	Val	Met	Gly	Asp	Val	Pro																					
				245					250					255																						
Phe	Gly	Glu	Glu	Pro	Val	Leu	Val	Arg	Val	His	Ser	Glu	Cys	Leu	Thr																					
			260					265					270																							
Gly	Asp	Val	Phe	Gly	Ser	His	Arg	Cys	Asp	Cys	Gly	Pro	Gln	Leu	His																					
		275					280					285																								
Ala	Thr	Leu	Asn	Arg	Ile	Ala	Thr	Glu	Gly	Arg	Gly	Val	Leu	Leu	Tyr																					
	290					295					300																									
Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	Ile	Asn	Lys	Leu	Lys	Ala																					
305					310					315					320																					
Tyr	Lys	Leu	Gln	Glu	Gln	Gly	Tyr	Asp	Thr	Val	Glu	Ala	Asn	Glu	Ala																					
				325					330					335																						
Leu	Gly	Phe	Leu	Pro	Asp	Leu	Arg	Asn	Tyr	Gly	Ile	Gly	Ala	Gln	Ile																					
			340					345					350																							
Leu	Arg	Asp	Leu	Gly	Val	Arg	Asn	Ile	Lys	Leu	Leu	Thr	Asn	Asn	Pro																					
		355					360					365																								
Arg	Lys	Ile	Ala	Gly	Leu	Glu	Gly	Tyr	Gly	Leu	Ser	Ile	Ser	Glu	Arg																					
	370					375					380																									
Val	Pro	Leu	Gln	Met	Glu	Ala	Lys	Glu	His	Asn	Lys	Lys	Tyr	Leu	Gln																					
	385				390					395					400																					
Thr	Lys	Met	Asn	Lys	Leu	Gly	His	Leu	Leu	His	Phe																									
				405						410																										

<210> 17  
 <211> 1239  
 <212> ADN  
 5 <213> *Bacillus subtilis*  
 <400> 17

ES 2 435 990 T3

atgagaggat ctcacatca ccatcacat gggatcgatc atatgtttca tccgatagaa 60  
 gaagcactgg acgctttaa aaaaggcgaa gtcatcatcg ttgtagatga tgaagacaga 120  
 gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
  
 atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg 240  
 ctgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta 300  
 agcatagacc atcgtgaaac gaagacaggt atcagcgcctc aagaaagatc ttttaccgtt 360  
 caagcattgc tggacagcaa atccgtgcc tctgattttc agcgtccggg gcacattttt 420  
 cactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
 gatcttgctg aagcttgtgg atctccagga gccggcgtca tttgtgaaat tatgaatgaa 540  
 gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
 atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtcgagcgt 660  
 gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
 gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccgtt cggagaagaa 780  
 ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgtgtttgg ctctcatcgc 840  
 tgtgattgcg gaccgcagct gcacgccact ctgaaccgaa ttgccacaga aggccgtgga 900  
 gtgctcctgt acttgcgcca agaaggacga ggcacgggtt taatcaataa attaaaagct 960  
 tataggcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
 ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgcccggaat 1080  
 ataaagcttt tgacgaataa tccgcgaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
 atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
 accaaaatga acaagctagg tcatttactt catttctaa 1239

<210> 18

<211> 412

5 <212> PRT

<213> *Bacillus subtilis*

<400> 18

ES 2 435 990 T3

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
1 5 10 15

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
20 25 30

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
50 55 60

ES 2 435 990 T3

Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
65 70 75 80

Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
85 90 95

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
195 200 205

Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
260 265 270

Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His  
275 280 285

Ala Thr Leu Asn Arg Ile Ala Thr Glu Gly Arg Gly Val Leu Leu Tyr  
290 295 300

ES 2 435 990 T3

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
305 310 315 320

Tyr Arg Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
325 330 335

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Ile Lys Leu Leu Thr Asn Asn Pro  
355 360 365

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
405 410

<210> 19

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 19

atgagaggat ctcaccatca ccatcaccat gggatcgate atatgtttca tccgatagaa 60  
gaagcactgg acgctttaa aaaaggcgaa gtcatcatcg ttgtagatga tgaagacaga 120  
gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgagacagg 240  
cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta 300  
agcatagacc atcgtgaaac gaagacaggt atcagcgcgc aagaaagatc ttttaccggt 360  
caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacattttt 420  
cactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
gatcttgctg aagcttgagg atctccagga gccggcgtca tttgtgaaat tatgaatgaa 540  
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
atgatcacca ttaaggattt gattcaatac cgttacaatc tgacaacact tgctgagcgt 660  
gaagttgaca ttacgctgcc tactgatttt gggacatttg aggtttatgg atacacaaat 720



ES 2 435 990 T3

gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa 780  
 ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgtgtttgg ctctcatcgc 840  
 tgtgattgcg caccgcagct gcacgccgcg ctgaacccaaa ttgccgcaga aggccgtgga 900  
 gtgctcctgt acttgcgcca agaaggacga ggcacatcggtt taatcaataa attaaaagct 960  
 tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
 ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat 1080  
 atgaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
 atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
 accaaaatga acaagctagg tcatttactt catttctaa 1239

<210> 20

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 20

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe  
 1 5 10 15

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile  
 20 25 30

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
 35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
 50 55 60

Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
 65 70 75 80

Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
 85 90 95

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
 100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
 115 120 125

ES 2 435 990 T3

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
195 200 205

Gln Tyr Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Glu Val Tyr Gly Tyr Thr Asn  
225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
260 265 270

Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Ala Pro Gln Leu His  
275 280 285

Ala Ala Leu Asn Gln Ile Ala Ala Glu Gly Arg Gly Val Leu Leu Tyr  
290 295 300

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
305 310 315 320

Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
325 330 335

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn Asn Pro  
355 360 365

ES 2 435 990 T3

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 21

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 21

atgagaggat ctcaccatca ccatcaccat gggatcgate atatgtttca tccgatagaa 60  
 gaagcactgg acgctttaa aaaaggcgaa gtcacatcgt ttgtagatga tgaagacaga 120  
 gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
 atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg 240  
 cttgatcttc accctatggg tgagcataat acagactctc accacactgc atttaccgta 300  
 agcatagacc atcgtgaaac gaagacaggt atcagcgcctc aagaaagatc ttttaccggt 360  
 caagcattgc tggacagcaa atccgtgcc tctgattttc agcgtccggg gcacattttt 420  
 ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480  
 gatcttgctg aagcttggtg atctccagga gccggcgtca tttgtgaaat tatgaatgaa 540  
 gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
 atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtcgagcgt 660  
 gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
 gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa 780  
 cgggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgtgtttgg ctctcatcgc 840  
 tgtgattgcg gaccgcagct gcacgccacg ctgaaccgaa ttgccacaga aggccgtgga 900  
 gtgctcctgt acttgcgcca agaaggacga ggcacgggtt taatcaataa attaaaagct 960  
 tataggcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggatacttg 1020  
 ccgatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgcccggaat 1080  
 ataaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
 atttcagaaa gagtgccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200

accaaaatga acaagctagg tcatttactt catttctaa

1239

<210> 22

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 22

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

ES 2 435 990 T3

195		200		205
Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile 210 215 220				
Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn 225 230 235 240				
Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro 245 250 255				
Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr 260 265 270				
Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His 275 280 285				
Ala Thr Leu Asn Arg Ile Ala Thr Glu Gly Arg Gly Val Leu Leu Tyr 290 295 300				
Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala 305 310 315 320				
Tyr Arg Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala 325 330 335				
Leu Gly Tyr Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile 340 345 350				
Leu Arg Asp Leu Gly Val Arg Asn Ile Lys Leu Leu Thr Asn Asn Pro 355 360 365				
Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg 370 375 380				
Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln 385 390 395 400				
Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe 405 410				

<210> 23

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

ES 2 435 990 T3

<400> 23

```

atgagaggat ctcacatca ccatcaccat gggatcgatc atatgtttca tccgatagaa      60
gaagcactgg acgctttaaa aaaaggcgaa gtcacatcgc ttgtagatga tgaagacaga      120
gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt      180
atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg      240
cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta      300
agcatagacc atcgtgaaac gaagacaggt atcagcgcctc aagaaagatc ttttaccgtt      360
caagcattgc tggacagcaa atccgtgcca tctgattttc agcgtccggg gcacatTTTT      420
ccactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt      480
gatcttgctg aagcttgtgg atctccagga gccggcgctca tttgtgaaat tatgaatgaa      540
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa      600
atgatcacca ttaaggattt gattcaatac cgttacaatc tgacaacact tgtcgcgctg      660
gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat      720
gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccggt cggagaagaa      780
ccggtattgg tccgggtgca ttcagaatgt ctcacagggt acgtgtttgg ctctcatcgc      840
tgtgattgca gaccgcagct gcacgccgcg ctgaaccaa ttgccgcaga aggccgtgga      900
gtgctcctgt acttgcgcca agaaggacga ggcacgggtt taatcaataa attaaaagct      960
tataagcttc aggaacaagg ctatgacacc gtagaagcca atgagggcgt tggattcttg     1020
ccggatcttc gcaactatgg catcggagca caaattttac gcgacctcgg tgtccggaat     1080
atgaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt     1140
atctcagaaa gagtgccgct tcaaattggag gcgaaagaac acaataaaaa atatttgcaa     1200
acaaaaatga acaagctagg tcatttactt catttctaa                                1239

```

<210> 24

<211> 412

5 <212> PRT

<213> *Bacillus subtilis*

<400> 24

```

Met Arg Gly Ser His His His His His His Gly Ile Asp His Met Phe
 1                               5                               10                               15

```

```

His Pro Ile Glu Glu Ala Leu Asp Ala Leu Lys Lys Gly Glu Val Ile
                20                               25                               30

```

ES 2 435 990 T3

Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe Val Ala  
 35 40 45

Leu Ala Glu His Ala Thr Pro Glu Val Ile Asn Phe Met Ala Thr His  
 50 55 60

Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Glu Ile Ala Asp Arg  
 65 70 75 80

Leu Asp Leu His Pro Met Val Glu His Asn Thr Asp Ser His His Thr  
 85 90 95

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
 100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
 115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
 130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
 145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
 165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
 180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
 195 200 205

Gln Tyr Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
 210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
 225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
 245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
 260 265 270

ES 2 435 990 T3

Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Arg Pro Gln Leu His  
 275 280 285

Ala Ala Leu Asn Gln Ile Ala Ala Glu Gly Arg Gly Val Leu Leu Tyr  
 290 295 300

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
 305 310 315 320

Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
 325 330 335

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
 340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn Asn Pro  
 355 360 365

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 25

<211> 1239

<212> ADN

5 <213> *Bacillus subtilis*

<400> 25

atgagaggat ctcacatca ccatcacat gggatcgatc atatgtttca tccgatagaa 60  
 gaagcactgg acgctttaa aaaaggcgaa gtcacatcg ttgtagatga tgaagacaga 120  
 gaaaatgaag gagactttgt ggctcttgcc gagcatgcaa cgccggaagt cattaacttt 180  
 atggcgacac atgggagagg actgatctgc acgccgctca gtgaggaaat cgcagacagg 240  
 cttgatcttc accctatggt tgagcataat acagactctc accacactgc atttaccgta 300  
 agcatagacc atcgtgaaac gaagacaggt atcagcgcctc aagaaagatc ttttaccggt 360  
 caagcattgc tggacagcaa atccgtgcca totgattttc agcgtccggg gcacatTTTT 420  
 ccaactgattg cgaaaaaagg aggtgtcctg aaaagagcgg gccatacaga agctgctggt 480



ES 2 435 990 T3

gatcttgctg aagcttgtgg atctccagga gccggcgctca tttgtgaaat tatgaatgaa 540  
gacggaacga tggcgagagt gcctgagctc attgaaattg cgaaaaagca tcaattaaaa 600  
atgatcacca ttaaggattt gattcaatgc cgttacaatc tgacaacact tgtcgcgagt 660  
gaagttgaca ttacgctgcc tactgatttt gggacattta aggtttatgg atacacaaat 720  
gaggtagatg gaaaagagca tgtcgcattt gtgatgggag atgtgccgtt cggagaagaa 780  
ccggtattgg tccgggtgca ttcagaatgt ctcacaggtg acgcgtttgg ctctcatcgc 840  
tgtgattgcg gaccgcagct gcacgccacg ctgaaccaa tggccgcaga aggccgtgga 900  
gtgctcctgt acttgcgcca agaaggacga ggcacgggtt taatcaataa attaaaagct 960  
tataagcttc aggaacaagg ctatgacacc gtagaagcca atgaggcgct tggattcttg 1020  
ccggatcttc gcaactatgg catcggagca caaatcttac gcgacctcgg tgtccggaat 1080  
atgaagcttt tgacgaataa tccgcgaaaa atcgcaggcc ttgaaggcta cggactcagt 1140  
atctcagaaa gaggccgct tcaaatggag gcgaaagaac acaataaaaa atatttgcaa 1200  
accaaatga acaagctagg tcatttactt catttctaa 1239

<210> 26

<211> 412

<212> PRT

5 <213> *Bacillus subtilis*

<400> 26

Met	Arg	Gly	Ser	His	His	His	His	His	His	Gly	Ile	Asp	His	Met	Phe
1			5					10						15	
His	Pro	Ile	Glu	Glu	Ala	Leu	Asp	Ala	Leu	Lys	Lys	Gly	Glu	Val	Ile
			20					25					30		
Ile	Val	Val	Asp	Asp	Glu	Asp	Arg	Glu	Asn	Glu	Gly	Asp	Phe	Val	Ala
			35				40					45			
Leu	Ala	Glu	His	Ala	Thr	Pro	Glu	Val	Ile	Asn	Phe	Met	Ala	Thr	His
	50					55					60				
Gly	Arg	Gly	Leu	Ile	Cys	Thr	Pro	Leu	Ser	Glu	Glu	Ile	Ala	Asp	Arg
65					70					75					80
Leu	Asp	Leu	His	Pro	Met	Val	Glu	His	Asn	Thr	Asp	Ser	His	His	Thr
				85					90					95	

ES 2 435 990 T3

Ala Phe Thr Val Ser Ile Asp His Arg Glu Thr Lys Thr Gly Ile Ser  
100 105 110

Ala Gln Glu Arg Ser Phe Thr Val Gln Ala Leu Leu Asp Ser Lys Ser  
115 120 125

Val Pro Ser Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu Ile Ala  
130 135 140

Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala Ala Val  
145 150 155 160

Asp Leu Ala Glu Ala Cys Gly Ser Pro Gly Ala Gly Val Ile Cys Glu  
165 170 175

Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Leu Ile Glu  
180 185 190

Ile Ala Lys Lys His Gln Leu Lys Met Ile Thr Ile Lys Asp Leu Ile  
195 200 205

Gln Cys Arg Tyr Asn Leu Thr Thr Leu Val Glu Arg Glu Val Asp Ile  
210 215 220

Thr Leu Pro Thr Asp Phe Gly Thr Phe Lys Val Tyr Gly Tyr Thr Asn  
225 230 235 240

Glu Val Asp Gly Lys Glu His Val Ala Phe Val Met Gly Asp Val Pro  
245 250 255

Phe Gly Glu Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu Thr  
260 265 270

Gly Asp Ala Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu His  
275 280 285

Ala Thr Leu Asn Gln Ile Ala Ala Glu Gly Arg Gly Val Leu Leu Tyr  
290 295 300

Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu Lys Ala  
305 310 315 320

Tyr Lys Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn Glu Ala  
325 330 335

ES 2 435 990 T3

Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala Gln Ile  
 340 345 350

Leu Arg Asp Leu Gly Val Arg Asn Met Lys Leu Leu Thr Asn Asn Pro  
 355 360 365

Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser Glu Arg  
 370 375 380

Val Pro Leu Gln Met Glu Ala Lys Glu His Asn Lys Lys Tyr Leu Gln  
 385 390 395 400

Thr Lys Met Asn Lys Leu Gly His Leu Leu His Phe  
 405 410

<210> 27

<211>61

<212> ADN

5 <213> Artificial

<220>

<223> Cebador RibA 1S

<400> 27

atgagaggat ctcaccatca ccacacacat gggatcgatc atatgtttca tccgatagaa 60

g 61

10 <210> 28

<211> 38

<212> ADN

<213> Artificial

<220>

15 <223> Cebador RibA 1AS

<400> 28

tataattgga tccttagaaa tgaagtaaat gacctagc 38

<210> 29

<211> 54

20 <212> ADN

<213> Artificial

<220>

<223> Cebador RibA 2S

<400> 29

attaatgaat tcattaaaga ggagaaatta actatgagag gatctcacca tcac 54

<210> 30

<211>31

<212> ADN

5 <213> Artificial

<220>

<223> Cebador RibANde+1

<400> 30

ggagggtttc atatgtttca tccgatagaa g 31

10 <210>31

<211>21

<212> ADN

<213> Artificial

<220>

15 <223> Cebador RibA4AS

<400> 31

taattaagct tggatcctta g 21

<210> 32

<211> 906

20 <212> ADN

<213> *Ashbya gossypii*

<400> 32

ES 2 435 990 T3

atgactgaat acacagtgcc agaagtgagg tgtgtcgcac ggcgcgcgat accgacggta	60
cagggcaccg atgtcttctt ccatctatac cacaactcga tcgacagcaa ggaacaccta	120
gcgattgtct tcggcgagaa catacgcctc cggagtctgt tccggtagcg gaaagacgac	180
acgcagcagg cgcggatggg gcggggcgcc tacgtgggccc agctgtaccc cgggcggacc	240
gagggcagacg cggatcggcg tcagggcctg gagctgcggg ttgatgagac agggcagctg	300
gtgggtggagc gggcgacgac gtggaccagg gagccgacac tgggtcgggct gcactcggag	360
tgttacacgg gcgagacggc gtggagcgcg cgggtgcgact gcggggagca gttcgaccag	420
gcgggtaagc tgatggctgc ggcgacagag ggcgaggtgg ttggcgggtgc ggggcacggc	480
gtgatcgtgt acctgcggca ggagggccgc ggcacgcggc taggcgagaa gctgaaggcg	540
tacaacctgc aggacctggg cgcggacacg gtgcaggcga acgagctgct caaccacct	600
gcggacgcgc gcgacttctc gttggggcgc gcaatcctac tggacctcgg tatcgaggac	660
atccggttgc tcacgaataa ccccgacaag gtgcagcagg tgcactgtcc gccggcgcta	720
cgctgcatcg agcgggtgcc catggtgccg ctttcatgga ctcagcccac acagggcgtg	780
cgctcgcgcg agctggacgg ctacctgcgc gccaaaggtcg agcgcgatggg gcacatgctg	840
cagcggccgc tgggtgetgca cacgtctgcg gcggccgagc tccccgcgc caacacacac	900
atataa	906

<210> 33

<211> 301

5 <212> PRT

<213> *Ashbya gossypii*

<400> 33

ES 2 435 990 T3

Met Thr Glu Tyr Thr Val Pro Glu Val Arg Cys Val Ala Arg Ala Arg  
 1 5 10 15

Ile Pro Thr Val Gln Gly Thr Asp Val Phe Leu His Leu Tyr His Asn  
 20 25 30

Ser Ile Asp Ser Lys Glu His Leu Ala Ile Val Phe Gly Glu Asn Ile  
 35 40 45

Arg Ser Arg Ser Leu Phe Arg Tyr Arg Lys Asp Asp Thr Gln Gln Ala  
 50 55 60

Arg Met Val Arg Gly Ala Tyr Val Gly Gln Leu Tyr Pro Gly Arg Thr  
 65 70 75 80

Glu Ala Asp Ala Asp Arg Arg Gln Gly Leu Glu Leu Arg Phe Asp Glu  
 85 90 95

Thr Gly Gln Leu Val Val Glu Arg Ala Thr Thr Trp Thr Arg Glu Pro  
 100 105 110

Thr Leu Val Arg Leu His Ser Glu Cys Tyr Thr Gly Glu Thr Ala Trp  
 115 120 125

Ser Ala Arg Cys Asp Cys Gly Glu Gln Phe Asp Gln Ala Gly Lys Leu  
 130 135 140

Met Ala Ala Ala Thr Glu Gly Glu Val Val Gly Gly Ala Gly His Gly  
 145 150 155 160

Val Ile Val Tyr Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Gly Glu  
 165 170 175

Lys Leu Lys Ala Tyr Asn Leu Gln Asp Leu Gly Ala Asp Thr Val Gln  
 180 185 190

ES 2 435 990 T3

Ala Asn Glu Leu Leu Asn His Pro Ala Asp Ala Arg Asp Phe Ser Leu  
 195 200 205

Gly Arg Ala Ile Leu Leu Asp Leu Gly Ile Glu Asp Ile Arg Leu Leu  
 210 215 220

Thr Asn Asn Pro Asp Lys Val Gln Gln Val His Cys Pro Pro Ala Leu  
 225 230 235 240

Arg Cys Ile Glu Arg Val Pro Met Val Pro Leu Ser Trp Thr Gln Pro  
 245 250 255

Thr Gln Gly Val Arg Ser Arg Glu Leu Asp Gly Tyr Leu Arg Ala Lys  
 260 265 270

Val Glu Arg Met Gly His Met Leu Gln Arg Pro Leu Val Leu His Thr  
 275 280 285

Ser Ala Ala Ala Glu Leu Pro Arg Ala Asn Thr His Ile  
 290 295 300

<210> 34

<211> 591

<212> ADN

5 <213> *Escherichia coli*

<400> 34

atgcagctta aacgtgtggc agaagccaaa ctgccaaccc catggggcga tttcctgatg 60

gtgggatttg aagaactggc aaccggacac gatcatgtcg cgctagtcta tggcgatatt 120

tccgggcata ccccggtact tgcgcgcgtc cattccgaat gtctgaccgg tgacgccttg 180

ttcagcttgc gctgcgattg tggcttccag ctccaagcgg cattgacgca aattgccgag 240

gaaggccgtg gtatthttgct gtatcacctg caggaaggtc gtaacattgg tctgctgaat 300

aaaatccgcg cttacgcact gcaggatcaa ggttacgata ccgtagaggc taaccaccag 360

ttaggcttcg ccgctgatga gcgcgacttc actctttgcg ctgatatggt caaactcctt 420

ggcgtcaatg aagtccgctt gttaaccaat aaccggaaaa aagtcgaaat tctgaccgaa 480

gcagggatta atattgttga acgcgtacca ttgattgtag gtcgtaacct caataacgaa 540

cattatctcg ataccaaagc cgagaaaatg ggccatttgc tgaacaaata a 591

<210> 35

<211> 196

10 <212> PRT

ES 2 435 990 T3

<213> *Escherichia coli*

<400> 35

Met Gln Leu Lys Arg Val Ala Glu Ala Lys Leu Pro Thr Pro Trp Gly  
1 5 10 15

Asp Phe Leu Met Val Gly Phe Glu Glu Leu Ala Thr Gly His Asp His  
20 25 30

Val Ala Leu Val Tyr Gly Asp Ile Ser Gly His Thr Pro Val Leu Ala  
35 40 45

Arg Val His Ser Glu Cys Leu Thr Gly Asp Ala Leu Phe Ser Leu Arg  
50 55 60

Cys Asp Cys Gly Phe Gln Leu Glu Ala Ala Leu Thr Gln Ile Ala Glu  
65 70 75 80

Glu Gly Arg Gly Ile Leu Leu Tyr His Arg Gln Glu Gly Arg Asn Ile  
85 90 95

Gly Leu Leu Asn Lys Ile Arg Ala Tyr Ala Leu Gln Asp Gln Gly Tyr  
100 105 110

Asp Thr Val Glu Ala Asn His Gln Leu Gly Phe Ala Ala Asp Glu Arg  
115 120 125

Asp Phe Thr Leu Cys Ala Asp Met Phe Lys Leu Leu Gly Val Asn Glu  
130 135 140

Val Arg Leu Leu Thr Asn Asn Pro Lys Lys Val Glu Ile Leu Thr Glu  
145 150 155 160

Ala Gly Ile Asn Ile Val Glu Arg Val Pro Leu Ile Val Gly Arg Asn  
165 170 175

Pro Asn Asn Glu His Tyr Leu Asp Thr Lys Ala Glu Lys Met Gly His  
180 185 190

Leu Leu Asn Lys  
195

<210> 36

5 <211> 1269

<212> ADN

<213> *Corynebacterium glutamicum*



ES 2 435 990 T3

<400> 36

```

gtgagtgaac atgagcaggc acacagccaa ttagattctg ttgaagaggc catcgctgac      60
atcgctgcgg gtaaagccgt cgtggtggta gatgatgaag atcgtgaaaa tgaaggcgac      120
atcatctttg cgcgccgaatt agccactcca gaattagtcg ctttcatggt gcgttattcc      180
tcgggataca tctgtgcgcc attaaccgca aaggatgcag atcgtcttga tctgcctccg      240
atgaccgcgc acaatcagga tgcccgcggc accgcttaca ccgtgaccgt tgatgccaac      300
accggcacca caggcatttc tgcaacagac cgcgcccaca ctttgcgctt gcttgctgat      360
ccagaagccg accgcacgga tttcacccgt cccggacacg ttgtgccact gcgtgctcgt      420
gaaggtggcg tcttgggtgcg cgctggacac accgaagcag ctgtcgattt ggctcgcgct      480
gcaggcctgc gccacgcagg tgttatctgc gaagtggta gtgaagagga cccaccggc      540
atggctcggg ttctgagct gcgcccttc tgcgatgagc acgatctgaa gctgatctct      600
attgagcagc tcattgagtg gcgtcgcaag aatgaaattt tggtggagcg ccaggtggaa      660
actgtgctgc ctaccgattt cggcacgttc aaggctggtg gttaccgttc catcatcgat      720
ggcaccgagc ttgttgccat tgttgccggc gacgtggcat ccgacggtgg cgaaaacgtc      780
ctggttcgag tccactctga gtgcttgact ggtgatgttt ttggatcccg gcgctgcgac      840
tgtggacagc agctgcacga gtctttgcgc ctgatccagg aagctggtcg gggagtagtg      900
gtgtacatgc gtgggcatga gggacgaggc attggtctgc tcgccaagct acgcgctac      960
caactccagg atgaaggtgc cgacaccgtc gatgcccaacc tcgcacttgg tcttcagcc     1020
gatgcccgcg aatttggcac cagcgcaccag attctctacg acttgggtgt gcgctcgctc     1080
aacttgatca gcaacaacc agccaagaag gtgggacttg aaggccacgg catttcatt     1140
gccagccgaa ccccatccc tgttgctggt catgaagaca atgttcgata cctgaaaacc     1200
aagcgtgacc gcatgggaca tgacctcca gatgtcgcac tgtgggaaca agagcaccca     1260
gaaaactaa                                     1269

```

<210> 37

<211> 422

5 <212> PRT

<213> *Corynebacterium glutamicum*

<400> 37

```

Val Ser Glu His Glu Gln Ala His Ser Gln Leu Asp Ser Val Glu Glu
1           5           10           15

```

```

Ala Ile Ala Asp Ile Ala Ala Gly Lys Ala Val Val Val Val Asp Asp

```

ES 2 435 990 T3

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

ES 2 435 990 T3

Gly Glu Asn Val Leu Val Arg Val His Ser Glu Cys Leu Thr Gly Asp  
 260 265 270

Val Phe Gly Ser Arg Arg Cys Asp Cys Gly Gln Gln Leu His Glu Ser  
 275 280 285

Leu Arg Leu Ile Gln Glu Ala Gly Arg Gly Val Val Val Tyr Met Arg  
 290 295 300

Gly His Glu Gly Arg Gly Ile Gly Leu Leu Ala Lys Leu Arg Ala Tyr  
 305 310 315 320

Gln Leu Gln Asp Glu Gly Ala Asp Thr Val Asp Ala Asn Leu Ala Leu  
 325 330 335

Gly Leu Pro Ala Asp Ala Arg Glu Phe Gly Thr Ser Ala Gln Ile Leu  
 340 345 350

Tyr Asp Leu Gly Val Arg Ser Leu Asn Leu Ile Ser Asn Asn Pro Ala  
 355 360 365

Lys Lys Val Gly Leu Glu Gly His Gly Ile Ser Ile Ala Ser Arg Thr  
 370 375 380

Pro Ile Pro Val Ala Val His Glu Asp Asn Val Arg Tyr Leu Lys Thr  
 385 390 395 400

Lys Arg Asp Arg Met Gly His Asp Leu Pro Asp Val Ala Leu Trp Glu  
 405 410 415

Gln Glu His Pro Glu Asn  
 420

<210> 38

<211> 1197

<212> ADN

5 <213> *Bacillus amyloliquefaciens*

<400> 38

ES 2 435 990 T3

atgtttcatc	cgatagaaga	ggcattagaa	gcgctgaaaa	aaggtgaagt	catcatcgtt	60
gtcgatgatg	aagacagaga	aaacgaagga	gatttcgtag	cgctcgctga	gcatgctacg	120
cctgaagtgg	tgaattttat	ggcgacccac	gggagaggcc	tgatctgcac	gccgctttct	180
gaagacatcg	ccggccggct	ggatcttcat	ccaatggctg	atcataatac	agactcgcac	240
gagaccgctg	ttacagtcag	cattgaccac	aagctgacaa	aaacgggaat	cagcgcctcag	300
gaacgttcct	ttacgattca	ggcgcttttg	gacgaagaat	ctgtgcctgg	cgattttcag	360
cgtccgggtc	atatttttcc	cttaatagca	aaaaaaggag	gcgtcctgaa	gcgggcccgc	420
cacacggaag	cagccgttga	cctggcaaaa	gcatgcgggt	ctcaaggagc	ggacgctcatt	480
tgtgaaatta	tgaatgaaga	cggcacaatg	gcgagagtgc	ctgagattag	cgagattgctg	540
aaaagccacc	agctgaaaat	gattacgata	aaagacttaa	tagaataccg	ctacaacatt	600
acaacacttg	tgaacagaga	agttgacatt	acgctgccga	ctgacttcgg	cacgttccgg	660
gtttacggat	atacaaacga	ggtggacgga	aaagaacatc	tcgcctttgt	catgggcat	720
gtcccgttta	acagcggacc	cgttcttgtc	agagtgcact	cagaatgcct	gaccggcgat	780
gtgtttgcat	cccaccgctg	tgattgctgg	cctcagcttc	atgccgcgct	gcgccaaatt	840
gccgaagaag	gccgcggcgt	tctattgtat	ttgcgtcagg	aaggcagagg	aatcggcttc	900
atcaataagc	tgaagcgta	tcgattgcag	gaacaagggt	acgacacggt	tgaagcgaac	960
gaagcgcctg	gctttctgcc	tgacttgccg	aactatggca	tcggcgccca	gattctccgc	1020
gatttagggg	ttcagcatat	gaaactttta	accaataacc	cccggaaaat	cgccggcctt	1080
gaagggtagc	gactaagcat	ttcagatcgg	gtgccgcttc	aaatggaagc	gagtgagcac	1140
aacaagcagt	atttacaac	caaaatgaaa	aaactcggac	acttgcttca	tttctaa	1197

<210> 39

<211> 398

5

<212> PRT

<213> *Bacillus amyloliquefaciens*

<400> 39

ES 2 435 990 T3

Met Phe His Pro Ile Glu Glu Ala Leu Glu Ala Leu Lys Lys Gly Glu  
1 5 10 15

Val Ile Ile Val Val Asp Asp Glu Asp Arg Glu Asn Glu Gly Asp Phe  
20 25 30

Val Ala Leu Ala Glu His Ala Thr Pro Glu Val Val Asn Phe Met Ala  
35 40 45

Thr His Gly Arg Gly Leu Ile Cys Thr Pro Leu Ser Glu Asp Ile Ala  
50 55 60

Gly Arg Leu Asp Leu His Pro Met Val Asp His Asn Thr Asp Ser His  
65 70 75 80

ES 2 435 990 T3

Glu Thr Ala Phe Thr Val Ser Ile Asp His Lys Leu Thr Lys Thr Gly  
 85 90 95  
 Ile Ser Ala Gln Glu Arg Ser Phe Thr Ile Gln Ala Leu Leu Asp Glu  
 100 105 110  
 Glu Ser Val Pro Gly Asp Phe Gln Arg Pro Gly His Ile Phe Pro Leu  
 115 120 125  
 Ile Ala Lys Lys Gly Gly Val Leu Lys Arg Ala Gly His Thr Glu Ala  
 130 135 140  
 Ala Val Asp Leu Ala Lys Ala Cys Gly Ser Gln Gly Ala Asp Val Ile  
 145 150 155 160  
 Cys Glu Ile Met Asn Glu Asp Gly Thr Met Ala Arg Val Pro Glu Ile  
 165 170 175  
 Ser Glu Ile Ala Lys Ser His Gln Leu Lys Met Ile Thr Ile Lys Asp  
 180 185 190  
 Leu Ile Glu Tyr Arg Tyr Asn Ile Thr Thr Leu Val Asn Arg Glu Val  
 195 200 205  
 Asp Ile Thr Leu Pro Thr Asp Phe Gly Thr Phe Arg Val Tyr Gly Tyr  
 210 215 220  
 Thr Asn Glu Val Asp Gly Lys Glu His Leu Ala Phe Val Met Gly Asp  
 225 230 235 240  
 Val Pro Phe Asn Ser Gly Pro Val Leu Val Arg Val His Ser Glu Cys  
 245 250 255  
 Leu Thr Gly Asp Val Phe Ala Ser His Arg Cys Asp Cys Gly Pro Gln  
 260 265 270  
 Leu His Ala Ala Leu Arg Gln Ile Ala Glu Glu Gly Arg Gly Val Leu  
 275 280 285  
 Leu Tyr Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Ile Asn Lys Leu  
 290 295 300  
 Lys Ala Tyr Arg Leu Gln Glu Gln Gly Tyr Asp Thr Val Glu Ala Asn  
 305 310 315 320

ES 2 435 990 T3

Glu Ala Leu Gly Phe Leu Pro Asp Leu Arg Asn Tyr Gly Ile Gly Ala  
 325 330 335

Gln Ile Leu Arg Asp Leu Gly Val Gln His Met Lys Leu Leu Thr Asn  
 340 345 350

Asn Pro Arg Lys Ile Ala Gly Leu Glu Gly Tyr Gly Leu Ser Ile Ser  
 355 360 365

Asp Arg Val Pro Leu Gln Met Glu Ala Ser Glu His Asn Lys Gln Tyr  
 370 375 380

Leu Gln Thr Lys Met Lys Lys Leu Gly His Leu Leu His Phe  
 385 390 395

<210> 40

<211> 1194

<212> ADN

5 <213> *Bacillus cereus*

<400> 40

atgtttcatc gtattgaaga agctctagaa gatttaaaaa aaggtaaagt cgttatcgtg 60  
 tgtgatgatg aaaaccgaga aatgaaggc gattttattg cttagcaga gtacattaca 120  
 ccagaaacaa taaattttat gattacacat ggccgtggtc tcgtttgtgt accgattacg 180  
 gaaggatagc cagaacgtct acaattagaa ccaatggtat ctcataatac agattcacat 240  
 catactgcgt ttacagtgag cattgacat gtctctacaa caacagggat tagcgctcac 300  
 gaacgtgcaa ctacgataca agaattgta aaccccgcat caaaagggtc tgatttcaat 360  
 cgacctggac atatctttcc attaattgcg aaagaaggcg gtgtcctgcg tcgtgcaggt 420  
 catacagaag ctgctgttga ttagcaaag ctatgcggtg ccgaaccagc tggagttatt 480  
 tgcgagatta taaatgagga cggcagatg gcacgtgtac ctgatttaat agaatgcgca 540  
 aaacaatttg atataaaaat gattacaata gaagatttaa ttgcttaccg ccgccatcat 600  
 gaaacacttg tgacgagaga agcggaaatt acattaccta cagatttcgg tactttccac 660  
 gcaattggct attctaactc attagatagc aaagaacata tcgcacttgt aaaagggtgat 720  
 atttcaacag gtgaaccggt acttgtagct gttcattctg aatgcttaac aggagatgta 780  
 ttcggttcac atcgctgca ttgcggacca caactccatg cagcacttgc tcaaattgag 840  
 cgtgaaggaa aagggtgtct tctttatag aggcaagaag gaagaggcat tgggcttctt 900  
 aataagcttc gtgcttataa attacaagaa gaaggattcg atactgtaga agcaaatgaa 960

ES 2 435 990 T3

aaactcggct tcctgctga tcttcgtgat tacggtatcg gtgctcaaat attaaaagat 1020  
 ttaggtttac agagttttacg attattaacg aataacccaa gaaaaattgc tggcttacia 1080  
 ggttacgatt tagaagtagt cgagcgtgta ccgttgcaaa tgccagcaaa agaagagaat 1140  
 aaatcgtatt taaaaacgaa agtaaacaaa ttaggacact tactaaactt ataa 1194

<210>41

<211> 397

<212> PRT

5 <213> *Bacillus cereus*

<400> 41

Met	Phe	His	Arg	Ile	Glu	Glu	Ala	Leu	Glu	Asp	Leu	Lys	Lys	Gly	Lys	1	5	10	15
Val	Val	Ile	Val	Cys	Asp	Asp	Glu	Asn	Arg	Glu	Asn	Glu	Gly	Asp	Phe	20	25	30	
Ile	Ala	Leu	Ala	Glu	Tyr	Ile	Thr	Pro	Glu	Thr	Ile	Asn	Phe	Met	Ile	35	40	45	
Thr	His	Gly	Arg	Gly	Leu	Val	Cys	Val	Pro	Ile	Thr	Glu	Gly	Tyr	Ala	50	55	60	
Glu	Arg	Leu	Gln	Leu	Glu	Pro	Met	Val	Ser	His	Asn	Thr	Asp	Ser	His	65	70	75	80
His	Thr	Ala	Phe	Thr	Val	Ser	Ile	Asp	His	Val	Ser	Thr	Thr	Thr	Gly	85	90	95	
Ile	Ser	Ala	His	Glu	Arg	Ala	Thr	Thr	Ile	Gln	Glu	Leu	Leu	Asn	Pro	100	105	110	
Ala	Ser	Lys	Gly	Ala	Asp	Phe	Asn	Arg	Pro	Gly	His	Ile	Phe	Pro	Leu	115	120	125	
Ile	Ala	Lys	Glu	Gly	Gly	Val	Leu	Arg	Arg	Ala	Gly	His	Thr	Glu	Ala	130	135	140	
Ala	Val	Asp	Leu	Ala	Lys	Leu	Cys	Gly	Ala	Glu	Pro	Ala	Gly	Val	Ile	145	150	155	160
Cys	Glu	Ile	Ile	Asn	Glu	Asp	Gly	Thr	Met	Ala	Arg	Val	Pro	Asp	Leu	165	170	175	



ES 2 435 990 T3

Ile Glu Cys Ala Lys Gln Phe Asp Ile Lys Met Ile Thr Ile Glu Asp  
 180 185 190

Leu Ile Ala Tyr Arg Arg His His Glu Thr Leu Val Thr Arg Glu Ala  
 195 200 205

Glu Ile Thr Leu Pro Thr Asp Phe Gly Thr Phe His Ala Ile Gly Tyr  
 210 215 220

Ser Asn Ser Leu Asp Thr Lys Glu His Ile Ala Leu Val Lys Gly Asp  
 225 230 235 240

Ile Ser Thr Gly Glu Pro Val Leu Val Arg Val His Ser Glu Cys Leu  
 245 250 255

Thr Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys Gly Pro Gln Leu  
 260 265 270

His Ala Ala Leu Ala Gln Ile Glu Arg Glu Gly Lys Gly Val Leu Leu  
 275 280 285

Tyr Met Arg Gln Glu Gly Arg Gly Ile Gly Leu Leu Asn Lys Leu Arg  
 290 295 300

Ala Tyr Lys Leu Gln Glu Glu Gly Phe Asp Thr Val Glu Ala Asn Glu  
 305 310 315 320

Lys Leu Gly Phe Pro Ala Asp Leu Arg Asp Tyr Gly Ile Gly Ala Gln  
 325 330 335

Ile Leu Lys Asp Leu Gly Leu Gln Ser Leu Arg Leu Leu Thr Asn Asn  
 340 345 350

Pro Arg Lys Ile Ala Gly Leu Gln Gly Tyr Asp Leu Glu Val Val Glu  
 355 360 365

Arg Val Pro Leu Gln Met Pro Ala Lys Glu Glu Asn Lys Ser Tyr Leu  
 370 375 380

Gln Thr Lys Val Asn Lys Leu Gly His Leu Leu Asn Leu  
 385 390 395

<210> 42

<211> 1215

<212> ADN

ES 2 435 990 T3

<213> *Bacillus halodurans*

<400> 42

```

atggacaaaa agctatattga tccgattgaa gaagcaatat atgaattaat gcaaggtcga      60
gtcgtgatcg tttgtgatga tgaggatcgg gaaaacgaag gggattttgt agcccttgct      120
gaaaaagcaa caccagaagt gattaacttc atgatcacgc atggccgtgg tctcgtttgc      180
acgccaatca cggaagagcg ggcaaaggaa ttagatcttg tccccatggg ggaccataat      240
accgatcccc atggtacggc gtttaccgtc agcattgatc atcaaatgac gaccacagga      300
atctctgccc atgaacgggc tatgacgatt caggcgtaa ttgataagaa aacgaaaaag      360
caccacttca aacgaccagg tcacattttc cccctaatag cgaaaaacgg aggagtactc      420
cgacggggccg gtcatacaga agcggccggt gatctagctc gtttgtcagg cgctgagccg      480
gcaggggtta tttgtgaaat cattaagaa gatggttcaa tggcacgagt tcctgatttg      540
cgaaaaatcg ccgatcagtt tgaactgaag atgatcacia ttaaagattt aatcgaatat      600
cgtcaccgta aagacaagct tgtcaagcgt gaagtagata tttccttacc gacggatttc      660
ggctcattcc gtgcaatcgg ttatacagat gtcattgatg gaaaagagag tgtcgtttta      720
gtgaaaggac agattgttga aggtgaacca acactcgttc gtgttcactc cgaatgttta      780
acaggtgatg tgttcgggtc tcaccgttgc gattgtggcc cacaactcca ggcagctctc      840
acacaaatcg agcaacaagg caaagggata ctcccttata tgcgtcaaga gggtcgtggt      900
atcggctctca tgaataagtt gaaggcatac aagcttcaag aagaaggcta tgatactgta      960
gaagcaaatg agaaattagg ctttcctgct gatcttcggg actatggaat gggcgcgcaa     1020
atcttacgcg acttaggtgt gtcaaaaatg cgcctcctta caaacatcc gcgaaaaaatt     1080
acgggcttga aagggtatgg ccttgaagtg gttgaacggg tgccgctcca attacctcat     1140
aacaaagata atgagcgcta tttgaaaaca aagcacgaaa agttaggaca tctgctaaat     1200
tttactcatt cgtaa                                                                1215

```

<210> 43

5 <211> 404

<212> PRT

<213> *Bacillus halodurans*

<400> 43

```

Met Asp Lys Lys Leu Phe Asp Pro Ile Glu Glu Ala Ile Tyr Glu Leu
 1           5           10           15

```

```

Met Gln Gly Arg Val Val Ile Val Cys Asp Asp Glu Asp Arg Glu Asn

```



ES 2 435 990 T3

Ser Glu Cys Leu Thr Gly Asp Val Phe Gly Ser His Arg Cys Asp Cys  
 260 265 270

Gly Pro Gln Leu Gln Ala Ala Leu Thr Gln Ile Glu Gln Gln Gly Lys  
 275 280 285

Gly Ile Leu Leu Tyr Met Arg Gln Glu Gly Arg Gly Ile Gly Leu Met  
 290 295 300

Asn Lys Leu Lys Ala Tyr Lys Leu Gln Glu Glu Gly Tyr Asp Thr Val  
 305 310 315 320

Glu Ala Asn Glu Lys Leu Gly Phe Pro Ala Asp Leu Arg Asp Tyr Gly  
 325 330 335

Met Gly Ala Gln Ile Leu Arg Asp Leu Gly Val Ser Lys Met Arg Leu  
 340 345 350

Leu Thr Asn Asn Pro Arg Lys Ile Thr Gly Leu Lys Gly Tyr Gly Leu  
 355 360 365

Glu Val Val Glu Arg Val Pro Leu Gln Leu Pro His Asn Lys Asp Asn  
 370 375 380

Glu Arg Tyr Leu Lys Thr Lys His Glu Lys Leu Gly His Leu Leu Asn  
 385 390 395 400

Phe Thr His Ser

## REIVINDICACIONES

- 5 1. Una GTP ciclohidrolasa II modificada seleccionada de *Bacillus* que muestra un incremento en la actividad específica de al menos alrededor de 10 % en comparación con la enzima no modificada correspondiente, en la que se han sustituido restos de aminoácidos en las posiciones que corresponden a las posiciones 261, 270, 276, 279, 308 y/o 347 de SEQ ID NO:2, conduciendo dichas sustituciones a restos de aminoácidos seleccionados del grupo que consiste en:
- (a) alanina en una posición que corresponde a la posición 261 de SEQ ID NO:2,
  - (b) alanina en una posición que corresponde a la posición 270 de SEQ ID NO:2,
  - (c) treonina en una posición que corresponde a la posición 276 de SEQ ID NO:2,
  - 10 (d) arginina en una posición que corresponde a la posición 279 de SEQ ID NO:2,
  - (e) arginina en una posición que corresponde a la posición 308 de SEQ ID NO:2, y
  - (f) isoleucina en una posición que corresponde a la posición 347 de SEQ ID NO:2.
- 15 2. Una GTP ciclohidrolasa II modificada según la reivindicación 1, que comprende mutaciones seleccionadas del grupo que consiste en V261A, G270A, A276T, Q279R, K308R, y M347I, en la que dichas posiciones corresponden a las posiciones respectivas de SEC ID NO:2.
3. Una GTP ciclohidrolasa II modificada según la reivindicación 2, seleccionada de *Bacillus subtilis*.
4. Una GTP ciclohidrolasa II modificada según una cualquiera de las reivindicaciones 1 a 3, que comprende una combinación de mutaciones 276T/279R/308R/347I, en la que las posiciones corresponden a las posiciones respectivas de SEC ID NO:2.
- 20 5. La GTP ciclohidrolasa II modificada según la reivindicación 1, en la que la secuencia de la GTP ciclohidrolasa II no modificada se selecciona del grupo que consiste en las secuencias ID NOs:2, 39, 41, y 43.
6. Un polinucleótido que comprende una secuencia nucleotídica que codifica una GTP ciclohidrolasa II modificada según una cualquiera de las reivindicaciones 1 a 4.
- 25 7. El polinucleótido que codifica la GTP ciclohidrolasa II modificada según la reivindicación 6, seleccionándose dicho polinucleótido del grupo que consiste en las secuencias ID NOs:6, 8, 10, 12, 14, 16, 18, 20, 22, 24 y 26.
8. Una célula hospedante que comprende un polinucleótido según la reivindicación 6 ó 7.
9. La célula hospedante según la reivindicación 8, que se selecciona de *Bacillus*.
10. Un procedimiento para producir riboflavina, un precursor de riboflavina, FMN, FAD, o un derivado de los mismos, que comprende:
- 30 (a) cultivar la célula hospedante según la reivindicación 8 ó 9 en un medio adecuado; y
- (b) opcionalmente separar riboflavina, un precursor de riboflavina, FMN, FAD, o un derivado de los mismos, del medio.
11. Un procedimiento para la producción de riboflavina, un precursor de riboflavina, FMN, FAD, o un derivado de los mismos, que comprende:
- 35 (a) proporcionar una célula hospedante seleccionada de *Bacillus* que comprende una GTP ciclohidrolasa II modificada según la reivindicación 1;
- (b) cultivar dicha célula hospedante en un medio adecuado, y
- (c) opcionalmente separar riboflavina, un precursor de riboflavina, FMN, FAD, o un derivado de los mismos, del medio.
- 40 12. Un procedimiento según una cualquiera de las reivindicaciones 10 a 11, en el que la célula hospedante se selecciona de *Bacillus subtilis*.
13. El uso de una GTP ciclohidrolasa II modificada según una cualquiera de las reivindicaciones 1 a 5, o un polinucleótido según la reivindicación 7, para incrementar la producción de riboflavina, un precursor de riboflavina, FMN, FAD, o un derivado de los mismos.

Figura 1

- gch2\_cangu: geneseqp:aay69776 (*Candida guilliermondii*)
- gch2\_ashgo: TrEMBL: CAA02912: (*Ashbya gossypii* (*Eremothecium gossypii*))
- gch2\_yeast: SWISS-PROT: gch2\_yeast (*Saccharomyces cerevisiae*)
- 5 gch2\_neucr: TrEMBL: Q871B3 (*Neurospora crassa*)
- gch2\_schpo: TrEMBL: Q9P7M9 (*Schizosaccharomyces pombe*)
- gch2\_arcfu: SWISS-PROT: gch2\_arcfu (*Archaeoglobus fulgidus*)
- gch2\_strco: SWISS-PROT: gch2\_strco (*Streptomyces coelicolor*)
- gch2\_helpj: SWISS-PROT: gch2\_helpj (*Helicobacter pylori* J99)
- 10 gch2\_helpy: SWISS-PROT: gch2\_helpy (*Helicobacter pylori*)
- gch2\_pyrfu: TrEMBL: Q8U4L7 (*Pyrococcus furiosus*)
- gch2\_thema: SWISS-PROT: gch2\_thema (*Thermotoga maritima*)
- gch2\_chlmu: SWISS-PROT: gch2\_chlmu (*Chlamydia muridarum*)
- gch2\_chltr: SWISS-PROT: gch2\_chltr (*Chlamydia trachomatis*)
- 15 gch2\_chlca: TrEMBL: AAP05635 (*Chlamydia caviae* GPIC)
- gch2\_chlpn: SWISS-PROT: gch2\_chlpn (*Chlamydia pneumoniae*)
- gch2\_arath: SWISS-PROT: gch2\_arath (*Arabidopsis thaliana*)
- gch2\_lyces: TrEMBL: CAC09119 (*Lycopersicon esculentum*)
- gch2\_orysa: TrEMBL: AA072560 (*Oryza sativum*)
- 20 gch2\_alceu: TrEMBL: Q9F184 (*Alcaligenes eutrophus*)
- gch2\_neima: SWISS-PROT: gch2\_neima (*Neisseria meningitidis* (serogrupo A))
- gch2\_neimb: SWISS-PROT: gch2\_neimb (*Neisseria meningitidis* (serogrupo B))
- gch2\_psepk: SWISS-PROT: gch2\_psepk (*Pseudomonas putida* (cepa KT2440))
- gch2\_psesm: SWISS-PROT: gch2\_psesm (*Pseudomonas syringae* (pv. tomate ))
- 25 gch2\_actac: TrEMBL: Q9JRR0 (*Actinobacillus actinomycetemcomitans* (*Haemophilus actinomycetemcomitans*))
- gch2\_haein: SWISS-PROT: gch2\_pasmu gch2\_haein (*Haemophilus influenzae*)
- gch2\_pasmu: SWISS-PROT: ( *Pasteurella multocida* )
- gch2\_ec06: TrEMBL: Q8FHU5 (*Escherichia coli* 06)
- gch2\_ecoli: SWISS-PROT: gch2\_ecoli (*Escherichia coli*)
- 30 gch2\_salty: TrEMBL: Q8XFY7 (*Salmonella typhimurium*)
- gch2\_yerpe: TrEMBL: Q8ZEF0 (*Yersinia pestis*) \_
- gch2\_bucal: SWISS-PROT: gch2\_bucal (*Buchnera aphidicola* (subsp. *Acyrtosiphon pisum*) (*Acyrtosiphon pisum* bacteria simbiótica ))
- gch2\_bucap: SWISS-PROT: gch2\_bucap (*Buchnera aphidicola* (subsp. *Schizaphis graminum*))
- 35 gch2\_wigbr: SWISS-PROT: gch2\_wigbr (*Wigglesworthia glossinidia brevipalpis*)
- gch2\_bucbp: SWISS-PROT: gch2\_wigbr (*Buchnera aphidicola* (subsp. *Baizongia pistaciae*))
- gch2\_mytle: TrEMBL: Q9CCP4 (*Mycobacterium leprae*)

## ES 2 435 990 T3

- gch2\_myctu: SWISS-PROT: gch2\_myctu (*Mycobacterium tuberculosis*)  
gch2\_coref: TrEMBL: Q8FT57 (*Corynebacterium efficiens*)  
gch2\_corgl: GENESEQP: AAB79913 (*Corynebacterium glutamicum*)  
gch2\_coram: SWISS-PROT: gch2\_coram (*Corynebacterium ammoniagenes* (*Brevibacterium ammoniagenes*))
- 5 gch2\_staa: TrEMBL: Q8NW14 (*Staphylococcus aureus* (cepa MW2))  
gch2\_staep: GENESEQP: ABP40248 (*Staphylococcus epidermidis*)  
gch2\_actpl: SWISS-PROT: gch2\_actpl (*Actinobacillus pleuropneumoniae*)  
gch2\_lacla: TrEMBL: Q9CGU7 (*Lactococcus lactis* (subsp. *lactis*) (*Streptococcus lactis*))  
gch2\_stcag: TrEMBL: Q8E658 (*Streptococcus agalactiae* (serotipo III))
- 10 gch2\_stcpn: TrEMBL: Q8DRF1 (*Streptococcus pneumoniae* (cepa ATCC BAA-255 / R6))  
gch2\_cloac: TrEMBL: Q97LG9 (*Clostridium acetobutylicum*)  
gch2\_fusnu: TrEMBL: Q8RIR1 (*Fusobacterium nucleatum* (subsp. *nucleatum*))  
gch2\_anasp: TrEMBL: Q8RIR1 (*Anabaena* sp. (cepa PCC 7120))  
gch2\_syny3: SWISS-PROT: gch2\_syny3 (*Synechocystis* sp. (cepa PCC 6803))
- 15 gch2\_synel: TrEMBL: Q8DI64 *Synechococcus elongatus* (*Thermosynechococcus elongatus*)  
gch2\_bacam: SWISS-PROT: gch2\_bacam (*Bacillus amyloliquefaciens*)  
gch2\_bacsu: SWISS-PROT: gch2\_bacsu (*Bacillus subtilis*)  
gch2\_bacce: TrEMBL: AAP11030 (*Bacillus cereus* ATCC 14579)  
gch2\_bacha: TrEMBL: Q9KCL5 (*Bacillus halodurans*)
- 20 gch2\_clope: TrEMBL: Q8XMX0 (*Clostridium perfringens*)  
gch2\_clote: TrEMBL: Q897Q8 (*Clostridium tetani*)  
gch2\_chlte: TrEMBL: Q8KC35 (*Chlorobium tepidum*)  
gch2\_aquae: SWISS-PROT: gch2\_aquae (*Aquifex aeolicus*)  
gch2\_lepin: TrEMBL: Q8F701 (*Leptospira interrogans*)
- 25 gch2\_deira: TrEMBL: Q9RXZ9 (*Deinococcus radiodurans*)  
gch2\_bacth: TrEMBL: Q8A528 (*Bacteroides thetaiotaomicron*)  
gch2\_caucr: TrEMBL: Q9A9S5 (*Caulobacter crescentus*)  
gch2\_coxbu: TrEMBL: AA090191 (*Coxiella burnetii* RSA 493)  
gch2\_rhiet: TrEMBL: Q8KL38 (*Rhizobium etli*)
- 30 gch2\_lacpl: TrEMBL: Q88X17 (*Lactobacillus plantarum*)  
gch2\_psegl: TrEMBL: Q8RS38 (*Pseudomonas glumae*)  
gch2\_strav: TrEMBL: BAC71833 (*Streptomyces avermitilis*)  
gch2\_phopo: SWISS-PROT: gch2\_phopo (*Photobacterium phosphoreum*)  
gch2\_azobr: SWISS-PROT: gch2\_azobr (*Azospirillum brasilense*)
- 35 gch2\_agrtu: TrEMBL: Q8UHC9 (*Agrobacterium tumefaciens* (cepa C58 / ATCC 33970))  
gch2\_rhime: TrEMBL: Q92RH2 (*Rhizobium meliloti* (*Sinorhizobium meliloti*))  
gch2\_brume: TrEMBL: Q8YFL5 (*Brucella melitensis*)

## ES 2 435 990 T3

- gch2\_brusu: TrEMBL: Q8G298 (*Brucella suis*)
- gch2\_rhilo: TrEMBL: Q985Z3 (*Rhizobium loti* (*Mesorhizobium loti*))
- gch2\_braja: TrEMBL: Q89RZ7 (*Bradyrhizobium japonicum*)
- gch2\_niteu: TrEMBL: CAD86468 (*Nitrosomonas europaea* ATCC 19718)
- 5 gch2\_ralso: TrEMBL: Q8Y1H7 (*Ralstonia solanacearum* (*Pseudomonas solanacearum*))
- gch2\_neime: TrEMBL: Q9JZ77 (*Neisseria meningitidis* (serogrupo B, segunda enzima encontrada))
- gch2\_xanax: TrEMBL: Q8PPD7 (*Xanthomonas axonopodis* (pv. citri))
- gch2\_xanca: TrEMBL: Q8PCM8 (*Xanthomonas campestris* (pv. campestris))
- gch2\_vibpa: TrEMBL: Q87RU5 (*Vibrio parahaemolyticus*)
- 10 gch2\_vibvu: TrEMBL: Q8DF98 (*Vibrio vulnificus*)
- gch2\_vibch: TrEMBL: Q9KPU3 (*Vibrio cholerae*)
- gch2\_vibfi: TrEMBL: Q8G9G5 (*Vibrio fischeri*)
- gch2\_sheon: TrEMBL: Q8EBP2 (*Shewanella oneidensis*)
- gch2\_phoph: TrEMBL: Q8G9H7 (*Photobacterium phosphoreum*)
- 15 ribb\_phole: TrEMBL: Q93E93 (*Photobacterium leiognathi*)
- gch2\_psepu: TrEMBL: Q88GB1 (*Pseudomonas putida* (cepa KT2440, segunda enzima encontrada))
- gch2\_psesy: TrEMBL: Q882G0 (*Pseudomonas syringae* (pv. Tomate, segunda enzima encontrada))
- gch2\_pseae: TrEMBL: Q9HWX4 (*Pseudomonas aeruginosa*)
- ribb\_dehmu: SWISS-PROT: ribb\_dehmu (*Dehalospirillum multivorans*)
- 20 gch2\_xylfa: TrEMBL: Q87D69 (*Xylella fastidiosa* (cepa Temecula / ATCC 700964))







```

gch2_caucr      ~~~~MSR LSEALTRLKA GGMIVLVVDE DRENEGDLVM AAEFADAAA AFMAKRASGL
gch2_coxbu     ~~~~MQKPFN IETALVALRQ GKMIILVDEE SRENEGDLII AABHATPEHI NFMVLYGRGL
gch2_rhiet     ~~~~MSTESS MPRVAPSAHN KEEDSMSFSS IDEALIAIEA GEMVIVVDEE NRENEGDLIV AAEKITAHEI AFMMKYARGL
gch2_lacp1     ~~~~MDTMKK VAAALAAALKR GELIIVADDE TREAREGDMVG LAAKATATV NRMITTSARGL
gch2_pseg1     ~~~~
gch2_strav     ~~~~
gch2_phopo     ~~~~MYADAPSD SAPPEAVLPM DEAMRAVDR
gch2_azobr     ~~~~MAYDQKR VVDAIRAFEA GEIVVVTDDD DRENEGDLII SAVHCTPEKM ALIVRHTSGI
gch2_agrtu     ~~~~MSYDQKR VWEAIRAFEA GEIVVVTDDG GRENEGDLIV AAVHCTPEKM AFIVRHTSGI
gch2_brume     ~~~~MSYNQKQ VVDALCAFER GEIVVVMDDD GRENEGDLIV AAVHCTPEKM AFIVRHTSGI
gch2_brusu     ~~~~MSYNQKQ VVDAIRAFER GEIVVVMDDD GRENEGDLIV AAVHCTPEKM AFIVRHTSGI
gch2_rhilo     ~~~~MPYDQKK IVEAIRAFER GEIVVVMDDD GRENEGDLIV AAVHCTPEKM AFIVRHTSGI
gch2_braja     ~~~~MPDT VQEVILQAFAR GELVVVTDDDE DRENEGDLIV AASLCTAEKM AFIRHTSGI
gch2_niteu     ~~~~MT. ISS TEEILADFRN GKVVILIDEE DRENEGDLVL AADFVTPEAI NFMARYGRGL
gch2_ralso     ~~~~QLPAMS. IAT VEEITABIRA GRMVILVDEE DRENEGDLIL AADFVTPEAI NFMARFGRGL
gch2_neime     ~~~~MSHISP IPEILADIKA GKVVIIIDAE DRENEGDLIM AAOQVVTPEAI NFMILKHARGL
gch2_xanax     ~~~~MNFAP VPELIEELRA GRMVIVVDEE DRENEGDLIM AAEVLKPSDI NFMVTHARGL
gch2_xanca     ~~~~MNFAP VPELIEELRA GRMVIVVDEE DRENEGDLIM AAEVLKPSDI NFMVTHARGL
gch2_vibpa     ~~~~MPIST PQEIIIDIRA GKVVILMDDDE DRENEGDLIM AAHEITPEAI NFMATYGRGL
gch2_vibvu     ~~~~MPIST PQEIIEDIRL GKVVILMDDDE DRENEGDLIM AAHEITPEAI NFMATYGRGL
gch2_vibch     ~~~~MPIST PQEIIEDIRQ GKVVILMDDDE DRENEGDLIM AAHEITPEAI NFMATYGRGL
gch2_vibfi     ~~~~MAISS AKDIIEDIRL GKVVILMDDDE DRENEGDLII AAEKITPEAI NFMATHGRGL
gch2_sheon     ~~~~MALHS IEEIIEDIRQ GKVVILMDDDE DRENEGDLII AAEKITPEAI NFMATYGRGL
gch2_phoph     ~~~~MALSS AKEIIDDIRQ GKVVILMDDDE DRENEGDLII ASEKITPETI NFMAMYGRGL
ribb_phole     ~~~~MTLSS AQEIIINDIRL GKVVILMDDDE NRENEGDLII AADMITPEAI NFMATYGRGL
gch2_psepu     ~~~~MAFNT IEELLEEDYRQ GKVVILVDDDE DRENEGDLII AAFRCDAQAI NFMARFARGL
gch2_psesy     ~~~~MAFDR IEDIIEDYRQ GKVVILVDDDE DRENEGDLII AADCCTAQAI SFMAREARGL
gch2_pseae     ~~~~MALNT IDELIEDIRQ GKVVILMDDDE DRENEGDLIM AAEKVRTEDI NFMVKHARGL
ribb_dehmu     ~~~~MNAIIS DQKTFQAIIR VNOAIEDIRQ GKVVVMVDDDE DRENEGDLIV AASFSTPQKV NFMASHAKGL
gch2_xylfa     ~~~~
gch2_cangu     ~~~~MASKYIVHP QPERRHGET HEFTMPLLSP TLTPSHIPSQ
gch2_ashgo     ~~~~
gch2_yeast     ~~~~
gch2_neucr     ~~~~PPQTPGTATP PPSLLSPSFT STSITSLSL E... PRGIP VD... SS
gch2_schpo     ~~~~KKEGDQVHHD KTLQLDLSLP ENVASEKRLP IDDVQCVNSP
gch2_arcfu     ~~~~MD ISAEERAKFI RELVSGNA.. GDLKYPGRIF VEETKEMGVL
gch2_strco     ~~~~

```

91

180



gch2\_anasp ICLAMTGDRL DELDLEPLMVT NIT.....DP NQTAFTVSID ASPHLGVSTG ISAE DRARTI QITLNPATKP SDLRRPGHIF PLRAKAGGVL  
gch2\_syny3 ICLAMTGDRL DELDLEPLMVS KNT.....DS NQTAFTVSID AAPHLGVTTG ISAE DRARTI QIAINPVTRP EDLSRPGHIF PLRAKTGGVL  
gch2\_synel ICLAMEGDRL DELDLEPLMVT TNT.....DS NQTAFTVSVD AGARWGVTTG ISAE DRARTI QALIDPSTQP QDLRRPGHVF PLRSRPGGVL  
gch2\_bacam ICTPLSEEDIA GRDLHPMVD HNT.....DS HETAFTVSID .HKL..TKTG ISAQERSFTI QALLDEESVP GDFQRPGHIF PLIAKKGGL  
gch2\_bacsu ICTPLSEIEIA DRDLHPMVE HNT.....DS HHTAFTVSID .HRE..TKTG ISAQERSFTV QALLDSKSVF SDFQRPGHIF PLIAKKGGL  
gch2\_bacce VCVPIITEGYA ERLQLEPMVS HNT.....DS HHTAFTVSID .HVS..TTTG ISAHERATTI QELLPASKG ADFNRPGHIF PLIAKEGGVL  
gch2\_bacha VCTPITEERA KELDLVPMVD HNT.....DP HGTAFTVSID .HQM..TTTG ISAEHERAMTI QALIDKTKK HHFKRPGHIF PLIAKNGGVL  
gch2\_clope ICAPVSEESIA ENLGLNPMVS HNT.....DN HETAFTVSID .YKD..TTTG ISAEFRARTI KELINEESKA EYFRRPGHVF PLIAKNGGVL  
gch2\_clote VCMPIIGERL KELNLNQMVD INT.....DT NGTAFTVSID .FID..TTTG ISAEYERARTI SKVLDSYKGF EDFKRPGHVF PLEAKEGGVL  
gch2\_chlte LCVAIPMERA RELQLEPMVQ RNT.....SQ HETNFTVSID .AIAEGVTTG ISAYDRYMTL KMLADPSSTA DDFSRPGHIF PLRAMDGGVL  
gch2\_aquae ICLSLTPERC EQDLHPMTP MNT.....DP KGTYFCVSID AHPKHGTTTG ISAYDRALTI KLAISPDAKP SDFVRRPGHVF PLKARPGGVL  
gch2\_lepin ICIPMEVERL KKLGLNRMVD DYSLG...DK HGTAFTVSVD A..KHGTSTG ISAQDRALTI QVLLDDKTYS ADLMRPGHLF PLQAVPGGVL  
gch2\_deira ICVTLTPERA RRLDLTPMVG AGAYGQGTDP NGTAFTVSVD ...HSSNSTG ISAYDRAATI AALLEDTSQP TDFRRPGHIF PLVARPGGVL  
gch2\_bacth LCAPITVSRC KELDLPHQVS DNTSVL.... .GTPFTVTID KL..EGCSTG VSASDRAATI QALADPASTP ATFGRPGHIN PLYAQEKGVL  
gch2\_caucr ICLTLEAETI DRLCLAPMVS DN.....RFS RQTAFTVSIE AA..TGVDTG ISAFDRARTI AAAIAPDAQA ADLVSFGHVF PLRARPGGVL  
gch2\_coxbu ICLPMCSSDF ERLNIPMNTS RN.....CSR YETPPFVSIE AA..YDVTTG ISAHDRARTI QIAIDEKSTP EDIIMPGHMF PLKANDAGVL  
gch2\_rhiet ICVPLPAERL DKLEIFLMVT RN.....SDS LQTAFTVSVD CK..HGTTTG ISAE DRAATV KSLIDPQIRP EDLSRPGHIF PLRANRLGVL  
gch2\_lacpl LCVPMAPSIA TRHLTPMNTS EH.....DA FCTAFTVSID ...HHTTSTG ISAE DRARTI QALANPNSQA ADFYHPGHVF PLIARTGGLL  
gch2\_pseg1  
gch2\_strav  
gch2\_phopo  
gch2\_azobr ATAAARRGEA VAIEFADGSV GAAVSVESVA IDAVQRLVQL TGAAPVLAVT RRRATVLIKIM GEGTGVVALS LPRCLTADEA HALADPEHRP  
gch2\_agrtu VCAPMPREEA KRLNLNAMVA ENDSA..... HTTAFTVSVD FK..HGTTTG ISADDRTLTV RNLANPNVGA SDFTRPGHIF PLISREGGVL  
gch2\_rhime VCTPMPREEA KRLNLNAMVA ENDSA..... HTTAFTVSVD FK..HGTTTG ISADDRTLTV RNLANPNVGP TDFVRRPGHIF PLVAREGGVL  
gch2\_brume VCTPMTREEA KRLNLTPMVA ENESA..... HTTAFTVTVD YR..HGTTTG ISAE DRTLTV RNLANPNAGA SDFVRRPGHIF PLVAREGGVL  
gch2\_brusu VCTPMTREEA KRLNLTPMVA ENESA..... HTTAFTVTVD YR..HGTTTG ISAE DRTLTV RNLANPNAGA SDFVRRPGHIF PLVAREGGVL  
gch2\_rhilo VCTPMPREEA KRLNLQPMVA DNDSA..... HTTAFTVSVD FK..HGTTTG ISADDRTLTV RNLANGNVA SDFVRRPGHIF PLIAREGGVL  
gch2\_braja VCAPVTTEA RRLRLDPMVA HNDSA..... HTTAFTVSID YK..PDGGTG ISAEERASCC RALSNPNVGA NDFARPGHIF PLIAKDKGVL  
gch2\_niteu ICLTLTEERC HQKLEPMVA DNHSP..... LGTNFTVSIE AA..TGVTTG ISAE DRACTV QAAVKADARP EDLVQPGHIF PLKAQKGGVL  
gch2\_ralso ICLTLTAERC RQLDLPLMVT RNGTP..... HGTNFTVSIE AA..EGVTTG ISAE DRARTV QVAVARSASP DDLVQPGHIF PLMAQPGGVL  
gch2\_neime VCLPMDGEMV EKLGLPMTQ KNGAQ..... YGTNFTVSIE AA..HGITTG ISAE DRALTI QTAVSPTAKP EDIVQPGHIF PLRAQKGGVL  
gch2\_xanax VCLSLTRERC AOLGLAPMVR NNTAQ..... FQTNFTVSIE AA..EGVTTG ISAYDRARTV RTAVRPDARP QDLQPGHIF PLISQPGGVL  
gch2\_xanca VCLSLTRERC TOLGLAPMVR NNTAQ..... FHTNFTVSIE AA..EGVTTG ISAYDRARTV RTAVRPDAKP HDLSQPGHIF PLIAQPGGVL  
gch2\_vibpa ICLTMTKARC ENLGLPPMVQ DNNAQ..... YTNFTVSIE AA..EGVTTG ISAE DRARTV QAAVAPNAKA ADLVQPGHIF PLAAQDGGVL  
gch2\_vibvu ICLTMTKARC QRLGLPPMVQ DNNAQ..... YTNFTVSIE AA..EGVTTG ISAE DRARTV QAAVAKEAKA ADLVQPGHIF PLAAQDGGVL  
gch2\_vibch ICLTLTKERC RRLGLNPMVQ DNNAQ..... YTNFTVSIE AA..EGVTTG ISAE DRARTV QAAVAKEAKA ADLVQPGHIF PLAAQDGGVL  
gch2\_vibfi ICLTMTKDRC EKLGLPPMVQ DNNAQ..... FTTNFTVSIE AA..EGVTTG ISAE DRARTV EAAVAKNAVA ADLVQPGHIF PLAAQEGGVL  
gch2\_sheon ICQFTMKARC QQLNPLMVT NNNNAQ..... FSTNFTVSIE AA..EGVTTG ISAE DRARTV KAAVAKDAKA SDLVQPGHIF PLMAQDGGVL  
gch2\_phoph ICLTLKARC QKGLPLMVQ DNTEQ..... FGTPFTISIE AA..TGVTTG ISAE DRARTV QAAVAADATA ANIVMPGHIF PLMAQEGGVL  
gch2\_phole ICLTLTKDRC QTLNPLMVQ NNHDK..... FSTAFVSIE AA..ADVTG ISAYDRARTV QAAIGPNASA DDIVMPGHIF PLMAQDGGVL

gch2\_psepu ICLTLTDDHC QRLGLEQMPV ANGSV . . . . . FSTAFVTSIE AA . . TGVTTG ISAADRARTV AAAMAPNARP EDLVQPGHIF PLRAREGGVL  
 gch2\_psesy ICLTLTDEHC KSLGLEQMPV SNGSV . . . . . FATAFVTSIE AT . . TGVTTG ISAADRARTV QAAVNPQAVP DDLVQPGHIF PLRARDGGVL  
 gch2\_pseae VCMPTMTRERC ERLGLPLMVQ RNSGG . . . . . FGTFVTSIE AA . . EGVTG ISAADRARTV QAAAANKAVA ADIVSPGHIF PLMAQPGGTL  
 ribb\_dehmu ICVAISKKIA NRLQLEPMVK KNDSS . . . . . YETAFTITVD AR . . T . AAFG ISAGERDWTI KILADGGSHE SELVVRPGHIF PLIAKEGGAL  
 gch2\_xy1fa PFGSTATIRC ERAAAELRTG RPVLLIDAHA RRHA . VMADD SMTAQSFATF ANAVGNTHYL FLTPVRSNVL GLEAPQGARI PLATLSYDLSL

181 (196) 270

gch2\_cangu TPQIPPEVPA EVRDLPLPE TLPVVKCMAR ARIPTQGPE IFLHLYENNV DNKEHLAIVF GEDVRSKTLY QKRPNETQQD RMTRGAYVGR  
 gch2\_ashgo ~~~~~~MTEY TVPEVRCVAR ARIPIVQGTD VFLHLYFNISI DSKEHLAIVF GENIRSLSLF RYRKDDTQQA RMVRGAYVQ  
 gch2\_yeast QDSSKYEVS GTGDGRNGDG GLPLVQCVAR ARIPTQGPD IFLHLYSNR DNKEHLAIVF GEDIRSLSLF RRRQCETQQD RMIRGAYIGK  
 gch2\_neucr IPTGGCGTK . KPK . LL . . E SLPHVECIVR ARIPTNGAE MFLHLYTNDV DNKEHLAIVF GNNIRSESLD VVBEGETEMD RLVRGAYTGR  
 gch2\_schpo MLSPALDIEP KPN . VVGTAE RLPKVRQVVR ARIPIVKGTE IFLHLYKNDP DNKEHLAIVF CQQIRSDSLD AVQBEGETEMD RLVRGAYVGT  
 gch2\_arcfu ERPGIAEACV DLARMAGFAP SAVYAPLM . . TAEGGVAGE DYALKFAKE . HSMPEV . . . . . RIKDMIEF RIK . SEKIVE RVIEAUTLPTK  
 gch2\_strco ~~~~~~M . . . . . EKIGVLGK KTT . QRTDVE RIVVTPLPT .  
 gch2\_helpj ~~~~~~M . KRLE VSNQAKLPTQ  
 gch2\_helpy ~~~~~~M . KRLE VSNQAKLPTQ  
 gch2\_pyrfu KRKGHTEASL ELVEMAGFKR YALLIEIL . . . . . DDKGYSHNL DYAFHFAKE . HDLPLI . . . . . RIREIWKE FVRRK . QLIH VYSKAKLPTK  
 gch2\_thema RRKGHTEASL EISELAGFSR HAVIVEIL . . . . . DEKGNSHNL DYVLKLSK . FSLPVL . . . . . EMDDVWRE FVRRK . LLMK KKAETLPTD  
 gch2\_chlmu KRPGHTEASM DLMLAGMYP . CGIFAEALVN A . . DHSMMRQ QQILDFAEQ . HGFTVI . . . . . TVDDDLITY RMTF . DSLVE HVSSARIPTK  
 gch2\_chltr ORPGHTEAAM DLMLAGMQP . CGIFAEALVN P . . DHSMMRQ QOVLFAEQ . HDLTVI . . . . . TVDDDLITY RYTY . DSLVT KISSARIPTK  
 gch2\_chlca KRAGHTEAAL DLMLANMQP . CGVLSELVN E . . DHTMMRL AQIIEFAKE . HEISVI . . . . . SIDSLLIAY RMLS . ERLVI FVSSARLPTK  
 gch2\_chlbn KRAGHTESTV DLMELAGLQP . CGVLAELVN E . . DYSMMRL PQILEFAK . HNIAMI . . . . . PVTSLIIAH RMLS . DRJVS KISSARLPTK  
 gch2\_arath gch2\_lyces KRAGHTEASV DLAVLAGLDP . VGVICEVVD D . . DGSMARL PMLRQFAKE . HNLKVI . . . . . SVADLIRY RRRK . DKLVE HASSARIPTM  
 gch2\_orysa KRAGHTEASV DLVALAGLRP . VSVLSTVIN PV . DGSMAGM PVLKQWALE . HDIPIV . . . . . SVADLIRY RRRK . EKLVE LIAVSRILPTK  
 gch2\_alceu ~~~~~~M . . . . . MN . DTEIE LIDSARLPTR  
 gch2\_neima ~~~~~~M . SELLN YVASCRLPTE  
 gch2\_neimb ~~~~~~M . SEMLN HVASCRLPTE  
 gch2\_psepk ~~~~~~MPVV FVAASKLPTP  
 gch2\_psesm ~~~~~~MPVV FVAASKLPTP  
 gch2\_actac ~~~~~~MTKIE LVAQANLPTP  
 gch2\_haein ~~~~~~MAKIQ LVAQANLPTP  
 gch2\_pasmu ~~~~~~MAKIQ RVTEANLPTP  
 gch2\_ec06 ~~~~~~NKPIONQH YLEN . FMQLK RVAEAKLPTP  
 gch2\_ecoli ~~~~~~MQLK RVAEAKLPTP  
 gch2\_salty ~~~~~~MQLK RVAEAKLPTP  
 gch2\_yerpe ~~~~~~MQLK RVAEAKLPTP  
 gch2\_buca1 ~~~~~~MQL . IEKAMLPTP  
 gch2\_buca2 ~~~~~~MQLI QIEKALPTP  
 gch2\_wigbr ~~~~~~MQLK RITEANLPTP

gch2_bucbp	RRPGHTEAAV	DLARLAGLQP	AGAICEIVS	QKDEGSMWQOT	DELRFVADE	.HDLALI...	.TIADLIAW	RRKH.EKHIE	RVAEARLPTK	~~~~~MKLT	KISEAKLPTS
gch2_mycte	RRPGHTEAAV	DLARMAQLQP	AGAICEIVS	QKDEGSMWHT	DELRFVADE	.HGLALI...	.TIADLIEW	RRKH.EKHIE	RVAEARLPTK		
gch2_coref	VRAGHTEAAV	DLARAAGLRP	AGVICEVVS	EEDPTGMARL	PELRRFCFT.	.HGLKLI...	.SIEQLIAW	RRQN.EILVE	RVVETKLPTE		
gch2_corg1	VRAGHTEAAV	DLARAAGLRP	AGVICEVVS	EEDPTGMARV	PELRRFCDE.	.HDLKLI...	.SIEQLIEW	RRKN.EILVE	RQVETVLPTD		
gch2_coram	ERDGHTEAAI	DLARLAGLRP	AGVICEIVS	EEDPTMARS	EELRRFSDE.	.HDLKMI...	.SIEQLIEW	RRHN.ETQVR	RTVETQLPTD		
gch2_staau	ARNGHTEAAV	DLAKLTGAKP	AGVICEIMN	..DDGTMAKG	ODLQKFKBK.	.HOLKMI...	.TIDDLIEW	RKKL.EPEIE	FKAKVKMPTD		
gch2_staep	TRNGHTEAAV	DLARLTGAQP	AGVICEIMN	..DDGTMAKG	EDLQSFKER.	.HHLKMI...	.TIKSLVAF	RRKV.ELNVN	LKAKVKMPTD		
gch2_actp1	VRNGHTEATV	DLARLAGLK.	HAGLCEIMA	..DDGTMTM	PDLQKFAVE.	.HNMFFI...	.TIQQLQEY	RRKH.DSLVK	QISVVVKMPTK		
gch2_lacla	VRNGHTEATV	DLRLLAGLK.	EVGVCVEIMD	..EDGTMHK	EKLKKAQK.	.WGLKYI...	.TIKAIQEY	LKQN.DLQVE	QVTRAKLPTK		
gch2_stcag	ARNGHTEATV	DLRLLAGLK.	ECGLCCEIMA	..EDGSMRK	DELLAFAQK.	.HDLAJA...	.TIKOLQDY	RROE.EGGV	REIEIQLPTQ		
gch2_stcpn	ERNGHTEATV	DLKLKAGLK.	ECGLCCEIMN	..HDGKMMRT	DDLIQFSKK.	.HNIPLI...	.TIKELQEY	RKVY.DQLVE	RVSTVNMPTK		
gch2_cloac	DRNGHTEATV	DLVRLAGFYP	SVGVCCEIMN	..EDGTMARL	PKLMEFSRK.	.NDIKII...	.TVEKLIQY	RKKN.EALIK	REAEAKLPTK		
gch2_fusnu	EREGHTEATV	DLCKICGLTP	VSVICEILK	..DDGTWARM	DDLEIFAKE.	.HNLKII...	.TIADLIKY	RKKT.EQLMK	IDVVANMPTD		
gch2_anasp	KRAGHTEAAV	DL SRLAGLYP	AGVICEIQN	..PDGSMARL	QOLIEYAKH.	.HNLKII...	.SIADLISY	RLQH.DRLVY	REVVTKLPSQ		
gch2_syny3	KRAGHTEAAV	DL SRLAGLYP	AGVICEIQN	..ADGSMARL	PELVEYARK.	.YDLKLI...	.SIADLISY	RLQH.DRFVQ	RETICEFPSQ		
gch2_synel	KRAGHTEAAV	DLTRLAGLYP	AGVICEIQN	..PDGTMARL	POLMEYART.	.HOLKII...	.SIADLISY	RLQH.ERFVR	REAVAKLPTK		
gch2_bacam	KRAGHTEAAV	DLAKACGSPG	AGVICEIMN	..EDGTMARV	PEISEIAKS.	.HOLKMI...	.TIKDLIQY	RYNI.TTLVN	REVDITLPTD		
gch2_bacsu	KRAGHTEAAV	DLAEACGSPG	AGVICEIMN	..EDGTMARV	PELIEIAKK.	.HOLKMI...	.TIKDLIQY	RYNI.TTLVE	REVDITLPTD		
gch2_bacce	RRAGHTEAAV	DLAKLCSAEP	AGVICEIIN	..EDGTMARV	PDLIECAKQ.	.FDIKMI...	.TIEDLIAY	RRHH.ETLVT	REAEITLPTD		
gch2_bacha	RRAGHTEAAV	DLAKLCSAEP	AGVICEIIN	..EDGSMARV	PDLRKLADQ.	.FELKMI...	.TIKDLIAY	RHRK.DKLVK	REVDISLPTD		
gch2_clope	KRDGHTEAAI	DLAKLAGLKE	VGVICEIMK	..DDGTMART	DELLEFAKI.	.HNMKII...	.TIEDLIAY	RTEF.DNFVY	RVVETNMPTK		
gch2_clote	KRAGHTEASV	DLARLAGFYP	AGVICEIVG	..EDGKMARL	POLMEYSKE.	.HNLKII...	.NIADLIAY	RPKK.ETLVK	RVVEAKMPTK		
gch2_chlte	RRVGHTEAAV	DLCLRAGCQP	AGLICEIILH	..DDGSMARV	PELLKLEBK.	.LGMKLI...	.TIKDLVAY	RMQR.SKLVH	RAVESKLPTA		
gch2_aquae	ERAGHTEASV	DLARLAGLYP	AGVICEIMK	..DDGTMARV	PDLMEFAKK.	.HNLKII...	.TIADLIKY	RLRR.ETLVE	KVASAHLPTP		
gch2_lepin	RRAGHTEAAV	DL SKLAGLYP	AGVICEIMN	..DDGTMARL	PDLEKFAEK.	.HGLNIY...	.TIEDLIRY	RRAK.ENLIR	LEVESKLPTK		
gch2_deira	RRAGHTEAGC	DLARLAGFAP	VGVICEIMG	..DDGEMSR	PDLLEAFGEK.	.HGLKVG...	.SIEALIAY	RMEH.DPFME	PVAEAEELPTR		
gch2_bacth	RRAGHTEATI	DMARLAGLYP	AGALMEIMS	..DDGTMARL	PELRAMADE.	.HGLKLI...	.SIHDLIAY	RLKQ.ESIVE	KGVEVNMPTK		
gch2_caucr	VRNGHTEASV	DLMRMAGLRP	AGVICEIMK	..DDGSMARR	PDLEVFCRE.	.HDLPLL...	.TIAELVAH	RKAT.ELLVD	EVAARLPSD		
gch2_coxbu	ARAGHTEGSV	DLRLAGLKP	AAVLCIILN	..KDGSMARL	PDLQLLAKQ.	.FRLPLV...	.SIQDLITY	RINH.ETLID	EVSESLPLK		
gch2_rhiet	GRPGHTEAAV	DLARLAGLAP	AGVICEIAN	..DDGTMARL	PDLEKFAIE.	.HGLHIV...	.TIDALIEY	CRSR.DTRVK	RYAQSFMT.		
gch2_lacp1	ARRGHTEAAV	TLAQLAGVTP	AAVICEVIK	K..NGLMARR	ROLKLALABG.	.LQLPLI...	.TIEDITTY	VIQTGAQLT	TTPSVNLPTK		
gch2_psegl	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~MNQ	SDDHSGVSI	TRIKVPIRPR		
gch2_strav	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	MRDLPSATVR	TRVTVPLRFG		
gch2_phopo	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~MVNVR	ERVSLKVGIN		
gch2_azobr	DGDMPDGLTA	TAMPDGSRET	AAVDIARLAR	LLPAIVAPA	TDHTGSAAB.	.WAAEHD...	.LLJVRARD	IADYRVHVR	TLRRVAEARV		
gch2_agrtu	MRSGHTEAAV	DLCLRLAGLPP	IGVISELVN	..DDGSMVRG	POVLDFADK.	.HGMKHV...	.SVADLIAY	R.QRKETLVE	MESSFTIETP		
gch2_rhime	MRSGHTEAAV	DLCLRLASLPP	IGVICELVN	..DDGTVMRG	PQVEAFAET.	.HGLKQV...	.SVADLIAY	R.QRKETLIE	QGHSEFEMDTP		
gch2_brume	MRSGHTEAAV	DLCKLANLPP	IGVICELVN	..DDGTVMRG	PDVAFAFAE.	.HKLHRV...	.TVADLIAY	R.QRKETLVR	RIGDAPVETC		
gch2_brusu	MRSGHTEAAV	DLCKLANLPP	IGVICELVN	..DDGTVKRG	PDVAFAFAE.	.HKLHRV...	.TVADLIAY	R.QRKETLVR	RIGDAPVETC		

gch2_rhilo	MRSGHTEAAV	DLCKLAGLPP	IGVISELVN	..DDGTVKRG	PQVQAFABE.	.HGLKQV...	..SVADLIAY	R.QRKETLVE	RVACSDIDTL
gch2_braja	LRSGHTEAAV	DLCKLSGLPP	VGVISELMN	..DDGSMVKG	EQVAREFAAQ.	.HKLKHV...	..TIADMIAY	R.QVREKLIIE	RVSTFTVDS
gch2_niteu	ARAGHTEAGC	DLGRLAGLTE	AAVICEILK	..ENGEMARL	PDLEIFABG.	.HTLKIG...	..TIADLIHH	R.NLTESLVS	RFAERPLHTA
gch2_ralso	MRAGHTEAGC	DLASLAGLTP	ASVICEIMK	..DDGTWARL	PDLEVEFABE.	.HGLKIG...	..TIADLIQY	R.SRTESSIVE	RIGERNMETQ
gch2_neime	VRAGHTEAGV	DLAQMGGLIP	ASVICEIIN	..DDGTWARM	PELMKFAFE.	.HKLKIG...	..TIADLIEY	R.SRTESSILLE	DMGNAPVQTP
gch2_xanax	TRAGHTEAAS	DLEMLAGLEP	AGVLEVLN	..PDGSMARR	PQLEVFABE.	.HGLKMG...	..SIADLIAY	R.LANEHTVE	RVDEREITTE
gch2_xanca	TRAGHTEAAS	DLEMLAGLEP	AGVLEVLN	..PDGSMARR	PQLEVFABE.	.HGLKMG...	..SIADLIAY	R.LANEHTVE	RVDEREITTE
gch2_vibpa	TRAGHTEAGC	DLARLAGLEP	ASVIVEILN	..DDGTWARR	PDLEIFABK.	.HGLKLG...	..TIADLIEY	R.NNTEHTIE	RVAECKLPTE
gch2_vibvu	TRAGHTEAGC	DLARLAGFEP	ASVIVEILN	..DDGTWARR	PDLEVFABQ.	.HGLKLG...	..TIADLIEY	R.NNTEHTIE	RVAECKLPTE
gch2_vibch	TRAGHTEAGC	DLARLAGLEP	ASVIVEILN	..DDGTWARR	PDLEVFABK.	.HGLKLG...	..TIADLIEY	R.NNTEHTIE	RVAECKLPTE
gch2_vibfi	TRAGHTEAGC	DLARLAGYEP	ASVIVEILN	..DDGTWARR	PDLEIFABK.	.HGLKLG...	..TIADLIEY	R.NNTEHTIE	RVAECKLPTE
gch2_sheon	TRAGHTEAGC	DLARLAGLEP	SGVIVEILN	..EDGTWARR	PDLEIFSEK.	.HGKIG...	..TIAALIEY	R.NNKETTIV	REAKCKLPTR
gch2_phoph	TRAGHTEAGC	DIARLAGLEP	SSVIVEILN	..EDGTWARR	PQLEIFABK.	.HGLKLG...	..TIADLIEY	R.NNHEHTIE	RVGECKLNTE
ribb_phole	IRAGHTEAGC	DVARLAGLEA	SSVIVEILN	..EDGTWARR	PQLEVFABQ.	.HGLKLG...	..TIADLIEY	R.TQRESHIE	RLSESELCTE
gch2_psepu	TRAGHTEAGC	DLARLAGFSP	ASVIVEILN	..DDGTWARR	PDLEVFABR.	.HGKIG...	..TIADLIHY	R.LSTEQTIVK	RIGERELPTV
gch2_psesy	TRAGHTEAGC	DLARMAGFTP	ASVIVEVMN	..DDGTWARR	PDLELFABK.	.HGIRIG...	..TIADLIHY	R.LSTEKTIV	RIGERELPTV
gch2_pseae	ARAGHTEAAC	DLARMAGFEP	SGVIVEVMN	..DDGSMARR	PELEAFABE.	.HGKIG...	..TIADLIHY	R.LJHERTIVE	RIAEQPLDSE
ribb_dehmu	VRIGHTEGSV	DLCLRAGQGD	SAVICEIMK	..EDGTWARR	PDLDIFCAK.	.HELVNIV...	..YISDIVEY	R.NMNESLIR	VIAESTTQFL
gch2_xylfa	AKLAYLRQPT	RPTAWAPGDI	MDAARAETEITR	LALLLPAIVA	APLTOHTEHA	FADCQTL...	..DLADLDTA	AACASTTEYE	LVTRTPVPLR

  

gch2_cangu	LFPGRTEADY	DSESNRLNLF	DEN.GQLIRD	PSTT.CSGEP	ILARIHSECY	TGETAWSARC	(270)	(276)	(279)	360
gch2_ashgo	LYPGRTEADA	DRRQGLELRF	DE.TGQLAVE	RATW.TREP	TLVRLHSECY	TGETAWSARC	.DCGEQFDEA	GRLM.....	.....	GEAGH
gch2_yeast	LYPGRTVADE	DDRGLALEF	DDSTGELLAS	KATWDAHND	TLVRIHSECY	TGENAWSARC	.DCGEQFDQA	GKLMAAATE.	.GEVVGAGH	
gch2_neucr	LYPGRVTSRE	PGVGETKKEE	QEEKAE...	.....	EKKQV	PLVRIHSECY	.DCGEQLEDE	ARLMS.....	...LPSNTAG	
gch2_schpo	LYPGRKSSYI	DVNSSKTTEE	QKEAR...	.....	KTNFV	PLVRIHSECY	.DCGEQLEDE	ARLIS.....	...EEGN...	
gch2_arcfu	FY.GTFRA.V	GY.KTPL..G	EIVALVKGR	V.....	DEGD	VLVRIHSECL	.DCGDQLENA	LKMDID.....	.....	RBGK
gch2_strco	VY.GKFRA.F	GYFDHER..G	DEQVALVHGD	L.....	GAED	VLTRLHSECL	.ECGAQLASA	LRQVA.....	.....	DAGS
gch2_helpj	F..GEFCI.Q	CFREKGSNGS	KDHLVIFTPN	FP.....	QN	PLVRLHSECL	.DCGGALQMA	LERIS.....	.....	KEG.
gch2_helpy	F..GEFYI.Q	CFREKGSNGS	KDHLVVFTPN	FS.....	QN	PLVRLHSECL	.DCGGALQMA	LERIS.....	.....	KEG.
gch2_pyrfu	Y..GDFEI.Y	SF..KNLDLF	KDHVALVKKP	FK.....	GT	PLVRVHSECL	.DCGSQLENA	LRMIS.....	.....	QEG.
gch2_thema	F..GVFKV.V	SF..ENHLDG	KEHFAIVKEP	LE.....	DP	VAVRIHSECV	.DCGSQLANF	LRYS.....	.....	AHG.
gch2_chlmu	Y..GDFSI.H	VYKSI..IDG	TEHFALVKGD	IREQ.....	ES	VPRVHSECL	.DCGAQLDMA	MRVIA.....	.....	EQGL
gch2_chl1tr	Y..GDFSI.H	VYESI..IDG	TQHVALVKGD	IHEQ.....	EA	VPRVHSECL	.DCGAQLDMA	MRVIA.....	.....	EEGL
gch2_chl1ca	Y..GDFNI.H	VYESL..LDG	IQHIALVKGD	VSGK.....	EN	VLVRVHSECL	.DCGHQLRSA	MEYIG.....	.....	LEGE
gch2_chl1pn	Y..GDFTI.H	VYESL..LEG	MQHIALVKGN	VAGK.....	SN	VLVRVHSECV	.DCGEQLSSA	MSVIA.....	.....	EKGT
gch2_arath	W..GPFTA.Y	CYRSI..LDG	IEHIAMVKEG	I.GD...	GQD	ILVRVHSECL	.DCGNQLALS	MQQIE.....	.....	ATGR
gch2_lyces	W..GPFTA.H	CFKSI..IDG	IEHIAMVKGD	I.GD...	GQD	ILVRVHSECL	.DCGSQALATA	MKQIE.....	.....	AAGR
gch2_orysa	W..GLFRA.Y	CYQSK..LDG	TEHIAVAKGD	I.GD...	GED	VLVRVHSECL	.DCGNQLDLA	MQLID.....	.....	KAGR
gch2_alceu	F..GDFTA.H	AFKGA..DSG	IEHLALSVGD	VTGD...	G..	VLIRMHSECL	.DCGPQLELA	MEKIA.....	.....	ABGR



gch2\_neima W..GVFTM.H GF.EE..AGG QEHALVTGVD CSD...GN.P VLTRIHSECL TGDALFSKKC .DCGPQLEAA MKAVQ.....AEGR  
gch2\_neimb W..GVFTM.H GF.EE..ANG QEHALVTGVD FSD...GN.P VLTRIHSECL TGDALFSRKC .DCGPQLEAA MRAVQ.....TEGR  
gch2\_psepk F..ATFTM.H GFLE..ATG REHVLSLGD IAD...GQ.P VLGRHSECL TGDALFSQRC .DCGSQLEAA LQAI.....REGR  
gch2\_psem F..AEFAM.H GFIES..ATG REHVLSLGD VAD...GA.P VLGRVHSECL TGDALFSQRC .DCGSQLEAA MRAIA.....TEGR  
gch2\_actac F..GLEKI.V GFEP..DTK KEHALVVLGD IRN...SDEP VLARIHSECL TGDALHSLKC .DCGFQLEAA LRQIQ.....QEGL  
gch2\_haein Y..GIFKM.V GFEP..DTK KEHALVMGD ISN...ADEP VLARIHSECL TGDALHSLKC .DCGFQLEAA LKQIQ.....BEGR  
gch2\_pasmu F..GMFRI.V GFEP..DTK KEHALVLGVE VEN...TDEP ILARIHSECL TGDALYSKLC .DCGFQLEAA LRQIS.....QEGR  
gch2\_eco6 W..GDFLM.V GFEE..ATG HDHVALVVDG IS...GHTP VLARVHSECL TGDALFSLRC .DCGFQLEAA LTQIA.....EEGR  
gch2\_ecoli W..GDFLM.V GFEE..ATG HDHVALVVDG IS...GHTP VLARVHSECL TGDALFSLRC .DCGFQLEAA LTQIA.....EEGR  
gch2\_salty L..GDFLM.V GFEE..ATG HDHAALVFGD IS...GKTP VLARVHSECL TGDALFSLRC .DCGFQLEAA LTHIA.....EEGR  
gch2\_yerpe W..GDFLM.V GFEE..ATG HDHLALVFGD IS...GDKP VLSRVHSECL TGDALFSLRC .DCGFQLEAA LTRIA.....EEGR  
gch2\_bucal W..GNFLI.F GFEEK..KNG KNHIALVVDG IK...TNTP TLSRVHSECL TGDALFFSLRC .DCGSQLEMS MKRIS.....QEGS  
gch2\_bucap W..GDFLI.F GFEEK..KNG KNHIALVYVK IK...KNIP ILARIHSECL TGDALFFSLRC .DCGIQLEMA MSRIA.....SEGN  
gch2\_wigr W..TNFLM.I GFEEI..KTK CNHIALHGN MS...KNTS ILTRIHSECL TGDALFSLRC .DCGFQLESA LSHIA.....NEKC  
gch2\_bucbp F..GEFLM.I VFEE..KTD KNHIALVYGD IKD...TNNS VLSRIHSECL TGDALFSLRC .DCGFQLEKSA LMEIV.....KEGS  
gch2\_mytle H..GEFRA.I GYTSI..YDD VKHVALVRGE IAGSSGDGD ILVRVHSECL TGDVFGSRRRC .DCGTQLDAA MAMVA.....REGR  
gch2\_myctu H..GEFRA.I GYTSI..YED VEHALVRGE IAGPNADGDD VLVRVHSECL TGDVFGSRRRC .DCGPQLDAA LAMVA.....REGR  
gch2\_coref I..GEFTA.V GYRSK..VDG TELMAIVEGD PSS.DG.GEN VLVRVHSECL TGDVFGSRRRC .DCGQQLNFS LRMIH.....EAGR  
gch2\_corg1 F..GTFA.V GYRSI..IDG TELVAIVAGD VAS.DG.GEN VLVRVHSECL TGDVFGSRRRC .DCGQQLHES LRLIQ.....EAGR  
gch2\_coram F..GSFTA.L GYKHE..IDG QEHALVLAGG VEELNG.AED VFVRVHSECL TGDVFGSRRRC .DCGQQLHQS MEIIO.....EAGQ  
gch2\_staau F..GTFDM.Y GFKAT..YTD BEIVLTKGA IR..QHEN...VRHLSACL TGDIFHSQRC .DCGAQLSBS MKYIN.....EHG.  
gch2\_actap F..GHFDM.Y GFYTD..YSD BEIIVALVYGD LK..SNPN...VRMHSACL TGDIFHSQRC .DCGAQLSBS MKYID.....EHG.  
gch2\_lacla Y..GEFMA.H SFVEV..ISG KEHALVYKGD LT..D.GE.Q VLARIHSECL TGDALFFSLRC .DCGQQLFAA MTQIE.....QEGR  
gch2\_lacla Y..GMFDI.L GFVNK..ING EHVVALVYGD IG..D.GOA ILCRVHSECL TSDALFSLRC .DCGGQLEEA LKIN.....DEGR  
gch2\_stcag F..GHFTA.Y GYSEV..VAN KEHALVYKGD IS..S.GED VLCRLHSECL TGDVFGSRRRC .DCGEQLANA LQOIE.....AEGR  
gch2\_stcpn Y..GNFKA.I SYIDK..LNG EHLALVMGN IE..D.EAN VLCRVHSECL TGDVFGSRRRC .DCGQQLDAA MKMIV.....ENGS  
gch2\_cloac Y..GEFKI.V GYKSL..IDN KENLAIVKGE LK...AEEE VLVRVHSECL TGDVFGSRRRC .DCGNQLSYA LSAIE.....KQGA  
gch2\_fusnu S..GTFKI.V GFDNL..IDG KEHALVYKGD VA...GKEN VTVRIHSECF TGDILGSLRC .DCGSQLEKTA MRRID.....KLGE  
gch2\_anasp F..GQFDI.Y AYRHT..LDN TEHVAIVKGN PE..NFRDEP VMVRMHSECL TGDALGSLRC .DCRMQLQAA LKMIE.....TAGQ  
gch2\_syny3 F..GEFKL.Y AYRNL..LDQ TEHVAIVKGD PS..EFGQOP VMVRMHSECL TGDALGSLRC .DCRMQLQAA LKMLE.....NHGL  
gch2\_synel F..GEFQI.Y GYRNS..LDQ SEHVAIVKGD PA..TFGEQP VLVRVHSECL TGDALGSLRC .DCRMQLQAA LKMIN.....AAGR  
gch2\_bacam F..GTFRV.Y GYTNE..VDG KEHALVYKGD VP..F.NSGP VLVRVHSECL TGDVFGSRRRC .DCGPQLHAA LRQIA.....EEGR  
gch2\_bacsu F..GTFKV.Y GYTNE..VDG KEHALVYKGD VP..F.NSGP VLVRVHSECL TGDVFGSRRRC .DCGPQLHAA LRQIA.....EEGR  
gch2\_bacce F..GTFHA.I GYSNS..LDT KEHALVYKGD IS..T.GE.P VLVRVHSECL TGDVFGSRRRC .DCGPQLHAA LAQIE.....AEGR  
gch2\_bacha F..GSFRA.I GYTDV..IDG KESVALVYKQ IV..E.GE.P TLVRVHSECL TGDVFGSRRRC .DCGPQLQAA LTQIE.....QGGK  
gch2\_clope Y..GDFKA.I GFIDK..LTG EHVVALVYKGE IK..S.EE.P VLLRVHSECL TGDVFGSRRRC .DCGEQFSA MEKIN.....EEGK  
gch2\_clote W..GEFKI.I GYENK..ING EHVVALVYKGD IE..N.GE.D VLVRMHSECL TGDALGVSRC .DCGYQYBAA MKAIS.....EERR  
gch2\_chlte Y..GEFKL.I AYETI..VDQ QNHMAFVYKGD VG..N.GE.P VLARVHSECL TGDVFGSRRRC .DCGHQLETA LRMI.....KEGR  
gch2\_aquae W..GVFKI.H AYRHK..LTG EQVALTMGE WK..E.DE.P VLVRVHSECL TGDVFRSFRRC .DCRPQLEKA LEMIA.....KEGK  
gch2\_lepin Y..GEFTI.R AYSTL..IDD KIHIALIKGD IK..K.EE.T IMVRVHSECL TGDIFSSNRC .DCGPQLHAA LEMIS.....KEGR

gch2\_deira F..GDFRL.I GFQDT..LSG AEHVALVMGE VT...PEP LLVRVHSECL TGDAFHSLRC .DCGPOLDAA LQAIA.....EEGR  
gch2\_bacth H..GMFRL.I PFRQK..SNG LEHMAIFKGT WN..E..DEP LLVRVHSSCA TGDILGSHRC .DCGEQLHKA MEMIE.....KEGK  
gch2\_caucr HAGQAFSV.R AFRSR..IDG GEHLAMVAGP L.T...GAP .LVRVHSECL TGDALGSRRC .DCGEQLRES LRQIS.....DSGN  
gch2\_coxbu NRGR.FTI.K TFRSR..VDH LEHVALVNQK IEE...DKP CLVRVHSECL TGDVFGSARC .DCGGOLEAA LDEIA.....TOG.  
gch2\_rhiet .RFGDFKA.I AYRDS..LTG TEHLALTLGE LCE...KED VLVRVHSECL TGEAFRSMRC .DCGQOLEMA LRAVQ.....LAGA  
gch2\_lacpl ..DQFTM.T GFEVP..GHT ALPVALQAGD IS...GDEP VLVRVHSECL TGDVFGSARC .DCGQQLHAA LQKIS.....QAQR  
gch2\_pseg1 SSATAFET.E MVTFNGLCDS AEHIALVFGK ....IGEA.. PLVRLHSECL TGDVFGSARC .DCGEQLDES IALFG.....ERG.  
gch2\_strav DGYTV.TG.E LVTFHGLADS QEHVAVLGE ....PCPN.. PLVRLHSECL TGDVFGSARC .DCGPQLREA VERIA.....DRG.  
gch2\_phopo SNAPA...E LLSFNGLSEG KEHIALIFKD ....AYKAFV PLVRMHSECL TGDGFHSTRC .DCGEQLDET IEIMA.....ELG.  
gch2\_azobr PLSGAENT.S IAAFRPIDGG PEHLAIIVGN ....PVAGEP VLARLHSECF TGDLLAGLRC .DCGQQLRGA IABIA.....RHGS  
gch2\_agrtu F..GPAMA.Q TYSL..PMDP MQHMAVIFGD IRD...GVD IPVRLHTEHV ADDVFGDR.Q P.....VRHY MQKIA.....EEGR  
gch2\_rhime Y..GKAKG.H TYSL..PMDT MQHLAVVFGD IRD...GVD IPVRLHLENV GADVFGRG.R Q.....IDEI MKRIA.....AEGR  
gch2\_brume A..GKAHA.Y SYEL..PWEP MQHVAVVFGD IRD...GEE VVRLHREDDV LNDVFGKGS N.....LDAI MQRMG.....QEGR  
gch2\_brusu A..GKAHA.Y SYEL..PWEP MRHVAVVFGD IRD...GEE VVRLHREDDV LNDVFGKGS N.....LDAI MQRMG.....QEGR  
gch2\_rhilo G..GKAQV.F TYTL..PMDT MHHVAVVFGD IRD...GEE VVRLHSEDDV VTDVFGT.SH R.....LDGI MKAMG.....ERKR  
gch2\_braja I..GPLOQ.Y AYRS..PFDS IAHVAVFYNG VGD...GKN VLFREHKPNI VKDIF.TGHK R.....MAAV LEHFE.....RSGR  
gch2\_niteu H..GEFRL.I AYHD..KIAG VTHIALVKGT WNT...GEE VLVRVHEPLS VIDLLEADDH S.HWSIHDA MAVIA.....QAGK  
gch2\_ralzo F..GTFRA.V AFRD..HAAG RAHLALVKGT PTP...ATE TLVRVHEPLS VIDLLECCRT T.HSWSVPAK LRAVQ.....AAPT  
gch2\_neime W..GEFQQ.H VYVD..KLSG ETHALALVKGT PAA...DTE TLVRVHEPFS VMDFIQANPR .HSWSLPPA LEHIQ.....QAES  
gch2\_xanax F..GPFKL.V TYRD..RIAH DLHFALLRGR ADP...QTP TLVRVQVENP LADLLHWQRD .DFGVAATDA LRAIA.....AEGQ  
gch2\_xanca F..GPFHL.V TYRD..RIAH DLHFALVRGT PDP...TTP TLVRVQVENP LADLLHWQRD .DFGVAATDA LRAIA.....AEGQ  
gch2\_vibpa H..GEFTL.V TYKD..TIDS QVHYAMCKGD LAG...EA. PLVRVHLQDV FTDVLRSDRN AERSWTLDKA MKRIG.....EEG.  
gch2\_vibvu F..GEFDL.V TYRD..TIDN QIHFALRKGT MGD...QT. PLVRVHLQDT FTDLLRSDRN AERSWTLDKA MKRIG.....QEG.  
gch2\_vibch Y..GEFEL.V TYRD..IIDK QIHFALRKGD ITS...AP. TLVRVHLQDT FTDLLHSNRA AERSWPLATA MORIG.....QEG.  
gch2\_vibfi F..GEFDL.V TYRD..TIDN QLHVALKKEN TSS...EA. PLVRVHLQDT FTDLLHSDRN AERSWTLPTA MQRIA.....NEG.  
gch2\_sheon F..GEFDM.V TYRD..TIDN QLHFALVKGE VKP...DC. .LVRVHLQNT FNDDLHSEDDV QORSWPLEKA MARIS.....AEG.  
gch2\_phoph Y..GEFDM.I TYRD..KIDD QIHVALCKGD IEA...DAA TLVRVHLQDT FKDILQSGAT Q...WTLPAQ MORIS.....AEN.  
ribb\_phole Y..GVFNL.I TYRD..TIDN QLHYALCKGD INP...DSE TLVRVHVKDT LKDLILHTGAT Q...WSLQAA MORIQ.....ADG.  
gch2\_psepu H..GTFRL.V TYED..RIEG GVMHAMVMGD IRR...EOP TLVRVHVDP LRDILVGAEYA GPANWTLWAA LQKVA.....EEGT  
gch2\_psesy H..GTFRL.V TYED..RIEG GVMHAMVMGD IRR...EDP TLVRVHVDP LRDILVGAEYT GPANWTLWAA LQKVA.....EEGC  
gch2\_pseae L..GHFNL.I TYRD..SVEG DVHLALTIK VCA...EOP TLVRVHMDDP LRDIL...QVN QPGRWSLRAA MTKVA.....EAGS  
ribb\_dehmu G..KDARR.Y DFVD..HNDN H.HIAYAFGN IKN...RSA .VKFHSIMP DNELL...AD TKKYNLSIOA IHYLQ.....KSG.  
gch2\_xylfa .DLGMSEF.I VFR..GGIAQ RDQVAILIG. ...RPDLSSA VPVRVHSSCL TGDLFVGSLLK .DCGDQLRHS LATLK.....ALGG

361 (308) 450 (347)  
gch2\_cangu GCIVYL.RQE GR.GIGLGEK LKAYNLQ.DL GADTVQANLM LRHPADARF SLATAILLDL GLNEIKLLTN NPD.KIAAVE GRNREVKVVE  
gch2\_ashgo GVIVYL.RQE GR.GIGLGEK LKAYNLQ.DL GADTVQANEL LNHPADARDF SLGRAILLDL GIEDIRLLTN NPD.KVQQVH .CPPALRGIE  
gch2\_yeast GVIVYL.RQE GR.GIGLGEK LKAYNLQ.DL GADTVQANLM LKHPVDARDF SLGKAILLDL GIGNVRLTN NPE.KIKQVD .HAPYLKVE  
gch2\_neucr GIIVYL.RQE GR.GIGLGEK LKAYNLQ.DL GSDTVEANLL LRHPADARSY GLATAMRLD GQKEIRLLTN NPD.KIRAVE GPNREIVYTE  
gch2\_schpo GVIIYL.RQE GR.GIGLSEK LKAYNLQ.DL GNDTVQANLL LQHPADGRSY AVATAMLDL GQDIRLLTN NPE.KVNSIE GPNKEVRYVE

gch2\_arcfu GVAIVMRGHE GR. GIGLINK LMAKQLQ.EE GKDTVDANIE LGFPPDMRSY GIAAQILMDL KVKSIIRLLTN NP. LKIEELK KYG..FKIVR  
gch2\_strco GIVVYLRGHE GR. GIGLLAK LRAMALQ.AE GLDTVEANLA LGLPVDARDY GVAARILDDL GVRSVRLMSN NP. RKREALV RHG..IRVAE  
gch2\_helpj GLVIY.LRQE GR. GIGLFNK VNAYALQ.DK GYDTIQANEM IGFKDDERDY SIAGEILEY RIKKMRLLTN NP. LKIAALE KYA...EVTR  
gch2\_helpy GLVIY.LRQE GR. GIGLFNK VNAYALQ.DK GYDTIQANEM IGFKDDERDY SVAGEILEY RIKKMRLLTN NP. KKIAALE KYA...EVTR  
gch2\_pyrfu GVLIIY.LDQE GR. GIGLREK IRAYEFQ.DR GLDTVEANKA LGFAPERTY EVAFOILRAL GIEKIKLLTN NP. SKVKALR ELG..IEIED  
gch2\_thema GILIIY.LRQE GR. GIGLSNK IAAYSLQ.DK GLDTVEANRV LGFSEDERDY APAAQILKAL GIERVLLFTN NQ. RKTVGLK KYG..IEVVE  
gch2\_chlmu GVIIVLRGQE GR. GIGFGHK IQAYALQ.DL GYDTVEANLQ LSFPIDAREY GVAQAQILKDL RLTSVRLITH NP. KKFELQ RLG..IHILD  
gch2\_chltr GVIIVLRGQE GR. GIGFGHK IRAYALQ.DL GYDTVDANLQ LGFPIDAREY GMAAQVLKDL QLTSVRLITH NP. RKFELQ RLG..IHVLD  
gch2\_chlca GVIIVLRGQE GR. GIGLGHK VRAYALQ.DR GYDTVDANLE IGFPVDSREY GIGAQILVDL GLTKIKLITH NP. HKYFGLQ GFG..LEIVD  
gch2\_chlpn GVLVYLRGQE GR. GIGLGHK VRAYALQ.DN GYDTVDANLA MGFPVDSREY GIGAQILVDL KLTIKLITH NP. QKYFGLQ GFG..LSITE  
gch2\_arath GVLVYLRGHE GR. GIGLGHK LRAYNLQ.DA GRDTVEANEE LGLPVSREY GIGAQILRDL GVRTMKLMTN NP. AKVVGLK GYG..LAIVG  
gch2\_orysa GVLVYLRGHD GR. GIGLGHK LRAYNLQ.DA GRDTVEANED LGLPVSREY GVGAQILRDL GVRTMKFITN NP. AQYSWLK GYG..LAISG  
gch2\_alceu GILLVLRGHE GR. GIGLQK LRAYNLQ.DD GHDTVQANVE LGLAVDSREY GIGAQILRDM GVRTMRLMTN NP. AKFVGLK GYG..LAVVG  
gch2\_neima GIIVYL.RQE GR. GIGLAEK IRAYRLQ.DG GLDTVDANLA LGQAIDARNY AFAADLLHQW GVKSVRLMSN NP. LKAAALE KAG..ITVTE  
gch2\_neimb GIIVYL.RQE GR. GIGLINK IRAYHLQ.DQ GMDTVEANVA LGLPVDARDF RLAQSIYEV LRAQSIYEV GIRSVKLLTN NP. EKIQTLLK DAG..INVVE  
gch2\_psepk GVLVYL.RQE GR. GIGLINK IRAYHLQ.DQ GMDTVEANLA LGLPVDARDF RLAQSIYEV LRAQSIYEV GIRSVKLLTN NP. EKIQTLLK DAG..INVVE  
gch2\_psesm GVLVYL.RQE GR. GIGLGNK IRAYELQ.DG GADTVEANER LGFADARDF AICLPMLEHL GVKSLRLMTN NP. RKVKALT DMN..IVVAE  
gch2\_actac GVLIIYH.RQE GR. GIGLINK IRAYSLQ.DQ GMDTVEANLA LGFADARDF GVCADIFDLL GVTQIRLLTN NP. NKIKTMK OAG..INWVA  
gch2\_haein GVLIIYH.RQE GR. GIGLINK IRAYSLQ.DK GMDTVEANLA LGFADARDF EVCADMFELL GVKKVRLLTN NP. EKVETMK KAG..INVVE  
gch2\_pasmu GVLIIYH.RQE GR. GIGLINK IRAYSLQ.DK GMDTVEANLA LGFADARDF EVCADMFELL GVKKVRLLTN NP. NKIDTMK KAG..INVVE  
gch2\_eco6 GILLVYH.RQE GR. NIGLINK IRAYALQ.DQ GYDTVEANHQ LGFADARDF TLCDMFKLL GVNVRLLTN NP. KKVEILT EAG..INIE  
gch2\_ecoli GILLVYH.RQE GR. NIGLINK IRAYALQ.DQ GYDTVEANHQ LGFADARDF TLCDMFKLL GVNVRLLTN NP. KKVEILT EAG..INIVE  
gch2\_salty GILLVYH.RQE GR. NIGLINK IRAYALQ.DQ GYDTVEANHQ LGFADARDF TLCDMFKLL GVNVRLLTN NP. KKVEILT EAG..INIVE  
gch2\_yerpe GVLIIYH.RQE GR. NIGLINK IRAYALQ.DL GADTVEANHQ LGFADARDF TLCDMFKLL GIKAVRLLTN NP. KKVEILT OAG..INIVE  
gch2\_bucai GVLIIYH.RQE GR. NIGLINK IKAYALQ.DR GLDTVQANQK LGFADARDF SLCADIFNLL NIKKIRLLTN NP. FKVQMLS DAG..INIVE  
gch2\_bucap GILIIYH.RQE GR. NIGLINK IKAYSLQ.DK GLDTVEANQK LGFADARDF SLCADIFNLL NIKKIRLLTN NP. FKVDMLI ASG..IEIVE  
gch2\_wigbr GILIIYH.RQE GR. NIGLINK IKAYSLQ.DK GIDTVDANYH LGFADARDF TVCVEILNLL QIKSIRLLTN NP. NKVNVLT SAG..ILITE  
gch2\_bucbp GILIIYH.RQE GR. NIGLSNK IRAYALQ.DI GLDTVEANHH LGFADARDF SVCIDIFNLL NIKKIKLLTN NP. SKVTVLN NAG..IQITE  
gch2\_mycle GVVLYMRGHE GR. GIGLMHK LQAYQLQ.DA GADTVDANIK LGFPPSDARDY GIGAQILVDL GVRSMRLLTN NP. AKRVGLD GYG..LQIEE  
gch2\_myctu GVVLYMRGHE GR. GIGLMHK LQAYQLQ.DA GADTVDANLQ LGLPADARDY GIGAQILVDL GVRSMRLLTN NP. AKRVGLD GYG..LHIEE  
gch2\_coref GVVVYMRGHE GR. GIGLLAK LRAYQLQ.DE GADTVDANLS LGLPADAREF GTGAQILYDL GVRINLMSN NP. TKRMGLE GHG..ISIAG  
gch2\_corgl GVVVYMRGHE GR. GIGLLAK LRAYQLQ.DE GADTVDANLA LGLPADAREF GTGAQILYDL GVRINLMSN NP. AKKVGLE GHG..ISIAS  
gch2\_coram GIIIVLRGHE GR. GIGLLAK LKAYSQ.DS GLDTVDANLE LGLPEAREY SVAGQILRDL GIKSANLLTN NP. HKGEGLR GFG..VEASA  
gch2\_staup GMIIY.LPQE GR. GIGLINK LRAYELI.EQ GYDTVTANLA LGFDELRDY HIAAQILKYF NIEHINLLSN NP. SKFEGLK QYG..IDIAE  
gch2\_actpl GMIIY.LPQE GR. GIGLINK LRAYELI.EK GYDTVTANLA LGFDELRDY HVAEILKYF DISEINLLSN NP. KKFEGLK DYG..IEIVD  
gch2\_lacla GVLIIY.LRQE GR. GIGLINK LRAYELQ.DK GMDTVEANVA LGFKEBEREY YIGAQMFOQL GVKSIIRLLTN NP. AKIEGLK EQG..LNIVA  
gch2\_stcag GVLIIY.LRQE GR. GIGLINK LRAYSLQ.DE GLDTVEANLA LGFEEBEREY SIGAQILKII GVKSLKMLTN NP. QKINDFO KYG..LVVEE  
gch2\_stcpn GVLIIY.LRQE GR. GIGLINK LKAYHLQ.EE GLDTLEANLA LGFEGDERDY GVSQAQLKDL GINSINLLTN NP. DKIQOLE AEG..ICVKN  
gch2\_stcpn GVLIIY.LRQE GR. GIGLINK LKAYHLQ.DQ GMDTLDANLA LGFEGDLRKY HIGAQMCKDL GLQSLHLITN NP. DKVEQLE KYG..ITISS

gch2_cloac	GVVLY.MDQE	GR.GIGLANK	LKAYMLQ.EK	GILDTVEANLK	LGFEADLREY	WAAAEMLKDL	GVKVKVLMTN	NP.EKINELN	KYG..IDVVS
gch2_fusnu	GILLY.LRQE	GR.GIGLANK	LRAYNLQ.EE	GMDTLDANLH	LGFGADMRYD	AVAAQMLKAL	GVKSIKLLTN	NP.LKINGLE	EYG..IPVVK
gch2_anasp	GVVYV.LRQE	GR.GIGLANK	LKAYSLO.DM	GLDTVEANER	LGFPADLRDY	GMGAQMLMDL	GVKKIRLITN	NP.RKIAGVK	GYG..LEVVD
gch2_syny3	GVVYV.LRQE	GR.GIGLVNK	LKAYSLO.DL	GYDTVEANER	LGFPADLRDY	GMGAQMLNDL	GVKQIRLITN	NP.RKIAGLK	GYG..LEIVE
gch2_synel	GVVYV.LRQE	GR.GIGLVNK	LRAYSLO.DL	GMDTVEANEH	LGFPADLRNY	GVGAQILNDL	GVKQIRLITN	NP.RKIAGLK	GYG..LEVVD
gch2_bacam	GVLLY.LRQE	GR.GIGLANK	LKAYRLQ.EQ	GYDTVEANEA	LGFLPDLRNY	GIGAQILRDL	GVQHMKLLTN	NP.RKIAGLE	GYG..LSISD
gch2_bacsu	GVLLY.LRQE	GR.GIGLANK	LKAYKLQ.EQ	GYDTVEANEA	LGFLPDLRNY	GIGAQILRDL	GVRNMKLLTN	NP.RKIAGLE	GYG..LSISE
gch2_bacce	GVLLY.MRQE	GR.GIGLANK	LKAYKLQ.EE	GYDTVEANEK	LGFPADLRDY	GIGAQILKDL	GLQSLRLLTN	NP.RKIAGLQ	GYD..LEWE
gch2_bacha	GILLY.MRQE	GR.GIGLANK	LKAYKLQ.EE	GYDTVEANEA	LGFPADLRDY	GVGAQILRDL	GVSKMRLTN	NP.RKITGLK	GYG..LEWE
gch2_clope	GILLY.MRQE	GR.GIGLANK	LKAYKLQ.DE	GLDTVEANEA	LGFPEDLRDY	GVGAQILKDL	GVREIRLMTN	NP.KKITGLE	GYG..LKIVE
gch2_clope	GVLVY.MRQE	GR.GIGLANK	LKAYNLQ.DK	GMDTVEANIA	LGFPEDLRDY	GIGAQILNDL	GIKKINLMTN	NP.KKITSLS	GYG..IKIVK
gch2_chlte	GVLLY.LMQE	GR.GIGLANK	LKAYNLQ.DE	GFDTVEANEK	LGFKPDLRDY	GIGAQILQDL	GVRKMLMTN	NP.KKIVGLE	GYG..LEIVE
gch2_aquae	GVLVYILGHE	GR.GIGLANK	IKAYELQ.EK	GYDTVEANEK	LGYPDLRDY	GIGAQILRDL	GVRKMLMTN	NP.RKIVALE	GYG..LEWE
gch2_lepin	GVLLY.MRQE	GR.GIGLANK	LKAYNIQ.DK	GYDTVEANEK	LGFPADLRDY	GIGAQILREI	GVGKMKILT	NP.RKIVGLD	GYG..LEWE
gch2_deira	GAVIY.LRQE	GR.GIGLANK	IRAYALQ.DG	GADTVDANLQ	LGFLPADARDF	GIGAQILHL	GARKLRIMTN	NP.RKLHLS	GYG..LEWE
gch2_bacth	GVVYV.LNQE	GR.GIGLMEK	MKAYKLQ.ED	GMDTVDANLC	LGHADERDY	GVGAQILREL	GVHKMKLLTN	NP.VKRVGLE	AYG..LEIVE
gch2_caucr	GALIVLRNHE	GR.GIGLANK	IAAYALQ.DQ	GLDTVQANHA	LGFPADQRDY	GVAQAQILRAL	GVGEVKLLSN	NP.RKAAGLR	ANG..VTVRR
gch2_coxbu	GVLLYLR.OE	GR.GIGLANK	IKAYALQ.EA	GLDTVEANHH	LGFGADLRDY	GLAAQMLRVL	GIREIRLMTN	NP.NKVESLT	RYG..ISVAE
gch2_rhiet	GVVIYMRQE	GR.GIGLANK	IAAYALQ.DH	GRDTLEANRE	LGFPADSREY	NAAADITRDL	GITSVRLTN	NP.TKIEALR	TSG..VIVSS
gch2_lacpl	GVFVY.LRQE	GR.GIGLANK	LAAYALQ.EH	GVDTVDANQ	LGFPQDERRY	DIAALVLRQM	GITSVRLTN	NP.DKVEQLR	AAG..IDWE
gch2_pseg1	GILLY.LRQE	GR.GIGLANK	LDAYRLQISQ	GLDTFANRA	LNFPDDLRF	VAAAQMLKAL	GVGEVSLVTN	NP.DKTAQLT	RHG..IKVRQ
gch2_strav	GVLLY.LRQE	GR.GIGLVNK	LDAYALQ.DQ	GLDTVEANAA	LGFLPEDRDY	TAAAQMLGAL	GIDALDLSN	NP.DKAGQLR	DLG..LTVTD
gch2_phopo	GIILY.LRQE	GRWGLGVNK	IDAYKLQ.SQ	GMDTVEANNH	LGFPDDLREF	SEAAQMLTAL	GIKDIRLVTN	NP.KKIKDLQ	QNG..INIVE
gch2_azobr	GVLLY.LAQE	GR.GIGLVNK	LRAYRIQ.DR	GFDTVDANEI	LGFEADERVY	LPAEMLRQL	GFTAVRLMTN	NP.EKLRQLA	RCG..IEWE
gch2_agrtu	GVIVLREGS	...VGVGQV	AGRRKARD..	..PDEAHAQA	RSRENWLEI	GLGAQILKDL	GISSIKLLT.	TREHHYVGL	GYG..IQISA
gch2_rhime	GVIVLREGS	...VGVGIS	QTARKGKH..	..EREVHSEA	QRESEWLEI	GLGAQILKDL	GITSIRLLS.	SREHHYVGL	GYG..IKIAA
gch2_brume	GILVYLRREGS	...VGVRS	HRDARAREAL	QGEHEAHAEA	VAREEWQRQI	GLGAQILKDL	GVSSIVLLA.	SREHHYVGL	GYG..ICITR
gch2_brusu	GILVYLRREGS	...VGVRS	HRDARAREAL	QGEHEAHAEA	VAREEWQRQI	GLGAQILKDL	GVSSIVLLA.	SREHHYVGL	GYG..ICITR
gch2_rhilo	GVIVYLRREGS	...VGVHQA	ER...KRPL	SGDREDHEEA	RRRENEWREI	GLGAQILKDL	GISSINLIA.	SREHHYVGL	GYG..IHIAK
gch2_braja	GVIVYLRDGA	...AGV...	...PVA	PLPDETSTEA	DRNRQWREI	GVGAQILRDL	GVTSIRHLT.	SSVHDYKGLS	GYG..IEIVA
gch2_niteu	GVVLLH.RK	...NDA	SVLTDRLIRA	RPRQD....	KVPYDLRQH	GIGAQILKDL	GVSKMRLLA.	TP.RRMPSMT	GYG..LEVAG
gch2_ralxo	GVVLLNCGD	...SPE	RFLTQFSALD	APQERP....	RSKPPRIY	GIGAQILKDL	GVGKMRVLA.	SP.LKLPSMA	GYD..LEVTA
gch2_neime	GVVILLHRTE	...DGA	SLL.DRTLPK	GANQAY....	..KW.DSKSY	GIGAQILAGL	NVKKLRVLG.	QP.SSFTGLT	GYG..LEVVG
gch2_xanax	GVMVVLSAPR	...DSE	SLLARLRQ..	...QPEAAAN	AKDVSQWRRN	GAGAQILADL	GLGKLRVLG.	TPRRQV.GLA	GYG..LEWE
gch2_xanca	GVMVVLSAPR	...DSE	SLLARLRQ..	...QPEVAAN	AKDVSQWRRN	GAGAQILADL	GLGKLRVLG.	TPRRQV.GLA	GYG..LEWE
gch2_vibpa	GVLVVLGNEE	...STE	LLIHRVKMFE	AQDKGEAPTL	AKKQTSRRV	GVGSQILADL	GVHDMRLLS.	STNKYHALG	GYG..LNVVE
gch2_vibvu	GVLVVLGNEE	...SPE	LLIHRVKMFE	AQDKKEAPKL	AKKQTSRRV	GVGSQILADL	GVHDMRLLS.	SSNKYHALG	GYG..LNVVE
gch2_vibch	GVLVVLGHEE	...SSE	LLIHRVKMFE	LQDKGEAPAM	AKKQTSRRV	GVGSQILADL	GVKEMRLLS.	SPNKYHALG	GYG..LNVVE
gch2_vibfi	GVLVVLGNEE	...STD	NIVHKVKVLE	RQDSGEAPTM	AKKQTSRRV	GVGSQILADI	GIDMRLLS.	STSKRYHALS	GYG..LNVVE
gch2_sheon	GVLVLLGNOE	...HSS	EILAKVKAFE	AEDQGTTPVS	AKWQTSRRV	GVGSQILASL	GVTKMRLLS.	SP.KRYHLS	GYG..LEVTE

gch2\_phoph GVLVILSKQE . . . . . STD SLINKVRNLA AEKEGRPQVK MSPYNPSRQV GVGSQILSDL GIGKMRLLS. SSTQRYHSLG GFG. . LEVE  
 ribb\_phole GLLVILISQIE . . . . . PSA MILNQINHLD SEHQASLPLP ITP. . OSRQI GLGSQILSEL GLCKIRLLS. SOSQOYRSLS GFD. . LEVE  
 gch2\_psepu GVVVILANHE . . . . . SSQ ALLERVPQLT . . . . . HPVRPYQ RQSKVYSEV GTGAQILQDL GIGKLRHLG. PP.LKYAGLA GYE. . LEVQ  
 gch2\_pseey GVVVILANHE . . . . . SSQ ALLERVPQLT . . . . . OPPROYT RQSKRIYSEV GTGAQILQDL GIGKLRHLG. PP.LKYAGLT GYD. . LEVIE  
 gch2\_pseae GVVLLGHQI . . . . . GGD DLLAHVRETA . . . . . SAPAPAP KA.TTYSITV GAGSAILRDL GVRKMRLLS. AP.MRFNAIS GFD. . LEVE  
 ribb\_dehmu GVLIFMDNGT . . . . . QDMS. . . . . GR.GNGIAAK IRAYGYQH.V GLDITIDADAQ LGFGPDERRY TGAVMMLRAL GITRIQLLSN NPA.KAERLR AAG. . IIVKQ  
 gch2\_xy1fa GVLLYL.DQE

540

451  
 gch2\_cangu RVPWVPLAW. RSKNGIKSKE IEGYLSAKIE RMGHLLEKPL KI . . . . . LPRANTHI . . . . . NSALSSTSL AI . . . . .  
 gch2\_ashgo RVPWVPLSWT QPTQVRSRE LDGYLRAKVE RMGHMLQRPL VLHTS.AAAE LPRANTHI . . . . . NSALSSTSL AI . . . . .  
 gch2\_yeast RVPWVPIHWT NSSEGIDSKE IEGYLRTKIE RMGHLLTEPL KLHTNPQPT TSEAQONRM NSALSSTSL AI . . . . .  
 gch2\_neucr RVAMVPLSW. KGKGFRAPE VEGYLTKKIE KMGHMLDMGA LPQ . . . . . LPDS . . . . .  
 gch2\_schpo RISMIPKSW. RGEKGIESAE VVLYLTKKIE KMGHLLD . . . . . LPDS . . . . .  
 gch2\_arcfu E.PIEVEP . . . . . CEV NLPYLAKAKD KMGHLICFND . . . . . AAMAHVSS . . . . .  
 gch2\_strco QVPLLIPP . . . . . CES NITYLRTKRE RLDHHLPHLD AAMAHVSS . . . . .  
 gch2\_helpj .ESLIVCA . . . . . NEH NQYLEVKKL KMGHLL . . . . .  
 gch2\_helpy .ESLIVCA . . . . . NEH NQYLEVKKL KMGHLL . . . . .  
 gch2\_pyrfu VIPLPGEI . . . . . NRY NFFYLKTKVE KLGHKIPGV. .EV . . . . .  
 gch2\_thema TKRLYGRV . . . . . TPH NRYLSTKMK KLGHLEEIF REVNS . . . . .  
 gch2\_chlmu RIVLPVIV . . . . . SSE NERYLRTKRD RMGHMLNFPV LNESEDEYET VERTSCC . . . . .  
 gch2\_chltr RILPVISI . . . . . STE NEGYLRTKKE RMGHMLDLPV LDDVEEYET VERMSCR . . . . .  
 gch2\_chlca RIALPIHV . . . . . SEE NEHYLRTKKE RMGHWIDLPI AEEVNIR . . . . .  
 gch2\_chlpn RVPLPVRI . . . . . SED NEQYLRTKQE RMGHMLDLPK CNN.RVQ . . . . .  
 gch2\_arath RVPLLSLI . . . . . TKE NKRYLETKRT KMGHMYGLXF KGDVVEKIES ESES . . . . .  
 gch2\_lyces MVPVVTFF . . . . . TNH YDTYLETTRA KMGHVIYGL . . . . . NLIRPATS TSTTNGKPN ENTSTIR . . . . .  
 gch2\_orysa RVPVISPI . . . . . TKE NQRYLETKRT KMGHVIYGL . . . . . NLIRPATS TSTTNGKPN ENTSTIR . . . . .  
 gch2\_alceu QLRIHVPS . . . . . NAE NEKYLATKRT RMGHLLD . . . . . PGNVPEEFLN PDDIAGDQDE DDTHN . . . . .  
 gch2\_neima RIPLHVGE . . . . . NLE NKRYLQTKAD KLGHLMSE . . . . .  
 gch2\_neimb RIPLHVGE . . . . . NLE NKRYLQTKAD KLGHLMSE . . . . .  
 gch2\_psepk RVPLHTGH . . . . . NPH NRYLATKAG KLGHMLGNEH QGEVFOA . . . . .  
 gch2\_pseem RVPLHTGQ . . . . . NPH NRYLATKAG KLGHMLGNEH QGEVFOA . . . . .  
 gch2\_actac RVPLNVGE . . . . . NRY NTAYLDTKAK KMGHFIVHNG .EQHLMCEPY CQEEVPK . . . . .  
 gch2\_haein RVPLNVGE . . . . . NRY NTKYLDTKAK KMGHYIVHNN DEQHLMTCPH CQEEI . . . . .  
 gch2\_pasmu RVALNVGE . . . . . NRY NTEYLDTKAK KMGHFIIHN. QQKYPLECPY CSEEVPOEK . . . . .  
 gch2\_ec06 RVPLIVGR . . . . . NPN NEHYLDTKAE KMGHLLNK . . . . .  
 gch2\_ecol1 RVPLIVGR . . . . . NPN NEHYLDTKAE KMGHLLNK . . . . .  
 gch2\_salty RVPLIVGR . . . . . NPN NEHYLDTKAA KMGHLLSK . . . . .  
 gch2\_yerpe RVPLIVGE . . . . . NPK NEHYLATKAA KMGHLLTK . . . . .  
 gch2\_buca1 RVPLIVKK . . . . . NPK NAYLTKKAE KMGHLLSE . . . . .



gch2\_brume  
 gch2\_brusu  
 gch2\_rhilo  
 gch2\_braja  
 gch2\_niteu  
 gch2\_ralso  
 gch2\_neime  
 gch2\_xanax  
 gch2\_xanca  
 gch2\_vibpa  
 gch2\_vibvu  
 gch2\_vibch  
 gch2\_vibfi  
 gch2\_sheon  
 gch2\_phoph  
 ribb\_phole  
 gch2\_psepu  
 gch2\_psesy  
 gch2\_pseae  
 ribb\_dehmu  
 gch2\_xylfa

TEII~  
 TEII~  
 TEII~  
 NELLEI~  
 YLEPGNRI..  
 YQSMADTP..  
 FEAEK~  
 YLECHAGE..  
 FLECHPGA..  
 YVCE~  
 YVCE~  
 YVTK~  
 CVAE~  
 YICN~  
 SIPFPGM..  
 SIPF.PG~  
 YLPAE~  
 STWNETV..  
 RISVIGQI..

.....RK~  
 .....QAT TAATH~  
 .....GVR .SVPPSAVP~  
 .....ASR VAAAAP~  
 .....DG~  
 .....A~  
 .....TKQ NEHYLRTKVS RAGHDLIDIDA LIMTSORPOD PSETVDGETV KPIAKTGHA~