



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0806790-2 A2**



(22) Data de Depósito: 11/01/2008
(43) Data da Publicação: 13/09/2011
(RPI 2123)

(51) *Int.Cl.:*
C12N 1/21
C12N 15/09
C12P 13/06
C12P 13/08
C12P 13/10
C12P 13/12
C12P 13/14
C12P 13/22
C12P 13/24

(54) **Título:** MICROORGANISMO, E, MÉTODO PARA PRODUZIR UM L-AMINOÁCIDO

(30) **Prioridade Unionista:** 22/01/2007 JP 2007-011392, 17/05/2007 JP 2007-131763, 17/05/2007 JP 2007-131763

(73) **Titular(es):** Ajinomoto CO., INC.

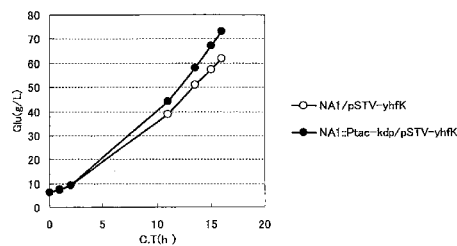
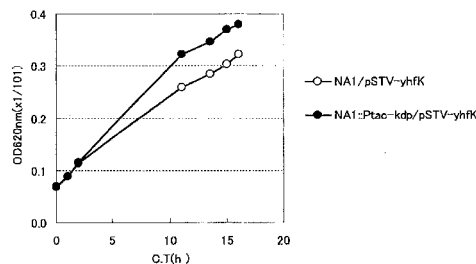
(72) **Inventor(es):** Rie Takikawa, Yoshihiko Hara

(74) **Procurador(es):** Momsen, Leonardos & Cia.

(86) **Pedido Internacional:** PCT JP2008050246 de 11/01/2008

(87) **Publicação Internacional:** WO 2008/090770de 31/07/2008

(57) **Resumo:** MICROORGANISMO, E, MÉTODO PARA PRODUZIR UM L-AMINOÁCIDO. Um L-aminoácido pode ser produzido por: cultivar, em um meio de cultura, um microorganismo pertencente à família Enterobacteriaceae, a qual é capaz de produzir o L-aminoácido e é modificada de modo que o sistema kdp possa ser potencializado, por esse meio produzindo e acumulando o L-aminoácido no meio de cultura ou em uma célula do microorganismo; e coletando-se o L-aminoácido do meio de cultura ou da célula.





“MICROORGANISMO, E, MÉTODO PARA PRODUZIR UM L-AMINOÁCIDO”

CAMPO TÉCNICO

5 A presente invenção diz respeito a um método para produzir um L-aminoácido com o uso de um microorganismo, em particular métodos para produzir um L-aminoácido em que dito L-aminoácido seja o ácido L-glutâmico, L-lisina, L-treonina, L-triptofano ou similar. Estes são L-aminoácidos industrialmente úteis, isto é, o ácido L-glutâmico serve como um aditivo para condimento, e a L-lisina, a L-treonina e o L-triptofano servem
10 como aditivos para a alimentação animal, ingredientes alimentares para a saúde, infusões de aminoácidos, e assim por diante.

FUNDAMENTOS DA TÉCNICA

Os L-aminoácidos são industrialmente produzidos por fermentação, com o uso de vários microorganismos. Por exemplo, o ácido L-glutâmico é produzido principalmente pela fermentação utilizando-se
15 bactérias produtoras do ácido L-glutâmico das assim chamadas bactérias corineiforme pertencente aos gêneros *Brevibacterium*, *Corynebacterium* ou *Microbacterium*, ou suas cepas mutantes (ver, por exemplo, o documento não patente 1). Como métodos para produzir ácido L-glutâmico pela fermentação com o uso de outras cepas bacterianas, métodos de uso de um
20 microorganismo pertencente ao gênero *Bacillus*, *Streptomyces*, *Penicillium* ou similar (referir-se, por exemplo, ao documento de Patente 1), métodos de uso de um microorganismo pertencente ao gênero *Pseudomonas*, *Arthrobacter*, *Serratia*, *Candida* ou similar (referir-se, por exemplo, ao documento de Patente 2), métodos de uso de um microorganismo pertencente ao gênero
25 *Bacillus*, *Pseudomonas*, *Serratia*, *Aerobacter aerogenes* (correntemente referido como *Enterobacter aerogenes*) ou similar (referir-se, por exemplo, ao documento de Patente 3), métodos de uso de uma cepa mutante de *Escherichia coli* (referir-se, por exemplo, ao documento de Patente 1), e assim

por diante, são conhecidos. Além disso, métodos para produzir o ácido L-glutâmico com o uso de um microorganismo pertencente aos gêneros *Klebsiella*, *Erwinia*, *Pantoea* ou *Enterobacter*, são também apresentados (referir-se, por exemplo, aos documentos de Patente 2 a 4).

5 Tais métodos para produzir substâncias alvo tais como os L-aminoácidos por fermentação com o uso de um microorganismo como descrito acima, incluem métodos que usam um microorganismo do tipo selvagem (cepa tipo selvagem), métodos de uso de uma cepa auxotrófica derivada de uma cepa do tipo selvagem, métodos de uso de uma cepa mutante
10 de regulação metabólica derivada de uma cepa do tipo selvagem como uma cepa resistente a qualquer dos vários medicamentos, métodos de uso de uma cepa tendo propriedades tanto da cepa auxotrófica quanto da cepa mutante de regulação metabólica, e assim por diante.

 Nos últimos anos, as técnicas de DNA recombinantes são
15 usadas na produção de substâncias alvo mediante fermentação. Por exemplo, a produtividade do L-aminoácido de um microorganismo é melhorada pela intensificação da expressão de um gene codificando uma enzima biossintética do L-aminoácido (Documentos de Patentes 5 e 6), ou pela intensificação do influxo de uma fonte de carbono dentro de um sistema de biossíntese de L-
20 aminoácido (Documento de Patente 7).

 O sistema *kdp* é uma ATPase do tipo P que funciona para absorver íons de potássio (Documento não Patente 2). O sistema *kdp* é codificado pelo óperon *kdp*, e sua expressão é induzida quando a concentração dos íons de potássio em um meio é baixa, ou quando a cultura é
25 realizada sob condições hiperosmótica (Documento não Patente 3). Além disso, sabe-se que a expressão é controlada por KdpD e KdpE que constituem um dos sistema de controle binário (Documento não Patente 4). Entretanto, a relação entre a intensificação do sistema *kdp* e a produção do L-aminoácido não foi até agora pesquisada.

Documento de Patente 1: Patente Japonesa Aberta ao Público
(KOKAI) nº 5-244970

Documento de Patente 2: Patente U.S. nº 3.563.857

Documento de Patente 3: Publicação da Patente Japonesa
5 (KOKOKU) nº 32-9393

Documento de Patente 4: Patente Japonesa Aberta ao Público
nº 2000-189175

Documento de Patente 5: Patente U.S. nº 5.168.056

Documento de Patente 6: Patente U.S. nº 5.776.736

10 Documento de Patente 7: Patente U.S. nº 5.906.925

Documento não Patente 1: Kunihiko Akashi *et al.*, “*Amino acid fermentation*”, pp.195-215, 1986, Japan Scientific Societies Press

Documento não Patente 2: Laimonis A. Laimins, *Proc. Natl. Acad. Sci. USA*, julho de 1978, 75(7): 3216-3219

15 Documento não Patente 3: Laimonis A. Laimins, *Proc. Natl. Acad. Sci. USA*, janeiro de 1981, 78(1): 464-468

Documento não Patente 4: Mark O. Walderhaug, *J. Bacteriol.*,
abril de 1992, 174 (7): 2152-2159

APRESENTAÇÃO DA INVENÇÃO

20 OBJETO PARA SER ALCANÇADO PELA INVENÇÃO

Um objeto da presente invenção é prover um microorganismo que pertença à família *Enterobacteriaceae* e seja capaz de eficientemente produzir um L-aminoácido, e prover um método de produzir eficientemente um L-aminoácido com o uso de um tal microorganismo.

25 MEIOS PARA SE ALCANÇAR O OBJETO

Os inventores da presente invenção conduziram várias pesquisas a fim de alcançar o objeto acima mencionado e, como um resultado, observaram que os L-aminoácidos poderiam ser eficientemente produzidos mediante o uso de um microorganismo do qual o sistema *kdp* foi

intensificado, e assim concluíram a presente invenção.

Isto é, a presente invenção provê os seguintes itens:

(1) um microorganismo pertencente à família Enterobacteriaceae, o qual tem a capacidade de produzir L-aminoácidos e foi modificado de modo que o sistema *kdp* seja intensificado.

(2) O microorganismo acima mencionado, em que o sistema *kdp* é intensificado pelo aumento da expressão do óperon de *kdp* ou um ou mais genes constituindo o óperon de *kdp*, e/ou aumentando a translação do óperon de *kdp* ou os genes.

(3) O microorganismo acima mencionado, em que o sistema *kdp* é intensificado pelo aumento do número de cópias do óperon de *kdp* ou um ou mais genes constituindo o óperon de *kdp*, ou modificando uma sequência de controle da expressão do óperon.

(4) O microorganismo acima mencionado, em que o óperon de *kdp* contenha pelo menos os genes *kdpA*, *kdpB* e *kdpC*.

(5) O microorganismo acima mencionado, em que o gene *kdpA* é um gene codificando uma proteína tendo a sequência de aminoácidos apresentada na SEQ ID NO: 2 ou 8 ou a subunidade A do sistema *kdp* tendo a sequência de aminoácidos de SEQ ID NO: 2 ou 8 incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

(6) O microorganismo acima mencionado, em que o gene *kdpB* é um gene codificando uma proteína tendo a sequência de aminoácidos apresentada na SEQ ID NO: 3 ou 9 ou a subunidade B do sistema *kdp* tendo a sequência de aminoácidos de SEQ ID NO: 3 ou 9 incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

(7) O microorganismo acima mencionado, em que o gene *kdpC* é um gene codificando uma proteína tendo a sequência de aminoácidos apresentada na SEQ ID NO: 4 ou 10 ou a subunidade C do sistema *kdp* tendo a sequência de aminoácidos de SEQ ID NO: 4 ou 10 incluindo substituições,

deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

(8) O microorganismo acima mencionado, em que o óperon de *kdp* é um DNA definido em qualquer um dos seguintes itens (a) a (d):

5 (a) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1,

(b) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1 ou uma sonda preparada com a sequência de nucleotídeos sob condições estringentes e codificando o sistema *kdp*,

10 (c) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7,

(b) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7 ou uma sonda preparada da sequência de nucleotídeos sob condições estringentes e
15 codificando o sistema *kdp*.

(9) O microorganismo acima mencionado, em que o L-aminoácido é uma ou mais espécies de L-aminoácidos selecionados do grupo consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoileucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-
20 triptofano e L-cisteína.

(10) O microorganismo acima mencionado, em que o microorganismo pertencente à família *Enterobacteriaceae* é uma bactéria *Escherichia*, uma bactéria *Enterobacter* ou uma bactéria *Pantoea*.

(11) Um método para produzir um L-aminoácido, que
25 compreende cultivar o microorganismo acima mencionado em um meio para produzir e acumular um L-aminoácido no meio ou nas células, e coletar o L-aminoácido do meio ou das células.

(12) O método supramencionado, em que o L-aminoácido constitui uma ou mais espécies de L-aminoácidos selecionados do grupo

consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoleucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-triptofano e L-cisteína.

BREVE DESCRIÇÃO DOS DESENHOS

5 A Figura 1 é um desenho que mostra a estrutura do plasmídeo auxiliar RSF-Red-TER.

A Figura 2 é um desenho que mostra a construção do plasmídeo auxiliar RSF-Red-TER.

10 A Figura 3 é um desenho que mostra a estrutura da região do cromossoma de *P. ananatis* localizando-se a montante do gene *LacZ*.

A Figura 4 é um gráfico que mostra o crescimento da cepa substituída pelo promotor do óperon de *kdp* na cultura sob condição acídica no tubo de teste.

15 A Figura 5 é um gráfico que mostra a produtividade do ácido L-glutâmico da cepa substituída pelo promotor do óperon de *kdp*.

A Figura 6 é um desenho que mostra o alinhamento das sequências de aminoácidos do KdpA de *Pantoea ananatis* (SEQ ID NO: 8) e de *Escherichia coli* (SEQ ID NO: 2), e da sequência de consenso delas (SEQ ID NO: 57).

20 A Figura 7 é um desenho que mostra o alinhamento das sequências de aminoácidos do KdpB de *Pantoea ananatis* (SEQ ID NO: 9) e de *Escherichia coli* (SEQ ID NO: 3), e da sequência de consenso delas (SEQ ID NO: 58).

25 A Figura 8 é um desenho que mostra o alinhamento das sequências de aminoácidos do KdpC de *Pantoea ananatis* (SEQ ID NO: 10) e de *Escherichia coli* (SEQ ID NO: 4), e da sequência de consenso delas (SEQ ID NO: 59).

MELHOR MODO PARA REALIZAR A INVENÇÃO

Daqui por diante a presente invenção será explanada em

detalhes.

<1> MICROORGANISMO DA PRESENTE INVENÇÃO

O microorganismo da presente invenção é um microorganismo pertencente à família *Enterobacteriaceae*, a qual tem uma capacidade de produzir o L-aminoácido e foi modificado de modo que o sistema *kdp* seja intensificado. A capacidade de produzir o L-aminoácido aqui referida significa uma capacidade do microorganismo da presente invenção para produzir e acumular um L-aminoácido em um meio ou células em uma quantidade em que o L-aminoácido possa ser coletado do meio ou das células, quando o microorganismo seja cultivado no meio. O microorganismo da presente invenção pode ter uma capacidade de produzir duas ou mais espécies de L-aminoácidos. O microorganismo tendo capacidade de produzir L-aminoácidos pode ser um microorganismo inerentemente tendo uma capacidade de produzir L-aminoácidos, ou um microorganismo obtido pela modificação de tais microorganismos como descrito abaixo, de modo a ter uma capacidade de produzir L-aminoácidos com o uso de um método de mutação ou de técnicas de DNA recombinantes.

O tipo do L-aminoácido não é particularmente limitado, e os exemplos incluem aminoácidos básicos tais como L-lisina, L-ornitina, L-arginina, L-histidina e L-citrulina, aminoácidos alifáticos tais como a L-isoleucina, L-alanina, L-valina, L-leucina e L-glicina, aminoácidos que sejam ácidos hidroximonoaminocarboxílicos tais como a L-treonina e a L-serina, aminoácidos cíclicos tais como a L-prolina, aminoácidos aromáticos tais como a L-fenilalanina, L-tirosina e L-triptofano, aminoácidos contendo enxofre tais como a L-cisteína, L-cistina e L-metionina, e aminoácidos acídicos tais como o ácido L-glutâmico, o ácido L-aspártico, a L-glutamina e a L-asparagina. O ácido L-glutâmico, a L-lisina, a L-treonina e o L-triptofano são especialmente preferidos. O microorganismo da presente invenção pode ter uma capacidade de produzir duas ou mais espécies de aminoácidos.

<1-1> TRANSMISSÃO DA CAPACIDADE DE PRODUZIR L-AMINOÁCIDOS

Exemplos dos métodos para transmitir a capacidade de produzir L-aminoácidos e microorganismos utilizáveis na presente invenção, aos quais a capacidade de produzir L-aminoácidos é comunicada, serão descritos abaixo. Entretanto, o microorganismo não fica limitado a estes, contanto que um microorganismo tendo a capacidade de produzir L-aminoácidos seja usado.

Microorganismos usados para a presente invenção incluem opcionalmente substituído microorganismos pertencentes aos gêneros *Escherichia*, *Enterobacter*, *Pantoea*, *Klebsiella*, *Serratia*, *Erwinia*, *Salmonella*, *Morganella*, ou similar, contanto que eles pertençam à família *Enterobacteriaceae* e tenham uma capacidade de produzir L-aminoácido. Em particular as bactérias classificadas dentro da família *Enterobacteriaceae* de acordo com a taxonomia usada pela base de dados da NCBI (National Center for Biotechnology Information) (<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=91347>) podem ser usadas. Como cepas precursoras das *Enterobacteriaceae* que devam ser modificadas, as bactérias pertencentes aos gêneros *Escherichia*, *Enterobacter*, *Pantoea*, *Erwinia*, *Enterobacter* ou *Klebsiella* são preferivelmente usadas.

A cepa precursora das bactérias de *Escherichia* usadas de modo a se obter uma bactéria de *Escherichia* da presente invenção, não é particularmente limitada. Aquelas descritas no trabalho de Neidhardt *et al.* (Backmann, B. J., 1996. *Derivations and Genotypes of some mutant derivatives of Escherichia coli K-12*, pp. 2460-2488, Tabela 1, em F. D. Neidhardt (ed.), *Escherichia coli and Salmonella Cellular and Molecular Biology/Segunda Edição*, American Society for Microbiology Press, Washington, D.C.) podem ser utilizadas. Entre elas, por exemplo, a

Escherichia coli é exemplificada. Exemplos de *Escherichia coli* incluem a cepa W3110 (ATCC nº 27325), a cepa MG1655 (ATCC nº 47076), e assim por diante, as quais são derivadas de uma cepa protótipo do tipo selvagem, a cepa K12.

5 Estas cepas acham-se disponíveis, por exemplo, da American Type Culture Collection (ATCC) (Endereço: 12301 Parklawn Drive, Rockville, Maryland 20852, P.O. Box 1549, Manassas, VA 20108, Estados Unidos da América) Isto é, cada cepa recebe um único número de registro (<http://www.atcc.org/>). As cepas podem ser ordenadas pelo uso deste número
10 de registro. O número de registro de cada cepa acha-se listado no catálogo da ATCC.

 Exemplos das bactérias *Enterobacter* incluem, *Enterobacter agglomerans*, *Enterobacter aerogenes* e assim por diante. Exemplos das bactérias *Pantoea* incluem a *Pantoea ananatis*. Nos últimos anos, algumas
15 bactérias de *Enterobacter agglomerans* foram reclassificadas como *Pantoea agglomerans*, *Pantoea ananatis* ou *Pantoea stewartii*, com base na análise da sequência de nucleotídeos dos rRNA 16S etc. Na presente invenção, o microorganismo pode pertencer ou ao gênero *Enterobacter* ou ao *Pantoea*, contanto que o microorganismo seja classificado dentro da família
20 *Enterobacteriaceae*.

 Em particular, as bactérias *Pantoea*, as bactérias *Erwinia* e as bactérias *Enterobacter* são classificadas como γ -proteobactérias, e elas se acham taxonomicamente muito próximas uma da outra (*J. Gen. Appl. Microbiol.*, 1997, 43, 355-361; *Int. J. Syst. Bacteriol.*, 1997, 43, 1061-1067).
25 Nos últimos anos, algumas bactérias pertencentes ao gênero *Enterobacter* foram reclassificadas como *Pantoea agglomerans*, *Pantoea dispersa*, ou similar, com base nas experiências de hibridização de DNA-DNA etc. (*International Journal of Systematic Bacteriology*, julho de 1989, 39: 337-345). Além disso, algumas bactérias pertencentes ao gênero *Erwinia* foram

reclassificadas como *Pantoea ananas* ou *Pantoea stewartii* (referir-se ao *Int. J. Syst. Bacteriol.*, 1993, 43: 162-173).

Exemplos das bactérias de *Enterobacter* incluem *Enterobacter agglomerans*, *Enterobacter aerogenes*, e assim por diante. Especificamente, as cepas exemplificadas na Patente Européia Aberta ao Público nº 952221 podem ser usadas.

Uma cepa típica do gênero *Enterobacter* é a cepa *Enterobacter agglomerans* ATCC 12287.

Cepas típicas das bactérias *Pantoea* incluem as *Pantoea ananatis*, *Pantoea stewartii*, *Pantoea agglomerans* e *Pantoea citrea*. Exemplos específicos incluem as seguintes cepas:

Pantoea ananatis AJ13355 (FERM BP-6614, Patente Européia Aberta ao Público nº 0952221)

Pantoea ananatis AJ13356 (FERM BP-6615, Patente Européia Aberta ao Público nº 0952221)

Pantoea ananatis AJ13601 (FERM BP-7207, Patente Européia Aberta ao Público nº 0952221)

Não obstante estas cepas terem sido identificadas e depositadas como *Enterobacter agglomerans* quando foram isoladas, elas são correntemente classificadas como *Pantoea ananatis* com base na análise da sequência de nucleotídeos do rRNA 16S etc., como descrito acima.

Exemplos das bactérias *Erwinia* incluem as *Erwinia amylovora* e as *Erwinia carotovora*, e exemplos das bactérias *Klebsiella* incluem a *Klebsiella planticola*. Exemplos específicos incluem as seguintes cepas:

Erwinia amylovora ATCC 15580

Erwinia carotovora ATCC 15713

Klebsiella planticola AJ13399 (FERM BP-6600, Patente Européia Aberta ao Público nº 955368)

Klebsiella planticola AJ13410 (FERM BP-6617, Patente Européia Aberta ao Público nº 955368).

Daqui por diante, os métodos para comunicar uma capacidade de produção de L-aminoácidos às bactérias de *Enterobacteriaceae*, ou métodos para intensificar uma capacidade de produção de L-aminoácidos de tais bactérias, são descritos.

Para comunicar uma capacidade de produzir um L-aminoácido, os métodos convencionalmente empregados na reprodução das bactérias corineformes ou das bactérias do gênero *Escherichia* (ver “*Amino Acid Fermentation*”, Gakkai Shuppan Center (Ltd.), 1ª Edição, publicada em 30 de maio de 1986, pp. 77-100) podem ser usados. Tais métodos incluem a aquisição de um mutante auxotrófico, uma cepa resistente analógica, ou um mutante de regulação metabólica, a construção de uma cepa recombinante em que a expressão de uma biossíntese de L-aminoácido seja intensificada, e assim por diante. Aqui, na reprodução de uma bactéria produtora de L-aminoácidos, as propriedades transmitidas, tais como uma mutação auxotrófica, a resistência analógica, ou a mutação de regulação metabólica, podem ser uma ou mais. A expressão da(s) enzima(s) de biossíntese de L-aminoácidos pode ser intensificada isoladamente ou em combinações de duas ou mais. Além disso, os métodos de comunicar propriedades tais como uma mutação auxotrófica, resistência analógica, ou mutação de regulação metabólica, podem ser combinados com os métodos de intensificar as enzimas de biossíntese.

Uma cepa mutante auxotrófica, a cepa resistente analógica de L-aminoácido, ou a cepa mutante de regulação metabólica com uma capacidade de produzir um L-aminoácido, podem ser obtidas submetendo-se uma cepa precursora ou uma cepa do tipo selvagem a mutagênese convencional, tal como a exposição dos raios-X ou a irradiação de UV, ou o tratamento com um mutágeno tal como N-metil-N'-nitro-N-nitrosoguanidina

(NTG), metanossulfonato de etila (EMS), etc., depois selecionando-se aqueles que apresentem autotrofismo, resistência analógica, ou uma mutação de regulação metabólica, e que também tenham uma capacidade de produzir um L-aminoácido.

5 As bactérias produtoras de L-aminoácido ou os métodos de construção para tal, são exemplificados abaixo.

BACTÉRIAS PRODUTORAS DE ÁCIDO L-GLUTÂMICO

Primeiramente, as bactérias produtoras de ácido L-glutâmico são explanadas como bactérias produtoras de L-aminoácido.

10 Exemplos das cepas precursoras para originar bactérias produtoras de ácido L-glutâmico da presente invenção, incluem, sem limitar, as cepas pertencentes ao gênero *Escherichia*, tais como *E. coli* VL334thrC⁺ (Patente Européia nº 1172433). *E. coli* VL334 (VKPM B-1641) é uma cepa auxotrófica de L-isoleucina e L-treonina tendo mutações nos genes *thrC* e
15 *ilvA* (Patente U.S. nº 4.278.765). Um alelo do tipo selvagem do gene *thrC* foi transferido pelo método de transdução geral com o uso de um bacteriófago P1 crescido nas células da cepa K12 de *E. coli* do tipo selvagem (VKPM B-7). Como resultado, uma cepa auxotrófica de L-isoleucina VL334thrC⁺ (VKPM B-8961) foi obtida.

20 Exemplos dos métodos para comunicar a capacidade de produzir o ácido L-glutâmico a uma bactéria, ou intensificar a capacidade de uma bactéria, incluem, por exemplo, modificar uma bactéria de modo que a expressão de um gene codificando uma enzima envolvida na biossíntese do ácido L-glutâmico seja intensificada.

25 Exemplos de enzimas envolvidas na biossíntese do ácido L-glutâmico incluem a glutamato desidrogenase (doravante também referida como “GDH”) (*gdh*), a glutamina sintetase (*glnA*), a glutamato sintetase (*gltAB*), a isocitrato desidrogenase (*icdA*), a aconitato hidratase (*acnA*, *acnB*), a citrato sintase (daqui por diante também referida como “CS”) (*gltA*),

metilcitrato sintase (daqui por diante também referida como “PRPC” (*prpC*), fosfoenolpiruvato carboxilase (daqui por diante também referida como “PEPC” (*ppc*), a piruvato carboxilase (*pyc*), a piruvato desidrogenase (*aceEF*, *lpdA*), piruvato cinase (*pykA*, *pykF*), fosfoenolpiruvato sintase (*ppsA*), enolase
 5 (*eno*), fosfogliceromutase (*pgmA*, *pgmI*), fosfoglicerato cinase (*pgk*), gliceraldeído-3-fosfato desidrogenase (*gapA*), triose fosfato isomerase (*tpiA*), frutose bisfosfato aldolase (*fbp*), fosfofrutocinase (*pfkA*, *pfkB*) e glicose fosfato isomerase (*pgi*), e assim por diante. As abreviaturas entre parênteses são os nomes dos genes que correspondem às enzimas, e esta convenção é
 10 usada na totalidade deste relatório descritivo. Entre estas enzimas, um ou mais dentre CS ou PRPC, PEPC e GDH são preferidas, e todas as três enzimas são mais preferidas (referir-se à WO 2006/051660).

Métodos para modificar uma bactéria para aumentar a expressão dos genes alvo serão elucidados abaixo.

15 O primeiro método é um método de aumentar o número de cópias de um gene alvo. Por exemplo, o número de cópias de um gene alvo pode ser aumentado pela clonagem do gene alvo em um plasmídeo apropriado e transformando-se uma bactéria hospedeira com o plasmídeo obtido. Por exemplo, quando o gene alvo é o gene que codifica CS (gene *gltA*), o gene
 20 que codifica PRPC (gene *prpC*), o gene que codifica PEPC (gene *ppc*) ou o gene que codifica GDH (gene *gdhA*), as sequências de nucleotídeos destes genes das bactérias *Escherichia* e das bactérias *Corynebacterium* já terão sido elucidadas (*Biochemistry*, vol. 22, pp. 5243-5249, 1983; *J. Biochem.*, vol. 95, pp. 909-916, 1984; *Gene*, vol. 27, pp. 193-199, 1984; *Microbiology*, vol. 140,
 25 pp. 1817-1828, 1994; *Mol. Gen. Genet.*, volume 218, pp. 330-339, 1989; *Molecular Microbiology*, volume 6, pp. 317-326, 1992) e, portanto, elas poderão ser obtidas pela sintetização de iniciadores com base nas respectivas sequências de nucleotídeos, e realizando-se a PCR com o uso de DNA cromossômico de uma bactéria pertencente à família *Enterobacteriaceae*

como o padrão.

Exemplos do plasmídeo usado para a transformação incluem um plasmídeo que autonomamente replica na bactéria hospedeira pertencente à família *Enterobacteriaceae*, tal como pUC19, pUC18, pBR322, RSF1010, pHSG299, pHSG298, pHSG399, pHSG398, pSTV28, pSTV29 (pHSG e pSTV acham-se disponíveis da Takara Bio Inc.), pMW119, pMW118, pMW219, pMW218 (os vetores pMW acham-se disponíveis da Nippon Gene Co., Ltd.), e assim por diante. Além disso, um DNA de fago pode também ser usado como o vetor, ao invés de um plasmídeo. Exemplos de plasmídeo para simultaneamente intensificar as atividades de CS ou PRPC, PEPC e CDH descritos acima, incluem RSFCPG incorporados com o gene *gltA*, o gene *ppc* e o gene *gdhA* (referir-se à Patente Européia Aberta ao Público nº 0952221), e RSFPPG correspondendo a RSFCPG, em que o gene *gltA* é substituído pelo gene *prpC* (referir-se aos exemplos).

Exemplos de métodos de transformação incluem o tratamento de células receptoras com cloreto de cálcio de modo a aumentar a permeabilidade do DNA, o que foi relatado quanto a *Escherichia coli* K-12 (Mandel, M. e Higa, A., 1970, *J. Mol. Biol.*, 53: 159-162), e preparando-se as células competentes a partir das células que se achem na fase de crescimento, seguido pela transformação com DNA, o que foi relatado quanto ao *Bacillus subtilis* (Duncan, C. H., Wilson, G. A. e Young, F. E. 1977, *Gene*, 1: 153-167). Alternativamente, um método de produzir células receptoras de DNA em protoplastos ou esferoplastos, que podem facilmente absorver DNA recombinante, seguido pela introdução do DNA recombinante nas células, o que é conhecido como sendo aplicável a *Bacillus subtilis*, actinomicetos e leveduras (Chang, S. e Choen, S. N., 1979, *Mol. Gen. Genet.*, 168: 111-115; Bibb, M. J. *et al.*, 1978, *Nature*, 274: 398-400; Hinnen, A., Hicks, J. B. e Fink, G. R. 1978, *Proc. Natl. Sci., USA*, 75: 1929-1933), pode também ser empregado. Além disso, os microorganismos podem também ser

transformados pelo método do pulso elétrico (Patente Japonesa Aberta ao Público nº 2-207791).

O número de cópias de um gene pode também ser aumentado pela introdução de cópias múltiplas do gene no DNA cromossômico do microorganismo, o que pode ser realizado por recombinação homóloga (MillerI, J. H. *Experiments in Molecular Genetics*, 1972, Cold Spring Harbor Laboratory) com o uso de cópias múltiplas de uma sequência como alvos no DNA cromossômico. As sequências presentes nas cópias múltiplas no DNA cromossômico incluem os DNAs respectivos, e as repetições invertidas presentes na extremidade de um elemento transponível. Igualmente, como apresentado na Patente Japonesa Aberta ao Público nº 2-109985, é possível incorporar o gene alvo em um transpóson, e possibilitar que ele seja transferido para introduzir cópias múltiplas do gene para o DNA cromossômico. O gene alvo pode também ser introduzido no cromossoma bacteriano por fago Mu (Patente Japonesa Aberta ao Público nº 2-109985), ou similar.

O segundo método é aumentar a expressão do gene alvo pela substituição de uma sequência reguladora da expressão do gene alvo, tal como um promotor, sobre o DNA cromossômico ou plasmídeo com um promotor mais forte. Por exemplo, o promotor lac, o promotor trp, o promotor trc, o promotor PR, o promotor lacUV, etc., são conhecidos como promotores fortes. Além disso, é também possível substituir vários nucleotídeos na região do promotor de um gene, de modo que o promotor seja modificado para ser mais forte, como apresentado na Publicação da Patente Internacional WO 00/18935. Exemplos de promotores fortes e de métodos para avaliar a intensidade dos promotores, são descritos em um artigo de Goldstein *et al.* (*Prokaryotic promoters in biotechnology. Biotechnol. Annu. Rev.*, 1995, 1, 105-128), etc.

A substituição de uma sequência reguladora da expressão pode

ser realizada, por exemplo, da mesma maneira como na substituição do gene com o uso de um plasmídeo sensível à temperatura. Exemplos de vetores que tenham uma origem de replicação sensível à temperatura e utilizável para a bactéria da presente invenção pertencente à família *Enterobacteriaceae* incluem por exemplo, o plasmídeo pMAN997 descrito na Publicação Internacional WO 99/03988, e assim por diante.

Além disso, sabe-se que as substituições de vários nucleotídeos em um espaçador entre o sítio de ligação ribossômica (RMS) e o códon de partida, em particular uma sequência imediatamente a montante do códon de partida, afetam grandemente a eficiência da translação do mRNA. Pela modificação destas, a quantidade de translação pode ser melhorada.

A modificação de uma sequência de controle da expressão pode ser combinada com o método de aumentar o número de cópias de um gene descrito acima.

Exemplos dos métodos para a substituição de genes conforme descrito acima, incluem os métodos que empregam DNA linear, tais como o “*Red-driven integration*” [Datsenko, K. A. e Wanner, B. L., *Proc. Natl. Acad. Sci. USA.*, 97: 6640-6645 (2000)], e o *Red-driven integration* em combinação com o sistema de excisão de fago λ (Cho, E. H., Gumport, R. I., Gardner, J. F. 2002, *J. Bacteriol.*, 184: 5200-5203) (WO 2005/010175), e assim por diante, métodos que usam um plasmídeo contendo uma origem de replicação sensível à temperatura, métodos que usam um plasmídeo capaz de transferência conjugativa, métodos que utilizam um vetor suicida que não tenha uma origem de replicação utilizável no hospedeiro escolhido (Patente U.S. nº 6.303.383, Patente Japonesa Aberta ao Público nº 05-007491) etc.

Como mostrado no Exemplo de Referência 1, uma cepa resistente ao produto de gene *Red* λ , por exemplo a cepa de *Pantoea ananatis* SC17 (0), pode ser adequadamente usada para a integração conduzida de *Red*. A cepa SC17 (0) foi depositada na Russian National Collection of Industrial

Microorganisms (VKPM), GNII Genetika (Rússia, 117545 Moscou 1, Dorozhny proezd. 1) em 21 de setembro de 2005, sob o número de acesso da VKPM B-9246.

Exemplos de microorganismos modificados pelo método descrito acima, de modo que a expressão do gene da citrato sintase, do gene da citrato de metila sintase, do gene da fosfoenolpiruvato carboxilase e/ou do gene de glutamato desidrogenase seja intensificada, incluem os microorganismos descritos nas Patentes Japonesas Abertas ao Público n^{os} 2001-333769, 2000-106869, 2000-189169 2000-333769, 2006-129840, WO 10 2006/051660, e assim por diante.

Além disso, a capacidade de produzir o ácido L-glutâmico pode também ser transmitida pela intensificação da atividade da 6-fosfogliconato desidratase, da atividade da 2-ceto-3-desóxi-6-fosfogliconato aldolase, ou de ambas estas atividades. Exemplos do microorganismo do qual a atividade da 6-fosfogliconato desidratase e a atividade da 2-ceto-3-desóxi-6-fosfogliconato aldolase são aumentadas, incluem o microorganismo apresentado na Patente Japonesa Aberta ao Público n^o 2003-274988.

A modificação para transmitir a capacidade de produzir o ácido L-glutâmico ou de intensificá-la, pode também ser alcançada pela redução ou eliminação da atividade de uma enzima que catalise uma reação que se ramifique da via da biossíntese do ácido L-glutâmico, e produção de um composto outro que não o ácido L-glutâmico. Exemplos de uma tal enzima que catalise uma reação que se ramifique da via de biossíntese do ácido L-glutâmico e produza um composto outro que não o ácido L-glutâmico, incluem 2-oxoglutarato desidrogenase [α -cetoglutarato desidrogenase (*sucA*)], isocitrato liase (*aceA*), fosfato acetiltransferase (*pta*), acetatocinase (*ack*), ácido acetoidróxi sintase (*ilvG*), acetolactato sintase (*ilvI*), formiato acetiltransferase (*pfl*), lactato desidrogenase (*ldh*), glutamato descarboxilase (*gadAB*), 1-pirrolina-5-carboxilato desidrogenase (*putA*), e

assim por diante. É particularmente preferível reduzir ou eliminar a atividade da 2-oxoglutarato desidrogenase entre estas enzimas.

A fim de reduzir ou eliminar as atividades das enzimas acima mencionadas, as mutações para reduzir ou eliminar as atividades intracelulares das enzimas podem ser introduzidas nos genes das enzimas supramencionadas mediante um tratamento usual de mutagênese ou uma técnica de engenharia genética. Exemplos do tratamento de mutagênese incluem, por exemplo, métodos que utilizam a irradiação de raio-x ou o raio ultravioleta, métodos que utilizam o tratamento com um mutágeno tal como a N-metil-N'-nitro-N-nitrosoguanidina, e assim por diante. O local do gene em que a mutação é introduzida pode ser uma região codificadora codificando uma proteína enzimática ou uma região para regular a expressão tal como um promotor. Exemplos das técnicas de engenharia genética incluem métodos que usam recombinação genética, transdução, fusão celular, e assim por diante.

O decréscimo ou a deficiência da atividade intracelular de uma enzima alvo e o grau de decréscimo na atividade podem ser confirmados pela medição da atividade enzimática em um extrato celular ou uma sua fração purificada obtida de uma cepa candidata e comparando-a com aquela de uma cepa do tipo selvagem. Por exemplo, a atividade da 2-oxoglutarato desidrogenase pode ser medida pelo método de Reed *et al.* [L. J. Reed e B. B. Mukherjee, *Methods in Enzymology*, 13, pp. 55-61 (1969)].

As bactérias pertencentes ao gênero *Escherichia* deficientes na atividade de 2-oxiglutarato desidrogenase ou tendo um atividade reduzida de 2-oxiglutarato desidrogenase, incluem as seguintes cepas (Patentes U.S. n^{os} 5.378.616 e 5.573.945):

E. coli W3110sucA::Kmr

E. coli AJ12624 (FERM BP-3853)

E. coli AJ12628 (FERM BP-3854)

E. coli AJ12949 (FERM BP-4881)

E. coli W3110sucA::Kmr é uma cepa obtida pelo rompimento do gene de 2-oxoglutarato desidrogenase (gene *sucA*) de *E. coli* W3110. Esta cepa é completamente deficiente na α -cetoglutarato desidrogenase.

5 Especificamente, os exemplos de bactéria em que a atividade da 2-oxoglutarato desidrogenase é deletada ou reduzida, incluem as seguintes cepas

Pantoea ananatis AJ13601 (FERM BP-7207, Patente Européia Aberta ao Público nº 1078989)

10 *Pantoea ananatis* AJ13356 (FERM BP-6615, Patente dos Estados Unidos nº 6.331.419)

Pantoea ananatis SC17sucA (FERM BP-8646, WO 2005/085419)

15 *Klebsiella planticola* cepa AJ13410 (FERM BP-6617, Patente dos Estados Unidos nº 6.197.559)

A cepa SC17sucA é uma cepa obtida pela seleção de uma cepa mutante (SC17) de baixa produção de phlegm da cepa AJ13355, que foi isolada da natureza como uma cepa que poderia proliferar em um meio contendo o ácido L-glutâmico e uma fonte de carbono em baixo pH, e rompendo o gene da 2-oxoglutarato desidrogenase (*sucA*) da cepa mutante. A cepa AJ13601 foi obtida pela introdução de um plasmídeo RSFCPG contendo os genes *gltA*, *ppc* e *gdhA* derivados de *Escherichia coli* e um plasmídeo pSTVCB contendo o gene *gltA* derivado da *Brevibacterium lactofermentum* dentro da cepa SC17sucA para se obter a cepa
25 SC17sucA/RSFCPG+pSTVCB, e pela seleção de uma cepa resistente ao ácido L-glutâmico de alta concentração em baixo pH, e uma cepa apresentando um elevado grau de proliferação e uma elevada capacidade de produzir o ácido L-glutâmico da cepa SC17sucA/RSFCPG+pSTVCB (Patente Européia Aberta ao Público nº 0952221). A cepa AJ13356 foi obtida pela

eliminação do gene da subunidade α KGDH-E1 (*sucA*) da cepa AJ13355. Além disso, a cepa NP106 descrita nos exemplos corresponde à cepa AJ13601, da qual o plasmídeo RSFCPG+pSTVCB é eliminado.

As cepas *Pantoea ananatis* AJ13355 e AJ13356 foram depositadas em 19 de fevereiro de 1998 no National Institute of Bioscience and Human Technology da Agency of Industrial Science and Technology (correntemente agência administrativa independente, o National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository, Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japão), com os números de depósito FERM P-16644 e FERM P-16645 respectivamente, e transferidas do depósito original para o depósito internacional com base no Tratado de Budapeste em 11 de janeiro de 1999, e foi depositados sob os números de acesso FERM BP-6644 e FERM BP-6615, respectivamente. À cepa SC17*sucA* foi atribuído um número individual AJ417, e foi depositado no National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, Japão, código postal: 305-8566) em 26 de fevereiro de 2004, tendo sido atribuído um número de acesso FERM BP-08646. A cepa de *Pantoea ananatis* AJ13601 foi depositada em 18 de agosto de 1999 no National Institute of Bioscience and Human Technology of Agency of Industrial Science and Technology (correntemente uma agência administrativa independente, o National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository, Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japão), com o número de depósito FERM P-17156, e transferida do depósito original para o depósito internacional com base no Tratado de Budapeste em 6 de julho de 2000, e foi depositada sob o número de acesso FERM BP-7207.

As cepas de *Pantoea ananatis* AJ13355, AJ13356 e AJ13601 e

a cepa de *Klebsiella planticola* AJ13399 têm uma capacidade de acumular ácido L-glutâmico em uma quantidade que excede a quantidade que fornece a concentração de saturação do ácido L-glutâmico em um meio líquido quando é cultivado sob condições acídicas.

5 Além disso, de modo a melhorar a capacidade de produzir ácido L-glutâmico das bactérias de *Enterobacteriaceae*, o método de eliminar o gene *arcA* (Patente U.S. nº 7.090.998), e o método de amplificar o gene *yhfK*, que é um gene de secreção do ácido glutâmico (WO 2005/085419 panfleto), também pode ser usado.

10 O método acima mencionado de intensificar ou eliminar a atividade enzimática, é de forma semelhante aplicável a outros aminoácidos produtores das bactérias, descritos abaixo.

BACTÉRIAS PRODUTORES DA L-TREONINA

Exemplos de cepas precursoras para originar as bactérias produtoras da L-treonina da presente invenção incluem, porém sem limitar, as cepas pertencentes ao gênero *Escherichia*, tais como *E. coli* TDH-6/pVIC40 (VKPM B-3996) (Patentes U.S. nºs 5.175.107, Patentes U.S. nºs 5.705.371), *E. coli* 472T23/pYN7 (ATCC 98081) (Patente U.S. nº 5.631.157), *E. coli* NRRL-21593 (Patente U.S. nº 5.939.307), *E. coli* FERM BP-3756 (Patente U.S. nº 5.474.918), *E. coli* FERM BP-3519 e FERM BP-3520 (Patente U.S. no 5.376.538), *E. coli* MG442 [Gusyatiner *et al.*, *Genetika* (em Russo), 14, 947-956 (1978)], *E. coli* VL643 e VL2055 (Patente Européia Aberta ao Público nº 1149911), e outras.

25 A cepa TDH-6 é deficiente no gene *thrC*, bem como assimilativa a sacarose, e o seu gene *ilvA* tem uma mutação permeável. Esta cepa também tem uma mutação no gene *rhtA*, que comunica resistência às altas concentrações de treonina ou de homoserina. A cepa B-3996 contém o plasmídeo pVIC40, que foi obtido pela inserção de um óperon *thrA*BC* que inclui um gene mutante *thrA* em um vetor originado de RSF1010. Este gene

mutante *thrA* codifica a homoserina aspartocinase desidrogenase I que substancialmente dessensibilizou a inibição da retroalimentação pela treonina. A cepa B-3996 foi depositada em 19 de novembro de 1987 no All-Union Scientific Center of Antibiotics (Nagatinskaya Street 3-A, 117105, Moscou, Federação Russa) sob o número de acesso RIA 1867. A cepa também foi depositada na Russian National Collection of Industrial Microorganisms (VKPM) (Rússia, 117545 Moscou, 1 Dorozhny proezd. 1) em 7 de abril de 1987, sob o número de acesso VKPM B-3996.

A *E. coli* VKPM B-5318 (Publicação da Patente Européia nº 0593792) também pode ser usada como uma cepa precursora para originar as bactérias produtoras da L-treonina da presente invenção. A cepa B-5318 é prototrófica com respeito à isoleucina, e um repressor de λ -fago C1 sensível à temperatura e o promotor PR substituem a região reguladora do óperon da treonina no plasmídeo pVIC40. A cepa VKPM B-5318 foi depositada na Russian National Collection of Industrial Microorganisms (VKPM) em 3 de maio de 1990, sob o número de acesso VKPM B-5318.

Preferivelmente a bactéria da presente invenção é adicionalmente modificada para intensificar a expressão de um ou mais dos seguintes genes:

- gene mutante *thrA* que codifica a homoserina aspartocinase desidrogenase I resistente à inibição da retroalimentação pela treonina;
- gene *thrB* que codifica a homoserina cinase;
- o gene *thrC* que codifica a treonina sintase;
- o gene *rhtA* que codifica uma proteína putativa da transmembrana;
- o gene *asd* gene que codifica a aspartato- β -semialdeído desidrogenase; e
- o gene *aspC* que codifica a aspartato aminotransferase (aspartato transaminase).

O gene *thrA* que codifica a homoserina aspartocinase desidrogenase I de *Escherichia coli* foi elucidado (posições dos nucleotídeos 337 a 2799, acesso do GenBank NC_000913.2, gi: 49175990). O gene *thrA* se localiza entre os genes *thrL* e *thrB* no cromossoma de *E. coli* K-12. O gene *thrB* que codifica a homoserina cinase de *Escherichia coli* foi esclarecido (posições dos nucleotídeos 2801 a 3733, acesso do GenBank NC_000913.2, gi: 49175990). O gene *thrB* se localiza entre os genes *thrA* e *thrC* no cromossoma de *E. coli* K-12. O gene *thrC* que codifica a treonina sintase de *Escherichia coli* foi esclarecido (posições dos nucleotídeos 3734 a 5020, acesso do GenBank NC_000913.2, gi: 49175990). O gene *thrC* se localiza entre o gene *thrB* e a matriz de leitura aberta *yaaX* no cromossoma de *E. coli* K-12. Todos os três genes funcionam como um óperon de treonina único. Para intensificar a expressão do óperon da treonina, a região atenuadora que afeta a transcrição é desejavelmente removida do óperon (WO 2005/049808, WO 2003/097839).

Um gene mutante *thrA*, que codifica a homoserina aspartocinase desidrogenase um resistente à inibição da retroalimentação pela treonina, bem como os genes *thrB* e *thrC*, podem ser obtidos como um óperon do bem conhecido plasmídeo pVIC40 que se acha presente na cepa VKPM B-3996 de *E. coli* produtora da treonina. O plasmídeo pVIC40 é descrito em detalhes na Patente U.S. nº 5.705.371.

O gene *rhtA* existe em 18 min sobre o cromossoma de *E. coli* perto do óperon *glnHPQ*, o qual codifica os componentes do sistema de transporte da glutamina. O gene *rhtA* é idêntico ao ORF1 (gene *ybiF*, posições dos nucleotídeos 764 a 1651, número de acesso do GenBank AAA218541, gi:440181) e se localiza entre os genes *pexB* e *ompX*. A unidade expressando uma proteína codificada pelo ORF1 foi designada como o gene *rhtA* (*rht*: resistência às homoserina e treonina). Igualmente, foi revelado que a mutação de *rhtA23* é uma substituição G-por-A na posição -1 em relação ao

códon de partida ATG (EXTRATOS do 17^o Congresso Internacional de Bioquímica e Biologia Molecular em combinação com o Encontro Anual da American Society for Biochemistry and Molecular Biology, São Francisco, Califórnia, em 24 a 29 de agosto de 1997, extrato nº 457, Patente Européia Aberta ao Público nº 1013765).

O gene *asd* de *E. coli* já foi esclarecido (posições dos nucleotídeos 3572511 a 3571408, acesso do GenBank NC_000913.1, gi:16131307), e pode ser obtido por PCR [reação em cadeia da polimerase; referir-se a White, T. J. *et al.*, *Trends Genet.*, 5, 185 (1989)] utilizando os iniciadores preparados com base na sequência de nucleotídeos do gene. Os genes *asd* de outros microorganismos podem ser obtidos de uma maneira semelhante.

Igualmente, o gene *aspC* de *E. coli* já foi elucidado (posições dos nucleotídeos 983742 a 984932, acesso do GenBank NC_000913.1, gi:16128895), e pode ser obtido por PCR. Os genes *aspC* de outros microorganismos podem ser obtidos de uma maneira semelhante.

BACTÉRIAS PRODUTORAS DA L-LISINA

Exemplos de bactérias produtoras da L-lisina, pertencentes ao gênero *Escherichia*, incluem mutantes tendo resistência a um análogo da L-lisina. O análogo da L-lisina inibe o crescimento de bactérias pertencentes ao gênero *Escherichia*, porém esta inibição é completa ou parcialmente dessensibilizada quando a L-lisina coexiste em um meio. Exemplos do análogo da L-lisina incluem, porém sem limitar, a oxalisina, o hidroxamato de lisina, a S-(2-aminoetil)-L-cisteína (AEC), a γ -metil-lisina, a α -clorocaprolactama, e assim por diante. Os mutantes que têm resistência a estes análogos da lisina podem ser obtidos submetendo-se as bactérias pertencentes ao gênero *Escherichia* a um tratamento de mutagênese artificial convencional. Exemplos específicos de cepas bacterianas úteis para produzir a L-lisina incluem a *Escherichia coli* AJ11442 (FERM BP-1543, NRRL B-

12185; ver a Patente U.S. nº 4.346.170) e *Escherichia coli* VL611. Nestes microorganismos, a inibição de retroalimentação da aspartocinase pela L-lisina é dessensibilizada.

A cepa WC196 pode ser usada como uma bactéria de *Escherichia coli* produtora da L-lisina. Esta cepa bacteriana foi reproduzida conferindo-se resistência de AEC à cepa W3110, a qual foi derivada de *Escherichia coli* K-12. A cepa resultante foi designada de cepa de *Escherichia coli* AJ13069 e foi depositada no National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (correntemente National Institute of Advanced Industrial Science and Technology, International Patent Organism Depositary, Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japão) em 6 de dezembro de 1994, e recebeu um número de acesso de FERM P-14690. Posteriormente, o depósito foi transformado em um depósito internacional sob as disposições do Tratado de Budapeste, em 29 de setembro de 1995, e recebeu um número de acesso FERM BP-5252 (Patente U.S. nº 5.827.698)

Exemplos de cepas precursoras para originar bactérias produtoras de L-lisina da presente invenção também incluem cepas em que a expressão de um ou mais genes codificando uma enzima biossintética de L-lisina é intensificada. Exemplos de tais genes incluem, porém sem limitar, os genes codificando a diidrodipicolinato sintase (*dapA*), a aspartocinase (*lysC*), a diidrodipicolinato redutase (*dapB*), diaminopimelato descarboxilase (*lysA*), diaminopimelato desidrogenase (*ddh*) (Patente U.S. no 6.040.160), a fosfoenolpirvato carboxilase (*ppc*), aspartato semialdeído desidrogenase (*asd*), e aspartase (*aspA*) (Patente Européia Aberta ao Público nº 1253195). Além disso, as cepas precursoras podem ter um nível aumentado de expressão do gene envolvido na eficiência de energia (*cyo*) (Patente Européia Aberta ao Público nº 1170376), do gene codificando a nucleotídeo nicotinamida transidrogenase (*pntAB*) (Patente U.S. nº 5.830.716), o gene *ybjE* (WO

2005/073390), ou combinações destes.

Exemplos de cepas precursoras para dar origem a bactérias produtoras de L-lisina da presente invenção também incluem as cepas com atividade reduzida ou eliminada de uma enzima que catalise uma reação para
5 gerar um composto outro que não a L-lisina pela ramificação da via biossintética da L-lisina. Exemplos das enzimas que catalisam uma reação para gerar um composto outro que não a L-lisina mediante ramificação da via biossintética da L-lisina, incluem a homoserina desidrogenase, a lisina descarboxilase (Patente U.S. nº 5.827.698), e a enzima málica (WO
10 2005/010175).

BACTÉRIAS PRODUTORAS DA L-CISTEÍNA

Exemplos de cepas precursoras para originar bactérias produtoras da L-cisteína da presente invenção incluem, porém sem limitar, as cepas pertencentes ao gênero *Escherichia*, tal como *E. coli* JM15 que é
15 transformado com diferentes alelos de *cysE* codificando as serina acetiltransferases resistentes à realimentação (Patente U.S. nº 6.218.168, Pedido de Patente Russa nº 2003121601); *E. coli* W3110 tendo genes superexpressos que codificam proteínas adequadas para secretar substâncias tóxicas para as células (Patente U.S. nº 5.972.663); cepas de *E. coli* tendo
20 atividade de cisteína dessulfidrase reduzida (JP 11155571A2); *E. coli* W3110 com atividade aumentada de um regulador transcricional positivo para regular a cisteína codificada pelo gene *cysB* (WO 0127307A1), etc.

BACTÉRIAS PRODUTORAS DA L-LEUCINA

Exemplos de cepas precursoras para originar bactérias produtoras da L-leucina da presente invenção incluem, sem limitar, as cepas
25 pertencentes ao gênero *Escherichia*, tais como as cepas de *E. coli* resistentes à leucina [por exemplo, a cepa 57 (VKPM B-7386, Patente U.S. nº 6.124.121)] ou análogos de leucina incluindo β -2-tienilalanina, 3-hidroxileucina, 4-azaleucina, 5,5,5-trifluoroleucina (Publicação da Patente Japonesa nº 62-

34397 e a Patente Japonesa Aberta ao Público nº 8-70879); as cepas de *E. coli* obtidas pelo método de engenharia genética descrito na WO 96/06926; *E. coli* H-9068 (Patente Japonesa Aberta ao Público nº 8-70879), e outras.

A bactéria da presente invenção pode ser melhorada mediante a intensificação da expressão de um ou mais genes envolvidos na biossíntese da L-leucina. Exemplos de tais genes incluem os genes do óperon *leuABCD*, do qual o exemplo típico preferido é um gene mutante *leuA* codificando a isopropilmalato sintase dessensibilizada para a inibição da retroalimentação pela L-leucina (Patente U.S. nº 6.403.342). Além disso, a bactéria da presente invenção pode ser melhorada pela intensificação da expressão de um ou mais genes codificando as proteínas que excretam o L-aminoácido da célula bacteriana. Exemplos de tais genes incluem os genes *B2682* e *b2683* (genes *ygaZH*) (Patente Européia Aberta ao Público nº 1239041 A2).

BACTÉRIAS PRODUTORAS DA L-HISTIDINA

Exemplos de cepas precursoras para originar bactérias produtoras da L-histidina da presente invenção incluem, sem limitar, as cepas pertencentes ao gênero *Escherichia*, tais como a cepa de *E. coli* 24 (VKPM B-5945, RU2003677); a cepa de *E. coli* 80 (VKPM B-7270, RU2119536); *E. coli* NRRL B-12116 a B12121 (Patente U.S. nº 4.388.405); *E. coli* H-9342 (FERM BP-6675) e H-9343 (FERM BP-6676) (Patente U.S. nº 6.344.347); *E. coli* H-9341 (FERM BP-6674) (Patente Européia nº 1085087); *E. coli* AI80/pFM201 (Patente U.S. nº 6.258.554) etc.

Exemplos de cepas precursoras para originar bactérias produtoras da L-histidina da presente invenção também incluem as cepas em que a expressão de um ou mais genes codificando uma enzima biossintética de L-histidina é intensificada. Exemplos de tais genes incluem os genes que codificam ATP fosforribosiltransferase (*hisG*), fosforribosil AMP cicloidrolase (*hisI*), fosforribosil-ATP pirofosfoidrolase (*hisIE*), fosforribosil-formimino-5-aminoimidazol carboxamida ribotídeo isomerase (*hisA*), amido-

transferase (*hisH*), histidinol fosfato aminotransferase (*hisC*), histidinol fosfatase (*hisB*), histidinol desidrogenase (*hisD*), e assim por diante.

É conhecido o fato de que enzimas biossintéticas de L-histidina codificadas por *hisG* e *hisBHAFI* são inibidas pela L-histidina e, portanto, uma capacidade de produzir L-histidina pode também ser eficientemente intensificada pela introdução de uma mutação conferindo resistência à inibição da retroalimentação no gene de fosforribosiltransferase ATP (*hisG*) (Patentes Russas nºs 2003677 e 2119536).

Exemplos específicos das cepas que tenham uma capacidade de produzir L-histidina incluem *E. coli* FERM P-5038 e 5048 que tenham sido introduzidas com um vetor carregando um DNA codificando uma enzima biossintética de L-histidina (Patente Japonesa Aberta ao Público nº 56-005099), as cepas de *E. coli* introduzidas com *rht*, um gene para uma exportação de aminoácido (Patente Européia Aberta ao Público nº 1016710), a cepa de 80 doada com sulfaguanidina, DL-1,2,4-triazol-3-alanina, e resistência à estreptomicina (VKPM B-7270, Patente Russa nº 2119536), e assim por diante.

BACTÉRIAS PRODUTORAS DE L-FENILALANINA

Exemplos de cepas precursoras para dar origem às bactérias produtoras da L-fenilalanina da presente invenção incluem, porém sem limitar, as cepas pertencentes ao gênero *Escherichia*, tais como a *E. coli* AJ12739 (*tyrA::Tn10*, *tyrR*) (VKPM B-8197); *E. coli* HW1089 (ATCC 55371) abrigando um gene mutante *pheA34* (Patente U.S. nº 5.354.672); *E. coli* MWEC101-b (KR8903681); *E. coli* NRRL B-12141, NRRL B-12145, NRRL B-12146 e NRRL B-12147 (Patente U.S. nº 4.407.952). Igualmente, como uma cepa precursora, *E. coli* K12 [W3110 (*tyrA*)/pPHAD] (FERM BP-12659), *E. coli* K-12 [W3110 (*tyrA*)/pBR-aroG4, pACMAB] denominada como AJ12604 (FERM BP-3579), podem ser usadas (Publicação da Patente Européia nº 488424 B1). Além disso, as bactérias produtoras da L-

fenilalanina pertencentes ao gênero *Escherichia* com uma atividade intensificada da proteína codificada pelo gene *yedA* ou pelo gene *yddG*, podem também ser usadas (Publicações dos Pedidos de Patentes U.S. n^{os} 2003/0148473 A1 e 2003/0157667 A1, respectivamente).

5 BACTÉRIAS PRODUTORAS DO L-TRIPTOFANO

Exemplos de cepas precursoras para dar origem às bactérias produtoras do L-triptofano da presente invenção incluem, porém sem limitar, as cepas pertencentes ao gênero *Escherichia*, tais como as *E. coli* JP4735/pMU3028 (DSM10122) e JP6015/pMU91 (DSM10123) deficientes na triptofanil-tRNA sintetase codificada pelo gene mutante *trpS* (Patente U.S. n^o 5.756.345); *E. coli* SV164 (pGH5) tendo um alelo *serA* codificando a fosfoglicerato desidrogenase livre da inibição da retroalimentação pela serina e um alelo *trpE* codificando a antranilato sintase livre da inibição da retroalimentação pelo triptofano (Patente U.S. n^o 6.180.373); *E. coli* AGX17 (pGX44) (NRRL B-12263) e AGX6(pGX50)aroP (NRRL B-12264) deficiente na enzima triptofanase (Patente U.S. n^o 4.371.614); *E. coli* AGX17/pGX50, pACKG4-pps em que a capacidade de produzir fosfoenolpiruvato é intensificada (WO 97/08333, Patente U.S. n^o 6.319.696), etc. As bactérias produtoras do L-triptofano pertencentes ao gênero

10

15

20

Escherichia no qual a atividade da proteína codificada pelo gene *yedA* ou pelo gene *yddG* é aumentada, podem também ser usadas (Publicações dos Pedidos de Patentes U.S. n^{os} 2003/0148473 A1 e 2003/0157667 A1).

Exemplos de cepas precursoras para dar origem às bactérias produtoras do L-triptofano da presente invenção também incluem as cepas em

25

que uma ou mais atividades das enzimas selecionadas da antranilato sintase (*trpE*), da fosfoglicerato desidrogenase (*serA*), e do triptofano sintase (*trpAB*), são intensificadas. A antranilato sintase e a fosfoglicerato desidrogenase acham-se, ambas, sujeitas à inibição da retroalimentação pelo L-triptofano e pela L-serina e, portanto, uma mutação dessensibilizando a

inibição da retroalimentação pode ser introduzida nestas enzimas. Exemplos específicos das cepas tendo uma tal mutação incluem uma *E. coli* SV164 que abriga a antranilato sintase dessensibilizada e uma cepa transformante obtida pela introdução, na *E. coli* SV164, do plasmídeo pGH5 (WO 94/08031), que contém um gene *serA* mutante codificando a fosfoglicerato desidrogenase dessensibilizada pela inibição da retroalimentação.

Exemplos de cepas precursoras para originar as bactérias produtoras do L-triptofano da presente invenção também incluem as cepas dentro das quais o óperon de triptofano contendo um gene que codifica a antranilato sintase dessensibilizada, tenha sido introduzido (Patente Japonesa Aberta ao Público nºs 57-71397, 62-244382, Patente U.S. nº 4.371.614). Além disso, a capacidade produtora do L-triptofano pode ser transmitida pela intensificação da expressão de um gene que codifique a triptofano sintase, entre os óperons de triptofano (*trpBA*). A triptofano sintase consiste das subunidades α e β que são codificadas pelos genes *trpA* e *trpB*, respectivamente. Além disso, a capacidade produtora do L-triptofano pode ser melhorada pelo aumento da expressão do óperon da isocitrato liase-malato sintase (WO 2005/103275).

BACTÉRIAS PRODUTORAS DA L-PROLINA

Exemplos de cepas precursoras para dar origem às bactérias produtoras de L-prolina da presente invenção incluem, porém sem limitar, as cepas pertencentes ao gênero *Escherichia*, tal como *E. coli* 702ilvA (VKPM B-8012) que é deficiente no gene *ilvA* e é capaz de produzir L-prolina (Patente Européia nº 1172433).

A bactéria da presente invenção pode ser melhorada pela intensificação da expressão de um ou mais genes envolvidos na biossíntese da L-prolina. Exemplos de tais genes para as bactérias produtoras da L-prolina que são preferidos, incluem o gene *proB* codificando a glutamato cinase, da qual a inibição da retroalimentação pela L-prolina é dessensibilizada (Patente

Alemã nº 3127361). Além disso, a bactéria da presente invenção pode ser melhorada pela intensificação da expressão de um ou mais genes codificando proteínas que excretam L-aminoácido da célula bacteriana. Exemplos de tais genes são os genes *b2682* e *b2683* (genes *ygaZH*) (Patente Européia Aberta ao Público nº 1239041 A2).

Exemplos de bactérias pertencentes ao gênero *Escherichia*, que tenham uma atividade para produzir L-prolina, incluem as seguintes cepas de *E. coli*: NRRL B-12403 e NRRL B-12404 (Patente Britânica nº 2075056), VKPM B-8012 (Pedido de Patente Russa nº 2000124295), 10 mutantes plasmídeos descritos na Patente Alemã nº 3127361, plasmídeos mutantes descritos por Bloom F. R. *et al.* (15º simpósio de inverno de Miami, 1983, p. 34), e outras.

BACTÉRIAS PRODUTORAS DA L-ARGININA

Exemplos de cepas precursoras para dar origem às bactérias produtoras da L-arginina da presente invenção incluem, porém sem limitar, as 15 cepas pertencentes ao gênero *Escherichia*, tais como a cepa 237 de *E. coli* (VKPM B-7925) (Publicação do Pedido de Patente U.S. nº 2002/058315 A1) e suas cepas derivadas que abriguem a N-acetilglutamato sintase mutante (Pedido de Patente Russa nº 2001112869), a cepa 382 de *E. coli* (VKPM B-20 7926) (Patente Européia Aberta ao Público nº 1170358 A1), uma cepa produtora da arginina em que o gene *argA* codificando a N-acetilglutamato sintetase é introduzido (Patente Européia Aberta ao Público nº 1170361 A1), e outras

Exemplos de cepas precursoras para originar as bactérias produtoras da L-arginina da presente invenção também incluem as cepas em 25 que a expressão de um ou mais genes codificando uma enzima biossintética da L-arginina são intensificados. Exemplos de tais genes incluem opcionalmente substituído genes que codificam a N-acetilglutamil fosfato redutase (*argC*), a ornitina acetil transferase (*argJ*), a N-acetilglutamato

cinase (*argB*), a acetilornitina transaminase (*argD*), a ornitina carbamoil transferase (*argF*), a ácido argininossuccínico sintetase (*argG*), a ácido argininossuccínico liase (*argH*), e a carbamoil fosfato sintetase (*carAB*).

BACTÉRIAS PRODUTORAS DA L-VALINA

5 Exemplos de cepas precursoras para dar origem a bactérias produtoras da L-valina da presente invenção incluem, sem limitar, as cepas que tenham sido modificadas para superexpressar o óperon *ilvGMEDA* (Patente U.S. nº 5.998.178). É desejável remover a região do óperon *ilvGMEDA*, o que é necessário para a atenuação de modo que a expressão do
10 óperon não seja atenuada pela L-valina que é produzida. Além disso, o gene *ilvA* no óperon é desejavelmente rompido de modo que a atividade da treonina desaminasa seja reduzida.

Exemplos de cepas precursoras para dar origem às bactérias produtoras da L-valina da presente invenção também incluem as cepas
15 mutantes tendo uma mutação de aminoacil t-RNA sintetase (Patente U.S. nº 5.658.766). Por exemplo, *E. coli* VL1970, que tem uma mutação no gene *ileS* que codifica a isoleucina tRNA sintetase, pode ser usado. *E. coli* VL1970 foi depositado na Russian National Collection of Industrial Microorganisms (VKPM) (Rússia, 113545 Moscou, 1 Dorozhny Proezd, 1) em 24 de junho de
20 1988, sob o número de acesso VKPM B-4411.

Além disso, os mutantes requerendo ácido lipóico para o crescimento e/ou carência de H⁺-ATPase, podem também ser usados nas cepas precursoras (WO 96/06926).

BACTÉRIAS PRODUTORAS DA L-ISOLEUCINA

25 Exemplos de cepas precursoras para dar origem às bactérias produtoras da L-isoleucina da presente invenção incluem, porém sem limitar, as cepas mutantes tendo resistência à 6-dimetilaminopurina (Patente Japonesa Aberta ao Público nº 5-304969), cepas mutantes tendo resistência a um análogo da isoleucina, tal como a tiaisoleucina e ao hidroxamato de

isoleucina, e cepas mutantes adicionalmente tendo resistência à DL-etionina e/ou ao hidroxamato de arginina (Patente Japonesa Aberta ao Público nº 5-130882). Além disso, as cepas recombinantes transformadas com um gene codificando uma proteína envolvida na biossíntese da L-isoleucina, tal como a
5 treonina desaminase e a acetoidroxato sintase, podem também ser usadas como cepas precursoras (Patente Japonesa Aberta ao Público nº 2-458, Patente Francesa nº 0356739, e Patente U.S. nº 5.998.178).

<1-2> INTENSIFICAÇÃO DO SISTEMA *kdp*

O microorganismo da presente invenção pode ser obtido pela
10 modificação de um tal microorganismo pertencente à família da *Enterobacteriaceae* e tendo uma capacidade produtora do L-aminoácido como descrito acima, de modo que o sistema *kdp* seja intensificado. Entretanto, após um microorganismo ter sido modificado de modo que o sistema *kdp* seja intensificado, a capacidade produtora do L-aminoácido pode
15 ser transmitida ao microorganismo.

O sistema *kdp* pode ser intensificado por modificação que aumente a expressão do óperon *kdp* ou um ou mais genes que constituam o óperon *kdp*, e tal aumento da expressão pode basear-se na intensificação da expressão de um gene endógeno mediante modificação de uma região de
20 controle da expressão, tal como a modificação de um promotor ou similar, ou intensificação da expressão de um gene exógeno pela introdução de um plasmídeo contendo o óperon ou qualquer dos genes, ou similar. Estes métodos podem ser realizados em combinação. O sistema *kdp* pode também ser intensificado pelo aumento da translação do óperon *kdp* ou de qualquer
25 dos genes que constituem o óperon *kdp*.

Na presente invenção, o “sistema *kdp*” significa uma ATPase do tipo P (ATPase do tipo P transportando potássio) que atua sobre o sistema de transporte de potássio de alta afinidade (EC 3.6.3.12).

A condição de “ser modificado de modo que o sistema *kdp*

seja intensificado” significa uma condição de que o transporte de potássio, acima mencionado, pela ATPase tipo P é intensificado, mais especialmente uma condição de que o microorganismo seja modificado de modo que a sua atividade de ATPase do tipo P seja intensificada. Uma tal condição

5 corresponde, por exemplo, à condição de que o número das moléculas da proteína de ATPase do tipo P por célula seja aumentado em comparação com aquele da cepa precursora ou de uma cepa selvagem, ou a condição de que a atividade da ATPase do tipo P por molécula seja aumentada em comparação com aquela da cepa precursora ou de uma cepa selvagem. A modificação é

10 preferivelmente realizada de modo que a atividade da ATPase do tipo P por célula seja melhorada até 150 % ou mais, preferivelmente 200 % ou mais, mais preferível 300 % ou mais, da atividade da cepa precursora ou de uma cepa selvagem. O microorganismo do tipo selvagem pertencente à família *Enterobacteriaceae* usado como uma referência para a comparação é, por

15 exemplo, o *Escherichia coli* MG1655 (ATCC 47076), *Pantoea ananatis* AJ13355 (FERM BP-6615), ou similar.

O aumento da expressão do óperon *kdp* pode ser confirmado pela comparação da quantidade do seu mRNA com aquela de uma cepa do tipo selvagem ou não modificada. Exemplos do método para confirmar a

20 expressão incluem a hibridização Northern e RT-PCR (*Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, 2001). O grau do aumento na expressão não é particularmente limitado, contanto que ele aumente em comparação com aquele de uma cepa selvagem ou cepa não modificada. Entretanto, ela é desejavelmente aumentada, por exemplo, 1,5

25 vez ou mais, preferivelmente 2 vezes ou mais, mais preferível 3 vezes ou mais, em comparação com aquela de uma cepa selvagem ou cepa não modificada.

A atividade da ATPase do tipo P pode ser medida, por exemplo, pela extração do sistema *kdp* de um microorganismo que o possua,

purificando-o [referir-se a Siebers, A. *et al.*, *Eur. J. Biochem.*, 178, 131 (1988)] e medindo-se a atividade da ATPase do tipo P do sistema *kdp* purificado [referir-se a Arnold, A. *et al.*, *Anal. Biochem.*, 71, 209 (1976)].

O sistema *kdp* consiste de três subunidades codificadas pelo óperon *kdp*, e, no que diz respeito a *E. coli*, as seguintes anotações são fornecidas para os genes da subunidade:

kdpA: ATPase do sistema de transporte de potássio de alta afinidade, cadeia A

kdpB: ATPase do sistema de transporte de potássio de alta afinidade, cadeia B

kdpC: ATPase do tipo P, sistema de transporte de potássio de alta afinidade, cadeia C

O “óperon *kdp*” referido na presente invenção é um aglomerado de genes codificando as subunidades A, B e C da ATPase do tipo P descrita acima, em que a subunidade A é codificada pelo gene *kdpA*, a subunidade B é codificada pelo gene *kdpB*, e a subunidade C é codificada pelo gene *kdpC*. Na presente invenção, o óperon *kdp* pode conter um gene que não os genes *kdpA*, *kdpB* e *kdpC*.

A sequência de nucleotídeos do óperon *kdp* de *Escherichia coli* é mostrada na SEQ ID NO: 1. Este óperon contém os seguintes seis genes, e as regiões codificadoras (incluindo o códon de parada) dos genes na SEQ ID NO: 1 são como segue. As sequências de aminoácidos codificadas por *kdpA*, *kdpB*, *kdpC*, *kdpD* e *kdpE* são mostradas nas SEQ ID NOS: 2 a 6, respectivamente.

kdpF: 457 a 546
kdpA: 546 a 2219
kdpB: 2242 a 4290
kdpC: 4299 a 4871
kdpD: 4864 a 7548

kdpE: 7545 a 8222

A sequência de nucleotídeos do óperon *kdp* de *Pantoea ananatis* é mostrada na SEQ ID NO: 7. Este óperon contém os seguintes quatro genes, e as regiões codificadoras (incluindo o códon de parada) dos genes na SEQ ID NO: 7 são como segue. As sequências de aminoácidos codificadas por *kdpA*, *kdpB*, *kdpC* e *kdpD* são mostradas nas SEQ ID NOS: 8 a 11, respectivamente.

kdpA: 543 a 2225

kdpB: 2228 a 4273

10 *kdpC*: 4284 a 4853

kdpD: 4867 a 7542

Além disso, a sequência de nucleotídeos do gene *kdpE* de *pantoea ananatis* e a sequência de aminoácidos codificada por este gene, são mostradas nas SEQ ID NOS: 12 e 13, respectivamente.

15 Neste relatório descritivo, as proteínas codificadas por *kdpA*, *kdpB*, *kdpC*, *kdpD* e *kdpE* podem ser indicadas como KdpA, KdpB, KdpC, KdpD e KdpE, respectivamente.

Os alinhamentos das sequências de aminoácidos de KdpA, KdpB e KdpC de *Pantoea ananatis* e *Escherichia coli* são mostrados nas Figuras 6 a 8. As sequências de consenso das sequências de *Pantoea ananatis* e *Escherichia coli* são mostradas nas linhas inferiores dos alinhamentos. Além disso, as sequências de consenso de KdpA, KdpB e KdpC são mostradas nas SEQ ID NOS: 57 a 59, respectivamente.

25 As homologias de KdpA, KdpB e KdpC de *Pantoea ananatis* e *Escherichia coli* são de 75,36 %, 81,35 % e 59,57 %, respectivamente.

No que diz respeito às bactérias de *Escherichia*, o gene *kdpA* é registrado no GenBank NP_415226.1 Reports potassium-transpo ... To [gi:16128674], o gene *kdpB* em NP_415225. Reports potassium-transpo ... [gi:16128673], o gene *kdpC* em NP_415224. Reports potassium-transpo ...

[gi:16128672], o gene *kdpD* em NP_415223. Reports fused sensory his ... [gi:16128671], e o gene *kdpE* em NP_415222. Reports DNA-binding respo ... [gi:16128670].

5 Além disso, o óperon *kdp* pode ser um clonado de um microorganismo pertencente à família *Enterobacteriaceae* tal como as bactérias *Escherichia*, *Pantoea*, *Enterobacter*, *Klebsiella*, *Serratia*, *Erwinia* e *Yersinia* à base das homologias aos genes exemplificados acima.

10 Como o óperon *kdp* utilizável para a presente invenção, o óperon *kdp* e as suas regiões de flanco, incluindo uma região de controle da expressão que se localize a montante do óperon, podem ser obtidos por PCR [reação em cadeia da polimerase; referir-se a White, T. J. *et al.*, *Trends Genet.*, 5, 185 (1989)] usando iniciadores preparados com base em uma sequência de nucleotídeos já elucidada de um microorganismo pertencente à família *Enterobacteriaceae*, e DNA cromossômico de um microorganismo
15 pertencente à família *Enterobacteriaceae* como o padrão. Os homólogos do óperon *kdp* de outros microorganismos também podem ser obtidos de uma maneira semelhante.

Um homólogo do óperon *kdp* significa um gene codificando uma ATPase do tipo P, que incorpore íons de potássio, derivada de outro
20 microorganismo e apresentando uma alta homologia ao óperon *kdp* de *Escherichia coli* ou de *Pantoea ananatis*. O gene *kdpA*, o gene *kdpB* e o gene *kdpC* derivados de outro microorganismo significa aqueles que apresentam homologias de 80 % ou mais, preferivelmente de 90 % ou mais, mais preferível de 95 % ou mais, particularmente preferível de 97 % ou mais, às
25 sequências totais de aminoácidos das SEQ ID NOS: 2, 3, 4, 8, 9 e 10, e codificando as subunidades que constituem uma proteína tendo a atividade da ATPase do tipo P.

Cada um dos genes pode codificar variante conservativa tendo as sequências de aminoácidos das SEQ ID NOS: 2, 3, 4, 8, 9 ou 10 incluindo

substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos em uma ou em várias posições, contanto que a atividade da ATPase do tipo P constituída destas subunidades não seja degradada. Não obstante o número denotado pelo termo “várias” possa diferir dependendo da
5 posição na estrutura tridimensional ou dos tipos de resíduos de aminoácidos das proteínas, ele é preferivelmente de 2 a 20, mais preferível de 2 a 10, particularmente preferível de 2 a 5. As substituições, deleções, inserções, adições, inversões etc. dos aminoácidos descritos acima incluem aquelas causadas pelas mutações de ocorrência natural, dependendo das diferenças
10 individuais ou das diferenças nas espécies de microorganismos.

Estas substituições são de preferência substituições conservativas que são mutações neutras não proporcionando nenhuma mudança funcional. Uma mutação conservativa é uma mutação em que a substituição tem lugar mutuamente entre Phe, Trp, Tyr, se o sítio da
15 substituição for um aminoácido aromático; entre Leu, Ile, Val, se o sítio da substituição for um aminoácido hidrofóbico; entre Gln, Asn, se ele for um aminoácido polar; entre Lys, Arg, His, se ele for um aminoácido básico; entre Asp e Glu, se ele for um aminoácido ácido; e entre Ser e Thr, se ele for um aminoácido tendo um grupo hidroxila. Exemplos específicos de substituições
20 conservativas incluem: a substituição de Ala por Ser ou Thr; a substituição de Arg por Gln, His ou Lys; a substituição de Asn por Glu, Gln, Lys, His ou Asp; a substituição de Asp por Asn, Glu ou Gln; a substituição de Cys por Ser ou Ala; a substituição de Gln por Asn, Glu, Lys, His, Asp ou Arg; a substituição de Glu por Gly, Asn, Gln, Lys ou Asp; a substituição de Gly por
25 Pro; a substituição de His por Asn, Lys, Gln, Arg ou Tyr; a substituição de Ile por Leu, Met, Val ou Phe; a substituição de Leu por Ile, Met, Val ou Phe; a substituição de Lys por Asn, Glu, Gln, His ou Arg; a substituição de Met por Ile, Leu, Val ou Phe; a substituição de Phe por Trp, Tyr, Met, Ile ou Leu; a substituição de Ser por Thr ou Ala; a substituição de Thr por Ser ou Ala for

Thr; a substituição de Trp por Phe ou Tyr; a substituição de Tyr por His, Phe ou Trp; e a substituição de Val por Met, Ile ou Leu.

Além disso, o óperon *kdp* com o uso dos códons que possam ser facilmente usados em um microorganismo hospedeiro escolhido, no qual o gene seja introduzido, pode também ser usado, desde que a degeneração do gene varie dependendo do microorganismo hospedeiro. De forma semelhante, contanto que a produção do L-aminoácido possa ser melhorada pela amplificação do óperon *kdp*, o óperon *kdp* pode ser alongado ou encurtado ou no término N e/ou no término C de cada subunidade codificada pelo óperon, mediante, por exemplo, 50 ou menos, preferivelmente 20 ou menos, mais preferível 10 ou menos, particularmente preferível 5 ou menos, da quantidade de resíduos de aminoácidos. Mais especificamente, cada subunidade pode ter uma sequência de aminoácidos que seja encurtada em 5 a 50 resíduos de aminoácidos ou no término N e/ou no término C na sequência de aminoácidos das SEQ ID NOS: 2, 3, 4, 8, 9 ou 10.

Além disso, o óperon *kdp* pode ser um DNA que hibridize sob condições estridentes com a sequência de nucleotídeos apresentada nas SEQ ID NOS: 1 ou 7, ou uma sequência complementar a cada região codificadora na sequência de nucleotídeos das SEQ ID NOS: 1 ou 7, ou uma sonda que possa ser preparada destas sequências, e que codifique o sistema *kdp*, isto é, proteína tendo atividade de ATPase do tipo P para incorporar íons de potássio.

As “condições estridentes” aqui referida significam condições em que um assim chamado híbrido específico seja formado e um híbrido não específico não seja formado. É difícil definir claramente as condições com valores numéricos, mas exemplos destas incluem condições em que os DNAs tendo alta homologia, por exemplo homologia de 70 % ou mais, preferível de 80 % ou mais, mais preferível de 90 % ou mais, ainda mais preferível de 95 % ou mais, particularmente preferível de 97 % ou mais,

hibridizam um com o outro e os DNAs tendo uma homologia menor do que o valor não hibridizam um com o outro; e especificamente incluem condições correspondentes à concentração de sal e à temperatura das condições de lavagem na hibridização Southern típica, por exemplo, 1 x SSC, 0,1 % de SDA, preferivelmente 0,1 x SSC, 0,1 % de SDS, em 60°C.

A sonda pode ser uma sonda tendo uma sequência parcial do óperon *kdp*. Uma tal sonda pode ser preparada por PCR com o uso de oligonucleotídeos preparados com base na sequência de nucleotídeos do gene de acordo com um método bem conhecido de uma pessoa habilitada na técnica, como iniciadores, e um fragmento de DNA contendo o gene como o padrão. Quando o fragmento de DNA de um comprimento de cerca de 300 pares de base seja usado como a sonda, a lavagem após a hibridização sob as condições acima mencionadas pode ser, por exemplo, a lavagem de uma vez ou duas vezes ou três vezes sob as condições de 50°C, 2 x SSC, 0,1 % de SDS.

Um tal gene homólogo ao óperon *kdp* pode ser obtido, por exemplo, pela modificação da região codificadora na sequência de nucleotídeos das SEQ ID NOS: 1 ou 7 pela mutagênese específica do sítio, de modo que a proteína codificada contenha substituições, deleções, inserções ou adições dos resíduos de aminoácidos de um sítio específico. Um tal gene pode também ser obtido pela seguinte mutagênese convencionalmente conhecida. No que diz respeito à mutagênese, um óperon codificando um sistema *kdp* altamente ativo pode ser obtido introduzindo-se artificialmente uma mutação no óperon *kdp* mediante o tratamento das sequências de nucleotídeos das SEQ ID NOS: 1 ou 7, ou uma região codificadora nestas sequências de nucleotídeos *in vitro* com hidroxilamina ou similar, ou tratando-se um microorganismo tendo o gene, por exemplo, um tal microorganismo pertencente à família *Enterobacteriaceae*, com irradiação ultravioleta ou um mutágeno usado para a mutagênese usual, tal como a N-metil-N'-nitro-N-

nitrosoguanidina (NTG) ou metanossulfonato de etila (EMS), ou por recombinação de genes com base na PCR propensa a erro [Cadwell, R. C., *PCR Meth. Appl.*, 2, 28 (1992)], embaralhamento do DNA [Stemmer, W. P., *Nature*, 370, 389 (1994)], ou StEP-PCR [Zhao, H., *Nature Biotechnol.*, 16, 258 (1998)]. O fato de o homólogo do óperon *kdp* codificar a ATPase do tipo P pode ser confirmado, por exemplo, introduzindo-se o gene em um microorganismo pertencente à família *Enterobacteriaceae* e tendo capacidade de produzir L-aminoácido, e determinando-se se a capacidade de produzir L-aminoácido é melhorada ou medindo-se a atividade da ATPase do tipo P pelo método acima mencionado.

As descrições acima concernentes às variantes e homólogos são também aplicadas ao gene *kdpD* e ao gene *kdpE* descrito mais abaixo.

Tal modificação de um microorganismo pertencente à família *Enterobacteriaceae*, de acordo com a qual a expressão do óperon *kdp* ou de um ou mais genes constituindo o óperon é aumentada, pode ser alcançada pelo método acima mencionado de modificar uma bactéria de modo que a expressão de um gene alvo seja intensificada. A saber, pelo aumento do número de cópias do óperon *kdp* ou cada gene constituindo o óperon a ser expresso, e/ou substituindo-se a sequência de controle da expressão do óperon por uma sequência de controle da expressão mais forte, ou pelo controle de cada gene que constitui o óperon por uma sequência de controle da expressão mais forte, a expressão do óperon ou de cada gene pode ser intensificada. Não obstante a intensificação da expressão dos genes que constituem o óperon *kdp* possa ser realizada quanto ao óperon inteiro ou a cada gene, a intensificação é preferivelmente realizada quanto ao óperon inteiro. Quando a expressão é intensificada para cada gene individual, o gene a ser intensificado pode ser qualquer um dos genes que constituem o óperon *kdp*, mas é preferível intensificar a expressão de pelo menos uma ou mais espécies de genes entre os genes *kdpA*, *kdpB* e *kdpC*, mais preferível todos os genes *kdpA*, *kdpB* e

kdpC.

O sistema *kdp* pode também ser intensificado pela modificação de uma sequência de espaçador entre o sítio de ligação ribossômica (RBS) e o códon de partida de cada gene, de modo que a translação de cada gene que constitua o óperon *kdp* seja aumentada.

Além disso, sabe-se que a expressão do óperon *kdp* é controlada pelo sistema de controle binário KdpDE codificado pelo gene *kdpD* e pelo gene *kdpE* (*J. Bacteriol.*, abril de 1992, 174 (7): 2152-2159), e a expressão do óperon *kdp* pode também ser aumentada pelo aumento da expressão do gene *kdpD* e do gene *kdpE*.

<2> MÉTODO PARA PRODUZIR O L-AMINOÁCIDO DA PRESENTE INVENÇÃO

Pelo cultivo do microorganismo da presente invenção em um meio para produzir e acumular um L-aminoácido no meio e coletar o L-aminoácido do meio, o L-aminoácido pode ser produzido.

Como o meio usado para o cultivo, um meio usual contendo uma fonte de carbono, uma fonte de nitrogênio e sais minerais, assim como nutrientes orgânicos de traço como os aminoácidos e as vitaminas, quando necessário, podem ser usados. Qualquer espécie de fonte de carbono e de fonte de nitrogênio pode ser usada, contanto que elas possam ser utilizadas por uma cepa a ser cultivada.

Açúcares tais como glicose, glicerol, frutose, sacarose, maltose, manose, galactose, hidrolisados de amido e melaços podem ser usados como fontes de carbono. Além disso, ácidos orgânicos tais como o ácido acético e o ácido cítrico, e álcoois tais como o etanol, podem também ser usados, cada um isoladamente ou em combinação com outras fontes de carbono. A amônia, os sais de amônio tais como o sulfato de amônio, o carbonato de amônio, o cloreto de amônio, o fosfato de amônio e o acetato de amônio, sais de ácido nítrico, e assim por diante, podem ser usados como a

fonte de nitrogênio. Aminoácidos, vitaminas, ácidos graxos, ácidos nucléico, esses contendo aquelas substâncias tais como peptona, casaminoácido, extrato de levedura e produto da decomposição da proteína de soja, e assim por diante, podem ser usados como os nutrientes orgânicos de traço. Quando é usada uma cepa mutante auxotrófica que requeira um aminoácido ou similar para seu crescimento, é preferível suplementar o nutriente requerido.

Em particular, quando um meio líquido preparado de modo a satisfazer uma condição para precipitar o ácido L-glutâmico é usado, a adição do ácido pantotênico ao meio provê precipitação mais eficiente do ácido L-glutâmico (WO 2004/111258). Como sais inorgânicos, os sais de ácido fosfórico, os sais de magnésio, os sais de cálcio, os sais de ferro, o sal de manganês, e assim por diante, podem ser usados.

A cultura é preferivelmente realizada como cultura aeróbica, enquanto a temperatura de fermentação é controlada para ser de 20 a 45°C, e o pH para ser de 3 a 9. Quando o pH é reduzido durante a cultura, carbonato de cálcio pode ser adicionado, ou a cultura é neutralizada com uma substância alcalina tal como gás amoníaco. O L-aminoácido alvo é acumulado no meio de cultura após preferivelmente 10 a 120 horas de cultura sob condições tais como descrito acima.

Além disso, a cultura pode ser realizada com a precipitação de ácido L-glutâmico em um meio, pelo uso, como o meio, de um meio líquido ajustado para satisfazer a uma condição sob a qual o ácido L-glutâmico é precipitado. Os exemplos da condição sob a qual o ácido L-glutâmico é precipitado incluem o pH de 5,0 a 4,0, preferivelmente de 4,5 a 4,0, mais preferível de 4,3 a 4,0, particularmente preferível de 4,0.

Quando o ácido L-glutâmico é precipitado no meio, a adição preliminar dos cristais de ácido L-glutâmico ou de L-lisina como cristais sementes podem proporcionar cristalização mais eficiente (Patente Européia nº 1233069, Patente Européia Aberta ao Público nº 1624069).

A coleta do L-aminoácido do caldo de cultura após a cultura pode ser realizada por um método de coleta conhecido. Por exemplo, após as células terem sido removidas do meio de cultura, o L-aminoácido pode ser coletado mediante concentração do meio para cristalizar o L-aminoácido, por cromatografia de troca de íons, ou similar. Quando a cultura é realizada sob uma condição sob a qual o ácido L-glutâmico é precipitado, o ácido L-glutâmico precipitado no meio pode ser coletado por centrifugação ou por filtração. Neste caso, o ácido L-glutâmico dissolvendo-se no meio pode ser precipitado e depois separado junto com o ácido L-glutâmico já precipitado.

10 Quando um aminoácido básico é produzido, a produção pode ser realizada por um método no qual a fermentação seja realizada mediante o controle do pH do meio durante a cultura de modo a ser de 6,5 a 9,0, e do pH do meio após a conclusão da cultura de modo a ser de 7,2 a 9,0, e o controle da pressão no tanque de fermentação durante a fermentação para que seja
15 positiva, ou provendo dióxido de carbono ou um gás misto contendo dióxido de carbono ao meio de modo que exista um período em que os íons de bicarbonato e/ou os íons de carbonato estejam presentes em uma quantidade de pelo menos 2 g/litro no meio de cultura durante a cultura, e estes íons de bicarbonato e/ou íons de carbonato sirvam como contra-íons de cátions
20 principalmente consistindo do aminoácido básico, e o aminoácido básico alvo seja coletado (referir-se à Patente Japonesa Aberta ao Público no 2002-065287, Publicação do Pedido de Patente U.S. nº 2002025564).

EXEMPLOS

25 Daqui por diante, a presente invenção será descrita em mais detalhes mediante referência aos exemplos.

EXEMPLO DE REFERÊNCIA 1

Construção da cepa de *Pantoea ananatis* resistente ao produto do gene λ Red

Para amplificar o óperon *kdp* em *Pantoea ananatis*, foi

construída uma cepa receptora que realiza o método denominado “integração conduzida por Red” ou “integração mediada por Red” [*Proc. Natl. Acad. Sci. USA*, 97, 6640-6645 (2000)].

Primeiramente, o novo plasmídeo auxiliar RSF-Red-TER, que expressa os genes *gam*, *bet* e *exo* de λ (doravante referidos como “genes λ Red”) foi construído (Figura 1). Os detalhes destes serão descritos no Exemplo de Referência 2.

Este plasmídeo pode ser usado em uma ampla faixa de hospedeiros tendo diferentes fundamentos genéticos. Isto é porque 1) este plasmídeo tem o réplicon do plasmídeo RSF1010 de amplo espectro de hospedeiros (Scholz *et al.*, 1989; Buchanan-Wollaston *et al.*, 1987), o qual pode ser estavelmente mantido por muitos tipos de bactérias Gram-negativas e Gram-positivas, e ainda células de plantas, 2) os genes λ Red, *gam*, *bet* e *exo*, se acham sob o controle do promotor PlacUV5, o qual é reconhecido pelas RNA polimerases de muitos tipos de bactérias [por exemplo, Brunschwig, E. e Darzins, A., *Gene*, 111, 1, 35-41 (1992); Dehio, M. *et al*, *Gene*, 215, 2, 223-229 (1998)], e 3) o fator de autorregulação P_{lacUV5}-lacI e o terminador da transcrição não dependente de ρ (TrnB) do óperon *rrnB* de *Escherichia coli* inferior ao nível de expressão basal dos genes λ Red [Skorokhodova, A. Yu *et al*, *Biotekhnologiya (Rus.)*, 5, 3-21 (2004)]. Além disso, o plasmídeo RSF-Red-TER contém o gene levansacarase e, mediante o uso deste gene, o plasmídeo pode ser coletado das células em um meio contendo sacarose.

Em *Escherichia coli*, a frequência de integração de um fragmento de DNA gerado por PCR junto com a região de flaqueio curta provida pelo plasmídeo RSF-Red-TER, é tão elevada quanto a frequência obtível com o uso do plasmídeo auxiliar pKD46 [Datsenko, K. A., Wanner, B. L., *Proc. Natl. Acad. Sci. USA*, 97, 6640-6645 (2000)]. Entretanto, a expressão dos genes λ Red é tóxica à *Pantoea ananatis*. As células

transformadas com o plasmídeo auxiliar RSF-Red-TER crescem de forma extremamente lenta no meio LB contendo IPTG (isopropil- β -D-tiogalactopiranosídeo, 1 mM) e um antibiótico apropriado (25 μ g/ml de cloranfenicol ou 40 μ g/ml de canamicina), e a eficiência da recombinação mediada por λ Red é extremamente lenta (10^{-8}), se observada sob qualquer condição.

Uma cepa variante de *Pantoea ananatis*, que seja resistente à expressão de todos os três genes λ Red, foi selecionada. Para este fim, o plasmídeo RSF-Red-TER foi introduzido na cepa SC17 de *Pantoea ananatis* (Patente U.S. nº 6.596.517) por eletroporação. Após uma cultura de 18 horas, cerca de 10^6 transformantes foram obtidos e, entre estes, 10 clones formaram colônias de grande dimensão, e todo o remanescente formaram colônias extremamente pequenas. Após uma cultura de 18 horas, as grandes colônias tinham um tamanho de cerca de 0,2 mm, ao passo que as pequenas colônias não cresceram nada, mesmo que a cultura fosse estendida por outras 24 horas, as grandes colônias continuaram a crescer. Uma das cepas mutantes de *Pantoea ananatis* das grandes colônias e resistentes à expressão de todos os três genes λ Red (*gam*, *bet* e *exo*), foi usada para a outra análise.

O DNA plasmídeo RSF-Red-TER foi isolado de um clone dos clones das grandes colônias, e de vários clones das colônias pequenas, e transformado novamente no *Escherichia coli* MG1655 para se examinar a capacidade do plasmídeo para sintetizar um produto ativo do gene Red. Por uma experiência de controle para integração dependente de Red nos transformantes obtidos, foi demonstrado que apenas o plasmídeo isolado do clone de grandes colônias induziu a expressão dos genes λ Red requeridos para a integração dependente de Red. De modo a pesquisar se a integração mediada por Red ocorrera no clone selecionado de grandes colônias, a eletroporação foi realizada com o uso de um fragmento linear de DNA produzido por PCR. Este fragmento foi projetado de modo contivesse um

marcador Km^R e uma região de flanco de 40 pares de base homóloga ao gene *hisD*. Este fragmento é integrado no gene *hisD* de *Pantoea ananatis* no sítio de reconhecimento *SmaI*. Dois clones de colônias pequenas foram usados como controle. A sequência de nucleotídeos do gene *hisD* de *Pantoea ananatis* é mostrada na SEQ ID NO: 14. Para a PCR, os oligonucleotídeos das SEQ ID NOS: 15 e 16 foram usados como iniciadores, e o plasmídeo pMW118-(λ att- Km^r - λ att) foi usado como o padrão. Os dois clones de colônias pequenas que não foram resistentes aos genes λ Red foram usados como um controle. A construção do plasmídeo pMW118-(λ attL- Km^r - λ attR) será explicada em detalhes no Exemplo de Referência 3.

O plasmídeo RSF-Red-TER pode induzir a expressão dos genes Red pelo gene *lacI* exercida no plasmídeo. Duas espécies de condições de indução foram pesquisadas. No primeiro grupo, IPTG (1 mM) foi adicionado 1 hora antes da eletroporação, e no segundo grupo, IPTG foi adicionado no início da cultura para a preparação das células das quais a eletroporação é possível. O índice de crescimento das células abrigando o RSF-Red-TER derivado do clone das grandes colônias, não foi significativamente inferior àquele de uma cepa não tendo o plasmídeo SC17. A adição de IPTG apenas levemente reduziu o índice de crescimento destas culturas. Por outro lado, a progênie dos clones das pequenas colônias cresceu de forma extremamente lenta, mesmo sem a adição de IPTG, e, após a indução, o crescimento foi substancialmente interrompido. Após a eletroporação das células da progênie do clone de grandes colônias, muitos clones Km^R cresceram (18 clones após um curto tempo de indução, e cerca de 100 clones após um período de indução prolongado). Todos os 100 clones que foram pesquisados tinham um fenótipo His^- , e cerca de 20 clones foram confirmados por PCR como tendo a estrutura de cromossoma esperada nas células. Por outro lado, mesmo quando a eletroporação foi realizada com a progênie dos clones de pequenas colônias, uma cepa integrada não foi obtida.

O clone de grandes colônias obtido foi cultivado sobre uma placa contendo 7 % de sacarose para eliminar o plasmídeo, e transformado novamente com RSF-Red-TER. A cepa sem o plasmídeo foi designada de SC17 (0). Esta cepa foi depositada na Russian National Collection of Industrial Microorganisms [VKPM, GNII *Genetica* (1 Dorozhny proezd., 1 Moscou 117545, Rússia) em 21 de setembro de 2005, e recebeu um número de acesso de VKPM B-9246.

Todos os clones que cresceram após a retransformação acima mencionada apresentaram grandes tamanhos de colônia como o clone SC17 (0) da cepa precursora. A experiência de integração mediada por Red foi realizada na cepa SC17 (0) retransformada com o plasmídeo RSF-Red-TER. Três dos transformantes independentes foram pesquisados com o uso do mesmo fragmento de DNA como aquele usado para a experiência anterior. O tempo curto de indução (1 hora antes da eletroporação) foi empregado. Os clones Km^R que excederam dez clones, cresceram em cada experiência. Todos os clones examinados. Todos os clones examinados tinham o fenótipo His^r. Desta maneira, uma cepa mutante designada SC17 (0) resistente à expressão dos genes λ Red foi selecionada. Esta cepa pôde ser usada como uma cepa receptora adequada para a integração dependente de Red no cromossoma de *Pantoea ananatis*.

EXEMPLO DE REFERÊNCIA 2

CONSTRUÇÃO DO PLASMÍDEO AUXILIAR RSF-Red-TER

O esquema de construção do plasmídeo auxiliar RSF-Red-TER é mostrado na Figura 2.

Como uma primeira etapa da construção, um vetor RSFsacBPlacMCS foi projetado. Para este fim, os fragmentos de DNA contendo o gene *cat* do plasmídeo pACYC184 e a região de gene estrutural do gene *sacB* de *Bacillus subtilis* foram amplificados por PCR com o uso dos

oligonucleotídeos das SEQ ID NOS: 17 e 18, e 19 e 20, respectivamente. Estes oligonucleotídeos continham os sítios de enzimas de restrição *Bgl*III, *Sac*I, *Xba*I e *Bam*HI, necessários e convenientes para outra clonagem, nas regiões de extremidade 5', respectivamente. O fragmento *sacB* obtido de 1,5 kb foi clonado no vetor pMW119- P_{lac} lacI previamente obtido no sítio *Xba*I-*Bam*HI. Este vetor foi construído da mesma maneira daquele descrito quanto ao vetor pMW118- P_{lac} lacI [Skorokhodova, A. Yu *et al.*, *Biotekhnologiya* (Rus.), 5, 3-21 (2004)]. Entretanto, este vetor continha um componente poliligador derivado de pMW219 ao invés do plasmídeo pMW218.

10 Então, o fragmento *cat* acima mencionado de 1,0 kb foi tratado com *Bgl*III e *Sac*I, e clonado no plasmídeo RSF- P_{lac} lacIsacB obtido na etapa anterior no sítio *Bam*HI-*Sac*I. O plasmídeo obtido pMW- P_{lac} lacIsacBcat continha o fragmento P_{lac} UV5-lacI-sacB-cat. A fim de subclonar este fragmento no vetor RSF1010, pMW- P_{lac} lacIsacBcat foi digerido com *Bgl*III, a
15 extremidade embotada com o fragmento Klenow da DNA polimerase I, e sucessivamente digerido com *Sac*I. Um fragmento *Bgl*III-*Sac*I de 3,8 kb do plasmídeo pMWP $_{lac}$ lacIsacBcat foi eluído de 1 % de gel de agarose, e ligado com o vetor RFS1010 que havia sido tratado com *Pst*I e *Sac*I. *Escherichia coli* TG1 foi transformado com a mistura de ligação, e plaqueado sobre o
20 meio LB contendo cloranfenicol (50 mg/litro). Os plasmídeos isolados dos clones desenvolvidos foram analisados com enzimas de restrição para se obter um plasmídeo RSFsacB. De modo a construir um vetor RSFsacB P_{lac} MCS, um fragmento de DNA contendo o promotor P_{lac} UV5 foi amplificado por PCR com o uso dos oligonucleotídeos de SEQ ID NOS: 21 e 22 como iniciadores e o
25 plasmídeo pMW119- P_{lac} lacI como o padrão. O fragmento obtido de 146 pares de base foi digerido com *Sac*I e *Not*I, e ligado com o fragmento grande *Sac*I-*Not*I do plasmídeo RSFsacB. Depois, por PCR com o uso dos oligonucleotídeos de SEQ ID NOS: 23 e 24 como iniciadores, e o plasmídeo pKD46 [Datsenko, K. A., Wanner, B. L., *Proc. Natl. Acad. Sci. USA*, 97,

6640-6645 (2000)] como padrão, um fragmento de DNA de 2,3 kb contendo os genes λ Red $\alpha\beta\gamma$ e o terminador de transcrição tL3, foi amplificado. O fragmento obtido foi clonado no vetor RSFsacBP_{lac}MCS no sítio PvuI-NotI. Deste modo, o plasmídeo RSFRed foi projetado.

5 De modo a eliminar a leitura através da transcrição dos genes Red, um terminador da transcrição dependente de ρ do óperon *rrnB* de *Escherichia coli* foi introduzido em uma posição entre o gene *cat* e o promotor P_{lacUV5}. Para esta finalidade, um fragmento de DNA contendo o promotor P_{lacUV5} e o terminador TrnB foi amplificado por PCR com o uso
10 dos oligonucleotídeos das SEQ ID NOS: 25 e 22 como iniciadores e o cromossoma de *Escherichia coli* BW3350 como um padrão. Estes fragmentos obtidos foram tratados com KpnI e ligados. Depois, o fragmento de 0,5 kb contendo tanto P_{lacUV5} quanto TrnB foi amplificado por PCR com o uso dos oligonucleotídeos das SEQ ID NOS: 22 e 26 como iniciadores. O fragmento
15 de DNA obtido foi digerido com EcoRI, a extremidade embotada por um tratamento com o fragmento Klenow da DNA polimerase I, digerido com BamHI, e ligado com o grande fragmento Ecl136II-BamHI do vetor RSFsacBPlacMCS. O plasmídeo obtido foi designado RSF-Red-TER.

EXEMPLO DE REFERÊNCIA 3

20 CONSTRUÇÃO DO PLASMÍDEO pMW118-(λ attL-Km^r- λ attR)

O plasmídeo pMW118-(λ attL-Km^r- λ attR) foi construído do plasmídeo pMW118-attL-Tc-attR (WO 2005/010175) mediante substituição do gene marcador de resistência à tetraciclina com o gene de resistência à
25 canamicina do plasmídeo pUC4K. Com essa finalidade, o fragmento grande *EcoRI-HindIII* do plasmídeo pMW118-attL-Tc-attR foi ligado a dois fragmentos do plasmídeo pUC4K: o fragmento *HindIII-PstI* (676 pares de base) e o fragmento *EcoRI-HindIII* (585 pares de base). O pMW118-attL-Tc-attR básico foi obtido pela ligação dos seguintes quatro fragmentos:

1) O fragmento *BglIII-EcoRI* (114 pares de base) incluindo *attL* (SEQ ID NO: 29) que foi obtido por amplificação de PCR da região correspondente ao attL do cromossoma de *Escherichia coli* W3350 (contendo o profago λ) com o uso dos iniciadores P1 e P2 (SEQ ID NOS: 27 e 28) (estes iniciadores continham os sítios de reconhecimento subsidiários para *BglIII* e *EcoRI*).

2) O fragmento *PstI-HindIII* (182 pares de base) incluindo *attR* (SEQ ID NO: 32) que foi obtido por amplificação de PCR da região correspondente ao attR do cromossoma de *Escherichia coli* W3350 (contendo o profago λ) com o uso dos iniciadores P3 e P4 (SEQ ID NOS: 30 e 31) (estes iniciadores continham os sítios de reconhecimento subsidiários para *PstI* e *HindIII*).

3) O grande fragmento *BglIII-HindIII* (3916 pares de base) do pMW118-ter_rrnB. O plasmídeo pMW118-ter_rrnB foi obtido pela ligação dos seguintes três fragmentos de DNA:

- O grande fragmento de DNA (2359 pares de base) incluindo o fragmento *AatII-EcoRI* de pMW118 que foi obtido pela digestão do pMW118 com *EcoRI*, tratamento com o fragmento Klenow de DNA polimerase I, e depois digestão com *AatII*;

- O pequeno fragmento *AatII-BglIII* (1194 pares de base) de pUC19 incluindo o gene *bla* para resistência à ampicilina (Ap^R), que foi obtido por amplificação de PCR da região correspondente do plasmídeo pUC19 com o uso dos iniciadores P5 e P6 (SEQ ID NOS: 33 e 34) (estes iniciadores continham os sítios de reconhecimento subsidiários para *PstI*, *AatII* e *BglIII*);

- O pequeno fragmento *BglIII-PstI*pol (363 pares de base) do terminador da transcrição ter_rrnB, que foi obtido por amplificação de PCR da região correspondente do cromossoma de *Escherichia coli* MG1655 com o uso dos iniciadores P7 e P8 (SEQ ID NOS: 35 e 36) (estes iniciadores

continham os sítios de reconhecimento subsidiários para *Pst*I, *Bgl*II e *Pst*I).

4) O pequeno fragmento *Eco*RI-*Pst*I (1388 pares de base) (SEQ ID NO: 37) de pML-Tc-ter_thrL incluindo o gene de resistência à tetraciclina e o terminador da transcrição ter_thrL; o plasmídeo pML-Tc-ter_thrL foi obtido pelas seguintes duas etapas:

- o plasmídeo pML-ter_thrL foi obtido pela digestão do plasmídeo pML-MCS [Mashko, S. V. *et al.*, *Biotekhnologiya* (em Russo), 2001, nº 5, 3-20) com *Xba*I e *Bam*HI, seguido pela ligação do grande fragmento (3342 pares de base) com o fragmento *Xba*I-*Bam*HI (68 pares de base) carregando o terminador ter_thrL obtido por amplificação de PCR da região correspondente do cromossoma de *Escherichia coli* MG1655 com o uso dos iniciadores P9 e P10 (SEQ ID NOS: 38 e 39) (estes iniciadores continham os sítios de reconhecimento subsidiários *Pst*I, *Xba*I e *Bam*HI);

- o plasmídeo pML-Tc-ter_thrL foi obtido pela digestão do plasmídeo pML-ter_thrL com *Kpn*I e *Xba*I, seguida pelo tratamento com o fragmento Klenow da DNA polimerase I, e ligado com o pequeno fragmento *Eco*RI-*Van*91I (1317 pares de base) de pBR322 incluindo o gene de resistência à tetraciclina (pBR322 foi digerido com *Eco*RI e *Van*91I e depois tratado com o fragmento Klenow da DNA polimerase I).

EXEMPLO 1

AQUISIÇÃO DA CEPA SUBSTITUÍDA PELO PROMOTOR DO ÓPERON *kdp*

(1) CONSTRUÇÃO DO PLASMÍDEO RSFPPG PRODUTOR DO ÁCIDO GLUTÂMICO

Um plasmídeo RSFPPG foi construído, no qual os genes do sistema da biossíntese do ácido L-glutâmico, o gene *prpC* (Publicação da Patente Internacional WO 2006/051660), o gene *ppc* e o gene *gdhA* (EP 0999282A), foram amplificados.

O iniciador 1 (SEQ ID NO: 40) e o iniciador 2 (SEQ ID NO:

41) para amplificar uma parte do RSFCPG (EP 1233068A) outro que não ORF do gene *gltA*, foram projetados. Mediante o uso destes iniciadores e RSFCPG como o padrão, a PCR foi realizada para se obter um fragmento de cerca de 14,9 kb. No que diz respeito ao *prpC*, a PCR foi realizada com o uso do iniciador 3 (SEQ ID NO: 42) e do iniciador 4 (SEQ ID NO: 43) e do DNA cromossômico da cepa W3110 de *E. coli* como o padrão, para se obter um fragmento de cerca de 1,2 kb. Ambos os produtos de PCR foram tratados com *Bgl*III e *Kpn*I, ligados, e depois usados para transformar a cepa JM109 de *E. coli*. Todas as colônias formadas foram coletadas e os plasmídeos foram extraídos das colônias como uma mistura. A cepa ME8330 de *E. coli*, que é uma cepa deficiente de citrato sintase (CS), foi transformada com a mistura de plasmídeos, e a suspensão celular foi aplicada no meio mínimo M9 (contendo 5 g de glicose, sulfato de magnésio 2 mM, 3 g de fosfato monopotássico, 0,5 g de cloreto de sódio, 1 g de cloreto de amônio e 6 g de fosfato dissódico, em 1 litro de água pura) contendo 50 mg/litro de uracila e 5 mg/litro de HCl de tiamina. Das cepas formadas, um plasmídeo foi extraído e designado RSFPPG. Este plasmídeo RSFPPG foi introduzido na cepa NP106 de *Pantoea ananatis*, a qual é uma cepa produtora de ácido L-glutâmico, para construir uma cepa produtora de ácido L-glutâmico, NP106/RSFPPG (esta cepa é referida como a “cepa NA1”).

A cepa NP106 foi obtida como segue. A cepa AJ13601 de *Pantoea ananatis* descrita acima foi cultivada durante a noite em 34°C no meio líquido LBG9 com agitação, e depois o meio foi diluído de modo que 100 a 200 colônias aparecessem por placa e foram aplicadas a uma placa de LBG9 contendo 12,5 mg/litro de tetraciclina. As colônias que apareceram foram replicadas em uma placa de LBG9 contendo 12,5 mg/litro de tetraciclina e 25 mg/litro de cloranfenicol, e uma cepa que foi sensível ao cloranfenicol foi selecionada para se obter uma cepa da qual o pSTVCB fosse eliminado, a qual foi designada G106S. A cepa G106S foi ainda cultivada

durante a noite em 34°C no meio líquido de LBGM9 com agitação, e o meio foi diluído de modo que 100 a 200 colônias aparecessem por placa, e aplicado a uma placa de LBGM9 sem medicamentos. As colônias que apareceram foram replicadas em uma placa LBGM9 contendo 12,5 mg/litro de tetraciclina, e uma placa de LBGM9 sem medicamentos, e uma cepa que resultou sensível à tetraciclina foi selecionada para se obter uma cepa da qual o RSFCPG fosse eliminado, a qual foi designada NP106. A NP106 obtida como descrito acima é uma cepa não contendo nenhum dos dois plasmídeos RSFCPG e pSTVCB, os quais são abrigados pela cepa AJ13601.

10 (2) AQUISIÇÃO DA CEPA NA QUAL O PROMOTOR DO ÓPERON *kdp* FOI SUBSTITUÍDO PELO PROMOTOR *tac*

i) A construção da cepa SC17 (0) de *P. ananatis* na qual a sequência contendo λ attL-Km^r- λ attR e o promotor Ptac ligado a jusante (λ attL-Km^r- λ attR-Ptac) foi integrada a montante do gene *lacZ*

15 O promotor Ptac foi integrado no cromossoma da cepa SC17 (0) de *P. ananatis* em uma posição a montante do gene *lacZ*. A estrutura da região do cromossoma de *P. ananatis* a montante do gene *LacZ* é mostrada na Figura 3. As sequências de nucleotídeos dos genes *yghU*, *scrK* e *lacZ* de *Pantoea ananatis* são apresentadas nas SEQ ID NOS: 44, 45 e 46. A
20 sequência da região -35 do promotor Ptac é ttgaca.

O fragmento do promotor Ptac foi amplificado por PCR com o uso do iniciador 1 de 5' (SEQ ID NO: 47) e iniciador 2 de 3' (SEQ ID NO: 48) correspondente ao promotor Ptac, e o plasmídeo pDR540 (Pharmacia, Suécia) como um padrão. Ambos os iniciadores continham uma sequência de reconhecimento BglII na extremidade 5'. O iniciador 2 continha 46
25 nucleotídeos da parte de extremidade 3' de Ptac, sequência SD, e uma parte de início da região codificadora do gene *lacZ*.

Um fragmento de DNA contendo um gene de resistência Km removível flanqueando os sítios attL e attR de λ foi também amplificado por

PCR com o uso de pMW118-(λ attL-Km^r- λ attR) como um padrão, e o iniciador 3 (SEQ ID NO: 49) e o iniciador 4 (SEQ ID NO: 50). O fragmento obtido tinha um sítio de reconhecimento *Bgl*III para ligação com o fragmento do promotor *tac* em uma extremidade, e um sítio correspondente a uma sequência homóloga ao cromossoma de *Pantoea ananatis* e localizando-se a montante do gene *scrK* para integração no genoma bacteriano na outra extremidade (Figura 3). Dois dos fragmentos do produto de PCR foram tratados com *Bgl*III, e ligados *in vitro* com a DNA T4 ligase.

A mistura de reação da ligação foi usada para a integração dependente de λ no cromossoma de *Pantoea ananatis*. O plasmídeo auxiliar RSF-Red-TER foi usado como um portador dos genes Red de fago λ . De modo a obter-se células eletrocompetentes da *Pantoea ananatis*, a cepa SC17 (0) foi transformada com o plasmídeo RSF-Red-Ter, e cultivada durante a noite em 34°C em meio LB contendo 50 μ g/ml de cloranfenicol. Depois, o caldo de cultura foi diluído 100 vezes com meio LB recente contendo 50 μ g/ml de cloranfenicol, e o crescimento das células foi deixado em 34°C sob aeração até que a OD₆₀₀ se tornasse de 0,3. Então, IPTG 1 mM foi acrescentado e a cultura prosseguiu até que a OD₆₀₀ se tornasse de 0,7. Uma amostra de 10 mM foi lavada por 3 vezes com um volume igual de água deionizada, e as células foram colocadas em suspensão em 40 μ l de glicerol frio a 10 %. Imediatamente antes da eletroporação, 100 a 200 ng de fragmento de DNA amplificado *in vitro* dissolvido em 5 μ l de água deionizada, foram adicionados à suspensão celular. A eletroporação foi feita pelo uso de um aparelho de eletroporação de bactérias (BioRad, Estados Unidos, Catálogo número 165-2089, Versão 2-89). Os parâmetros do pulso usado foram de uma intensidade de campo de 20 kV/cm, e um tempo de pulso de 5 milissegundos.

Após a eletroporação, 1 ml de meio LB suplementado com glicose (0,5 %) foi imediatamente adicionado à suspensão celular. Depois, as células foram deixadas crescer em 34°C por 2 horas sob aeração, plaqueadas

em meio sólido de mLB contendo 40 µg/ml de cloranfenicol, e incubadas durante a noite em 34°C. O integrante Km^R selecionado foi estriado sobre placa de meio LB a qual IPTG (1 mM) e sacarose (5 g/litro) foram adicionados, e cultivado em 34°C para possibilitar a formação de colônias isoladas. De modo a remover o plasmídeo auxiliar RSF-Red-TER do integrante, as variantes Km^R e Cm^S foram isoladas.

As estruturas cromossômicas das colônias selecionadas Km^R e Cm^S foram confirmadas por sequenciação de nucleotídeos.

10 ii) SUBSTITUIÇÃO DO PROMOTOR DO ÓPERON *kdp* PELO PROMOTOR *tac*

Dois iniciadores sintéticos de DNA apresentados nas SEQ ID NOS: 51 e 52 foram sintetizados de uma maneira convencional. O iniciador mostrado na SEQ ID NO: 51 tinha uma estrutura que uma sequência homóloga da região de montante do óperon *kdp* de *Pantoea ananatis* foi seguida por uma sequência homóloga da extremidade 5' de λ attL-Km^f- λ attR-Ptac. O iniciador da SEQ ID NO: 52 tinha uma estrutura que uma sequência complementar de extremidade 5' contendo o primeiro códon de partida do óperon *kdp* de *Pantoea ananatis* foi seguida por uma sequência complementar de extremidade 3' de λ attL-Km^f- λ attR-Ptac. Pela realização da PCR com o uso destes iniciadores e o DNA cromossômico da cepa selecionada em i) como o padrão, um fragmento de 1,6 quilopares de base da sequência λ attL-Km^f- λ attR-Ptac tendo a sequência homóloga da região a montante do óperon *kdp* na extremidade 5' e a sequência homóloga na extremidade 5' contendo o primeiro códon de partida do óperon *kdp* na extremidade 3', foi amplificado.

25 O fragmento de PCR acima mencionado foi purificado e introduzido no SC17(0)/RSF-Red-TER por eletroporação de uma maneira convencional.

A cepa SC17(0)/RSF-Red-TER, na qual o fragmento de PCR foi introduzido, foi selecionada sobre meio L (meio contendo 10 g de

Bactotripton, 5 g de extrato de levedura, 5 g de NaCl, e 15 g de ágar, em 1 litro de água purificada, pH 7,0) contendo 40 mg/litro de canamicina para se obter cerca de 20 colônias como transformantes. A introdução do fragmento acima mencionado na região de montante do óperon *kdp*, foi confirmada por PCR com o uso de dois iniciadores sintéticos de DNA apresentados nas SEQ ID NOS: 53 e 54, e uma cepa para a qual a inserção do fragmento pôde ser confirmada foi designada SC17(0)::Ptac-kdp. O DNA genômico foi extraído desta cepa, e usado para transformar a cepa NA1/pSTV-yhfK por eletroporação. A cepa NA1/pSTV-yhfK foi obtida da cepa AJ13601 (referir-se à Patente Japonesa Aberta ao Público nº 2001-333769) pela eliminação de dois plasmídeos, o RSFCPG e o pSTVCB, e introdução de dois plasmídeos, o plasmídeo para a produção do ácido L-glutâmico, RSFPPG, e pSTV-yhfK (referir-se à Patente Japonesa Aberta ao Público nº 2005-278643).

Ambos os plasmídeos RSFCPG e pSTVCB são apresentados na Patente Japonesa Aberta ao Público no 2001-333769. O RSFCPG é um plasmídeo contendo os genes *gltA*, *ppc* e *gdhA* derivados de *Escherichia coli*. O pSTVCB é um plasmídeo obtido pela introdução do gene *gltA* derivado de *Brevibacterium lactofermentum* no pSTV29 (Takara Shuzo). O pSTV-yhfK é um plasmídeo obtido pela introdução do gene *yhfK* derivado de *Pantoea ananatis* no pSTV29 (Takara Shuzo).

A cepa NA1/pSTV-yhfK, dentro da qual o DNA genômico de SC17(0)::Ptac-kdp foi introduzido, foi selecionada em uma placa do meio L (meio contendo 10 g de Bactotripton, 5 g de extrato de levedura, 5 g de NaCl e 15 g de ágar, em 1 litro de água purificada, pH 7,0) que foi suplementado com ingredientes de meio mínimo (meio contendo 0,5 g de glicose, sulfato de magnésio 2 mM, 3 g de fosfato monopotássico, 0,5 g de cloreto de sódio, 1 g de cloreto de amônio e 6 g de fosfato dissódico, em 1 litro de água purificada), 40 mg/litro de canamicina, 12,5 mg/litro de cloridreto de tetraciclina e 25 mg/litro de cloranfenicol. Como resultado, cerca de 20

colônias foram obtidas como transformantes. Em todas estas cepas, o fragmento de λ attL-Km^r- λ attR-Ptac foi introduzido a montante do óperon *kdp*, e um clone entre elas foi selecionado e designado NA1::Ptac-kdp/pSTV-yhfK.

5 (3) AVALIAÇÃO DA CULTURA DA CEPA SUBSTITUÍDA PELO PROMOTOR DO ÓPERON *kdp* NO TUBO DE TESTE

Em seguida, de modo a examinar-se o efeito da intensificação do óperon *kdp* no crescimento, a cultura foi realizada nos tubos de teste. A cepa NA1:Ptac-kdp/pSTV-yhfK e a cepa NA1/pSTV-yhfK como um controle
10 foram usadas, e o crescimento delas sob uma condição acídica foi examinado.

[Composição do meio para a cultura do tubo de teste]

	D-glicose	0,5 %
	Na ₂ HPO ₄	6,0 g/litro
	KH ₂ PO ₄	3,0 g/litro
15	NaCl	0,5 g/litro
	NH ₄ Cl	1,0 g/litro
	MgSO ₄ ·7H ₂ O	2,0 mM
	Ácido ϵ -diaminopimélico	200 mg/litro
	Cloridreto de L-lisina	200 mg/litro
20	DL-Metionina	200 mg/litro
	Ácido L-glutâmico	30 g/litro
	Cloridreto de tetraciclina	12,5 mg/litro
	Cloranfenicol	25 mg/litro

O meio foi ajustado a pH 4,5 ou pH 4,9 com amônia aquosa, e
25 depois foi filtrado.

As cepas NA1::Ptac-kdp/pSTV-yhfK e NA1/pSTV-yhfK foram, cada uma, pré-cultivadas no meio L (meio contendo 10 g de Bactotripton, 5 g de extrato de levedura, 5 g de NaCl e 15 g de ágar, em 1 litro de água purificada, pH 7,0) que foi suplementado com ingredientes de

meio mínimo (meio contendo 0,5 g de glicose, sulfato de magnésio 2 mM, 3 g de fosfato monopotássico, 0,5 g de cloreto de sódio, 1 g de cloreto de amônio e 6 g de fosfato dissódico, em 1 litro de água purificada), 12,5 mg/litro de cloridreto de tetraciclina e 25 mg/litro de cloranfenicol, e as células correspondentes a 1/8 da placa foram raspadas, lavadas duas vezes com solução salina fisiológica, e finalmente colocadas em suspensão em 1 ml de solução salina fisiológica. A suspensão, em um volume de 20 µl, foi inoculada em 5 ml do meio para cultura no tubo de teste contida em um tubo de teste, e cultivada em 34°C com agitação. Durante a cultura, a OD (660 nm) foi medida a cada 30 minutos com o uso de um medidor de OD automático (TN1506 BIO PHOTORECORDER, ADVANTEC). Os resultados são mostrados na Figura 4.

Em comparação com a cepa NA1/pSTV-yhfK como um controle, a cepa intensificada por *kdp*, a cepa NA1::Ptac-kdp/pSTV-yhfK, apresentaram melhoramento do crescimento sob condições ácidas de pH 4,5 ou pH 4,9. Assim, o efeito da intensificação do óperon *kdp* para melhorar o crescimento e o índice de produção de ácido L-glutâmico, foi demonstrado por estes resultados.

(4) AVALIAÇÃO DA CULTURA DA CEPA SUBSTITUÍDA PELO PROMOTOR DO ÓPERON *KDP* EM S-jar

Então, de modo a examinar o efeito da intensificação do óperon *kdp* sobre a produção do ácido L-glutâmico, a cultura de produção do ácido L-glutâmico foi realizada pelo uso da cepa NA1:Ptac-kdp/pSTV-yhfK e da cepa NA1/pSTV-yhfK.

A cultura foi realizada por duas etapas de cultura de sementes para permitir a formação das células e a cultura principal para produzir ácido L-glutâmico.

A cultura das sementes foi realizada com a seguinte composição de meio.

[Composição do meio de cultura das sementes]

	Sacarose	50 g/litro
	MgSO ₄ ·7H ₂ O	0,4 g/litro
	GD113 (antiespuma)	0,1 ml/litro
5	(NH ₄) ₂ SO ₄	4,0 g/litro
	KH ₂ PO ₄	2,0 g/litro
	Extrato de levedura	4,0 g/litro
	FeSO ₄ ·7H ₂ O	0,01 g/litro
	MnSO ₄ ·5H ₂ O	0,01 g/litro
10	Ácido cítrico	0,02 g/litro
	Cloridreto de L-lisina	0,4 g/litro
	DL-Metionina	0,4 g/litro
	Ácido ε-diaminopimélico	0,4 g/litro
	Pantotenato de cálcio	18 mg/litro
15	Cloridreto de tetraciclina	12,5 mg/litro
	Cloranfenicol	25 mg/litro

O meio foi esterilizado com vapor em 120°C por 20 minutos.

As cepas NA1::Ptac-kdp/pSTV-yhfK e NA1/pSTV-yhfK foram, cada uma, pré-cultivadas no meio L (meio contendo 10 g de Bactotripton, 5 g de extrato de levedura, 5 g de NaCl e 15 g de ágar, em 1 litro de água purificada, pH 7,0) que foi suplementado com ingredientes de meio mínimo (meio contendo 0,5 g de glicose, sulfato de magnésio 2 mM, 3 g de fosfato monopotássico, 0,5 g de cloreto de sódio, 1 g de cloreto de amônio e 6 g de fosfato dissódico, em 1 litro de água purificada), 12,5 mg/litro de tetraciclina e 25 mg/litro de cloranfenicol, e as células correspondentes a uma placa foram inoculadas em 300 ml do meio da composição acima mencionada contida em uma minijarra de 1 litro de volume, e a agitação foi controlada em 34°C e pH de 6,0 por cerca de 12 horas de modo que a aeração de 1/1 vvm e uma concentração de oxigênio de 3 % ou mais fosse obtida. Durante a cultura,

o pH foi controlado de modo a ser de 6,0 com a adição de gás amoníaco. A cultura das sementes terminou no momento da depleção do sacarídeo no meio observado como um índice.

A composição do meio de cultura principal é mostrada abaixo.

5 [Composição do meio de cultura] (As concentrações são aquelas após a inoculação de 20 % do meio de cultura das sementes)

	Sacarose	100 g/litro
	MgSO ₄ ·7H ₂ O	0,4 g/litro
	GD113	0,1 ml/litro
10	(NH ₄) ₂ SO ₄	5,0 g/litro
	KH ₂ PO ₄	6,0 g/litro
	Extrato de levedura	6,0 g/litro
	FeSO ₄ ·7H ₂ O	0,02 g/litro
	MnSO ₄ ·5H ₂ O	0,02 g/litro
15	Ácido cítrico	0,02 g/litro
	Betáína*	2,0 g/litro
	Cloridreto de L-lisina	0,8 g/litro
	DL-Metionina	0,6 g/litro
	Ácido ε-diaminopimélico	0,6 g/litro
20	Pantotenato de cálcio	18 mg/litro
	Cloridreto de tetraciclina	25 mg/litro
	Cloranfenicol	25 mg/litro

* N,N,N-trimetilglicina

25 As células obtidas pela cultura das sementes em um volume de 60 ml foram inoculadas em 240 ml de meio tendo a composição acima mencionada contida em um minijarro de 1 litro de volume, e cultivadas em pH 4,7. A cultura foi terminada 16 horas após o início da cultura principal. A densidade celular e a concentração de ácido L-glutâmico no meio de cultura foram medidas através do tempo. A densidade celular foi examinada pela

medição da turbidez do meio de cultura diluído 101 vezes com água em 620 nm com o uso de um espectrofotômetro (U-2000A, Hitachi). A concentração do ácido L-glutâmico foi medida quanto ao sobrenadante da cultura apropriadamente diluído com água mediante o uso do Biotech Analyzer (AS-5 210, Sakura SI).

Os resultados são apresentados na Tabela 1 e na Figura 5. Ficou claro que o crescimento, bem como o acúmulo do ácido L-glutâmico e o índice de produção do ácido L-glutâmico da cepa intensificada pelo óperon *kdp*, a cepa NA1::Ptac-*kdp*/pSTV-yhfK, foram melhorados em comparação com a cepa comparativa, a cepa NA1/pSTV-yhfK.

TABELA 1

	NA1/pSTVyhfK	NA1::Ptac- <i>kdp</i> /pSTV-yhfK
Ácido L-glutâmico produzido (g/Jar)	15,8	19,0
Índice de produção do ácido L-glutâmico (g/litro/hora)	3,30	3,96

EXEMPLO 2

AMPLIFICAÇÃO DO ÓPERON *kdp* NO *Escherichia coli* DE ACUMULAÇÃO DA L-TREONINA15 (1) CONSTRUÇÃO DO PLASMÍDEO PARA AMPLIFICAÇÃO DO ÓPERON *kdp*

De modo a introduzir o óperon *kdp* em uma bactéria de *Escherichia*, um plasmídeo para amplificação do óperon *kdp* é construído mediante o uso de um plasmídeo conhecido pMW218 (Takara Shuzo).

20 O pMW218 é primeiro digerido com as enzimas de restrição *Hind*III e *Bam*HI, e a reação é terminada pela adição de uma solução de fenol/clorofórmio e misturando-os. A mistura de reação é centrifugada, depois a camada superior é coletada, e os DNAs são coletados pela precipitação de etanol. O óperon *kdp* é separadamente amplificado por PCR com o uso do cromossoma extraído do *Escherichia coli* MG1655 como o padrão e os iniciadores de DNA apresentados nas SEQ ID NOS: 55 e 56 (desnaturação

em 94°C por 10 segundos, recozendo-se em 60°C por 30 segundos, e extensão em 72°C por 120 segundos). Para a PCR, Pyrobest DNA polimerase (Takara Shuzo) é usado. O fragmento de óperon *kdp* obtido é digerido com as enzimas de restrição *Hind*III e *Bam*HI, e a reação é terminada pela adição de
5 uma solução de fenol/clorofórmio, e misturando-os.

O digerido pMW218 e o fragmento da região do gene *kdpABC* preparados como descrito acima, são ligados pelo uso do Kit de Ligação de DNA Versão 2 (Takara Shuzo). *Escherichia coli* (células competentes JM109 de *E. coli*, Takara Shuzo) é transformado com a solução
10 de ligação, aplicada ao meio de ágar de LB contendo 50 mg/litro de canamicina, e incubada durante a noite em 37°C. As colônias que apareceram sobre o meio de ágar são inoculadas no meio líquido de LB contendo 50 mg/litro de canamicina, e cultivadas em 37°C por 8 horas com agitação. O DNA plasmídeo é extraído de cada meio de cultura pelo método de SDS
15 alcalino, e sua estrutura é confirmada pela digestão com as enzimas de restrição para se obter pMW218kdp.

(2) INTRODUÇÃO DE pMW218kdp NO *Escherichia coli* B-3996 DE ACÚMULO DA TREONINA E PRODUÇÃO DE AMINOÁCIDO.

O pMW218kdp obtido como descrito acima é introduzido na
20 cepa VKPM B-3996 pelo método de eletroporação [*Canadian Journal of Microbiology*, 43, 197 (1997)].

O transformante obtido (esta cepa é denominada “B-3996/pMW218kdp”) e a cepa em que pMW218 é introduzido como um controle (esta cepa é denominada “B-3996/pMW218”) são cultivados como
25 segue, e as concentrações de L-treonina nos sobrenadantes da cultura são examinados.

Cada transformante é inoculado em 3 ml do meio líquido de LB contendo 50 mg/litro de canamicina e 20 mg/litro de estreptomicina, e cultivado durante a noite em 37°C em um tubo de teste, depois 200 µl do

meio de cultura são inoculados em um meio de produção de treonina (20 ml) contendo 50 mg/litro de canamicina e 20 mg/litro de estreptomicina, e a cultura é realizada em 37°C por 24 horas com agitação. Após conclusão da cultura, as células são removidas por centrifugação, e a concentração da L-treonina no sobrenadante da cultura é medida pelo uso de um analisador de aminoácido (L-8500, Hitachi). Pode-se observar que a quantidade de acúmulo da treonina no meio é melhorada na cepa amplificada pelo óperon *kdp*, B-3996/pMW218kdp em comparação com a cepa B-3996/pMW218 como um controle.

10	[MEIO DE PRODUÇÃO DA TREONINA]	
	D-glicose	40 g/litro
	(NH ₄) ₂ SO ₄	16 g/litro
	KH ₂ PO ₄	1,0 g/litro
	MgSO ₄ ·7H ₂ O	1,0 g/litro
15	FeSO ₄ ·7H ₂ O	0,01 g/litro
	MnSO ₄ ·7H ₂ O	0,01 g/litro
	L-Isoleucina	50 mg/litro
	DL-Metionina	500 mg/litro
	Carbonato de cálcio	0,6 g/litro
20	Estreptomicina	20 mg/litro
	Canamicina	50 mg/litro

O meio é ajustado a pH 7,5 com hidróxido de potássio.

O meio é esterilizado com vapor em 115°C por 10 minutos.

EXPLANAÇÃO DA LISTAGEM DE SEQUÊNCIAS

25 SEQ ID NO: 1: Sequência de nucleotídeos de óperon *kdp* de *Escherichia coli* (sequências de aminoácidos de *kdpA*, *kdpB* e *kdpC* são também mostradas)

kdpF: 457 a 546

kdpA: 546 a 2219

kdpB: 2242 a 4290

kdpC: 4299 a 4871

kdpD: 4864 a 7548

kdpE: 7545 a 8222

- 5 SEQ ID NO: 2: Sequência de aminoácidos de KdpA
 SEQ ID NO: 3: Sequência de aminoácidos de KdpB
 SEQ ID NO: 4: Sequência de aminoácidos de KdpC
 SEQ ID NO: 5: Sequência de aminoácidos de KdpD
 SEQ ID NO: 6: Sequência de aminoácidos de KdpE
- 10 SEQ ID NO: 5: Sequência de nucleotídeos do óperon KdpD (a
 sequência de aminoácidos de SDHC é também mostrada)

SEQ ID NO: 7: Sequência de nucleotídeos de óperon *kdp* de
Pantoea ananatis (as sequências de aminoácidos de *kdpA*, *kdpB*, *kdpC* e *kdpD*
 são também mostradas)

15 *kdpA*: 543 a 2225

kdpB: 2228 a 4273

kdpC: 4284 a 4853

kdpD: 4867 a 7542

- 20 SEQ ID NO: 8: Sequência de aminoácidos de KdpA
 SEQ ID NO: 9: Sequência de aminoácidos de KdpB
 SEQ ID NO: 10: Sequência de aminoácidos de KdpC
 SEQ ID NO: 11: Sequência de aminoácidos de KdpD
 SEQ ID NO: 12: Sequência de nucleotídeos de gene *kdpE* de
Pantoea ananatis

25 SEQ ID NO: 13: Sequência de aminoácidos de KdpE

SEQ ID NO: 14: Sequência de nucleotídeos de do gene *hisD*
 de *Pantoea ananatis*

SEQ ID NO: 15: Iniciador para amplificação do fragmento
 para integração do gene Km^r no gene *hosD*

SEQ ID NO: 16: Iniciador para amplificação do fragmento para integração do gene Km^r no gene *hosD*

SEQ ID NO: 17: Iniciador para amplificação do gene *cat*

SEQ ID NO: 18: Iniciador para amplificação do gene *cat*

5 SEQ ID NO: 19: Iniciador para amplificação do gene *sacB*

SEQ ID NO: 20: Iniciador para amplificação do gene *sacB*

SEQ ID NO: 21: Iniciador para amplificação do fragmento de DNA contendo o promotor PlacUV5

10 SEQ ID NO: 22: Iniciador para amplificação do fragmento de DNA contendo o promotor PlacUV5

SEQ ID NO: 23: Iniciador para amplificação do fragmento de DNA contendo os genes λ Red $\alpha\beta\gamma$ e tL3

SEQ ID NO: 24: Iniciador para amplificação do fragmento de DNA contendo os genes λ Red $\alpha\beta\gamma$ e tL3

15 SEQ ID NO: 25: Iniciador para amplificação do fragmento de DNA contendo o promotor PlacUV5 e TrnB

SEQ ID NO: 26: Iniciador para amplificação de fragmento de DNA contendo o promotor PlacUV5 e TrnB

SEQ ID NO: 27: Iniciador para amplificação de *attL*

20 SEQ ID NO: 28: Iniciador para amplificação de *attL*

SEQ ID NO: 29: Sequência de nucleotídeos de *attL*

SEQ ID NO: 30: Iniciador para amplificação de *attR*

SEQ ID NO: 31: Iniciador para amplificação de *attR*

SEQ ID NO: 32: Sequência de nucleotídeos de *attR*

25 SEQ ID NO: 33: Iniciador para amplificação de fragmento de DNA contendo o gene *bla*

SEQ ID NO: 34: Iniciador para amplificação de fragmento de DNA contendo o gene *bla*

SEQ ID NO: 35: Iniciador para amplificação de fragmento de

DNA contendo *ter_rrnB*

SEQ ID NO: 36: Iniciador para amplificação de fragmento de

DNA contendo *ter_rrnB*

SEQ ID NO: 37: Sequência de nucleotídeos de fragmento de

5 DNA contendo o terminador *ter_thrL*

SEQ ID NO: 38: Iniciador para amplificação de fragmento de

DNA contendo o terminador *ter_thrL*

SEQ ID NO: 39: Iniciador para amplificação de fragmento de

DNA contendo o terminador *ter_thrL*

10 SEQ ID NO: 40: Iniciador para amplificar parte do gene *gltA*
outro que não ORF

SEQ ID NO: 41: Iniciador para amplificar parte do gene *gltA*

outro que não ORF

SEQ ID NO: 42: Iniciador para amplificação do gene *prpC*

15 SEQ ID NO: 43: Iniciador para amplificação do gene *prpC*

SEQ ID NO: 44: Sequência de nucleotídeos do gene *yghU* de

Pantoea ananatis

SEQ ID NO: 45: Sequência de nucleotídeos do gene *scrK* de

Pantoea ananatis

20 SEQ ID NO: 46: Sequência de nucleotídeos do gene *lacZ* de

Pantoea ananatis

SEQ ID NO: 47: Iniciador para amplificação do fragmento de

DNA contendo o promotor *Ptac*

SEQ ID NO: 48: Iniciador para amplificação do fragmento de

25 DNA contendo o promotor *Ptac*

SEQ ID NO: 49: Iniciador para amplificação do fragmento de

DNA contendo o gene de resistência *Km*

SEQ ID NO: 50: Iniciador para amplificação do fragmento de

DNA contendo o gene de resistência *Km*

SEQ ID NO: 51: Iniciador para amplificação da sequência de montante do óperon *kdp* ligada ao promotor tac

SEQ ID NO: 52: Iniciador para amplificação da sequência de montante do óperon *kdp* ligado ao promotor tac

5 SEQ ID NO: 53: Iniciador para confirmar a estrutura de montante do óperon *kdp*

SEQ ID NO: 54: Iniciador para confirmar a estrutura de montante do óperon *kdp*

SEQ ID NO: 55: Iniciador para amplificação do óperon *kdp*

10 SEQ ID NO: 56: Iniciador para amplificação do óperon *kdp*

SEQ ID NO: 57: Sequência de consenso das sequências de aminoácidos KdpA de *Pantoea ananatis* e *Escherichia coli*

SEQ ID NO: 58: Sequência de consenso das sequências de aminoácidos KdpB de *Pantoea ananatis* e *Escherichia coli*

15 SEQ ID NO: 59: Sequência de consenso das sequências de aminoácidos KdpC de *Pantoea ananatis* e *Escherichia coli*

APLICABILIDADE INDUSTRIAL

Mediante o uso do microorganismo da presente invenção, um L-aminoácido tal como o ácido L-glutâmico, L-lisina, L-treonina, L-arginina, 20 L-histidina, L-isoileucina, L-valina, L-leucina, L-treonina, L-fenilalanina, L-tirosina, L-triptofano ou L-cisteína, ácido L-glutâmico, ou similar, pode ser eficientemente produzido por fermentação. Em uma forma de realização preferida, o microorganismo da presente invenção apresenta tanto uma quantidade de produção de L-aminoácido quanto o índice de produção, 25 superiores.

TABELA DE MICROORGANISMOS

Apêndice 3

Página 14

Tratado de Budapeste sobre o reconhecimento internacional do depósito de microorganismos para os fins de procedimento de patente

Para Ajonomoto FORMULÁRIO INTERNACIONAL recibo no caso de um depósito original Expedido de acordo com a Regra 7.1 pela AUTORIDADE DEPOSITÁRIA INTERNACIONAL identificada no final desta página.

1-15-1 Kyobashi, Chuo-ku
Tóquio 104-8315
Japão
nome e endereço
do depositante

I. IDENTIFICAÇÃO DO MICROORGANISMO	
Referência de identificação dada pelo DEPOSITANTE Escherichia coli TAR61 _{pVIC40}	Número de acesso dado pela AUTORIDADE DEPOSITÁRIA INTERNACIONAL: B-3996
II. DESCRIÇÃO CIENTÍFICA E/OU DESIGNAÇÃO TAXONÔMICA PROPOSTA	
O microorganismo especificado em I acima foi acompanhado por: -x- uma descrição científica produtor de treonina — uma designação taxonômica proposta — (Marcar com uma cruz onde aplicável)	
III. RECIBO E ACEITAÇÃO	
Nossa autoridade depositária internacional aceita o microorganismo especificado na coluna I, recebido em 07 de abril de 1987 (data de depósito original).	
IV. RECIBO DE SOLICITAÇÃO PARA CONVERSÃO	
O microorganismo identificado em I acima foi recebido por esta Autoridade Depositária Internacional em (data do depósito original) e uma solicitação para converter o depósito original em um depósito sob o Tratado de Budapeste foi recebida pela mesma em (data de recibo da solicitação para conversão)	
V. AUTORIDADE DEPOSITÁRIA INTERNACIONAL	
NOME: Russian National Collection of Industrial Microorganisms (VKPM) Funcionário(s) FGUP GosNIlgenetika: Endereço: Russia 117545 Moscou 1 Dorozhny proezd 1	Assinatura(s) de pessoa(s) tendo o poder de representar a Autoridade Depositária Internacional Data: 19/12/002 Sineoky S.P.

FORMA INTERNACIONAL

TRATADO DE BUDAPESTE SOBRE O RECONHECIMENTO
INTERNACIONAL DO DEPÓSITO DE MICROORGANISMOS
PARA OS FINS DE PROCEDIMENTO DE PATENTE

RECIBO NO CASO DE UM DEPÓSITO ORIGINAL

expedido de acordo com a Regra 7.1 pela AUTORIDADE
DEPOSITÁRIA INTERNACIONAL identificada no final
desta página.

NOME DO DEPOSITANTE: Ajinomoto Co. Inc.
Presidente, Kunio, EGASHIRA
ENDEREÇO: 15-1, Kyobashi 1-chome, Chuo-ku,
Tóquio, Japão

I. Indicação de um Microorganismo	
Indicação pelo depositante para especificar o microorganismo <i>Enterobacter agglomerans</i> AJ13601	Número do Depósito FERM BP-7207
II. Natureza científica e uma posição por taxonomia	
O microorganismo especificado na coluna I foi anexado a documentos em que as seguintes matérias foram descritas.	
<input type="checkbox"/> natureza científica <input checked="" type="checkbox"/> uma posição por taxonomia	
III. Recibo e Aceitação	
Nossa autoridade depositária internacional aceita o microorganismo especificado na coluna I, recebido em 18 de agosto de 1999 (data de depósito original).	
IV. Recibo de solicitação de transferência	
Nossa autoridade depositária internacional recebeu o microorganismo especificado na coluna I em 18 de agosto de 1999 (data de depósito original). E nós aceitamos uma solicitação de transferência para depósito sob o Tratado de Budapeste do depósito original em 6 de julho de 2000. (Transferência de FERM P-17516 depositado em 18 de agosto de 1999)	
V. Autoridade Depositária Internacional	
NOME:	Institute of Bioscience and Human-Technology Agency of Industrial Science and Technology Dr. Shinichi OHASHI, DIRETOR GERAL
ENDEREÇO	1-3, Higashi 1-chome, Tsukuba-shi Ibaraki-ken, 305-8566 Japão (vedado)
Datado de 6 de julho de 2000	

LISTAGEM DE SEQUÊNCIAS

<110> Ajinomoto Co., Inc.

<120> Uma Bactéria que produz L-Aminoácido e um Método para Produzir um L-Aminoácido

<130> C874-C7269

<150> JP2007-011392
<151> 2007-01-22

<150> JP2007-131763
<151> 2007-05-17

<160> 59

<170> PatentIn version 3.3

<210> 1
<211> 8501
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (546)..(2219)

<220>
<221> CDS
<222> (2242)..(4290)

<220>
<221> CDS
<222> (4299)..(4871)

<400> 1

cgaccagacg	cgatgcaggt	agctgtaaaa	gcgggcgcg	tccttccaga	aaagcatgaa	60
aggcagcgcc	agaacgccc	ctgcaaccgc	ataagtacgg	cgtaaaaaga	taagccatgc	120
aggatattct	ctgtagagtt	ccatactttt	ctcccccttt	tgtatctacc	cggtgaatgg	180
caccggaaaa	atgaatttgt	ttatctgatg	aaaatagtac	cgctttttgt	gtaattttac	240
tactcatccg	accacttatt	tttgcttatt	gatggtttat	ttacattcat	cctgtaatta	300
agttacacaa	aagttaaatt	aatactaaac	attagttaaa	tcattggcttt	tgccattttt	360
atactttttt	tacaccccgc	ccgcagattt	ttgcgaaatc	ttgcagcca	gaattctacc	420
cttccgggat	cacttttagg	ccactggagg	tgcactgtga	gtgcaggcgt	gataaccggc	480
gtattgctgg	tgtttttatt	actgggttat	ctggtttatg	ccctgatcaa	tgcgaggcgg	540
ttctg atg gct gcg caa ggg ttc tta ctg atc gcc acg ttt tta ctg gtg						590
	Met Ala Ala Gln Gly Phe Leu Leu Ile Ala Thr Phe Leu Leu Val					
	1 5 10 15					
tta atg gtg ctg gcg cgt cct tta ggc agc ggg ctg gcg cgg ctg att						638
Leu Met Val Leu Ala Arg Pro Leu Gly Ser Gly Leu Ala Arg Leu Ile						
	20 25 30					
aat gac att cct ctt ccc ggt aca acg ggc gtt gag cgc gta ctt ttt						686
Asn Asp Ile Pro Leu Pro Gly Thr Thr Gly Val Glu Arg Val Leu Phe						
	35 40 45					
cgc gca ctt ggc gtc tct gac cgt gag atg aac tgg aag caa tat ctt						734
Arg Ala Leu Gly Val Ser Asp Arg Glu Met Asn Trp Lys Gln Tyr Leu						
	50 55 60					
tgt gcc att ctc ggc ctg aac atg ctg ggg ctg gcg gtg ctg ttt ttt						782
Cys Ala Ile Leu Gly Leu Asn Met Leu Gly Leu Ala Val Leu Phe Phe						
	65 70 75					

atg ttg ctc ggt cag cac tat ctg ccg ctt aat cca cag cag ttg cca	830
Met Leu Leu Gly Gln His Tyr Leu Pro Leu Asn Pro Gln Gln Leu Pro	
80 85 90 95	
ggg ctg tcg tgg gat ctg gcg ctg aat acc gcc gtc agc ttt gtc acc	878
Gly Leu Ser Trp Asp Leu Ala Leu Asn Thr Ala Val Ser Phe Val Thr	
100 105 110	
aat acc aac tgg caa tct tat agc ggt gaa acc acg ttg agc tat ttc	926
Asn Thr Asn Trp Gln Ser Tyr Ser Gly Glu Thr Thr Leu Ser Tyr Phe	
115 120 125	
agc cag atg gcg ggc tta acg gtg caa aac ttt ctt tct gcc gcc agc	974
Ser Gln Met Ala Gly Leu Thr Val Gln Asn Phe Leu Ser Ala Ala Ser	
130 135 140	
ggg att gcg gtg att ttt gcc ctc atc cgt gcg ttt acc cgc cag agc	1022
Gly Ile Ala Val Ile Phe Ala Leu Ile Arg Ala Phe Thr Arg Gln Ser	
145 150 155	
atg agc acg ctc ggg aat gcc tgg gtc gat ctg cta cgc atc acg tta	1070
Met Ser Thr Leu Gly Asn Ala Trp Val Asp Leu Leu Arg Ile Thr Leu	
160 165 170 175	
tgg gtg cta gtc cct gtg gcg ttg ttg att gca ctg ttt ttt att caa	1118
Trp Val Leu Val Pro Val Ala Leu Leu Ile Ala Leu Phe Phe Ile Gln	
180 185 190	
caa ggt gcg ctg caa aac ttt ctg cct tat cag gct gtg aat acc gtt	1166
Gln Gly Ala Leu Gln Asn Phe Leu Pro Tyr Gln Ala Val Asn Thr Val	
195 200 205	
gaa gga gcg caa cag ctg tta ccc atg ggg cct gta gct tct cag gaa	1214
Glu Gly Ala Gln Gln Leu Leu Pro Met Gly Pro Val Ala Ser Gln Glu	
210 215 220	
gcg atc aag atg ctc ggt act aac ggc ggt ggc ttc ttt aat gcc aac	1262
Ala Ile Lys Met Leu Gly Thr Asn Gly Gly Gly Phe Phe Asn Ala Asn	
225 230 235	
tcg tcg cat ccg ttt gaa aac cca acc gca ctg acc aac ttc gtg cag	1310
Ser Ser His Pro Phe Glu Asn Pro Thr Ala Leu Thr Asn Phe Val Gln	
240 245 250 255	
atg ctg gcg atc ttc ttg atc cca acg gcg ctg tgc ttt gcc ttt ggt	1358
Met Leu Ala Ile Phe Leu Ile Pro Thr Ala Leu Cys Phe Ala Phe Gly	
260 265 270	
gaa gtg atg ggc gat cgc cgc cag ggg cgc atg ttg ctg tgg gcg atg	1406
Glu Val Met Gly Asp Arg Arg Gln Gly Arg Met Leu Leu Trp Ala Met	
275 280 285	
tca gtg att ttt gtc atc tgc gta ggc gtg gtg atg tgg gca gaa gtt	1454
Ser Val Ile Phe Val Ile Cys Val Gly Val Val Met Trp Ala Glu Val	
290 295 300	
cag ggt aat cct cat ctg ctg gca ctg ggc acg gac agc agc atc aat	1502
Gln Gly Asn Pro His Leu Leu Ala Leu Gly Thr Asp Ser Ser Ile Asn	
305 310 315	
atg gaa ggt aaa gag agc cgt ttc ggc gtg ctg gtc agt agc ctg ttt	1550
Met Glu Gly Lys Glu Ser Arg Phe Gly Val Leu Val Ser Ser Leu Phe	
320 325 330 335	
gcg gtc gtg acg acg gcg gct tcc tgt ggc gcg gtg att gcg atg cat	1598
Ala Val Val Thr Thr Ala Ala Ser Cys Gly Ala Val Ile Ala Met His	
340 345 350	
gat tcg ttt acc gct ctc ggt ggc atg gtg ccg atg tgg ctg atg caa	1646
Asp Ser Phe Thr Ala Leu Gly Gly Met Val Pro Met Trp Leu Met Gln	
355 360 365	
att ggt gaa gtg gtg ttc ggc ggt gtc ggt tct ggt ctt tac ggc atg	1694
Ile Gly Glu Val Val Phe Gly Gly Val Gly Ser Gly Leu Tyr Gly Met	
370 375 380	
atg ctg ttt gtc ctg ctg gcg gtg ttt att gcc ggc ctg atg att ggt	1742
Met Leu Phe Val Leu Leu Ala Val Phe Ile Ala Gly Leu Met Ile Gly	
385 390 395	
cgt aca ccg gaa tat ctg ggt aaa aaa atc gac gta cgc gag atg aaa	1790
Arg Thr Pro Glu Tyr Leu Gly Lys Lys Ile Asp Val Arg Glu Met Lys	
400 405 410 415	

ctg act gca ctg gca att ctg gtt acc ccg acg ctg gtg ctg atg ggc	1838
Leu Thr Ala Leu Ala Ile Leu Val Thr Pro Thr Leu Val Leu Met Gly	
420 425 430	
gcg gcg ttg gcg atg atg acc gac gcc gga cgt agc gcc atg ctc aac	1886
Ala Ala Leu Ala Met Met Thr Asp Ala Gly Arg Ser Ala Met Leu Asn	
435 440 445	
cct ggc ccg cat ggt ttt agc gaa gtg ctg tac gcc gtg tca tcc gcc	1934
Pro Gly Pro His Gly Phe Ser Glu Val Leu Tyr Ala Val Ser Ser Ala	
450 455 460	
gct aac aac aac ggc agc gcc ttt gcc gga tta agc gcc aac tct ccg	1982
Ala Asn Asn Asn Gly Ser Ala Phe Ala Gly Leu Ser Ala Asn Ser Pro	
465 470 475	
ttc tgg aac tgt tta ctg gcg ttc tgc atg ttt gtc ggt cgc ttc ggg	2030
Phe Trp Asn Cys Leu Leu Ala Phe Cys Met Phe Val Gly Arg Phe Gly	
480 485 490 495	
gtg att atc ccg gtg atg gca att gcc ggt tcg ctg gtg agt aaa aag	2078
Val Ile Ile Pro Val Met Ala Ile Ala Gly Ser Leu Val Ser Lys Lys	
500 505 510	
agc caa gcc gcc agc tcc ggc acg ctg cca acg cac ggc ccg ctg ttt	2126
Ser Gln Ala Ala Ser Ser Gly Thr Leu Pro Thr His Gly Pro Leu Phe	
515 520 525	
gtt ggc ctg tta atc ggc acc gtg ttg ctg gtt ggc gca ctg acc ttt	2174
Val Gly Leu Leu Ile Gly Thr Val Leu Leu Val Gly Ala Leu Thr Phe	
530 535 540	
atc cct gcc ctg gcg ctt ggt ccg gtg gcg gaa tat ctc tcc tga	2219
Ile Pro Ala Leu Ala Leu Gly Pro Val Ala Glu Tyr Leu Ser	
545 550 555	
tgatattgag tgagcactga at atg agt cgt aaa caa ctg gcg cta ttc gaa	2271
Met Ser Arg Lys Gln Leu Ala Leu Phe Glu	
560 565	
cca aca ctt gtc gtt cag gcg ctg aaa gaa gcg gtg aaa aaa tta aac	2319
Pro Thr Leu Val Val Gln Ala Leu Lys Glu Ala Val Lys Lys Leu Asn	
570 575 580	
ccg cag gcg caa tgg cgc aat ccg gtg atg ttt atc gtc tgg atc ggc	2367
Pro Gln Ala Gln Trp Arg Asn Pro Val Met Phe Ile Val Trp Ile Gly	
585 590 595	
agt ctg ctg acc acc tgt att agc atc gcg atg gca agc ggt gcg atg	2415
Ser Leu Leu Thr Thr Cys Ile Ser Ile Ala Met Ala Ser Gly Ala Met	
600 605 610 615	
ccc ggc aat gcg ctg ttt agc gcg gcc att agc ggt tgg ctg tgg atc	2463
Pro Gly Asn Ala Leu Phe Ser Ala Ala Ile Ser Gly Trp Leu Trp Ile	
620 625 630	
acc gta ctg ttc gct aat ttc gcc gag gcg ctg gca gaa ggc cgc agt	2511
Thr Val Leu Phe Ala Asn Phe Ala Glu Ala Leu Ala Glu Gly Arg Ser	
635 640 645	
aaa gcg cag gcc aac agt ctg aaa ggg gtg aaa aaa act gcc ttt gcc	2559
Lys Ala Gln Ala Asn Ser Leu Lys Gly Val Lys Lys Thr Ala Phe Ala	
650 655 660	
cgc aag ctg cgt gag ccg aaa tat ggc gct gcg gcg gac aaa gtt cct	2607
Arg Lys Leu Arg Glu Pro Lys Tyr Gly Ala Ala Ala Asp Lys Val Pro	
665 670 675	
gcc gac caa ctt cgt aaa ggc gat atc gta ctg gta gaa gct ggc gat	2655
Ala Asp Gln Leu Arg Lys Gly Asp Ile Val Leu Val Glu Ala Gly Asp	
680 685 690 695	
att atc ccc tgc gat ggt gaa gtt att gaa ggg ggt gca tcg gtc gat	2703
Ile Ile Pro Cys Asp Gly Glu Val Ile Glu Gly Gly Ala Ser Val Asp	
700 705 710	
gaa agc gcc atc acc ggg gaa tcg gca ccg gtg atc cgt gaa tcc ggc	2751
Glu Ser Ala Ile Thr Gly Glu Ser Ala Pro Val Ile Arg Glu Ser Gly	
715 720 725	
ggc gat ttt gcc tcc gtc acc ggc ggc acg cgt att ctt tct gac tgg	2799
Gly Asp Phe Ala Ser Val Thr Gly Gly Thr Arg Ile Leu Ser Asp Trp	
730 735 740	

ctg gtg att gag tgt agc gtt aac ccc ggc gag aca ttt ctg gat cgg	2847
Leu Val Ile Glu Cys Ser Val Asn Pro Gly Glu Thr Phe Leu Asp Arg	
745 750 755	
atg atc gcg atg gtg gaa ggc gca cag cga cgc aaa acg ccg aac gaa	2895
Met Ile Ala Met Val Glu Gly Ala Gln Arg Arg Lys Thr Pro Asn Glu	
760 765 770 775	
att gcc ctg acc att ctg ctg att gcc ctg act atc gtc ttt tta ctg	2943
Ile Ala Leu Thr Ile Leu Leu Ile Ala Leu Thr Ile Val Phe Leu Leu	
780 785 790	
gca acc gcc acg ctg tgg ccg ttt tcc gcg tgg ggc ggt aat gca gtc	2991
Ala Thr Ala Thr Leu Trp Pro Phe Ser Ala Trp Gly Gly Asn Ala Val	
795 800 805	
agc gta acg gta ctg gtg gcg ctg ctg gtc tgt ctg atc cca acc act	3039
Ser Val Thr Val Leu Val Ala Leu Leu Val Cys Leu Ile Pro Thr Thr	
810 815 820	
att ggc ggc ctg ttg tca gcg atc ggc gtc gcc ggg atg agc cgg atg	3087
Ile Gly Gly Leu Leu Ser Ala Ile Gly Val Ala Gly Met Ser Arg Met	
825 830 835	
cta ggc gcg aat gtg att gcc acc agc gga cgt gca gtt gaa gcg gca	3135
Leu Gly Ala Asn Val Ile Ala Thr Ser Gly Arg Ala Val Glu Ala Ala	
840 845 850 855	
ggt gac gtt gac gtt ctg cta ctg gat aaa acc ggc acc atc aca ctc	3183
Gly Asp Val Asp Val Leu Leu Leu Asp Lys Thr Gly Thr Ile Thr Leu	
860 865 870	
ggt aac cgt cag gcg tcg gag ttt atc ccc gcg cag ggc gtg gat gaa	3231
Gly Asn Arg Gln Ala Ser Glu Phe Ile Pro Ala Gln Gly Val Asp Glu	
875 880 885	
aaa acg ctg gct gac gcc gca caa ctg gct tcg ctg gct gat gaa acg	3279
Lys Thr Leu Ala Asp Ala Ala Gln Leu Ala Ser Leu Ala Asp Glu Thr	
890 895 900	
ccg gaa ggc cgc agt att gtg atc ctc gcc aag cag cgt ttt aac ctg	3327
Pro Glu Gly Arg Ser Ile Val Ile Leu Ala Lys Gln Arg Phe Asn Leu	
905 910 915	
cgc gag cgc gat gtg cag tcg ctc cat gcc acc ttt gta ccg ttt act	3375
Arg Glu Arg Asp Val Gln Ser Leu His Ala Thr Phe Val Pro Phe Thr	
920 925 930 935	
gcg caa agc cgg atg agc ggg atc aac atc gac aac cgc atg atc cgt	3423
Ala Gln Ser Arg Met Ser Gly Ile Asn Ile Asp Asn Arg Met Ile Arg	
940 945 950	
aaa ggt tct gtc gat gcc att cgt cgc cat gtt gag gct aac ggt ggt	3471
Lys Gly Ser Val Asp Ala Ile Arg Arg His Val Glu Ala Asn Gly Gly	
955 960 965	
cac ttc cct acc gat gtt gat caa aaa gtc gat cag gtt gcg cgt cag	3519
His Phe Pro Thr Asp Val Asp Gln Lys Val Asp Gln Val Ala Arg Gln	
970 975 980	
gga gcc acg ccg ctg gtg gtg gtg gaa ggt tct cgt gtg ctg ggc gtt	3567
Gly Ala Thr Pro Leu Val Val Val Glu Gly Ser Arg Val Leu Gly Val	
985 990 995	
att gcg ctg aaa gat atc gtc aaa ggc ggt att aaa gag cgc ttc	3612
Ile Ala Leu Lys Asp Ile Val Lys Gly Gly Ile Lys Glu Arg Phe	
1000 1005 1010	
gcc cag ctg cgc aaa atg ggc att aaa acg gtg atg att acc ggc	3657
Ala Gln Leu Arg Lys Met Gly Ile Lys Thr Val Met Ile Thr Gly	
1015 1020 1025	
gat aac cgt ctg act gcc gcc gcg att gct gcg gaa gcg ggt gtc	3702
Asp Asn Arg Leu Thr Ala Ala Ala Ile Ala Ala Glu Ala Gly Val	
1030 1035 1040	
gat gat ttt ctc gcc gaa gcg aca ccg gag gcc aag ctg gca ttg	3747
Asp Asp Phe Leu Ala Glu Ala Thr Pro Glu Ala Lys Leu Ala Leu	
1045 1050 1055	
atc cgt cag tat cag gcg gaa ggt cgt ttg gta gcg atg acc ggc	3792
Ile Arg Gln Tyr Gln Ala Glu Gly Arg Leu Val Ala Met Thr Gly	
1060 1065 1070	

cgt aat ctc agc gtt gaa cag ctc acg caa ctg atc gca aaa tac	4784
Arg Asn Leu Ser Val Glu Gln Leu Thr Gln Leu Ile Ala Lys Tyr	
1390 1395 1400	
agc caa caa ccg ctg gtg aaa tat atc ggc cag ccg gtt gtc aac	4829
Ser Gln Gln Pro Leu Val Lys Tyr Ile Gly Gln Pro Val Val Asn	
1405 1410 1415	
att gtt gaa ctc aat ctg gcg ctg gat aaa ctt gat gaa taa	4871
Ile Val Glu Leu Asn Leu Ala Leu Asp Lys Leu Asp Glu	
1420 1425	
cgaaccctta cgtcccgacc ccgatcgtct gctggaacaa actgccgcgc cgcacgccc	4931
gaagctgaaa gttttcttcg gtgcctgtgc aggcgtcggg aagacctggg cgcacgccc	4991
agaagcccag cgactgcggg cgcaagggct ggatattgtg gttggcgtgg tagaaacca	5051
cgggcgaaaa gataccgccg ccatgctgga agggctggct gttctgccc taaaacgcca	5111
ggcgtaccgt gggcggcata tcagcagatt tgatctcgat gccgccctcg cccgcgcgcc	5171
ggcgtctgac ttaatggacg aactggcgca cagtaatgcg ccaggttccc gtcacccaa	5231
acgctggcag gatatcgaag aactgctgga agctggcatt gatgttttca ctaccgtcaa	5291
cgttcagcat ctggaaaagtc tgaatgatgt ggtcagcggc gtcaccggaa ttcaggtacg	5351
ggaaaaccgtg cccgatcctt ttttcgatgc cgccgacgac gtgggtgctgg tggacttgcc	5411
cccggacgat ctgcgccagc ggctgaaaaga aggcaaagtc tatattgccg ggcagggcga	5471
gcgcgccatt gaacattttt tccgcaaagg taatctgatc gccctgcgcg aactggcact	5531
gcgccgtact gccgatcgcg ttgatgagca aatgcgcgc tggcgggggc atcctggcga	5591
agagaaaagtg tggcacacgc gcgacgcgat ccttttatgc atcgccata acaccggcag	5651
cgaaaaactg gtccgcgcag cggcgcggct ggcgtcacgg ctgggtagcg tctggcacgc	5711
ggtgtatggt gaaaccctcg cctgcaccg cttaccggaa aaaaaacgtc gggcaattct	5771
cagcgctta gactcggcgc aggaactggg cgcgagacg gcaaaccttt ctgatccagc	5831
ggaagagaaa gcggtagtgc gttatgccc tgaacataat ctgggcaaga ttattctcgg	5891
tcgcccggcc tcgcccgcgt ggtggcgtcg ggaacgttt gctgaccgac tggcgcgcat	5951
cgccccgat ctcgatcagg tgctggctcg gcttgatgaa ccaccgccgc gcacgattaa	6011
caacgcgcgc gataaccgct ctttaaaaga caagtggcgt gtacaaattc agggatgcgt	6071
ggttgccgcc gcgttatgcg ccgttatcac ctttaattgcc atgcagtggc tgatggcgtt	6131
tgatgccgcc aacctgggta tgctgtatct gcttggcgtg gtgggtgggg cgctatttta	6191
tggacgctgg ccttcagtgg ttgccaccgt cattaatgta gtgagtttcg atctctttt	6251
tatcgcccca cgcggcacgc tcgccgtctc tgatgtgcaa tatctgctga ccttcgcggt	6311
gatgttaacc gtcgggctgg tgatcgggaa ccttactgct ggcgtgcgtt atcaggcgcg	6371
ggtagcccgt taccgcgagc aacgcacacg gcacttata gaaatgtcga aagctctggc	6431
ggtgggcccgc agtccgcagg atatcgtgc caccagcgaa caatttattg cctccacgtt	6491
tcatgcccgc agtcaggtgt tgttgcgca tgacaacggg aaattgcagc cgtaacaca	6551
tccgcaagga atgacgccgt gggacgatgc catcgcgcag tggagttttg ataaagcct	6611
gcctgcgggc gggggcaccg acacgttacc cgggtgaccg taccagattt tgcgctaaa	6671
aagcggcgag aaaacctacg ggctgggtgg ggtggagccg gggaaatctgc gccagttgat	6731
gatcccggaa cagcagcgcg tgctggagac gtttacgctg ttagtcgcca atgccctga	6791
gcggtgacg ctaaccgcca gcgaagaaca gcgcgggatg gcaagcgaac gtgaacagat	6851
ccgcaacgcc ctgctggcgc cgctttcgca tgatttacgc acgcccctta cgggtgctgtt	6911
tggtcagga gaaatcttaa cgctcgatct ggcaagcga ggatcacccc acgcccgcca	6971
ggccagcgag atccgctcagc atgtgctgaa cactaccgca ctggtgaata atctactgga	7031
tatggcgcga attcagtcgc gcggctttaa tttgaagaaa gagtgggtaa cgctggaaga	7091
agtagtcggc agcgcgctgc aaatgctgga accgggttta tcgtcgcca tcaatctttc	7151
tctgccagaa ccgctgacct taatccacgt tgacgggcca ctctttgaac ggggtgctgat	7211
taatctgctg gagaacgcgc tgaaatatgc ggggtgcgcag gccgaaattg gtatcgatgc	7271
ccacgttgag ggcgaaaatc tacaactgga tgtctgggat aacggccccg gtcttcgcgc	7331
aggccaggag cagacgatat ttgataagtt tgctcgcggg aataaagagt cggcagttac	7391
gggggtaggg cttggactgg caatttgcgc ggcgatagtg gatgtacacg ggggcaactat	7451
taccgcgttc aaccgaccgc aaggtggtgc ctgttttctg gttacacttc cccagcaaac	7511
tgcccctgaa cttgaagaat ttcattgagga tatgtgacaa acgttctgat tgttgaagat	7571
gaacaggcta ttcgctgcct tctgcgcacg gcgctggagg gcgacgggat gcgctcttt	7631
gaggccgaaa cgctgcaacg cggtttgctg gaagcggcaa cccgtaagcc agatttgatt	7691
attctcgatc tcggcctgcc cgatggtgat gggattgagt ttatccgcga cctgcgccag	7751
tggagcgcgc tgccggtgat tgtgctttcc gcacgcgcg aagagagcga caaaatcgcc	7811
gcgctggatg ccggagcggg tgattatctg agtaagccgt ttggcattgg cgaattgcag	7871
gcccgtctgc gcgtcgcat acgcccacac tctgaccaca ccgcgcccga tccgctggta	7931
aaattttccg atgttaccgt cgatttagcc ccccgcgtga ttcaccgggg tgaggaagag	7991
gtgatctca caccaattga gttccgctcg ctggcggctg tgctcaacaa tgccggaaaa	8051
gtactcacc agcgcagct ccttaaccag gtgtgggggc caaacgcggt cgaacacagt	8111

```

cactatttgc gtatttatat gggacatctg cgacaaaaac tggaacagga tcccgccgcg 8171
ccacgccatt tcattactga aaccggtatt ggctatcggg ttatgctttg aatattaatt 8231
ttaatacagc ctgcctttta ttaattaaag ccgtaataat aaatacggct ttttatctta 8291
aacaacacac aaaaataaca attcaatatt ttatattact gagtaaaaag ctcatcatta 8351
aataaattaa gattgatcat tttttattga tcaccttcac agttcaaccg tttttcctgg 8411
atagaatata ttcaccttca ttcacatcag gaaaggtaaa ttaaattggaa aataacagcc 8471
gcactatgcc ccatataagg cggacaactc 8501

```

```

<210> 2
<211> 557
<212> PRT
<213> Escherichia coli

```

```

<400> 2
Met Ala Ala Gln Gly Phe Leu Leu Ile Ala Thr Phe Leu Leu Val Leu
1 5 10 15
Met Val Leu Ala Arg Pro Leu Gly Ser Gly Leu Ala Arg Leu Ile Asn
20 25 30
Asp Ile Pro Leu Pro Gly Thr Thr Gly Val Glu Arg Val Leu Phe Arg
35 40 45
Ala Leu Gly Val Ser Asp Arg Glu Met Asn Trp Lys Gln Tyr Leu Cys
50 55 60
Ala Ile Leu Gly Leu Asn Met Leu Gly Leu Ala Val Leu Phe Phe Met
65 70 75 80
Leu Leu Gly Gln His Tyr Leu Pro Leu Asn Pro Gln Gln Leu Pro Gly
85 90 95
Leu Ser Trp Asp Leu Ala Leu Asn Thr Ala Val Ser Phe Val Thr Asn
100 105 110
Thr Asn Trp Gln Ser Tyr Ser Gly Glu Thr Thr Leu Ser Tyr Phe Ser
115 120 125
Gln Met Ala Gly Leu Thr Val Gln Asn Phe Leu Ser Ala Ala Ser Gly
130 135 140
Ile Ala Val Ile Phe Ala Leu Ile Arg Ala Phe Thr Arg Gln Ser Met
145 150 155 160
Ser Thr Leu Gly Asn Ala Trp Val Asp Leu Leu Arg Ile Thr Leu Trp
165 170 175
Val Leu Val Pro Val Ala Leu Leu Ile Ala Leu Phe Phe Ile Gln Gln
180 185 190
Gly Ala Leu Gln Asn Phe Leu Pro Tyr Gln Ala Val Asn Thr Val Glu
195 200 205
Gly Ala Gln Gln Leu Leu Pro Met Gly Pro Val Ala Ser Gln Glu Ala
210 215 220
Ile Lys Met Leu Gly Thr Asn Gly Gly Gly Phe Phe Asn Ala Asn Ser
225 230 235 240
Ser His Pro Phe Glu Asn Pro Thr Ala Leu Thr Asn Phe Val Gln Met
245 250 255
Leu Ala Ile Phe Leu Ile Pro Thr Ala Leu Cys Phe Ala Phe Gly Glu
260 265 270
Val Met Gly Asp Arg Arg Gln Gly Arg Met Leu Leu Trp Ala Met Ser
275 280 285
Val Ile Phe Val Ile Cys Val Gly Val Val Met Trp Ala Glu Val Gln
290 295 300
Gly Asn Pro His Leu Leu Ala Leu Gly Thr Asp Ser Ser Ile Asn Met
305 310 315 320
Glu Gly Lys Glu Ser Arg Phe Gly Val Leu Val Ser Ser Leu Phe Ala
325 330 335
Val Val Thr Thr Ala Ala Ser Cys Gly Ala Val Ile Ala Met His Asp
340 345 350
Ser Phe Thr Ala Leu Gly Gly Met Val Pro Met Trp Leu Met Gln Ile
355 360 365
Gly Glu Val Val Phe Gly Gly Val Gly Ser Gly Leu Tyr Gly Met Met
370 375 380

```

Leu Phe Val Leu Leu Ala Val Phe Ile Ala Gly Leu Met Ile Gly Arg
 385 390 395 400
 Thr Pro Glu Tyr Leu Gly Lys Lys Ile Asp Val Arg Glu Met Lys Leu
 405 410 415
 Thr Ala Leu Ala Ile Leu Val Thr Pro Thr Leu Val Leu Met Gly Ala
 420 425 430
 Ala Leu Ala Met Met Thr Asp Ala Gly Arg Ser Ala Met Leu Asn Pro
 435 440 445
 Gly Pro His Gly Phe Ser Glu Val Leu Tyr Ala Val Ser Ser Ala Ala
 450 455 460
 Asn Asn Asn Gly Ser Ala Phe Ala Gly Leu Ser Ala Asn Ser Pro Phe
 465 470 475 480
 Trp Asn Cys Leu Leu Ala Phe Cys Met Phe Val Gly Arg Phe Gly Val
 485 490 495
 Ile Ile Pro Val Met Ala Ile Ala Gly Ser Leu Val Ser Lys Lys Ser
 500 505 510
 Gln Ala Ala Ser Ser Gly Thr Leu Pro Thr His Gly Pro Leu Phe Val
 515 520 525
 Gly Leu Leu Ile Gly Thr Val Leu Leu Val Gly Ala Leu Thr Phe Ile
 530 535 540
 Pro Ala Leu Ala Leu Gly Pro Val Ala Glu Tyr Leu Ser
 545 550 555

<210> 3
 <211> 682
 <212> PRT
 <213> Escherichia coli

<400> 3
 Met Ser Arg Lys Gln Leu Ala Leu Phe Glu Pro Thr Leu Val Val Gln
 1 5 10 15
 Ala Leu Lys Glu Ala Val Lys Lys Leu Asn Pro Gln Ala Gln Trp Arg
 20 25 30
 Asn Pro Val Met Phe Ile Val Trp Ile Gly Ser Leu Leu Thr Thr Cys
 35 40 45
 Ile Ser Ile Ala Met Ala Ser Gly Ala Met Pro Gly Asn Ala Leu Phe
 50 55 60
 Ser Ala Ala Ile Ser Gly Trp Leu Trp Ile Thr Val Leu Phe Ala Asn
 65 70 75 80
 Phe Ala Glu Ala Leu Ala Glu Gly Arg Ser Lys Ala Gln Ala Asn Ser
 85 90 95
 Leu Lys Gly Val Lys Lys Thr Ala Phe Ala Arg Lys Leu Arg Glu Pro
 100 105 110
 Lys Tyr Gly Ala Ala Ala Asp Lys Val Pro Ala Asp Gln Leu Arg Lys
 115 120 125
 Gly Asp Ile Val Leu Val Glu Ala Gly Asp Ile Ile Pro Cys Asp Gly
 130 135 140
 Glu Val Ile Glu Gly Gly Ala Ser Val Asp Glu Ser Ala Ile Thr Gly
 145 150 155 160
 Glu Ser Ala Pro Val Ile Arg Glu Ser Gly Gly Asp Phe Ala Ser Val
 165 170 175
 Thr Gly Gly Thr Arg Ile Leu Ser Asp Trp Leu Val Ile Glu Cys Ser
 180 185 190
 Val Asn Pro Gly Glu Thr Phe Leu Asp Arg Met Ile Ala Met Val Glu
 195 200 205
 Gly Ala Gln Arg Arg Lys Thr Pro Asn Glu Ile Ala Leu Thr Ile Leu
 210 215 220
 Leu Ile Ala Leu Thr Ile Val Phe Leu Leu Ala Thr Ala Thr Leu Trp
 225 230 235 240
 Pro Phe Ser Ala Trp Gly Gly Asn Ala Val Ser Val Thr Val Leu Val
 245 250 255
 Ala Leu Leu Val Cys Leu Ile Pro Thr Thr Ile Gly Gly Leu Leu Ser

			260						265						270
Ala	Ile	Gly	Val	Ala	Gly	Met	Ser	Arg	Met	Leu	Gly	Ala	Asn	Val	Ile
		275					280					285			
Ala	Thr	Ser	Gly	Arg	Ala	Val	Glu	Ala	Ala	Gly	Asp	Val	Asp	Val	Leu
	290					295					300				
Leu	Leu	Asp	Lys	Thr	Gly	Thr	Ile	Thr	Leu	Gly	Asn	Arg	Gln	Ala	Ser
305					310					315					320
Glu	Phe	Ile	Pro	Ala	Gln	Gly	Val	Asp	Glu	Lys	Thr	Leu	Ala	Asp	Ala
				325					330					335	
Ala	Gln	Leu	Ala	Ser	Leu	Ala	Asp	Glu	Thr	Pro	Glu	Gly	Arg	Ser	Ile
			340					345					350		
Val	Ile	Leu	Ala	Lys	Gln	Arg	Phe	Asn	Leu	Arg	Glu	Arg	Asp	Val	Gln
	355						360					365			
Ser	Leu	His	Ala	Thr	Phe	Val	Pro	Phe	Thr	Ala	Gln	Ser	Arg	Met	Ser
	370					375						380			
Gly	Ile	Asn	Ile	Asp	Asn	Arg	Met	Ile	Arg	Lys	Gly	Ser	Val	Asp	Ala
385					390					395					400
Ile	Arg	Arg	His	Val	Glu	Ala	Asn	Gly	Gly	His	Phe	Pro	Thr	Asp	Val
				405					410					415	
Asp	Gln	Lys	Val	Asp	Gln	Val	Ala	Arg	Gln	Gly	Ala	Thr	Pro	Leu	Val
			420					425					430		
Val	Val	Glu	Gly	Ser	Arg	Val	Leu	Gly	Val	Ile	Ala	Leu	Lys	Asp	Ile
	435						440					445			
Val	Lys	Gly	Gly	Ile	Lys	Glu	Arg	Phe	Ala	Gln	Leu	Arg	Lys	Met	Gly
	450					455					460				
Ile	Lys	Thr	Val	Met	Ile	Thr	Gly	Asp	Asn	Arg	Leu	Thr	Ala	Ala	Ala
465					470					475					480
Ile	Ala	Ala	Glu	Ala	Gly	Val	Asp	Asp	Phe	Leu	Ala	Glu	Ala	Thr	Pro
				485					490					495	
Glu	Ala	Lys	Leu	Ala	Leu	Ile	Arg	Gln	Tyr	Gln	Ala	Glu	Gly	Arg	Leu
			500					505					510		
Val	Ala	Met	Thr	Gly	Asp	Gly	Thr	Asn	Asp	Ala	Pro	Ala	Leu	Ala	Gln
		515					520					525			
Ala	Asp	Val	Ala	Val	Ala	Met	Asn	Ser	Gly	Thr	Gln	Ala	Ala	Lys	Glu
	530					535					540				
Ala	Gly	Asn	Met	Val	Asp	Leu	Asp	Ser	Asn	Pro	Thr	Lys	Leu	Ile	Glu
545					550					555					560
Val	Val	His	Ile	Gly	Lys	Gln	Met	Leu	Met	Thr	Arg	Gly	Ser	Leu	Thr
				565					570					575	
Thr	Phe	Ser	Ile	Ala	Asn	Asp	Val	Ala	Lys	Tyr	Phe	Ala	Ile	Ile	Pro
			580					585					590		
Ala	Ala	Phe	Ala	Ala	Thr	Tyr	Pro	Gln	Leu	Asn	Ala	Leu	Asn	Ile	Met
		595					600					605			
Cys	Leu	His	Ser	Pro	Asp	Ser	Ala	Ile	Leu	Ser	Ala	Val	Ile	Phe	Asn
	610					615					620				
Ala	Leu	Ile	Ile	Val	Phe	Leu	Ile	Pro	Leu	Ala	Leu	Lys	Gly	Val	Ser
625					630					635					640
Tyr	Lys	Pro	Leu	Thr	Ala	Ser	Ala	Met	Leu	Arg	Arg	Asn	Leu	Trp	Ile
				645					650					655	
Tyr	Gly	Leu	Gly	Gly	Leu	Leu	Val	Pro	Phe	Ile	Gly	Ile	Lys	Val	Ile
			660					665					670		
Asp	Leu	Leu	Leu	Thr	Val	Cys	Gly	Leu	Val						
	675						680								

<210> 4
 <211> 190
 <212> PRT
 <213> Escherichia coli

<400> 4
 Met Ser Gly Leu Arg Pro Ala Leu Ser Thr Phe Ile Phe Leu Leu Leu
 1 5 10 15

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

<210> 5
 <211> 894
 <212> PRT
 <213> Escherichia coli

<400> 5
 Met Asn Asn Glu Pro Leu Arg Pro Asp Pro Asp Arg Leu Leu Glu Gln
 1 5 10 15
 Thr Ala Ala Pro His Arg Gly Lys Leu Lys Val Phe Phe Gly Ala Cys
 20 25 30
 Ala Gly Val Gly Lys Thr Trp Ala Met Leu Ala Glu Ala Gln Arg Leu
 35 40 45
 Arg Ala Gln Gly Leu Asp Ile Val Val Gly Val Val Glu Thr His Gly
 50 55 60
 Arg Lys Asp Thr Ala Ala Met Leu Glu Gly Leu Ala Val Leu Pro Leu
 65 70 75 80
 Lys Arg Gln Ala Tyr Arg Gly Arg His Ile Ser Glu Phe Asp Leu Asp
 85 90 95
 Ala Ala Leu Ala Arg Arg Pro Ala Leu Ile Leu Met Asp Glu Leu Ala
 100 105 110
 His Ser Asn Ala Pro Gly Ser Arg His Pro Lys Arg Trp Gln Asp Ile
 115 120 125
 Glu Glu Leu Leu Glu Ala Gly Ile Asp Val Phe Thr Thr Val Asn Val
 130 135 140
 Gln His Leu Glu Ser Leu Asn Asp Val Val Ser Gly Val Thr Gly Ile
 145 150 155 160
 Gln Val Arg Glu Thr Val Pro Asp Pro Phe Phe Asp Ala Ala Asp Asp
 165 170 175
 Val Val Leu Val Asp Leu Pro Pro Asp Asp Leu Arg Gln Arg Leu Lys
 180 185 190
 Glu Gly Lys Val Tyr Ile Ala Gly Gln Ala Glu Arg Ala Ile Glu His
 195 200 205
 Phe Phe Arg Lys Gly Asn Leu Ile Ala Leu Arg Glu Leu Ala Leu Arg
 210 215 220
 Arg Thr Ala Asp Arg Val Asp Glu Gln Met Arg Ala Trp Arg Gly His
 225 230 235 240
 Pro Gly Glu Glu Lys Val Trp His Thr Arg Asp Ala Ile Leu Leu Cys
 245 250 255
 Ile Gly His Asn Thr Gly Ser Glu Lys Leu Val Arg Ala Ala Ala Arg

Thr Leu Ile His Val Asp Gly Pro Leu Phe Glu Arg Val Leu Ile Asn
 770 775 780
 Leu Leu Glu Asn Ala Val Lys Tyr Ala Gly Ala Gln Ala Glu Ile Gly
 785 790 795 800
 Ile Asp Ala His Val Glu Gly Glu Asn Leu Gln Leu Asp Val Trp Asp
 805 810 815
 Asn Gly Pro Gly Leu Pro Pro Gly Gln Glu Gln Thr Ile Phe Asp Lys
 820 825 830
 Phe Ala Arg Gly Asn Lys Glu Ser Ala Val Pro Gly Val Gly Leu Gly
 835 840 845
 Leu Ala Ile Cys Arg Ala Ile Val Asp Val His Gly Gly Thr Ile Thr
 850 855 860
 Ala Phe Asn Arg Pro Glu Gly Gly Ala Cys Phe Arg Val Thr Leu Pro
 865 870 875 880
 Gln Gln Thr Ala Pro Glu Leu Glu Glu Phe His Glu Asp Met
 885 890

<210> 6
 <211> 225
 <212> PRT
 <213> Escherichia coli

<400> 6
 Val Thr Asn Val Leu Ile Val Glu Asp Glu Gln Ala Ile Arg Arg Phe
 1 5 10 15
 Leu Arg Thr Ala Leu Glu Gly Asp Gly Met Arg Val Phe Glu Ala Glu
 20 25 30
 Thr Leu Gln Arg Gly Leu Leu Glu Ala Ala Thr Arg Lys Pro Asp Leu
 35 40 45
 Ile Ile Leu Asp Leu Gly Leu Pro Asp Gly Asp Gly Ile Glu Phe Ile
 50 55 60
 Arg Asp Leu Arg Gln Trp Ser Ala Val Pro Val Ile Val Leu Ser Ala
 65 70 75 80
 Arg Ser Glu Glu Ser Asp Lys Ile Ala Ala Leu Asp Ala Gly Ala Asp
 85 90 95
 Asp Tyr Leu Ser Lys Pro Phe Gly Ile Gly Glu Leu Gln Ala Arg Leu
 100 105 110
 Arg Val Ala Leu Arg Arg His Ser Ala Thr Thr Ala Pro Asp Pro Leu
 115 120 125
 Val Lys Phe Ser Asp Val Thr Val Asp Leu Ala Ala Arg Val Ile His
 130 135 140
 Arg Gly Glu Glu Glu Val His Leu Thr Pro Ile Glu Phe Arg Leu Leu
 145 150 155 160
 Ala Val Leu Leu Asn Asn Ala Gly Lys Val Leu Thr Gln Arg Gln Leu
 165 170 175
 Leu Asn Gln Val Trp Gly Pro Asn Ala Val Glu His Ser His Tyr Leu
 180 185 190
 Arg Ile Tyr Met Gly His Leu Arg Gln Lys Leu Glu Gln Asp Pro Ala
 195 200 205
 Arg Pro Arg His Phe Ile Thr Glu Thr Gly Ile Gly Tyr Arg Phe Met
 210 215 220
 Leu
 225

<210> 7
 <211> 7620
 <212> DNA
 <213> Pantoea ananatis

<220>
 <221> CDS

<222> (543)..(2225)

<220>

<221> CDS

<222> (2228)..(4273)

<220>

<221> CDS

<222> (4284)..(4853)

<220>

<221> CDS

<222> (4867)..(7542)

<400> 7

```

aggctggacc ccagccaaaa ccctgcaaat gatggtgagc tggccggggc atttcgacag      60
taccttactg cacgcttttcg tcagtacgat agggatgtat cccgtcgggt ccctggttcg      120
cctggcgtct ggacgcattg cgctggtggt gaagggcggc gataagtcat tacagcgacc      180
tacgggtgat gtcttctggt cactgcacgc gcagcgggaa atcaaaccgc aggtgctgga      240
tctgggcgac agtttttgta ctgacagcat tgtgtgtgcc gaagataacg gccgttgagg      300
caacgtcgat ctgcgccgaa tctggttgcg ggacgccgcc tgatattgcg ggctgttttt      360
atacattttt tacatccccg acccgcattt taaccctttc tttatgtgcg ccagcgcaag      420
ctactcagca aatccaatcc tctggagggt gctgtgagtg caggcgtaat aaccggcatt      480
gtgctggttag tgttgttgct gggctatctg atctatgcc tgttaaattgc ggaggccttc      540
tg atg gcg gcc aat gcg ttt tta ctg atc gcg gtt tat ctg ctg ttg      587
  Met Ala Ala Asn Ala Phe Leu Leu Ile Ala Val Tyr Leu Leu Leu
    1             5             10            15
ctg atg gtg atg gcg caa ccg ctg ggg cgt ggg ctg gcc gcg ctg gtt      635
Leu Met Val Met Ala Gln Pro Leu Gly Arg Gly Leu Ala Ala Leu Val
             20             25             30
gcc gat aaa ccc ctc ttt gca cgt gct gaa gcc ctg ctg tgg cgt ttt      683
Ala Asp Lys Pro Leu Phe Ala Arg Ala Glu Ala Leu Leu Trp Arg Phe
             35             40             45
tcg ggt gta caa gaa ggc ggt atg cgc tgg cag cac tac ctg ctg gca      731
Ser Gly Val Gln Glu Gly Gly Met Arg Trp Gln His Tyr Leu Leu Ala
             50             55             60
att ttg gtg ttc aac ctg ctt ggc ttc gtg gtg ctg ctc gcc atc cta      779
Ile Leu Val Phe Asn Leu Leu Gly Phe Val Val Leu Leu Ala Ile Leu
             65             70             75
atg ttt cag gga gcg ttg ccg ctc aat ccg caa cat ctt ccc gga ctg      827
Met Phe Gln Gly Ala Leu Pro Leu Asn Pro Gln His Leu Pro Gly Leu
             80             85             90             95
agc tgg gat ttg gcg ctg aat acc gct atc agt ttt gtc acc aac acc      875
Ser Trp Asp Leu Ala Leu Asn Thr Ala Ile Ser Phe Val Thr Asn Thr
             100            105            110
aac tgg cag tct tat gcc ggt gaa agc acc ctg agt tac ttc agc cag      923
Asn Trp Gln Ser Tyr Ala Gly Glu Ser Thr Leu Ser Tyr Phe Ser Gln
             115            120            125
atg gtc ggg ctg acg gtg cag aac ttc gtt tcc gcc gcc acc ggc atc      971
Met Val Gly Leu Thr Val Gln Asn Phe Val Ser Ala Ala Thr Gly Ile
             130            135            140
gcc gtg gcg ttt gcg ctg att cgc ggt ttt gct aat cgt tcg gtg gca      1019
Ala Val Ala Phe Ala Leu Ile Arg Gly Phe Ala Asn Arg Ser Val Ala
             145            150            155
acc ctg ggc aac gcc tgg cgc gat tta acg cgc att aca ctc tat gtc      1067
Thr Leu Gly Asn Ala Trp Arg Asp Leu Thr Arg Ile Thr Leu Tyr Val
             160            165            170            175
ctg ttg ccg atc agc ctg ctg atg gcg ctg ttt ttt gtc agc cag ggc      1115
Leu Leu Pro Ile Ser Leu Leu Met Ala Leu Phe Phe Val Ser Gln Gly
             180            185            190
agc atc cag aac ttc ctg ccg tat cac aac gtc acc agc ctg gaa ggt      1163
Ser Ile Gln Asn Phe Leu Pro Tyr His Asn Val Thr Ser Leu Glu Gly
             195            200            205

```

gcg cag caa acg ctg gca atg ggg ccg gtt gcc tct cag gaa gcc atc	1211
Ala Gln Gln Thr Leu Ala Met Gly Pro Val Ala Ser Gln Glu Ala Ile	
210 215 220	
aaa atg ctg ggc acc aac ggc ggc ggc ttt ttc aac gtt aac tct gcg	1259
Lys Met Leu Gly Thr Asn Gly Gly Gly Phe Phe Asn Val Asn Ser Ala	
225 230 235	
cat ccg ttt gag aac cct acc gcg ctg agc aat ttc gta cag atg ctt	1307
His Pro Phe Glu Asn Pro Thr Ala Leu Ser Asn Phe Val Gln Met Leu	
240 245 250 255	
agt atc ttc ctg att cct gca gca ctc tgc ttt gcc ttt ggc gaa agc	1355
Ser Ile Phe Leu Ile Pro Ala Ala Leu Cys Phe Ala Phe Gly Glu Ser	
260 265 270	
gtt aaa gat cgg cgc cag ggc tca atg ttg ctc tgg tcc atg acg ttg	1403
Val Lys Asp Arg Arg Gln Gly Ser Met Leu Leu Trp Ser Met Thr Leu	
275 280 285	
atg ttt gtc gtg gct gct gcg ctg gtg atg tgg gct gaa cta cgt ggc	1451
Met Phe Val Val Ala Ala Ala Leu Val Met Trp Ala Glu Leu Arg Gly	
290 295 300	
aac ccg cac ttt ctg acg cta ggg gct gac agc gcc atc aat atg gaa	1499
Asn Pro His Phe Leu Thr Leu Gly Ala Asp Ser Ala Ile Asn Met Glu	
305 310 315	
ggc aaa gaa acg cgc ttc ggc att ctc aac tcc agc ctg ttt gcg gtg	1547
Gly Lys Glu Thr Arg Phe Gly Ile Leu Asn Ser Ser Leu Phe Ala Val	
320 325 330 335	
att acg acg gcg gcg tcc tgc ggt gcg gta aac gcg atg cat gac tcg	1595
Ile Thr Thr Ala Ala Ser Cys Gly Ala Val Asn Ala Met His Asp Ser	
340 345 350	
ttt acg gcg ctg ggc ggt atg gtg ccg atg ctg ctg atg caa ctg ggc	1643
Phe Thr Ala Leu Gly Gly Met Val Pro Met Leu Leu Met Gln Leu Gly	
355 360 365	
gag gtg gtg ttt ggc ggc gtg ggt gcc ggt ctg tac ggg atg ctg ctg	1691
Glu Val Val Phe Gly Gly Val Gly Ala Gly Leu Tyr Gly Met Leu Leu	
370 375 380	
ttt gtc tta ctg gcg gtg ttt att gcc ggg ttg atg att ggc cgc aca	1739
Phe Val Leu Leu Ala Val Phe Ile Ala Gly Leu Met Ile Gly Arg Thr	
385 390 395	
ccg gaa ttc ctc ggc aag aaa atc gac gta tgg gaa atg aaa atg acg	1787
Pro Glu Phe Leu Gly Lys Lys Ile Asp Val Trp Glu Met Lys Met Thr	
400 405 410 415	
gcc ctg gcg att ctg gtc acg ccc gcg ctg gtg ttg atc ggt acg gcg	1835
Ala Leu Ala Ile Leu Val Thr Pro Ala Leu Val Leu Ile Gly Thr Ala	
420 425 430	
att gcg atg atg acc gac gcc gga cgc gca ggt atg gca aac ccc gga	1883
Ile Ala Met Met Thr Asp Ala Gly Arg Ala Gly Met Ala Asn Pro Gly	
435 440 445	
acg cat ggc ttt agt gaa gtc ctg tat gcc gtt tcg tcg gcc gcc aat	1931
Thr His Gly Phe Ser Glu Val Leu Tyr Ala Val Ser Ser Ala Ala Asn	
450 455 460	
aac aat ggc agc gcc ttt gcg ggc ctg aac gcc aat acg ccg ttc tgg	1979
Asn Asn Gly Ser Ala Phe Ala Gly Leu Asn Ala Asn Thr Pro Phe Trp	
465 470 475	
aac ctg ctg ctg gcg gtg tgt atg ttc gta ggt cgc ttc ggc atc att	2027
Asn Leu Leu Leu Ala Val Cys Met Phe Val Gly Arg Phe Gly Ile Ile	
480 485 490 495	
att ccg gtc atg gcg att gcg ggg gca atg gcg gtg aaa aaa gtg cag	2075
Ile Pro Val Met Ala Ile Ala Gly Ala Met Ala Val Lys Lys Val Gln	
500 505 510	
ccg gta ggt aac ggc acg ctc cct acg cac ggt ccg ctg ttt atc gca	2123
Pro Val Gly Asn Gly Thr Leu Pro Thr His Gly Pro Leu Phe Ile Ala	
515 520 525	
ctg ctg gtc ggt acc gtc ttg ttg gtc ggc gcg ctg acc ttt att cct	2171
Leu Leu Val Gly Thr Val Leu Leu Val Gly Ala Leu Thr Phe Ile Pro	
530 535 540	

gct ctg gcg ctg ggt ccg gtc gcc gag cac ctg caa ctt att cag gga	2219
Ala Leu Ala Leu Gly Pro Val Ala Glu His Leu Gln Leu Ile Gln Gly	
545 550 555	
caa taa tc atg agt cgt caa caa cag gtg ttt gac gca gcg ctg tta	2266
Gln Met Ser Arg Gln Gln Gln Val Phe Asp Ala Ala Leu Leu	
560 565 570	
cgt acc tca gcg atc gat gcg gta aaa aaa ctc gat cct cgc gtg cag	2314
Arg Thr Ser Ala Ile Asp Ala Val Lys Lys Leu Asp Pro Arg Val Gln	
575 580 585	
ttt cgc aat ccg gtc atg ttt gtg gtt tac ctg ggc agt atc ctg acc	2362
Phe Arg Asn Pro Val Met Phe Val Val Tyr Leu Gly Ser Ile Leu Thr	
590 595 600 605	
tcg att ctg gcc ata atg atg ttt acc gga cac cag agc ggc agc gcc	2410
Ser Ile Leu Ala Ile Met Met Phe Thr Gly His Gln Ser Gly Ser Ala	
610 615 620	
agc ttt acc ggc gcg att gcc ctg tgg tta tgg ttc acc gtg ctg ttt	2458
Ser Phe Thr Gly Ala Ile Ala Leu Trp Leu Trp Phe Thr Val Leu Phe	
625 630 635	
gcc aac atg gca gaa gcc ctg gcg gaa ggg cgc agt aaa gcc cag gca	2506
Ala Asn Met Ala Glu Ala Leu Ala Glu Gly Arg Ser Lys Ala Gln Ala	
640 645 650	
aac agc ctg aaa ggc gtt aaa aag acc agc ttc gcc aaa aaa ctg tcg	2554
Asn Ser Leu Lys Gly Val Lys Lys Thr Ser Phe Ala Lys Lys Leu Ser	
655 660 665	
gcg gcg cac tac ggt gca gcg tgg cag cag gtg gcg gcc gat gcg ctg	2602
Ala Ala His Tyr Gly Ala Ala Trp Gln Gln Val Ala Ala Asp Ala Leu	
670 675 680 685	
cgt aaa ggg gat gcc gtg ctg gta gag gcc ggt gat gtg atc ccc tgc	2650
Arg Lys Gly Asp Ala Val Leu Val Glu Ala Gly Asp Val Ile Pro Cys	
690 695 700	
gac ggt gaa gtc gtg gaa ggg ggc gca tcg gta gac gag agc gcg atc	2698
Asp Gly Glu Val Val Glu Gly Gly Ala Ser Val Asp Glu Ser Ala Ile	
705 710 715	
acc ggt gaa tcg gca ccg gtg atc cgt gaa tcg ggc ggg gat ttc gcc	2746
Thr Gly Glu Ser Ala Pro Val Ile Arg Glu Ser Gly Gly Asp Phe Ala	
720 725 730	
tcg gtg acc ggc ggg aca cgc att ctg tct gac tgg ctg gtc att acc	2794
Ser Val Thr Gly Gly Thr Arg Ile Leu Ser Asp Trp Leu Val Ile Thr	
735 740 745	
tgc agc gcc aac cca ggc gaa acc ttc ctg gac cgg atg atc gcc atg	2842
Cys Ser Ala Asn Pro Gly Glu Thr Phe Leu Asp Arg Met Ile Ala Met	
750 755 760 765	
gtc gaa ggc gca cag cgt cgt aaa acc ccg aat gag att gcc ctg acc	2890
Val Glu Gly Ala Gln Arg Arg Lys Thr Pro Asn Glu Ile Ala Leu Thr	
770 775 780	
att ctg ctg gtg tcg ctc acc att gtg ttt ctg tta gcc acc gtc acg	2938
Ile Leu Leu Val Ser Leu Thr Ile Val Phe Leu Leu Ala Thr Val Thr	
785 790 795	
ctg tgg cct ttt tca gcc tgg ggc ggc acg ccg gtc acc atc acc gta	2986
Leu Trp Pro Phe Ser Ala Trp Gly Gly Thr Pro Val Thr Ile Thr Val	
800 805 810	
ctg gtg gcg ctg ctg gta tgc ctg atc ccg acc acc att ggc ggt ctg	3034
Leu Val Ala Leu Leu Val Cys Leu Ile Pro Thr Thr Ile Gly Gly Leu	
815 820 825	
ctg tcc gct atc ggc gtg gcc ggg atg agc cgg atg ctg ggc gct aac	3082
Leu Ser Ala Ile Gly Val Ala Gly Met Ser Arg Met Leu Gly Ala Asn	
830 835 840 845	
gtc att gcc acc agt ggc cgc gct gtt gaa gcc gct ggc gac gtg gat	3130
Val Ile Ala Thr Ser Gly Arg Ala Val Glu Ala Ala Gly Asp Val Asp	
850 855 860	
gtg ctg atg ctg gat aaa acc ggc acc atc acg ctg ggt aac cgt cag	3178
Val Leu Met Leu Asp Lys Thr Gly Thr Ile Thr Leu Gly Asn Arg Gln	
865 870 875	

gca acg cag ttt tta ccg gct ccc ggc gtc acg gaa gaa cag ctg gcg Ala Thr Gln Phe Leu Pro Ala Pro Gly Val Thr Glu Glu Gln Leu Ala 880 885 890	3226
gat gcg gcg cag ctg gcg tcc ctg gcg gat gaa acg ccg gaa ggg cgc Asp Ala Ala Gln Leu Ala Ser Leu Ala Asp Glu Thr Pro Glu Gly Arg 895 900 905	3274
agc atc gtg gtg ctg gcg aag caa aag ttt aac ctg cgt gaa cgt gac Ser Ile Val Val Leu Ala Lys Gln Lys Phe Asn Leu Arg Glu Arg Asp 910 915 920 925	3322
ctg agc agc atg ggc gcc agc ttt att ccc ttc tcg gct caa acc cgt Leu Ser Ser Met Gly Ala Ser Phe Ile Pro Phe Ser Ala Gln Thr Arg 930 935 940	3370
atg agc ggc gtc aac gta cag gac cgc ctg atc cgt aaa ggt gcg gtc Met Ser Gly Val Asn Val Gln Asp Arg Leu Ile Arg Lys Gly Ala Val 945 950 955	3418
gat gcg gtg cgc cgt cat att gaa gcc agc cac ggt gcc ttt ccg gct Asp Ala Val Arg Arg His Ile Glu Ala Ser His Gly Ala Phe Pro Ala 960 965 970	3466
gag gtg aac gcc cgg gtt gaa gag gtg gcg cgg gcc ggt ggc aca ccg Glu Val Asn Ala Arg Val Glu Glu Val Ala Arg Ala Gly Gly Thr Pro 975 980 985	3514
ctg gtg gtg gcg gaa ggc gca aag gtg ctg ggc gtg gtg gcg cta aaa Leu Val Val Ala Glu Gly Ala Lys Val Leu Gly Val Val Ala Leu Lys 990 995 1000 1005	3562
gat atc gtt aaa ggt ggc atc aaa gaa cgt ttt gcg gaa ctg cgc Asp Ile Val Lys Gly Gly Ile Lys Glu Arg Phe Ala Glu Leu Arg 1010 1015 1020	3607
aag atg ggc att aaa acc gtg atg atc acc ggt gat aac ccg ctc Lys Met Gly Ile Lys Thr Val Met Ile Thr Gly Asp Asn Pro Leu 1025 1030 1035	3652
acc gcc gcc gcc atc gcg gca gag gca ggg gtg gat gac ttt ctg Thr Ala Ala Ala Ile Ala Ala Glu Ala Gly Val Asp Asp Phe Leu 1040 1045 1050	3697
tca gaa gcg acg ccg gaa gcc aag ctg gca ctg att cgt cag tat Ser Glu Ala Thr Pro Glu Ala Lys Leu Ala Leu Ile Arg Gln Tyr 1055 1060 1065	3742
cag gcg gag ggg cgc tta gtc gcg atg acc ggt gac ggc acc aat Gln Ala Glu Gly Arg Leu Val Ala Met Thr Gly Asp Gly Thr Asn 1070 1075 1080	3787
gac gcg cca gcg ttg gcc cag gcg gac gtg gcg gtc gcc atg aac Asp Ala Pro Ala Leu Ala Gln Ala Asp Val Ala Val Ala Met Asn 1085 1090 1095	3832
tcg ggt act cag gcg gcg aaa gaa gcg ggc aac atg gtc gat ctg Ser Gly Thr Gln Ala Ala Lys Glu Ala Gly Asn Met Val Asp Leu 1100 1105 1110	3877
gac tcg aac ccc acc aaa ctg ctg gaa gtg gtg cac att ggt aag Asp Ser Asn Pro Thr Lys Leu Leu Glu Val Val His Ile Gly Lys 1115 1120 1125	3922
cag atg ctg atg acg cgc ggt tcc ttg acc acc ttc agt atc gcc Gln Met Leu Met Thr Arg Gly Ser Leu Thr Thr Phe Ser Ile Ala 1130 1135 1140	3967
aat gac gtt gcc aag tat ttc gcc atc atc ccg gcc gcc ttt gct Asn Asp Val Ala Lys Tyr Phe Ala Ile Ile Pro Ala Ala Phe Ala 1145 1150 1155	4012
gca acc tat cca cag tta aat atg ctc aac gtg atg cag ctg cac Ala Thr Tyr Pro Gln Leu Asn Met Leu Asn Val Met Gln Leu His 1160 1165 1170	4057
tcg ccc gca tcg gcc atc ctg tcg gcc gtg att ttt aac gcc cta Ser Pro Ala Ser Ala Ile Leu Ser Ala Val Ile Phe Asn Ala Leu 1175 1180 1185	4102
gtg att gta ttc ctg atc cct ctg gcg ctt aaa ggt gtc agc tat Val Ile Val Phe Leu Ile Pro Leu Ala Leu Lys Gly Val Ser Tyr 1190 1195 1200	4147

cgc ccc ttg agt gcc gca tcg ctg ttg cgc cgt aat tta ctg att	4192
Arg Pro Leu Ser Ala Ala Ser Leu Leu Arg Arg Asn Leu Leu Ile	
1205 1210 1215	
tat ggg tta ggt ggg ctg ctg gtg ccc ttt gtc ggc atc aaa gcg	4237
Tyr Gly Leu Gly Gly Leu Leu Val Pro Phe Val Gly Ile Lys Ala	
1220 1225 1230	
att gat atg ttg ctc gtg ctg tct ggt atg gcc tga ggagattaaa atg	4286
Ile Asp Met Leu Leu Val Leu Ser Gly Met Ala Met	
1235 1240	
agt cag tta cgt ccg gcg att ttc ctg ctt ttg ctg cta acg gtt	4331
Ser Gln Leu Arg Pro Ala Ile Phe Leu Leu Leu Leu Leu Thr Val	
1245 1250 1255	
gtg tgc ggc gtc gtt tat cct ttg ctt acc acg gga ctg tcg caa	4376
Val Cys Gly Val Val Tyr Pro Leu Leu Thr Thr Gly Leu Ser Gln	
1260 1265 1270	
ctg ctg ttt ccc tgg cag gct aac ggg tca gta ttg aat gtc gat	4421
Leu Leu Phe Pro Trp Gln Ala Asn Gly Ser Val Leu Asn Val Asp	
1275 1280 1285	
ggc gaa gaa cgg gga tca gcg ctg att ggt cag aat ttt agc cag	4466
Gly Glu Glu Arg Gly Ser Ala Leu Ile Gly Gln Asn Phe Ser Gln	
1290 1295 1300	
ccc ggt tat ttc tgg ggg cgt cct tct gca acc ggt gat aag cct	4511
Pro Gly Tyr Phe Trp Gly Arg Pro Ser Ala Thr Gly Asp Lys Pro	
1305 1310 1315	
tat aat cct ctg gcc tcc agc ggc agc aac ctg gcg gcc agc aac	4556
Tyr Asn Pro Leu Ala Ser Ser Gly Ser Asn Leu Ala Ala Ser Asn	
1320 1325 1330	
ccg gcg ctg gat aag gct gta gcc gag cgc gtg gcg gct ttg cgc	4601
Pro Ala Leu Asp Lys Ala Val Ala Glu Arg Val Ala Ala Leu Arg	
1335 1340 1345	
acg gcg aat ccg cag gct aac ggg gcg gta ccg gtc gag ctg gta	4646
Thr Ala Asn Pro Gln Ala Asn Gly Ala Val Pro Val Glu Leu Val	
1350 1355 1360	
acc acc tcg gcc agt gga ctg gac ccg gag att tcg cct gag gct	4691
Thr Thr Ser Ala Ser Gly Leu Asp Pro Glu Ile Ser Pro Glu Ala	
1365 1370 1375	
gcg ctg tgg cag gca ccg cgt atc gcg gcg gca cgt cag ctg ccg	4736
Ala Leu Trp Gln Ala Pro Arg Ile Ala Ala Ala Arg Gln Leu Pro	
1380 1385 1390	
ctg gcg aaa gtg gat gcc ctg gta gac agc atg acg cag cgc ccg	4781
Leu Ala Lys Val Asp Ala Leu Val Asp Ser Met Thr Gln Arg Pro	
1395 1400 1405	
ctg ctg ccc ttt atc ggc gaa ccg act gtc aat gtg ctg cag ctt	4826
Leu Leu Pro Phe Ile Gly Glu Pro Thr Val Asn Val Leu Gln Leu	
1410 1415 1420	
aat ctg gcg ctg aac gac ctc aaa taa ctgtaaggat gct atg aac cac	4875
Asn Leu Ala Leu Asn Asp Leu Lys Met Asn His	
1425 1430	
gaa ccg ctg cgc ccc gat ccg gat gcg ctg ctg cag acc agc agc	4920
Glu Pro Leu Arg Pro Asp Pro Asp Ala Leu Leu Gln Thr Ser Ser	
1435 1440 1445	
gac agc cat cgc ggc aaa ctg aaa atc tat ttt ggc gcc tgt gcg	4965
Asp Ser His Arg Gly Lys Leu Lys Ile Tyr Phe Gly Ala Cys Ala	
1450 1455 1460	
ggc gta gga aaa acc tat gcc atg ttg cag gag gcg cag ccg ctg	5010
Gly Val Gly Lys Thr Tyr Ala Met Leu Gln Glu Ala Gln Arg Leu	
1465 1470 1475	
cgt gcc cag ggg ctg gat gtg ctg gtg ggc gta gtg gaa acg cac	5055
Arg Ala Gln Gly Leu Asp Val Leu Val Gly Val Val Glu Thr His	
1480 1485 1490	
gaa cgt ccg gaa aca gcg cag ctt ctt aac gga ctg gtg ctg ttg	5100
Glu Arg Pro Glu Thr Ala Gln Leu Leu Asn Gly Leu Val Leu Leu	
1495 1500 1505	

ccg cgc	cgg gcg acg ggc cgt	tcg cgg cat gcg gag	ttc gac ctt	5145
Pro Arg	Arg Ala Thr Gly Arg	Ser Arg His Ala Glu	Phe Asp Leu	
1510	1515	1520		
gat gcc	gcg ctg gcg cgc cat	ccg gca gta att ttg	atg gat gag	5190
Asp Ala	Ala Leu Ala Arg His	Pro Ala Val Ile Leu	Met Asp Glu	
1525	1530	1535		
ctg gcg	cac acg aac gtg aag	ggc tca cgt cat ccc	aaa cgc tgg	5235
Leu Ala	His Thr Asn Val Lys	Gly Ser Arg His Pro	Lys Arg Trp	
1540	1545	1550		
cag gat	att gag gaa ctg ctg	gag gcg ggc att gat	gtc ctg acg	5280
Gln Asp	Ile Glu Glu Leu Leu	Glu Ala Gly Ile Asp	Val Leu Thr	
1555	1560	1565		
aca gtg	aat gtt cag cat ctg	gaa agt ctg aat gat	gtg gtc ggt	5325
Thr Val	Asn Val Gln His Leu	Glu Ser Leu Asn Asp	Val Val Gly	
1570	1575	1580		
ggc gtc	acc ggc att cag gtg	cgt gaa acc gtt ccc	gat ccc ttt	5370
Gly Val	Thr Gly Ile Gln Val	Arg Glu Thr Val Pro	Asp Pro Phe	
1585	1590	1595		
ttc gac	gct gcc gat gaa gtg	gta ctg gtt gat ctc	ccg cct gac	5415
Phe Asp	Ala Ala Asp Glu Val	Val Leu Val Asp Leu	Pro Pro Asp	
1600	1605	1610		
gat ctc	cgc cag cgc ctg aaa	gag ggc aag gtc tac	att ggc gat	5460
Asp Leu	Arg Gln Arg Leu Lys	Glu Gly Lys Val Tyr	Ile Gly Asp	
1615	1620	1625		
cgt gcc	gaa cgc gcc atc gaa	aat ttc ttt cgt aag	ggc aac ctg	5505
Arg Ala	Glu Arg Ala Ile Glu	Asn Phe Phe Arg Lys	Gly Asn Leu	
1630	1635	1640		
tat gcc	ctg cgt gag ctg gcg	ctg cgc cgc act gcc	gac cgg gtc	5550
Tyr Ala	Leu Arg Glu Leu Ala	Leu Arg Arg Thr Ala	Asp Arg Val	
1645	1650	1655		
gat gac	cag atg cgc gcc tgg	cgc gac agt caa ggc	cgc gat cgg	5595
Asp Asp	Gln Met Arg Ala Trp	Arg Asp Ser Gln Gly	Arg Asp Arg	
1660	1665	1670		
gtc tgg	cac acg cgt gat gcc	att tta ttg tgt att	ggg gac gat	5640
Val Trp	His Thr Arg Asp Ala	Ile Leu Leu Cys Ile	Gly Asp Asp	
1675	1680	1685		
acc ggc	agt gaa aaa ctg gtg	cgg acg gcg gcg cgt	ctg gcc gcc	5685
Thr Gly	Ser Glu Lys Leu Val	Arg Thr Ala Ala Arg	Leu Ala Ala	
1690	1695	1700		
agg ctg	ggc agc gaa tgg cat	gcc gtt tac gtg gaa	acg ccc cgg	5730
Arg Leu	Gly Ser Glu Trp His	Ala Val Tyr Val Glu	Thr Pro Arg	
1705	1710	1715		
ctt aac	cgg cta ccg gaa gcg	cgg cgt cgg gcc att	tta cgc acg	5775
Leu Asn	Arg Leu Pro Glu Ala	Arg Arg Arg Ala Ile	Leu Arg Thr	
1720	1725	1730		
ctg aag	ctg gcg cag gat atg	ggg gcg gag acg gcg	acg ctg tcc	5820
Leu Lys	Leu Ala Gln Asp Met	Gly Ala Glu Thr Ala	Thr Leu Ser	
1735	1740	1745		
gac cct	gat gag gcg cag gcg	gtc ctg cgt tac gcg	cgg gaa cat	5865
Asp Pro	Asp Glu Ala Gln Ala	Val Leu Arg Tyr Ala	Arg Glu His	
1750	1755	1760		
aat ctg	ggt aag att gtg aca	ggc cga cgc ccg gcg	cgc cgc tgg	5910
Asn Leu	Gly Lys Ile Val Thr	Gly Arg Arg Pro Ala	Arg Arg Trp	
1765	1770	1775		
cgg cgt	gac agc ttt gcc gag	cgg ctg ggg cag ttg	ggt ccc gat	5955
Arg Arg	Asp Ser Phe Ala Glu	Arg Leu Gly Gln Leu	Gly Pro Asp	
1780	1785	1790		
ctg gat	ctg ttg gtc gtg gcg	ctc aat gag cct atc	cag gat gcg	6000
Leu Asp	Leu Leu Val Val Ala	Leu Asn Glu Pro Ile	Gln Asp Ala	
1795	1800	1805		
ccc cat	ccg tta gcc gag gat	cgg gtt aac agc gac	aaa tgg cgg	6045
Pro His	Pro Leu Ala Glu Asp	Arg Val Asn Ser Asp	Lys Trp Arg	
1810	1815	1820		

ctg cag	ctg cgc	ggc gtc	ctg atg	gcg ctg	gtg ctg	tgt att	gtg	6090
Leu Gln	Leu Arg	Gly Val	Leu Met	Ala Leu	Val Leu	Cys Ile	Val	
1825			1830			1835		
gtc acc	gcc gca	ggg cag	tcg gtg	ctg atc	agc ttc	gat ccg	gcc	6135
Val Thr	Ala Ala	Gly Gln	Ser Val	Leu Ile	Ser Phe	Asp Pro	Ala	
1840			1845			1850		
aac tgt	gtg atg	atc tat	tta ctg	gcg gtg	gtg atc	gtc gcg	ttg	6180
Asn Cys	Val Met	Ile Tyr	Leu Leu	Ala Val	Val Ile	Val Ala	Leu	
1855			1860			1865		
cgc tat	gga cga	tgg ccc	tcc gtt	atc gcc	acc gtc	atg aac	atc	6225
Arg Tyr	Gly Arg	Trp Pro	Ser Val	Ile Ala	Thr Val	Met Asn	Ile	
1870			1875			1880		
att gcc	ttt gac	ctg ttt	ttc gtc	gca cct	acc ggc	acg gtc	gcg	6270
Ile Ala	Phe Asp	Leu Phe	Phe Val	Ala Pro	Thr Gly	Thr Val	Ala	
1885			1890			1895		
gtc tcg	gat ttg	caa tac	ctg gtg	acc ttt	ggg gtg	atg ctg	gcg	6315
Val Ser	Asp Leu	Gln Tyr	Leu Val	Thr Phe	Gly Val	Met Leu	Ala	
1900			1905			1910		
gtc ggg	gtc att	gtt ggc	aac ctg	acg gcc	ggc gtt	cgc tac	cag	6360
Val Gly	Val Ile	Val Gly	Asn Leu	Thr Ala	Gly Val	Arg Tyr	Gln	
1915			1920			1925		
gcg cgg	gtt gcc	cgc tac	cgg gag	cag cg	acg cg	cag ctt	tac	6405
Ala Arg	Val Ala	Arg Tyr	Arg Glu	Gln Arg	Thr Arg	Gln Leu	Tyr	
1930			1935			1940		
gaa atg	gcc aag	tcg ctg	gga agc	ggc ctg	acg cct	gaa gat	atc	6450
Glu Met	Ala Lys	Ser Leu	Gly Ser	Gly Leu	Thr Pro	Glu Asp	Ile	
1945			1950			1955		
gcc gcg	acc agc	cag cg	gtg ttg	gag gcg	acc tta	cag gcg	cga	6495
Ala Ala	Thr Ser	Gln Arg	Val Leu	Glu Ala	Thr Leu	Gln Ala	Arg	
1960			1965			1970		
tgc ctg	ctg ctg	cta ccc	gat gag	cag ggc	gaa ctg	cac acg	ctg	6540
Cys Leu	Leu Leu	Leu Pro	Asp Glu	Gln Gly	Glu Leu	His Thr	Leu	
1975			1980			1985		
ggc aac	gcg ctg	ccg ggc	aac gaa	ccg gat	tgg gct	atc gcg	aaa	6585
Gly Asn	Ala Leu	Pro Gly	Asn Glu	Pro Asp	Trp Ala	Ile Ala	Lys	
1990			1995			2000		
tgg agc	ttc agc	aag ggc	cag cca	gca ggc	gca ggc	acg gac	acc	6630
Trp Ser	Phe Ser	Lys Gly	Gln Pro	Ala Ala	Gly Ala	Gly Thr	Asp Thr	
2005			2010			2015		
tta ccg	gcg gtg	ccc tat	cag att	ctg ccg	ctg aaa	gtg ggc	gat	6675
Leu Pro	Ala Val	Pro Tyr	Gln Ile	Leu Pro	Leu Lys	Val Gly	Asp	
2020			2025			2030		
ctg tgt	cg	gga ttg	ctg gtg	gtt gaa	ccg cag	aat gtg	cgt cag	6720
Leu Cys	Arg Gly	Leu Leu	Val Val	Glu Pro	Gln Asn	Val Arg	Gln	
2035			2040			2045		
ctg atg	gtg ccg	gaa cag	caa cg	ctg ctg	gaa acc	ttc acc	gtg	6765
Leu Met	Val Pro	Glu Gln	Gln Arg	Leu Leu	Glu Thr	Phe Thr	Val	
2050			2055			2060		
ctg att	gcc aat	gcc ctg	gag cg	atg gcg	ctg tcc	cag agt	gag	6810
Leu Ile	Ala Asn	Ala Leu	Glu Arg	Met Ala	Leu Ser	Gln Ser	Glu	
2065			2070			2075		
gcg gct	tcc cg	ctg tca	gct gaa	cgt gag	cag ctg	cgt aat	gct	6855
Ala Ala	Ser Arg	Leu Ser	Ala Glu	Arg Glu	Gln Leu	Arg Asn	Ala	
2080			2085			2090		
ttg ctg	tcg gcg	ctc tcc	cat gat	tta cgt	acc ccg	ctg acg	gtg	6900
Leu Leu	Ser Ala	Leu Ser	His Asp	Leu Arg	Thr Pro	Leu Thr	Val	
2095			2100			2105		
ctg ttt	ggt cag	gca gaa	atg ctg	atg ctg	gac ctg	gcc agc	gat	6945
Leu Phe	Gly Gln	Ala Glu	Met Leu	Met Leu	Asp Leu	Ala Ser	Asp	
2110			2115			2120		
aac tca	aag tat	gtg ccc	cag gcc	agc cag	att cgt	gaa caa	acc	6990
Asn Ser	Lys Tyr	Val Pro	Gln Ala	Ser Gln	Ile Arg	Glu Gln	Thr	
2125			2130			2135		

ctg agt acc att cgt ctg gtc agc aac atg ctg gat atg gcg cgt 7035
 Leu Ser Thr Ile Arg Leu Val Ser Asn Met Leu Asp Met Ala Arg
 2140 2145 2150
 att cag tca ggc ggc ctg aat tta cgc gaa gag tgg ctg gcg ctg 7080
 Ile Gln Ser Gly Gly Leu Asn Leu Arg Glu Glu Trp Leu Ala Leu
 2155 2160 2165
 gaa gag gtg att ggt ggc gcg ctc agt agc atg gcg ccg tcg ctc 7125
 Glu Glu Val Ile Gly Gly Ala Leu Ser Ser Met Ala Pro Ser Leu
 2170 2175 2180
 aag gga aga gag gtc gaa ctc gat ctg cct gaa gat att gtc ctg 7170
 Lys Gly Arg Glu Val Glu Leu Asp Leu Pro Glu Asp Ile Val Leu
 2185 2190 2195
 atc aaa ggc gac agt acg ttg ctg gag cgg gta ttt acc aac ctg 7215
 Ile Lys Gly Asp Ser Thr Leu Leu Glu Arg Val Phe Thr Asn Leu
 2200 2205 2210
 att gaa aac agc ctg aag tac gct ggc aac tgt gcg ccc cgc ggc 7260
 Ile Glu Asn Ser Leu Lys Tyr Ala Gly Asn Cys Ala Pro Arg Gly
 2215 2220 2225
 ata cgt gcc tgg tgt gaa aat acc cgg ctg gaa atc gcc atc tgg 7305
 Ile Arg Ala Trp Cys Glu Asn Thr Arg Leu Glu Ile Ala Ile Trp
 2230 2235 2240
 gac ggc ggg ccg ggc atc gcc caa aac gac ctg acg cgg att ttc 7350
 Asp Gly Gly Pro Gly Ile Ala Gln Asn Asp Leu Thr Arg Ile Phe
 2245 2250 2255
 gac aaa ttt tca cgc ggt gat aaa gaa tcg gcc gta ccg ggc gtt 7395
 Asp Lys Phe Ser Arg Gly Asp Lys Glu Ser Ala Val Pro Gly Val
 2260 2265 2270
 ggg ctg gga ctg gcg att tgt aaa acg att atc gaa agc cac ggc 7440
 Gly Leu Gly Leu Ala Ile Cys Lys Thr Ile Ile Glu Ser His Gly
 2275 2280 2285
 ggt cag atc tgg gcg gaa aat cgt gct gaa ggc ggt gcc tgc ttt 7485
 Gly Gln Ile Trp Ala Glu Asn Arg Ala Glu Gly Gly Ala Cys Phe
 2290 2295 2300
 cgt ctc tct tta cca ctt cca ccc gtt cct gaa att tct cct gaa 7530
 Arg Leu Ser Leu Pro Leu Pro Pro Val Pro Glu Ile Ser Pro Glu
 2305 2310 2315
 ggc ttg aaa taa cttcacagat gatcggttat aatgcgcgac cttactgatt 7582
 Gly Leu Lys
 2320
 atgattggga aattatggaa cgttttaccg aaaacctg 7620

<210> 8
 <211> 560
 <212> PRT
 <213> Pantoea ananatis

<400> 8
 Met Ala Ala Asn Ala Phe Leu Leu Ile Ala Val Tyr Leu Leu Leu Leu
 1 5 10 15
 Met Val Met Ala Gln Pro Leu Gly Arg Gly Leu Ala Ala Leu Val Ala
 20 25 30
 Asp Lys Pro Leu Phe Ala Arg Ala Glu Ala Leu Leu Trp Arg Phe Ser
 35 40 45
 Gly Val Gln Glu Gly Gly Met Arg Trp Gln His Tyr Leu Leu Ala Ile
 50 55 60
 Leu Val Phe Asn Leu Leu Gly Phe Val Val Leu Leu Ala Ile Leu Met
 65 70 75 80
 Phe Gln Gly Ala Leu Pro Leu Asn Pro Gln His Leu Pro Gly Leu Ser
 85 90 95
 Trp Asp Leu Ala Leu Asn Thr Ala Ile Ser Phe Val Thr Asn Thr Asn
 100 105 110
 Trp Gln Ser Tyr Ala Gly Glu Ser Thr Leu Ser Tyr Phe Ser Gln Met

Met Ser Arg Gln Gln Gln Val Phe Asp Ala Ala Leu Leu Arg Thr Ser
 1 5 10 15
 Ala Ile Asp Ala Val Lys Lys Leu Asp Pro Arg Val Gln Phe Arg Asn
 20 25 30
 Pro Val Met Phe Val Val Tyr Leu Gly Ser Ile Leu Thr Ser Ile Leu
 35 40 45
 Ala Ile Met Met Phe Thr Gly His Gln Ser Gly Ser Ala Ser Phe Thr
 50 55 60
 Gly Ala Ile Ala Leu Trp Leu Trp Phe Thr Val Leu Phe Ala Asn Met
 65 70 75 80
 Ala Glu Ala Leu Ala Glu Gly Arg Ser Lys Ala Gln Ala Asn Ser Leu
 85 90 95
 Lys Gly Val Lys Lys Thr Ser Phe Ala Lys Lys Leu Ser Ala Ala His
 100 105 110
 Tyr Gly Ala Ala Trp Gln Gln Val Ala Ala Asp Ala Leu Arg Lys Gly
 115 120 125
 Asp Ala Val Leu Val Glu Ala Gly Asp Val Ile Pro Cys Asp Gly Glu
 130 135 140
 Val Val Glu Gly Gly Ala Ser Val Asp Glu Ser Ala Ile Thr Gly Glu
 145 150 155 160
 Ser Ala Pro Val Ile Arg Glu Ser Gly Gly Asp Phe Ala Ser Val Thr
 165 170 175
 Gly Gly Thr Arg Ile Leu Ser Asp Trp Leu Val Ile Thr Cys Ser Ala
 180 185 190
 Asn Pro Gly Glu Thr Phe Leu Asp Arg Met Ile Ala Met Val Glu Gly
 195 200 205
 Ala Gln Arg Arg Lys Thr Pro Asn Glu Ile Ala Leu Thr Ile Leu Leu
 210 215 220
 Val Ser Leu Thr Ile Val Phe Leu Leu Ala Thr Val Thr Leu Trp Pro
 225 230 235 240
 Phe Ser Ala Trp Gly Gly Thr Pro Val Thr Ile Thr Val Leu Val Ala
 245 250 255
 Leu Leu Val Cys Leu Ile Pro Thr Thr Ile Gly Gly Leu Leu Ser Ala
 260 265 270
 Ile Gly Val Ala Gly Met Ser Arg Met Leu Gly Ala Asn Val Ile Ala
 275 280 285
 Thr Ser Gly Arg Ala Val Glu Ala Ala Gly Asp Val Asp Val Leu Met
 290 295 300
 Leu Asp Lys Thr Gly Thr Ile Thr Leu Gly Asn Arg Gln Ala Thr Gln
 305 310 315 320
 Phe Leu Pro Ala Pro Gly Val Thr Glu Glu Gln Leu Ala Asp Ala Ala
 325 330 335
 Gln Leu Ala Ser Leu Ala Asp Glu Thr Pro Glu Gly Arg Ser Ile Val
 340 345 350
 Val Leu Ala Lys Gln Lys Phe Asn Leu Arg Glu Arg Asp Leu Ser Ser
 355 360 365
 Met Gly Ala Ser Phe Ile Pro Phe Ser Ala Gln Thr Arg Met Ser Gly
 370 375 380
 Val Asn Val Gln Asp Arg Leu Ile Arg Lys Gly Ala Val Asp Ala Val
 385 390 395 400
 Arg Arg His Ile Glu Ala Ser His Gly Ala Phe Pro Ala Glu Val Asn
 405 410 415
 Ala Arg Val Glu Glu Val Ala Arg Ala Gly Gly Thr Pro Leu Val Val
 420 425 430
 Ala Glu Gly Ala Lys Val Leu Gly Val Val Ala Leu Lys Asp Ile Val
 435 440 445
 Lys Gly Gly Ile Lys Glu Arg Phe Ala Glu Leu Arg Lys Met Gly Ile
 450 455 460
 Lys Thr Val Met Ile Thr Gly Asp Asn Pro Leu Thr Ala Ala Ala Ile
 465 470 475 480
 Ala Ala Glu Ala Gly Val Asp Asp Phe Leu Ser Glu Ala Thr Pro Glu
 485 490 495
 Ala Lys Leu Ala Leu Ile Arg Gln Tyr Gln Ala Glu Gly Arg Leu Val

500 505 510
 Ala Met Thr Gly Asp Gly Thr Asn Asp Ala Pro Ala Leu Ala Gln Ala
 515 520 525
 Asp Val Ala Val Ala Met Asn Ser Gly Thr Gln Ala Ala Lys Glu Ala
 530 535 540
 Gly Asn Met Val Asp Leu Asp Ser Asn Pro Thr Lys Leu Leu Glu Val
 545 550 555 560
 Val His Ile Gly Lys Gln Met Leu Met Thr Arg Gly Ser Leu Thr Thr
 565 570 575
 Phe Ser Ile Ala Asn Asp Val Ala Lys Tyr Phe Ala Ile Ile Pro Ala
 580 585 590
 Ala Phe Ala Ala Thr Tyr Pro Gln Leu Asn Met Leu Asn Val Met Gln
 595 600 605
 Leu His Ser Pro Ala Ser Ala Ile Leu Ser Ala Val Ile Phe Asn Ala
 610 615 620
 Leu Val Ile Val Phe Leu Ile Pro Leu Ala Leu Lys Gly Val Ser Tyr
 625 630 635 640
 Arg Pro Leu Ser Ala Ala Ser Leu Leu Arg Arg Asn Leu Leu Ile Tyr
 645 650 655
 Gly Leu Gly Gly Leu Leu Val Pro Phe Val Gly Ile Lys Ala Ile Asp
 660 665 670
 Met Leu Leu Val Leu Ser Gly Met Ala
 675 680

<210> 10
 <211> 189
 <212> PRT
 <213> *Pantoea ananatis*

<400> 10
 Met Ser Gln Leu Arg Pro Ala Ile Phe Leu Leu Leu Leu Leu Thr Val
 1 5 10 15
 Val Cys Gly Val Val Tyr Pro Leu Leu Thr Thr Gly Leu Ser Gln Leu
 20 25 30
 Leu Phe Pro Trp Gln Ala Asn Gly Ser Val Leu Asn Val Asp Gly Glu
 35 40 45
 Glu Arg Gly Ser Ala Leu Ile Gly Gln Asn Phe Ser Gln Pro Gly Tyr
 50 55 60
 Phe Trp Gly Arg Pro Ser Ala Thr Gly Asp Lys Pro Tyr Asn Pro Leu
 65 70 75 80
 Ala Ser Ser Gly Ser Asn Leu Ala Ala Ser Asn Pro Ala Leu Asp Lys
 85 90 95
 Ala Val Ala Glu Arg Val Ala Ala Leu Arg Thr Ala Asn Pro Gln Ala
 100 105 110
 Asn Gly Ala Val Pro Val Glu Leu Val Thr Thr Ser Ala Ser Gly Leu
 115 120 125
 Asp Pro Glu Ile Ser Pro Glu Ala Ala Leu Trp Gln Ala Pro Arg Ile
 130 135 140
 Ala Ala Ala Arg Gln Leu Pro Leu Ala Lys Val Asp Ala Leu Val Asp
 145 150 155 160
 Ser Met Thr Gln Arg Pro Leu Leu Pro Phe Ile Gly Glu Pro Thr Val
 165 170 175
 Asn Val Leu Gln Leu Asn Leu Ala Leu Asn Asp Leu Lys
 180 185

<210> 11
 <211> 891
 <212> PRT
 <213> *Pantoea ananatis*

<400> 11
 Met Asn His Glu Pro Leu Arg Pro Asp Pro Asp Ala Leu Leu Gln Thr

1				5					10					15			
Ser	Ser	Asp	Ser	His	Arg	Gly	Lys	Leu	Lys	Ile	Tyr	Phe	Gly	Ala	Cys		
			20					25					30				
Ala	Gly	Val	Gly	Lys	Thr	Tyr	Ala	Met	Leu	Gln	Glu	Ala	Gln	Arg	Leu		
		35					40					45					
Arg	Ala	Gln	Gly	Leu	Asp	Val	Leu	Val	Gly	Val	Val	Glu	Thr	His	Glu		
	50					55					60						
Arg	Pro	Glu	Thr	Ala	Gln	Leu	Leu	Asn	Gly	Leu	Val	Leu	Leu	Pro	Arg		
65					70				75					80			
Arg	Ala	Thr	Gly	Arg	Ser	Arg	His	Ala	Glu	Phe	Asp	Leu	Asp	Ala	Ala		
				85				90					95				
Leu	Ala	Arg	His	Pro	Ala	Val	Ile	Leu	Met	Asp	Glu	Leu	Ala	His	Thr		
			100					105					110				
Asn	Val	Lys	Gly	Ser	Arg	His	Pro	Lys	Arg	Trp	Gln	Asp	Ile	Glu	Glu		
		115					120					125					
Leu	Leu	Glu	Ala	Gly	Ile	Asp	Val	Leu	Thr	Thr	Val	Asn	Val	Gln	His		
	130					135					140						
Leu	Glu	Ser	Leu	Asn	Asp	Val	Val	Gly	Gly	Val	Thr	Gly	Ile	Gln	Val		
145					150					155					160		
Arg	Glu	Thr	Val	Pro	Asp	Pro	Phe	Phe	Asp	Ala	Ala	Asp	Glu	Val	Val		
				165					170					175			
Leu	Val	Asp	Leu	Pro	Pro	Asp	Asp	Leu	Arg	Gln	Arg	Leu	Lys	Glu	Gly		
			180					185					190				
Lys	Val	Tyr	Ile	Gly	Asp	Arg	Ala	Glu	Arg	Ala	Ile	Glu	Asn	Phe	Phe		
		195					200					205					
Arg	Lys	Gly	Asn	Leu	Tyr	Ala	Leu	Arg	Glu	Leu	Ala	Leu	Arg	Arg	Thr		
	210					215					220						
Ala	Asp	Arg	Val	Asp	Asp	Gln	Met	Arg	Ala	Trp	Arg	Asp	Ser	Gln	Gly		
225					230					235				240			
Arg	Asp	Arg	Val	Trp	His	Thr	Arg	Asp	Ala	Ile	Leu	Leu	Cys	Ile	Gly		
				245					250					255			
Asp	Asp	Thr	Gly	Ser	Glu	Lys	Leu	Val	Arg	Thr	Ala	Ala	Arg	Leu	Ala		
			260					265						270			
Ala	Arg	Leu	Gly	Ser	Glu	Trp	His	Ala	Val	Tyr	Val	Glu	Thr	Pro	Arg		
		275					280					285					
Leu	Asn	Arg	Leu	Pro	Glu	Ala	Arg	Arg	Arg	Ala	Ile	Leu	Arg	Thr	Leu		
	290					295					300						
Lys	Leu	Ala	Gln	Asp	Met	Gly	Ala	Glu	Thr	Ala	Thr	Leu	Ser	Asp	Pro		
305					310					315					320		
Asp	Glu	Ala	Gln	Ala	Val	Leu	Arg	Tyr	Ala	Arg	Glu	His	Asn	Leu	Gly		
				325					330					335			
Lys	Ile	Val	Thr	Gly	Arg	Arg	Pro	Ala	Arg	Arg	Trp	Arg	Arg	Asp	Ser		
			340					345						350			
Phe	Ala	Glu	Arg	Leu	Gly	Gln	Leu	Gly	Pro	Asp	Leu	Asp	Leu	Leu	Val		
		355					360					365					
Val	Ala	Leu	Asn	Glu	Pro	Ile	Gln	Asp	Ala	Pro	His	Pro	Leu	Ala	Glu		
		370				375					380						
Asp	Arg	Val	Asn	Ser	Asp	Lys	Trp	Arg	Leu	Gln	Leu	Arg	Gly	Val	Leu		
385					390					395				400			
Met	Ala	Leu	Val	Leu	Cys	Ile	Val	Val	Thr	Ala	Ala	Gly	Gln	Ser	Val		
				405					410					415			
Leu	Ile	Ser	Phe	Asp	Pro	Ala	Asn	Cys	Val	Met	Ile	Tyr	Leu	Leu	Ala		
			420					425					430				
Val	Val	Ile	Val	Ala	Leu	Arg	Tyr	Gly	Arg	Trp	Pro	Ser	Val	Ile	Ala		
		435					440					445					
Thr	Val	Met	Asn	Ile	Ile	Ala	Phe	Asp	Leu	Phe	Phe	Val	Ala	Pro	Thr		
		450				455					460						
Gly	Thr	Val	Ala	Val	Ser	Asp	Leu	Gln	Tyr	Leu	Val	Thr	Phe	Gly	Val		
465					470					475				480			
Met	Leu	Ala	Val	Gly	Val	Ile	Val	Gly	Asn	Leu	Thr	Ala	Gly	Val	Arg		
				485					490					495			
Tyr	Gln	Ala	Arg	Val	Ala	Arg	Tyr	Arg	Glu	Gln	Arg	Thr	Arg	Gln	Leu		
				500				505						510			

Tyr Glu Met Ala Lys Ser Leu Gly Ser Gly Leu Thr Pro Glu Asp Ile
 515 520 525
 Ala Ala Thr Ser Gln Arg Val Leu Glu Ala Thr Leu Gln Ala Arg Cys
 530 535 540
 Leu Leu Leu Leu Pro Asp Glu Gln Gly Glu Leu His Thr Leu Gly Asn
 545 550 555 560
 Ala Leu Pro Gly Asn Glu Pro Asp Trp Ala Ile Ala Lys Trp Ser Phe
 565 570 575
 Ser Lys Gly Gln Pro Ala Gly Ala Gly Thr Asp Thr Leu Pro Ala Val
 580 585 590
 Pro Tyr Gln Ile Leu Pro Leu Lys Val Gly Asp Leu Cys Arg Gly Leu
 595 600 605
 Leu Val Val Glu Pro Gln Asn Val Arg Gln Leu Met Val Pro Glu Gln
 610 615 620
 Gln Arg Leu Leu Glu Thr Phe Thr Val Leu Ile Ala Asn Ala Leu Glu
 625 630 635 640
 Arg Met Ala Leu Ser Gln Ser Glu Ala Ala Ser Arg Leu Ser Ala Glu
 645 650 655
 Arg Glu Gln Leu Arg Asn Ala Leu Leu Ser Ala Leu Ser His Asp Leu
 660 665 670
 Arg Thr Pro Leu Thr Val Leu Phe Gly Gln Ala Glu Met Leu Met Leu
 675 680 685
 Asp Leu Ala Ser Asp Asn Ser Lys Tyr Val Pro Gln Ala Ser Gln Ile
 690 695 700
 Arg Glu Gln Thr Leu Ser Thr Ile Arg Leu Val Ser Asn Met Leu Asp
 705 710 715 720
 Met Ala Arg Ile Gln Ser Gly Gly Leu Asn Leu Arg Glu Glu Trp Leu
 725 730 735
 Ala Leu Glu Glu Val Ile Gly Gly Ala Leu Ser Ser Met Ala Pro Ser
 740 745 750
 Leu Lys Gly Arg Glu Val Glu Leu Asp Leu Pro Glu Asp Ile Val Leu
 755 760 765
 Ile Lys Gly Asp Ser Thr Leu Leu Glu Arg Val Phe Thr Asn Leu Ile
 770 775 780
 Glu Asn Ser Leu Lys Tyr Ala Gly Asn Cys Ala Pro Arg Gly Ile Arg
 785 790 795 800
 Ala Trp Cys Glu Asn Thr Arg Leu Glu Ile Ala Ile Trp Asp Gly Gly
 805 810 815
 Pro Gly Ile Ala Gln Asn Asp Leu Thr Arg Ile Phe Asp Lys Phe Ser
 820 825 830
 Arg Gly Asp Lys Glu Ser Ala Val Pro Gly Val Gly Leu Gly Leu Ala
 835 840 845
 Ile Cys Lys Thr Ile Ile Glu Ser His Gly Gly Gln Ile Trp Ala Glu
 850 855 860
 Asn Arg Ala Glu Gly Gly Ala Cys Phe Arg Leu Ser Leu Pro Leu Pro
 865 870 875 880
 Pro Val Pro Glu Ile Ser Pro Glu Gly Leu Lys
 885 890

<210> 12
 <211> 1215
 <212> DNA
 <213> *Pantoea ananatis*

<220>
 <221> CDS
 <222> (301)..(975)

<400> 12
 gcgtacggcg caaaatgatt ttatgacgag aataagtccg gtacatggat tcctcctgac 60
 gaaaagtccg ttaaattgta tctgcatcca atattacgcc tgattgtgta attttactac 120
 tcactctgacc actaaatttg acgctttttt gtactgttaa tgacaatctg aaaattagtt 180

acatttttgggt taaataatgg ttttgacgat gatcttttagc gcgattcgcgatt ggcaacacgg 240
 acgaaaggct gaaaagttga tacagtgtat tttccctttt caggcaatgg caggattggc 300
 gtg acc acg gtt tta atc att gaa gat gag aaa gaa att cgc cgc ttc 348
 Val Thr Thr Val Leu Ile Ile Glu Asp Glu Lys Glu Ile Arg Arg Phe
 1 5 10 15
 gtg cgc atc gcg ttg gaa agc gaa ggc ctg aag gtt ttc gat gcc gaa 396
 Val Arg Ile Ala Leu Glu Ser Glu Gly Leu Lys Val Phe Asp Ala Glu
 20 25 30
 acg cta caa cgt ggg ttg att gag gcg gcg acg cga aaa ccc gat ctg 444
 Thr Leu Gln Arg Gly Leu Ile Glu Ala Ala Thr Arg Lys Pro Asp Leu
 35 40 45
 gtc att ctc gat ctc ggc ctg ccc gat ggc gat ggc aaa acc ttt att 492
 Val Ile Leu Asp Leu Gly Leu Pro Asp Gly Asp Gly Lys Thr Phe Ile
 50 55 60
 ggc gag ctg cgt cag tgg agc acg ctg ccc gtg att gtg ctg tcg gcc 540
 Gly Glu Leu Arg Gln Trp Ser Thr Leu Pro Val Ile Val Leu Ser Ala
 65 70 75 80
 cga atc gac gaa cag gat aaa att gac gcg ctg gat gca ggg gcc gac 588
 Arg Ile Asp Glu Gln Asp Lys Ile Asp Ala Leu Asp Ala Gly Ala Asp
 85 90 95
 gat tac ctg acg aaa ccc ttc ggt att ggt gaa ctg ctg gca cgc gtt 636
 Asp Tyr Leu Thr Lys Pro Phe Gly Ile Gly Glu Leu Leu Ala Arg Val
 100 105 110
 cgc gtc gcc ttg cgc cgt cat gcc gga caa cat acc gat ccc aag gtc 684
 Arg Val Ala Leu Arg Arg His Ala Gly Gln His Thr Asp Pro Lys Val
 115 120 125
 agc ttc gcc gac gtt acc gtg gat att gcg gcc cgc aga gtg ctg cgc 732
 Ser Phe Ala Asp Val Thr Val Asp Ile Ala Ala Arg Arg Val Leu Arg
 130 135 140
 gct ggc gag gaa gtg cac ctt acg ccg ata gag ttt cgt ttg ctg acg 780
 Ala Gly Glu Glu Val His Leu Thr Pro Ile Glu Phe Arg Leu Leu Thr
 145 150 155 160
 acg ctg ctg aac aac gcg ggc aaa gtg ctg acc cag cgg cag ctg ttg 828
 Thr Leu Leu Asn Asn Ala Gly Lys Val Leu Thr Gln Arg Gln Leu Leu
 165 170 175
 agc cag gtg tgg gga cca aac gcc gtt gaa cac agc cac tat ctg cgg 876
 Ser Gln Val Trp Gly Pro Asn Ala Val Glu His Ser His Tyr Leu Arg
 180 185 190
 atc tat atg ggg cac ctg cgg caa aag ctg gag gcg aat cct acc cag 924
 Ile Tyr Met Gly His Leu Arg Gln Lys Leu Glu Ala Asn Pro Thr Gln
 195 200 205
 ccg gta cat ctg ctc acg gaa acc ggc atc ggc tac cgg ttt atg cca 972
 Pro Val His Leu Leu Thr Glu Thr Gly Ile Gly Tyr Arg Phe Met Pro
 210 215 220
 taa aaaaagcgcc acttaggcgc ttttttcatt taacaggcaa atcaggcggt 1025
 tttcagcact tcgctgacaa tctctaccgc ttttttttct atctgcgcgc ggtgtttctgc 1085
 gccaggaaa ctttcacaat agattttgta tgcattctcg gtgcctgaag gacgggcccgc 1145
 aaaccagccg ttttccgtca tcactttcag gccgccgata gacgcgcat tgcccggcgc 1205
 agcggtcaga 1215

<210> 13
 <211> 224
 <212> PRT
 <213> *Pantoea ananatis*

<400> 13
 Val Thr Thr Val Leu Ile Ile Glu Asp Glu Lys Glu Ile Arg Arg Phe
 1 5 10 15
 Val Arg Ile Ala Leu Glu Ser Glu Gly Leu Lys Val Phe Asp Ala Glu
 20 25 30
 Thr Leu Gln Arg Gly Leu Ile Glu Ala Ala Thr Arg Lys Pro Asp Leu
 35 40 45

Val Ile Leu Asp Leu Gly Leu Pro Asp Gly Asp Gly Lys Thr Phe Ile
 50 55 60
 Gly Glu Leu Arg Gln Trp Ser Thr Leu Pro Val Ile Val Leu Ser Ala
 65 70 75 80
 Arg Ile Asp Glu Gln Asp Lys Ile Asp Ala Leu Asp Ala Gly Ala Asp
 85 90 95
 Asp Tyr Leu Thr Lys Pro Phe Gly Ile Gly Glu Leu Leu Ala Arg Val
 100 105 110
 Arg Val Ala Leu Arg Arg His Ala Gly Gln His Thr Asp Pro Lys Val
 115 120 125
 Ser Phe Ala Asp Val Thr Val Asp Ile Ala Ala Arg Arg Val Leu Arg
 130 135 140
 Ala Gly Glu Glu Val His Leu Thr Pro Ile Glu Phe Arg Leu Leu Thr
 145 150 155 160
 Thr Leu Leu Asn Asn Ala Gly Lys Val Leu Thr Gln Arg Gln Leu Leu
 165 170 175
 Ser Gln Val Trp Gly Pro Asn Ala Val Glu His Ser His Tyr Leu Arg
 180 185 190
 Ile Tyr Met Gly His Leu Arg Gln Lys Leu Glu Ala Asn Pro Thr Gln
 195 200 205
 Pro Val His Leu Leu Thr Glu Thr Gly Ile Gly Tyr Arg Phe Met Pro
 210 215 220

<210> 14
 <211> 1308
 <212> DNA
 <213> Pantoea ananatis

<400> 14
 atgagcagaa tcatgacgcc cgtgaactgg gaagcctgca gcagcgaggc gcagcaggcg 60
 ctgtttggcac gccctgcgct cgcctcgtct gacagcatca gccagatcgt gcgcatgtg 120
 ttggctcagag tgaagagga aggcgatgag gctttacgag aattcagcgc gcgctttgac 180
 aaggttgaaa cagacgacct gcgcgttacg ccacagcaga tgcaggcggc cagcgatcgc 240
 cttggtgacg agctgaaaca ggcatggcc gtggccattg gcaatattga aacctttcac 300
 cgtgcgcaga tcctgcccgc ggtggatgtg gaaacgcagc ccggcgtgcg ctgtcagcaa 360
 attacgcgcc cgatgaaatc ggtgggcttg tatattccgg gcggttctgc cccgctgttt 420
 tctaccgttc tgatgctggc taccocggcg cggattgcgg gctgtggtcg cgtgggtgctg 480
 tgctgcgcc cgcgattgc tgatgaaatt ctctacgocg ccaaactttg cgggtggaa 540
 gaagtgttcc aggtgggtgg atcacaggcg attgccgcc tggcttttg caccgaaagc 600
 atccctaagg tagataaaat ttttggctcc ggcaacgcgt gggttaccga agccaaacgt 660
 caggtcagcc agcgccttga tggcgcggcg attgatatgc ccgctggccc gtcggaagtg 720
 ctgggtgatt cgatgaagg tgccacaccg gccttcgttg cctctgatct gctgtcgcag 780
 gcggaacacg gccctgactc gcaggtgatt ttactgacgc cttcgcctggc gctggccgag 840
 cgcgctcgcc aggcggtgga ggatcagctg gccagttgc cacgtgcggc gacagcccgc 900
 caggcactgg aaagcagccg cctgatcgtc gcccgggata tgcagcaatg cattgcgatc 960
 tccaaccgct atggtccgga gcacctgatt ctgcaaacc gcacgccagc ggatctgggtg 1020
 gaacagatta ccagcgcggc ttcggttttc ctggcgact ggtcaccgga atccgcagga 1080
 gattatgctt cgggcaccaa ccacgtgctg ccgacctag gctataccgc gacatgctcc 1140
 agcctgggccc tggccgactt tcagaaacgc atgacggtac aggagctgac gccgcagggc 1200
 ttctgaacc tggcggcgac catcgaacc ctggcggccg ctgaacagct gcacgcccac 1260
 aaaaatgccg tcacgttgcg cgttgccgca ctcaaggagc aagcatga 1308

<210> 15
 <211> 68
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 15
 ccatagcggg ttgagatcgc aatgcattgc tgcatatccc tgaagcctgc ttttttatac 60

taagttgg 68

<210> 16
 <211> 68
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 16
 gcccgccagg cactggaaag cagccgcctg atcgtcgccc cgctcaagtt agtataaaaa 60
 agctgaac 68

<210> 17
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 17
 tagcgagatc tctgatgtcc ggcggtgctt ttg 33

<210> 18
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 18
 aaaaagagct cttacgcccc gccctgccac tc 32

<210> 19
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 19
 caggatctag aaggagacat gaacgatgaa catc 34

<210> 20
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 20
 gataaggatc cgaaataaaa gaaaatgccca atagga 36

<210> 21
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 21
 cctttgagct cgcgggcagt gagcgcaacg c 31

<210> 22
 <211> 48
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 22
 ctagagcggc cgccgatcgg gatcctcctg tgtgaaattg ttatccgc 48

<210> 23
 <211> 42
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 23
 ctctacgatc gaggagggtta taaaaaatgg atattaatac tg 42

<210> 24
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 24
 tcaaagcggc cgctttctcg tctgtttcta ctggta 36

<210> 25
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 25
 cctttggtac cgcgggcagt gagcgcaacg c 31

<210> 26
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

 <400> 26
 aacaggaatt ctttgctgg cggcagtagc gcgg 34

 <210> 27
 <211> 40
 <212> DNA
 <213> Artificial

 <220>
 <223> iniciador P1

 <400> 27
 ctagtaagat cttgaagcct gcttttttat actaagttgg 40

 <210> 28
 <211> 41
 <212> DNA
 <213> Artificial

 <220>
 <223> iniciador P2

 <400> 28
 atgatcgaat tcgaaatcaa ataatgattt tattttgact g 41

 <210> 29
 <211> 120
 <212> DNA
 <213> Artificial

 <220>
 <223> attL contendo fragmento de DNA

 <400> 29
 agatcttgaa gcctgctttt ttataactaag ttggcattat aaaaaagcat tgcttatcaa 60
 tttgttgcaa cgaacaggtc actatcagtc aaaataaaat cattatttga ttcgaattc 120

 <210> 30
 <211> 41
 <212> DNA
 <213> Artificial

 <220>
 <223> iniciador P3

 <400> 30
 atgccactgc agtctgttac aggtcactaa taccatctaa g 41

 <210> 31
 <211> 46
 <212> DNA
 <213> Artificial

 <220>

<223> iniciador P4

<400> 31

accgttaagc tttctagacg ctcaagttag tataaaaaag ctgaac

46

<210> 32

<211> 184

<212> DNA

<213> Artificial

<220>

<223> attL contendo fragmento de DNA

<400> 32

ctgcagtctg ttacaggtca ctaataccat ctaagtagtt gattcatagt gactgcatat
gttgtgtttt acagtattat gtagtctgtt ttttatgcaa aatctaattt aatatattga
tatttatatc attttacgtt tctcgttcag cttttttata ctaacttgag cgtctagaaa
gctt

60

120

180

184

<210> 33

<211> 38

<212> DNA

<213> Artificial

<220>

<223> iniciador P5

<400> 33

ttcttagacg tcaggtggca cttttcgggg aaatgtgc

38

<210> 34

<211> 37

<212> DNA

<213> Artificial

<220>

<223> iniciador P6

<400> 34

taacagagat ctgcgcgaga aaaaaaggat ctcaaga

37

<210> 35

<211> 46

<212> DNA

<213> Artificial

<220>

<223> iniciador P7

<400> 35

aacagagatc taagcttaga tcctttgcct ggcggcagta gcgcgg

46

<210> 36

<211> 35

<212> DNA

<213> Artificial

<220>

<223> iniciador P8

<400> 36
ataaactgca gcaaaaagag tttgtagaaa cgcaa

35

<210> 37
<211> 1388
<212> DNA
<213> Artificial

<220>
<223> ter_thrL e gene Tc contendo fragmento de DNA

<400> 37
gaattctcat gtttgacagc ttatcatcga taagctttaa tgcggtagtt tatcacagtt 60
aaattgctaa cgcagtcagg caccgtgtat gaaatctaac aatgcgctca tcgtcatcct 120
cggcaccgtc accctggatg ctgtaggcat aggcttgggt atgccggtac tgccgggcct 180
cttgcgggat atcgtccatt ccgacagcat cgccagtcac tatggcgtgc tgctagcgct 240
atatgcggtg atgcaatttc tatgcgacc cgttctcgga gcaactgtccg accgcttgg 300
ccgccgcca gtcctgctcg cttcgtact tggagccact atcgactacg cgatcatggc 360
gaccacacc gtcctgtgga tcctctacgc cggacgcac gtggccggca tcaccggcgc 420
cacaggtgcg gttgctggcg cctatatcgc cgacatcacc gatggggaag atcgggctcg 480
ccacttcggg ctcatgagcg cttgtttcgg cgtgggtatg gtggcaggcc ccgtggccgg 540
gggactgtg ggcgccatct ccttgcacgc accattcctt gcggcgggcg tgctcaacgg 600
cctcaaccta ctactgggct gcttcctaat gcaggagtgc cataagggag agcgtcgacc 660
gatgcccttg agagccttca acccagtcag ctcttccgg tgggcgcggg gcatgactat 720
cgtcgccgca cttatgactg tcttctttat catgcaactc gtaggacagg tgccggcagc 780
gctctgggtc attttcggcg aggaccgctt tcgctggagc gcgacgatga tcggcctgct 840
gcttgcggtg ttcggaatct tgcacgccct cgetcaagcc ttcgtcactg gtcccggcac 900
caaacgtttc ggcgagaagc aggcattat cgccggcatg gcggccgacg cgctgggcta 960
cgtcttgetg gcgttcgcga cgcgaggctg gatggccttc cccattatga ttcttctcgc 1020
ttccggcggc atcgggatgc ccgcgttgca ggccatgctg tccaggcagg tagatgacga 1080
ccatcagga cagcttcaag gatcgctcgc ggctcttacc agcctaactt cgatcactgg 1140
accgctgatc gtcacggcga tttatgccgc ctcggcgagc acatggaacg ggttggcatg 1200
gattgtaggc gccgccctat accttgtctg cctccccgcg ttgcgtcgcg gtgcatggag 1260
ccgggccacc tcgacctgaa tggaagccgg cggcacctcg ctaacggatt caccactcca 1320
actagaaagc ttaacacaga aaaaagcccg cacctgacag tgcgggcttt ttttttcgac 1380
cactgcag 1388

<210> 38
<211> 36
<212> DNA
<213> Artificial

<220>
<223> iniciador P9

<400> 38
agtaattcta gaaagcttaa cacagaaaa agccccg

36

<210> 39
<211> 43
<212> DNA
<213> Artificial

<220>
<223> iniciador P10

<400> 39
ctagtaggat ccctgcagtg gtcgaaaaaa aaagcccgcg ctg

43

<210> 40
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador 1

<400> 40
 ggaagatcta tttgccttcg cacatcaacc tgg 33

<210> 41
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador 2

<400> 41
 cggggtacct tgtaaatatt ttaacccgcc 30

<210> 42
 <211> 56
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador 3

<400> 42
 ggaagatcta aggagacctt aatgagcga cacaacgatc ctgcaaaaca gtaccc 56

<210> 43
 <211> 39
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador 4

<400> 43
 cggggtacct cgtagagggt tactggcgct tatccagcg 39

<210> 44
 <211> 864
 <212> DNA
 <213> *Pantoea ananatis*

<400> 44
 atgtctgaac aacactatca gcccgctaaa gtctggaagt gggaccgga agcgaaaggt 60
 aatggtgcca aaactaaccg ccctaccgct ggccaaccc atgaaaaagc cctgccggtt 120
 ggtaagcatc cgctgcagct ctactctctg ggcagccta acggccagaa agtcactatt 180
 ctgctggaag agttgctggc gctgaacgtg aacgatgcag agtacgatgc ctggttaatc 240
 aatattgggtg aaggcgacca gttcagcagc ggttttgttg agatcaaccc caactccaaa 300
 atccccggcac tgtgcgatca ttcagccaca ccaccgattc gcgtatttga atccggtaac 360
 atcctgctct atctggcgga aaaatacggg tatttctctgc cgaaagatcc ggccaacgc 420

accgaaacgc	tgaactggct	gttctggctg	caaggctcag	ccccttacct	tggcggcggc	480
tttggctact	tctatcatta	cgcgccagaa	aagattgaat	acgccatcaa	ccgcttctca	540
ctggaagcca	aacgtcagtt	tgacgtgctg	gaccgtcagc	ttgccgataa	ccgttatctg	600
gcgggtgacg	attacacat	cgccgataac	gccacctggc	cgtggtacgg	cagcatgggtg	660
ctgtataacc	agtacaatgc	ggcagaattc	ctcgaccttc	agtcctacaa	aaatgtgggtg	720
cgctggggccg	aagagatcgc	tctgctccg	gccgttatgc	gcggccgcaa	ggtgaaccgt	780
gtgatgggcg	aacccgccga	tcagetgcgc	gagcgccatg	acgcatcgga	ctttgatagc	840
caaacccaag	acaagcaagc	ctga				864

<210> 45
 <211> 918
 <212> DNA
 <213> *Pantoea ananatis*

<400> 45						
atgtcagcac	gagtctgggtg	tctgggtgat	gccgtgggtg	acctgcttcc	ggacggggccg	60
gggcatttaa	tacagtgtgc	aggcggggcg	cccgccaatg	tggcgggtggg	cattgcccgc	120
ttacagggcc	gcagcgggtt	tattggccgg	gttggggacg	atccttttgg	tcactttatg	180
cagcacacgc	tggcgactga	acaggttgat	acccgctata	tgacgctgga	cagcggcccag	240
cgcacctcaa	cgggtgggtg	ggcgtggat	caggaaggtg	agcggacttt	tacctttatg	300
gtgcgcccc	gtgcagatct	gtttctggaa	caaggcgatc	tccccagggt	tgagcaaggt	360
gaatggcttc	actgctgctc	aattgccctg	gcggcagaac	cttcgcgctc	caccaccttt	420
tctgccatgc	agcagatcag	cgatgccggt	ggctttgtga	gctttgatcc	caatattcgt	480
cacgatctgt	ggcacgacga	tgcccaactg	cgggactgtg	tgaaccgggc	gttacagctg	540
gccgatgtgg	tcaagctgtc	tgaggaagag	ctggcttttc	tgactccggg	ggcgcaaac	600
gctgacagca	tcagggcgt	ggcggaaacg	tttgcgatta	gcctgctgat	ggtcaccag	660
ggcaaggcag	gagtgaaagt	ctggcatcag	ggtaaaccatt	atcactatcc	cacgctgcct	720
gtggtgagcg	tggacaccac	cggcgcaggg	gatgcgtttg	tcgcccggct	gctatggggg	780
ctggcggaaa	aggggatgcc	cgctaataag	gccgagctgg	cggcacgact	cagcagcgca	840
cagcagtggtg	gggcgctggc	gacgacggca	aaaggggcca	tgaccgcggt	gccttatcgt	900
caccaaattg	aaggatga					918

<210> 46
 <211> 3095
 <212> DNA
 <213> *Pantoea ananatis*

<400> 46						
agactgccat	gaccctggac	agagattcat	tagcggcogt	actcgcccgg	cgcgactggg	60
aaaaccccgg	cgtcagcga	cataaccggc	tggaagccca	tccgcccgtt	tacagctggc	120
gcagcgtgta	ggcggcccat	aacaacgcgc	catcggcgca	gcgaaaaagc	ctgagcggcg	180
aatggacggt	tgctttttc	cctgcgcccg	aggcgggtgc	ggatagctgg	cgcacccagg	240
atthgcaggc	ggcagcagc	attaccgtgc	cgtcagctctg	gcaaatgcag	ggctatgatg	300
ttccgattta	caccaatggt	acctatccca	ttccggttga	tcccccgcc	gttccggctg	360
aaaatcctac	gggatgttat	tcgctcacat	ttaatgtgga	tgacagactg	ctgcaacatg	420
gacaaacccg	aattatthtt	gacggcgtga	attcagcctt	ctatctctgg	tgcaaccggc	480
gctgggtggg	ctacgggcag	gacagccgct	tgccgtctga	atthgatctg	agcgaatttc	540
tgcgcgaagg	tgaaaatcgc	ctggcgggtga	tgggtgtgcg	ctggagcgat	ggcagctatc	600
tggaagatca	ggatagtgg	cgcatgagtg	gtatthtccg	cgatgtttcc	ctgctgcata	660
agcctgccag	ccatcttcgc	gatctgcgca	ttcgtacgca	tttcaatgac	gatttcagtc	720
gtgcgcggct	ggaagctgag	gtgcgggttg	ccggagcaact	ggatgacgat	ttacgggtca	780
gcgtgcagct	cttcgcgggt	gacacgctaa	ccggagaagc	gacgtcgcct	ttgggcagcg	840
cgattattga	cgagcgcggc	gcgtggagcg	atcggacaac	gctgtgcac	aacgttgcta	900
accctgcgct	gtggagtgcg	gaaacgcgc	atctctaccg	ggcggttgtg	cagttacacc	960
ggacggacgg	tacgctgatt	gaggcggagg	cctgcgacgt	gggattccgg	cacgtcagca	1020
tcgaaaatgg	cctgctgctg	ctcaacggtc	agccactgct	catccgccc	accaaccgcc	1080
acgagcatca	tctgaacgc	ggtcaggtga	tggatcgtga	cactatgggtg	caggatattt	1140
tgctgatgaa	gcagaataac	ttcaacgcgg	tgcgctgctc	ccattaccct	aacgatcccc	1200
tgtggtatag	cctgtgtgac	cactacggct	tgtacgtcgt	ggatgaagcc	aacatcgaaa	1260
cgcatggcat	ggtgccgatg	aatgcgctga	gcgacgatcc	cgtctggctt	cccgccatga	1320
gccagcgcgt	cacgcgcatg	gtgcagcgcg	atcgtaatca	cccctgcatt	attatctggt	1380

cactgggtaa	cgaatcgggc	cacggtgcta	accatgatgc	gctctaccgc	tggtgaaaa	1440
gtgaagatcc	ttcccgcccg	gtccagtacg	aagggtggcg	ggccaatacc	gcagcgaccg	1500
atattatctg	tccgatgtat	gcgcggtcg	atgaggatca	acctttcccg	gccgtgccga	1560
aatgggccat	taaaaaatgg	ctgtctatgc	caggcgagca	gcgtccgctt	attctctgtg	1620
aatatgctca	tgccatggga	aacagccttg	gtggctacgc	aaaatactgg	caggcatttc	1680
gtcagtatcc	tcgctgcag	ggcggttttg	tctgggactg	ggtcgatcag	tcgctcataa	1740
aatatgacga	taacggcgaa	ccctgggctg	cctaccggtg	cgactttggt	gatacgccta	1800
atgatcgcca	gttctgcatg	aacgggctgg	tctttgccga	ccggacgcc	catccctcgc	1860
tctacgaagc	ccgccatgcg	cagcagttct	tccagttcag	gctgctaccg	ggcagcgagc	1920
gcacgctgga	agtgaccagc	gaatacctgt	tccgccacag	tgataatgaa	atcctgcact	1980
ggtcggtgcc	tcaggatggc	aacctgctgg	ccgccggtga	agtcacgctg	gatatcgcg	2040
cacagggccg	ccagcagatc	gcactcccgg	aggtgccgct	gcctcaatct	gccccccagc	2100
tctggctgac	ggtgcgggtg	gaacagcctc	agccgacggc	ctggtcagaa	gccggtcata	2160
tcagcgcctg	gcagcagtgg	ccactggagg	cgatcctgaa	tgtggccctg	ccgcctcagg	2220
cggcgagtgc	gccacagctc	agccgcggcg	aagacacctt	cagcgtggcg	gtgaacaacc	2280
agcgtgggc	attcagccgt	cagcagggcg	tgctcacgca	gtactggatc	gacgatcagc	2340
cgcagctcct	gtcgcgctg	cgcgatcagt	ttaccctgtg	gccgctggat	aacgatctcg	2400
gcgtcagca	agtcacgccc	atcgacccta	atgcctgggt	cgagcgtgg	aaagcggcgg	2460
ggcactatca	gtctgaggtc	accaccctgc	agtcaccgc	cgaagcgtc	tccgacgccg	2520
tcgtgatcaa	caccgttcat	gcctggcagt	tccagggcaa	aacgctgttt	attagccgta	2580
aagtgtaccg	tattgatgga	tttggcgaga	tggcggtgac	ggtgaacgta	gagatcgcg	2640
gcggcacgcc	ttatccggca	cgcatcggt	tgagctgtca	gctcacccag	atcgtcgagc	2700
gagtgaactg	gctaggcctg	gggcccgatg	aaaattacc	ggatcgctg	acctcggcct	2760
gctttgaccg	ctgggattta	ccgctgagtg	aaatgtacac	ccggtatggt	ttccccaccg	2820
aaaacggcct	gcgctgcggc	acgcgtgaac	tcaactatgg	tgcgcaccag	tggcgtggcg	2880
atttccagtt	caacatcagt	cgctacagcc	agacacagct	gatggaaacc	tgccatcgcc	2940
acctgttacg	gcctgaggcg	ggcacgtggc	tcaatattga	tggcttccac	atgggcgtcg	3000
gcggtgacga	ttcctggagc	ccgtcgggat	caccggaatt	cctgctcagt	gcaggacggt	3060
acagctacca	gtttatctgg	gggcaacaaa	aataa			3095

<210> 47
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 47
 gcttaagatc tccctggtga caattaatca tcgg

34

<210> 48
 <211> 70
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 48
 agtacggccg ctaatgaatc tctgtccagg gtcattggcag tctccttgtg tgaattgtt
 atccgctcac

60
70

<210> 49
 <211> 39
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 49
 ccgttagatc tcgctcaagt tagtataaaa aagctgaac 39

<210> 50
 <211> 70
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 50
 tagcgggctg atagtgttgt tcagacatga tgaggttcgc cttgaagcct gcttttttat 60
 actaagttgg 70

<210> 51
 <211> 78
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 51
 cgttgggaca acgtcgatct gcgccgaatc tggttgctgg acgccgcctg tgaagcctgc 60
 tttttatac taagttgg 78

<210> 52
 <211> 80
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 52
 caacagcaga taaaccgca tcagtaaaaa cgcattggcc gccatggcag tctccttgtg 60
 tgaaattggt atccgctcac 80

<210> 53
 <211> 17
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 53
 catgtcttct ggtcact 17

<210> 54
 <211> 16
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 54

caagcaggtt gaacac

16

<210> 55
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 55
 cccaagcttc cctgatcaat gaggaggcgt tc

32

<210> 56
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> iniciador

<400> 56
 cgggatccga cgatcggggt cgggacgtaa ggg

33

<210> 57
 <211> 557
 <212> PRT
 <213> Artificial

<220>
 <223> sequência de consenso de KdpA de P. anantis e E. coli

<220>
 <221> variação
 <222> 4, 5, 11, 12, 15, 19, 21, 25, 29, 31, 32, 34, 37, 38, 39, 40, 41,
 42, 44, 45, 47, 49, 50, 53, 54, 55, 56, 58, 60, 61, 64, 6, 69, 71, 74,
 75, 78, 79, 80, 82, 83, 85, 86, 93, 107, 119, 122, 131, 139, 143, 148,
 154, 156, 157, 158, 1690, 161, 168, 171, 176, 179, 182, 185, 190, 191,
 194, 195, 202, 203, 205, 207, 213, 215, 238, 241, 251, 258, 264, 273,
 274, 275, 281, 286, 288, 289, 290, 293, 294, 295, 296, 297, 303, 304,
 309, 311, 314, 317, 325, 329, 331, 338, 348, 364, 368, 378, 384, 404, 412,
 416, 426, 430, 432, 434, 443, 446, 450, 475, 478, 483, 487, 496, 506,
 512, 514, 528, 532, 555, 557

<223> Xaa pode ser qualquer aminoácido que ocorre naturalmente

<400> 57
 Met Ala Ala Xaa Xaa Phe Leu Leu Ile Ala Xaa Xaa Leu Leu Xaa Leu
 1 5 10 15
 Met Val Xaa Ala Xaa Pro Leu Gly Xaa Gly Leu Ala Xaa Leu Xaa Xaa
 20 25 30
 Asp Xaa Pro Leu Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa Xaa Leu Xaa Arg
 35 40 45
 Xaa Xaa Gly Val Xaa Xaa Xaa Xaa Met Xaa Trp Xaa Xaa Tyr Leu Xaa
 50 55 60
 Ala Ile Leu Xaa Xaa Asn Xaa Leu Gly Xaa Xaa Val Leu Xaa Xaa Xaa
 65 70 75 80
 Leu Xaa Xaa Gln Xaa Xaa Leu Pro Leu Asn Pro Gln Xaa Leu Pro Gly
 85 90 95
 Leu Ser Trp Asp Leu Ala Leu Asn Thr Ala Xaa Ser Phe Val Thr Asn
 100 105 110
 Thr Asn Trp Gln Ser Tyr Xaa Gly Glu Xaa Thr Leu Ser Tyr Phe Ser

115 120 125
 Gln Met Xaa Gly Leu Thr Val Gln Asn Phe Xaa Ser Ala Ala Xaa Gly
 130 135 140
 Ile Ala Val Xaa Phe Ala Leu Ile Arg Xaa Phe Xaa Xaa Xaa Ser Xaa
 145 150 155 160
 Xaa Thr Leu Gly Asn Ala Trp Xaa Asp Leu Xaa Arg Ile Thr Leu Xaa
 165 170 175
 Val Leu Xaa Pro Xaa Xaa Leu Leu Xaa Ala Leu Phe Phe Xaa Xaa Gln
 180 185 190
 Gly Xaa Xaa Gln Asn Phe Leu Pro Tyr Xaa Xaa Val Xaa Xaa Xaa Glu
 195 200 205
 Gly Ala Gln Gln Xaa Leu Xaa Met Gly Pro Val Ala Ser Gln Glu Ala
 210 215 220
 Ile Lys Met Leu Gly Thr Asn Gly Gly Gly Phe Phe Asn Xaa Asn Ser
 225 230 235 240
 Xaa His Pro Phe Glu Asn Pro Thr Ala Leu Xaa Asn Phe Val Gln Met
 245 250 255
 Leu Xaa Ile Phe Leu Ile Pro Xaa Ala Leu Cys Phe Ala Phe Gly Glu
 260 265 270
 Xaa Xaa Xaa Asp Arg Arg Gln Gly Xaa Met Leu Leu Trp Xaa Met Xaa
 275 280 285
 Xaa Xaa Phe Val Xaa Xaa Xaa Xaa Xaa Val Met Trp Ala Glu Xaa Xaa
 290 295 300
 Gly Asn Pro His Xaa Leu Xaa Leu Gly Xaa Asp Ser Xaa Ile Asn Met
 305 310 315 320
 Glu Gly Lys Glu Xaa Arg Phe Gly Xaa Leu Xaa Ser Ser Leu Phe Ala
 325 330 335
 Val Xaa Thr Thr Ala Ala Ser Cys Gly Ala Val Xaa Ala Met His Asp
 340 345 350
 Ser Phe Thr Ala Leu Gly Gly Met Val Pro Met Xaa Leu Met Gln Xaa
 355 360 365
 Gly Glu Val Val Phe Gly Gly Val Gly Xaa Gly Leu Tyr Gly Met Xaa
 370 375 380
 Leu Phe Val Leu Leu Ala Val Phe Ile Ala Gly Leu Met Ile Gly Arg
 385 390 395 400
 Thr Pro Glu Xaa Leu Gly Lys Lys Ile Asp Val Xaa Glu Met Lys Xaa
 405 410 415
 Thr Ala Leu Ala Ile Leu Val Thr Pro Xaa Leu Val Leu Xaa Gly Xaa
 420 425 430
 Ala Xaa Ala Met Met Thr Asp Ala Gly Arg Xaa Xaa Met Xaa Asn Pro
 435 440 445
 Gly Xaa His Gly Phe Ser Glu Val Leu Tyr Ala Val Ser Ser Ala Ala
 450 455 460
 Asn Asn Asn Gly Ser Ala Phe Ala Gly Leu Xaa Ala Asn Xaa Pro Phe
 465 470 475 480
 Trp Asn Xaa Leu Leu Ala Xaa Cys Met Phe Val Gly Arg Phe Gly Xaa
 485 490 495
 Ile Ile Pro Val Met Ala Ile Ala Gly Xaa Xaa Xaa Xaa Lys Lys Xaa
 500 505 510
 Gln Xaa Xaa Xaa Xaa Gly Thr Leu Pro Thr His Gly Pro Leu Phe Xaa
 515 520 525
 Xaa Leu Leu Xaa Gly Thr Val Leu Leu Val Gly Ala Leu Thr Phe Ile
 530 535 540
 Pro Ala Leu Ala Leu Gly Pro Val Ala Glu Xaa Leu Xaa
 545 550 555

<210> 58
 <211> 681
 <212> PRT.
 <213> Artificial

<220>

370						375						380			
Gly	Xaa	Asn	Xaa	Xaa	Xaa	Arg	Xaa	Ile	Arg	Lys	Gly	Xaa	Val	Asp	Ala
385						390				395					400
Xaa	Arg	Arg	His	Xaa	Glu	Ala	Xaa	Xaa	Gly	Xaa	Phe	Pro	Xaa	Xaa	Val
			405						410					415	
Xaa	Xaa	Xaa	Val	Xaa	Xaa	Val	Ala	Arg	Xaa	Gly	Xaa	Thr	Pro	Leu	Val
			420					425					430		
Val	Xaa	Glu	Gly	Xaa	Xaa	Val	Leu	Gly	Val	Xaa	Ala	Leu	Lys	Asp	Ile
		435					440					445			
Val	Lys	Gly	Gly	Ile	Lys	Glu	Arg	Phe	Ala	Xaa	Leu	Arg	Lys	Met	Gly
	450					455					460				
Ile	Lys	Thr	Val	Met	Ile	Thr	Gly	Asp	Asn	Xaa	Leu	Thr	Ala	Ala	Ala
465					470					475					480
Ile	Ala	Ala	Glu	Ala	Gly	Val	Asp	Asp	Phe	Leu	Xaa	Glu	Ala	Thr	Pro
			485						490					495	
Glu	Ala	Lys	Leu	Ala	Leu	Ile	Arg	Gln	Tyr	Gln	Ala	Glu	Gly	Arg	Leu
			500					505					510		
Val	Ala	Met	Thr	Gly	Asp	Gly	Thr	Asn	Asp	Ala	Pro	Ala	Leu	Ala	Gln
		515					520					525			
Ala	Asp	Val	Ala	Val	Ala	Met	Asn	Ser	Gly	Thr	Gln	Ala	Ala	Lys	Glu
	530					535					540				
Ala	Gly	Asn	Met	Val	Asp	Leu	Asp	Ser	Asn	Pro	Thr	Lys	Leu	Xaa	Glu
545					550					555					560
Val	Val	His	Ile	Gly	Lys	Gln	Met	Leu	Met	Thr	Arg	Gly	Ser	Leu	Thr
			565						570					575	
Thr	Phe	Ser	Ile	Ala	Asn	Asp	Val	Ala	Lys	Tyr	Phe	Ala	Ile	Ile	Pro
			580					585					590		
Ala	Ala	Phe	Ala	Ala	Thr	Tyr	Pro	Gln	Leu	Asn	Xaa	Leu	Asn	Xaa	Met
		595					600					605			
Xaa	Leu	His	Ser	Pro	Xaa	Ser	Ala	Ile	Leu	Ser	Ala	Val	Ile	Phe	Asn
	610					615					620				
Ala	Leu	Xaa	Ile	Val	Phe	Leu	Ile	Pro	Leu	Ala	Leu	Lys	Gly	Val	Ser
625					630					635					640
Tyr	Xaa	Pro	Leu	Xaa	Ala	Xaa	Xaa	Xaa	Leu	Arg	Arg	Asn	Leu	Xaa	Ile
			645						650					655	
Tyr	Gly	Leu	Gly	Gly	Leu	Leu	Val	Pro	Phe	Xaa	Gly	Ile	Lys	Xaa	Ile
			660					665					670		
Asp	Xaa	Leu	Leu	Xaa	Xaa	Xaa	Gly	Xaa							
		675					680								

<210> 59

<211> 189

<212> PRT

<213> Artificial

<220>

<223> sequência de consenso de KdpC de P. anantis e E. coli

<220>

<221> variação

<222> 3, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 20, 28, 30, 32, 33, 42, 43, 44, 45, 46, 47, 48, 49, 60, 61, 62, 66, 73, 74, 75, 80, 83, 89, 93, 97, 98, 100, 107, 111, 113, 114, 115, 123, 130, 131, 133, 135, 138, 141, 144, 146, 149, 151, 152, 153, 154, 155, 156, 157, 159, 160, 161, 162, 163, 165, 168, 169, 170, 173, 175, 178, 178, 179, 180, 186, 187, 188, 189

<223> Xaa pode ser qualquer aminoácido que ocorre naturalmente

<400> 59

Met	Ser	Xaa	Leu	Arg	Pro	Ala	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Leu	Xaa	Xaa
1			5						10				15		

Xaa	Xaa	Gly	Xaa	Val	Tyr	Pro	Leu	Leu	Thr	Thr	Xaa	Leu	Xaa	Gln	Xaa
			20						25					30	

Xaa Phe Pro Trp Gln Ala Asn Gly Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 35 40 45
 Xaa Arg Gly Ser Ala Leu Ile Gly Gln Asn Phe Xaa Xaa Xaa Gly Tyr
 50 55 60
 Phe Xaa Gly Arg Pro Ser Ala Thr Xaa Xaa Xaa Pro Tyr Asn Pro Xaa
 65 70 75 80
 Ala Ser Xaa Gly Ser Asn Leu Ala Xaa Ser Asn Pro Xaa Leu Asp Lys
 85 90 95
 Xaa Xaa Ala Xaa Arg Val Ala Ala Leu Arg Xaa Ala Asn Pro Xaa Ala
 100 105 110
 Xaa Xaa Xaa Val Pro Val Glu Leu Val Thr Xaa Ser Ala Ser Gly Leu
 115 120 125
 Asp Xaa Xaa Ile Xaa Pro Xaa Ala Ala Xaa Trp Gln Xaa Pro Arg Xaa
 130 135 140
 Ala Xaa Ala Arg Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa
 145 150 155 160
 Xaa Xaa Xaa Gln Xaa Pro Leu Xaa Xaa Xaa Ile Gly Xaa Pro Xaa Val
 165 170 175
 Asn Xaa Xaa Xaa Leu Asn Leu Ala Leu Xaa Xaa Leu Xaa
 180 185

REIVINDICAÇÕES

1. Microorganismo pertencente à família *Enterobacteriaceae*, caracterizado pelo fato de que tem uma capacidade de produção de L-aminoácidos e foi modificado de modo que o sistema *kdp* fosse intensificado.

5 2. Microorganismo de acordo com a reivindicação 1, caracterizado pelo fato de que o sistema *kdp* é intensificado pelo aumento da expressão do óperon *kdp* ou de um ou mais genes que constituem o óperon *kdp*, e/ou pelo aumento da translação do óperon *kdp* ou dos genes.

10 3. Microorganismo de acordo com a reivindicação 2, caracterizado pelo fato de que o sistema *kdp* é intensificado pelo aumento do número de cópias do óperon *kdp* ou de um ou mais genes que constituem o óperon *kdp*, ou pela modificação de uma sequência de controle da expressão do óperon.

15 4. Microorganismo de acordo com as reivindicações 2 ou 3, caracterizado pelo fato de que o óperon *kdp* contém pelo menos os genes *kdpA*, *kdpB* e *kdpC*.

20 5. Microorganismo de acordo com a reivindicação 4, caracterizado pelo fato de que o gene *kdpA* é um gene que codifica uma proteína tendo a sequência de aminoácidos mostrada nas SEQ ID NOS: 2 ou 8, ou a subunidade A do sistema *kdp* tendo a sequência de aminoácidos das SEQ ID NOS: 2 ou 8 incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

25 6. Microorganismo de acordo com as reivindicações 4 ou 5, caracterizado pelo fato de que o gene *kdpB* é um gene que codifica uma proteína tendo a sequência de aminoácidos mostrada nas SEQ ID NOS: 3 ou 9, ou a subunidade B do sistema *kdp* tendo uma sequência de aminoácidos de SEQ ID NOS: 3 ou 9 incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

7. Microorganismo de acordo com qualquer uma das

reivindicações 4 a 6, caracterizado pelo fato de que o gene *kdpC* é um gene que codifica uma proteína tendo a sequência de aminoácidos mostrada nas SEQ ID NOS: 4 ou 10, ou a subunidade C do sistema *kdp* tendo uma sequência de aminoácidos das SEQ ID NOS: 4 ou 10, incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

8. Microorganismo de acordo com qualquer uma das reivindicações 4 a 7, caracterizado pelo fato de que o óperon *kdp* é um DNA definido em qualquer um dos seguintes itens (a) a (d):

(a) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1,

(b) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1 ou uma sonda preparada da sequência de nucleotídeos sob condições estringentes e codificando o sistema *kdp*,

(c) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7,

(b) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7 ou uma sonda preparada da sequência de nucleotídeos sob condições estringentes e codificando o sistema *kdp*.

9. Microorganismo de acordo com qualquer uma das reivindicações 1 a 8, caracterizado pelo fato de que o L-aminoácido é uma ou mais espécies de L-aminoácidos selecionados do grupo consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoleucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-triptofano e L-cisteína.

10. Microorganismo de acordo com qualquer uma das reivindicações 1 a 9, caracterizado pelo fato de que o microorganismo pertencente à família *Enterobacteriaceae* é uma bactéria *Escherichia*, uma bactéria *Enterobacter* ou uma bactéria *Pantoea*.

11. Método para produzir um L-aminoácido, caracterizado pelo fato de que compreende cultivar o microorganismo como definido em qualquer uma das reivindicações 1 a 10 em um meio para produzir e acumular um L-aminoácido no meio ou nas células, e coletar o L-aminoácido do meio ou das células.

12. Método de acordo com a reivindicação 11, caracterizado pelo fato de que o L-aminoácido é uma ou mais espécies de L-aminoácidos selecionadas do grupo consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoleucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-triptofano e L-cisteína.

FIG. 1

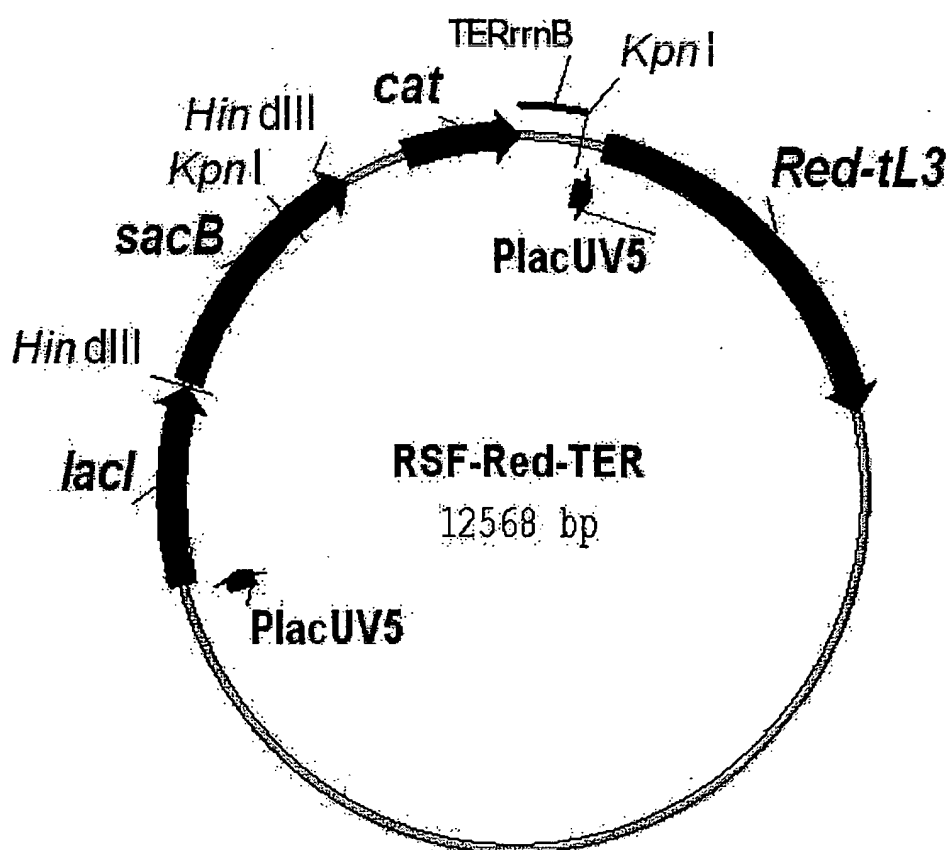


FIG. 2

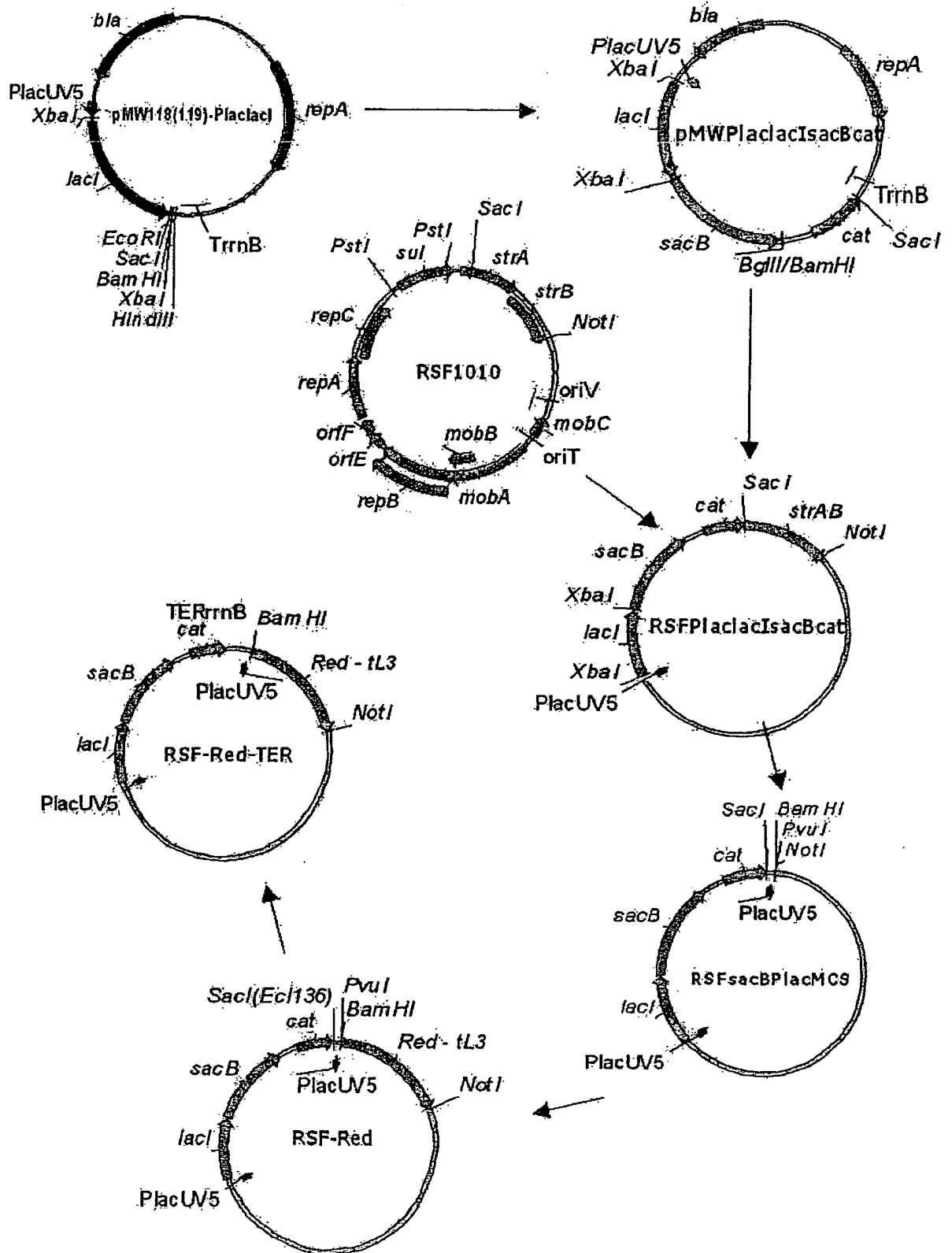


FIG. 3

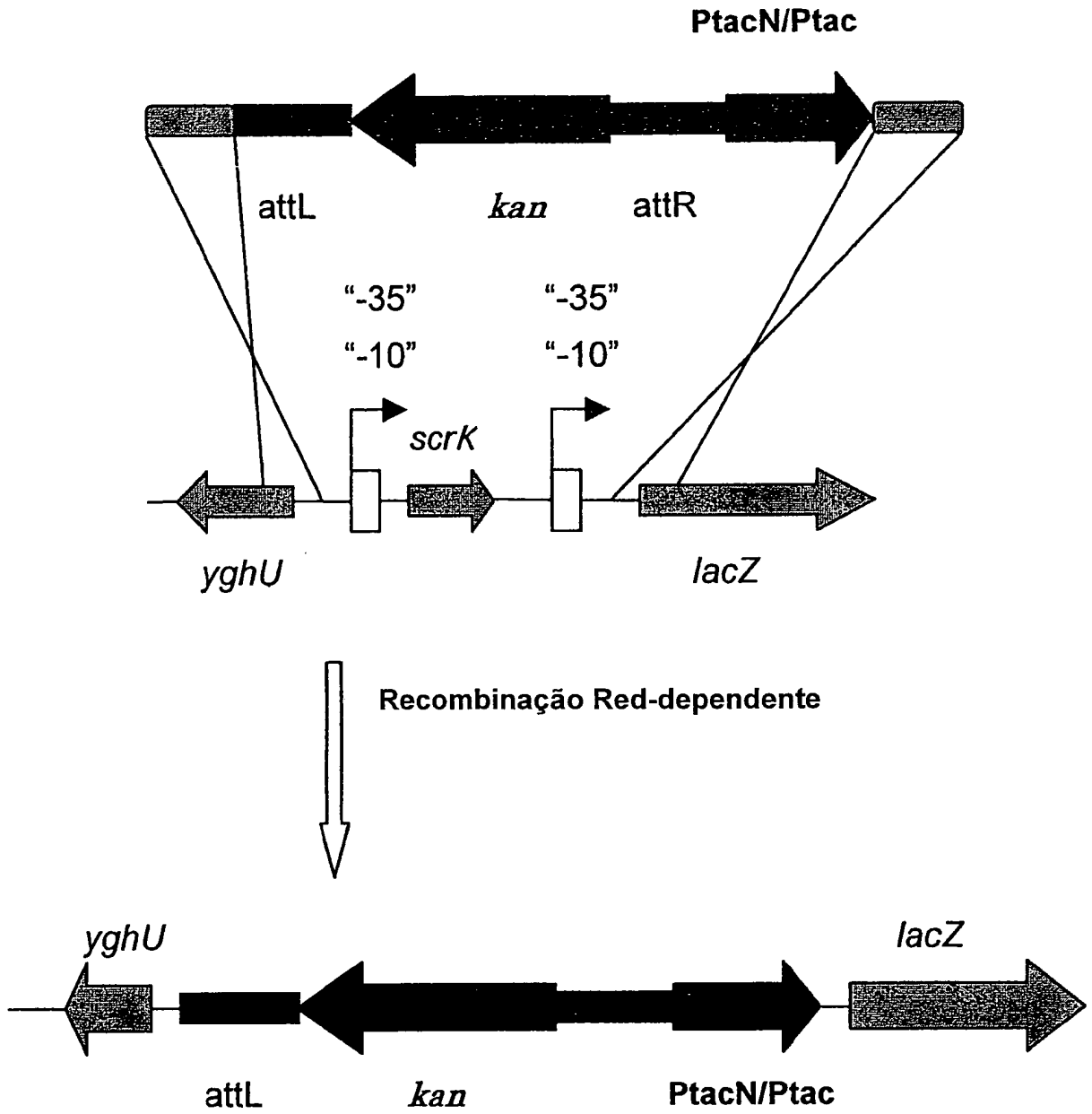


FIG. 4

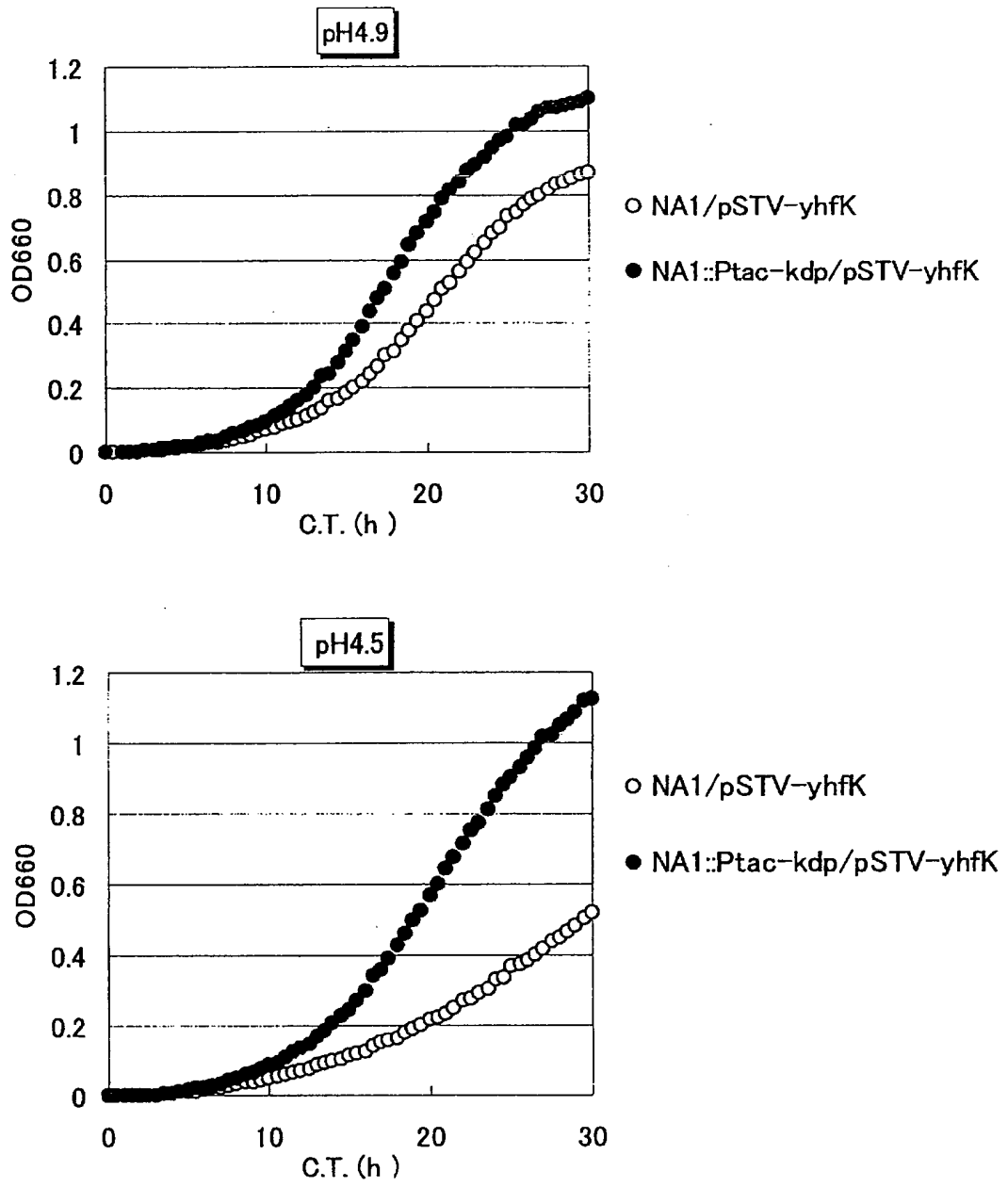


FIG. 5

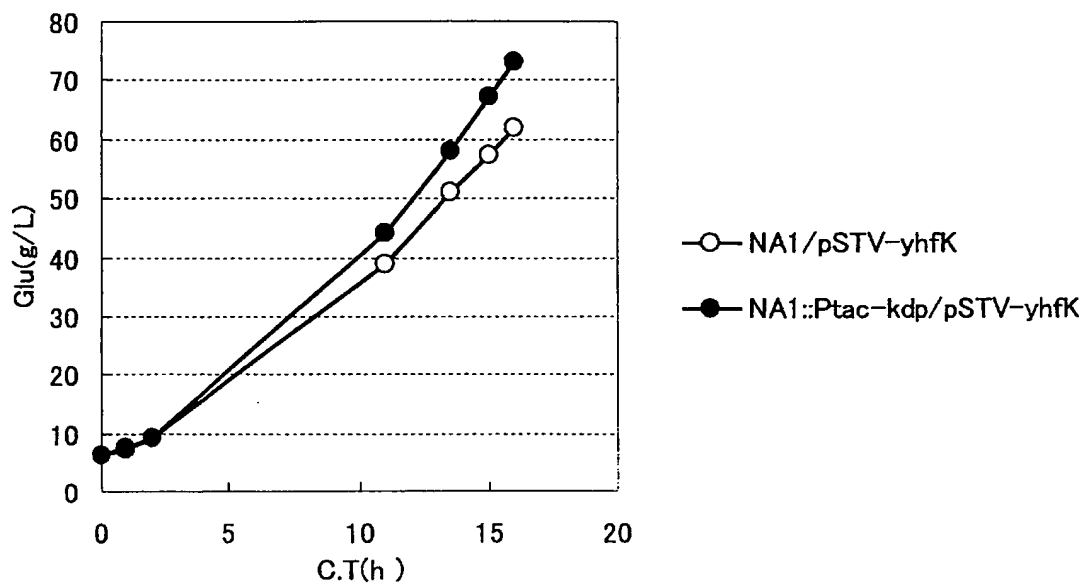
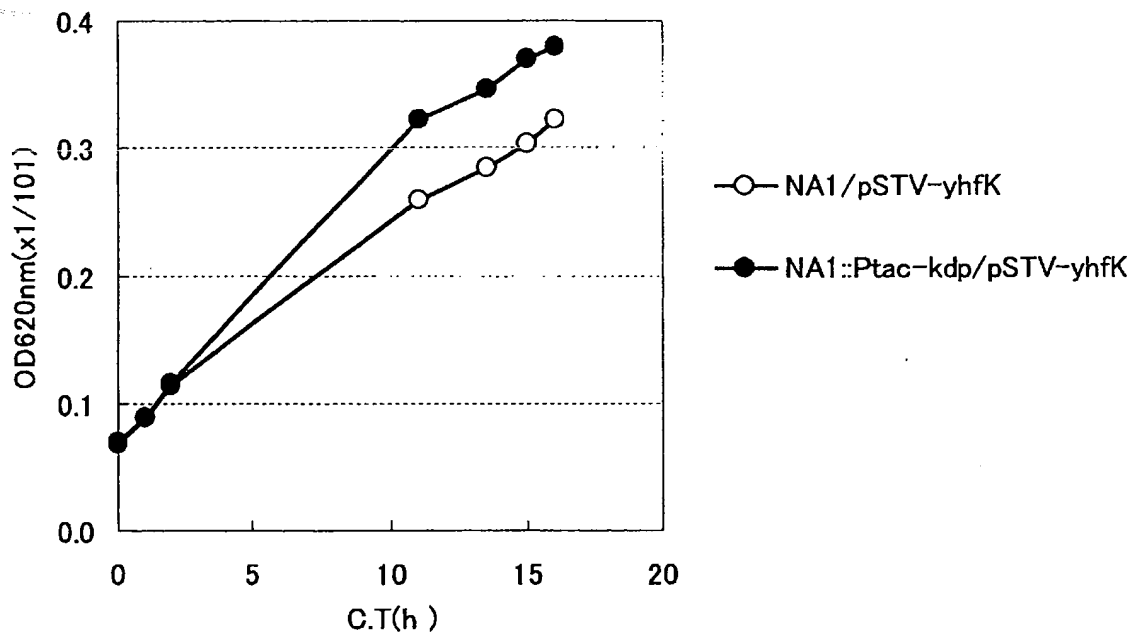


FIG. 6

		1	50
KdpA_P. ananatis	(1)	MAANAFLLIAVYLLLLMVMAQPLGRGLAALVADKPLFAR---AEALLWRFS	
KdpA_E. coli	(1)	MAAQGFLLIATFLLVLMVLARPLGSGLARLINDIPLPGTTGVERVLFRAL	
Consenso		MAAXFLLIAXLLXLMVXAXPLGXGLAXLXXDXPLXXXXXXEXLXRX	
		51	100
KdpA_P. ananatis	(49)	GVQEGMRWQHLYLLATLVFNLLGFVVLLAILMFQALPLNPQHLPLGLSWD	
KdpA_E. coli	(51)	GVSDREMWNWKYLCAILGLNMLGLAVLFFMLLGQHYLPLNPQQLPGLSWD	
Consenso		GVXXXXMXWXXYLXAILXXNXLGGXXVXXXXLXXQXXLPLNPQXLPLGLSWD	
		101	150
KdpA_P. ananatis	(99)	LALNTAISFVTNTNWQSYAGESTLSYFSQMVGLTVQNFVSAATGIAVAF	
KdpA_E. coli	(101)	LALNTAVSFVTNTNWQSYSGETLSYFSQAGLTVQNFVSAASGIAVIFA	
Consenso		LALNTAXSFVTNTNWQSYXGEXTLSYFSQMXGLTVQNFVSAAXGIAVXFA	
		151	200
KdpA_P. ananatis	(149)	LIRGFANRSVATLGNAWRDLTRITLVLLPISLLMALFFVSQGSIQNPLP	
KdpA_E. coli	(151)	LIRAFTRQSMSTLGNAWVDLLRITLVVLPVALLIALFFIQQALQNPLP	
Consenso		LIRFXFXSXXTLGNAWDXLRITLVLPVALLIALFFXXQGXQXNPLP	
		201	250
KdpA_P. ananatis	(199)	YHNVTSLEGAQQTLAMGPVASQEAIKMLGTNGGGFFNVNSAHPFENPTAL	
KdpA_E. coli	(201)	YQAVNTVEGAQQLPMGPVASQEAIKMLGTNGGGFFNANSSHPFENPTAL	
Consenso		YXXVXXEQAQQLXMGPVASQEAIKMLGTNGGGFFNXSHPFENPTAL	
		251	300
KdpA_P. ananatis	(249)	SNFVQMLSIFLIPAALCFAGFESVKDRRQGSMLLWSMTLMFVVAALVMW	
KdpA_E. coli	(251)	TNFVQMLAIFLIPTALCFAGFVEMGDRRQGRMLLWAMSVIFVICVGVVMW	
Consenso		XNFVQMLXIFLIPXALCFAGFEXXDRRQGXMLLWXXXXFVXXXXVMW	
		301	350
KdpA_P. ananatis	(299)	AELRGNPHFLTLAGDSAINMEGKETRFGILNSSLFAVITTAASCGAVNAM	
KdpA_E. coli	(301)	AEVQGNPHLLALGTDSSINMEGKESRFGVLVSSLFAVITTAASCGAVIAM	
Consenso		AEXXGNPHXLXGDSXINMEGKEXRFGXLXSSLFAVITTAASCGAVXAM	
		351	400
KdpA_P. ananatis	(349)	HDSFTALGGMVPMMLMQLGEVVFGGAGLYGMLLFVLLAVFIAGLMIGR	
KdpA_E. coli	(351)	HDSFTALGGMVPMWMLQIGEVVFGGVSGLYGMMLFVLLAVFIAGLMIGR	
Consenso		HDSFTALGGMVPMXLMQXGEVVFGGAGLYGMMLFVLLAVFIAGLMIGR	
		401	450
KdpA_P. ananatis	(399)	TPEFLGKKIDVWEMKMTALAILVTPALVLIPTAIAMMTDAGRAGMANPGT	
KdpA_E. coli	(401)	TPEYLGKKIDVREMKLTAALVTPALVLMGAALAMMTDAGRSAMLNPGP	
Consenso		TPEXLGKKIDVXEMKXTALAILVTPXVLXGXAXAMMTDAGRXXMXNPGX	
		451	500
KdpA_P. ananatis	(449)	HGFSEVLYAVSSAANNNGSAFAGLNANTPFWNLLAVCMFVGRFGVIIIPV	
KdpA_E. coli	(451)	HGFSEVLYAVSSAANNNGSAFAGLSANSPPFNCLLAFCMFVGRFGVIIIPV	
Consenso		HGFSEVLYAVSSAANNNGSAFAGLXANXPPFNXLLAXCMFVGRFGXIIIPV	
		501	550
KdpA_P. ananatis	(499)	MAIAGAMAVKKVQPVNGTLPHTGPLFIALLVGTVLLVGALTFIPALALG	
KdpA_E. coli	(501)	MAIAGSLVSKKSQAASSGTLPTHGPLFVGLLIGTVLLVGALTFIPALALG	
Consenso		MAIAGXXXXKXQXXXGTLPTHGPLFXXLLXGTVLLVGALTFIPALALG	
		551	562
KdpA_P. ananatis	(549)	PVAEHLQLIQGQ	
KdpA_E. coli	(551)	PVAEYLS	
Consenso		PVAEXLX	

FIG. 7

KdpB_P. ananatis	(1)	1	MSR-QQVFDAALLRRTSAIDAVKKLDPRVQFRNPVMFVVYLGSILTSILA	50
KdpB_E. coli	(1)		MSRKQLALFEPTLVVQALKEAVKKLNPAQWRNPVMFIVWIGSLLTTCIS	
Consenso			MSRXQXXXFXXXLXXXXXXXXAVKKLXPXXQXRNPMFVXXGSXLTXXX	
KdpB_P. ananatis	(50)	51	IMMFTGHQSGSASFTGAIALWLWFTVLFANMAEALAEGRSKAQANSLKGV	100
KdpB_E. coli	(51)		IAMASGAMPGNALFSAAISGWLWITVLFANFAEALAEGRSKAQANSLKGV	
Consenso			IXMXXGXXXGXAFXXAIXXWLWXTVLFANXAEALAEGRSKAQANSLKGV	
KdpB_P. ananatis	(100)	101	KKTSFAKKLSAAHYGAAWQQAADALRKGDVAVLVEAGDVIPCDGEVVEGG	150
KdpB_E. coli	(101)		KKTAFARKLREPKYGAAADKVPADQLRKGDIVLVEAGDIIPCDGEVIEGG	
Consensus			KKTXFAXKLXXXXYGAAXXVXADXLRKGDVXVLEAGDXIPCDGEVXEGG	
KdpB_P. ananatis	(150)	151	ASVDESAITGESAPVIRESGGDFASVTGGTRILSDWLVIITCSANPGETFL	200
KdpB_E. coli	(151)		ASVDESAITGESAPVIRESGGDFASVTGGTRILSDWLVIIECSVNPGETFL	
Consenso			ASVDESAITGESAPVIRESGGDFASVTGGTRILSDWLVIIXCSXNPGETFL	
KdpB_P. ananatis	(200)	201	DRMIAMVEGAQRRKTPNEIALTILLVSLTIVFLLATVTLWPFSAWGGTPV	250
KdpB_E. coli	(201)		DRMIAMVEGAQRRKTPNEIALTILLIALTIVFLLATATLWPFSAWGGNAV	
Consenso			DRMIAMVEGAQRRKTPNEIALTILLXXLTIVFLLATXTLWPFSAWGGXXV	
KdpB_P. ananatis	(250)	251	TITVLVALLVCLIPTTIGLLSAIGVAGMSRMLGANVIATSGRAVEAAGD	300
KdpB_E. coli	(251)		SVTVLVALLVCLIPTTIGLLSAIGVAGMSRMLGANVIATSGRAVEAAGD	
Consenso			XXTVLVALLVCLIPTTIGLLSAIGVAGMSRMLGANVIATSGRAVEAAGD	
KdpB_P. ananatis	(300)	301	VDVLMLDKTGTITLGNRQATQFLPAPGVTEEQLADAAQLASLADETPEGR	350
KdpB_E. coli	(301)		VDVLLLDKTGTITLGNRQASEFIPAQGVDEKTLADAAQLASLADETPEGR	
Consenso			VDVXLDKTGTITLGNRQAXXFXPAXGVXEXXLADAAQLASLADETPEGR	
KdpB_P. ananatis	(350)	351	SIVVLAKQKFNLRERDLSSMGASFIPFSAQTRMSGVNVQDRLIRKGAVDA	400
KdpB_E. coli	(351)		SIVILAKQRFNLRERDVQSLHATFVPFTAQSRMSGINIDNRMIRKGSVDA	
Consenso			SIVXLAKQXFNLRERDXXSXXAXFPFXAQXRMSGNXXXRXIRKGVDA	
KdpB_P. ananatis	(400)	401	VRRHIEASHGAFPAEVNARVEEVARAGGTPLVVAEGAKVLGVVALKDIVK	450
KdpB_E. coli	(401)		IRRHVEANGGHFPTVDQKVDQVARQGATPLVVVEGSRVLGVIALKDIVK	
Consenso			RRRHXEAXXGXFPXXVXXVXXVARXGXTPLVVXEGXXVLGVXALKDIVK	
KdpB_P. ananatis	(450)	451	GGIKERFAELRKMGIKTVMITGDNPLTAAAIAAEAGVDDFLSEATPEAKL	500
KdpB_E. coli	(451)		GGIKERFAQLRKMGIKTVMITGDNRLTAAAIAAEAGVDDFLAEATPEAKL	
Consenso			GGIKERFAXLRKMGIKTVMITGDNXLTAAAIAAEAGVDDFLXEATPEAKL	
KdpB_P. ananatis	(500)	501	ALIRQYQAEGRVAMTGDGTNDAPALAQADVAVAMNSGTQAAKEAGNMVD	550
KdpB_E. coli	(501)		ALIRQYQAEGRVAMTGDGTNDAPALAQADVAVAMNSGTQAAKEAGNMVD	
Consenso			ALIRQYQAEGRVAMTGDGTNDAPALAQADVAVAMNSGTQAAKEAGNMVD	
KdpB_P. ananatis	(550)	551	LDSNPTKLEVVHIGKQMLMTRGSLTTFSIANDVAKYFAIIPAAFAATYP	600
KdpB_E. coli	(551)		LDSNPTKLEVVHIGKQMLMTRGSLTTFSIANDVAKYFAIIPAAFAATYP	
Consenso			LDSNPTKLEVVHIGKQMLMTRGSLTTFSIANDVAKYFAIIPAAFAATYP	
KdpB_P. ananatis	(600)	601	QLNMLNVMQLHSPASAILS AVIFNALVIVFLIPLALKGVSYPPLSASLL	650
KdpB_E. coli	(601)		QLNALNIMCLHSPDSAILS AVIFNALIIVFLIPLALKGVSYPPLTASAML	
Consenso			QLNXLNXMXLHSPXSAAILS AVIFNALXIVFLIPLALKGVSYPPLXAXXL	
KdpB_P. ananatis	(650)	651	RRNLLIYGLGGLLVPFVGIKVIDMLLVLSGMA	683
KdpB_E. coli	(651)		RRNLWIYGLGGLLVPFVGIKVIDLLLTVCGLV	
Consenso			RRNLXIYGLGGLLVPFVGIKVIDXLLXXXGX	

FIG. 8

		1		50
KdpC_P. ananatis	(1)	MSQLRPAIFLLLLLTVVCGVVYPLLTTGLSQLLFPWQANGSVLNVDGEER		
KdpC_E. coli	(1)	MSGLRPALSTFIFLLLI TGGVYPLLTTVLGQWFFPWQANGSLIREGDTVR		
Consenso		MSXLRPAXXXXXLXXXXGXVYPLLTTXLXQXXFPWQANGSXXXXXXXX		
		51		100
KdpC_P. ananatis	(51)	GSALIGQNFSQPGYFWGRPSATGDKPYNPLASSGSNLAASNPA LDKAVAE		
KdpC_E. coli	(51)	GSALIGQNFTGNGYFHGRPSATAEMPYNPQASGGSNLAVSNPEL DKLIAA		
Consenso		GSALIGQNFXXXGYFXGRPSATXXXPYNPXASXGSNLAXSNPXL DXXXAX		
		101		150
KdpC_P. ananatis	(101)	RVAALRTANPQANGAVPVELVTTSASGLDPEISPEAALWQAPRIAAARQL		
KdpC_E. coli	(101)	RVAALRAANPDASASVPVELVTASASGLDNNITPQAAAWQIPRVAKARNL		
Consenso		RVAALRXANPAXXXVPVELVTXSASGLDXXIXPAAAXWQXPRXAXARXL		
		151		191
KdpC_P. ananatis	(151)	PLAKVDALVDSMTQRPLLPFIGEPTVNVLQLNLALNDLK		
KdpC_E. coli	(151)	SVEQLTQLIAKYSQQPLVKYIGQPVVNIVELNLALDKLDE		
Consenso		XXXXXXXXLXXXXXQXPLXXXIGXPXVNXNLALXXLX		

RESUMO

“MICROORGANISMO, E, MÉTODO PARA PRODUZIR UM L-AMINOÁCIDO”

Um L-aminoácido pode ser produzido por: cultivar, em um
5 meio de cultura, um microorganismo pertencente à família
Enterobacteriaceae, a qual é capaz de produzir o L-aminoácido e é
modificada de modo que o sistema kdp possa ser potencializado, por esse
meio produzindo e acumulando o L-aminoácido no meio de cultura ou em
uma célula do microorganismo; e coletando-se o L-aminoácido do meio de
10 cultura ou da célula.

A requerente apresenta novas vias das páginas 5, 8, 15 e 68 do relatório descritivo e também novas vias das reivindicações para conformar o presente pedido com o pedido internacional.

deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

(8) O microorganismo acima mencionado, em que o óperon de *kdp* é um DNA definido em qualquer um dos seguintes itens (a) a (d):

5 (a) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1,

(b) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1 ou uma sonda preparada com a sequência de nucleotídeos sob condições estringentes e codificando o sistema *kdp*,

10 (c) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7,

(d) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7 ou uma sonda preparada da sequência de nucleotídeos sob condições estringentes e codificando o sistema *kdp*.

15 (9) O microorganismo acima mencionado, em que o L-aminoácido é uma ou mais espécies de L-aminoácidos selecionados do grupo consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoleucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-triptofano e L-cisteína.

(10) O microorganismo acima mencionado, em que o microorganismo pertencente à família *Enterobacteriaceae* é uma bactéria *Escherichia*, uma bactéria *Enterobacter* ou uma bactéria *Pantoea*.

25 (11) Um método para produzir um L-aminoácido, que compreende cultivar o microorganismo acima mencionado em um meio para produzir e acumular um L-aminoácido no meio ou nas células, e coletar o L-aminoácido do meio ou das células.

(12) O método supramencionado, em que o L-aminoácido constitui uma ou mais espécies de L-aminoácidos selecionados do grupo

<1-1> TRANSMISSÃO DA CAPACIDADE DE PRODUZIR L-AMINOÁCIDOS

Exemplos dos métodos para transmitir a capacidade de produzir L-aminoácidos e microorganismos utilizáveis na presente invenção, aos quais a capacidade de produzir L-aminoácidos é comunicada, serão descritos abaixo. Entretanto, o microorganismo não fica limitado a estes, contanto que um microorganismo tendo a capacidade de produzir L-aminoácidos seja usado.

Microorganismos usados para a presente invenção incluem opcionalmente substituído microorganismos pertencentes aos gêneros *Escherichia*, *Enterobacter*, *Pantoea*, *Klebsiella*, *Serratia*, *Erwinia*, *Salmonella*, *Morganella*, ou similar, contanto que eles pertençam à família *Enterobacteriaceae* e tenham uma capacidade de produzir L-aminoácido. Em particular as bactérias classificadas dentro da família *Enterobacteriaceae* de acordo com a taxonomia usada pela base de dados da NCBI (National Center for Biotechnology Information) (<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=91347>) podem ser usadas. Como cepas precursoras das *Enterobacteriaceae* que devam ser modificadas, as bactérias pertencentes aos gêneros *Escherichia*, *Enterobacter*, *Pantoea*, *Erwinia* ou *Klebsiella* são preferivelmente usadas.

A cepa precursora das bactérias de *Escherichia* usadas de modo a se obter uma bactéria de *Escherichia* da presente invenção, não é particularmente limitada. Aquelas descritas no trabalho de Neidhardt *et al.* (Backmann, B. J., 1996. *Derivations and Genotypes of some mutant derivatives of Escherichia coli K-12*, pp. 2460-2488, Tabela 1, em F. D. Neidhardt (ed.), *Escherichia coli and Salmonella Cellular and Molecular Biology/Segunda Edição*, American Society for Microbiology Press, Washington, D.C.) podem ser utilizadas. Entre elas, por exemplo, a

transformados pelo método do pulso elétrico (Patente Japonesa Aberta ao Público nº 2-207791).

5 O número de cópias de um gene pode também ser aumentado pela introdução de cópias múltiplas do gene no DNA cromossômico do microorganismo, o que pode ser realizado por recombinação homóloga (Miller, J. H. *Experiments in Molecular Genetics*, 1972, Cold Spring Harbor Laboratory) com o uso de cópias múltiplas de uma sequência como alvos no DNA cromossômico. As sequências presentes nas cópias múltiplas no DNA cromossômico incluem os DNAs respectivos, e as repetições invertidas presentes na extremidade de um elemento transponível. Igualmente, como apresentado na Patente Japonesa Aberta ao Público nº 2-109985, é possível incorporar o gene alvo em um transpóson, e possibilitar que ele seja transferido para introduzir cópias múltiplas do gene para o DNA cromossômico. O gene alvo pode também ser introduzido no cromossoma bacteriano por fago Mu (Patente Japonesa Aberta ao Público nº 2-109985), ou similar.

O segundo método é aumentar a expressão do gene alvo pela substituição de uma sequência reguladora da expressão do gene alvo, tal como um promotor, sobre o DNA cromossômico ou plasmídeo com um promotor mais forte. Por exemplo, o promotor lac, o promotor trp, o promotor trc, o promotor PR, o promotor lacUV5, etc., são conhecidos como promotores fortes. Além disso, é também possível substituir vários nucleotídeos na região do promotor de um gene, de modo que o promotor seja modificado para ser mais forte, como apresentado na Publicação da Patente Internacional WO 00/18935. Exemplos de promotores fortes e de métodos para avaliar a intensidade dos promotores, são descritos em um artigo de Goldstein *et al.* (*Prokaryotic promoters in biotechnology. Biotechnol. Annu. Rev.*, 1995, 1, 105-128), etc.

A substituição de uma sequência reguladora da expressão pode

SEQ ID NO: 51: Iniciador para amplificação da sequência de montante do óperon *kdp* ligada ao promotor *tac*

SEQ ID NO: 52: Iniciador para amplificação da sequência de montante do óperon *kdp* ligado ao promotor *tac*

5 SEQ ID NO: 53: Iniciador para confirmar a estrutura de montante do óperon *kdp*

SEQ ID NO: 54: Iniciador para confirmar a estrutura de montante do óperon *kdp*

SEQ ID NO: 55: Iniciador para amplificação do óperon *kdp*

10 SEQ ID NO: 56: Iniciador para amplificação do óperon *kdp*

SEQ ID NO: 57: Sequência de consenso das sequências de aminoácidos KdpA de *Pantoea ananatis* e *Escherichia coli*

SEQ ID NO: 58: Sequência de consenso das sequências de aminoácidos KdpB de *Pantoea ananatis* e *Escherichia coli*

15 SEQ ID NO: 59: Sequência de consenso das sequências de aminoácidos KdpC de *Pantoea ananatis* e *Escherichia coli*

APLICABILIDADE INDUSTRIAL

Mediante o uso do microorganismo da presente invenção, um L-aminoácido tal como o ácido L-glutâmico, L-lisina, L-treonina, L-arginina, 20 L-histidina, L-isoleucina, L-valina, L-leucina, L-treonina, L-fenilalanina, L-tirosina, L-triptofano ou L-cisteína ou similar, pode ser eficientemente produzido por fermentação. Em uma forma de realização preferida, o microorganismo da presente invenção apresenta tanto uma quantidade de produção de L-aminoácido quanto o índice de produção, superiores.

REIVINDICAÇÕES

1. Microorganismo pertencente à família *Enterobacteriaceae*, caracterizado pelo fato de que tem uma capacidade de produção de L-aminoácidos e foi modificado de modo que o sistema *kdp* fosse intensificado.

5 2. Microorganismo de acordo com a reivindicação 1, caracterizado pelo fato de que o sistema *kdp* é intensificado pelo aumento da expressão do óperon *kdp* ou de um ou mais genes que constituem o óperon *kdp*, e/ou pelo aumento da translação do óperon *kdp* ou dos genes.

10 3. Microorganismo de acordo com a reivindicação 2, caracterizado pelo fato de que o sistema *kdp* é intensificado pelo aumento do número de cópias do óperon *kdp* ou de um ou mais genes que constituem o óperon *kdp*, ou pela modificação de uma sequência de controle da expressão do óperon.

15 4. Microorganismo de acordo com as reivindicações 2 ou 3, caracterizado pelo fato de que o óperon *kdp* contém pelo menos os genes *kdpA*, *kdpB* e *kdpC*.

20 5. Microorganismo de acordo com a reivindicação 4, caracterizado pelo fato de que o gene *kdpA* é um gene que codifica uma proteína tendo a sequência de aminoácidos mostrada nas SEQ ID NOS: 2 ou 8, ou a subunidade A do sistema *kdp* tendo a sequência de aminoácidos das SEQ ID NOS: 2 ou 8 incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

25 6. Microorganismo de acordo com as reivindicações 4 ou 5, caracterizado pelo fato de que o gene *kdpB* é um gene que codifica uma proteína tendo a sequência de aminoácidos mostrada nas SEQ ID NOS: 3 ou 9, ou a subunidade B do sistema *kdp* tendo uma sequência de aminoácidos de SEQ ID NOS: 3 ou 9 incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

7. Microorganismo de acordo com qualquer uma das

reivindicações 4 a 6, caracterizado pelo fato de que o gene *kdpC* é um gene que codifica uma proteína tendo a sequência de aminoácidos mostrada nas SEQ ID NOS: 4 ou 10, ou a subunidade C do sistema *kdp* tendo uma sequência de aminoácidos das SEQ ID NOS: 4 ou 10, incluindo substituições, deleções, inserções ou adições de um ou vários resíduos de aminoácidos.

8. Microorganismo de acordo com qualquer uma das reivindicações 4 a 7, caracterizado pelo fato de que o óperon *kdp* é um DNA definido em qualquer um dos seguintes itens (a) a (d):

(a) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1,

(b) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 546 a 4871 da SEQ ID NO: 1 ou uma sonda preparada da sequência de nucleotídeos sob condições estridentes e codificando o sistema *kdp*,

(c) um DNA compreendendo a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7,

(d) um DNA que hibridiza com a sequência de nucleotídeos dos nucleotídeos números 543 a 4853 da SEQ ID NO: 7 ou uma sonda preparada da sequência de nucleotídeos sob condições estridentes e codificando o sistema *kdp*.

9. Microorganismo de acordo com qualquer uma das reivindicações 1 a 8, caracterizado pelo fato de que o L-aminoácido é uma ou mais espécies de L-aminoácidos selecionados do grupo consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoleucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-triptofano e L-cisteína.

10. Microorganismo de acordo com qualquer uma das reivindicações 1 a 9, caracterizado pelo fato de que o microorganismo pertencente à família *Enterobacteriaceae* é uma bactéria *Escherichia*, uma bactéria *Enterobacter* ou uma bactéria *Pantoea*.

11. Método para produzir um L-aminoácido, caracterizado pelo fato de que compreende cultivar o microorganismo como definido em qualquer uma das reivindicações 1 a 10 em um meio para produzir e acumular um L-aminoácido no meio ou nas células, e coletar o L-aminoácido do meio ou das células.

5

12. Método de acordo com a reivindicação 11, caracterizado pelo fato de que o L-aminoácido é uma ou mais espécies de L-aminoácidos selecionadas do grupo consistindo de ácido L-glutâmico, L-lisina, L-treonina, L-arginina, L-histidina, L-isoleucina, L-valina, L-leucina, L-fenilalanina, L-tirosina, L-triptofano e L-cisteína.

10