



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

(11) PI 0507083-0 B1

(22) Data do Depósito: 18/01/2005

(45) Data de Concessão: 21/06/2016



(54) Título: PROCESSO PARA PREPARAÇÃO DE L-TREONINA E L-LISINA POR FERMENTAÇÃO DE MICRO-ORGANISMOS RECOMBINANTES DA FAMÍLIA ENTEROBACTERIACEAE

(51) Int.Cl.: C12N 15/60; C12N 9/88; C12P 13/08

(30) Prioridade Unionista: 23/01/2004 DE 10 2004 003 410.9

(73) Titular(es): EVONIK DEGUSSA GMBH

(72) Inventor(es): MECHTHILD RIEPING

Relatório Descritivo da Patente de Invenção para "**PROCESSO PARA PREPARAÇÃO DE L-TREONINA E L-LISINA POR FERMENTAÇÃO DE MICRO-ORGANISMOS RECOMBINANTES DA FAMÍLIA *ENTEROBACTERIACEAE***".

5 Campo da Invenção

A presente invenção se refere a um processo para preparação de L-aminoácidos, especificamente L-treonina, usando microorganismos recombinantes da família *Enterobacteriaceae*, onde o gene *eno* é melhorado, especificamente superexpresso e aos seus microorganismos.

10 Antecedentes da Invenção

L-aminoácidos, especificamente L-treonina são usados na medicina humana e na indústria farmacêutica, na indústria alimentícia e muito especificamente na nutrição animal.

É sabido que os L-aminoácidos são preparados por fermentação de cepas da *Enterobacteriaceae*, especificamente *Escherichia coli* (*E. coli*) e *Serratia marcescens*. Devido a sua grande importância, tem havido um trabalho constante de aperfeiçoamento dos processos de preparação. Os aperfeiçoamentos ao processo podem se relacionar aos meios de fermentação, tais como, por exemplo, agitação e fornecimento de oxigênio ou à composição dos meios nutrientes, tais como, a concentração de açúcar durante a fermentação ou a elaboração da forma do produto, por exemplo, por cromatografia de troca de íon ou as propriedades de output intrínsecas do microorganismo propriamente.

Os métodos de mutagênese, seleção e seleção de mutante são empregados para aperfeiçoar as propriedades de saída desses microorganismos. As cepas que são resistentes aos antimetabólitos, tais como, por exemplo, ácido α -amino- β -hidroxivalérico análogo a treonina (AVH) ou são auxotróficas para metabólitos ou de importância reguladora e produzem L-aminoácido, tais como, por exemplo, L-treonina são obtidas dessa maneira.

Os métodos da técnica de DNA recombinante também foram empregados por alguns anos para aperfeiçoar a resistência das cepas da família *Enterobacteriaceae* que produz L-aminoácidos, por ampliação dos

genes de biossíntese do aminoácido individuais e investigação do efeito na produção. Informação resumida da célula e biologia molecular do *Escherichia coli* e *Salmonella* devem ser obtidas em Neidhardt (ed): *Escherichia coli and Salmonella, Cellular and Molecular Biology*, 2ª Edição, ASM Press, Washington, D.C., USA (1996).

Objetivo da Invenção

O objetivo da invenção é prover novos meios para aperfeiçoar a preparação fermentativa de L-aminoácidos, especificamente da L-treonina.

Sumário da Invenção

10 A invenção provê microorganismos recombinantes da família *Enterobacteriaceae* que contêm um gene *eno* melhorado ou superexpresso, que codifica a enolase ou alelos da mesma, e que produz L-aminoácidos, especificamente L-treonina em um modo aperfeiçoado.

15 O ponto de partida para a comparação se constitui, em cada caso, nos microorganismos que não são recombinantes para o gene *eno* e não contêm o gene *eno* melhorado.

Esses microorganismos incluem, especificamente, microorganismos da família *Enterobacteriaceae* nos quais um polinucleotídeo que codifica um polipeptídeo, a sequência de aminoácido do qual é idêntica a extensão de pelo menos 90%, especificamente à extensão de pelo menos 95%, preferivelmente à extensão de 98% ou pelo menos 99%, específica e preferivelmente à extensão de 99,7% e muito específica e preferivelmente à extensão de 100% em relação a uma sequência de aminoácido escolhida do grupo consistindo essencialmente na SEQ ID Nº. 4, SEQ ID Nº. 6, e SEQ ID Nº. 8 é melhorado. As sequências de aminoácido que são idênticas aquelas das SEQ ID Nº. 4, SEQ ID Nº. 6 ou SEQ ID Nº.8 são preferidas.

Os microorganismos mencionados contêm polinucleotídeos melhorados ou superexpressos escolhidos do grupo consistindo em:

a) polinucleotídeo com a sequência de nucleotídeo SEQ ID Nº. 3, SEQ ID Nº. 5 ou SEQ ID Nº. 7;

b) polinucleotídeo com uma sequência de nucleotídeo que corresponde às SEQ ID Nº. 3, SEQ ID Nº. 5 ou SEQ ID Nº. 7 dentro do contex-

to de degeneração do código genético;

c) seqüência de polinucleotídeo com uma seqüência que hibridiza sob condições estridentes com a seqüência complementar às SEQ ID Nº. 3, SEQ ID Nº. 5 ou SEQ ID Nº. 7;

5 d) polinucleotídeo com uma seqüência SEQ ID Nº. 3, SEQ ID Nº. 5 ou SEQ ID Nº. 7 que contém mutantes de sentido de função neutra, onde os polinucleotídeos codificam enolase.

A invenção também provê um processo para a preparação fermentativa de L-aminoácidos, especificamente L-treonina, usando microor-
 10 ganismos recombinantes da família da *Enterobacteriaceae* que, especificamente, já produzem L-aminoácidos e nos quais, pelo menos o gene eno ou as seqüências de nucleotídeo que codificam o produto de gene dos mesmos é ou são melhoradas.

Descrição Detalhada da Invenção

15 Os microorganismos descritos são preferivelmente empregados.

Quando os L-aminoácidos ou aminoácidos forem mencionados a seguir, isso significa um ou mais aminoácidos incluindo seus sais, escolhidos do grupo consistindo em L-asparagina, L-treonina, L-serina, L-glutamato, L-glicina, L-alanina, L-cisteína, L-valina, L-metionina, L-prolina,
 20 L-isoleucina, L-leucina, L-tirosina, L-fenilalanina, L-histidina, L-lisina, L-triptofano, L-arginina e L-homosserina. L-Treonina é especificamente preferida.

O termo "melhoramento" com relação a isso descreve o aumento na atividade intracelular ou concentração de uma ou mais enzimas ou prote-
 25 ínas em um microorganismo que são codificadas pelo DNA correspondente, por exemplo, por aumento do número de cópias do gene ou genes em pelo menos um (1) cópia, ligando funcionalmente um promotor potente ao gene ou usando um gene ou alelo que codifica uma enzima correspondente ou proteína com uma atividade alta e opcionalmente combinando esses meios.

30 Os alelos são em geral entendidos como significando formas alternativas de um dado gene. As formas são distinguidas por diferenças na seqüência de nucleotídeo.

A proteína codificada por uma sequência de nucleotídeo, isto é, uma ORF, um gene ou um alelo ou o ácido ribonucléico codificado é em geral denominado o produto do gene.

Os meios de melhoramento da atividade ou concentração da
5 proteína correspondente são em geral aumentados em pelo menos 10%, 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400% ou 500%, até um máximo de 1.000% ou 2.000%, com base naqueles da proteína do tipo selvagem ou na atividade ou concentração da proteína no microorganismo ou cepa de origem que não é recombinante para a enzima ou proteína correspondente.
10 O microorganismo ou cepa de origem que não é recombinante é entendido como significando o microorganismo no qual os meios da invenção são realizados.

A invenção também provê um processo para preparação de L-aminoácidos por fermentação de microorganismos recombinantes da família
15 da *Enterobacteriaceae*, caracterizado pelo fato de que:

a) os microorganismos da família da *Enterobacteriaceae* que produzem o aminoácido desejado e nos quais o gene *eno* ou alelos dos mesmos é/são melhorados, especificamente superexpressos, são cultivados em um meio sob condições apropriadas para a formação do produto de gene
20 *eno*,

b) o aminoácido desejado é concentrado no meio ou nas células dos microorganismos, e

c) o L-aminoácido desejado é isolado, os constituintes do caldo de fermentação e/ou a biomassa nessa totalidade ou porções (≥ 0 a 100%)
25 dos mesmos opcionalmente permanecendo no produto que foi isolado ou sendo completamente removidos.

Os microorganismos, especificamente microorganismos recombinantes, com um gene *eno* melhorado que a presente invenção também provê podem produzir L-aminoácidos de glicose, sacarose, lactose, frutose,
30 maltose, melado, opcionalmente amido, opcionalmente celulose ou de glicérol etanol. Eles são representativos da família *Enterobacteriaceae* escolhida do gênero *Escherichia*, *Erwinia*, *Providencia* e *Serratia*. Os gêneros *Escheri-*

chia e *Serratia* são os preferidos. Do gênero *Escherichia* especificamente, as espécies *Escherichia coli* e do gênero *Serratia* especificamente as espécies *Serratia marcescens* devem ser mencionadas.

Microorganismos recombinantes são em geral produzidos por
 5 transformação, transdução ou conjugação ou uma combinação desses métodos com um vetor que contém o gene desejado, um alelo desse gene ou partes do mesmo e/ou um promotor que melhora a expressão do gene. Esse promotor pode ser o promotor originário de melhora da mutação da sequência reguladora endógena a montante do gene ou um promotor eficiente
 10 que foi fundido com o gene.

Cepas apropriadas, que produzem L-treonina especificamente do gênero *Escherichia*, especificamente das espécies *Escherichia coli*, são por exemplo:

- *Escherichia coli* H4581 (EP 0 301 572)
- 15 - *Escherichia coli* KY10935 (Bioscience Biotechnology and Biochemistry 61(11):1877-1882 (1997))
- *Escherichia coli* VNIIgenetika MG442 (US-A-4278.765)
- *Escherichia coli* VNIIgenetika M1 (US-A-4.321.325)
- *Escherichia coli* VNIIgenetika 472T23 (US-A-5.631.157)
- 20 - *Escherichia coli* BKIIM B-3996 (US-A-5,175,107)
- *Escherichia coli* kat 13 (WO 98/04715)
- *Escherichia coli* KCCM-10132 (WO 00/09660).

Cepas que produzem L-treonina apropriadas do gênero *Serratia*, especificamente das espécies *Serratia marcescens* são, por exemplo:

- 25 - *Serratia marcescens* HNr21 (Applied and Environmental Microbiology 38(6): 1045-1051 (1979))
- *Serratia marcescens* TLR156 (Gene 57(2-3): 151-158 (1987))
- *Serratia marcescens* T-2000 (Applied Biochemistry and Biotechnology 37(3): 255-265 (1992)).

30 Cepas da família *Enterobacteriaceae* que produzem L-treonina preferivelmente possuem, entre outros, um ou mais aspectos fenotípicos ou genotípicos escolhidos do grupo consistindo em: resistência ao ácido

α -amino- β -hidroxivalérico, resistência a tialisina, resistência a etionina, resistência a α -metilserina, resistência ao ácido diaminosuccínico, resistência ao ácido α -aminobutírico, resistência à borrelidina, resistência ao ácido ciclo-
 pentano-carboxílico, resistência a rifampicina, resistência aos análogos de
 5 valina, tais como, por exemplo, hidroxamato de valina, resistência aos análogos da purina, tais como, por exemplo, 6-dimetilaminopurina, uma necessidade de L-metionina, opcionalmente uma necessidade parcial e compensável de L-isoleucina, uma necessidade de ácido mesodiaminopimélico, auxotrofia com relação aos dipeptídeos contendo treonina, resistência à
 10 L-treonina, resistência ao rafinato de treonina, resistência a L-homosserina, resistência à L-lisina, resistência a L-metionina, resistência ao ácido L-glutâmico, resistência ao L-aspartato, resistência a L-leucina, resistência a L-fenilalanina, resistência L-serina, resistência a L-cisteína, resistência a L-valina, sensibilidade ao fluorpiruvato, treonina desidrogenase defectiva,
 15 opcionalmente uma capacidade para utilização de sacarose, melhora da treonina óperon, melhora da homoserina desidrogenase l-aspartato quinase I, preferivelmente da forma resistente a retroalimentação, melhora da homoserina quinase, melhora da treonina sintase, melhora da asparato quinase, opcionalmente da forma resistente de retroalimentação, melhora da
 20 aspartato semialdeído desidrogenase, melhora do fosfoenol piruvato carboxilase, opcionalmente da forma resistente da retroalimentação, melhora do fosfoenol piruvato de sintase, melhora da transidrogenase, melhora do produto de gene RhtB, melhora do produto de gene RhtC, melhora do produto de gene YfiK, melhora de uma piruvato carboxilase e atenuação da formação de ácido acético.
 25

Foi verificado que os microorganismos da família da *Enterobacteriaceae* produzem L-aminoácidos, especificamente L-treonina, de um modo aperfeiçoado após melhora, especificamente superexpressão do gene eno ou alelos dos mesmos.

30 As seqüências de nucleotídeo dos genes do *Escherichia coli* pertencem a técnica anterior (vide as referências ao texto que se seguem) e podem também ser encontradas na seqüência do genoma do *Escherichia coli*

publicada por Blattner e outros (Science 277: 1453-1462 (1997)). É sabido que as enzimas endógenas no hospedeiro (metionina aminopeptidase) podem separar a metionina de aminoácido do terminal N.

As seqüências de nucleotídeo para o gene eno são da mesma forma conhecidas da *Shigella flexneri* e *Salmonella typhimurium*, que também pertencem a família Enterobacteriaceae.

O gene eno é descrito, entre outros, pelos seguintes dados:

Descrição: enolase, fosfopiruvato hidratase

Função: enolização (divisão da água): 2-fosfoglicerato para fosfoenol piruvato na glicólise

Número EC: EC 4.2.1.11

Referência: Spring TG. e Wold F.; The Journal of Biological Chemistry 246(22): 6797-6802 (1971), Klein e outros; DNA Sequence 6(6): 351-355 (1996), Gulick e outros; Biochemistry 40(51): 15716 -15724 (2001), Kaga e outros; Bioscience, Biotechnology and Biochemistry 66(10): 2216-2220 (2002)

Número de acesso: AE000361

O gene eno da *Salmonella typhimurium* é descrito, entre outros, na referência que se segue: Garrido-Pertierra A.; Revista Espanola de Fisiologia 36(1): 33-39 (1980).

As seqüências de ácido nucléico podem ser encontradas nos bancos de dados do National Center for Biotechnology Information (NCBI) of the National Library of Medicine (Bethesda, MD, USA), o banco de dados da seqüência de nucleotídeo do European Molecular Biologies Laboratories (EMBL, Heidelberg, Alemanha ou Cambridge, Reino Unido) ou o DNA data-bank of Japan (DDBJ, Mishima, Japão).

Para melhor clareza, a seqüência de nucleotídeo conhecida do gene eno do *Escherichia coli* é encontrada na SEQ ID Nº. 3 e as seqüências conhecidas para o gene eno de *Shigella flexneri* (AE015293) e *Salmonella typhimurium* (AE008835) são encontradas sob SEQ ID Nº. 5 ou SEQ ID Nº. 7. As seqüências de aminoácido das proteínas codificadas por essas estruturas de leitura são mostradas como SEQ ID Nº. 4, SEQ ID Nº. 6 ou a SEQ ID Nº. 8.

Os genes descritos nas referência de texto mencionadas podem ser usados de acordo com a invenção. Os alelos dos genes que resultam da degeneração do código genético ou em razão das "mutações de sentido" de função neutra podem adicionalmente ser usados. O uso dos genes endógenos é preferido.

"Genes endógenos" ou "seqüências de nucleotídeo endógenas" são entendidas com significando os genes ou alelos ou seqüências de nucleotídeo presentes na população de uma espécie.

Alelos apropriados do gene eno que contêm mutações de sentido de função neutra incluem, entre outros, aqueles que conduzem a no máximo 50 ou no máximo 40 ou a no máximo 30 ou a no máximo 20, preferivelmente a no máximo 10 ou a no máximo 5, muito específica e preferivelmente a no máximo 3 ou a no máximo 2 ou a no mínimo uma (1) troca de aminoácido de conservação na proteína codificada pelos menos.

No caso de aminoácidos aromáticos, as trocas de conservação são referidas quando fenilalanina, triptofano e trisona são trocadas uma pela outra. No caso de aminoácidos hidrófobos, trocas de conservação são referidas quando leucina, isoleucina e valina são trocadas uma pela outra. No caso de aminoácidos polares, as trocas de conservação são referidas quando a glutamina e asparagina são trocadas uma pela outra. No caso dos aminoácidos básicos, as trocas de conservação são referidas quando arginina, lisina e histidina são trocadas uma pela outra. No caso dos aminoácidos ácidos, as trocas de conservação são referidas quando ácido aspártico e ácido glutâmico são trocados um pelo outro. No caso de aminoácidos contendo grupos hidroxila, as trocas de conservação são referidas quando serina e treonina são trocadas uma pela outra. Todas as outras trocas de aminoácido são denominadas trocas de aminoácido de não conservação.

Da mesma forma, aquelas seqüências de nucleotídeo que codificam variantes das proteínas mencionadas que contêm adicionalmente um alongamento ou encurtamento de pelo menos um (1) aminoácido no término N ou C podem também ser usados. Esse alongamento ou encurtamento não é superior a 50, 40, 30, 20, 10, 5, 3 ou 2 aminoácidos ou radicais de aminoácido.

Alelos apropriados também incluem aqueles que codificam proteínas onde pelo menos um (1) aminoácido é inserido (inserção) ou removido (deleção). O número máximo de tais alterações, denominadas indels, pode se referir a 2, 3, 5, 10, 20, porém no caso mais de 30 aminoácidos.

5 Alelos apropriados incluem, adicionalmente, aqueles que são obtidos por hibridização, especificamente sob condições estringentes, usando a SEQ ID Nº. 3, SEQ ID Nº. 5 ou SEQ ID Nº. 7 ou partes das mesmas, especificamente as regiões de codificação ou as seqüências complementares as mesmas.

10 As instruções para identificação das seqüências de DNA por meio de hibridização podem ser encontradas pelo versado na arte, entre outros, no livro de bolso "The DIG System Users Guide for Filter Hybridization" da Boehringer Mannheim GmbH (Mannheim, Alemanha, 1993) e em Liebl e outros (International Journal of Systematic Bacteriology 41: 255-260
15 (1991)). A hibridização acontece sob condições estringentes, isto é, apenas os híbridos nos quais a sonda e a seqüência alvo, isto é, os polinucleotídeos tratados com a sonda, são pelo menos 80% idênticos sendo formados. É sabido que a estrigência da hibridização, incluindo as etapas de lavagem, é influenciada ou determinada por variação da composição tampão, a temperatura e a concentração do sal. A reação de hibridização é, em geral, realizada sob uma estrigência relativamente baixa, em comparação as etapas de lavagem (Hybaid Hybridisation Guide, Hybaid Limited, Teddington, Reino Unido, 1996).

25 Um tampão correspondendo a um tampão 5 x SSC a uma temperatura de cerca de 50°C - 68°C, por exemplo, pode ser empregado para reação de hibridização. As sondas também podem hibridizar aqui com polinucleotídeos que são menos de 80% idênticos à seqüência da sonda. Tais híbridos são menos estáveis e são removidos por lavagem sob condições estringentes. Isso pode ser obtido, por exemplo, por diminuição da concentração de sal para 2x SSC e opcional e subsequente 0,5x SSC (The DIG System User's Guide for Filter Hybridisation, Boehringer Mannheim, Mannheim, Alemanha, 1995) uma temperatura de cerca de 50°C - 68°C,
30

cerca de 52°C - 68°C, cerca de 54°C - 68°C, cerca de 56°C - 68°C, cerca de 58°C - 68°C, cerca de 60°C - 68°C, cerca de 62 - 68°C, cerca de 64°C - 68°C, cerca de 66°C - 68°C sendo estabelecida. Faixas de temperatura de cerca de 64°C - 68°C ou cerca de 66°C - 68°C são preferidas. É opcionalmente possível abaixar a concentração do sal a uma concentração correspondente de 0,2x SSC ou 0,1x SSC. Os fragmentos de polinucleotídeo que são, por exemplo, pelo menos 80% ou pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98% ou pelo menos 99% idênticos à sequência da sonda empregada ou às sequências de nucleotídeo mostradas na SEQ ID N°. 3, SEQ ID N°. 5 ou SEQ ID N°. 7 podem ser isolados por aumento escalonado da temperatura de hibridização de 50°C para 68°C em etapas de cerca de 1 - 2°C. Instruções adicionais sobre hibridização estão disponíveis comercialmente na forma dos assim chamados kits (por exemplo, DIG Easy Hyb da Roche Diagnostics GmbH, Mannheim, Alemanha, Número do catálogo 1603558). As sequências de nucleotídeo obtidas desse modo codificam polipeptídeos que são pelo menos 90% idênticos, especificamente idênticos a extensão de pelo menos 95%, preferivelmente a extensão de pelo menos 98% ou pelo menos 99%, muito específica e preferivelmente a extensão de 99,7% em relação às sequências de aminoácido mostradas na SEQ ID N°. 4, SEQ ID N°. 6 ou SEQ ID N°. 8.

Para obter uma melhora, por exemplo, a expressão dos genes ou propriedades catalíticas das proteínas de enzima podem ser aumentadas. As duas medidas podem ser opcionalmente combinadas.

Assim, por exemplo, várias cópias dos genes correspondentes podem ser aumentadas ou o promotor e a região de regulação ou o sítio de ligação do ribossoma a montante do gene estrutural podem ser mutados. Os cassetes de expressão que são incorporados à montante do gene estrutural atuam da mesma forma. Pelos promotores induzíveis, é adicionalmente possível aumentar a expressão no curso da produção de L-treonina fermentativa, e pode adicionalmente ser vantajoso também usar promotores para a expressão do gene que permitem outra expressão do gene com relação ao

tempo. A expressão é da mesma forma aperfeiçoada por medidas de prolongamento da vida do RNA-m. Adicionalmente, a atividade da enzima e também aumentada por prevenção da degradação da proteína enzima. Os genes ou construções de gene podem tanto estar presentes nos plasmídeos com um número variado de cópias quanto podem ser integrados e ampliados no cromossoma. Alternativamente, uma superexpressão dos genes em questão pode adicionalmente ser obtida por alteração da composição dos meios e do procedimento de cultura.

Os métodos de superexpressão são adequadamente descritos na técnica anterior – por exemplo, por Makrides e outros (*Microbiological Reviews* 60 (3), 512-538 (1996)). Utilizando-se vetores, o número de cópias é aumentado em pelo menos (1) cópia. Os vetores que podem ser usados são plasmídeos, tais como os que são descritos, por exemplo, na US 5.538.873. Os vetores que podem também ser usados são fagos, por exemplo, o fago Mu, conforme descrito na EP 0332448, ou o fago lambda (λ). Um aumento no número de cópias pode também ser obtido por incorporação de uma cópia adicional a um sítio adicional do cromossoma – por exemplo, dentro do sítio do fago λ (Yu and Court, *Gene* 223, 77-81 (1998)). A US 5.939.307 revela que por incorporação dos cassetes de expressão ou promotores, tais como, por exemplo, o promotor tac, o promotor trp, o promotor lpp ou o promotor P_L e o promotor P_R do fago λ por exemplo a montante do óperon de treonina cromossômica, foi possível obter um aumento na expressão. Os promotores do fago T7, os promotores da "gear-box" ou o promotor nar podem ser usados da mesma forma. Tais cassetes de expressão ou promotores podem também ser usados, conforme descrito na EP 0.593.792, para superexpressar os genes ligados aos plasmídeos. Por uso do alelo $lacI^Q$, a expressão dos genes ligados aos plasmídeos pode por sua vez ser controlada (Glascok and Weickert, *Gene* 223, 221-231 (1998)). É adicionalmente possível que a atividade dos promotores seja aumentada por modificação de sua sequência por meio de uma ou mais trocas de nucleotídeo, por inserção(ões) e/ou deleção(ões). Outra expressão de gene com relação ao tempo pode ser obtida, por exemplo, conforme descrito por

Walker e outros (Journal of Bacteriology 181: 1269-80 (1999)) empregando o promotor fis dependente de fase de crescimento.

O versado na técnica pode encontrar várias instruções a esse respeito, entre outros em Chang and Cohen (Journal of Bacteriology 134: 1141-1156 (1978)), em Hartley and Gregori (Gene 13: 347-353 (1981)), em Amann and Brosius (Gene 40: 183-190 (1985)), em de Broer e outros (Proceedings of the National Academy of Sciences of the United States of America 80: 21-25 (1983)), em LaVallie e outros (BIO/TECHNOLOGY 11: 187-193 (1993)), in PCT/US97/13359, em Llosa e outros (Plasmid 26: 222-224 (1991)), em Quandt and Klipp (Gene 80: 161-169 (1989)), in Hamilton e outros (Journal of Bacteriology 171: 4617-4622 (1989)), em Jensen and Hammer (Biotechnology and Bioengineering 58: 191-195 (1998)) e nos livros de texto conhecidos de genética e biologia molecular.

Os vetores de plasmídeo que podem replicar na *Enterobacteriaceae*, tais como, por exemplo, vetores de clonagem derivados de pACYC184 (Bartolomé e outros; Gene 102: 75-78 (1991)), pTrc99A (Amann e outros; Gene 69: 301-315 (1988)) ou derivados de pSC101 (Vocke and Bastia; Proceedings of the National Academy of Sciences USA 80(21): 6557-6561 (1983)) podem ser usados. Uma cepa transformada com um vetor de plasmídeo, onde o vetor de plasmídeo porta pelo menos um gene eno ou uma sequência de nucleotídeo que codifica o produto do gene do mesmo ou um alelo, pode ser empregada em um processo de acordo com a invenção.

O termo transformação é entendido como significando a absorção do ácido nucléico isolado por um hospedeiro (microorganismo).

É também possível transferir mutações que afetam a expressão de genes específicos em várias cepas por troca de sequência (Hamilton e outros; (Journal of Bacteriology 171: 4617-4622 (1989)), conjugação ou transdução.

Explicações mais detalhadas, em termos de genética e biologia molecular são encontradas nos livros texto conhecidos de genética e biologia molecular, tais como, por exemplo, o livro texto de Birge (Bacterial and

Bacteriophage Genetics, 4th ed., Springer Verlag, New York (USA), 2000) ou o livro texto de Berg, Tymoczko and Stryer (Biochemistry, 5th ed., Freeman and Company, New York (USA), 2002) ou o livro de bolso de Sambrook e outros (Molecular Cloning, A Laboratory Manual, (3 volume set), Cold Spring Harbor Laboratory Press, Cold Spring Harbor (USA), 2001).

Adicionalmente pode ser vantajoso para a produção de L-aminoácidos, especificamente L-treonina, com as cepas da família da Enterobacteriaceae, melhorar uma ou mais via de biossíntese de treonina conhecida ou enzimas do metabolismo anaplerótico ou enzimas para a produção de fosfato de dinucleotídeo de adenina nicotinamida ou enzimas de glicólise ou enzimas PTS ou enzimas de metabolismo de enxofre, além da melhora do gene eno. O uso de genes endógenos em geral é preferido.

Assim, por exemplo, podem ser melhorados ao mesmo tempo, um ou mais dos genes escolhidos do grupo consistindo em:

- pelo menos um gene do óperon thrABC óperon que codifica aspartato quinase, homoserina desidrogenase, homoserina quinase e treonina sintase (US-A-4.278.765),
- o gene pyc de *Corynebacterium glutamicum* que codifica piruvato carboxilase (WO 99/18228),
- o gene pps que codifica fosfoenol piruvato sintase (Molecular and General Genetics 231(2): 332-336 (1992)),
- o gene ppc que codifica fosfoenol piruvato carboxilase (WO 02/064808),
- os genes pntA e pntB genes que codificam subunidades de piridina transidrogenase (European Journal of Biochemistry 158: 647-653 (1986)),
- o gene rhtB que codifica a proteína que fornece resistência a homoserina (EP-A-0 994 190),
- o gene rhtC que codifica a proteína que fornece resistência à treonina (EP-A-1 013 765),
- o gene thrE da *Corynebacterium glutamicum* que codifica a proteína veículo de exportação de treonina (WO 01/92545),

- o gene *gdhA* que codifica glutamato desidrogenase (Nucleic Acids Research 11: 5257-5266 (1983); Gene 23: 199-209 (1983)),
- o gene *pgm* que codifica fosfoglucomutase (WO 03/004598),
- o gene *fba* que codifica bifosfato aldolase de glicose (WO 03/004664),
- o gene *ptsH* para o óperon *ptsHlcr* que codifica a proteína fosfohistidina de hexose fosfotransferase do sistema de fosfotransferase PTS (WO 03/004674),
- o gene *ptsI* do óperon *ptsHlcr* que codifica a enzima I do sistema de fosfotransferase PTS (WO 03/004674),
- o gene *crr* do óperon *ptsHlcr* que codifica o componente IIA específico da glicose do sistema de fosfotransferase PTS (WO 03/004674),
- o gene *ptsG* que codifica o componente IIBC específico para glicose (WO 03/004670),
- o gene *lrp* que codifica o regulador do régulon de leucina (WO 03/004665),
- o gene *fadR* que codifica o regulador do régulon (WO 03/038106),
- o gene *iclR* que codifica o regulador do metabolismo intermediário central (WO 03/038106),
- o gene *ahpC* do óperon *ahpCF* que codifica a subunidade pequena da alquil hidroxiperoxido reductase (WO 03/004663),
- o gene *ahpF* do óperon *ahpCF* que codifica a subunidade grande da alquil hidroxiperoxido reductase (WO 03/004663),
- o gene *cysK* que codifica cisteína sintase A (WO 03/006666),
- o gene *cysB* que codifica o regulador do régulon (WO 03/006666),
- o gene *cysJ* do óperon *cysJIH* que codifica a flavoproteína da NADPH sulfito reductase (WO 03/006666),
- o gene *cysI* do óperon *cysJIH* que codifica a hemoproteína da NADPH sulfito reductase (WO 03/006666),

- o gene *cysH* do óperon *cysJIH* que codifica adenilil sulfato reductase (WO 03/006666),
- o gene *rseA* do óperon *rseABC* que codifica uma proteína de membrana com atividade anti-sigmaE (WO 03/008612),
- 5 • o gene *rseC* do óperon *rseABC* que codifica um regulador global do fator sigmaE (WO 03/008612),
- o gene *sucA* do óperon *sucABCD* que codifica a subunidade de descarboxilase da 2-cetoglutarato desidrogenase (WO 03/008614),
- o gene *sucB* do óperon *sucABCD* que codifica a subunidade
- 10 diidrolipolitranssuccinase E2 da 2-cetoglutarato desidrogenase (WO 03/008614),
- o gene *sucC* do óperon *sucABCD* que codifica a subunidade β da succinil-CoA sintetase (WO 03/008615),
- o gene *sucD* do óperon *sucABCD* que codifica a subunidade
- 15 α da succinil-CoA sintetase (WO 03/008615),
- o gene *aceE* que codifica o componente E1 do complexo de piruvato desidrogenase (WO 03/076635),
- o gene *aceF* que codifica o componente E2 do complexo de piruvato desidrogenase (WO 03/076635), e
- 20 • o gene *rseB* que codifica o regulador da atividade do fator sigmaE (Molecular Microbiology 24(2): 355-371 (1997)),
- o produto de gene do quadro aberto de leitura (ORF) *yodA* da *Escherichia coli* (Número de acesso AE000288 do National Center for Biotechnology Information (NCBI, Bethesda, MD, USA), DE10361192.4),
- 25 • o produto de gene do quadro aberto de leitura (ORF) *yaaU* da *Escherichia coli* (Número de acesso AE005181 do National Center for Biotechnology Information (NCBI, Bethesda, MD, USA), DE10361268.8) e
- o gene *malT* que codifica o ativador transcricional positivo do régulon de maltose (Gene 42: 201-208 (1986))
- 30 Adicionalmente pode ser vantajoso para a produção de L-aminoácidos, especificamente L-treonina, além de melhorar o gene *eno* para um ou mais dos genes escolhidos do grupo consistindo em:

- o gene *tdh* que codifica a treonina desidrogenase (Journal of Bacteriology 169: 4716-4721 (1987)),
 - o gene *mdh* que codifica malato desidrogenase (E.C. 1.1.1.37) (Archives in Microbiology 149: 36-42 (1987)),
 - 5 • o produto de gene do quadro aberto de leitura (ORF) (*orf*) *yj-fA* da *Escherichia coli* (Número de acesso AAC77180 do National Center for Biotechnology Information (NCBI, Bethesda, MD, USA), WO 02/29080)),
 - o produto de gene do quadro aberto de leitura (ORF) *ytfP* da *Escherichia coli* (Número de acesso AAC77179 do National Center for Bio-
 - 10 technology Information (NCBI, Bethesda, MD, USA), WO 02/29080)),
 - o gene *pckA* que codifica a enzima fosfoenol piruvato carboxiquinase (WO 02/29080),
 - o gene *poxB* que codifica piruvato oxidase (WO 02/36797),
 - o gene *dgsA* que codifica o regulador DgsA do sistema fos-
 - 15 fotransferase (WO 02/081721) e é também conhecido sob a denominação de gene *mlc*,
 - o gene *fruR* que codifica o repressor de frutose (WO 02/081698) e é também conhecido sob a denominação de gene *cra*,
 - o gene *rpoS* que codifica o fator sigma³⁸ (WO 01/05939) e é
 - 20 também conhecido sob a denominação de gene *katF*, e
 - o gene *aspA* que codifica aspartato amônio liase (WO 03/008603)
- a serem atenuados, especificamente eliminados ou para a expressão dos mesmos a ser reduzida.
- 25 O termo "atenuação" nesse contexto descreve a redução ou eliminação da atividade intracelular ou concentração de uma ou mais enzimas ou proteínas em um microorganismo que são codificadas pelo DNA correspondente, por exemplo, por uso de um promotor mais fraco que no microorganismo ou cepa de origem, que não é recombinante para a enzima correspondente ou proteína ou um gene ou alelo que codifica uma enzima correspondente ou proteína com uma atividade inferior ou inativação da enzima
- 30 correspondente (proteína) ou o quadro aberto de leitura ou o gene e opcio-

nalmente combinando essas medidas.

Pelas medidas de atenuação, a atividade ou concentração da proteína correspondente é em geral reduzida para 0 a 75%, 0 a 50%, 0 a 25%, 0 a 10% ou 0 a 5% da atividade ou concentração da proteína do tipo selvagem ou da atividade ou concentração da proteína no microorganismo ou cepa de origem que não é recombinante para a enzima ou proteína correspondente. O microorganismo ou cepa de origem que não é recombinante é entendido como significando o microorganismo no qual as medidas da invenção são realizadas.

Para obter uma atenuação, por exemplo, a expressão dos genes ou estruturas abertas de leitura ou as propriedades catalíticas das proteínas enzimas podem ser reduzidas ou eliminadas. As duas enzimas podem opcionalmente ser combinadas.

A redução na expressão genética pode acontecer por cultivo apropriado, por modificação genética (mutação) das estruturas de sinal da expressão genética ou também pela técnica de antisenso-RNA. As estruturas de sinal da expressão genética são, por exemplo, genes repressores, genes ativadores, operadores, promotores, atenuadores, sítios de ligação de ribossoma, o códon de partida e terminadores. O versado na arte pode encontrar informações a esse respeito, entre outros, por exemplo, em Jensen and Hammer (Biotechnology and Bioengineering 58: 191-195 (1998)), em Carrier und Keasling (Biotechnology Progress 15: 58-64 (1999)), Franch and Gerdes (Current Opinion in Microbiology 3: 159-164 (2000)) e nos livros texto conhecidos de genética e biologia molecular, tais como, por exemplo, o livro texto de Knippers ("Molekulare Genetik [Molecular Genetics]", 6ª edição, Georg Thieme Verlag, Stuttgart, Alemanha, 1995) ou em Winnacker ("Gene und Klone [Genes and Clones]", VCH Verlagsgesellschaft, Weinheim, Alemanha, 1990).

As mutações que conduzem a uma alteração ou redução nas propriedades catalíticas das proteínas da enzima são conhecidas da técnica anterior. Exemplos que podem ser mencionados são os trabalhos de Qiu and Goodman (Journal of Biological Chemistry 272: 8611-8617 (1997)), Ya-

no e outros (Proceedings of the National Academy of Sciences of the United States of America 95: 5511-5515 (1998)), Wenthe and Schachmann (Journal of Biological Chemistry 266: 20833-20839 (1991)). Descrições resumidas podem ser encontradas nos livros texto conhecidos de genética e biologia molecular, tais como, por exemplo, aquele de Hagemann ("Allgemeine Genetik [General Genetics]", Gustav Fischer Verlag, Stuttgart, 1986).

Possíveis mutações são transições, transversões, inserções e deleções de pelo menos um (1) par de base ou nucleotídeo. Dependendo do efeito da troca do aminoácido causado pela mutação na atividade da enzima, "mutações sem sentido" são também referidas. A mutação sem sentido conduz a uma troca de um dado aminoácido em uma proteína para outra, isso sendo, especificamente, uma troca de aminoácido não conservadora. A capacidade funcional ou atividade da proteína é prejudicada por isso e reduzida a um valor de 0 a 75%, 0 a 50%, 0 a 25%, 0 a 10% ou 0 a 5%. A mutação sem sentido conduz a uma parada do códon na região de codificação do gene e portanto a uma interrupção prematura na tradução. As inserções e deleções de pelo menos um par de base em um gene conduzem as mutações de deslocamento de estrutura, o que conduz aos aminoácidos incorretos sendo incorporados ou a tradução sendo interrompida prematuramente. Se um códon de parada for formado na região de codificação como uma consequência da mutação, isso também conduz a um término prematuro da tradução. As deleções de pelo menos um (1) ou mais códons tipicamente também conduzem a uma perda completa da atividade da enzima.

As instruções de geração de tais mutações se encontram na técnica anterior e podem ser encontradas em livros texto conhecidos de genética e biologia molecular, tais como, por exemplo, o livro texto de Knippers ("Molekulare Genetik [Molecular Genetics]", 6ª edição, Georg Thieme Verlag, Stuttgart, Alemanha, 1995), aquele de Winnacker ("Gene und Klone [Genes and Clones]", VCH Verlagsgesellschaft, Weinheim, Alemanha, 1990) ou aquele de Hagemann ("Allgemeine Genetik [General Genetics]", Gustav Fischer Verlag, Stuttgart, 1986).

As mutações apropriadas nos genes podem ser incorporadas

em cepas apropriadas por substituição do gene ou alelo.

Um método comum é o método descrito por Hamilton e outros (Journal of Bacteriology 171: 4617–4622 (1989)), de substituição do gene com a ajuda de um derivado pSC101 condicionalmente replicante
 5 pMAK705. Outros métodos descritos na técnica anterior, tais como, por exemplo, aquele de Martinez-Morales e outros (Journal of Bacteriology 181: 7143-7148 (1999)) ou aquele de Boyd e outros (Journal of Bacteriology 182: 842-847 (2000)), podem da mesma forma ser usados.

Também é possível transferir as mutações nos genes específicos
 10 cos ou mutações que afetam a expressão dos genes específicos ou quatros abertos de leitura em várias cepas por conjugação ou transdução.

Além de melhorar o gene eno isso pode também ser vantajoso para a produção de L-aminoácidos, especificamente L-treonina, para eliminar reações colaterais indesejáveis (Nakayama: "Breeding of Amino Acid
 15 Producing Microorganisms", em: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, Londres, Reino Unido, 1982).

Os microorganismos produzidos de acordo com a invenção podem ser cultivados em processos de batelada (cultura de batelada) na batelada de alimentação (processo de alimentação) no processo de alimentação
 20 de batelada repetida (processo de alimentação repetida) ou em um processo contínuo (DE 102004028859.3 ou US 5,763,230). Um resumo dos métodos de cultura conhecidos é descrito no livro texto de Chmiel (Bioprozesstechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) ou no livro texto de Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)).
 25

O meio de cultura a ser usado deve satisfazer os requisitos das cepas específicas de modo apropriado. As descrições do meio de cultura para vários microorganismos estão contidas no livro de bolso "Manual of
 30 Methods for General Bacteriology" da American Society for Bacteriology (Washington D. C., USA, 1981).

Açúcares e carboidratos, tais como, glicose, sacarose, lactose,

frutose, maltose, melado, amido e opcionalmente celulose, óleos e gorduras, tais como, por exemplo, óleo de soja, óleo de girassol, óleo de amendoim e gordura de coco, ácidos graxos, tais como, por exemplo, ácido palmítico, ácido esteárico e ácido linoléico, álcoois, tais como, por exemplo, glicerol
5 etanol e ácidos orgânicos, tais como, por exemplo, ácido acético podem ser usados como a fonte de carbono. Essas substâncias podem ser usadas individualmente ou como uma mistura.

Compostos contendo nitrogênio orgânico, tais como, peptonas, extrato de levedura, extrato de carne, extrato de malte, licor de milho, fari-
10 nha de soja e uréia ou compostos inorgânicos, tais como, sulfato de amônio, cloreto de amônio, fosfato de amônio, carbonato de amônio e nitrato de amônio podem ser usados como uma fonte de nitrogênio. As fontes de nitrogênio podem ser usadas individualmente ou como um mistura.

Ácido fosfórico, diidrogeno fosfato de potássio ou hidrogeno fos-
15 fato de dipotássio ou os sais contendo sódio correspondentes podem ser usados como uma fonte de fósforo. O meio de cultura deve compreender, adicionalmente, sais de metais, tais como, por exemplo, sulfato de magnésio ou sulfato de ferro, que são necessários para o crescimento. Finalmente, as substâncias de crescimento essenciais, tais como, aminoácidos e vitaminas
20 podem ser empregados além das substâncias mencionadas acima. Os precursores apropriados podem, além disso, ser adicionados ao meio de cultura. As substâncias de partida mencionadas podem ser adicionadas à cultura na forma de uma batelada simples ou podem ser alimentadas durante a cultura de modo apropriado.

25 A fermentação em geral é realizada em um pH de 5,5 a 9,0, especificamente de 6,0 a 8,0. Os compostos básicos, tais como, hidróxido de sódio, hidróxido de potássio, amônia ou amônia aquosa ou compostos ácidos, tais como, ácido fosfórico ou ácido sulfúrico, podem ser empregados de modo apropriado para controlar o pH da cultura. Antiespumantes, tais como,
30 por exemplo, poliglicol ésteres de ácido graxo, podem ser empregados para controlar o desenvolvimento da espuma. Substâncias apropriadas possuindo uma ação seletiva, por exemplo, antibióticos, podem ser adicionados ao

meio para manter a estabilidade dos plasmídeos. Para manter as condições aeróbicas, oxigênio ou misturas de gás contendo oxigênio, tais como, por exemplo ar, são introduzidas à cultura. A temperatura da cultura é geralmente de 25°C a 45°C e preferivelmente 30°C a 40°C. A cultura continua até um

5 máximo de L-aminoácidos ou L-treonina se formarem. Esse alvo é geralmente alcançado dentro de 10 horas a 160 horas.

A análise dos L-aminoácidos pode ser realizada por cromatografia de troca de ânion com derivatização de ninidrina subsequente, conforme descrito por Spackman e outros (Analytical Chemistry, 30: 1190-1206

10 (1958)) ou pode ser realizada por HPLC de fase reversa, conforme descrito por Lindroth e outros (Analytical Chemistry 51: 1167-1174 (1979)).

O processo de acordo com a invenção é usado para a preparação fermentativa de L-aminoácidos, tais como, por exemplo, L-treonina, L-isoleucina, L-valina, L-metionina, L-homosserina, L-triptofano e L-lisina,

15 especificamente L-treonina.

Os microorganismos que se seguem foram depositados no Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSMZ = German Collection of Microorganisms and Cell Cultures, Braunschweig, Alemanha) de acordo com o Tratado de Budapeste:

- 20
- cepa de *Escherichia coli*, *E. coli* MG442 como DSM 16574.

A presente invenção é explicada em maiores detalhes no que se segue com a ajuda dos exemplos da concretização.

Os meios mínimo (M9) e completo (LB) para *Escherichia coli* usados são descritos por J.H. Miller (A Short Course in Bacterial Genetics (1992), Cold Spring Harbor Laboratory Press). O isolamento de DNA de

25 plasmídeo do *Escherichia coli* e todas as técnicas de restrição, ligação, Klenow e tratamento com fosfatase alcalina são realizados pelo método de Sambrook e outros (Molecular Cloning - A Laboratory Manual (1989) Cold Spring Harbor Laboratory Press). A menos que descrito de outra forma, a

30 transformação do *Escherichia coli* é realizada pelo método de Chung e outros (Proceedings of the National Academy of Sciences of the United States of America 86: 2172-2175 (1989)).

A temperatura de incubação para a preparação das cepas e transformantes é de 37°C.

Exemplo 1

Construção do plasmídeo de expressão pTrc99A_{eno}

- 5 O gene *eno* do *E.coli* K12 é ampliado usando a reação da cadeia da polimerase (PCR) e oligonucleotídeos sintéticos. Partindo da sequência de nucleotídeo do gene *eno* no *E.coli* K12 MG1655 (Número de acesso AE000361, Blattner e outros (Science 277: 1453-1474 (1997))), os iniciadores da PCR são sintetizados (MWG Biotech, Ebersberg, Alemanha).
- 10 Os iniciadores contêm seqüências para enzimas de restrição que são marcadas por sublinhamento na seqüência de nucleotídeo mostrada a seguir. O iniciador *eno1* contém o sítio de clivagem de restrição para XbaI, o iniciador *eno2* contém o mesmo para o HindIII.

eno1:

- 15 5' – GTTTGCTAGAGTTTCAGTTTAACTAGTGAC – 3' (SEQ ID Nº. 1)

eno2:

5' – CCGGAGGCTGGCAAGCTTAAATCAG – 3' (SEQ ID Nº. 2)

- O DNA cromossômico do *E.coli* K12 MG1655 DNA empregado para a PCR é isolado de acordo com as instruções dos fabricantes com "Qiagen Genomic-tips 100/G" (QIAGEN, Hilden, Alemanha). Um fragmento de
- 20 DNA com tamanho de cerca de 1.381 pares de base pode ser ampliado com os iniciadores específicos sob condições de PCR padrão (Innis e outros (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press) com polimerase de Vent-DNA (New England BioLabs, Frankfurt, Alemanha) (SEQ ID Nº. 3).
- 25

- O fragmento *eno* ampliado é restrito com as enzimas de restrição HindIII e XbaI e após purificação (Purification Kit, QIAGEN, Hilden, Alemanha) verificado em um gel de agarose a 0,8%. O vetor pTrc99A (Pharmacia Biotech, Uppsala, Suécia) é clivado com enzimas HindIII e XbaI e a
- 30 ligação é realizada com o fragmento *eno* restrito. A cepa *E.coli* TOP10 One Shot (TOPO TA Cloning Kit, Invitrogen, Groningen, Países-Baixos) é transformada com a batelada de ligação e células transportando plasmídeo são

selecionadas no ágar LB, as quais são adicionados 50 µg/mL de ampicilina. A clonagem de sucesso pode ser demonstrada após isolamento de DNA de plasmídeo por clivagem de controle com as enzimas HindIII/XbaI e PvuI. O plasmídeo é denominado pTrc99Aeno (figura 1).

5 Exemplo 2

Preparação de L-treonina com a cepa MG442/pTrc99Aeno

A cepa *E.coli* de produção de L-treonina MG442 é descrita no pedido de patente US-A-4.278.765 e depositada como CMIM B-1628 no Russian National Collection for Industrial Microorganisms (VKPM, Moscow, Rússia) e como DSM 16574 no Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSMZ = German Collection of Microorganisms and Cell Cultures, Braunschweig, Alemanha) de acordo com o tratado de Budapeste.

A cepa MG442 é transformada com o plasmídeo de expressão pTrc99Aeno descrito no Exemplo 1 e com o vetor pTrc99Aeno células transportando plasmídeo são selecionadas no ágar LB com 50 µg/mL de ampicilina. Transformações de sucesso podem ser demonstradas após isolamento de DNA de plasmídeo por clivagens de controle com as enzimas

HindIII/XbaI. As cepas MG442/pTrc99Aeno e MG442/pTrc99A são formadas dessa maneira. Colônias individuais selecionadas são então multiplicadas adicionalmente em um meio mínimo com a composição de partida: 3,5 g/L de $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$, 1,5 g/L de KH_2PO_4 , 1 g/L de NH_4Cl , 0,1 g/L de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2 g/L de glicose, 20 g/L de ágar, 50 mg/L de ampicilina. A formação da L-treonina é verificada nas culturas de batelada de 10 mL contidas em frascos cônicos de 100 mL. Para tanto, 10 mL de meio de pré-cultura da seguinte composição: 2 g/L de extrato de levedura, 10 g/L de $(\text{NH}_4)_2\text{SO}_4$, 1 g/L de KH_2PO_4 , 0,5 g/L de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 15 g/L de CaCO_3 , 20 g/L de glicose, 50 mg/L de ampicilina são isolados e inoculados e a batelada é incubada por 16 horas a 37°C e 180 rpm em um incubador da Kühner AG (Birsfelden, Suíça).

Porções de 250 µl dessa pré-cultura são trans-inoculadas em 10 mL de meio de produção (25 g/L de $(\text{NH}_4)_2\text{SO}_4$, 2 g/L de KH_2PO_4 , 1 g/L de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,03 g/L de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,018 g/L de $\text{MnSO}_4 \cdot 1\text{H}_2\text{O}$, 30 g/L

de CaCO_3 , 20 g/L de glicose, 50 mg/L de ampicilina) e a batelada é incubada por 48 horas a 37°C. Para indução completa da expressão do gene *eno*, 100 mg/L de isopropil β -D-tiogalactopiranosido (IPTG) são adicionados em bateladas paralelas. A formação de L-treonina pela cepa de partida MG442 é investigada a mesma maneira, porém sem adição de ampicilina ao meio. Após a incubação a densidade óptica (OD) da suspensão de cultura é determinada com um fotômetro LP2W da Dr. Lange (Düsseldorf, Alemanha) em uma medição de comprimento de onda de 660 nm.

A concentração de L-treonina formada é então determinada no sobrenadante de cultura filtrado e estéril com um analisador de aminoácido da Eppendorf-BioTronik (Hamburg, Alemanha) por cromatografia de troca de ion e reação de pós-coluna com detecção de ninidrina.

O resultado do experimento é mostrado na Tabela 1.

Tabela 1

Cepa	Aditivos	OD (660 nm)	L-Treonina g/L
MG442	-	5,6	1,4
MG442/pTrc99A	-	3,8	1,3
MG442/pTrc99A <i>eno</i>	-	2,6	1,7
MG442/pTrc99A <i>eno</i>	IPTG	5,1	2,6

Breve Descrição da Figura:

Figura 1: Mapeamento do plasmídeo pTrc99A*eno* contendo o gene *eno*

Os dados de comprimento devem ser entendidos como dados aproximados. As abreviaturas e designações usadas possuem o seguinte significado:

- Amp: gene resistente a ampicilina
- *lacI*: gene para proteína repressora do promotor *trc*
- P*trc*: região promotora de *trc*, induzível por IPTG

- eno: região de codificação para o gene eno
- 5S: região 5S RNAr
- rmbT: região terminadora de RNAr

As abreviaturas para as enzimas de restrição possuem o signifi-

5 cado que se segue:

- HindIII: endonuclease de restrição de *Haemophilus influenzae* R₆
- PvuI: endonuclease de restrição de *Proteus vulgaris*
- XbaI: endonuclease de restrição de *Xanthomonas cam-*
10 *pestris*

LISTAGEM DE SEQÜÊNCIA

<110> Degussa AG

<120> PROCESSO PARA PREPARAÇÃO DE L-AMINOÁCIDOS USANDO CEPAS DA
FAMÍLIA DAS *ENTEROBACTERIACEAE*

15 <130> 020479 BT

<160> 8

<170> PatentIn version 3.2

<210> 1

<211> 31

20 <212> DNA

<213> Seqüência artificial

<220>

<221> Iniciador

<222> (1)..(31)

25 <223> eno1

<400> 1

gtttgtctag agtttcagtt taactagtga c

31

<210> 2

<211> 25

30 <212> DNA

<213> Seqüência artificial

<220>

REIVINDICAÇÕES

1. Processo para preparação de L-treonina e L-lisina por fermentação de micro-organismos recombinantes da família *Enterobacteriaceae*, caracterizado pelo fato de que:

5 (a) os micro-organismos que produzem o L-aminoácido desejado contêm um polinucleotídeo melhorado ou superexpresso correspondendo ao gene eno escolhido do grupo consistindo em SEQ ID N° 3, SEQ ID N° 5 ou SEQ ID N° 7, ou uma sequência degenerada das mesmas codificando a
10 ções nas quais o L-aminoácido desejado é concentrado no meio ou nas células, e

(b) o L-aminoácido desejado é isolado, os constituintes do caldo de fermentação e/ou a biomassa em sua totalidade ou porções (≥ 0 a 100%) dos mesmos permanecendo no produto, os quais foram isolados ou comple-
15 tamente removidos,

sendo que o melhoramento descreve a atividade intracelular aumentada ou concentração da referida proteína em pelo menos 10%, com base na proteína selvagem ou da cepa parental.

2. Processo, de acordo com a reivindicação 1, caracterizado pelo
20 fato de que os micro-organismos são produzidos por transformação, transdução ou conjugação ou uma combinação desses métodos, com um vetor que contém o gene eno, um alelo desse gene ou partes do mesmo e/ou um promotor.

3. Processo, de acordo com a reivindicação 1 ou 2, caracterizado
25 pelo fato de que o número de cópias do gene eno ou dos alelos é aumentado em pelo menos 1.

4. Processo, de acordo com a reivindicação 3, caracterizado pelo
fato de que o aumento no número de cópias do gene eno em pelo menos 1 é efetuado por integração da ORF ou dos alelos no cromossoma dos micro-
30 organismos.

5. Processo, de acordo com a reivindicação 3, caracterizado pelo
fato de que o aumento do número de cópias do gene eno em pelo menos 1 é

obtido por um vetor que replica extracromossomicamente.

6. Processo, de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que para obter o melhoramento:

5 (a) o promotor e a região de regulação ou o sítio de ligação do ribossoma a montante do gene *eno* são mutados, ou

(b) os cassetes de expressão ou promotores são incorporados a montante do gene *eno*.

7. Processo, de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que o gene *eno* está sob controle de um promotor que melhora a expressão do gene.

8. Processo, de acordo com qualquer uma das reivindicações 1 a 7, caracterizado pelo fato de que a concentração ou atividade do produto do gene *eno* é aumentada pela superexpressão do gene *eno* ou sequências de nucleotídeos ou alelos que codificam para o produto gênico do mesmo.

15 9. Processo, de acordo com qualquer uma das reivindicações 1 a 8, caracterizado pelo fato de que os micro-organismos são selecionados do gênero *Escherichia*.

10. Processo, de acordo com qualquer uma das reivindicações 1 a 9, caracterizado pelo fato de que os micro-organismos são selecionados da espécie *Escherichia coli*.

11. Processo, de acordo com qualquer uma das reivindicações 1 a 10, caracterizado pelo fato de que, para a preparação da L-treonina, os micro-organismos da família *Enterobacteriaceae* são fermentados nos quais adicionalmente, ao mesmo tempo, um ou mais genes escolhidos do grupo consistem em:

i. pelo menos um gene do óperon *thrABC* óperon que codifica aspartato quinase, homoserina desidrogenase, homoserina quinase e treonina sintase,

30 ii. o gene *pyc* de *Corynebacterium glutamicum* que codifica piruvato carboxilase,

iii. o gene *pps* que codifica fosfoenol piruvato sintase,

iv. o gene *ppc* que codifica fosfoenol piruvato carboxilase,

- v. os genes pntA e pntB genes que codificam subunidades de piridina transidrogenase,
- vi. o gene rhtB que codifica a proteína que fornece resistência a homoserina,
- 5 vii. o gene rhtC que codifica a proteína que fornece resistência à treonina,
- viii. o gene thrE da *Corynebacterium* que codifica a proteína veículo de exportação de treonina,
- ix. o gene gdhA que codifica glutamato desidrogenase,
- 10 x. o gene pgm que codifica fosfoglucomutase,
- xi. o gene fba que codifica bifosfato aldolase de frutose,
- xii. o gene ptsH que codifica a proteína fosfohistidina de hexose fosfotransferase
- xiii. o gene ptsI que codifica a enzima I do sistema de fosfotransfe-
- 15 rase,
- xiv. o gene crr que codifica o componente IIA específico da glicose,
- xv. o gene ptsG que codifica o componente IIBC específico para glicose,
- xvi. o gene lrp que codifica o regulador do regulon de leucina,
- 20 xvii. o gene fadR que codifica o regulador do regulon,
- xviii. o gene iclR que codifica o regulador do metabolismo intermediário central,
- xix. o gene ahpC que codifica a subunidade pequena da hidroxiperoxido reductase de alquila,
- 25 xx. o gene ahpF que codifica a subunidade grande da hidroperóxido reductase de alquila,
- xxi. o gene cysK que codifica cisteína sintase A,
- xxii. o gene cysB que codifica o regulador do regulon cys,
- xxiii. o gene cysJ que codifica a flavoproteína da NADPH sulfito re-
- 30 ductase,
- xxiv. o gene cysI que codifica a hemoproteína da NADPH sulfito reductase,

- xxv. o gene *cysH* que codifica sulfato reductase de adenilila,
- xxvi. o gene *rseA* que codifica uma proteína de membrana com atividade anti-sigmaE,
- xxvii. o gene *rseC* que codifica um regulador global do fator sigmaE,
- 5 xxviii. o gene *sucA* que codifica a subunidade de descarboxilase da 2-cetoglutarato desidrogenase,
- xxix. o gene *sucB* que codifica a subunidade diidrolipolitranssuccinase E2 da 2-cetoglutarato desidrogenase,
- xxx. o gene *sucC* que codifica a subunidade β da succinil-CoA sintetase,
- 10 tase,
- xxxi. o gene *sucD* que codifica a subunidade α da succinil-CoA sintetase,
- xxxii. o gene *aceE* que codifica o componente E1 do complexo de piruvato desidrogenase,
- 15 xxxiii. o gene *aceF* que codifica o componente E2 do complexo de piruvato desidrogenase,
- xxxiv. o gene *rseB* que codifica o regulador da atividade do fator sigmaE,
- xxxv. o produto de gene do quadro aberto de leitura (ORF) *yodA* da
- 20 *Escherichia coli*, e
- xxxvi. o produto de gene do quadro aberto de leitura (ORF) *yaaU* da *Escherichia coli*
- é ou são melhorados, especificamente, superexpressos.
12. Processo, de acordo com qualquer uma das reivindicações 1
- 25 a 11, caracterizado pelo fato de que, para a preparação de L-treonina, microorganismos da família *Enterobacteriaceae* são fermentados, nos quais adicionalmente ao mesmo tempo um ou mais genes escolhidos do grupo consistindo em:
- i. o gene *tdh* que codifica a treonina desidrogenase,
- 30 ii. o gene *mdh* que codifica malato desidrogenase,
- iii. o produto de gene do quadro aberto de leitura (ORF) *yjfA* da *Escherichia coli*,

iv. o produto de gene do quadro aberto de leitura (ORF) ytfP da *Escherichia coli*,

v. o gene pckA que codifica a enzima fosfoenol piruvato carboxiquinase,

5 vi. o gene poxB que codifica piruvato oxidase,

vii. o gene dgsA que codifica o regulador DgsA do sistema fosfo-transferase

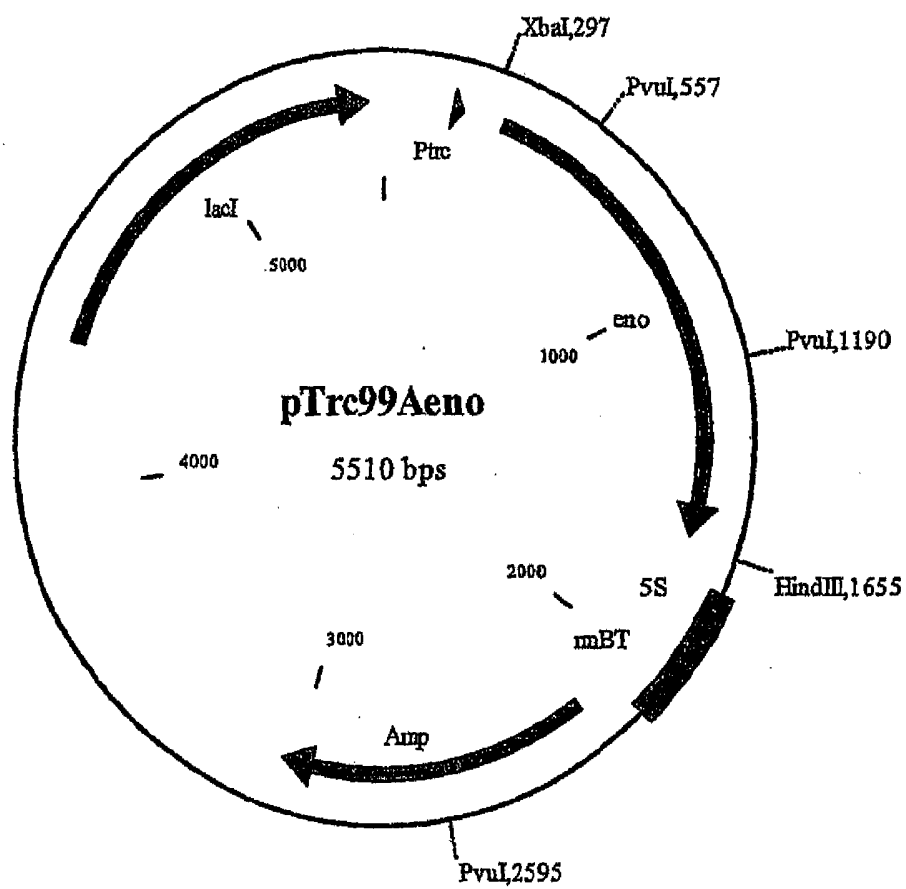
viii. o gene fruR que codifica o repressor de frutose,

ix. o gene rpoS que codifica o fator sigma³⁸, e

10 x. o gene aspA que codifica aspartato amônio liase

é ou são atenuados, especificamente, eliminados ou reduzidos na expressão.

Fig. 1: Mapeamento do plasmídeo pTrc99Aeno



Patente de Invenção: "PROCESSO PARA PREPARAÇÃO DE L-TREONINA E L-LISINA POR FERMENTAÇÃO DE MICRO-ORGANISMOS RECOMBINANTES DA FAMÍLIA *ENTEROBACTERIACEAE*".

5

- 10

- b) concentração do L-aminoácido desejado no meio ou nas células dos microorganismos, e
- c) isolamento do L-aminoácido desejado.

LISTAGEM DE SEQUÊNCIAS

<110> Evonik Degussa GmbH

<120> PROCESSO PARA PREPARAÇÃO DE L-TREONINA E L-LISINA POR
FERMENTAÇÃO DE MICRO-ORGANISMOS RECOMBINANTES DA FAMÍLIA
ENTEROBACTERIACEAE

<130> 020479 BT

<140> PI0507083-0

<141> 18.01.2005

<150> DE 10 2004 003 410.9

<151> 23.01.2004

<160> 8

<170> PatentIn versão 3.2

<210> 1

<211> 31

<212> DNA

<213> Sequência Artificial

<220>

<221> Iniciador

<222> (1)..(31)

<223> eno1

<400> 1

gtttgtctag agtttcagtt taactagtga c

31

<210> 2

<211> 25

<212> DNA

<213> Sequência Artificial

<220>

<221> Iniciador

<222> (1)..(25)

<223> eno2

<400> 2

ccggaggctg gcaagcttaa atcag

25

<210> 3
 <211> 1381
 <212> DNA
 <213> Escherichia coli

<220>
 <221> PCR product
 <222> (1)..(1381)

<220>
 <221> CDS
 <222> (46)..(1344)
 <223> eno gene

<400> 3
 gtttgtctag agtttcagtt taactagtga cttgaggaaa accta atg tcc aaa atc 57
 Met Ser Lys Ile
 1

gta aaa atc atc ggt cgt gaa atc atc gac tcc cgt ggt aac ccg act 105
 Val Lys Ile Ile Gly Arg Glu Ile Ile Asp Ser Arg Gly Asn Pro Thr
 5 10 15 20

gtt gaa gcc gaa gta cat ctg gag ggt ggt ttc gtc ggt atg gca gct 153
 Val Glu Ala Glu Val His Leu Glu Gly Gly Phe Val Gly Met Ala Ala
 25 30 35

gct ccg tca ggt gct tct act ggt tcc cgt gaa gct ctg gaa ctg cgc 201
 Ala Pro Ser Gly Ala Ser Thr Gly Ser Arg Glu Ala Leu Glu Leu Arg
 40 45 50

gat ggc gac aaa tcc cgt ttc ctg ggt aaa ggc gta acc aaa gct gtt 249
 Asp Gly Asp Lys Ser Arg Phe Leu Gly Lys Gly Val Thr Lys Ala Val
 55 60 65

gct gcg gta aac ggc ccg atc gct cag gcg ctg att ggc aaa gat gct 297
 Ala Ala Val Asn Gly Pro Ile Ala Gln Ala Leu Ile Gly Lys Asp Ala
 70 75 80

aaa gat cag gct ggc att gac aag atc atg atc gac ctg gac ggc acc 345
 Lys Asp Gln Ala Gly Ile Asp Lys Ile Met Ile Asp Leu Asp Gly Thr
 85 90 95 100

gaa aac aaa tcc aaa ttc ggc gcg aac gca atc ctg gct gta tct ctg	393
Glu Asn Lys Ser Lys Phe Gly Ala Asn Ala Ile Leu Ala Val Ser Leu	
105 110 115	
gct aac gcc aaa gct gct gca gct gct aaa ggt atg ccg ctg tac gag	441
Ala Asn Ala Lys Ala Ala Ala Ala Ala Lys Gly Met Pro Leu Tyr Glu	
120 125 130	
cac atc gct gaa ctg aac ggt act ccg ggc aaa tac tct atg ccg gtt	489
His Ile Ala Glu Leu Asn Gly Thr Pro Gly Lys Tyr Ser Met Pro Val	
135 140 145	
ccg atg atg aac atc atc aac ggt ggt gag cac gct gac aac aac gtt	537
Pro Met Met Asn Ile Ile Asn Gly Gly Glu His Ala Asp Asn Asn Val	
150 155 160	
gat atc cag gaa ttc atg att cag ccg gtt ggc gcg aaa act gtg aaa	585
Asp Ile Gln Glu Phe Met Ile Gln Pro Val Gly Ala Lys Thr Val Lys	
165 170 175 180	
gaa gcc atc cgc atg ggt tct gaa gtt ttc cat cac ctg gca aaa gtt	633
Glu Ala Ile Arg Met Gly Ser Glu Val Phe His His Leu Ala Lys Val	
185 190 195	
ctg aaa gcg aaa ggc atg aac act gct gtt ggt gac gaa ggt ggc tat	681
Leu Lys Ala Lys Gly Met Asn Thr Ala Val Gly Asp Glu Gly Gly Tyr	
200 205 210	
gcg ccg aac ctg ggt tcc aac gct gaa gct ctg gct gtt atc gct gaa	729
Ala Pro Asn Leu Gly Ser Asn Ala Glu Ala Leu Ala Val Ile Ala Glu	
215 220 225	
gct gtt aaa gct gct ggt tat gaa ctg ggc aaa gac atc act ttg gcg	777
Ala Val Lys Ala Ala Gly Tyr Glu Leu Gly Lys Asp Ile Thr Leu Ala	
230 235 240	
atg gac tgc gca gct tct gaa ttc tac aaa gat ggt aaa tac gtt ctg	825
Met Asp Cys Ala Ala Ser Glu Phe Tyr Lys Asp Gly Lys Tyr Val Leu	
245 250 255 260	
gct ggc gaa ggc aac aaa gcg ttc acc tct gaa gaa ttc act cac ttc	873
Ala Gly Glu Gly Asn Lys Ala Phe Thr Ser Glu Glu Phe Thr His Phe	
265 270 275	
ctg gaa gaa ctg acc aaa cag tac ccg atc gtt tct atc gaa gac ggt	921
Leu Glu Glu Leu Thr Lys Gln Tyr Pro Ile Val Ser Ile Glu Asp Gly	
280 285 290	

ctg gac gaa tct gac tgg gac ggt ttc gca tac cag acc aaa gtt ctg	969
Leu Asp Glu Ser Asp Trp Asp Gly Phe Ala Tyr Gln Thr Lys Val Leu	
295 300 305	
ggc gac aaa atc cag ctg gtt ggt gac gac ctg ttc gta acc aac acc	1017
Gly Asp Lys Ile Gln Leu Val Gly Asp Asp Leu Phe Val Thr Asn Thr	
310 315 320	
aag atc ctg aaa gaa ggt atc gaa aaa ggt atc gct aac tcc atc ctg	1065
Lys Ile Leu Lys Glu Gly Ile Glu Lys Gly Ile Ala Asn Ser Ile Leu	
325 330 335 340	
atc aaa ttc aac cag atc ggt tct ctg acc gaa act ctg gct gca atc	1113
Ile Lys Phe Asn Gln Ile Gly Ser Leu Thr Glu Thr Leu Ala Ala Ile	
345 350 355	
aag atg gcg aaa gat gct ggc tac act gca gtt atc tct cac cgt tct	1161
Lys Met Ala Lys Asp Ala Gly Tyr Thr Ala Val Ile Ser His Arg Ser	
360 365 370	
ggc gaa act gaa gac gct acc atc gct gac ctg gct gtt ggt act gct	1209
Gly Glu Thr Glu Asp Ala Thr Ile Ala Asp Leu Ala Val Gly Thr Ala	
375 380 385	
gca ggc cag atc aaa act ggt tct atg agc cgt tct gac cgt gtt gct	1257
Ala Gly Gln Ile Lys Thr Gly Ser Met Ser Arg Ser Asp Arg Val Ala	
390 395 400	
aaa tac aac cag ctg att cgt atc gaa gaa gct ctg ggc gaa aaa gca	1305
Lys Tyr Asn Gln Leu Ile Arg Ile Glu Glu Ala Leu Gly Glu Lys Ala	
405 410 415 420	
ccg tac aac ggt cgt aaa gag atc aaa ggc cag gca taa gactgacttt	1354
Pro Tyr Asn Gly Arg Lys Glu Ile Lys Gly Gln Ala	
425 430	
atctgattta agcttgccag cctccgg	1381

<210> 4

<211> 432

<212> PRT

<213> Escherichia coli

<400> 4

5

Met Ser Lys Ile Val Lys Ile Ile Gly Arg Glu Ile Ile Asp Ser Arg
 1 5 10 15

Gly Asn Pro Thr Val Glu Ala Glu Val His Leu Glu Gly Gly Phe Val
 20 25 30

Gly Met Ala Ala Ala Pro Ser Gly Ala Ser Thr Gly Ser Arg Glu Ala
 35 40 45

Leu Glu Leu Arg Asp Gly Asp Lys Ser Arg Phe Leu Gly Lys Gly Val
 50 55 60

Thr Lys Ala Val Ala Ala Val Asn Gly Pro Ile Ala Gln Ala Leu Ile
 65 70 75 80

Gly Lys Asp Ala Lys Asp Gln Ala Gly Ile Asp Lys Ile Met Ile Asp
 85 90 95

Leu Asp Gly Thr Glu Asn Lys Ser Lys Phe Gly Ala Asn Ala Ile Leu
 100 105 110

Ala Val Ser Leu Ala Asn Ala Lys Ala Ala Ala Ala Ala Lys Gly Met
 115 120 125

Pro Leu Tyr Glu His Ile Ala Glu Leu Asn Gly Thr Pro Gly Lys Tyr
 130 135 140

Ser Met Pro Val Pro Met Met Asn Ile Ile Asn Gly Gly Glu His Ala
 145 150 155 160

Asp Asn Asn Val Asp Ile Gln Glu Phe Met Ile Gln Pro Val Gly Ala
 165 170 175

Lys Thr Val Lys Glu Ala Ile Arg Met Gly Ser Glu Val Phe His His
 180 185 190

Leu Ala Lys Val Leu Lys Ala Lys Gly Met Asn Thr Ala Val Gly Asp
 195 200 205

Glu Gly Gly Tyr Ala Pro Asn Leu Gly Ser Asn Ala Glu Ala Leu Ala
 210 215 220

Val Ile Ala Glu Ala Val Lys Ala Ala Gly Tyr Glu Leu Gly Lys Asp
 225 230 235 240

Ile Thr Leu Ala Met Asp Cys Ala Ala Ser Glu Phe Tyr Lys Asp Gly
 245 250 255

Lys Tyr Val Leu Ala Gly Glu Gly Asn Lys Ala Phe Thr Ser Glu Glu
 260 265 270

Phe Thr His Phe Leu Glu Glu Leu Thr Lys Gln Tyr Pro Ile Val Ser
 275 280 285

Ile Glu Asp Gly Leu Asp Glu Ser Asp Trp Asp Gly Phe Ala Tyr Gln
 290 295 300

Thr Lys Val Leu Gly Asp Lys Ile Gln Leu Val Gly Asp Asp Leu Phe
 305 310 315 320

Val Thr Asn Thr Lys Ile Leu Lys Glu Gly Ile Glu Lys Gly Ile Ala
 325 330 335

Asn Ser Ile Leu Ile Lys Phe Asn Gln Ile Gly Ser Leu Thr Glu Thr
 340 345 350

Leu Ala Ala Ile Lys Met Ala Lys Asp Ala Gly Tyr Thr Ala Val Ile
 355 360 365

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

Val Gly Thr Ala Ala Gly Gln Ile Lys Thr Gly Ser Met Ser Arg Ser
385 390 395 400

Asp Arg Val Ala Lys Tyr Asn Gln Leu Ile Arg Ile Glu Glu Ala Leu
405 410 415

Gly Glu Lys Ala Pro Tyr Asn Gly Arg Lys Glu Ile Lys Gly Gln Ala
420 425 430

<210> 5

<211> 1299

<212> DNA

<213> *Shigella flexneri*

<220>

<221> CDS

<222> (1)..(1299)

<223> eno coding region

<400> 5

atg tcc aaa atc gta aaa atc atc ggt cgt gaa atc atc gac tcc cgt 48
Met Ser Lys Ile Val Lys Ile Ile Gly Arg Glu Ile Ile Asp Ser Arg
1 5 10 15

ggt aac ccg act gtt gaa gcc gaa gta cat ctg gag ggt ggt ttc gtc 96
Gly Asn Pro Thr Val Glu Ala Glu Val His Leu Glu Gly Gly Phe Val
20 25 30

ggt atg gca gct gct ccg tca ggt gct tct act ggt tcc cgt gaa gct 144
Gly Met Ala Ala Ala Pro Ser Gly Ala Ser Thr Gly Ser Arg Glu Ala
35 40 45

ctg gaa ctg cgc gat ggc gac aaa tcc cgt ttc ctg ggt aaa ggc gta 192
Leu Glu Leu Arg Asp Gly Asp Lys Ser Arg Phe Leu Gly Lys Gly Val
50 55 60

acc aaa gct gtt gct gcg gta aac ggc ccg atc gct cag gcg ctg att 240
Thr Lys Ala Val Ala Ala Val Asn Gly Pro Ile Ala Gln Ala Leu Ile
65 70 75 80

ggc aaa gat gct aaa gat cag gct ggc att gac aag atc atg atc gac 288

Gly	Lys	Asp	Ala	Lys	Asp	Gln	Ala	Gly	Ile	Asp	Lys	Ile	Met	Ile	Asp	
				85					90					95		
ctg	gac	ggc	acc	gaa	aac	aaa	tcc	aaa	ttc	ggc	gcg	aac	gca	atc	ctg	336
Leu	Asp	Gly	Thr	Glu	Asn	Lys	Ser	Lys	Phe	Gly	Ala	Asn	Ala	Ile	Leu	
			100					105					110			
gct	gta	tct	ctg	gct	aac	gcc	aaa	gct	gct	gca	gct	gct	aaa	ggc	atg	384
Ala	Val	Ser	Leu	Ala	Asn	Ala	Lys	Ala	Ala	Ala	Ala	Ala	Lys	Gly	Met	
			115				120						125			
ccg	ctg	tac	gag	cac	atc	gct	gaa	ctg	aac	ggc	act	ccg	ggc	aaa	tac	432
Pro	Leu	Tyr	Glu	His	Ile	Ala	Glu	Leu	Asn	Gly	Thr	Pro	Gly	Lys	Tyr	
	130					135				140						
tcc	atg	ccg	gtt	ccg	atg	atg	aac	atc	atc	aac	ggc	ggc	gag	cac	gct	480
Ser	Met	Pro	Val	Pro	Met	Met	Asn	Ile	Ile	Asn	Gly	Gly	Glu	His	Ala	
145					150					155				160		
gac	aac	aac	gtt	gat	atc	cag	gaa	ttc	atg	att	cag	ccg	gtt	ggc	gcg	528
Asp	Asn	Asn	Val	Asp	Ile	Gln	Glu	Phe	Met	Ile	Gln	Pro	Val	Gly	Ala	
			165						170					175		
aaa	act	gtg	aaa	gaa	gcc	atc	cgc	atg	ggc	tct	gaa	gtt	ttc	cat	cac	576
Lys	Thr	Val	Lys	Glu	Ala	Ile	Arg	Met	Gly	Ser	Glu	Val	Phe	His	His	
			180					185					190			
ctg	gca	aaa	gtt	ctg	aaa	gcg	aaa	ggc	atg	aac	act	gct	gtt	ggc	gac	624
Leu	Ala	Lys	Val	Leu	Lys	Ala	Lys	Gly	Met	Asn	Thr	Ala	Val	Gly	Asp	
			195				200					205				
gaa	ggc	ggc	tat	gcg	ccg	aac	ctg	ggc	tcc	aac	gct	gaa	gct	ctg	gct	672
Glu	Gly	Gly	Tyr	Ala	Pro	Asn	Leu	Gly	Ser	Asn	Ala	Glu	Ala	Leu	Ala	
	210					215					220					
gtt	atc	gct	gaa	gct	gtt	aaa	gct	gct	ggc	tat	gaa	ctg	ggc	aaa	gac	720
Val	Ile	Ala	Glu	Ala	Val	Lys	Ala	Ala	Gly	Tyr	Glu	Leu	Gly	Lys	Asp	
225					230				235					240		
atc	act	ctg	gcg	atg	gac	tgc	gca	gct	tct	gaa	ttc	tac	aaa	gat	ggc	768
Ile	Thr	Leu	Ala	Met	Asp	Cys	Ala	Ala	Ser	Glu	Phe	Tyr	Lys	Asp	Gly	
			245						250					255		
aaa	tac	gtt	ctg	gct	ggc	gaa	ggc	aac	aaa	gcg	ttc	acc	tct	gaa	gaa	816
Lys	Tyr	Val	Leu	Ala	Gly	Glu	Gly	Asn	Lys	Ala	Phe	Thr	Ser	Glu	Glu	
			260					265					270			

ttc act cac ttc ctg gaa gaa ctg acc aaa cag tac ccg atc gtt tct 864
Phe Thr His Phe Leu Glu Glu Leu Thr Lys Gln Tyr Pro Ile Val Ser
275 280 285

atc gaa gac ggt ctg gac gaa tct gac tgg gac ggt ttc gca tac cag 912
Ile Glu Asp Gly Leu Asp Glu Ser Asp Trp Asp Gly Phe Ala Tyr Gln
290 295 300

acc aaa gtt ctg ggc gac aaa atc cag ctg gtt ggt gac gac ctg ttc 960
Thr Lys Val Leu Gly Asp Lys Ile Gln Leu Val Gly Asp Asp Leu Phe
305 310 315 320

gta acc aac acc aag atc ctg aaa gaa ggt atc gaa aaa ggt atc gct 1008
Val Thr Asn Thr Lys Ile Leu Lys Glu Gly Ile Glu Lys Gly Ile Ala
325 330 335

aac tcc atc ctg atc aaa ttc aac cag atc ggt tct ctg acc gaa act 1056
Asn Ser Ile Leu Ile Lys Phe Asn Gln Ile Gly Ser Leu Thr Glu Thr
340 345 350

ctg gct gca atc aag atg gcg aaa gat gct ggc tac act gca gtt atc 1104
Leu Ala Ala Ile Lys Met Ala Lys Asp Ala Gly Tyr Thr Ala Val Ile
355 360 365

tct cac cgt tct ggt gaa act gaa gat gct acc atc gct gat ctg gct 1152
Ser His Arg Ser Gly Glu Thr Glu Asp Ala Thr Ile Ala Asp Leu Ala
370 375 380

gtt ggt act gct gca ggc cag atc aaa act ggt tct atg agc cgt tct 1200
Val Gly Thr Ala Ala Gly Gln Ile Lys Thr Gly Ser Met Ser Arg Ser
385 390 395 400

gac cgt gtt gct aaa tac aac cag ctg att cgt atc gaa gaa gct ctg 1248
Asp Arg Val Ala Lys Tyr Asn Gln Leu Ile Arg Ile Glu Glu Ala Leu
405 410 415

ggc gaa aaa gca ccg tac aac ggt cgt aaa gag atc aaa ggc cag gca 1296
Gly Glu Lys Ala Pro Tyr Asn Gly Arg Lys Glu Ile Lys Gly Gln Ala
420 425 430

taa 1299

<210> 6

<211> 432

<212> PRT

<213> *Shigella flexneri*

<400> 6

Met	Ser	Lys	Ile	Val	Lys	Ile	Ile	Gly	Arg	Glu	Ile	Ile	Asp	Ser	Arg
1				5					10					15	

Gly	Asn	Pro	Thr	Val	Glu	Ala	Glu	Val	His	Leu	Glu	Gly	Gly	Phe	Val
			20					25					30		

Gly	Met	Ala	Ala	Ala	Pro	Ser	Gly	Ala	Ser	Thr	Gly	Ser	Arg	Glu	Ala
		35						40				45			

Leu	Glu	Leu	Arg	Asp	Gly	Asp	Lys	Ser	Arg	Phe	Leu	Gly	Lys	Gly	Val
	50					55						60			

Thr	Lys	Ala	Val	Ala	Ala	Val	Asn	Gly	Pro	Ile	Ala	Gln	Ala	Leu	Ile
65					70					75					80

Gly	Lys	Asp	Ala	Lys	Asp	Gln	Ala	Gly	Ile	Asp	Lys	Ile	Met	Ile	Asp
				85					90					95	

Leu	Asp	Gly	Thr	Glu	Asn	Lys	Ser	Lys	Phe	Gly	Ala	Asn	Ala	Ile	Leu
			100					105					110		

Ala	Val	Ser	Leu	Ala	Asn	Ala	Lys	Ala	Ala	Ala	Ala	Ala	Lys	Gly	Met
		115					120						125		

Pro	Leu	Tyr	Glu	His	Ile	Ala	Glu	Leu	Asn	Gly	Thr	Pro	Gly	Lys	Tyr
	130					135					140				

Ser	Met	Pro	Val	Pro	Met	Met	Asn	Ile	Ile	Asn	Gly	Gly	Glu	His	Ala
145					150					155					160

Asp	Asn	Asn	Val	Asp	Ile	Gln	Glu	Phe	Met	Ile	Gln	Pro	Val	Gly	Ala
				165					170					175	

Lys Thr Val Lys Glu Ala Ile Arg Met Gly Ser Glu Val Phe His His
 180 185 190

Leu Ala Lys Val Leu Lys Ala Lys Gly Met Asn Thr Ala Val Gly Asp
 195 200 205

Glu Gly Gly Tyr Ala Pro Asn Leu Gly Ser Asn Ala Glu Ala Leu Ala
 210 215 220

Val Ile Ala Glu Ala Val Lys Ala Ala Gly Tyr Glu Leu Gly Lys Asp
 225 230 235 240

Ile Thr Leu Ala Met Asp Cys Ala Ala Ser Glu Phe Tyr Lys Asp Gly
 245 250 255

Lys Tyr Val Leu Ala Gly Glu Gly Asn Lys Ala Phe Thr Ser Glu Glu
 260 265 270

Phe Thr His Phe Leu Glu Glu Leu Thr Lys Gln Tyr Pro Ile Val Ser
 275 280 285

Ile Glu Asp Gly Leu Asp Glu Ser Asp Trp Asp Gly Phe Ala Tyr Gln
 290 295 300

Thr Lys Val Leu Gly Asp Lys Ile Gln Leu Val Gly Asp Asp Leu Phe
 305 310 315 320

Val Thr Asn Thr Lys Ile Leu Lys Glu Gly Ile Glu Lys Gly Ile Ala
 325 330 335

Asn Ser Ile Leu Ile Lys Phe Asn Gln Ile Gly Ser Leu Thr Glu Thr
 340 345 350

Leu Ala Ala Ile Lys Met Ala Lys Asp Ala Gly Tyr Thr Ala Val Ile
 355 360 365

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

Val Gly Thr Ala Ala Gly Gln Ile Lys Thr Gly Ser Met Ser Arg Ser
 385 390 395 400

Asp Arg Val Ala Lys Tyr Asn Gln Leu Ile Arg Ile Glu Glu Ala Leu
 405 410 415

Gly Glu Lys Ala Pro Tyr Asn Gly Arg Lys Glu Ile Lys Gly Gln Ala
 420 425 430

<210> 7
 <211> 1299
 <212> DNA
 <213> Salmonella typhimurium

<220>
 <221> CDS
 <222> (1)..(1299)
 <223> eno coding region

<400> 7
 atg tcc aaa atc gtt aaa gtc atc ggt cgt gaa atc atc gac tcc cgt 48
 Met Ser Lys Ile Val Lys Val Ile Gly Arg Glu Ile Ile Asp Ser Arg
 1 5 10 15

ggt aac ccg act gtt gaa gct gaa gta cac ctg gaa ggt ggt ttc gta 96
 Gly Asn Pro Thr Val Glu Ala Glu Val His Leu Glu Gly Gly Phe Val
 20 25 30

ggt atg gcg gcg gct ccg tca ggt gct tct act ggt tcc cgc gaa gcg 144
 Gly Met Ala Ala Ala Pro Ser Gly Ala Ser Thr Gly Ser Arg Glu Ala
 35 40 45

ctg gaa ctg cgc gat ggc gac aaa tcc cgt ttc ctg ggt aaa ggc gta 192
 Leu Glu Leu Arg Asp Gly Asp Lys Ser Arg Phe Leu Gly Lys Gly Val
 50 55 60

acc aaa gct gtt ggc gcg gtt aac ggc ccg atc gct cag gct att ctt 240
 Thr Lys Ala Val Gly Ala Val Asn Gly Pro Ile Ala Gln Ala Ile Leu

65	70	75	80	
ggc aaa gac gct aaa gac cag gct ggc atc gac aaa atc atg atc gac				288
Gly Lys Asp Ala Lys Asp Gln Ala Gly Ile Asp Lys Ile Met Ile Asp				
	85	90	95	
ctg gac ggt act gaa aac aaa tct aac ttc ggt gca aac gcc att ctg				336
Leu Asp Gly Thr Glu Asn Lys Ser Asn Phe Gly Ala Asn Ala Ile Leu				
	100	105	110	
gct gtc tct ctg gct aac gcc aaa gct gct gct gcc gct aaa ggt atg				384
Ala Val Ser Leu Ala Asn Ala Lys Ala Ala Ala Ala Ala Lys Gly Met				
	115	120	125	
ccg ctg tac gag cac att gct gaa ctg aat ggc acg ccg ggc aaa tac				432
Pro Leu Tyr Glu His Ile Ala Glu Leu Asn Gly Thr Pro Gly Lys Tyr				
	130	135	140	
tcc atg ccg gtt ccg atg atg aac atc atc aac ggc ggc gag cac gct				480
Ser Met Pro Val Pro Met Met Asn Ile Ile Asn Gly Gly Glu His Ala				
	145	150	155	160
gac aac aac gtc gac atc cag gaa ttc atg atc cag ccg gtt ggc gcg				528
Asp Asn Asn Val Asp Ile Gln Glu Phe Met Ile Gln Pro Val Gly Ala				
	165	170	175	
aaa acg gtt aaa gaa gcc atc cgt atg ggt tct gaa gtt ttc cat cac				576
Lys Thr Val Lys Glu Ala Ile Arg Met Gly Ser Glu Val Phe His His				
	180	185	190	
ctg gca aaa gtg ctg aaa ggc aaa ggc atg aac acc gct gtg ggt gac				624
Leu Ala Lys Val Leu Lys Gly Lys Gly Met Asn Thr Ala Val Gly Asp				
	195	200	205	
gaa ggc ggc tat gcg ccg aac ctg ggt tcc aac gct gaa gcg ctg gct				672
Glu Gly Gly Tyr Ala Pro Asn Leu Gly Ser Asn Ala Glu Ala Leu Ala				
	210	215	220	
gtt atc gct gaa gcg gtt aaa gcg gct ggt tac gag ctg ggt aaa gac				720
Val Ile Ala Glu Ala Val Lys Ala Ala Gly Tyr Glu Leu Gly Lys Asp				
	225	230	235	240
atc acc ctg gcg atg gac tgc gca gca tct gaa ttc tac aaa gac ggt				768
Ile Thr Leu Ala Met Asp Cys Ala Ala Ser Glu Phe Tyr Lys Asp Gly				
	245	250	255	
aaa tac gtt ctg gct ggc gaa ggc aac aaa gcg ttc acc tcc gaa gaa				816

Lys Tyr Val Leu Ala Gly Glu Gly Asn Lys Ala Phe Thr Ser Glu Glu	
260 265 270	
ttc acc cac ttc ctg gaa gaa ctg acc aaa cag tac ccg atc gtt tcc	864
Phe Thr His Phe Leu Glu Glu Leu Thr Lys Gln Tyr Pro Ile Val Ser	
275 280 285	
atc gaa gat ggt ctg gac gag tct gac tgg gac ggt ttt gca tac cag	912
Ile Glu Asp Gly Leu Asp Glu Ser Asp Trp Asp Gly Phe Ala Tyr Gln	
290 295 300	
acc aaa gta ctg ggc gac aaa atc cag ctg gtt ggt gac gac ctg ttc	960
Thr Lys Val Leu Gly Asp Lys Ile Gln Leu Val Gly Asp Asp Leu Phe	
305 310 315 320	
gta acc aac acc aaa atc ctg aaa gaa ggc atc gag aaa ggc atc gct	1008
Val Thr Asn Thr Lys Ile Leu Lys Glu Gly Ile Glu Lys Gly Ile Ala	
325 330 335	
aac tcc atc ctg atc aaa ttc aac cag atc ggt tct ctg acc gaa act	1056
Asn Ser Ile Leu Ile Lys Phe Asn Gln Ile Gly Ser Leu Thr Glu Thr	
340 345 350	
ctg gct gca atc aag atg gcg aaa gac gct ggc tat act gct gtc atc	1104
Leu Ala Ala Ile Lys Met Ala Lys Asp Ala Gly Tyr Thr Ala Val Ile	
355 360 365	
tct cac cgt tct ggc gaa act gaa gac gct acc atc gct gac ctg gct	1152
Ser His Arg Ser Gly Glu Thr Glu Asp Ala Thr Ile Ala Asp Leu Ala	
370 375 380	
gtt ggt acc gct gca ggc cag atc aaa acc ggt tct atg agc cgt tct	1200
Val Gly Thr Ala Ala Gly Gln Ile Lys Thr Gly Ser Met Ser Arg Ser	
385 390 395 400	
gac cgt gtt gct aaa tac aac cag ctg att cgt atc gaa gaa gct ctg	1248
Asp Arg Val Ala Lys Tyr Asn Gln Leu Ile Arg Ile Glu Glu Ala Leu	
405 410 415	
ggc gaa aaa gca ccg tat aac ggt cgt aaa gag atc aaa ggc cag gcg	1296
Gly Glu Lys Ala Pro Tyr Asn Gly Arg Lys Glu Ile Lys Gly Gln Ala	
420 425 430	
taa	1299

<211> 432

<212> PRT

<213> Salmonella typhimurium

<400> 8

Met Ser Lys Ile Val Lys Val Ile Gly Arg Glu Ile Ile Asp Ser Arg
 1 5 10 15

Gly Asn Pro Thr Val Glu Ala Glu Val His Leu Glu Gly Gly Phe Val
 20 25 30

Gly Met Ala Ala Ala Pro Ser Gly Ala Ser Thr Gly Ser Arg Glu Ala
 35 40 45

Leu Glu Leu Arg Asp Gly Asp Lys Ser Arg Phe Leu Gly Lys Gly Val
 50 55 60

Thr Lys Ala Val Gly Ala Val Asn Gly Pro Ile Ala Gln Ala Ile Leu
 65 70 75 80

Gly Lys Asp Ala Lys Asp Gln Ala Gly Ile Asp Lys Ile Met Ile Asp
 85 90 95

Leu Asp Gly Thr Glu Asn Lys Ser Asn Phe Gly Ala Asn Ala Ile Leu
 100 105 110

Ala Val Ser Leu Ala Asn Ala Lys Ala Ala Ala Ala Lys Gly Met
 115 120 125

Pro Leu Tyr Glu His Ile Ala Glu Leu Asn Gly Thr Pro Gly Lys Tyr
 130 135 140

Ser Met Pro Val Pro Met Met Asn Ile Ile Asn Gly Gly Glu His Ala
 145 150 155 160

Asp Asn Asn Val Asp Ile Gln Glu Phe Met Ile Gln Pro Val Gly Ala

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

Leu Ala Ala Ile Lys Met Ala Lys Asp Ala Gly Tyr Thr Ala Val Ile
355 360 365

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

Val Gly Thr Ala Ala Gly Gln Ile Lys Thr Gly Ser Met Ser Arg Ser
385 390 395 400

Asp Arg Val Ala Lys Tyr Asn Gln Leu Ile Arg Ile Glu Glu Ala Leu
405 410 415

Gly Glu Lys Ala Pro Tyr Asn Gly Arg Lys Glu Ile Lys Gly Gln Ala
420 425 430

Este anexo apresenta o código de controle da listagem de sequências biológicas.

Código de Controle

Campo 1



EAFCF27154BA474D

Campo 2



F920E88CE64436A4

Outras Informações:

- Nome do Arquivo: p139655 25184993_1.txt
- Data de Geração do Código: 02-07-2015
- Hora de Geração do Código: 13:37:42
- Código de Controle:
 - Campo 1: EAFCF27154BA474D
 - Campo 2: F920E88CE64436A4