

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

(21) **PI 0710412-0 A2**

(22) Data de Depósito: 02/05/2007
(43) Data da Publicação: 18/09/2012
(RPI 2176)



(51) *Int.Cl.:*
C12P 7/16
C12P 7/26
C12N 1/21
C12N 15/52

(54) Título: "CELULAS HOSPEDEIRAS MICROBIANAS RECOMBINANTES, MÉTODO PARA PRODUÇÃO DE 2-BUTANOL, MÉTODO PARA PRODUÇÃO DE 2-BUTANONA, 2-BUTANOL E 2-BUTANONA"

(30) Prioridade Unionista: 02/05/2006 US 60/796,816, 21/12/2006 US 60/871,156

(73) Titular(es): E.I. DU PONT DE NEMOURS AND COMPANY

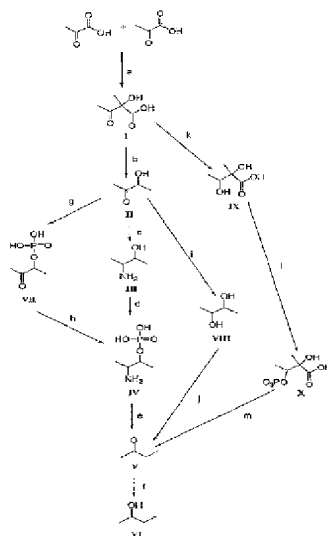
(72) Inventor(es): Andrew C. Eliot, Charles E. Nakamura, Gail K. Donaldson, JEAN-FRANCOIS TOMB, Vasantha Nagarajan

(74) Procurador(es): Paola Calabria Mattioli

(86) Pedido Internacional: PCT US2007010744 de 02/05/2007

(87) Publicação Internacional: WO 2007/130521 de 15/11/2007

(57) Resumo: CÉLULAS HOSPEDEIRAS MICROBIANAS RECOMBINANTES, MÉTODO PARA PRODUÇÃO DE 2-BUTANOL, MÉTODO PARA PRODUÇÃO DE 2-BUTANONA, 2-BUTANOL E 2-BUTANONA. A presente invenção apresenta métodos para a produção fermentativa de álcoois de quatro carbonos. Especificamente butanol, de preferência 2-butanol, é produzido pelo crescimento fermentativo de bactérias recombinantes que expressam 2-butanol por via biossintética. Os microorganismos recombinantes e os métodos da invenção podem também ser adaptados para produzir 2-butanona, um intermediário nas vias biossintéticas do 2-butanol aqui apresentadas.



**"CÉLULAS HOSPEDEIRAS MICROBIANAS RECOMBINANTES, MÉTODO
PARA PRODUÇÃO DE 2-BUTANOL, MÉTODO PARA PRODUÇÃO DE 2-
BUTANONA, 2-BUTANOL E 2-BUTANONA"**

REFERÊNCIA REMISSIVA A PEDIDO CORRELATO

5 Este pedido reivindica a prioridade sujeita a 35 U.S.C. §119 do
pedido provisório U.S. Nº de série 60/796816, depositado em 2 de maio de
2006 e pedido provisório U.S. Nº de série 60/871156, depositado em 21 de
dezembro de 2006.

CAMPO DA INVENÇÃO

10 A invenção refere-se ao campo de microbiologia industrial e à
produção de álcoois. Mais especificamente, 2-butanol é produzido via
fermentação industrial de um microorganismo recombinante. Os
microorganismos recombinantes e os métodos da invenção podem também
ser adaptados para produzir 2-butanona, um intermediário nas vias
15 biossintéticas do 2-butanol aqui apresentadas.

ANTECEDENTES DA INVENÇÃO

O butanol é um importante produto químico, útil como aditivo
para combustível, como matéria-prima química na indústria de plásticos, e
como um extrator de grau alimentício na indústria de alimentos e de sabor e
20 aroma. Cada ano entre 4,5 bilhões de quilogramas (10 bilhões de libras) e 5,4
bilhões de quilogramas (12 bilhões de libras) de butanol são produzidos por
meios petroquímicos e a necessidade deste produto químico provavelmente
irá aumentar. A 2-butanona, também chamada de metil etil cetona (MEK), é
um solvente usado amplamente e é a mais importante cetona produzida
25 comercialmente depois de acetona. Ela é usada como solvente para tintas,
resinas e adesivos, e como um extrator seletivo e ativador de reações
oxidantes.

Métodos para a síntese química de 2-butanona são conhecidos,

como a desidrogenação de 2-butanol ou em um processo onde butano líquido é oxidado cataliticamente formando 2-butanona e ácido acético (*Ullmann's Encyclopedia of Industrial Chemistry*, 6th edition, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Germany, Vol. 5, pp.727-732). A 2-butanona pode também ser convertida quimicamente a 2-butanol por hidrogenação (Breen et al., J. or Catalysis 236: 270-281 (2005)). Métodos para a síntese química de 2-butanol são conhecidos, como a hidratação do n-buteno (*Ullmann's Encyclopedia of Industrial Chemistry*, 6th edition, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Germany, Vol. 5, pp. 716-719). Estes processos usam matérias-primas derivadas de produtos petroquímicos e geralmente são caros, e não são ambientalmente amigáveis. A produção de 2-butanona e 2-butanol a partir de matérias-primas derivadas de plantas minimizaria as emissões de gases responsáveis pelo efeito estufa, o que representaria um avanço na técnica.

Os métodos de produção de 2-butanol por biotransformação de outros produtos químicos orgânicos também são conhecidos. Por exemplo, Stampfer et al. (WO 03/078615) descrevem a produção de álcoois secundários, como 2-butanol, pela redução de cetonas, que são catalisadas por uma enzima álcool desidrogenase obtida a partir de *Rhodococcus ruber*. Semelhantemente, Kojima et al. (EP 0645453) descreve um método para a preparação de álcoois secundários, como 2-butanol, pela redução de cetonas, que são catalisadas por uma enzima álcool desidrogenase secundária obtida a partir de *Candida parapsilosis*. Adicionalmente, Kuehnle et al. (EP 1149918) descrevem um processo que produz 1-butanol e 2-butanol pela oxidação de hidrocarbonetos por diferentes cepas de *Rhodococcus ruber*. O processo favoreceu a produção de 1-butanol com uma seletividade de 93,8%.

A produção de 2-butanol por certas cepas de *Lactobacilli* também é conhecida (Speranza et. al. *J. Agric. Food Chem.* (1997) 45:3476-3480). O 2-butanol é produzido pela transformação do *meso*-2,3-butanodiol. A produção

de 2-butanol a partir de acetolactato e acetoína por estas cepas de *Lactobacilli* também foi demonstrada. No entanto, não houve relatos de um microorganismo recombinante projetado para produzir 2-butanol.

Existe uma necessidade, portanto, de processos para a produção de 2-butanol e 2-butanona ambientalmente responsáveis e de baixo custo. A presente invenção aborda esta necessidade e revela hospedeiros para produção microbiana recombinante que expressam 2-butanol e 2-butanona por vias biossintéticas.

DESCRIÇÃO RESUMIDA DA INVENÇÃO

A invenção apresenta um microorganismo recombinante tendo uma via biossintética do 2-butanol projetada. Também é fornecido um microorganismo recombinante tendo uma via biossintética da 2-butanona projetada, que é a mesma que a via biossintética de 2-butanol com a omissão da última etapa. Os microorganismos projetados podem ser usados para a produção comercial de 2-butanol ou 2-butanona.

Conseqüentemente, a invenção apresenta uma célula hospedeira microbiana recombinante que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol;
- iv) 2,3-butanodiol para 2-butanona; e
- v) 2-butanona para 2-butanol;

sendo que pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana e sendo que a dita célula hospedeira microbiana produz 2-butanol.

Em outra realização, a invenção apresenta uma célula hospedeira

microbiana recombinante que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol; e
- iv) 2,3-butanodiol para 2-butanona;

sendo que pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana e sendo que a dita célula hospedeira microbiana produz 2-butanona.

Em outra modalidade a invenção apresenta um método para a produção de 2-butanol que compreende:

1) fornecer uma célula hospedeira microbiana recombinante que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol;
- iv) 2,3-butanodiol para 2-butanona; e
- v) 2-butanona para 2-butanol;

sendo que pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana;

2) colocar a célula hospedeira de (1) em contato com um substrato fermentável de carbono em um meio de fermentação nas condições em que 2-butanol é produzido.

Semelhantemente a invenção apresenta um método para a produção de 2-butanona que compreende:

1) fornecer uma célula hospedeira microbiana recombinante que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol; e
- iv) 2,3-butanodiol para 2-butanona;

sendo que pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana;

2) colocar a célula hospedeira de (1) em contato com um substrato fermentável de carbono em um meio de fermentação nas condições em que a 2-butanona é produzida.

Em outra modalidade a invenção apresenta um 2-butanol ou 2-butanona contendo um meio de fermentação de produto produzido pelo método da invenção.

BREVE DESCRIÇÃO DAS FIGURAS, TABELAS E

DESCRIÇÕES DE SEQÜÊNCIAS

A invenção pode ser mais plenamente entendida a partir da seguinte descrição detalhada, figuras e descrições de seqüências anexas, que fazem parte do presente pedido.

A figura 1 mostra quatro vias diferentes para biossíntese de 2-butanona e 2-butanol.

A figura 2 mostra uma árvore filogenética de comprimento inteiro das subunidades grandes das diol/glicerol desidratases, com >95% de seqüências idênticas removidas (exceto que todas as seqüências de funções experimentalmente verificadas foram retidas), e um código lista a identidade de cada seqüência na árvore. As seqüências com função experimentalmente

determinada como diol ou glicerol desidratases são ressaltadas em cinza escuro ou claro, respectivamente.

A figura 3 mostra uma árvore filogenética de comprimento inteiro das subunidades médias das diol/glicerol desidratases, com >95% de seqüências idênticas removidas, e um código lista a identidade de cada seqüência na árvore. As seqüências com função experimentalmente determinada como diol ou glicerol desidratases são ressaltadas em cinza escuro ou claro, respectivamente.

A figura 4 mostra uma árvore filogenética de comprimento inteiro das subunidades pequenas das diol/glicerol desidratases, com >95% de seqüências idênticas removidas, e um código lista a identidade de cada seqüência na árvore. As seqüências com função experimentalmente determinada como diol ou glicerol desidratases são ressaltadas em cinza escuro ou claro, respectivamente.

A tabela 12 é uma tabela do perfil HMM da subunidade grande alfa para a enzima de diol/glicerol desidratase. A tabela 12 é submetida juntamente eletronicamente e está aqui incorporada, a título de referência.

A tabela 13 é uma tabela do perfil HMM da subunidade média beta para a enzima de diol/glicerol desidratase. A tabela 13 é submetida juntamente eletronicamente e está aqui incorporada, a título de referência.

A tabela 14 é uma tabela do perfil HMM da subunidade pequena gama para a enzima de diol/glicerol desidratase. A tabela 14 é submetida juntamente eletronicamente e está aqui incorporada, a título de referência.

As seqüências a seguir obedecem a 37 C.F.R. 1,821-1,825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") e são consistentes com World Intellectual Property Organization (WIPO) [Organização mundial de propriedade intelectual] padrão ST.25 (1998) e os

requisitos para sequência listados pelo EPO e PCT (regras 5.2 e 49.5 (a-bis), e seção 208 e anexo C das instruções administrativas). Os símbolos e formato usados para os dados de sequência de aminoácidos e de nucleotídeos estão de acordo com as regras estabelecidas no 37 CFR § 1,822.

5

TABELA 1**SUMÁRIO DE NÚMEROS DE SEQ ID DE ÁCIDO NUCLÉICO E PROTEÍNA**

Descrição	SEQ ID Ácido nucléico	SEQ ID Proteína
<i>budA</i> , acetolactato descarboxilase a partir de <i>Klebsiella pneumoniae</i> ATCC 25955	1	2
<i>alsD</i> , acetolactato descarboxilase a partir de <i>Bacillus subtilis</i>	80	81
<i>budA</i> , acetolactato descarboxilase a partir de <i>Klebsiella terrigena</i>	82	83
<i>budB</i> , acetolactato sintase a partir de <i>Klebsiella pneumoniae</i> ATCC 25955	3	4
<i>alsS</i> , acetolactato sintase a partir de <i>Bacillus subtilis</i>	76	77
<i>budB</i> , acetolactato sintase a partir de <i>Klebsiella terrigena</i>	78	79
<i>budC</i> butanodiol desidrogenase a partir de <i>Klebsiella pneumoniae</i> IAM1063	5	6
Butanodiol desidrogenase a partir de <i>Bacillus cereus</i>	84	85
Butanodiol desidrogenase a partir de <i>Bacillus cereus</i>	86	87
<i>butB</i> , butanodiol desidrogenase a partir de <i>Lactococcus lactis</i>	88	89
<i>pddA</i> , subunidade alfa de butanodiol desidratase a partir de <i>Klebsiella oxytoca</i> ATCC 8724	7	8
<i>pddB</i> , subunidade beta de desidratase de butanodiol a partir de <i>Klebsiella oxytoca</i> ATCC 8724	9	10

Descrição	SEQ ID Ácido nucléico	SEQ ID Proteína
<i>pddC</i> , subunidade gama de desidratase de butanodiol a partir de <i>Klebsiella oxytoca</i> ATCC 8724	11	12
<i>pduC</i> , subunidade grande de diol desidratase dependente de B12a partir de <i>Salmonella typhimurium</i>	92	93
<i>pduD</i> , subunidade média de diol desidratase dependente de B12a partir de <i>Salmonella typhimurium</i>	94	95
<i>pduE</i> , subunidade pequena de diol desidratase dependente de B12a partir de <i>Salmonella typhimurium</i>	96	97
<i>pduC</i> , subunidade grande de diol desidratase dependente de B12a partir de <i>Lactobacillus collinoides</i>	98	99
<i>pduD</i> , subunidade média de diol desidratase dependente de B12a partir de <i>Lactobacillus collinoides</i>	100	101
<i>pduE</i> , subunidade pequena de diol desidratase dependente de B12a partir de <i>Lactobacillus collinoides</i>	102	103
<i>pddC</i> , subunidade alfa de diol desidratase dependente de adenosilcobalamina a partir de <i>Klebsiella pneumoniae</i>	104	105
<i>pddD</i> , subunidade beta de diol desidratase dependente de adenosilcobalamina a partir de <i>Klebsiella pneumoniae</i>	106	107
<i>pddD</i> , subunidade gama de diol desidratase dependente de adenosilcobalamina a partir de <i>Klebsiella pneumoniae</i>	108	109
<i>ddrA</i> , subunidade grande de fator de reativação de diol desidratase a partir de <i>Klebsiella oxytoca</i>	110	111

Descrição	SEQ ID Ácido nucléico	SEQ ID Proteína
<i>ddrB</i> , subunidade pequena de fator de reativação de diol desidratase a partir de <i>Klebsiella oxytoca</i>	112	113
<i>pduG</i> , subunidade grande de fator de reativação de diol desidratase a partir de <i>Salmonella typhimurium</i>	114	115
<i>pduH</i> , subunidade pequena de fator de reativação de diol desidratase a partir de <i>Salmonella typhimurium</i>	116	117
<i>pduG</i> , subunidade grande de fator de reativação de diol desidratase a partir de <i>Lactobacillus collinoides</i>	118	119
<i>pduH</i> , subunidade pequena de fator de reativação de diol desidratase a partir de <i>Lactobacillus collinoides</i>	120	121
<i>sadH</i> , butanol desidrogenase a partir de <i>Rhodococcus ruber</i> 219	13	14
<i>adhA</i> , butanol desidrogenase a partir de <i>Pyrococcus furiosus</i>	90	91
<i>chnA</i> , ciclohexanol desidrogenase a partir de <i>Acinetobacter sp.</i>	71	72
<i>yqhD</i> , butanol desidrogenase a partir de <i>Escherichia coli</i>	74	75
amina:piruvato transaminase a partir de <i>Vibrio fluvialis</i> (uma acetoina aminase)	144 códon opt.	122
aminoálcool quinase a partir de <i>Erwinia carotovora</i> subespécie <i>atroseptica</i>	123	124
aminoálcool O-fosfato liase a partir de <i>Erwinia carotovora</i> subespécie <i>atroseptica</i>	125	126
<i>budC</i> , acetoina redutase (butanodiol desidrogenase) a partir de <i>Klebsiella terrigena</i> (agora <i>Raoultella terrigena</i>)	133	134

Descrição	SEQ ID Ácido nucléico	SEQ ID Proteína
subunidade alfa de glicerol desidratase a partir de <i>Klebsiella pneumoniae</i>	145	146
subunidade beta de glicerol desidratase a partir de <i>Klebsiella pneumoniae</i>	147	148
subunidade gama de glicerol desidratase a partir de <i>Klebsiella pneumoniae</i>	149	150
subunidade grande de glicerol desidratase reativase a partir de <i>Klebsiella pneumoniae</i>	151	152
subunidade pequena de glicerol desidratase reativase a partir de <i>Klebsiella pneumoniae</i>	153	154

SEQ ID nº 15 a 65 são as seqüências de nucleotídeos de oligonucleotídeo PCR, clonagem, triagem e iniciadores de seqüenciamento usados nos exemplos.

5 SEQ ID nº 66 é a seqüência de nucleotídeos da região apagada do gene *yqhD* em *E. coli* cepa MG1655 $\Delta yqhCD$, descrita no exemplo 11.

SEQ ID nº 67 é a seqüência de nucleotídeos de uma variante do promotor de glicose isomerase 1,6GI.

SEQ ID nº 68 é a seqüência de nucleotídeos do promotor 1,5GI.

10 SEQ ID nº 69 é uma seqüência de nucleotídeos do operon de diol desidratase a partir de *Klebsiella oxytoca*.

SEQ ID nº 70 é a seqüência de nucleotídeos do fator de reativação operon de diol desidratase a partir de *Klebsiella oxytoca*.

SEQ ID nº 73 é a seqüência de nucleotídeos de pDCQ2, que é descrita no exemplo 9.

SEQ ID nº 127 a 132 são as seqüências de nucleotídeos de adicional oligonucleotídeo PCR e iniciadores de clonagem usados nos exemplos.

SEQ ID nº 155 é a região de codificação do códon otimizado para a aminoálcool quinase de *Erwinia carotovora subespécie atroseptica*.

5 SEQ ID nº 156 é a região de codificação do códon otimizado para a aminoálcool O-fosfato liase de *Erwinia carotovora subespécie atroseptica*.

SEQ ID nº 157 a 163 são as seqüências de nucleotídeos de adicional oligonucleotídeo PCR e iniciadores de clonagem usados nos exemplos.

10 SEQ ID nº 275 é a seqüência de nucleotídeos de um operon de *Erwinia carotovora subespécie atroseptica*.

TABELA 2.

GRANDES (L), MÉDIAS (M) E PEQUENAS (S) SUBUNIDADES ADICIONAIS DE

GLICEROL E DIOL DESIDRATASE

<u>^a Descrição</u>	<u>^b subunidade</u>	<u>SEQ ID de</u> <u>proteína</u>
Correspondentes subunidades do mesmo organismo ^c		
Subunidade alfa de glicerol desidratase a partir de <i>Clostridium pasteurianum</i>	L	135
Subunidade beta de glicerol desidratase a partir de <i>Clostridium pasteurianum</i>	M	136
Subunidade gama de glicerol desidratase a partir de <i>Clostridium pasteurianum</i>	S	137
Subunidade alfa de glicerol desidratase a partir de <i>Escherichia blattae</i>	L	138
Subunidade beta de glicerol desidratase a partir de <i>Escherichia blattae</i>	M	139
Subunidade gama de glicerol desidratase a partir de <i>Escherichia blattae</i>	S	140

^a Descrição	^b subunidade	SEQ ID de proteína
Subunidade alfa de glicerol desidratase a partir de <i>Citrobacter freundii</i>	L	141
Subunidade beta de glicerol desidratase a partir de <i>Citrobacter freundii</i>	M	142
Subunidade gama de glicerol desidratase a partir de <i>Citrobacter freundii</i>	S	143
diol desidratase subunidade alfa a partir de <i>Lactobacillus brevis</i>	L	164
Diol desidratase subunidade beta a partir de <i>Lactobacillus brevis</i>	M	165
Diol desidratase subunidade gama a partir de <i>Lactobacillus brevis</i>	S	166
diol desidratase subunidade alfa a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> serovar <i>Choleraesuis</i> str. SC-B67	L	167).
diol desidratase, subunidade beta a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> serovar <i>Choleraesuis</i> str. SC-B67	M	168
Diol desidratase subunidade gama a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> serovar <i>Choleraesuis</i> str. SC-B67	S	169
Propanodiol desidratase, subunidade grande a partir de <i>Escherichia coli</i> E24377A	L	170
Diol/Glicerol desidratase média a partir de <i>Escherichia coli</i> E24377A	M	171
Propanodiol desidratase, pequena subunidade a partir de <i>Escherichia coli</i> E24377A	S	172
Diol desidratase subunidade grande a partir de <i>Shigella sonnei</i> Ss046	L	173
Diol desidratase média a partir de <i>Shigella sonnei</i> Ss046	M	174
Diol desidratase pequena subunidade a partir de <i>Shigella sonnei</i> Ss046	S	175
Propanodiol desidratase subunidade grande a partir de <i>Yersinia bercovieri</i> ATCC 43970	L	176

28

^a Descrição	^b subunidade	SEQ ID de proteína
Proteína hipotética YberA_01000484 a partir de <i>Yersinia bercovieri</i> ATCC 43970	M	177
Propanodiol desidratase subunidade pequena a partir de <i>Yersinia bercovieri</i> ATCC 43970	S	178
Propanodiol desidratase subunidade grande a partir de <i>Yersinia mollaretii</i> ATCC 43969	L	179
Proteína hipotética YmolA_01001292 a partir de <i>Yersinia mollaretii</i> ATCC 43969	M	180
Propanodiol desidratase pequena subunidade a partir de <i>Yersinia mollaretii</i> ATCC 43969	S	181
Diol desidratase subunidade grande a partir de <i>Yersinia enterocolitica</i> subespécie <i>enterocolitica</i> 8081	L	182
Diol desidratase subunidade média a partir de <i>Yersinia enterocolitica</i> subespécie <i>enterocolitica</i> 8081	M	183
Diol desidratase subunidade pequena a partir de <i>Yersinia enterocolitica</i> subespécie <i>enterocolitica</i> 8081	S	184
Propanodiol desidratase subunidade grande a partir de <i>Yersinia intermedia</i> ATCC 29909	L	185
Diol/glicerol desidratase média a partir de <i>Yersinia intermedia</i> ATCC 29909	M	186
Propanodiol desidratase subunidade pequena a partir de <i>Yersinia intermedia</i> ATCC 29909	S	187
Glicerol desidratase subunidade grande a partir de <i>Listeria welshimeri</i> serovar 6b str. SLCC5334	L	188

22

^a Descrição	^b subunidade	SEQ ID de proteína
Propanodiol utilização de desidratase subunidade média a partir de <i>Listeria welshimeri</i> sorovar 6b str. SLCC5334	M	189
Utilização de oropropanodiol: desidratase subunidade pequena a partir de <i>Listeria welshimeri</i> serovar 6b str. SLCC5334	S	190
Proteína hipotética lin1117 a partir de <i>Listeria innocua</i> Clip11262	L	191
Proteína hipotética lin1118 a partir de <i>Listeria innocua</i> Clip11262	M	192
Proteína hipotética lin1119 a partir de <i>Listeria innocua</i> Clip11262	S	193
Proteína hipotética lmo1153 a partir de <i>Listeria monocytogenes</i> EGD-e	L	194
Proteína hipotética lmo1154 a partir de <i>Listeria monocytogenes</i> EGD-e	M	195
Proteína hipotética lmo1155 a partir de <i>Listeria monocytogenes</i> EGD-e	S	196
Glicerol desidratase subunidade grande a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> sorovar <i>Typhi</i> str. CT18	L	197
diol desidratase subunidade média a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> sorovar <i>Typhi</i> str. CT18	M	198
diol desidratase subunidade pequena a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> sorovar <i>Typhi</i> str. CT18	S	199
Glicerol desidratase putativo subunidade grande a partir de <i>Escherichia coli</i>	L	200
Glicerol desidratase putativo subunidade média a partir de <i>Escherichia coli</i>	M	201
Glicerol desidratase putativo subunidade pequena a partir de <i>Escherichia coli</i>	S	202
Glicerol desidratase subunidade grande a partir de <i>Listeria monocytogenes</i> str. 4b F2365	L	203

^a Descrição	^b subunidade	SEQ ID de proteína
utilização de propanodiol: desidratase, subunidade média a partir de <i>Listeria monocytogenes</i> str. 4b F2365	M	204
utilização de propanodiol: desidratase, pequena subunidade a partir de <i>Listeria monocytogenes</i> str. 4b F2365	S	205
Glicerol desidratase subunidade grande pduC, putativa a partir de <i>Streptococcus sanguinis</i> SK36	L	206
utilização de propanodiol: Desidratase subunidade média, putativa a partir de <i>Streptococcus sanguinis</i> SK36	M	207
Diol desidratase dependente de B12 pequena subunidade, putativa a partir de <i>Streptococcus sanguinis</i> SK36	S	208
DhaB a partir de <i>Escherichia blattae</i>	L	209
DhaC a partir de <i>Escherichia blattae</i>	M	210
DhaE a partir de <i>Escherichia blattae</i>	S	211
Coenzima dependente de B12 glicerol desidrogenase subunidade grande a partir de <i>Clostridium perfringens</i> str. 13	L	212
Coenzima dependente de B12 glicerol desidrogenase subunidade média a partir de <i>Clostridium perfringens</i> str. 13	M	213
Coenzima dependente de B12 glicerol desidrogenase subunidade pequena a partir de <i>Clostridium perfringens</i> str. 13	S	214
Propanodiol desidratase subunidade grande a partir de <i>Yersinia frederiksenii</i> ATCC 33641	L	215
Proteína hipotética YfreA_01000478 a partir de <i>Yersinia frederiksenii</i> ATCC 33641]	M	216
Propanodiol desidratase, pequena subunidade a partir de <i>Yersinia frederiksenii</i> ATCC 33641	S	217

^a <u>Descrição</u>	^b <u>subunidade</u>	<u>SEQ ID de</u> <u>proteína</u>
Glicerol desidratase a partir de <i>Thermoanaerobacter ethanolicus</i> X514	L	218
Desidratase subunidade média a partir de <i>Thermoanaerobacter ethanolicus</i> X514	M	219
Desidratase subunidade pequena a partir de <i>Thermoanaerobacter ethanolicus</i> X514	S	220
Glicerol desidratase subunidade grande GldC a partir de <i>Lactobacillus hilgardii</i>	L	221
Glicerol desidratase subunidade média GldD a partir de <i>Lactobacillus hilgardii</i>	M	222
Glicerol desidratase pequena subunidade GldE a partir de <i>Lactobacillus hilgardii</i>	S	223
Glicerol desidratase a partir de <i>Lactobacillus reuteri</i> JCM 1112	L	224
similar a diol desidratase subunidade gama a partir de <i>Lactobacillus reuteri</i> JCM 1112	M	225
utilização de propanodiol: desidratase pequena subunidade de <i>Lactobacillus reuteri</i> JCM 1112	S	226
glicerol desidratase subunidade grande GldC a partir de <i>Lactobacillus diolivorans</i>	L	227
glicerol desidratase subunidade média GldD a partir de <i>Lactobacillus diolivorans</i>	M	228
Glicerol desidratase pequena subunidade GldE a partir de <i>Lactobacillus diolivorans</i>	S	229
propanodiol desidratase subunidade grande a partir de <i>Lactobacillus reuteri</i>	L	230

^a Descrição	^b subunidade	SEQ ID de proteína
propanodiol desidratase subunidade média a partir de <i>Lactobacillus reuteri</i>	M	231
propanodiol desidratase pequena subunidade a partir de <i>Lactobacillus reuteri</i>	S	232
glicerol desidratase subunidade grande a partir de <i>Mesorhizobium loti</i> MAFF303099	L+M	233
glicerol desidratase pequena subunidade a partir de <i>Mesorhizobium loti</i> MAFF303099	S	234
glicerol desidratase a partir de <i>Mycobacterium vanbaalenii</i> PYR-1	L+M	235
utilização de propanodiol: desidratase pequena subunidade a partir de <i>Mycobacterium vanbaalenii</i> PYR-1	S	236
glicerol desidratase a partir de <i>Mycobacterium</i> sp. MCS	L+M	237
desidratase pequena subunidade a partir de <i>Mycobacterium</i> sp. MCS	S	238
Desidratase subunidade grande: Desidratase subunidade média a partir de <i>Mycobacterium flavescens</i> PYR-GCK	L+M	239
utilização de propanodiol: desidratase, pequena subunidade a partir de <i>Mycobacterium flavescens</i> PYR-GCK	S	240
glicerol desidratase a partir de <i>Mycobacterium</i> sp. JLS	L+M	241
desidratase pequena subunidade a partir de <i>Mycobacterium</i> sp. JLS	S	242
glicerol desidratase subunidade grande a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	L	243
desidratase subunidade média a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	M	244

^a Descrição	^b subunidade	SEQ ID de proteína
diol desidratase subunidade gama a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	S	245
Subunidades adicionais		
glicerol desidratase subunidade grande a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	L+M	246
glicerol desidratase subunidade grande a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	L+M	247
Coenzima dependente de B12 glicerol desidrogenase subunidade pequena a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	S	248
Coenzima dependente de B12 glicerol desidrogenase subunidade pequena a partir de <i>Mycobacterium smegmatis</i> str. MC2 155	S	249
diol desidratase subunidade média a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> serovar <i>Paratyphi</i> A str. ATCC 9150	M	250
diol desidratase subunidade pequena a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> serovar <i>Paratyphi</i> A str. ATCC 9150	S	251
glicerol desidratase, subunidade beta a partir de <i>Clostridium perfringens</i> SM101	M	252
glicerol desidratase, subunidade gama a partir de <i>Clostridium perfringens</i> SM101	S	253
PduC a partir de <i>Salmonella enterica</i> subespécie <i>enterica</i> serovar <i>Typhimurium</i>	L	254
glicerol desidratase subunidade grande a partir de <i>Listeria monocytogenes</i> str. 4b H7858	L	255

34
A

^a Descrição	^b subunidade	SEQ ID de proteína
DhaB a partir de <i>Escherichia blattae</i>	L	256
DhaB a partir de bactérias não cultivadas	L	257
DhaB a partir de bactérias não cultivadas	L	258
glicerol desidratase subunidade grande GldC a partir de <i>Lactobacillus collinoides</i>	L	259
PduD a partir de bactérias não cultivadas	M	260
PduD a partir de bactérias não cultivadas	M	261
DhaC a partir de bactérias não cultivadas	M	262
DhaC a partir de bactérias não cultivadas	M	263
DhaC a partir de bactérias não cultivadas	M	264
Coenzima dependente de B12 glicerol desidratase, subunidade média a partir de <i>Clostridium perfringens</i> ATCC 13124	M	265
desconhecido	M	266
Subunidade beta de glicerol desidratase a partir de <i>Escherichia blattae</i>	M	267
PduE a partir de bactérias não cultivadas	S	268
PduE a partir de bactérias não cultivadas	S	269
desidratase, pequena subunidade a partir de <i>Listeria monocytogenes</i> str. 1/2a F6854	S	270
DhaE a partir de bactérias não cultivadas	S	271
DhaE a partir de bactérias não cultivadas	S	272
DhaE a partir de bactérias não cultivadas	S	273
desidratase pequena subunidade a partir de <i>Listeria monocytogenes</i> FSL N1-017	S	274

- a Descrição: da anotação Genbank da seqüência e pode não ser correto incluir a designação do glicerol ou diol, ou pode não incluir informação da subunidade.
- b Subunidade: identificados por homologia de seqüência com a grande, média ou pequena subunidade da enzima *Klebsiella oxytoca*.
- c Subunidades que são do mesmo organismo são mencionadas juntas e são consideradas a mesma enzima, ou têm números Genbank próximos indicando proximidade no genoma.

DESCRIÇÃO DETALHADA DA INVENÇÃO

A presente invenção refere-se aos métodos para a produção de 2-butanol usando microorganismos recombinantes. A presente invenção atende inúmeras necessidades comerciais e industriais. Butanol é um importante produto químico industrial com uma variedade de aplicações, onde o seu potencial como combustível ou aditivo de combustível é particularmente significativo. Embora apenas um álcool de quatro carbonos, o butanol tem um conteúdo energético semelhante ao da gasolina e pode ser misturado com qualquer combustível fóssil. Butanol é favorecido como combustível ou aditivo de combustível, uma vez que só produz CO₂ e pouco ou nenhum SO_x ou NO_x quando queimado no motor padrão de combustão interna. Adicionalmente butanol é menos corrosivo que etanol, o aditivo preferencial de combustível até agora.

Em adição a seu uso como biocombustível ou aditivo para combustível, butanol tem o potencial de impactar problemas de distribuição de hidrogênio na emergente indústria de células de combustível. As células de combustível hoje são atormentadas pelas questões de segurança associadas ao transporte e distribuição de hidrogênio. Butanol pode ter seu conteúdo de hidrogênio facilmente mudado e pode ser distribuído através de estações de gás já existentes, na pureza exigida para os motores de combustão ou células de combustível em veículos.

Finalmente, a presente invenção produz 2-butanol a partir de fontes de carbono derivadas de plantas, evitando o impacto ambiental negativo associado aos processos petroquímicos padrão para produção de butanol.

A presente invenção também fornece os microorganismos recombina-
5 tes e os métodos para produzir 2-butanona, um intermediário nas vias biossintéticas do 2-butanol aqui apresentadas. 2-butanona, também conhecido como metil etil cetona (MEK), é útil como um solvente em tintas e outros revestimentos. Ele é também usado na indústria de borracha sintética e na produção de cera de parafina.

10 As seguintes definições e abreviações são para ser usadas para a interpretação das reivindicações e no relatório descritivo.

O termo "invenção" ou "presente invenção" para uso na presente invenção é um termo não-limitador e não pretende se referir a uma única modalidade da invenção particular, mas abrange todas as possíveis
15 modalidades conforme descrito no relatório descritivo e nas reivindicações.

O termo "via biossintética de 2-butanol" refere-se a vias enzimáticas para produzir 2-butanol a partir de piruvato.

O termo "via biossintética de 2-butanona" refere-se a vias da enzima para produzir 2-butanona a partir de piruvato.

20 O termo "acetolactato sintase", também conhecido como "ácido acetohidróxi sintase", refere-se a um polipeptídeo (ou polipeptídeos), com uma atividade enzimática que catalisa a conversão de duas moléculas de ácido pirúvico em uma molécula de alfa-acetolactato. Acetolactato sintase, conhecida como EC 2.2.1.6 [anteriormente 4.1.3.18] (*Enzyme Nomenclature* 1992, Academic Press, San Diego) pode ser dependente do cofator tiamina pirofosfato para sua atividade. As enzimas adequadas de acetolactato sintase
25 estão disponíveis junto à inúmeras fontes, por exemplo, *Bacillus subtilis* [GenBank Nº: AAA22222 NCBI (National Center for Biotechnology Information)]

seqüência de aminoácidos (SEQ ID nº 77), L04470 NCBI seqüência de nucleotídeo (SEQ ID nº 76)], *Klebsiella terrigena* [GenBank Nº: AAA25055 (SEQ ID nº 79), L04507 (SEQ ID nº 78)], e *Klebsiella pneumoniae* [GenBank Nº: AAA25079 (SEQ ID nº 4), M73842 (SEQ ID nº 3)].

5 O termo "acetolactato descarboxilase" refere-se a um polipeptídeo (ou polipeptídeos) tendo uma atividade enzimática que catalisa a conversão de alfa-acetolactato em acetoína. As acetolactato descarboxilases são conhecidas como EC 4.1.1.5 e estão disponíveis, por exemplo, a partir de *Bacillus subtilis* [GenBank Nº: AAA22223 (SEQ ID nº 81), L04470 (SEQ ID nº 80)], *Klebsiella terrigena* [GenBank Nº: AAA25054 (SEQ ID nº 83), L04507 (SEQ ID nº 82)] e *Klebsiella pneumoniae* [GenBank Nº: AAU43774 (SEQ ID nº 2), AY722056 (SEQ ID nº 1)].

O termo "acetoína aminase" refere-se a um polipeptídeo (ou polipeptídeos), tendo uma atividade enzimática que catalisa a conversão de acetoína em 3-amino-2-butanol. Acetoína aminase pode utilizar o cofator piridoxal 5'-fosfato ou NADH (nicotinamida adenina dinucleotide reduzida) ou NADPH (nicotinamida adenina dinucleotide fosfato reduzida). O produto resultante pode ter (*R*) ou (*S*) estereoquímica na posição 3. As enzimas dependentes do fosfato piridoxal podem utilizar um aminoácido como alanina ou glutamato como doador de amino. As enzimas dependentes de NADH e NADPH podem utilizar amônia como um segundo substrato. Um exemplo de uma adequada aminase acetoína dependente de NADH, também conhecida como amino álcool desidrogenase, é descrito por Ito et al. (Patente U.S. Nº 6.432.688). Um exemplo de uma aminase acetoína dependente de piridoxal é a amina:piruvato aminotransferase (também chamada amina:piruvato transaminase) e descrita por Shin Kim (*J. Org. Chem.* 67:2848-2853 (2002)).

O termo "butanol desidrogenase" refere-se a um polipeptídeo (ou polipeptídeos) tendo uma atividade enzimática que catalisa a interconversão de

2-butanona e 2-butanol. Butanol desidrogenases são um subconjunto de uma vasta família de álcool desidrogenases. Butanol desidrogenase pode ser dependente de NAD ou NADP. As enzimas dependentes de NAD são conhecidas como EC 1.1.1.1 e estão disponíveis, por exemplo, a partir de
5 *Rhodococcus ruber* [GenBank N°: CAD36475 (SEQ ID n° 14), AJ491307 (SEQ ID n° 13)]. As enzimas dependentes de NADP são conhecidas como EC 1.1.1.2 e estão disponíveis, por exemplo, a partir de *Pyrococcus furiosus* [GenBank N°: AAC25556 (SEQ ID n° 91), AF013169 (SEQ ID n° 90)]. Adicionalmente, uma butanol desidrogenase está disponível a partir de *Escherichia coli* [GenBank
10 N°:NP_417484 (SEQ ID n° 75), NC_000913 (SEQ ID n° 74)] e uma cicloexanol desidrogenase está disponível a partir de *Acinetobacter* sp. [GenBank N°: AAG10026 (SEQ ID n° 72), AF282240 (SEQ ID n° 71)].

O termo "acetoína quinase" refere-se a um polipeptídeo (ou polipeptídeos), tendo uma atividade enzimática que catalisa a conversão de
15 acetoína em fosfato acetoina. Acetoína quinase pode usar ATP (adenosina trifosfato) ou fosfoenolpiruvato como o doador de fosfato na reação. Apesar de não haver relatos de enzimas catalisando essa reação em acetoína, existem enzimas que catalisam a reação análoga no substrato semelhante dihidroxiacetona, por exemplo, enzimas conhecida como CE 2.7.1.29 (Garcia-
20 Alles et al. (2004) *Biochemistry* 43: 13037-13046).

O termo "acetoína fosfato aminase" refere-se a um polipeptídeo (ou polipeptídeos) com uma atividade enzimática que catalisa a conversão de fosfoacetoína em 3-amino-2-butanol O-fosfato. Acetoína fosfato aminase pode utilizar o cofator piridoxal 5'-fosfato, NADH ou NADPH. O produto resultante
25 pode ter (*R*) ou (*S*) estereoquímica na posição 3. As enzimas dependente do fosfato piridoxal podem utilizar um aminoácido como alanina ou glutamato. As enzimas dependentes de NADH e NADPH podem utilizar amônia como um segundo substrato. Apesar de não haver relatos de enzimas catalisando essa

reação em fosfoacetoína, existe uma enzima dependente de piridoxal fosfato que se propõe para realizar a reação análoga no substrato semelhante serinol fosfato (Yasuta et al. (2001) *Appl. Environ. Microbiol.* 67:4999-5009).

O termo "aminobutanol fosfato fosfo-liase", também chamado de "álcool amino O -fosfato liase", refere-se a um polipeptídeo (ou polipeptídeos), tendo uma atividade enzimática que catalisa a conversão de 3-amino-2-butanol O-fosfato em 2 - butanona. Aminobutanol fosfato fosfo-liase pode utilizar o cofator piridoxal 5'-fosfato. Não existem relatos anteriores de enzimas catalisando essa reação em aminobutanol fosfato, embora existam relatos de enzimas que catalisam a reação análoga semelhante sobre o substrato 1-amino-2-propanol fosfato (Jones et al. (1973) *Biochem J.* 134:167-182). A presente invenção descreve um recentemente identificado aminobutanol fosfato fosfo-liase (SEQ ID nº 126) a partir do organismo *Erwinia carotovora*, com a atividade demonstrada em exemplo da presente invenção.

O termo "aminobutanol quinase" refere-se a um polipeptídeo (ou polipeptídeos) tendo uma atividade enzimática que catalisa a conversão de 3-amino-2-butanol em 3-amino-2-butanol O-fosfato. Aminobutanol quinase pode usar ATP como o doador de fosfato. Apesar de não haver relatos de enzimas catalisarem essa reação em 3-amino-2-butanol, há relatos de enzimas que catalisam a reação análoga nos substratos semelhantes etanolamina e 1-amino-2-propanol (Jones et al., *acima*). A presente invenção descreve, no exemplo 14, um amino álcool quinase de *Erwinia carotovora* subespécie *atroseptica* (SEQ ID nº 124). O termo "Butanodiol desidrogenase" também conhecido como "acetoína redutase" refere-se a um polipeptídeo (ou polipeptídeos), tendo uma atividade das enzimas que catalisa a conversão de acetoína em 2,3-Butanodiol. Butanodiol desidrogenases são um subconjunto de uma vasta família de álcoois desidrogenases. Enzimas butanodiol desidrogenase podem ser específicas para a produção da estereoquímica (R) -

ou (S)- no álcool produto. (S)-butanodiol desidrogenases específicas são conhecidas como CE 1.1.1.76 e estão disponíveis, por exemplo, a partir de *Klebsiella pneumoniae* (GenBank n° s: BBA13085 (SEQ ID n° 6), D86412 (SEQ ID n° 5)). Butanodiol desidrogenase (R)-específico são conhecidas como EC 1.1.1.4 e estão disponíveis, por exemplo, a partir de *Bacillus cereus* [GenBank N° NP_830481 (SEQ ID n° 85), NC_004722 (SEQ ID n° 84); AAP07682 (SEQ ID n° 87), AE017000 (SEQ ID n° 86)], e *Lactococcus lactis* [GenBank N° AAK04995 (SEQ ID n° 89), AE006323 (SEQ ID n° 88)].

O termo "Butanodiol desidratase", também conhecido como "diol desidratase" ou "propanodiol desidratase" refere-se a um polipeptídeo (ou polipeptídeos), com uma atividade enzimática que catalisa a conversão de 2,3-butanodiol em 2-butanona. Butanodiol desidratase pode usar o cofator adenosil cobalamina (vitamina B12). Enzimas dependente de adenosil cobalamina são conhecidas como CE 4.2.1.28 e estão disponíveis, por exemplo, a partir de *Klebsiella oxytoca* [GenBank n° s: BAA08099 (subunidade alfa) (SEQ ID n° 8), D45071 (SEQ ID n° 7); BAA08100 (subunidade beta) (SEQ ID n° 10), D45071 (SEQ ID n° 9); e BBA08101 (subunidade gama) (SEQ ID n° 12), D45071 (SEQ ID n° 11) (note que todas as três subunidades são necessárias para atividade)] e *Klebsiella pneumoniae* [GenBank N°: AAC98384 (subunidade alfa) (SEQ ID n° 105), AF102064 (SEQ ID n° 104); GenBank N°: AAC98385 (subunidade beta) (SEQ ID n° 107), AF102064 (SEQ ID n° 106), GenBank N°: AAC98386 (subunidade gama) SEQ ID n° 109), AF102064 (SEQ ID n° 108)]. Outros diol desidratases adequados incluem, mas não se limitam a, diol desidratases dependentes de B12 disponíveis a partir de *Salmonella typhimurium* [GenBank N°: AAB84102 (subunidade grande) (SEQ ID n° 93), AF026270 (SEQ ID n° 92); GenBank N°: AAB84103 (subunidade média) (SEQ ID n° 95), AF026270 (SEQ ID n° 94); GenBank N°: AAB84104 (subunidade pequena) (SEQ ID n° 97), AF026270 (SEQ ID n° 96)] e *Lactobacillus collinoides* [GenBank N°: CAC82541

(subunidade grande) (SEQ ID nº 99), AJ297723 (SEQ ID nº 98); GenBank Nº: CAC82542 (subunidade média) (SEQ ID nº 101); AJ297723 (SEQ ID nº 100); GenBank Nº: CAD01091 (subunidade pequena) (SEQ ID nº 103), AJ297723 (SEQ ID nº 102)] e enzimas a partir de *Lactobacillus brevis* (particularmente cepas CNRZ 734 e CNRZ 735, Speranza et al., *acima*), e seqüência de nucleotídeo que codificam as enzimas correspondentes. Métodos de isolamento de gene de diol desidratase são bem conhecidos na técnica (por exemplo, patente U.S. Nº 5.686.276). Diol desidratases adicionais são mencionadas na tabela 2.

O termo "glicerol desidratase" refere-se a um polipeptídeo (ou polipeptídeos) tendo uma atividade enzimática que catalisa a conversão de glicerol a 3-hidroxi propionaldeído. Glicerol desidratases dependentes de adenosil cobalamina são conhecidas como CE 4.2.1.30. As glicerol desidratases da CE 4.2.1.30 são similares às diol desidratases na seqüência e no fato de terem três subunidades. As glicerol desidratases também podem ser usadas para converter 2,3-butanodiol em 2-butanona. Alguns exemplos de glicerol desidratases de EC 4.2.1.30 incluem as de *Klebsiella pneumoniae* (subunidade alfa, SEQ ID nº 145, região de codificação e SEQ ID nº 146, proteína; subunidade beta, SEQ ID nº 147, região de codificação e SEQ ID nº 148, proteína; e subunidade gama SEQ ID nº 149, região de codificação e SEQ ID nº 150, proteína); a partir de *Clostridium pasteurianum* [GenBank Nº: 3360389 (subunidade alfa, SEQ ID nº 135), 3360390 (subunidade beta, SEQ ID nº 136), e 3360391 (subunidade gama, SEQ ID nº 137)]; a partir de *Escherichia blattae* [GenBank Nº: 60099613 (subunidade alfa, SEQ ID nº 138), 57340191 (subunidade beta, SEQ ID nº 139), e 57340192 (subunidade gama, SEQ ID nº 140)]; e a partir de *Citrobacter freundii* [GenBank Nº: 1169287 (subunidade alfa, SEQ ID nº 141), 1229154 (subunidade beta, SEQ ID nº 142), e 1229155 (subunidade gama, SEQ ID nº 143)]. Note que todas as

três subunidades são necessárias para atividade. Glicerol desidratases adicionais estão listadas na tabela 2.

Diol e glicerol desidratases podem sofrer inativação suicida durante catálise. Uma proteína fator reativante, também aqui referida como "reativase", pode ser usada para reativar as enzimas inativas (Mori et al., *J. Biol. Chem.* 272:32034 (1997)). Preferencialmente, o fator de reativação é obtido a partir da mesma fonte que a diol ou glicerol desidratase usadas. Por exemplo, fatores de reativação de diol desidratase adequados estão disponíveis a partir de *Klebsiella oxytoca* [GenBank Nº: AAC15871 (subunidade grande) (SEQ ID nº 111), AF017781 (SEQ ID nº 110); GenBank Nº: AAC15872 (subunidade pequena) (SEQ ID nº 113), AF017781 (SEQ ID nº 112)]; *Salmonella typhimurium* [GenBank Nº: AAB84105 (subunidade grande) (SEQ ID nº 115), AF026270 (SEQ ID nº 114), GenBank Nº: AAD39008 (subunidade pequena) (SEQ ID nº 117), AF026270 (SEQ ID nº 116)] e *Lactobacillus collinoides* [GenBank Nº: CAD01092 (subunidade grande) (SEQ ID nº 119), AJ297723 (SEQ ID nº 118); GenBank Nº: CAD01093 (pequena subunidade) (SEQ ID nº 121), AJ297723 (SEQ ID nº 120)]. Ambas as pequenas e grandes subunidades são necessárias para atividade. Por exemplo, fatores de reativação de glicerol desidratase adequados estão disponíveis junto à *Klebsiella pneumoniae* (subunidade grande, SEQ ID nº 151, região de codificação e SEQ ID nº 152, proteína; e pequena subunidade, SEQ ID nº 153, região de codificação e SEQ ID nº 154, proteína).

O termo "um anaeróbico facultativo" refere-se a um microrganismo que pode crescer em ambos os ambientes aeróbico e anaeróbico.

O termo "substrato de carbono" ou "substrato de carbono fermentável" refere-se a uma fonte de carbono capaz de ser metabolizada pelos organismos hospedeiros da presente invenção e especialmente fontes de carbono selecionadas do grupo consistindo em monossacarídeos,

oligossacarídeos, polissacarídeos e substratos de um-carbono ou misturas dos mesmos.

O termo "gene" refere-se a um fragmento de ácido nucléico que é capaz de ser expresso como uma proteína específica, opcionalmente, incluindo seqüências reguladoras que precedem (seqüências 5' não codificantes) e seguem (seqüências 3' não codificantes) a seqüência de codificação. "Gene nativo" refere-se a um gene como encontrado na natureza com as suas próprias seqüências reguladoras. "Gene quimérico" refere-se a qualquer gene que não é um gene nativo, compreendendo seqüências reguladoras e de codificação que não são encontradas juntas na natureza. Conseqüentemente, um gene quimérico pode compreender seqüências reguladoras e de seqüências de codificação que são derivadas de fontes diferentes, ou seqüências reguladoras e seqüências de codificação obtidas a partir da mesma fonte, mas organizadas de uma forma diferente do que aquela encontrada na natureza. "Gene endógeno" refere-se a um gene nativo na sua natural localização no genoma de um organismo. Um gene "estrangeiro" ou "heterólogo" refere-se a um gene que normalmente não se encontra no organismo hospedeiro, mas que é introduzida no mesmo por transferência de genes. Os genes estrangeiros podem compreender os genes nativos inseridos em um organismo não-nativo, ou genes quiméricos. Um "transgene" é um gene que foi introduzido no genoma por um procedimento de transformação.

Para uso na presente invenção, um "fragmento isolado de ácido nucléico" ou "molécula isolada de ácido nucléico" ou "construção genética" serão utilizados indiferentemente e significarão um polímero de DNA ou RNA de fita simples- ou dupla, opcionalmente contendo bases de nucleotídeos alteradas, não naturais ou sintéticas. Um fragmento isolado de ácido nucléico, sob a forma de um polímero de DNA, pode compreender um ou mais segmentos do cDNA, DNA genômico ou DNA sintético.

Um fragmento de ácido nucléico é "hibridizável" com outro fragmento de ácido nucléico, como um cDNA, DNA genômico ou molécula RNA, quando um fragmento de ácido nucléico na forma de fita simples pode se anelar com outros fragmentos de ácido nucléico sob as condições apropriadas de temperatura e força iônica da solução. As condições de hibridização e lavagem são bem conhecidas e exemplificadas em Sambrook, J., Fritsch, E.F. e Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), especialmente capítulo 11 e tabela 11.1 neste (integralmente incorporadas neste documento por referência).

As condições de temperatura e de força iônica determinam a "estringência" da hibridização. As condições de estringência podem ser ajustadas para testar por fragmentos moderadamente semelhantes (como seqüências homólogas de organismos aparentados de forma distante), e até por fragmentos muito semelhantes (como os genes que duplicam enzimas funcionais de organismos estreitamente aparentados). As lavagens pós-hibridização determinam as condições de estringência. Um conjunto de condições preferenciais utiliza uma série de lavagens começando com 6X SSC, 0,5% SDS, à temperatura ambiente durante 15 min, em seguida, repetindo com 2X SSC, 0,5% SDS a 45°C por 30 min, e, em seguida, repetindo duas vezes com 0,2X CCD, 0,5% SDS, a 50°C por 30 min. Um conjunto preferencial de condições de estringência utiliza altas temperaturas nas quais as lavagens são idênticas às mostradas acima exceto que a temperatura das duas lavagens finais de 30 minutos em 0,2X CCD, 0,5% SDS, foi aumentada para 60°C. Outro conjunto preferencial de condições altamente estrigentes usa duas lavagens finais no 0,1X SSC, 0,1% SDS a 65°C. Um conjunto adicional de condições de estringência inclui hibridização a 0,1X SSC, 0,1% SDS, 65°C e lavagens com 2X SSC, 0,1% SDS 0,1X seguido pelo CCD, 0,1% SDS, por exemplo.

A hibridização exige que os dois ácidos nucléicos contêm

seqüências complementares, embora dependendo da estringência da hibridização, falta de correlação entre as bases são possíveis. A estringência adequada para hibridização de ácidos nucléicos depende do comprimento dos ácidos nucléicos e do grau de complementação, variáveis bem conhecidas na técnica. Quanto maior for o grau de semelhança ou de homologia entre duas seqüências de nucleotídeos, maior o valor de T_m para híbridos de ácidos nucléicos tendo essas seqüências. A relativa estabilidade (o que corresponde à maior T_m) de hibridizações de ácido nucléico decresce na seguinte ordem: RNA:RNA, DNA:RNA, DNA:DNA. Para os híbridos de mais de 100 nucleotídeos de comprimento, equações para calcular T_m foram derivadas (consulte Sambrook et al., *acima*, 9,50-9,51). Para hibridizações mais curtas com ácidos nucléicos, ou seja, oligonucleotídeos, a posição da falta de correlação se torna mais importante e, o comprimento do oligonucleotídeo determina a sua especificidade (ver Sambrook et al., *acima*, 11,7-11,8). Em uma modalidade o comprimento de um ácido nucléico hibridizável é pelo menos de cerca de 10 nucleotídeos. De preferência, um comprimento mínimo para um ácido nucléico hibridizável é de pelo menos cerca de 15 nucleotídeos; com mais preferência de pelo menos cerca de 20 nucleotídeos; e com a máxima preferência o comprimento é de pelo menos cerca de 30 nucleotídeos. Além disso, o técnico qualificado irá reconhecer que a temperatura e a concentração de sal da solução de lavagem podem ser ajustadas conforme necessário, de acordo com fatores como o comprimento da sonda.

Uma "parte substancial" de um aminoácido ou seqüência de nucleotídeos é a parte que compreende o suficiente da seqüência de aminoácidos de um polipeptídeo ou da seqüência de nucleotídeos de um gene para identificar putativamente que polipeptídeo ou gene, ou por avaliação manual da seqüência por um técnico no assunto, ou pela comparação e identificação seqüencial automática por computador, utilizando algoritmos

como BLAST (Altschul, SF, et al., *J. Mol. Mol. Biol.*, 215:403-410 (1993)). Em geral, uma seqüência de dez ou mais aminoácidos contíguos ou trinta ou mais nucleotídeos é necessária a fim de putativamente identificar uma seqüência de polipeptídeos ou de ácido nucléico como homóloga a uma proteína ou um gene conhecido. Além disso, no que diz respeito às seqüências de nucleotídeos, sondas de oligonucleotídeo de genes específicos compreendendo 20-30 nucleotídeos contíguos podem ser utilizadas nos métodos de identificação dependentes de identificação de gene da seqüência (por exemplo, "Southern hybridization") e de isolamento (por exemplo, hibridização "*in situ*" de colônias bacterianas ou placas bacteriófagos). Adicionalmente, oligonucleotídeos curtos de 12 a 15 bases podem ser utilizados como iniciadores de amplificação por PCR, a fim de obter um determinado fragmento do ácido nucléico compreendendo os iniciadores. Conseqüentemente, uma "parte substancial" de uma seqüência de nucleotídeos compreende bastante da seqüência para identificar especificamente e/ou isolar um fragmento do ácido nucléico compreendendo a seqüência. O presente relatório descritivo ensina a completa seqüência de aminoácidos e nucleotídeos que codificam proteínas fúngicas particulares. O técnico no assunto, tendo o benefício das seqüências como aqui relatadas, pode agora usar a totalidade ou uma parte substancial das seqüências divulgadas, para fins conhecidos pelos versados na técnica. Conseqüentemente, a presente invenção compreende as seqüências completas como relatado na lista de seqüências anexa, bem como porções significativas dessas seqüências como definido acima.

O termo "complementar" é utilizado para descrever a relação entre as bases de nucleotídeos que são capazes de hibridizar uma à outra. Por exemplo, no que diz respeito ao DNA, adenosina é complementar à timina e citosina é complementar à guanina.

Os termos "homologia" e "homólogo" são usados de maneira

intercambiável neste documento. Eles se referem a fragmentos de ácidos nucleicos em que mudanças em uma ou mais bases de nucleotídeos não afetam a habilidade do fragmento de ácido nucleico de mediar expressão gênica ou produzir um determinado fenótipo. Estes termos também se referem às alterações dos fragmentos do ácido nucleico da presente invenção, como a supressão ou inserção de um ou mais nucleotídeos que não alterem substancialmente as propriedades funcionais do fragmento resultante de ácido nucleico em relação ao fragmento inicial sem modificação. É, portanto, entendido, como aqueles técnicos no assunto irão apreciar, que a invenção abrange mais do que as seqüências exemplificadoras específicas.

Além disso, o técnico no assunto reconhece que as seqüências de ácidos nucleicos homólogas abrangidas por essa invenção também são definidas pela sua capacidade de hibridizar sob moderadamente severas condições (por exemplo, 0,5 X SSC, 0,1% SDS, 60°C) com as seqüências exemplificadas aqui, ou para qualquer parte do seqüências de nucleotídeos apresentadas aqui e que são funcionalmente equivalentes a qualquer uma das seqüências de ácidos nucleicos apresentadas neste documento.

O termo "porcentagem de identidade", conforme é conhecido na técnica, é um relacionamento entre duas ou mais seqüências de polipeptídeos ou duas ou mais seqüências de polinucleotídeos, conforme determinado por comparação das seqüências. Na técnica, a "identidade" significa também o grau de parentesco entre as seqüências de polipeptídeos ou polinucleotídeos, conforme o caso, como determinado pela correspondência entre cadeias de tais seqüências. "Identidade" e "similaridade" podem ser facilmente calculados por métodos conhecidos, incluindo, mas não limitados aos descritos em: 1.) *Computational Molecular Biology* (Lesk, A. M., Ed.) Oxford University: NY (1988); 2.) *Biocomputing: Informatics and Genome Projects* (Smith, D. W., Ed.) Academic: NY (1993); 3.) *Computer Analysis of Sequence Data, Part I* (Griffin,

A. M., e Griffin, H. G., Eds.) Humana: NJ (1994); 4.) *Sequence Analysis in Molecular Biology* (von Heinje, G., Ed.) Academic (1987); e 5.) *Sequence Analysis Primer* (Gribskov, M. e Devereux, J., Eds.) Stockton: NY (1991).

Os métodos preferenciais para determinar a identidade são projetados para proporcionar a melhor comparação entre as seqüências testadas. Os métodos para determinar a identidade e a similaridade são codificados em programas de computador disponíveis publicamente. As seqüência de alinhamentos e cálculos de percentagens de identidade podem ser realizados utilizando o programa MegAlign™ da LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Os múltiplos alinhamentos das seqüências são realizados usando-se "método de alinhamento Clustal", que abrange várias variedades do algoritmo incluindo o "método de alinhamento Clustal V" correspondente ao método de alinhamento identificado como Clustal V (descrito por Higgins e Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci.*, 8:189-191 (1992)) e encontrado no programa MegAlign™ do LASERGENE bioinformatics computing suite (DNASTAR Inc.). Para múltiplos alinhamentos, os valores pré-estabelecidos correspondem a GAP PENALTY=10 e GAP LENGTH PENALTY=10. Os parâmetros pré-estabelecidos para alinhamento em pares e cálculos de porcentagem de identidade de seqüências de proteína usando o método Clustal são KTUPLE=1, GAP PENALTY=3, WINDOW=5 e DIAGONALS SAVED=5. Para ácidos nucléicos esses parâmetros são KTUPLE=2, GAP PENALTY=5, WINDOW=4 e DIAGONALS SAVED=4. Após o alinhamento das seqüências usando o programa Clustal V, é possível obter uma "porcentagem de identidade" visualizando a tabela "distância da seqüência" no mesmo programa. Adicionalmente o "método de alinhamento Clustal W" está disponível e corresponde ao método de alinhamento identificado como Clustal W (descrito por Higgins e Sharp, *CABIOS*. 5:151-153

(1989); Higgins, D.G. et al., *Comput. Appl. Biosci.* 8:189-191(1992)) e encontrado no programa MegAlign™ v6.1 de LASERGENE bioinformatics computing suite (DNASTAR Inc.). Os parâmetros pré-estabelecidos para múltiplos alinhamentos (GAP PENALTY=10, GAP LENGTH PENALTY=0,2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0,5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB). Após o alinhamento das seqüências usando o programa Clustal W, é possível obter uma "porcentagem de identidade" visualizando a tabela "distância da seqüência" no mesmo programa.

É bem entendido pelo técnico no assunto que muitos níveis de identidade de seqüência são úteis na identificação de polipeptídeos, a partir de outras espécies, onde tais polipeptídeos têm a mesma ou similar função ou atividade. Exemplos úteis de porcentagens de identidades incluem, mas não estão limitados a: 24%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% ou 95%, ou qualquer inteiro percentual de 24% a 100% podem ser úteis na descrição da presente invenção, como 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99%. Os fragmentos adequados de ácido nucléico não somente tem as homologias acima, mas tipicamente codificam um polipeptídeo tendo pelo menos 50 aminoácidos, de preferência pelo menos 100 aminoácidos, com mais preferência pelo menos 150 aminoácidos, ainda com mais preferência pelo menos 200 aminoácidos e, com a máxima preferência, pelo menos 250 aminoácidos.

O termo "software de análise de seqüências" refere-se a qualquer

algoritmo computadorizado ou programa de software que é útil para a análise de seqüências de nucleotídeos ou de aminoácidos. "Software de análise de seqüências" pode estar disponível comercialmente ou desenvolvido independentemente. Típicos Softwares de análise de seqüências incluem, mas não estão limitados a: 1.) Os softwares integrados de programas GCG (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul et al., *J. Mol. Biol.*, 215:403-410 (1990)); 3.) DNASTAR (DNASTAR, Inc. Madison, WI); 4.) Sequencher (Gene Codes Corporation, Ann Arbor, MI); e 5.) O programa FASTA incorporando o algoritmo Smith-Waterman (W. R. Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Plenum: New York, NY). No contexto deste pedido, será entendido que onde o software de análise seqüencial é utilizado para a análise, os resultados da análise serão baseados nos "valores pré-determinados" do programa referenciado, salvo disposição em contrário. Para uso na presente invenção, "valores pré-determinados" significará qualquer conjunto de valores ou parâmetros que vem com o software quando o mesmo é inicializado pela primeira vez.

Para uso na presente invenção o termo "seqüência de codificação" ou "CDS" refere-se a uma seqüência de DNA que codifica uma específica seqüência de aminoácidos. "Seqüências reguladoras adequadas" referem-se a seqüências de nucleotídeos localizadas a montante (seqüências 5' não codificantes), no interior, ou a jusante (seqüências 3' não codificantes) de uma seqüência de codificação e que influenciam a transcrição, processamento ou estabilidade de RNA, ou tradução da associada seqüência de codificação. As seqüências reguladoras podem incluir promotores, seqüências de iniciação de tradução, introns, seqüências de reconhecimento de poliadenilação, sítio de processamento de RNA, sítio de ligação efetor e estrutura de haste e laço.

O termo "promotor" refere-se a uma seqüência de DNA capaz de controlar a expressão de uma seqüência de codificação ou RNA funcional. Em geral, uma seqüência de codificação está situada a 3' de uma seqüência promotora. Os promotores poderão ser derivados na sua totalidade a partir de um gene nativo, ou seja, composto por vários elementos derivados de diferentes promotores encontrados na natureza, ou mesmo compreende segmentos de DNA sintéticos. Entende-se por aqueles versados na técnica que diferentes promotores podem dirigir a expressão de um gene em diferentes tipos de células ou tecidos, ou em diferentes fases de desenvolvimento, ou em resposta a diferentes condições ambientais ou fisiológicas. Os promotores que causam a um gene ser expresso na maioria dos tipos de células, na maioria das vezes, são comumente denominados "promotores constitutivos". É ainda reconhecido que, uma vez que na maioria dos casos os exatos limites das seqüências reguladoras não foram completamente definidos, os fragmentos de DNA de diferentes comprimentos podem ter idêntica atividade promotora.

O termo "ligada de maneira funcional" refere-se a associação das seqüências de ácido nucléico em um único fragmento de ácidos nucléicos de modo que a função de um é afetada pelo outro. Por exemplo, um promotor está ligado de maneira funcional com uma seqüência de codificação quando é capaz de viabilizar a expressão da seqüência de codificação (ou seja, a seqüência de codificação está sob o controle transcricional do promotor). As seqüências de codificação podem ser ligados de maneira funcional a seqüências reguladoras na orientação senso ou anti-senso.

O termo "expressão", conforme usado neste documento, refere-se à transcrição e acúmulo estável de RNA senso (mRNA) ou anti-senso derivados do fragmento de ácido nucléico da invenção. Expressão pode também referir-se a tradução de mRNA em um polipeptídeo.

Conforme utilizado aqui o termo "transformação" refere-se à

transferência de um fragmento de ácido nucléico em um organismo hospedeiro, resultando em herança geneticamente estável. Os organismos hospedeiros que contêm os fragmentos de ácido nucléico transformados são referidos como organismos "transgênicos", "recombinantes" ou "transformados".

5 Os termos "plasmídeo" e "vetor" referem-se a um elemento extra cromossômico muitas vezes transportando genes que não fazem parte do metabolismo central da célula e, geralmente, sob a forma de duplas fitas circulares de fragmentos de DNA. Esses elementos podem ser seqüências livremente replicadas, seqüências integradoras de genoma, seqüências de nucleotídeos ou
10 de fagos, circulares ou lineares, de fita d e DNA ou RNA simples ou dupla, provenientes de qualquer fonte, na qual uma série de seqüências de nucleotídeos foram unidas ou recombinadas em uma única construção que é capaz de introduzir um fragmento de promotor e seqüência de DNA para um produto de um gene selecionado juntamente com uma seqüência 3' adequada não-traduzida em
15 uma célula. "Vetor de transformação" refere-se a um vetor específico contendo um gene estrangeiro e tendo elementos em adição ao gene estrangeiro que facilitam a transformação de uma determinada célula hospedeira.

Para uso na presente invenção o termo "degenerescência do códon" refere-se à natureza no código genético que permite a variação da
20 seqüência de nucleotídeos sem afetar a seqüência de aminoácidos de um polipeptídeo codificado. O técnico no assunto está bem consciente da "preferência do códon (codon-bias)" exibido por uma determinada célula hospedeira no uso de códons de nucleotídeos para especificar um determinado aminoácido. Por isso, quando faz a síntese de um gene para a melhoria da
25 expressão gênica em uma célula hospedeira, é desejável que o design do gene seja tal que a freqüência de uso do códon seja próxima da freqüência preferencial de uso do códon da célula hospedeira.

O termo "códon otimizado" ao se referir a genes ou regiões de

codificação de moléculas de ácidos nucléicos para transformação de vários hospedeiros, refere-se à alteração de códons no gene ou nas regiões de codificação das moléculas de ácido nucléico para refletir o uso típico do códon do organismo hospedeiro, sem alterar o polipeptídeo codificado pelo DNA.

5 O termo "meio de fermentação de produto" refere-se a um meio no qual a fermentação ocorreu de tal forma que o produto é presente no meio.

As técnicas padrão de DNA recombinante e clonagem molecular utilizadas aqui são bem conhecidas na técnica e são descritas por Sambrook, J., Fritsch, E. F. F. e Maniatis, T. *Molecular Cloning: A Laboratory Manual*,
10 Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) (mais adiante neste documento "Maniatis"); e por Silhavy, T. J., Bennan, M. L. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984) e por Ausubel, F. et al., *Current Protocols in Molecular Biology*, publicado por Greene Publishing
15 Assoc. e Wiley-Interscience (1987).

A VIA BIOSINTÉTICA DE 2-BUTANONA E 2-BUTANOL

Os carboidratos utilizando microorganismos empregam a via Embden-Meyerhof-Parnas (EMP), a via Entner-Doudoroff e o ciclo fosfato pentose como as vias centrais metabólicas para fornecer energia e precursores
20 celulares para o crescimento e manutenção. Estas vias têm em comum o intermediário gliceraldeído-3-fosfato e, finalmente, piruvato é formado diretamente ou em combinação com a via EMP. As reações combinados de conversão de açúcar a piruvato para produzir energia (por exemplo, adenosina 5'-trifosfato, ATP) e reduzindo equivalentes (por exemplo, reduzida
25 nicotinamida adenina dinucleotídeo, NADH [nicotinamide adenine dinucleotide], e reduzida nicotinamida adenina dinucleotídeo fosfato, NADPH [nicotinamide adenine dinucleotide phosphate]). NADH e NADPH precisam ser reciclados para suas formas oxidadas (NAD^+ e NADP^+ , respectivamente). Na presença de

elétrons inorgânicos receptores (por exemplo, O_2 , NO_3^- e SO_4^{2-}), os equivalentes reduzidos podem ser usados para aumentar a concentração de energia;

A invenção permite a produção de 2-butanona ou 2-butanol a partir de fontes de carboidratos com microrganismos recombinantes, fornecendo uma via biossintética completa a partir de piruvato até 2-butanona ou 2-butanol. Três vias adicionais são descritas. Embora 2-butanol, pelo que se sabe, não é o principal produto de qualquer fermentação bacteriana, há um número de possíveis vias para a produção de 2-butanol através de conhecidos tipos de reações bioquímicas. Estas vias são mostradas na figura 1. As letras e algarismos romanos citados abaixo correspondem às letras e algarismos romanos na figura 1, que são usados para descrever as etapas de conversão e produtos, respectivamente. Conforme descrito abaixo, 2-butanona é um intermediário em todas estas vias biossintética de 2-butanol.

Todas as vias iniciam-se com a reação de duas moléculas de piruvato para produzir alfa-acetolactato (I), mostrado como o substrato para o produto da conversão (a) na figura 1. A partir de alfa-acetolactato, existem 4 possíveis vias para 2-butanona (V), aqui referidas como vias biossintética de 2-butanona:

via 1) I--->II--->III--->IV--->V (substrato para conversões de produto b,c,d,e);

2) I--->II--->VII--->IV--->V (substrato para conversões de produto b,g,h,e)

3) I--->II--->VIII--->V (substrato para conversões de produto b,i,j):

Esta é a via da presente invenção.

4) I--->IX--->X--->V (substrato para conversões de produto k,l,m)

A via biossintética de 2-butanol conclui-se com a conversão de 2-butanona (V) em 2-butanol (VI). Uma discussão detalhada do substrato para

conversões de produto em cada via é dada abaixo.

VIA 1:

(A) PIRUVATO A ALFA-ACETOLACTATO:

O primeiro passo na via 1 é a conversão de duas moléculas de
5 piruvato para uma molécula de alfa-acetolactato (composto I na figura 1) e uma
molécula de dióxido de carbono, catalisada por uma enzima dependente de
tiamina pirofosfato. As enzimas catalisando este substrato para conversão do
produto (genericamente chamadas de ou acetolactato sintase ou de ácido
acetohidróxi sintase; EC 2.2.1.6 [comutadas de 4.1.3.18 em 2002]) são bem
10 conhecidos, e elas participam na via biossintética para os aminoácidos
proteínogênicos leucina e valina, bem como na via para produção fermentativa
de 2,3-butanodiol e acetoína de uma série de organismos.

O versado na técnica vai apreciar o fato que os polipeptídeos
tendo atividade da acetolactato sintase isolada de uma variedade de fontes
15 serão muito úteis na presente invenção independente da homologia da
seqüência. Alguns exemplos de enzimas adequadas de acetolactato sintase
estão disponíveis junto a inúmeras fontes, por exemplo, *Bacillus subtilis*
[GenBank Nº: AAA22222 NCBI (National Center for Biotechnology Information)
seqüência de aminoácidos (SEQ ID nº 77), L04470 NCBI seqüência de
20 nucleotídeo (SEQ ID nº 76)], *Klebsiella terrigena* [GenBank Nº: AAA25055
(SEQ ID nº 79), L04507 (SEQ ID nº 78)], e *Klebsiella pneumoniae* [GenBank
Nº: AAA25079 (SEQ ID nº 4), M73842 (SEQ ID nº 3)]. As enzimas preferenciais
para acetolactato sintase são aquelas que têm pelo menos 80% a 85% de
identidade com SEQ ID Nº 4, 77 e 79, onde pelo menos 85% a 90% de
25 identidade é mais preferível e onde pelo menos 95% de identidade baseada no
método de alinhamento Clustal W usando os parâmetros pré-estabelecidos
GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso
de proteína Gonnet 250, é o preferencial.

(B) ALFA-ACETOLACTATO PARA ACETOÍNA:

Alfa-acetolactato (I) é convertido para acetoína (II) pela ação de uma enzima como acetolactato descarboxilase (EC 4.1.1.5). Como a acetolactato sintase, esta enzima é dependente de tiamina pirofosfato e é, também, parte na produção de 2,3-butanodiol e acetoína por inúmeros organismos. As enzimas de fontes diferentes variam muito em tamanho (25 kDa a 50 kDa), oligomerização (dímero-hexâmero), localização (intracelular ou extracelular), e regulação alostérica (por exemplo, ativação por cadeia ramificada de aminoácidos). Para o propósito da presente invenção, uma localização intracelular é preferível a extracelular, mas outras variações são geralmente aceitáveis.

O técnico no assunto vai apreciar o fato que os polipeptídeos tendo atividade de acetolactato descarboxilase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da seqüência. Alguns exemplos de enzimas adequadas de acetolactato descarboxilase estão disponíveis junto a inúmeras fontes, por exemplo, *Bacillus subtilis* [GenBank N°: AAA22223 (SEQ ID n° 81), L04470 (SEQ ID n° 80)], *Klebsiella terrigena* [GenBank N°: AAA25054 (SEQ ID n° 83), L04507 (SEQ ID n° 82)] e *Klebsiella pneumoniae* [GenBank N°: AAU43774 (SEQ ID n° 2), AY722056 (SEQ ID n° 1)].

As enzimas preferenciais para acetolactato descarboxilase são aquelas que têm pelo menos 80% a 85% de identidade com SEQ ID N° 2, 81 e 83 onde pelo menos 85% a 90% de identidade é preferível e onde pelo menos 95% de identidade baseada no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250, é preferencial.

(C) ACETOÍNA A 3-AMINO-2-BUTANOL:

Existem dois tipos de reações bioquímicas conhecidas que poderiam afetar o substrato para conversão do produto de acetoína (II) a 3-

amino-2-butanol (III), especificamente, transaminação dependente de piridoxal fosfato utilizando um doador acessório de amino e aminação redutora direta com amônia. Neste último caso, os equivalentes reduzidos são oferecidos na forma de um cofator reduzido de nicotinamida (NADH ou NADPH). Um exemplo de uma enzima dependente de NADH que catalisa essa reação com acetoína como um substrato é relatado por Ito et al. (Patente US Nº 6.432.688). Qualquer estereoespecificidade desta enzima não foi avaliada. Um exemplo de um piridoxal dependente de fosfato transaminase que catalisa a conversão de acetoína em 3-amino-2-butanol tem sido relatado por Shin e Kim (*acima*). Esta enzima foi aqui mostrada no exemplo 13 para converter tanto o isômero (*R*) de acetoína para o isômero (*2R,3S*) de 3-amino-2-butanol e o isômero (*S*) de acetoína para o isômero (*2S,3S*) de 3-amino-2-butanol. Qualquer tipo de enzima (isto é, transaminase ou aminase reduzida) é considerado para ser uma acetoína aminase e pode ser utilizado na produção de 2-butanol. Outras enzimas neste grupo podem ter diferentes estereoespecificidades.

O versado na técnica vai apreciar o fato que os polipeptídeos tendo atividade da acetoína aminase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da seqüência. Um exemplo desta atividade foi aqui descrito e é identificado como SEQ ID nº 122.

Conseqüentemente, as enzimas preferenciais para acetoína aminase são aquelas que têm pelo menos 80% a 85% de identidade com SEQ ID Nº 2, e onde pelo menos 85% a 90% de identidade é mais preferível e onde pelo menos 95% de identidade baseada no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250, é de preferência máxima.

(D) 3-AMINO-2-BUTANOL A 3-AMINO-2-BUTANOL O-FOSFATO:

Não existem enzimas conhecidas na técnica que catalisam o substrato para conversão do produto de 3-amino-2-butanol (III) a fosfato de 3-

amino-2-butanol (IV). No entanto, foi demonstrado que algumas espécies *Pseudomonas* e *Erwinia* expressam etanolamina quinase dependente de ATP (EC 2.7.1.82) que lhes permitem utilizar etanolamina ou 1-amino-2-propanol como uma fonte de nitrogênio (Jones et al. (1973) *Biochem. J.* 134:167-182). É provável que esta enzima também tenha atividade para 3-amino-2-butanol ou poderia ser projetada para fazê-lo, proporcionando assim uma aminobutanol quinase. A presente invenção descreve, no exemplo 14, um gene de *Erwinia carotovora* subespécie *atroseptica* (SEQ ID nº 123) que codifica uma proteína (SEQ ID nº 24) que é identificada como uma amino álcool quinase. Esta enzima pode ser usada para converter 3-amino-2-butanol em 3-amino-2-butanol O-fosfato.

O versado na técnica vai apreciar o fato que os polipeptídeos tendo atividade de aminobutanol quinase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da sequência. Um exemplo desta atividade foi aqui descrito e é identificado como SEQ ID nº 124. Conseqüentemente, as enzimas preferenciais para aminobutanol quinase são aquelas que têm pelo menos 80% a 85% de identidade com SEQ ID nº 124, e onde pelo menos 85% a 90% de identidade é preferível e onde pelo menos 95% de identidade baseada no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250, é de máxima preferência.

(E) FOSFATO DE 3-AMINO-2-BUTANOL A 2-BUTANONA:

Embora não existam enzimas conhecidas para catalisar o substrato para produzir a conversão do 3-amino-2-butanol fosfato (IV) para 2-butanona (V), o substrato é muito semelhantes aos utilizados pela enzima piridoxal dependente de fosfato fosfoetanolamina fosfo liase, que tem sido encontrada em um pequeno número de espécies de *Pseudomonas* e *Erwinia*. Estas enzimas têm atividade dirigida a fosfoetanolamina e a ambos os enantiômeros de fosfo-1-2-

aminopropano (1973) *Biochem. J.* 134:167-182), e pode, também ter atividade dirigida a 3-amino-2-butanol O-fosfato. A presente invenção descreve um gene de *Erwinia carotovora* subespécie *atroseptica* (SEQ ID NO:125) que codifica uma proteína (SEQ ID NO:126) com homologia para aminotransferases classe III. O exemplo 15 demonstra que esta enzima é ativa em ambos substratos de fosfato de aminopropanol e fosfato de aminobutanol. A enzima recém-identificada e caracterizada foi capaz de catalisar a conversão de uma mistura de (R)-3-amino-(S)-2-butanol e (S)-3-amino-(R)-2-butanol O-fosfato, e uma mistura de (R)-3-amino-(R)-2-butanol e (S)-3-amino-(S)-2-butanol O-fosfato a 2-butanona. A enzima recém-identificada e caracterizada foi também capaz de catalisar a conversão de ambos (R) e (S)-2-amino-1-propanol fosfato a propanona, com uma preferência por (S)-2-amino-1-propanol fosfato. A maior atividade foi observada com o substrato natural proposto DL-1-amino-2-propanol fosfato, que foi convertido para propionaldeído.

O técnico no assunto vai apreciar o fato que os polipeptídeos tendo atividade de aminobutanol fosfato fosfo-liase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da seqüência. Um exemplo de uma enzima aminobutanol fosfato fosfo-liase adequada é descrita na presente invenção como SEQ ID nº 126. Conseqüentemente, as enzimas preferenciais para aminobutanol fosfato fosfo-liase são aquelas que têm pelo menos 80% a 85% de identidade com SEQ ID nº 126, e onde pelo menos 85% a 90% de identidade é preferível e onde pelo menos 95% de identidade baseada no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250, é de preferência máxima.

(F) 2-BUTANONA A 2-BUTANOL:

A etapa final em todas as vias para produzir 2 - butanol a partir de ácido pirúvico é a redução de 2-butanona (V) a 2-butanol (VI). Este

substrato para conversão do produto é catalisado por alguns membros da classe ampla de álcoois desidrogenases (formas utilizando tanto NADH ou NADPH como fonte de hidreto, dependendo da enzima) que podem ser chamadas de butanol desidrogenases. As enzimas de cada tipo que catalisam a redução de 2-butanona são bem conhecidas, como descrito acima, na definição de butanol desidrogenase.

O técnico no assunto vai apreciar o fato que os polipeptídeos tendo atividade de butanol desidrogenase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da seqüência.

Alguns exemplos de enzimas adequadas de butanol desidrogenase estão disponíveis junto à inúmeras fontes, por exemplo, *Rhodococcus ruber* [GenBank N°: CAD36475 (SEQ ID n° 14), AJ491307 (SEQ ID n° 13)]. As enzimas dependentes de NADP são conhecidas como EC 1.1.1.2 e estão disponíveis, por exemplo, a partir de *Pyrococcus furiosus* [GenBank N°: AAC25556 (SEQ ID n° 91), AF013169 (SEQ ID n° 90)]. Adicionalmente, uma butanol desidrogenase está disponível a partir de *Escherichia coli* [GenBank N°: NP_417484 (SEQ ID n° 75), NC_000913 (SEQ ID n° 74)] e uma cicloexanol desidrogenase está disponível a partir de *Acinetobacter* sp. [GenBank N°: AAG10026 (SEQ ID n° 72), AF282240 (SEQ ID n° 71)]. As enzimas preferenciais para butanol desidrogenase são aquelas que têm pelo menos 80% a 85% de identidade com SEQ ID N° 14, 91, 75 e 72, onde pelo menos 85% a 90% de identidade é preferível e onde pelo menos 95% de identidade baseada no método de alinhamento Clustal 72 W usando os parâmetros pré-estabelecidos GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250, é de máxima preferência.

VIA 2:

(A) PIRUVATO A ALFA-ACETOLACTATO:

Este substrato para conversão do produto é o mesmo descrito

acima para via 1.

(B) ALFA-ACETOLACTATO PARA ACETOÍNA:

Este substrato para conversão do produto é o mesmo descrito acima para via 1.

5

(G) ACETOÍNA A FOSFOACETOÍNA:

Embora as enzimas que catalisam o substrato à conversão do produto de acetoína (II) para fosfoacetoína (VII) não foram descritas, a estrutura do substrato acetoína é muito semelhante a da diidróxi acetona, e, portanto, acetoína pode ser um substrato aceitável para diidróxi acetona quinase (CE 2.7.1.29), uma enzima que catalisa fosforilação de diidróxi acetona. Técnicas de engenharia de proteína para a alteração de especificidade do substrato das enzimas são bem conhecidas (Antikainen e Martin (2005) *Bioorg. Med. Chem.* 13:2701-2716) e podem ser usadas para gerar uma enzima com a especificidade necessária. Nesta conversão, a porção de fosfato pode ser fornecida por qualquer doador de fosfato de alta energia biológica, com o substrato comum sendo fosfoenolpiruvato (como na *E. coli* diidroacetona quinase) e ATP (como na *Citrobacter freundii* diidroacetona quinase) (Garcia-Alles et al. (2004) *Biochemistry* 43:13037-13045).

10

15

(H) FOSFOACETOÍNA A 3-AMINO-2-BUTANOL O-FOSFATO:

20

Embora as enzimas que catalisam o substrato à conversão do produto de fosfoacetoína (VII) para 3-amino-2-butanol O-fosfato (IV) não foram descritas, a estrutura do substrato é muito semelhante a do fosfato de diidróxi acetona um substrato para o proposto serinol fosfato aminotransferase codificadas através da porção 5 ' do gene da *rtxA* em algumas espécies de *Bradyrhizobium* (Yasuta et al., *acima*). Assim, uma serinol fosfato aminotransferase pode ser funcional nesta etapa.

25

(E) 3-AMINO-2-BUTANOL O-FOSFATO A 2-BUTANONA:

Este substrato para conversão do produto é o mesmo descrito

acima para via 1.

(F) 2-BUTANONA A 2-BUTANOL:

Este substrato para conversão do produto é o mesmo descrito acima para via 1.

5

VIA 3:

(A) PIRUVATO A ALFA-ACETOLACTATO:

Este substrato para conversão do produto é o mesmo descrito acima para via 1.

(B) ALFA-ACETOLACTATO PARA ACETOÍNA:

10

Este substrato para conversão do produto é o mesmo descrito acima para via 1.

(I) ACETOÍNA A 2,3-BUTANODIOL:

15

O substrato para conversão de produto de acetoína (II) de 2,3-butanodiol (VIII) pode ser catalisado por uma butanodiol desidrogenase, que tanto podem utilizar NADH ou NADPH como a fonte de equivalentes reduzidos durante a realização de reduções. Enzimas com atividade dirigida a acetoína participam na via para a produção de 2,3-butanodiol em organismos que produzem esse composto. As enzimas relatadas aqui (por exemplo, BudC de *Klebsiella pneumoniae* (Ui et al. (2004) *Letters in Applied Microbiology* 39:533-537) geralmente utilizam NADH. Qualquer cofator é aceitável para utilização na produção de 2-butanol por esta via.

20

O versado na técnica vai apreciar o fato que os polipeptídeos tendo atividade de butanodiol desidrogenase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da sequência. Alguns exemplos de enzimas adequadas de butanodiol desidrogenase estão disponíveis junto à inúmeras fontes, por exemplo, *Klebsiella pneumoniae* [GenBank Nº: BBA13085 (SEQ ID nº 6), D86412 (SEQ ID nº 5)]. As butanodiol desidrogenases (*R*)-específicas são conhecidas como EC 1.1.1.4 e estão

25

disponíveis, por exemplo, a partir de *Bacillus cereus* [GenBank Nº NP_830481 (SEQ ID nº 85), NC_004722 (SEQ ID nº 84); AAP07682 (SEQ ID nº 87), AE017000 (SEQ ID nº 86)] e *Lactococcus lactis* [GenBank Nº AAK04995 (SEQ ID nº 89), AE006323 (SEQ ID nº 88)]. As enzimas preferenciais de butanodiol

5 desidrogenase são aquelas que têm pelo menos 80% a 85% identidade com SEQ ID Nº 6, 85, 87 e 89, onde pelo menos 85% - 90% identidade é mais preferencial e onde pelo menos 95% de identidade com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP

● PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de

10 proteína Gonnet 250, é de máxima preferência.

(J) 2,3-BUTANODIOL A 2-BUTANONA:

O substrato para conversão de produto de 2,3-butanodiol (VIII) em 2-butanona (V) pode ser catalisada por enzimas desidratase diol (CE 4.2.1.28) e enzimas glicerol desidratase (EC 4.2.1.30). A melhor caracterizada diol

15 desidratase é a coenzima dependente de B12 da enzima *Klebsiella oxytoca*, mas enzimas similares são encontradas em um número de bactérias entéricas Foi mostrado que a enzima *K. oxytoca* aceita *meso*-2,3-butanodiol como um substrato (Bachovchin *et al.* (1977) *Biochemistry* 16:1082-1092), produzindo o

produto desejado 2-butanona. O exemplo 17 mostra que a glicerol desidratase

20 *Klebsiella pneumoniae* foi capaz de converter *meso*-2,3-butanodiol em 2-butanona. As três subunidades da glicerol desidratase de *Klebsiella pneumoniae* (alfa: SEQ ID nº 145 (região de codificação) e 146 (proteína); beta: SEQ ID nº 147 (região de codificação) e 148 (proteína); e gama: SEQ ID nº 149 (região de codificação) e 150 (proteína)) foram expressos em conjunto com as

25 duas subunidades da *Klebsiella pneumoniae* glicerol desidratase reativase, SEQ ID nº 151 (região de codificação) e 152 (proteína); e pequena subunidade, SEQ ID nº 153 (região de codificação) e 154 (proteína)) para fornecer atividade.

Existem também relatos na literatura de um independente B12

diol desidratase a partir de *Clostridium glycolicum* (Hartmanis et al. (1986) *Arch. Biochem. Biophys.* 245:144-152). Esta enzima tem atividade para 2,3-butanodiol, embora esta atividade é inferior a 1% da atividade dirigida a etanodiol, mas a enzima pode ser projetada para melhorar essa atividade. A
5 melhor caracterizada desidratase independente de B12 é o glicerol desidratase de *Clostridium butyricum* (O'Brien et al. (2004) *Biochemistry* 43:4635-4645), que tem uma alta atividade para 1,2-propanodiol, bem como para glicerol. Esta enzima usa S-adenosil metionina como uma fonte de radical adenosil. Não há
relatos de atividade para 2,3-butanodiol, mas essa atividade, se já não estiver
10 presente, pode eventualmente ser manipulada.

O versado na técnica vai apreciar o fato que os polipeptídeos tendo atividade de butanodiol desidrogenase isolada de uma variedade de fontes serão muito úteis na presente invenção independente da homologia da sequência. Conforme observado acima uma variedade de diol e glicerol
15 desidratases tem sido descrita na literatura e será adequada para utilização na presente invenção. Conseqüentemente, em um aspecto preferencial da invenção enzimas de diol e glicerol desidratase são aquelas que tem pelo menos 80% a 85% de identidade com enzimas tendo as subunidades grande, média e pequena, respectivamente, das seqüências mencionadas abaixo:

- 20 a) SEQ ID nº 8, SEQ ID nº 10 e SEQ ID nº 12;
- b) SEQ ID nº 93, SEQ ID nº 95 e SEQ ID nº 97;
- c) SEQ ID nº 99, SEQ ID nº 101, e SEQ ID nº 103;
- d) SEQ ID nº 105, SEQ ID nº 107 e SEQ ID nº 109;
- e) SEQ ID nº 135, SEQ ID nº 136 e SEQ ID nº 137;
- 25 f) SEQ ID nº 138, SEQ ID nº 139 e SEQ ID nº 140;
- g) SEQ ID nº 146, SEQ ID nº 148 e SEQ ID nº 150;
- h) SEQ ID nº 141, SEQ ID nº 142 e SEQ ID nº 143; e
- i) SEQ ID nº 164, SEQ ID nº 165 e SEQ ID nº 166.

onde pelo menos 85% a 90% de identidade é mais preferível e onde pelo menos 95% de identidade baseada no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250, é de máxima preferência. Semelhantemente, enzimas de diol e glicerol desidratase preferenciais são aquelas que tem pelo menos 80% a 85% de identidade com enzimas tendo as subunidades grande, média e pequena, respectivamente, das seqüências mencionadas abaixo:

● Subunidade grande: SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164;

10 Subunidade média: SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165;

pequena subunidade: SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166; onde pelo menos 85% a 90% de identidade é mais preferencial e onde pelo menos 95% de identidade com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1, e séries de matriz de peso de proteína Gonnet 250, é da máxima preferência.

As enzimas adicionais de diol e glicerol desidratase que podem ser usadas na via biossintética 3 da presente invenção foram identificado através da estrutura bioinformática/função análise que é descrita a seguir e no exemplo 18.

(F) 2-BUTANONA A 2-BUTANOL:

Este substrato para conversão do produto é o mesmo descrito acima para via 1.

DIOL E GLICEROL DESIDRATASE PARA VIA BIOSSINTÉTICA 3

Qualquer enzima que é uma diol ou glicerol desidratases pode ser usada na presente invenção para a conversão de 2,3-butanodiol para 2-butanona. Uma relação de estrutura/função para diol e glicerol desidratases nas classes de enzima EC 4.2.1.28 e EC 4.2.1.30, respectivamente, foi estabelecida na presente invenção no exemplo 18. A função é fornecida através de dados experimentais e a

estrutura é fornecida pela análise bioinformática. Oito enzimas de diol e glicerol desidratase com atividades que têm sido experimentalmente demonstradas foram analisadas. Deste grupo (mencionados na tabela 10), foi mostrado que as enzimas diol desidratase de *Klebsiella oxytoca* e a glicerol desidratase de *Klebsiella pneumoniae* convertem 2,3-butanodiol para 2-butanona (Bachovchin et al. (1977) *Biochemistry* 16:1082-1092 e exemplo 17 da presente invenção, respectivamente), enquanto as atividades das seis enzimas adicionais foram demonstradas usando seus substratos naturais (referências dadas na tabela 10).

Este conjunto de oito diol e glicerol desidratases foi analisado usando o algoritmo de hmmsearch do pacote de software HMMER (Janelia Farm Research Campus, Ashburn, VA). O parâmetro Z do algoritmo de hmmsearch foi ajustado para 1 bilhão. A saída da análise HMMER usando um conjunto de seqüências de proteína é um perfil Hidden Markov modelo (perfil HMM). A teoria por trás dos perfis HMMs é descrita em Durbin et al., *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, Cambridge University Press, 1998; Krogh et al., 1994; *J. Mol. Biol.* 235:1501–1531, aqui incorporado, a título de referência) que caracteriza um conjunto de proteínas com base na probabilidade de que cada aminoácido ocorrer em cada posição no alinhamento das proteínas do conjunto.

Desde que cada uma das oito diol e glicerol desidratases (diol/glicerol desidratase) com função experimentalmente verificada que foi usada para a análise tem três subunidades (grande ou alfa, média ou beta e pequena ou gama), um perfil separado HMM foi preparado para cada subunidade. O perfil da subunidade grande HMM (tabela 12) foi construído usando proteínas com SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164 que são descritas nas tabelas 1 e 2. O perfil da subunidade média HMM (tabela 13) foi construído usando proteínas com SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165 que são descritas nas tabelas 1 e 2. O perfil da subunidade pequena HMM (tabela 13) foi construído usando proteínas com SEQ ID nº 12, 103, 109, 137,

140, 143, 150 e 166 que são descritas nas tabelas 1 e 2. As referências que fornecem os dados do ensaio funcional são apresentadas na tabela 10. O perfil HMM preparado para a subunidade grande dá uma caracterização estrutural para a subunidade grande funcional da diol/glicerol desidratases. Similarmente os perfis HMMs para as subunidades médias e pequenas dão caracterizações estruturais para as subunidades funcionais médias e pequenas, respectivamente, de diol/glicerol desidratases. Assim, qualquer proteína que tem uma compatibilidade significativa quer para o perfil HMMas de uma subunidade grande, média ou pequena está diretamente ligado à função da subunidade para qual o perfil foi preparado. Para ser significativa, a compatibilidade tem um E-valor de 0,01 ou menos, e maior uso de "compatibilidade" se entende que é com este critério de E-valor. Assim subunidades diol/glicerol desidratases que podem ser usadas na presente invenção são proteínas que são compatíveis com os perfis HMMs, que foram preparadas usando as proteínas com SEQ ID Nº mencionados acima, com um E-valor de 0,01 ou menos.

As proteínas que têm comprimento total e têm vínculo funcional com a subunidade grande da diol/glicerol desidratases, através de correspondência entre o Perfil HMM da subunidade grande, incluem, mas não estão limitados a, proteínas com SEQ ID Nº 93, 167, 170, 173, 176, 179, 182, 185, 188, 191, 194, 197, 200, 203, 206, 209, 212, 215, 218, 221, 224, 227, 130, 243, 254, 255, 256, 257, 258 e 259. As proteínas que têm comprimento total e têm vínculo funcional com a subunidade média da diol/glicerol desidratases, através de correspondência entre o Perfil HMM da subunidade média, incluem, mas não estão limitados a, proteínas com SEQ ID Nº 95, 168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 222, 225, 228, 231, 244, 250, 252, 260, 261, 262, 263, 364, 265, 266 e 267. As proteínas que têm comprimento total e têm vínculo funcional com a subunidade pequena da diol/glicerol desidratases, através de correspondência entre o Perfil HMM da subunidade pequena, incluem, mas

não estão limitados a, proteínas com SEQ ID N° 97, 169, 172, 175, 178, 181, 184, 187, 190, 193, 196, 199, 202, 205, 208, 211, 214, 217, 220, 223, 226, 229, 232, 234, 236, 238, 240, 242, 245, 248, 249, 251, 253, 268, 270, 271, 272, 273 e 274. Além disso, as proteínas das subunidades grandes e médias unidas e que têm comprimento total e vínculo funcional com a subunidades grandes e médias da diol/glicerol desidratases, através de correspondência entre o Perfil HMM da subunidade grande e média, incluem, mas não estão limitados a, proteínas com SEQ ID N° 233, 235, 237, 239, 241, 246 e 247.

Desde que os perfis HMMs descritos acima fornecem uma relação de estrutura/função de diol/glicerol desidratases, proteínas recém-identificadas que correspondem a esses perfis podem também ser utilizadas na presente invenção. Além disso, seqüências de proteínas de subunidades de diol/glicerol desidratases que podem ser usadas na presente invenção incluem proteínas com alterações de aminoácido que têm efeitos mínimos sobre a função da subunidade, e que são substancialmente semelhantes às seqüências das SEQ ID N ° mencionados acima. É bem conhecido na técnica que a substituição de um aminoácido quimicamente equivalente em um determinado local que não afeta as propriedades funcionais das proteínas codificadas é comum. Para os propósitos da presente invenção, as substituições fornecendo proteínas substancialmente similares são definidas como trocas dentro de um dos cinco grupos a seguir:

1. Resíduos alifáticos pequenos, não-polar ou levemente polar: Ala, Ser, Thr (Pro, Gly);

2. Resíduos polar, carregados negativamente e suas amidas: Asp, Asn, Glu, Gln;

3. Resíduos polar, carregados positivamente: His, Arg, Lys;

4. Resíduos alifáticos grandes, não-polar: Met, Leu, Ile, Val (Cys);

e

5. Resíduos aromáticos grandes: Phe, Tyr, Trp.

Assim, substituições de um aminoácido por outro nestes grupos podem ser esperadas para produzir uma proteína funcionalmente equivalente. Em muitos casos, as mudanças que resultam na modificação das porções N e C terminal da proteína também não seriam esperadas para alterar a atividade da proteína.

As proteínas substancialmente semelhantes às das SEQ ID que correspondem aos perfis HMMs podem ser 90% ou 95% idênticas na sequência de aminoácidos de uma das proteínas correspondentes e estas podem ser usadas na presente invenção.

O técnico no assunto pode facilmente identificar um conjunto de três subunidades que podem ser usadas juntas para proporcionar uma diol/glicerol desidratase funcional. Especialmente adequado, é uma combinação de uma subunidade grande, média e pequena da mesma cepa de organismo, cujas regiões de codificação estão localizadas perto uma da outra no genoma. Estas subunidades mais provavelmente formariam uma diol ou glicerol desidratase natural. Muitas subunidades grandes, médias e pequenas, estão agrupadas desta forma na tabela 2. Uma combinação de subunidades de cepas ou espécies intimamente relacionadas é adequada para compor uma diol desidratase ou uma glicerol desidratase. Qualquer combinação de subunidades que catalisam a conversão de 2,3-butanodiol para 2-butanona pode ser usada. As combinações de subunidade efetivas podem prontamente ser determinadas pelo versado na técnica através de comparações de seqüências de aminoácidos e/ou ensaios funcionais.

Conseqüentemente, a invenção fornece enzimas de diol e glicerol desidratase tendo seqüências de aminoácidos compreendendo subunidades grandes, médias e pequenas de comprimento total, cada uma com um parâmetro E-valor de 0,01 ou menos, quando pesquisado usando um perfil Hidden Markov Model preparado através do uso das subunidades grandes das SEQ ID N °: 8, 99,

105, 135, 138, 141, 146 e 164; a subunidade média de SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165; e as subunidades pequenas de SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166; cada pesquisa sendo executada usando o algoritmo hmmsearch sendo que o parâmetro Z está ajustado para um 1 bilhão.

5 Alternativamente, a invenção fornece enzimas de diol e glicerol desidratase tendo seqüências de aminoácidos identificadas pelo processo compreendendo a) gerar um perfil Hidden Markov Model a partir do alinhamento das seqüências de aminoácidos correspondendo às subunidades grandes, médias e pequenas de enzimas de diol e glicerol desidratase, sendo que:

10 i) A subunidade grande compreende uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164;

ii) A subunidade média compreende uma seqüência de aminoácidos selecionado do grupo consistindo em SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165; e

15 iii) A pequena subunidade compreende uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166;

b) pesquisar pelo menos uma base de dados pública de seqüências de proteína contendo seqüências de diol e glicerol desidratases com o perfil "Hidden Markov Model" de (a) usando o algoritmo hmmsearch sendo que o parâmetro Z está ajustado para um 1 bilhão e o parâmetro E-valor está ajustado para 0,01, para identificar um primeiro conjunto de dados de seqüências de aminoácidos de diol e glicerol desidratase; e

25 c) remover quaisquer seqüências parciais do primeiro conjunto de dados de (b) para gerar um segundo conjunto de dados de seqüências de aminoácidos de diol e glicerol desidratases, sendo que as enzimas de diol desidratase e de glicerol desidratase são identificadas.

Com relação às subunidades grandes de diol e glicerol desidratases da invenção, as enzimas podem compreender uma subunidade grande que compreende uma sequência de aminoácidos tendo pelo menos 95% de identidade com uma sequência de aminoácidos selecionada a partir do grupo consistindo de SEQ ID N °: 8, 93, 99, 105, 135, 138, 141, 146, 164, 167, 170, 173, 176, 179, 182, 185, 188, 191, 194, 197, 200, 203, 206, 209, 212, 215, 218, 221, 224, 227, 130, 243, 254, 255, 256, 257, 258 e 259, com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

Com relação às subunidades médias de diol e glicerol desidratases da invenção, as enzimas podem compreender uma subunidade média que compreende uma sequência de aminoácidos tendo pelo menos 95% de identidade com uma sequência de aminoácidos selecionada a partir do grupo consistindo de SEQ ID n° 10, 95, 101, 107, 136, 139, 142, 148, 165, 168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 222, 225, 228, 231, 244, 250, 252, 260, 261, 262, 263, 364, 265, 266 e 167 com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

Com relação às subunidades pequenas de diol e glicerol desidratases da invenção, as enzimas podem compreender uma subunidade média compreendendo uma pequena subunidade que compreende uma sequência de aminoácidos tendo pelo menos 95% de identidade com uma sequência de aminoácidos selecionada a partir do grupo consistindo de SEQ ID n° 12, 97, 103, 109, 137, 140, 143, 150, 166, 169, 172, 175, 178, 181, 184, 187, 190, 193, 196, 199, 202, 205, 208, 211, 214, 217, 220, 223, 226, 229, 232, 234, 236, 238, 240, 242, 245, 248, 249, 251, 253, 268, 270, 271, 272, 273 e 274,

com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

Alternativamente, a diol desidratase ou glicerol desidratase podem
5 compreender subunidades grande, média e pequena unidas, compreendendo uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo de SEQ ID N °: 233, 235, 237, 239, 241, 246 e 247, com base no método de alinhamento Clustal
● W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP
10 LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

Alternativamente as enzimas de diol desidratase ou glicerol desidratase podem compreender subunidades grande, média e pequena unidas, tendo pelo menos 95% de identidade com uma seqüência de aminoácidos que compreende todas as três das seqüências de aminoácidos
15 que codificam subunidades grandes, médias e pequenas selecionadas do grupo consistindo em:

- a) SEQ ID nº 8, SEQ ID nº 10 e SEQ ID nº 12;
- b) SEQ ID nº 93, SEQ ID nº 95 e SEQ ID nº 97;
- c) SEQ ID nº 99, SEQ ID nº 101, e SEQ ID nº 103;
- 20 d) SEQ ID nº 105, SEQ ID nº 107 e SEQ ID nº 109;
- e) SEQ ID nº 135, SEQ ID nº 136 e SEQ ID nº 137;
- f) SEQ ID nº 138, SEQ ID nº 139 e SEQ ID nº 140;
- g) SEQ ID nº 146, SEQ ID nº 148 e SEQ ID nº 150;
- h) SEQ ID nº 141, SEQ ID nº 142 e SEQ ID nº 143 e
- 25 i) SEQ ID nº 164, SEQ ID nº 165 e SEQ ID nº 166;

com base no método de alinhamento Clustal W usando os parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

VIA 4:**(A) PIRUVATO A ALFA-ACETOLACTATO:**

Este substrato para conversão do produto é o mesmo descrito acima para via 1.

5 (K) ALFA-ACETOLACTATO A ÁCIDO 2,3-DIIDRÓXI-2-METILBUTANÓICO:

O substrato para conversão do produto de acetolactato (I) a ácido 2,3-diidróxi-2-metilbutanóico (IX) não é conhecido na técnica. Entretanto, o produto desta conversão foi relatado como um componente dos meios de fermentação (Ziadi et al. (1973) *Comptes Rendus des Seances de l'Academie des Sciences, Serie D: Sciences Naturelles* 276:965-8), mas o mecanismo de formação é desconhecido. O provável mecanismo de formação é a redução de acetolactato com o NADH ou NADPH como o doador de elétrons. Para utilizar esta via para a produção de 2-butanol, uma enzima catalisando essa reação precisa ser identificada ou produzida. Entretanto, o precedente para redução enzimática de cetonas a álcoois está bem estabelecido.

10

15

(L) ÁCIDO 2,3-DIIDRÓXI-2-METILBUTANÓICO A ÁCIDO 2-HIDRÓXI-2-METIL-3-FOSFOBUTANÓICO:

Não existem enzimas conhecidas que catalisam o substrato para a conversão do produto de ácido 2,3-diidróxi-2-metilbutanóico (IX) a ácido 2-hidróxi-2-metil-3-fosfobutanóico (X). Entretanto, há um grande número de quinases na natureza que possuem especificidades variadas. Por conseguinte, é provável que uma enzima possa ser isolada ou produzida com essa atividade.

20

(M) ÁCIDO 2-HIDRÓXI-2-METIL-3-FOSFOBUTANÓICO A 2-BUTANONA:

Não existem enzimas conhecidas que catalisam o substrato para conversão de produto de ácido 2-hidróxi-2-metil-3-fosfobutanóico (X) em 2-butanona (V). A combinação desta reação com a anterior, é muito parecida com a reação multi-etapa catalisada por mevalonato-5-pirofosfato (M5PP)

25

descarboxilase, que consiste na fosforilação inicial da M5PP a 3-fosfomevalonato-5-PP, seguido de descarboxilação dependente da eliminação de fosfato (Alvear *et al.* (1982) *Biochemistry* 21:4646-4650).

(F) 2-BUTANONA A 2-BUTANOL:

5 Este substrato para conversão do produto é o mesmo descrito acima para a via 1.

Portanto, como existem múltiplas vias recombinantes de piruvato para 2-butanol, existe uma série de opções para executar as etapas de conversão individuais, e o versado na técnica será capaz de utilizar seqüências publicamente disponíveis e seqüências divulgadas aqui para construir as vias de interesse. Uma listagem de um número representativo de genes conhecidos na técnica e úteis na construção de vias biossintéticas de 2-butanol é dada, sobretudo nas tabelas 1 e 2.

HOSPEDEIROS MICROBIANOS PARA PRODUÇÃO DE 2-BUTANOL E 2-BUTANONA

15 Os hospedeiros microbianos para produção de 2-butanol e 2-butanona podem ser selecionados a partir de bactérias, cianobactérias, fungos filamentosos e leveduras. Os hospedeiros microbianos para produção de 2-butanol e 2-butanona devem ser tolerantes ao produto produzido, de modo que a produtividade não é limitada pela toxicidade do produto para o hospedeiro. A seleção de um hospedeiro microbiano para produção de 2-butanol é descrita em detalhe abaixo. Os mesmos critérios são aplicáveis à seleção de um hospedeiro para produção de 2-butanona.

Os micróbios que são metabolicamente ativos em altos níveis de titulação de 2-butanol não são bem conhecidos na técnica. Embora mutantes tolerantes a butanol foram isoladas de *Clostridia* solventogenic, existe pouca informação disponível sobre a tolerância de butanol de outras cepas bacterianas potencialmente úteis. A maioria dos estudos sobre a comparação da tolerância de álcool em bactérias sugerem que butanol é mais tóxico do que etanol (de Cavalho

et al., *Microsc. Res. Tech.* 64:215-22 (2004) e Kabelitz et al., *FEMS Microbiol. Lett.* 220:223-227 (2003)). Tomas et al. (*J. Bacteriol.* 186:2006-2018 (2004)) afirmam que o rendimento de 1-butanol durante a fermentação em *Clostridium acetobutylicum* pode ser limitada pela toxicidade ao butanol. O principal efeito do 1-butanol sobre *Clostridium acetobutylicum* é a interrupção das funções da membrana (Hermann et al., *Appl. Environ. Microbiol.* 50:1238-1243 (1985)).

Os hospedeiros microbianos selecionados para a produção de 2-butanol devem ser tolerantes ao 2-butanol e deverão ser capazes de converter carboidratos a 2-butanol utilizando a via biossintética introduzidas. Os critérios para seleção de hospedeiros microbianos adequados incluem o seguinte: a tolerância intrínseca ao 2-butanol, alta taxa de utilização de carboidratos, a disponibilidade de ferramentas genéticas para a manipulação genética e a capacidade de gerar alterações cromossômicas estáveis.

As cepas hospedeiras adequadas com uma tolerância para 2-butanol podem ser identificadas pelo ensaio com base na intrínseca tolerância da cepa. A tolerância intrínseca dos micróbios ao 2-butanol pode ser medida pela determinação da concentração de 2-butanol que é responsável por 50% de inibição da taxa de crescimento (IC50), quando cultivado em um meio mínimo. Os valores de IC50 podem ser determinados usando métodos conhecidos na técnica. Por exemplo, os micróbios de interesse podem ser cultivados na presença de várias quantidades de 2-butanol e a taxa de crescimento monitorada pela medição da densidade óptica em 600 nanômetros. O tempo para duplicar pode ser calculado a partir da parte logarítmica da curva de crescimento e utilizada como uma medida da taxa de crescimento. A concentração de 2-butanol que produz 50% de inibição de crescimento pode ser determinada a partir de um gráfico da porcentagem de inibição de crescimento versus a concentração de 2-butanol. Preferencialmente, cepa hospedeira deveria ter uma IC50 para 2-butanol de

mais de cerca de 0,5%. Mais adequado é uma cepa hospedeira com uma IC50 para 2-butanol que é maior do que cerca de 1,5%. Particularmente adequado é uma cepa hospedeira com uma IC50 para 2-butanol que é maior do que cerca de 2,5%.

5 O hospedeiro microbiano para produção de 2-butanol também deve utilizar glicose e/ou outros carboidratos em uma taxa alta. A maioria dos micróbios são capazes de utilizar carboidratos. Entretanto, certos micróbios ambientais não podem usar eficientemente os carboidratos, e, portanto, não seriam hospedeiros adequados.

10 A habilidade de modificar geneticamente o hospedeiro é essencial para a produção de qualquer microorganismo recombinante. Modos de tecnologia de transferência de gene que podem ser utilizados incluem por eletroporação, conjugação, transdução ou transformação natural. Uma ampla gama de plasmídeos conjugativos hospedeiros e de marcadores resistentes a
15 medicamentos estão disponíveis. Os vetores de clonagem usados com um organismo são personalizados para o organismo hospedeiro com base na natureza dos marcadores resistentes a antibióticos que podem funcionar neste hospedeiro.

O hospedeiro microbiano pode, também, ser manipulado de modo
20 à inativar vias concorrentes para fluxo de carbono pela inativação de vários genes. Isto exige a disponibilização de ou transpósons ou vetores de integração cromossômicos para inativação direta. Adicionalmente, os hospedeiros de produção que são propensos a mutagênese química pode sofrer aprimoramento na intrínseca tolerância a 2-butanol através de
25 mutagênese química e triagem de mutantes.

Com base nos critérios acima descritos, hospedeiros microbianos adequados para a produção de 2-butanol e 2-butanona incluem, mas não estão limitados a, os membros do gênero *Clostridium*, *Zymomonas*, *Escherichia*,

Salmonella, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Pediococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* e *Saccharomyces*. Hospedeiros preferenciais incluem: *Escherichia coli*, *Alcaligenes eutrophus*,
5 *Bacillus licheniformis*, *Paenibacillus macerans*, *Rhodococcus erythropolis*, *Pseudomonas putida*, *Lactobacillus plantarum*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus faecalis*, *Pediococcus pentosaceus*, *Pediococcus acidilactici*, *Bacillus subtilis* e *Saccharomyces cerevisiae*.

CONSTRUÇÃO DE HOSPEDEIRO DE PRODUÇÃO

10 Os organismos recombinantes contendo os genes necessários que codificam a via enzimática para a conversão de um substrato de carbono fermentável para 2-butanol ou 2-butanona podem ser construídos com técnicas bem conhecidas na técnica. Na presente invenção, genes que codificam as enzimas da via biossintética 3 de 2-butanol: acetolactato sintase, acetolactato
15 descarboxilase, butanodiol dehydrogenase, butanodiol desidratase e butanol desidrogenase; ou via biossintética 3 de 2-butanona omitindo o butanol desidrogenase, podem ser isoladas de várias fontes, conforme descrito acima.

Métodos para obtenção dos genes desejados de um genoma bacteriano são comuns e bem conhecidas na técnica de biologia molecular. Por
20 exemplo, se a seqüência do gene é conhecida, iniciadores podem ser projetados e a pretendida seqüência amplificada utilizando-se métodos de amplificação padrão de iniciador direcionados como a reação em cadeia da polimerase (Patente U.S. Nº 4.683.202) para obter quantidades de DNA adequadas para clonagem em vetores de expressão. Se um gene que é heterólogo para uma
25 seqüência conhecida deve ser isolado, bibliotecas genômicas adequadas podem ser criadas por endonuclease de restrição de digestão e pode ser testada com sondas tendo seqüência complementar para a desejada seqüência genética. Uma vez que a seqüência é isolada, o DNA pode ser amplificado utilizando-se

métodos de amplificação padrão de iniciador direcionados como a reação em cadeia da polimerase (Patente U.S. Nº 4.683.202) para obter quantidades de DNA adequadas para clonagem em vetores de expressão, que são então transformados em células hospedeiras adequadas.

5 Além disso, dada a seqüência de aminoácidos de uma proteína com atividade enzimática desejada, a seqüência de codificação pode ser verificada pela tradução reversa da seqüência de proteína. Um fragmento de DNA contendo a seqüência de codificação pode ser preparado sinteticamente e clonado em um vetor de expressão, e então transformado na célula hospedeira desejada.

10 Com a preparação de um fragmento de DNA sintético contendo uma seqüência de codificação, esta seqüência pode ser otimizada para expressão na célula hospedeira procurada. As ferramentas para otimização do códon para expressão em um hospedeiro heterólogo estão prontamente disponíveis. Algumas ferramentas de otimização do códon estão disponíveis

15 baseadas no conteúdo GC do organismo hospedeiro. O conteúdo GC de alguns hospedeiros microbianos exemplificadores é dado na tabela 3.

TABELA 3

CONTEÚDOS GC DE HOSPEDEIROS MICROBIANOS

Cepa	% GC
<i>B. licheniformis</i>	46
<i>B. subtilis</i>	42
<i>C. acetobutylicum</i>	37
<i>E. coli</i>	50
<i>P. putida</i>	61
<i>A. eutrophus</i>	61

Cepa	% GC
<i>Paenibacillus macerans</i>	51
<i>Rhodococcus erythropolis</i>	62
<i>Brevibacillus</i>	50
<i>Paenibacillus polymyxa</i>	50

Uma vez que os genes relevantes para a via são identificados e isolados eles podem ser transformados em hospedeiros de expressão adequados por meios bem conhecidos na técnica. Os vetores úteis para a transformação de uma grande variedade de células hospedeiras são comuns e comercialmente disponíveis a partir de empresas como EPICENTRE® (Madison, WI), Invitrogen Corp. (Carlsbad, CA), Stratagene (La Jolla, CA) e New England Biolabs, Inc. (Beverly, MA). Tipicamente o vetor contém um marcador selecionável e seqüências que permitem a replicação autônoma ou integração cromossômica no hospedeiro desejado. Além disso, vetores adequados compreendem uma região promotora que abriga controles de iniciação transcricional e uma região de controle transcricional de terminação, entre os quais um fragmento de DNA da região de codificação pode ser inserido, para fornecer a expressão da região de codificação inserida. Ambas as regiões de controle podem ser derivadas a partir de genes homólogos à célula hospedeira transformada, embora deva ser entendido que tais regiões de controle também podem ser derivadas de genes que não são nativos para as espécies específicas escolhidas como hospedeiros de produção.

As regiões de controle de iniciação ou promotoras, as quais são úteis para dirigir a expressão relevante nas regiões de codificação da via da célula hospedeira pretendida, são numerosos e familiares para os versados na técnica. Virtualmente qualquer promotor capaz de acionar estes elementos

genéticos é adequado para a presente invenção, incluindo, mas não limitado a, promotores derivados dos seguintes genes: *CYC1*, *HIS3*, *GAL1*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, *TPI*, *CUP1*, *FBA*, *GPD* e *GPM* (úteis para expressão em *Saccharomyces*); *AOX1* (úteis para expressão em *Pichia*); bem como os promotores *lac*, *ara*, *tet*, *trp*, *IP_L*, *IP_R*, *T7*, *tac* e *trc* (úteis para expressão em *Escherichia coli*, *Alcaligenes*, e *Pseudomonas*); os promotores *amy*, *apr*, e *npr*, e vários promotores fago úteis para expressão em *Bacillus subtilis*, *Bacillus licheniformis*, e *Paenibacillus macerans*; *nisA* (úteis para expressão em bactérias Gram-positivas, Eichenbaum et al. *Appl. Environ. Microbiol.* 64(8):2763-2769 (1998)); e o promotor P11 sintético (útil para expressão em *Lactobacillus plantarum*, Rud et al., *Microbiology* 152:1011-1019 (2006)).

As regiões de controle de terminação podem, também, ser derivadas de vários genes nativos para os hospedeiros preferenciais. Opcionalmente, um sítio de terminação pode ser desnecessário. Entretanto, se incluído, isto é de máxima preferência.

Certos vetores são capazes de replicarem em uma ampla gama de bactérias hospedeiras e podem ser transferidos por conjugação. A sequência completa e anotada de pRK404 e três vetores relacionados, pRK437, pRK442 e pRK442(H), estão disponíveis. Essas derivações provaram ser valiosas ferramentas para a manipulação genética em bactérias gram-negativas (Scott et al., *Plasmid* 50(1):74-79 (2003)). Vários plasmídeos derivados da ampla faixa de hospedeiros Inc P4 plasmídeo RSF1010 também estão disponíveis com os promotores que podem funcionar em uma variedade de bactérias gram-negativas. Os plasmídeos pAYC36 e pAYC37, têm ativos promotores junto com múltiplos sítios de clonagem para permitir a genes heterólogos a expressão genética em bactérias gram-negativas.

As ferramentas cromossômicas de substituição genética também

estão amplamente disponíveis. Por exemplo, uma variante termosensível do replicon pWV101 de ampla faixa de hospedeiros foi alterado de modo a construir um plasmídeo pVE6002 que pode ser utilizado para produzir substituição genética em uma faixa de bactérias Gram-positivas (Maguin et al.,
5 *J. Bacteriol.* 174(17):5633-5638 (1992)). Adicionalmente, transposomes in vitro estão disponíveis junto a fontes comerciais como EPICENTRE® para criar mutações aleatórias em diversos genomas.

A expressão de uma via biossintética de 2-butanol em vários hospedeiros preferenciais microbianos é descrita com mais detalhes abaixo.

10 Para a expressão de uma via biossintética de 2-butanona, a mesma descrição se aplica, mas o substrato final para conversão do produto de 2-butanona para 2-butanol é omitido.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM *E. COLI*

Os vetores úteis para a transformação de *E. coli* são comuns e
15 comercialmente disponíveis a partir das empresas mostradas acima. Por exemplo, os genes de uma via biossintética de 2-butanol podem ser isolados a partir de várias fontes, como descrito acima, clonados em um vetor pUC19 modificado e transformados em *E. coli* NM522, conforme descrito nos exemplos 6 e 7. Alternativamente, os genes que codificam uma via
20 biossintética de 2-butanol podem ser divididos em múltiplos operons, clonados para expressão de vetores, e transformados em diversas cepas *E. coli*, conforme descrito nos exemplos 9, 10 e 11. A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

RHODOCOCCUS ERYTHROPOLIS

25 Uma série de vetores de transferência de *E. coli*-*Rhodococcus* estão disponíveis para expressão em *R. erythropolis*, incluindo, mas não se limitando a pRhBR17 e pDA71 (Kostichka et al., *Appl. Microbiol. Biotechnol.*

62:61-68 (2003)). Adicionalmente, uma série de promotores estão disponíveis para expressão de gene heterólogo em *R. erythropolis* (consulte por exemplo Nakashima *et al.*, *Appl. Environ. Microbiol.* 70:5557-5568 (2004), e Tao *et al.*, *Appl. Microbiol. Biotechnol.* 2005, DOI 10.1007/s00253-005-0064). As rupturas de gene visadas nos genes cromossômicos do *R. erythropolis* podem ser criadas utilizando os métodos descritos por Tao *et al.* *acima*, e Brans *et al.* (*Appl. Environ. Microbiol.* 66: 2029-2036 (2000)).

Os genes heterólogos necessários para a produção de 2-butanol, como descrito acima, podem ser clonados inicialmente em pDA71 ou pRhBR71 e transformados em *E. coli*. Os vetores podem então ser transformados em *R. erythropolis* por eletroporação, conforme descrito por Kostichka *et al.*, *acima*. Os recombinantes podem ser cultivados em meio sintético contendo glicose e a produção de 2-butanol pode acontecer ao se usar métodos de fermentação conhecidos na técnica. A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM *B. SUBTILIS*

Os métodos para a expressão dos genes e criação de mutações em *B. subtilis* também são bem conhecidos na técnica. Por exemplo, os genes de uma via biossintética de 2-butanol podem ser isolados a partir de várias fontes, como descrito acima, clonados em um vetor de transformação modificada *E. coli*-*Bacillus* e transformados em *Bacillus subtilis* BE1010, conforme descrito no exemplo 8. Os genes desejados podem ser clonados em um vetor de expressão *Bacillus* e transformados em uma cepa para fazer um hospedeiro de produção. Alternativamente, os genes podem ser integrados no cromossomo *Bacillus* utilizando replicons condicionais ou vetores suicidas que são conhecidas para o versado na técnica. Por exemplo, o "Bacillus Genetic Stock Center" possui numerosos vetores de integração. A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM *B.*

LICHENIFORMIS

A maioria dos plasmídeos e vetores de transferência que replicam em *B. subtilis* podem ser usados para transformar *B. licheniformis* por ou
5 transformação por protoplastos ou por eletroporação. Os genes necessários para a produção de 2-butanol podem ser clonados em plasmídeos pBE20 ou pBE60 derivados (Nagarajan et al., *Gene* 114:121-126 (1992)). Métodos para transformar *B. licheniformis* são conhecidas na técnica (por exemplo, consulte
● Fleming et al. *Appl. Environ. Microbiol.*, 61(11):3775-3780 (1995)). Os
10 plasmídeos construídos para a expressão em *B. subtilis* podem ser transformados em *B. licheniformis* para produzir um hospedeiro microbiano recombinante que produz 2-butanol. A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

PAENIBACILLUS MACERANS

Os plasmídeos podem ser construídos como descrito acima para a expressão em *B. subtilis* e utilizados para transformar *Paenibacillus macerans* por transformação de protoplasta para produzir um hospedeiro microbiano recombinante que produz 2-butanol. A via biossintética de 2-butanona pode ser
20 expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

ALCALIGENES (RALSTONIA) EUTROPHUS

Os métodos para a expressão dos genes e criação de mutações na *Alcaligenes eutrophus* são conhecidos na técnica (consulte, por exemplo,
25 Taghavi et al., *Appl. Environ. Microbiol.*, 60(10):3585-3591 (1994)). Os genes para uma via biossintética de 2-butanol podem ser clonados em qualquer um dos vetores da ampla faixa de hospedeiros descrita acima, e através de eletroporação em *Alcaligenes eutrophus* para gerar recombinantes que

produzem 2-butanol. A via do poli(hidroxibutirato) em *Alcaligenes* tem sido descrita em pormenor, uma variedade de técnicas genéticas para modificar o genoma do *Alcaligenes eutrophus* são conhecidas, e essas ferramentas podem ser aplicadas para projetar uma via biossintética de 2-butanol. A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo a butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

PSEUDOMONAS PUTIDA

Os métodos para expressão genética em *Pseudomonas putida* são conhecidos na técnica (consulte, por exemplo, Ben-Bassat et al., Patente U.S. Nº 6.586.229, que é aqui incorporada, por referência). Os genes de uma via biossintética de 2-butanol podem ser inseridos em pPCU18, e este DNA ligado pode ser eletroporado em células eletrocompetentes de *Pseudomonas putida* DOT-T1 C5aAR1 para gerar recombinantes que produzem 2-butanol. A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

LACTOBACILLUS PLANTARUM

O gênero *Lactobacillus* pertence à família *Lactobacillales* e muitos plasmídeos e vetores utilizados na transformação de *Bacillus subtilis* e *Streptococcus* podem ser usados para *Lactobacillus*. Alguns exemplos não-limitadores de vetores adequados incluem pAMâ1 e derivados dessas substâncias (Renault et al., *Gene* 183:175-182 (1996); e O'Sullivan et al., *Gene* 137:227-231 (1993)); pMBB1 e pHW800, um derivado de pMBB1 (Wyckoff et al. *Appl. Environ. Microbiol.* 62:1481-1486 (1996)); pMG1, um plasmídeo conjugativo (Tanimoto et al., *J. Bacteriol.* 184:5800-5804 (2002)); pNZ9520 (Kleerebezem et al., *Appl. Environ. Microbiol.* 63:4581-4584 (1997)); pAM401 (Fujimoto et al., *Appl. Environ. Microbiol.* 67:1262-1267 (2001)); e pAT392

(Arthur et al., *Antimicrob. Agents Chemother.* 38:1899-1903 (1994)). Vários plasmídeo a partir de *Lactobacillus plantarum* têm também sido divulgados (van Kranenburg et al., *Appl. Environ. Microbiol.* 71(3):1223-1230 (2005)).

Os vários genes para uma via biossintética de 2-butanol podem ser montados em qualquer vetor adequado, como aqueles descritos acima. Os códon
5 podem ser otimizados para expressão com base no índice de códon deduzido a partir de seqüências de genoma de *Lactobacillus plantarum* ou *Lactobacillus arizonensis*. Os plasmídeos podem ser introduzidos em uma célula hospedeira usando-se métodos conhecidos na técnica, como eletroporação
10 (Cruz-Rodz et al. *Molecular Genetics and Genomics* 224:1252-154 (1990), Bringel, et al. *Appl. Microbiol. Biotechnol.* 33: 664-670 (1990), Alegre et al., *FEMS Microbiology letters* 241:73-77 (2004)), e conjugação (Shrago et al., *Appl. Environ. Microbiol.* 52:574-576 (1986)). Os genes da via biossintética de 2-butanol podem também ser integrados no cromossomo de *Lactobacillus* usando
15 vetores de integração (Hols et al., *Appl. Environ. Microbiol.* 60:1401-1403 (1990), Jang et al., *Micro. Lett.* 24:191-195 (2003)). A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

ENTEROCOCCUS FAECIUM, ENTEROCOCCUS GALLINARIUM E ENTEROCOCCUS FAECALIS

20 O gênero *Enterococcus* pertence à família *Lactobacillales* e muitos plasmídeos e vetores utilizados na transformação de *Lactobacillus*, *Bacillus subtilis* e *Streptococcus*, descritos acima, podem ser usados para *Enterococcus*. Os vetores de expressão para *E. faecalis* usando o gene *nisA* a partir de *Lactococcus* podem também ser usados (Eichenbaum et al., *Appl.*
25 *Environ. Microbiol.* 64:2763-2769 (1998). Adicionalmente, vetores de substituição genética no cromossomo *E. faecium* podem ser usados (Nallaapareddy et al., *Appl. Environ. Microbiol.* 72:334-345 (2006)).

Os vários genes para uma via biossintética de 2-butanol podem

ser montados em qualquer vetor adequado, como aqueles descritos acima. Os códon podem ser otimizados para expressão com base no índice de códon deduzido a partir de seqüências de genoma de *Enterococcus faecalis* ou *Enterococcus faecium*. Os plasmídeos podem ser introduzidos na célula hospedeira por meio de métodos conhecidos na técnica, como a eletroporação, conforme descrito por Cruz-Rodz et al. (*Molecular Genetics and Genomics* 224:1252-154 (1990)) ou conjugação, conforme descrito por Tanimoto et al. (*J. Bacteriol.* 184:5800-5804 (2002)) e Grohamann et al. (*Microbiol. Mol. Biol. Rev.* 67:277-301 (2003)). A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL OU 2-BUTANONA EM

PEDIOCOCCUS PENTOSACEUS E PEDIOCOCCUS ACIDILACTICI,

O gênero *Pediococcus* pertence à família *Lactobacillales* e muitos plasmídeos e vetores utilizados na transformação de *Bacillus subtilis* e *Streptococcus*, descritos acima, podem ser usados para *Pediococcus*. Um exemplo não-limitador de um vetor adequado é pHPS9 (Bukhtiyarova et al. *Appl. Environ. Microbiol.* 60:3405-3408 (1994)). Vários plasmídeos a partir de *Pediococcus* têm também sido divulgados (Alegre et al., *FEMS Microbiol. Lett.* 250:151-156 (2005); Shareck et al. *Crit. Rev Biotechnol.* 24:155-208 (2004)).

Os genes para uma via biossintética de 2-butanol podem ser montados em qualquer vetor adequado, como aqueles descritos acima. Os códon podem ser otimizados para expressão com base no índice de códon deduzido a partir de seqüências de genoma de *Pediococcus pentosaceus*. Os plasmídeos podem ser introduzidos na célula hospedeira por meio de métodos conhecidos na técnica, como eletroporação (consulte, por exemplo, Osmanagaoglu et al., *J. Basic Microbiol.* 40:233-241 (2000); Alegre et al., *FEMS Microbiol. Lett.* 250:151-156 (2005)) e conjugação (Gonzalez and Kunka, *Appl. Environ. Microbiol.* 46:81-89 (1983)). Os genes da via

biossintética de 2-butanol podem também ser integrados no cromossomo de *Pediococcus* usando vetores de integração (Davidson et al. *Antonie van Leeuwenhoek* 70:161-183 (1996)). A via biossintética de 2-butanona pode ser expressa de forma semelhante, omitindo o butanol desidrogenase.

5

MEIO DE FERMENTAÇÃO

O meio de fermentação da presente invenção precisa conter substratos adequados de carbono. Os substratos adequados podem incluir, mas não estão limitados a monossacarídeos como a glicose e a frutose, oligossacarídeos como a lactose ou sacarose, polissacarídeos como o amido ou
10 celulose, ou misturas dos mesmos, e misturas não purificadas a partir de fontes renováveis de matérias primas, como queijo e soro de leite permeado, água de maceração de milho (*comsteep liquor*), melaço de açúcar de beterraba e malte de cevada. Além disso, o substrato de carbono pode também ser um substrato de carbono como o dióxido de carbono, ou metanol para o qual a conversão
15 metabólica em intermediários bioquímicos vitais foi demonstrada. Além de substratos de um e dois carbonos, organismos metilotróficos são também conhecidos por utilizar uma série de outros compostos que contêm carbono, como a metilamina, glucosamina e uma variedade de aminoácidos para a atividade metabólica. Por exemplo, leveduras metilotróficas são conhecidas por utilizar o
20 carbono de metilamina para formar trealose ou glicerol (Bellion et al., *Microb. Growth C1 Compd.*, [Int. Symp.], 7th (1993), 415-32, Editor(s): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK). Semelhantemente, várias espécies de *Candida* metabolizarão alanina ou ácido oléico (Sulter et al., *Arch. Microbiol.* 153:485-489 (1990)). Por isso é que se espera uma fonte de carbono
25 para utilização na presente invenção que possa abranger uma grande variedade de substratos que contêm carbono e apenas seja limitada pela escolha do organismo.

Embora seja esperado que todos os substratos de carbono

acima mencionados e suas misturas sejam adequados para a presente invenção, substratos preferenciais de carbono são a glicose, frutose e sacarose, assim como as misturas de quaisquer destes açúcares. A sacarose pode ser obtida a partir de matérias primas como cana de açúcar, açúcares de beterraba, mandioca e sorgo doce. A glicose e a dextrose podem ser obtidas através de sacarificação do amido baseado em matérias primas incluindo grãos, como milho, trigo, centeio, cevada e aveia.

Além disso, açúcares fermentáveis podem ser obtidos a partir de celulose e biomassa lignocelulósica através de processos de tratamento prévio e sacarificação, como descrito, por exemplo, no pedido de patente US20070031918A1, que é co-propriedade e co-pendente, o qual está aqui incorporado por referência. A biomassa refere-se a qualquer material celulósico ou lignocelulósico e inclui materiais que compreendem celulose, e opcionalmente compreendendo, ainda, hemicelulose, lignina, amido, oligossacarídeos e/ou monossacarídeos. E la pode, também, compreender componentes adicionais, como proteína e/ou lipídio. A biomassa pode ser derivada de uma única fonte, ou a biomassa pode compreender uma mistura derivada de mais de uma fonte; por exemplo, a biomassa pode compreender uma mistura de sabugos de milho e palha de milho, ou uma mistura de grama e folhas. Ela inclui, mas não se limita a, colheitas usadas para bioenergia, resíduos agrícolas, resíduos sólidos urbanos, resíduos industriais sólidos, lamas provenientes de fábricas papel, resíduos de jardim, madeira e resíduos florestais. Exemplos de biomassa incluem, mas não estão limitados a, grãos de milho, sabugos de milho, restos de colheitas, como cascas de milho, palha de milho, gramíneas, trigo, palha de trigo, cevada, palha de cevada, feno, palha de arroz, capim "switchgrass", resíduos de papel, bagaço de cana de açúcar, sorgo, soja, componentes obtidos a partir de moagem de grãos, árvores, galhos e raízes, folhas, madeira, serradura, arbustos e moitas, legumes, frutos, flores e estrume animal.

Além de uma adequada fonte de carbono, meios de fermentação adequados devem conter minerais, sais, cofatores, tampões e outros componentes, conhecidas por aqueles versados na técnica, adequados para o crescimento das culturas e na promoção de uma via enzimática necessária para produção de 2-butanol ou 2-butanona.

CONDIÇÕES DA CULTURA

Tipicamente as células são cultivadas a uma temperatura na faixa de cerca de 25°C a cerca de 40°C em um meio adequado. Os meios de crescimento adequados na presente invenção são comuns e comercialmente preparados como meio Luria Bertani (LB), meio Sabouraud dextrose (SD) ou meio Yeast Medium [meio de levedura] (YM). Outros meios de crescimento definidos ou sintéticos podem também ser utilizados, bem como o meio adequado para o crescimento de microorganismos particulares será conhecida por um perito na matéria de microbiologia ou ciência da fermentação. O uso de agentes conhecidos para modular a repressão de catabólito direta ou indiretamente, por exemplo, adenosina cíclico 2':3'-monofosfato, também podem ser incorporadas na fermentação médio.

A faixa de pH adequada para a fermentação está entre pH 5,0 e pH 9,0, onde pH de 6,0 a pH 8,0 é preferencial como a condição inicial.

As fermentações podem ser realizadas sob condições aeróbicas e anaeróbicas, onde as condições anaeróbicas ou microaeróbicas são preferenciais.

FERMENTAÇÕES INDUSTRIAIS POR BATELADA E CONTÍNUA

O presente processo emprega um método de fermentação por batelada. Uma clássica fermentação por batelada é um sistema fechado em que a composição do meio é fixada no início da fermentação e não está sujeita a alterações artificiais durante a fermentação. Assim, no início da fermentação o meio é inoculado com o desejado organismo ou organismos, e a fermentação ocorre sem que se acrescente nada ao sistema. Entretanto, tipicamente uma

fermentação por "batelada" é batelada com relação à adição de fonte de carbono e muitas vezes são feitas tentativas para controlar fatores tais como pH e concentração de oxigênio. Nos sistemas de batelada, as composições de biomassa e metabólito do sistema mudam constantemente até o momento no qual a fermentação é interrompida. Dentro da batelada as culturas de células moderam através de uma fase lag estática para uma fase log de crescimento elevado e finalmente para uma fase estacionária, onde a taxa de crescimento é diminuída ou interrompida. Se não tratadas, as células em fase estacionária eventualmente morrerão. As células em fase log geralmente são responsáveis pela maior parte da produção do produto final ou intermediário.

Uma variação no sistema padrão por batelada é o sistema por batelada alimentado.- Os processos de fermentação por batelada alimentados também são adequados na presente invenção e compreendem um sistema típico por batelada com a exceção de que o substrato é adicionado em incrementos enquanto a fermentação progride.- Os sistemas alimentados por batelada são úteis quando a repressão de catabólito é capaz de inibir o metabolismo da célula e quando é desejável se ter limitadas quantidades de substrato no meio.- A medição da concentração real de substrato em sistemas por batelada alimentados é difícil e, portanto, é estimada com base nas variações dos fatores mensuráveis, como pH, oxigênio dissolvido e da pressão parcial de resíduos de gases como o CO₂. Bateladas e fermentações por bateladas alimentadas são comuns e bem conhecidas na técnica e exemplos podem ser encontrados em Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA., ou Despende, Mukund V., *Appl. Biochem. Biotechnol.*, 36:227, (1992), aqui incorporado por referência.

Embora a presente invenção seja realizada com a modalidade de batelada, se espera que o método seja adaptável aos métodos de fermentação contínua. A fermentação contínua é um sistema aberto em que o

meio de fermentação definido é adicionado continuamente para um biorreator e igual quantidade de meio condicionado é removido simultaneamente para processamento. A fermentação contínua geralmente mantém as culturas constantemente em alta densidade.

5 Ela permite a modulação de um fator ou de qualquer número de fatores que afetam o crescimento celular ou a concentração do produto final. Por exemplo, um método irá manter um nutriente limitante, como a fonte de carbono ou nível de nitrogênio a uma taxa fixa e permitir que todos os outros parâmetros a variar moderadamente. Em outros sistemas, uma série de fatores que afetam o
10 crescimento podem ser modificados continuamente enquanto a concentração de células, medida pela turvação do meio de cultura, é mantida constante. Os sistemas contínuos se esforçam para manter um regime permanente de crescimento e, portanto, a perda de células devido ao meio sendo retirado deve ser balanceada com a taxa de crescimento celular na fermentação. Os métodos
15 de modulação de nutrientes e fatores de crescimento para os processos de fermentação contínuos, como técnicas para maximizar a taxa de formação de produto são bem conhecidos na técnica de microbiologia industrial e uma variedade de métodos são detalhados por Brock, *acima*.

É contemplado que a presente invenção possa ser praticada
20 usando ou processos por batelada, ou processos por batelada alimentado, ou processos contínuos e qualquer modo conhecido de fermentação pode ser adequado. Além disso, espera-se que as células possam ser imobilizadas em um substrato como catalisador da célula inteira, e submetida às condições de fermentação para produção de 2-butanol e 2-butanona.

25 **MÉTODOS PARA ISOLAMENTO DE 2-BUTANOL E 2-BUTANONA DO MEIO DE**
FERMENTAÇÃO

O 2-butanol bioproduzido pode ser isolado do meio de fermentação usando métodos conhecidos na técnica de fermentações ABE(consulte por

exemplo, Durre, *Appl. Microbiol. Biotechnol.* 49:639-648 (1998), Groot et al., *Process Biochem.* 27:61-75 (1992), e referências neste). Por exemplo, sólidos podem ser removidos do meio de fermentação por centrifugação, filtração, decantação ou similares. Em seguida, o 2-butanol pode ser isolado do meio de fermentação usando métodos como a destilação, destilação azeotrópica, extração líquido-líquido, adsorção, arraste gasoso (*gas stripping*), evaporação por membrana ou pervaporação. Estes mesmos métodos podem ser adaptados para isolar 2-butanona bioproduzida do meio de fermentação

EXEMPLOS

A presente invenção está adicionalmente definida nos exemplos a seguir. Deve ficar entendido que estes exemplos, apesar de indicarem uma modalidade preferencial da invenção, são dados apenas a título de ilustração. A partir da discussão acima e destes exemplos, um versado na técnica pode conhecer as características essenciais desta invenção e, sem que se desvie do caráter e âmbito da mesma, podemos fazer várias alterações e modificações na invenção para adaptá-la às suas diferentes utilizações e condições.

MÉTODOS GERAIS

As técnicas padrão de DNA recombinante e clonagem molecular descritas nos exemplos são bem conhecidas na técnica e são descritas por Sambrook, J., Fritsch, E. F. F. e Maniatis, T. *Molecular Cloning: A Laboratory Manual*, nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, (1989) (Maniatis) e por T. J. Silhavy, M. L. Bannan e L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984) e por Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, publicado por Greene Publishing Assoc. e Wiley-Interscience (1987).

Os materiais e métodos adequados para a manutenção e o crescimento de culturas de bactérias são bem conhecidos na técnica. As técnicas adequadas para utilização nos seguintes exemplos podem ser

encontradas de acordo com o *Manual of Methods for General Bacteriology* (Phlipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg e G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994)) ou por Thomas D. Brock em *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989). Todos os reagentes, enzimas de restrição e materiais descritos para o crescimento e manutenção das células bacterianas foram obtidos de Aldrich Chemicals (Milwaukee, WI), BD Diagnostic Systems (Sparks, MD), Life Technologies (Rockville, MD) ou Sigma Chemical Company (St. Louis, MO) a não ser que indicado de outra forma. As cepas bacterianas são obtidas a partir da "American Type Culture Collection" (ATCC, Manassas, VA) exceto onde especificado em contrário.

OS INICIADORES OLIGONUCLEOTÍDEOS DESCRITOS NOS EXEMPLOS A SEGUIR SÃO APRESENTADOS NA TABELA 4. TODOS OS INICIADORES OLIGONUCLEOTÍDEOS FORAM SINTETIZADOS POR SIGMA-GENOSYS (WOODLANDS, TX). TABELA 4

CLONAGEM E TRIAGEM DE INICIADORES

Gene	Nome do iniciador	Seqüência	SEQ ID nº	Descrição
<i>budB</i>	B1	CACCATGGACAAACAGTATCCGGTACGCC	15	<i>budB</i> forward
<i>budB</i>	B2	CGAAGGGCGATAGCTTTACCAATCC	16	<i>budB</i> reverse
<i>budA</i>	B3	CACCATGAATCATTCTGCTGAATGCACCTG CG	17	<i>budA</i> forward
<i>budA</i>	B4	GATACTGTTTGTCCATGTGACC	18	<i>budA</i> reverse
<i>budC</i>	B5	CACCATGAAAAAAGTCGCACTTGTTACC	19	<i>budC</i> forward

Gene	Nome do iniciador	Seqüência	SEQ ID nº	Descrição
<i>budC</i>	B6	TTAGTTAAATACCAT	20	<i>budA</i> reverse
<i>pddA</i>	B7	CACCATGAGATCGA AAAGATTTG	21	<i>pddABC</i> forward
<i>pddC</i>	B8	CTTAGAGAAGTTAATCGTCGCC	22	<i>pddABC</i> reverse
<i>sadh</i>	B9	CACCATGAAAGCCCTCCAGTACACC	23	<i>sadh</i> forward
<i>sadh</i>	B10	CGTCGTGTCATGCCCGGG	24	<i>sadh</i> reverse
<i>budA</i>	B11	GATCGAATTCGTTTAAACTTAGTTTTCTACC GCACG	25	<i>budABC</i> forward
<i>budC</i>	B12	GATCGCATGCAAGCTTTCATATAGTCGGAA TTCC	26	<i>budABC</i> reverse
<i>pddA</i>	B13	GATCGAATTCGTTTAAACAAAGGAGGTCTG ATTCATGAGATCG	27	<i>pddABC</i> forward
<i>pddC</i>	B14	GATCGGATTCTTAATCGTCGCC	28	<i>pddABC</i> reverse
<i>sadh</i>	B15	GATCGGATCCAAAGGAGGTCGGGCGCATG AAAGCCC	29	<i>sadh</i> forward
<i>sadh</i>	B16	GATCTCTAGAAAGCTTTCAGCCCGGGACGA CC	30	<i>sadh</i> reverse

Gene	Nome do iniciador	Sequência	SEQ ID nº	Descrição
--	BenF	ACTTTCTTTTCGCCTGTTTCAC	31	--
--	BenBPR	CATGAAGCTTGTTTAACTCGGTGACCTTG AAAATAATGAAAACCTTATATTGTTTTGAAAA TAATGAAAACCTTATATTG	32	--
budAB	BABC F	GAGCTCGAATTCAAAGGAGGAAGTGTATAT GAATCATTC	33	budAB forward
budAB	BAB R	GGATCCTCTAGAATTAGTTAAATACCATCC CGCCG	34	budAB reverse
budC	BC Spe F	ACTAGTAAAGGAGGAAAGAGTATGAAGAAG GTCGCACT	40	budC forward
budC	BC Xba R	TCTAGAAAGCAGGGGCAAGCCATGTC	41	budA reverse
pddABC- ddrAB	DDo For	AAGCTTAAAGGAGGCTGATTCATGAGATCG AAAAGATT	44	pddABC- ddrAB forward
pddABC- ddrAB	DDo Rev	TCTAGATTATTCATCCTGCTGTTCTCC	45	pddABC- ddrAB reverse
chnA	ChnA F	CATCAATTGACTACGTAGTCGTACGTGTAA GGAGGTTTGAAATGGAAAAAATTATG	54	chnA forward
chnA	ChnA R	CATGCTAGCCCCGGGTATCTTCTACTCATT TTTTATTTCG	55	chnA reverse
--	Top ter F1	CTAGAAGTCAAAAGCCTCCGACCGGAGGC TTTTGA	58	forward
--	Top ter F2	CTGCTCGAGTTGCTAGCAAGTTTAAACAAA AAAAAGCCCGCTCATTAGGCGGGCTGAGC	59	forward

Gene	Nome do iniciador	Sequência	SEQ ID nº	Descrição
		T		
--	Bot ter R1	CAGCCCGCCTAATGAGCGGGCTTTTTTTTG TTTAAAC	60	reverse
--	Bot ter R2	TTGCTAGCAACTCGAGCAGTCAAAGCCTC CGGTCGGAGGCTTTTGACTT	61	reverse
KA-AT	OT872	CTCCGGAATTCATGTCTGACGGACGACTCA CCGCA	127	aminoálcool quinase/liase operon forward
KA-AT	OT873	TTCCAATGCATTGGCTGCAGTTATCTCTGT GCACGAGTGCCGATGA	128	aminoálcool quinase/liase operon reverse
KA	OT879	AACAGCCAAGCTTGGCTGCAGTCATCGCG CATTCTCCGGG	129	aminoálcool quinase reverse
AT	OT880	TCTCCGGAATTCATGACGTCTGAAATGACA GCGACAGAAG	130	aminoálcool liase forward
pBAD.HisB	OT909	GCTAACAGGAGGAAGAATTCATGGGGGGT TCTC	131	sítio Adds EcoRI para substituir o sítio NcoI

Gene	Nome do iniciador	Seqüência	SEQ ID nº	Descrição
pBAD.HisB	OT910	GAGAACCCCCCATGAATTCTTCCTCCTGTTA GC	132	sítio Adds EcoRI para substituir o sítio NcoI
budAB	N84seqR 3	GGACCTGCTTCGCTTTATCG	159	reverse
APT	APTfor	GCGCGCCCGGGAAGAAGGAGCTCTTCACC ATGAACAAACCACAGTCTTGG	162	APT forward
APT	APTrev	GCGCGCCCGGGTTCATGCCACCTCTGCG	163	APT reverse

TABELA 5**INICIADORES DE SEQUENCIAMENTO**

Nome	Seqüência	Gene-específico	SEQ ID nº
M13 forward	GTAAAACGACGGCCAGT	--	35
M13 reverse	AACAGCTATGACCATG	--	36
N83 SeqF2	GCTGGATTACCAGCTCGACC	--	37
N83 SeqF3	CGGACGCATTACCGGCAAAG	--	38
N84 Seq R2	GCATCGAGATTATCGGGATG	--	65
N84 SeqR4	CGAAGCGAGAGAAGTTATCC	--	39
Trc F	TTGACAATTAATCATCCGGC	todos	42
Trc R	CTTCTCTCATCCGCCAAAAC	todos	43
DDko seq F2	GCATGGCGCGGATTTGACGAAC	pddABC-ddrAB	46

Nome	Seqüência	Gene-específico	SEQ ID nº
DDko seq F5	CATTAAAGAGACCAAGTACGTG	pddABC-ddrAB	47
DDko seq F7	ATATCCTGGTGGTGTCTCGTCGGCGT	pddABC-ddrAB	48
DDko seq F9	TCTTTGTCACCAACGCCCTGCG	pddABC-ddrAB	49
DDko seq R1	GCCCACCGCGCTCGCCGCCGCG	pddABC-ddrAB	50
DDko seq R3	CCCCCAGGATGGCGGCTTCGGC	pddABC-ddrAB	51
DDko seq R7	GGGCCGACGGCGATAATCACTT	pddABC-ddrAB	52
DDko seq R10	TTCTTCGATCCACTCCTTAACG	pddABC-ddrAB	53
chnSeq F1	CTCAACAGGGTGTAAGTGTAGT	chnA	56
chnSeq R1	CGTTTTGATATAGCCAGGATGT	chnA	57
pCL1925 vec F	CGGTATCATCAACAGGCTTACC	todos	62
pCL1925 vec R1	AGGGTTTTCCCAGTCACGACGT	todos	63
pCL1925 vec R2	CGCAATAGTTGGCGAAGTAATC	todos	64
APTseqRev	GCTAGAGATGATAGC	APT	160
APTseqFor	GGAAGAGACTATCCAGCG	APT	161

MÉTODOS PARA A DETERMINAÇÃO DA CONCENTRAÇÃO DE 2-BUTANOL E 2-BUTANONA NO MEIO DE CULTURA

A concentração de 2-butanol e 2-butanona no meio de cultura pode ser determinada por inúmeros métodos conhecidos na técnica. Por exemplo, um método específico de cromatografia líquida de alta eficiência (HPLC-high performance liquid chromatography) utilizando uma coluna Shodex SH-1011 com uma coluna de segurança Shodex SH-G, ambas compradas de Waters Corporation (Milford, MA), com detecção de índice de refração (IR). A separação

cromatográfica foi obtida usando-se H_2SO_4 0,01 M como a fase móvel, com um fluxo de 0,5 mL/min e uma temperatura de coluna de 50°C. Nas condições usadas, 2-butanona e 2-butanol tiveram tempos de retenção de 39,5 e 44,3 min, respectivamente. Alternativamente, métodos de cromatografia gasosa (GC-gas chromatography) estão disponíveis. Por exemplo, um determinado método GC utilizado um coluna HP-INNOWax (30 m x 0,53 mm ID, 1 µm de espessura de filme, Agilent Technologies, Wilmington, DE), com um detector por ionização de chama (FID - flame ionization detector). O gás de arraste foi hélio a um fluxo de 4,5 mL/min, medido a 150°C com pressão na cabeça da coluna constante; "split" do injetor foi 1:25 a 200°C; temperatura do forno foi 45°C durante 1 minuto, 45 para 220°C a 10°C/min, e 220°C durante 5 minutos; e detecção FID foi empregada a 240°C com gás de compensação hélio a 26 mL/min. Os tempos de retenção de 2-butanona e 2-butanol foram de 3,61 e 5,03 min, respectivamente.

A 2-Butanona também pode ser detectada por derivatização com 3-metil-2-benzotiazolinona hidrazona (MBTH - *3-methyl-2-benzothiazolinone hydrazone*). Uma solução aquosa contendo 2-butanona é misturada com um volume igual de uma solução aquosa de 6 mg/mL de MBTH em 375 mM de glicina-HCl (pH 2,7) e incubados a 100°C por 3 min. As amostras resultantes de derivatizado de MBTH são analisadas em uma coluna(Supelco) com 25 cm x 4,6 mm (id) Supelosil LC-18-D5 5 µm utilizando uma fase móvel de 55% de acetonitrila em água a um fluxo de 1 mL/min. Os derivados de 2-butanona aparecem como dois picos (isômeros *cis* e *trans*), com tempos de retenção de cerca de 12,3 e 13,3 min e absorbância máxima de 230 e 307 nm.

O significado das siglas é o seguinte: "s" significa segundo(s), "min" significa minuto(s), "h" significa hora(s), "psi" significa libras por polegada quadrada, "nm" significa nanômetros, "d" significa dia(s), "µL" significa microlitro(s), "mL" significa mililitro(s), "L" significa litro(s), "mm" significa milímetros(s), "Mm" significa millimolar, "M" significa molar, "mmol" significa

millimole(s), " μ mol" significa micromole(s) "g", "significa grama(s)," "mg" "significa micrograma (s) e "Ng" significa nanogramas(s), "PCR" significa reação em cadeia da polimerase, "DO" significa densidade óptica, "DO₆₀₀", a densidade óptica medida no comprimento de onda de 600 nm, "kDa" significa 5 kilodaltons, "g", a constante gravitacional, "pb" significa par(es) de base, "kpb" significa kilopar(es) de base, "% w/v" significa porcentagem em peso/volume, "% v/v" significa porcentagem em volume/volume, "wt%" significa porcentagem em peso, "HPLC" significa cromatografia líquida de alta eficiência, e "GC" significa cromatografia gasosa. O termo "seletividade molar" é o número 10 de moles de produto produzidos por mol de substrato de açúcar consumido e são relatados como uma porcentagem.

EXEMPLO 1

CLONAGEM E EXPRESSÃO DE ACETOLACTATO SINTASE

O objetivo do presente exemplo foi clonar e expressar em *E. coli* 15 o gene *budB* que codifica a enzima acetolactato sintase. O gene *budB* foi amplificado a partir do DNA genômico de *Klebsiella pneumoniae* cepa ATCC 25955 usando PCR.

A seqüência *budB* que codifica acetolactato sintase foi amplificada a partir do DNA genômico de *Klebsiella pneumoniae* (ATCC 25955) por PCR 20 usando o par iniciador B1 (SEQ ID nº 15) e B2 (SEQ ID nº 16). Outros reagentes de amplificação de PCR (por exemplo, Kod HiFi DNA polimerase (Novagen Inc., Madison, WI; catálogo Nº 71805-3)) foram fornecidos em kits do fabricante e usados de acordo com o protocolo do fabricante. O DNA genômico de *Klebsiella pneumoniae* foi preparado usando o kit Gentra Puregene (Gentra 25 Systems, Inc., Minneapolis, MN; catalog number D-5000A). A amplificação foi executada em um termociclador de DNA GeneAmp 9700 (PE Applied Biosystems, Foster city, CA). A seqüência de nucleotídeos do quadro de leitura aberto (ORF - open reading frame) e da seqüência de aminoácidos predita da

Enzima são dadas como SEQ ID nº 3 e SEQ ID nº 4, respectivamente.

Para os estudos de expressão a tecnologia de clonagem Gateway (Invitrogen Corp., Carlsbad, CA) foi usada. O vetor de entrada pENTR/SD/D-TOPO permite clonagem direcional e forneceu uma sequência Shine-Dalgarno para o gene de interesse. O vetor de destino pDEST14 usou um promotor T7 para expressão do gene sem identificação. O iniciador forward incorporou quatro bases (CACC), imediatamente adjacente ao códon de partida translacional permitir a clonagem direcional da região de codificação *budB* de acetolactato sintase do produto da PCR em pENTR/SD/D-TOPO (Invitrogen), gerando o plasmídeo pENTRSDD-TOPO *budB*. O pENTR construído foi transformado em células *E. coli* Top10 (Invitrogen) e chapeado de acordo com as recomendações do fabricante. Os transformante foram cultivados de um dia para o outro e DNA do plasmídeo foi preparado usando o kit QIAprep Spin Miniprep (Qiagen, Valencia, CA; catálogo Nº 27106) de acordo com as recomendações do fabricante. Para criar um clone de expressão, a região codificante *budB* de pENTRSDD-TOPO *budB* foi transferida para o vetor pDEST 14 por recombinação *in vitro* usando a mistura LR Clonase (Invitrogen, Corp., Carlsbad, CA). O vetor resultante, pDEST14*budB*, foi transformado em células de BL-21-AI (Invitrogen Corp.). As células BL-21-AI transportam uma cópia cromossômica da polimerase T7 RNA sobre o controle do promotor arabinose-induzível *araBAD*.

Os transformantes são inoculados no meio LB suplementado com 50 µg/mL de ampicilina e cultivados de um dia para o outro. Uma alíquota da cultura é usada para inocular 50 mL de meio LB suplementado com 50 µg/mL de ampicilina. A cultura é incubada a 37°C com agitação até a DO₆₀₀ atingir 0,6 a 0,8. A cultura é dividida em duas porções de 25-mL e arabinose é adicionada a um dos frascos para uma concentração final de 0,2% de peso por volume. O frasco de controle negativo não é induzido com arabinose. Os frascos são incubados por 4 h a 37°C com agitação. As células são colhidas por

centrifugação e os péletes das células estão ressuspensos em Mops 50 mM, tampão pH 7,0. As células são perturbadas, quer por sonicação ou por uma passagem através de uma prensa francesa (French Pressure Cell). Cada lisado celular é centrifugado produzindo o sobrenadante e o pélete da fração insolúvel. Uma alíquota de cada fração (toda a célula lisada, a partir de células induzidas e de controle, é ressuspensa na SDS (MES) tampão de carga (Invitrogen), aquecida a 85°C por 10 min e submetidas à análise SDS-PAGE (NuPAGE 4 a 12% Bis Tris-Gel, catálogo número NP0322Box, Invitrogen). A proteína de peso molecular esperado, como deduzida a partir da seqüência de ácido nucléico, está presente na cultura induzida mas não no controle não induzido.

A atividade de acetolactato sintase nos extratos livres da célula é medida usando-se o método descrito por Bauerle et al. (Bauerle et al. (1964) *Biochim. Biophys. Acta* 92:142-149). A concentração protéica é medida ou pelo método Bradford ou pelo Kit Bicinchoninic (Sigma, catálogo N° BCA-1; St. Louis, MO) usando albumina sérica bovina (BSA) (Bio-Rad, Hercules, CA) como o padrão.

EXEMPLO 2

CLONAGEM E EXPRESSÃO DE ACETOLACTATO DESCARBOXILASE

O objetivo do presente exemplo foi clonar e expressar em *E. coli* o gene *budB* que codifica a enzima acetolactato descarboxilase. O gene *budB* foi amplificado a partir do DNA genômico de *Klebsiella pneumoniae* cepa ATCC 25955 usando PCR.

A seqüência *budA* que codifica acetolactato descarboxilase, foi clonada da mesma maneira descrita para *budB* em exemplo 1, com a exceção de que os iniciadores usados para amplificação de PCR foram B3 (SEQ ID n° 17) e B4 (SEQ ID n° 18). A seqüência de nucleotídeos do quadro de leitura aberto (ORF - open reading frame) e da seqüência de aminoácidos predita da

102
AS

Enzima são dadas como SEQ ID nº 1 e SEQ ID nº 2, respectivamente. O plasmídeo resultante foi chamado pENTRSDD-TOPObudA.

A atividade de acetolactato descarboxilase nos extratos livres da célula é medida usando-se o método descrito por Bauerle et al., *acima*.

5

EXEMPLO 3 (PROFILÁTICO)

CLONAGEM E EXPRESSÃO DE BUTANODIOL DESIDROGENASE

O propósito deste exemplo profilático é para descrever como clonar e expressar em *E. coli* o gene *budC* que codifica a enzima butanodiol desidrogenase. O gene *budC* foi amplificado a partir do DNA genômico de *Klebsiella pneumoniae* cepa IAM1063 usando PCR.

10

A sequência *budC* codificando butanodiol desidrogenase é clonada e expressa da mesma maneira conforme descrito para *budA* no exemplo 1, exceto que os iniciadores usados para amplificação de PCR são B5 (SEQ ID nº 19) e B6 (SEQ ID nº 20) e o modelo de DNA genômico é da *Klebsiella pneumoniae* IAM1063 (a qual é obtida na coleção de culturas do Instituto de microbiologia aplicada, Tóquio, Japão). O DNA genômico de *Klebsiella pneumoniae* IAM1063 é preparado usando o kit Gentra Puregene (Gentra Systems, Inc., Minneapolis, MN; catalog number D-5000A). A sequência de nucleotídeos do quadro de leitura aberto (ORF - open reading frame) e da sequência de aminoácidos predita da Enzima são dadas como SEQ ID nº 5 e SEQ ID nº 6, respectivamente.

15

20

A atividade de butanodiol desidrogenase nos extratos livres da célula é medida espectrofotometricamente através do consumo de NADH em uma absorvância de 340 nm.

EXEMPLO 4 (PROFILÁTICO)

CLONAGEM E EXPRESSÃO DE BUTANODIOL DESIDRATASE

25

O propósito deste exemplo profilático é para descrever como clonar e expressar em *E. coli* os genes *pddA*, *pddB* e *pddC* que codifica a enzima butanodiol desidratase. Os genes *pddA*, *pddB* e *pddC* foram amplificados a partir

do DNA genômico de *Klebsiella oxytoca* ATCC 8724 usando PCR.

As seqüências *pddA*, *pddB* e *pddC* codificando butanodiol desidratase são clonadas e expressas da mesma maneira conforme descrito para *budA* no exemplo 1, exceto que o modelo genômico de DNA é de *Klebsiella oxytoca* ATCC 8724, e os iniciadores usados são B7 (SEQ ID nº 21) e B8 (SEQ ID nº 22). *Klebsiella oxytoca* DNA genômico é preparado usando o kit Gentra Puregene Puregene (Gentra Systems, Inc., Minneapolis, MN; catalog number D-5000A). Um produto PCR único incluindo todos os três quadros de leitura aberto (ORFs) é clonado, de modo que todas as três regiões de codificação são expressas como um operon a partir de um promotor único no plasmídeo de expressão. As seqüências de nucleotídeos dos quadros de leitura aberto (ORFs) para as três subunidades são dadas como SEQ ID nº 7, 9 e 11, respectivamente, e as seqüências de aminoácidos preditas das três subunidades da enzima são dadas como SEQ ID nº 8, 10 e 12, respectivamente.

A atividade de butanodiol desidratase nos extratos livres da célula é medida pela derivação do produto de cetona com 2,4-dinitrofenilhidrazina (DNPH). Resumidamente, 100 µL da mistura de reação, extrato de células contendo aproximadamente 0,0005 unidades de enzima, tampão de fosfato de potássio 40 mM (pH 8,0), 2 µg de adenosilcobalamina, 5 µg de 2,3,-butanodiol e 1 µg de albumina sérica bovina, são bruscamente arrefecidas pela adição de um volume igual de 0,05%, em peso, de DNPH em HCl 1,0 N. Depois de 15 minutos em temperatura ambiente, a cor é desenvolvido pela adição de 100 µL de NaOH 4 N. A quantidade do produto é determinada a partir da absorbância da solução final a 550 nm, em comparação com uma curva padrão preparada com 2-butanona. Todas as reações são executadas a 37°C sobre luz vermelha mortiça.

EXEMPLO 5 (PROFILÁTICO)

CLONAGEM E EXPRESSÃO DE BUTANOL DESIDROGENASE

O propósito deste exemplo profilático é para descrever como

clonar e expressar em *E. coli* o gene *sadh* que codifica butanol desidrogenase. O gene *sadh* foi amplificado a partir do DNA genômico de *Rhodococcus ruber* cepa 219 usando PCR.

A sequência *sadh* codificando butanol desidrogenase é clonada e expressa da mesma maneira conforme descrito para *budA* no exemplo 1, exceto que o modelo de DNA genômico é *Rhodococcus ruber* cepa 219 (Meens, Institut fuer Mikrobiologie, Universitaet Hannover, Hannover, Alemanha) e os iniciadores são B9 (SEQ ID nº 23) e B10 (SEQ ID nº 24). O DNA genômico do *Rhodococcus ruber* é preparado usando o Kit Ultra Clean™ Microbial DNA Isolation (MO BIO Laboratories Inc., Carlsbad, Califórnia), de acordo com o protocolo do fabricante. A sequência de nucleotídeos do quadro de leitura aberto (ORF - open reading frame) e da sequência de aminoácidos predita da enzima são dadas como SEQ ID nº 13 e SEQ ID nº 14, respectivamente.

A atividade da butanol desidrogenase em extratos livres de células é medida através do aumento da absorbância a 340 nm resultante da conversão do NADH para NAD quando a enzima é incubada com NAD e 2-butanol.

EXEMPLO 6 (PROFILÁTICO)

CONSTRUÇÃO DE UM VETOR DE TRANSFORMAÇÃO PARA OS

GENES EM UMA VIA BIOSSINTÉTICA DE 2-BUTANOL

O propósito deste exemplo profilático é para descrever a preparação de um vetor de transformação para os genes na via biossintética de 2-butanol (isto é, via 3 conforme descrito acima). Como a maioria dos organismos, *E. coli* converte glicose inicialmente para ácido pirúvico. As enzimas necessárias para converter ácido pirúvico a 2-butanol seguindo a via 3, isto é, acetolactato sintase, acetolactato descarboxilase, butanodiol desidrogenase, butanodiol desidratase e butanol desidrogenase, são codificados pelos genes *budA*, *budB*, *budC*, *pddA*, *pddB*, *pddC* e *sadh*. Para simplificar a construção da via biossintética de 2-butanol em um organismo recombinante, os genes que codificam as 5 etapas da via são

divididos em dois operons. A parte superior da via compreende as três primeiras etapas catalisadas por acetolactato sintase, acetolactato descarboxilase e butanodiol desidrogenase. A parte inferior da via compreende as duas últimas etapas catalisadas por butanodiol desidratase e butanol desidrogenase.

5 As seqüências de codificação são amplificadas pela PCR com os iniciadores que incorporam sítios de restrição para clonagem posterior e os iniciadores forward contêm um sítio de ligação otimizado no ribossomo de *E. coli* (AAAGGAGG). Os produtos PCR são Topo clonados no vetor pCR4Blunt-TOPO e transformados em células Top10 (Invitrogen). O DNA do Plasmídeo é
10 preparado a partir dos clones TOPO, e a seqüência do fragmento clonado PCR é verificada. As enzimas de restrição e T4 DNA ligase (New England Biolabs, Beverly, MA) são usadas de acordo com as recomendações do fabricante. Para experimentos de clonagem, os fragmentos de restrição são purificados por gel utilizando-se kit de extração QIAquick Gel (Qiagen).

15 Após a confirmação da seqüência, as regiões de codificação são subclonadas em um vetor pUC19 modificado como uma plataforma de clonagem. O vetor pUC19 é modificado por uma digestão com HindIII/SAPI, seguido de tratamento com polimerase Klenow DNA para preencher as extremidades. O fragmento de vetor 2,4 kB é purificado por gel e religado
20 criando pUC19dHS. Alternativamente, o vetor pUC19 é modificado por uma digestão com SphI/SAPI, seguido de tratamento com polimerase Klenow DNA para embotar as extremidades. O fragmento de vetor 2,4 kB é purificado por gel e religado criando pUC19dSS. A digestão remove o promotor lac adjacente ao MCS (sítios de clonagem múltiplas [multiple cloning sites]),
25 impedindo a transcrição de operons do vetor.

VIA SUPERIOR:

As regiões de codificação do *budABC* são amplificadas a partir de DNA genômico da *Klebsiella pneumoniae* por PCR usando pares de iniciadores

B11 e B12 (tabela 4), dadas como SEQ ID nº 25 e 26, respectivamente. Os iniciadores forward incorporam um sítio de restrição EcoRI e um sítio de ligação de ribossomo (RBS - ribosome binding site). O iniciador reverse incorpora um sítio de restrição SphI. O produto PCR é clonado em pCR4 Blunt-TOPO criando pCR4 Blunt-TOPO-budABC.

Para construir a via superior, operon pCR4 Blunt-TOPO-budABC é digerido com EcoRI e SphI liberando um fragmento *budABC* 3,2 kbp. O vetor pUC19dSS também é digerido com EcoRI e SphI, liberando um fragmento de vetor kbp 2,0. O fragmento *budABC* e os fragmentos do vetor são ligados juntos usando-se T4 DNA ligase (New England Biolabs) para formar pUC19dSS-budABC.

VIA INFERIOR:

As regiões de codificação *pddABC* são amplificadas a partir de DNA genômico de *Klebsiella oxytoca* ATCC 8724 por PCR usando iniciadores B13 e B14 (tabela 4), dadas como SEQ ID Nº s: 27 e 28, respectivamente, criando um produto 2,9 kbp. Os iniciadores forward incorporam sítios de restrição EcoRI e PmeI e um sítio de ligação de ribossomo (RBS - ribosome binding site). O iniciador reverse incorpora um sítio de restrição BamHI. O produto PCR é clonado em pCRBlunt II-TOPO criando pCRBluntII-pdd.

O gene *sadh* gene é amplificado a partir de DNA genômico de *Rhodococcus ruber* cepa 219 por PCR usando iniciadores B15 e B16 (tabela 4), dadas como SEQ ID Nº s: 29 e 30, respectivamente, criando um produto 1,0 kbp. Os iniciador forward incorpora um sítio de restrição BamHI e um sítio de ligação de ribossomo (RBS - ribosome binding site). O iniciador reverse incorpora um sítio de restrição XbaI. O produto PCR é clonado em pCRBlunt II-TOPO criando pCRBluntII-sadh.

Para construir o operon do percurso inferior, um fragmento de BamHI de 2,9 kbp EcoRI e pCRBluntII-pdd, um fragmento de 1,0 kbp BamHI e

XbaI a partir de pCRBluntII-sadh, e o fragmento maior a partir de um digerido EcoRI e XbaI de pUC19dHS são ligados juntos. A ligação de três vias cria pUC19dHS-pdd-sadh.

O vetor pUC19dSS-budABC é digerido com PmeI e HindIII, liberando um fragmento de 3,2 kbp que é clonado em pBenBP, um vetor de transformação *E.coli-B. subtilis*. O plasmídeo pBenBP é criado pela modificação do vetor pBE93, que é descrito por Nagarajan (WO 93/2463, exemplo 4). Para gerar pBenBP, a sequência de sinais de *Bacillus amyloliquefaciens* promotor protease neutra (NPR) e o gene *phoA* são removidos do pBE93 com um NcoI/HindIII digerido. O promotor NPR é PCR amplificado a partir de pBE93 por iniciadores BenF e BenBPR, dados pela SEQ ID nº 31 e 32, respectivamente. O iniciador BenBPR incorpora BstEII, PmeI e HindIII, sítios a jusante do promotor. O produto PCR é digerido com NcoI e HindIII, e o fragmento é clonado nos sítios correspondentes do vetor pBE93 para criar pBenBP. O fragmento superior do operon é subclonado nos sítios PmeI e HindIII em pBenBP criando pBen-budABC.

O vetor pUC19dHS-pdd-sadh é digerido com PmeI e HindIII, liberando um fragmento de 3,9 kbp que é clonado em sítios PmeI e HindIII de pBenBP, criando pBen-pdd-sadh.

EXEMPLO 7 (PROFILÁTICO)

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL EM *E. COLI*

O propósito deste exemplo profilático é para descrever como clonar e expressar a via biossintética de 2-butanol em *E. coli*.

Os plasmídeos pBen-budABC e pBen-pdd-sadh, preparados conforme descrito no exemplo 6, são separadamente transformados em *E. coli* NM522 (ATCC Nº 47000), e expressão dos genes em cada operon é monitorada pela análise SDS-PAGE e ensaio da enzima. Depois da confirmação da expressão de todos os genes, pBen-budABC é digerido com

108

EcoRI e HindIII para liberar o promotor NPR e o fragmento budABC. O fragmento tem a extremidade embotada usando o fragmento Klenow de polimerase DNA (New England Biolabs, catalog no. M0210S). O plasmídeo pBen-pdd-sadh é digerido com EcoRI e semelhantemente embotado para criar um linearizado fragmento de vetor de extremidade embotada. O vetor e os fragmentos NPR-budABC são ligados, criando p2BOH. O plasmídeo é transformado em *E. coli* NM522 para fornecer *E. coli* NM522/p2BOH, e a expressão dos genes é monitorada conforme anteriormente descrito.

E. coli NM522/p2BOH é inoculado em um frasco de mistura de 250 mL contendo 50 mL do meio e agitado a 250 rpm e 35°C. O meio é composto de: dextrose, 5 g/L; MOPS, 0,05 M; sulfato de amônio, 0,01 M; fosfato de potássio, monobásico, 0,005 M; S10 mistura de metais, 1% (v/v); extrato de levedura, 0,1% (peso por volume); ácidos casamino, 0,1% (peso por volume); tiamina, 0,1 mg/L; prolina, 0,05 mg/L; e biotina 0,002 mg/L, e é titulado para pH 7,0 com KOH. S10 mistura de metais contém: MgCl₂, 200 mM; CaCl₂, 70 mM; MnCl₂, 5 mM; FeCl₃, 0,1 mM; ZnCl₂, 0,1 mM; cloridrato de tiamina, 0,2 mM; CuSO₄, 172 µM; CoCl₂, 253 µM; e Na₂MoO₄, 242 µM. Depois de 18 h, 2-butanol é detectado por HPLC ou análise por cromatografia a gás usando métodos que são bem conhecidos na técnica, por exemplo, conforme descrito na seção de métodos geral acima.

EXEMPLO 8 (PROFILÁTICO)

EXPRESSÃO DA VIA BIOSSINTÉTICA DE 2-BUTANOL EM *E. COLI*

O propósito deste exemplo profilático é para descrever como clonar e expressar a via biossintética de 2-butanol em *Bacillus subtilis*.

Os plasmídeos pBen-budABC e pBen-pdd-sadh, preparado conforme descrito no exemplo 6, são separadamente transformados em *Bacillus subtilis* BE1010 (*J. Bacteriol.* 173:2278-2282 (1991)) e a expressão dos genes em cada operon é monitorada conforme descrito no exemplo 7. O plasmídeo pBen-budABC é digerido com EcoRI e HindIII para liberar o

fragmento do promotor NPR e budABC. O fragmento tem a extremidade embotada usando o fragmento Klenow de polimerase DNA (New England Biolabs, catalog no. M0210S). O plasmídeo pBen-pdd-sadh é digerido com EcoRI e semelhantemente embotado para criar um linearizado fragmento de
5 vetor de extremidade embotada. O vetor e os fragmentos NPR-budABC são ligados, criando p2BOH. O plasmídeo é transformado em *Bacillus subtilis* BE1010 para fornecer *Bacillus subtilis* BE1010/p2BOH, e a expressão dos genes é monitorada conforme anteriormente descrito.

O *Bacillus subtilis* BE1010/p2BOH é inoculado em um frasco de
10 mistura de 250 mL contendo 50 mL do meio e agitado a 250 rpm e 35°C por 18 horas. O meio é composto de: dextrose, 5 g/L; MOPS, 0,05 M; ácido glutâmico, 0,02 M; sulfato de amônio, 0,01 M; tampão de fosfato de potássio, monobásico, 0,005 M; S10 mistura de metais (conforme descrito no exemplo 7), 1% (v/v); extrato de levedura, 0,1% (peso por volume); ácidos casamino, 0,1%
15 (peso por volume); triptofano, 50 mg/L; metionina, 50 mg/L; e lisina, 50 mg/L, e é titulado para pH 7,0 com KOH. Depois de 18 h, 2-butanol é detectado por HPLC ou análise por cromatografia a gás usando métodos que são bem conhecidos na técnica, por exemplo, conforme descrito na seção de métodos geral acima.

EXEMPLO 9

CONSTRUÇÃO DE UM VETOR DE TRANSFORMAÇÃO PARA OS GENES EM UMA

VIA BIOSINTÉTICA DE 2-BUTANOL

O propósito deste exemplo foi preparar um hospedeiro recombinante *E. coli* transportando os genes em uma via biosintética de 2-butanol (isto é, via 3 conforme descrito acima). Como a maioria dos organismos, *E. coli* converte
25 glicose inicialmente para ácido pirúvico. As enzimas necessárias para converter ácido pirúvico a 2-butanona seguindo a via 3, isto é, acetolactato sintase, acetolactato descarboxilase, butanodiol desidrogenase e butanodiol desidratase, são codificados pelos genes *budA*, *budB*, *budC*, *pddA*, *pddB* e *pddC*. Na última

etapa da via, uma butanol desidrogenase converte 2-butanona para 2-butanol. As desidrogenases que executam esta última etapa são promíscuas e pode ser encontradas em muitos organismos. Para simplificar a construção da via biossintética de 2-butanol em um organismo recombinante, os genes que codificam as 5 etapas da via foram divididos em múltiplos operons. A parte superior da via compreende as três primeiras etapas catalisadas por acetolactato sintase, acetolactato descarboxilase e butanodiol desidrogenase e foi clonada em um vetor de expressão. A parte inferior da via compreende as duas últimas etapas catalisadas por butanodiol desidratase incluindo o fator de reativação (Mori et al., *J. Biol. Chem.* 272:32034 (1997)) e uma butanol desidrogenase. A diol desidratase pode sofrer inativação suicida durante catálise. A proteína codificada pelo fator de reativação *ddrA* e *ddrB* (GenBank AF017781, SEQ ID nº 70) reativa a enzima inativa. Os genes *ddrA* e *ddrB* flanqueiam o operon diol desidratase. Os operons para a desidratase/fator de reativação e o butanol desidrogenase foram ou clonados em outro vetor de expressão ou o operon desidratase/fator de reativação foi clonado isoladamente para outro vetor de expressão e a última etapa foi fornecida por uma atividade endógena no hospedeiro demonstrativo.

CONSTRUÇÃO DO VETOR PTrc99A-BUDABC:

As regiões de codificação *budAB* são amplificadas a partir de DNA genômico de *K. pneumoniae* ATCC 25955 por PCR usando o par de iniciadores BABC F e BAB R, dados como SEQ ID nº 33 e 34, respectivamente (vide tabela 4), criando um produto 2,5 kpb. Os iniciadores forward incorporam sítios de restrição *SacI* e *EcoRI* e um sítio de ligação de ribossomo (RBS - ribosome binding site). O iniciador reverse incorpora um sítio de restrição *SpeI*. O produto PCR foi clonado em pCR4 Blunt-TOPO criando pCR4 Blunt-TOPO-*budAB*. O plasmídeo de DNA foi preparado a partir de clones TOPO e a sequência dos genes foi verificada com os iniciadores forward M13 (SEQ ID nº 35), reverse M13 (SEQ ID nº 36), N83 SeqF2 (SEQ ID nº 37), N83 SeqF3 (SEQ

ID nº 38) e N84 SeqR4 (SEQ ID nº 39) (ver Tabela 5).

As região de codificação *budC* foi amplificada a partir de DNA genômico de *K. pneumoniae* ATCC 25955 por PCR usando o par de iniciadores BC Spe F e BC Xba R dados como SEQ ID nº 40 e 41, respectivamente, criando
5 um produto 0,8 kbp. O iniciador forward incorporou um sítio de restrição SpeI, um RBS e modificou o CDS mudando o segundo e terceiro códons de AAA para AAG. O iniciador reverse incorpora um sítio de restrição XbaI. O produto PCR foi clonado em pCR4 Blunt-TOPO criando pCR4 Blunt-TOPO-*budC*. O plasmídeo de DNA foi preparado a partir de clones TOPO e a sequência dos genes foi verificada
10 com os iniciadores forward M13 (SEQ ID nº 35) e reverse M13 (SEQ ID nº 36).

Para construir o operon *budABC*, pCR4 Blunt-TOPO-*budC* foi digerido com SnaBI e XbaI liberando um fragmento *budC* 1,0 kbp. O vetor pTrc99a (Amann et al., *Gene* 69(2):301-315 (1988)) foi digerido com SmaI e XbaI criando um fragmento de vetor linearizado 4,2 kbp. O vetor e o
15 fragmento *budC* foram ligados para criar pTrc99a-*budC* e transformados em células *E. coli* Top 10 (Invitrogen). Os transformantes foram analisados pela amplificação de PCR com iniciadores Trc F (SEQ ID nº 42) e Trc R (SEQ ID nº 43) para um produto 1,2 kbp para confirmar a presença do elemento de inserção *budC*. Os genes *budAB* foram subclonados a partir de pCR4 Blunt-
20 TOPO-*budAB* como um fragmento 2,5 kbp EcoRI/SpeI. O vetor pTrc99a-*budC* foi digerido com EcoRI e SpeI e o fragmento de vetor resultante 5,0 kbp foi purificado por gel. O vetor purificado e o elemento de inserção *budAB* foram ligados e transformados em células *E. coli* Top 10. Os transformantes foram testados pela amplificação PCR com iniciadores Trc F
25 (SEQ ID nº 42) e N84 Seq R2 (SEQ ID nº 65) para confirmar a criação de pTrc99a-*budABC*. Neste plasmídeo, as regiões de codificação *bud A*, *B* e *C* são adjacentes uma a outra, nesta ordem, e entre o promotor Trc e a sequência de terminação *rrnB*.

112

RESULTADOS

Foram examinadas três *E. coli* Top 10/pTrc99a-budABC independente e isolados para a produção de butanodiol, usando-se *E. coli* Top 10/pCL1925-Kodd-ddr (descritas abaixo) como controle negativo. As cepas foram cultivadas em meio LB contendo 100 µg/mL de carbenicillina. As células resultantes foram usadas para inocular frascos de agitação (aproximadamente 175 mL de volume total) contendo 125 mL de TM3a/meio de glicose com 100 µg/mL de carbenicillina. Além disso, os frascos inoculados com cepas transportam pTrc99a-budABC contendo isopropil β-D-1-tiogalactopiranosideo (IPTG) 0,4 mM. O meio de TM3a/glicose contém (por litro): 10 g de glicose, 13,6 g de KH₂PO₄, 2,0 g de monidrato de ácido cítrico, 3,0 g de (NH₄)₂SO₄, 2,0 g de MgSO₄.7H₂O, 0,2 g de CaCl₂.2H₂O, 0,33 g de citrato férrico de amônio, 1,0 mg de tiaminaHCl, 0,50 g de extrato de levedura, e 10 mL de solução de traços de elementos, pH ajustado para 6,8 com NH₄OH. A solução de traços de elementos contém: ácido cítrico.H₂O (4,0 g/L), MnSO₄.H₂O (3,0 g/L), NaCl (1,0 g/L), FeSO₄.7H₂O (0,10 g/L), CoCl₂.6H₂O (0,10 g/L), ZnSO₄.7H₂O (0,10 g/L), CuSO₄.5H₂O (0,010 g/L), H₃BO₃ (0,010 g/L), e Na₂MoO₄.2H₂O (0,010 g/L). Os frascos, tampados com tampas com respiro, foram inoculados a uma DO₆₀₀ inicial de aproximadamente 0,03 unidades e incubado a 34°C com agitação a 300 rpm.

Aproximadamente 23 h depois da indução, uma alíquota do meio foi analisada por HPLC (coluna Shodex Sugar SH1011) e cromatografia gasosa GC (HP-INNOWax), usando os mesmos métodos descritos na seção geral de métodos para 2-butanol e 2-butanona. Os resultados da análise são dados na tabela 6. Os três clones de *E. coli* converteram glicose para acetoina e *meso*-2,3-butanodiol, os intermediários desejados da via, com uma seletividade molar de 14%. Esta seletividade foi aproximadamente 35 vezes mais alta que a observada com a cepa de controle *E. coli* sem budABC.

113

TABELA 6
PRODUÇÃO DE ACETOÍNA E MESO-2,3-BUTANODIOL
POR *E. COLI* TOP 10/PTRC99A-BUDABC

Cepa	DO ₆₀₀	Acetoína, mM	Meso-2,3-butanodiol, mM	Seletividade molar ^a , %
Controle negativo	1,4	0,07	0,03	0,4
Isolado nº1	1,5	0,64	1,3	14
Isolado nº2	1,4	0,70	1,2	14
Isolado nº3	1,4	0,74	1,3	15

^a Seletividade molar é (acetoína + meso-2,3-butanodiol)/(glicose consumida).

5

CONSTRUÇÃO DE VETOR PCL1925-KoDD-DDR:

Os operons diol desidratase (GenBank D45071, SEQ ID nº 69) e o fator de reativação (GenBank AF017781, SEQ ID nº 70) foram amplificados PCR a partir de *Klebsiella oxytoca* ATCC 8724 como uma única unidade com iniciadores DDo para (SEQ ID nº 44) e DDo Rev (SEQ ID nº 45). Os iniciadores forward incorporam um RBS otimizado de *E.coli* e um sítio de restrição HindIII. O iniciador reverse inclui um sítio de restrição XbaI. O produto PCR 5318 bp foi clonado em pCR4Blunt-TOPO e clones do resultante pCR4Blunt-TOPO-Kodd-ddr foram seqüenciados com iniciadores M13 forward (SEQ ID nº 35), M13 reverse (SEQ ID nº 36), DDko seq F2 (SEQ ID nº 46), DDko seq F5 (SEQ ID nº 47), DDko seq F7 (SEQ ID nº 48), DDko seq F9 (SEQ ID nº 49), DDko seq R1 (SEQ ID nº 50), DDko seq R3 (SEQ ID nº 51), DDko seq R7 (SEQ ID nº 52), e DDko seq R10 (SEQ ID nº 53). Um clone tendo o elemento de inserção com a

seqüência esperada foi identificado.

Para expressão, a diol desidratase e os genes de fator de reativação foram subclonados em pCL1925 (Patent US Nº 7074608), uma baixa cópia de um plasmídeo transportando o promotor de glicose isomerase Streptomycyes. pCR4Blunt-TOPO-Kodd-ddr foi digerido com HindIII e XbaI e o resultante fragmento Kodd-ddr 5,3 kbp foi purificado por gel. O vetor pCL1925 foi digerido com HindIII e XbaI e o fragmento de vetor resultante 4539 bp foi purificado por gel. O vetor e o fragmento Kodd-ddr foram ligados e transformados em *E. coli* Top10. Os transformantes foram testados por PCR com iniciadores DDko Seq F7 (Nº SEQ ID:48) e DDko Seq R7 (SEQ ID nº 52). A ampliação do plasmídeo (pCL1925-Kodd-ddr) carregando o elemento de inserção resultou em um produto de aproximadamente 797 pb.

A atividade da diol desidratase em senso a *meso* -2,3-butanodiol foi medida pela incubação de extrato de células (proteína total ~ 0,8 mg/ml) com butanodiol coenzima B₁₂ 10 mM e 12 mM em HEPES 80 mM (pH 8,2) por 17 h à temperatura ambiente. A formação do esperado produto, 2-butanona, foi determinado por HPLC como descrito nos métodos gerais.

CONSTRUÇÃO DO VETOR PCL1925-KODD-DDR::T5 CHNA TER:

Para proporcionar uma atividade heteróloga da álcool desidrogenase, o gene *chnA* que codifica cicloexanol desidrogenase a partir de *Acinetobacter* sp. (Cheng et al., *J. Bacteriol.* 182:4744-4751 (2000)) foi clonado no vetor pCL1925 com o operon da diol desidratase, pCL1925-Kodd-ddr. O gene *chnA*, dado como SEQ ID nº 71 (Genbank Nº: AF282240, SEQ ID nº 73) foi amplificado a partir pDCQ2, um cosmídeo transportando gene aglomerado de cicloexanol a partir de *Acinetobacter*, com iniciadores ChnA F (SEQ ID nº 54) e ChnA R (SEQ ID nº 55). O produto resultante da PCR 828 bp foi clonado em pCR4Blunt-TOPO para criar pCR4Blunt-Topo-chnA e transformantes foram testados pela colônia PCR com os iniciadores forward

M13 (SEQ ID nº 35) e reverse M13 (SEQ ID nº 36). Os clones corretos produziram um produto PCR de cerca de 1 kbp e foram seqüenciados com os iniciadores forward M13 (SEQ ID nº 35) e reverse M13 (SEQ ID nº 36).

Após o seqüenciamento de pCR4Blunt-Topo-chnA para confirmar a
5 correta seqüência, o gene *chnA* foi subclonado como um plasmídeo a partir de um fragmento 813 bp MfeI/SmaI. O vetor de expressão pQE30 (Qiagen) foi digerido com MfeI e SmaI e o fragmento do vetor resultante 3350 bp foi purificado por gel. O fragmento *chnA* e o vetor purificado foram unidos e transformados em células *E. coli* Top10. Os transformantes da colônia PCR
10 foram testados com iniciadores chnSeq F1 (SEQ ID nº 56) e chnseq R1 (SEQ ID nº 57) para verificar a presença de um produto PCR 494 bp. Esta clonagem colocou o gene *chnA* sob o controle do promotor T5 no plasmídeo, pQE30-chnA.

Para preparar o vetor pCL1925 para transportar dois operons, terminadores foram adicionados ao vetor. Um fragmento de terminador tonB
15 terminador-mcs- trpA foi preparado por anelamento de oligonucleotídeo com iniciadores Top ter F1 (SEQ ID nº 58), Top terF2 (SEQ ID nº 59), Bot ter R1 (SEQ ID nº 60) e Bot ter R2 (SEQ ID nº 61). O DNA anelado foi purificado por gel em PAGE gel 6% (Embi-tec, San Diego, CA). O vetor pCL1925 foi digerido com SacI e XbaI e purificado por gel. O DNA anelado e o fragmento de vetor
20 foram unidos para criar pCL1925-ter. Os transformantes foram testados pela amplificação da colônia de PCR com iniciadores pCL1925 vec F (SEQ ID nº 62) e pCL1925 vec R 1 (SEQ ID nº 63) para verificar a presença de um produto PCR de aproximadamente 400 bp. Os clones positivos a partir do ensaio com PCR foram seqüenciados com os mesmos iniciadores.

25 O vetor pCL1925-ter foi digerido com XhoI e PmeI e o fragmento resultante 4622 bp foi purificado por gel. O pQE30-chnA foi digerido com NcoI e do DNA foi tratado com Klenow DNA polimerase para embotar as extremidades. O pQE30-chnA foi, então, digerido com XhoI e o fragmento

promotor *chnA* de vetor resultante 1,2 kbp T5 foi purificado por gel. O vetor pCL1925-ter e o fragmento de operon *chnA* e foram unidos para resultar em pCL1925-ter-T5*chnA* e transformados em *E. coli* Top10. Os transformantes foram testados pela amplificação da colônia de PCR com iniciadores pCL1925
5 vec F (SEQ ID nº 64) e *chnseq* R1 (SEQ ID nº 59) para verificar a presença de um produto de aproximadamente 1 kpb.

Para concluir a construção do vetor da via, o plasmídeo pCL1925-KoDD-ddr foi digerido com XbaI e SacI e o fragmento do vetor resultante 9504
bp foi purificado por gel. O operon *chnA* flanqueado por terminadores, com o
10 terminator *trpA* (Koichi et al. (1997) Volume 272, Number 51, pp. 32034-32041) 3' para a sequência de codificação *chnA*, a partir de pCL1925-ter-T5*chnA* foi purificada por gel como um fragmento 1271 bp XbaI/SacI. Depois da ligação dos fragmentos e transformação em *E. coli* Top10, os transformantes foram testados pela colônia PCR. Os iniciadores *chnSeq* F1 (SEQ ID nº 58) e
15 pCL1925 vec R2 (SEQ ID nº 64) amplificaram o esperado produto 1107 bp PCR no plasmídeo resultante, pCL1925-KoDD-ddr::ter-T5*chnA*.

EXEMPLO 10

EXPRESSÃO DE UMA VIA BIOSSINTÉTICA DE 2-BUTANOL EM *E. COLI* COM

SUPEREXPRESSÃO ÁLCOOL DESIDROGENASE ENDÓGENO

20 O propósito deste exemplo foi para expressar uma via bio sintética de 2-butanol em várias cepas *E. coli*.

CONSTRUÇÃO DE CEPAS DE *E. COLI* EXPRESSANDO CONSTITUTIVAMENTE YQH D:

E. coli contém um gene nativo (*yqhD*) que foi identificado como uma 1,3-propanodiol desidrogenase (patente U.S. Nº 6,514,733). O gene *yqhD*,
25 dado como SEQ ID nº 74, tem 40% de identidade com o gene *adhB* em *Clostridium*, uma butanol desidrogenase dependente provavelmente de NADH. O gene *yqhD* foi colocado sob a expressão constitutiva de uma variante do promotor 1,6GI de glicose isomerase (SEQ ID nº 67) em cepa de *E. coli*

117
SA

MG1655 1,6yqhD::Cm (WO 2004/033646) usando a λ Red technology (Datsenko e Wanner, *Proc. Natl. Acad. Sci. U.S.A.* 97:6640 (2000)). Semelhantemente, o promotor nativo foi substituído pelo promotor 1,5GI (WO 2003/089621) (SEQ ID nº 68), criando cepas MG1655 1,5yqhD:: Cm, assim, substituindo o promotor 1,6GI de MG1655 1,6yqhD:: Cm com o promotor 1,5GI. Os promotores 1,5GI e 1,6GI diferem por 1 pb na região -35, alterando, assim, a resistência dos promotores (WO 2004/033646). Com a substituição do promotor nativo *yqhD* pelo promotor 1,5GI ou 1,6GI, o gene *yqhC* que codifica os reguladores transcricionais putativos para o operon *yqh* é eliminado. A atividade enzimática de butanol desidrogenase foi confirmada pelo ensaio usando métodos que são bem conhecidas na técnica.

TRANSFORMAÇÃO DE CEPAS *E. COLI*:

Os plasmídeos da via pCL1925-Kodd-ddr e pTrc99a-budABC, descritas no exemplo 9, foram co-transformadas em cepas *E. coli* MG1655, MG1655 1,6yqhD e MG1655 1,5yqhD. As duas últimas cepas superexpressam a 1,3-propanodiol desidrogenase, YqhD, que tem também atividade de butanol desidrogenase. As cepas foram examinadas para a produção de 2-butanona e 2-butanol, essencialmente como descrito acima. As células foram inoculadas em frascos de agitação (volume total de cerca de 175 mL), contendo 50 ou 150 mL de TM3a/meio de glicose (com vitamina B₁₂ 0,1 mg/L e antibióticos adequados e IPT) para representar o meio e condições de baixas concentrações de oxigênio, respectivamente. Espectinomicina (50 mg/mL) e carbenicilina (100 mg/mL) foram utilizadas para plasmídeos pCL1925 -Kodd ddr e pTrc99a-budABC, respectivamente. Os frascos foram inoculados a uma DO₆₀₀ inicial de $\leq 0,04$ unidades e incubados a 34°C com agitação a 300 rpm. Os frascos contendo 50 mL do meio foram tampados com tampas com respiro; os frascos contendo 150 mL foram tampados com tampas sem respiro para minimizar troca de ar. IPTG estava presente no tempo zero a uma

concentração de zero a 0,04 mM. Os resultados analíticos para produção de 2-butanona e 2-butanol são apresentados na tabela 7. Todas as cepas de *E. coli* que compreendem uma via biossintética de 2-butanol produzem 2-butanona sob condições de baixas e médias concentrações de oxigênio e produzem 2-butanol sob condições de baixas concentrações de oxigênio.

TABELA 7

PRODUÇÃO DE 2-BUTANONA E 2-BUTANOL PELO *E. COLI* MG1655 CEPAS

HOSPEDEIRAS VIA PLASMÍDEOS PCL1925-KODD-DDR E PTRC99A-BUDABC

<u>Cepa</u> ^{a,b}	<u>IPTG,</u> <u>mM</u>	<u>volume de meio,</u> <u>mL</u>	<u>2-butanona,</u> <u>mM</u>	<u>2-butanol,</u> <u>mM</u>
MG1655 n°1	0	50	0,08	Não detectado
MG1655 n°2	0	50	0,11	Não detectado
MG1655 n°1	0,04	50	0,12	Não detectado
MG1655 n°2	0,04	50	0,11	Não detectado
MG1655 n°1	0	150	0,15	0,047
MG1655 n°2	0	150	0,19	0,041
MG1655 n°1	0,04	150	0,10	0,015
MG1655 n°2	0,04	150	0,11	0,015
MG1655 1,5yqhD n°1	0	50	0,10	Não detectado
MG1655 1,5yqhD n°2	0	50	0,07	Não detectado

113

<u>Cepa</u> ^{a,b}	<u>IPTG,</u> <u>mM</u>	<u>volume de meio,</u> <u>mL</u>	<u>2-butanona,</u> <u>mM</u>	<u>2-butanol,</u> <u>mM</u>
MG1655 1,5yqhD nº1	0,04	50	0,12	Não detectado
MG1655 1,5yqhD nº2	0,04	50	0,18	Não detectado
MG1655 1,5yqhD nº1	0	150	0,16	0,030
MG1655 1,5yqhD nº2	0	150	0,18	0,038
MG1655 1,5yqhD nº1	0,04	150	0,10	0,021
MG1655 1,5yqhD nº2	0,04	150	0,09	0,017
MG1655 1,6yqhD nº1	0	50	0,08	Não detectado
MG1655 1,6yqhD nº2	0	50	0,07	Não detectado
MG1655 1,6yqhD nº1	0,04	50	0,12	Não detectado
MG1655 1,6yqhD nº2	0,04	50	0,15	Não detectado
MG1655 1,6yqhD nº1	0	150	0,17	0,019

120
A

<u>Cepa</u> ^{a,b}	<u>IPTG,</u> <u>mM</u>	<u>volume de meio,</u> <u>mL</u>	<u>2-butanona,</u> <u>mM</u>	<u>2-butanol,</u> <u>mM</u>
MG1655 1,6yqhD nº2	0	150	0,18	0,041
MG1655 1,6yqhD nº1	0,04	150	0,11	0,026
MG1655 1,6yqhD nº2	0,04	150	0,11	0,038
controle (meio não inoculado)			Não detectado	Não detectado

a Nº 1 e Nº 2 representam isolados independentes.

b MG1655 é MG1655/pCL1925-Kodd-ddr/pTrc99a-budABC

MG1655 1,6yqhD é MG1655 1,6yqhD/pCL1925-Kodd-ddr/pTrc99a- budABC

MG1655 1,6yqhD é MG1655 1,5yqhD/pCL1925-Kodd-ddr/pTrc99a-budABC.

5

EXEMPLO 11

EXPRESSÃO DE UMA VIA BIOSSINTÉTICA DE 2-BUTANOL EM *E. COLI* COM

HETERÓLOGO ÁLCOOL DESIDROGENASE

Os plasmídeos pCL1925-KoDD-ddr:: ter-T5chnA e pTrc99a-budABC, descritos no exemplo 9, foram transformados em cepas de *E. coli* MG1655 e MG1655 Δ yqhCD para uma demonstração da produção de 2-butanol.

MG1655 Δ yqhCD carrega uma inativação yqhCD que foi feito usando o método de Datsenko e Wanner (*Proc. Natl. Acad. Sci. EUA.*

97(12):6640-6645 (2000)). Após a substituição da região com o cassete FRT-CmR-FRT de pKD3, o marcador de resistência a cloranfenicol foi removido usando a recombinase FLP. A sequência da região apagada é dada como SEQ ID nº 66.

5 As cepas MG1655/pTrc99a-budABC/pCL1925KoDD-ddr:: ter-T5 chnA e MG1655 Δ yqhCD/pTrc99a-budABC/pCL1925KoDD-ddr:: ter-T5 chnA foram examinadas para a produção de 2-butanona e 2-butanol, essencialmente como descrito acima. A cepa MG1655 Δ yqhCD/pCL1925
foi usada como controle negativo. As células foram inoculadas em frascos
10 de agitação (volume total de cerca de 175 mL), contendo 50 ou 150 mL de TM3a/meio de glicose (com vitamina B₁₂ 0,1 mg/L e antibióticos adequados) para representar o meio e condições de baixas concentrações de oxigênio, respectivamente. Espectinomomicina (50 µg/mL) e ampicilina (100 µg/mL) foram usados para a seleção de plasmídeos baseados em
15 pCL1925 e pTrc99a-budABC, respectivamente. Os derivados da atividade enzimática a partir de pTrc99a-budABC foram detectados pelo ensaio da enzima na ausência de indutor IPTG, assim, IPTG não foi adicionado ao meio. Os frascos foram inoculados a uma OD₆₀₀ inicial de $\leq 0,01$ unidades e incubados a 34°C com agitação a 300 rpm durante 24 h. Os frascos
20 contendo 50 mL do meio foram tampados com tampas com respiro; os frascos contendo 150 mL, foram tampados com tampas sem respiro para minimizar troca de ar. Os resultados analíticos para produção de 2-butanona e 2-butanol são apresentados na tabela 8. Ambas as cepas de *E. coli* compreendendo uma via biossintética de 2-butanol produzindo 2-
25 butanona sob condições de médias e baixas concentrações de oxigênio e produzindo 2-butanol sob condições de baixas concentrações de oxigênio, enquanto a cepa de controle negativo não produziram níveis detectáveis nem de 2-butanona nem de 2 -butanol.

TABELA 8**PRODUÇÃO DE 2-BUTANONA E 2-BUTANOL POR CEPAS DE *E. COLI***

Cepa ^a	Volume, mL	2-butanona, mM	2-butanol, mM
Controle negativo, MG1655 $\Delta yqhCD/pCL1925$	50	Não detectado	Não detectado
MG1655/pTrc99a-budABC/pCL1925KoDD- ddr::T5 chnA ter	50	0,33	Não detectado
MG1655 $\Delta yqhCD/pTrc99a$ - budABC/pCL1925KoDD-ddr::T5 chnA ter nº1	50	0,23	Não detectado
MG1655 $\Delta yqhCD/pTrc99a$ - budABC/pCL1925KoDD-ddr::T5 chnA nº2	50	0,19	Não detectado
Controle negativo, MG1655 $\Delta yqhCD/pCL1925$	150	Não detectado	Não detectado
MG1655/pTrc99a-budABC/pCL1925KoDD- ddr::T5 chnA ter	150	0,41	0,12
MG1655 $\Delta yqhCD/pTrc99a$ - budABC/pCL1925KoDD-ddr::T5 chnA nº1	150	0,15	0,46
MG1655 $\Delta yqhCD/pTrc99a$ - budABC/pCL1925KoDD-ddr::T5 chnA nº2	150	0,44	0,14
Meio		Não detectado	Não detectado

^a Nº 1 e Nº 2 representam isolados independentes.

EXEMPLO 12

CLONAGEM DE AMINO:PIRUVATO TRANSAMINASE (APT)

Uma amino:piruvato transaminase (APT) a partir de *Vibrio Fluvialis JS17* foi identificada por Shin et al. (Appl. Microbiol Biotechnol. (2003) 61:463-471). Foi descoberto que a sequência de aminoácidos (SEQ ID nº 122) tem significativa homologia com ω -aminoácido:piruvato transaminase (Shin e Kim (J. Org. Chem. 67:2848-2853 (2002)). Foi mostrado que o *Vibrio Fluvialis* APT tem atividade transaminase com respeito a acetoína.

Para expressão da enzima APT em *E. coli*, um códon otimizado APT região de codificação (SEQ ID nº 144) foi projetado usando os códons *E. coli* preferenciais com adicionais considerações como estabilidade de códon e mRNA, e sintetizado (por DNA2.0; Redwood City, CA). O fragmento da região de codificação do DNA foi subclonado no vetor pBAD.HisB (Invitrogen) entre os sítios NcoI e HindIII e o plasmídeo resultante, a partir de agora chamado de pBAD.APT1, foi transformada em células TOP10.

EXEMPLO 13

CARACTERIZAÇÃO DE *VIBRIO FLUVIALIS* APT ATIVIDADE DE ALANINA:ACETOÍNA AMINOTRANSFERASE

Um volume de 5 mL de meio LB + 100 µg/mL de ampicilina foi inoculado com uma colônia nova de células TOP10/pBAD:APT1. A cultura foi incubada a 37°C por aproximadamente 16 h com agitação (225 rpm). Uma alíquota de 300 µL desta cultura foi usada para inocular 300 mL do mesmo meio, que foi incubada a 37°C com agitação (225 rpm). Quando a cultura atingiu uma DO₆₀₀ de 0,8, L-arabinose foi adicionado para uma concentração final de 0,2% (peso por volume). A cultura foi incubada por mais 16 h e então colhida. As células foram lavadas uma vez com tampão de fosfato de potássio 100 mM (pH 7,8) e, então, congeladas e armazenadas a -80°C.

Para isolar a enzima, o pélete de célula foi descongelado e ressuspenso em 8 mL de tampão de fosfato de potássio 100 mM (pH 7)

contendo etilenodiaminotetraacetato 0, 2 mM, ditioneitol 1 mM e 1 tablete de coquetel inibidor de protease (Roche; Indianapolis, IN). As células foram lisadas por duas passagens através de uma prensa francesa (French Pressure Cell) a 6,2 MPa (900 psi), e o lisado resultante foi clarificado por centrifugação por 30 min a $17000 \times g$. Sulfato de amônio foi adicionado até 35% de saturação, e a solução foi agitada por 30 min à temperatura ambiente, e então sólidos precipitados removidos por centrifugação (30 min, $17000 \times g$). Mais sulfato de amônio foi adicionado ao sobrenadante para se obter 55% de saturação, e a solução foi novamente agitada por 30 min à temperatura ambiente. Os sólidos precipitados foram removidos por centrifugação (30 min, $17000 \times g$) e, então, ressuspensos em 5 mL de tampão de fosfato de potássio 100 mM (pH 7) contendo piridoxal 5'-fosfato 10 μ M e ditioneitol 1 mM. A solução foi desalinizada pela passagem através de uma coluna PD10 equilibrada com tampão A (tampão bis-tris propano 50 mM (pH 6) contendo piridoxal 5'-fosfato 10 μ M e ditioneitol 1 mM). O extrato desalinizado foi, então, colocado em uma coluna de fluxo 20 mL Q-Fast pre-equilibrada com tampão A. O APT foi eluída com um gradiente linear de NaCl 0-0,1 M em tampão A. A enzima foi detectada em frações eluídas pela presença de uma proteína de tamanho de banda $\sim 8,3E-20$ g (50 kD) quando analisada pela eletroforese em gel de SDS-poliacrilamida e pela absorbância característica a 418 nm. Frações contendo a enzima eluída em NaCl $\sim 0,3$ M. Estas frações foram coligadas para produzir um total de 6 mL de uma solução de 5,45 mg/mL de enzima, que foi >90% de pureza, como julgado pela eletroforese em gel de SDS-poliacrilamida .

A atividade alanina:acetoína aminotransferase de APT foi examinada usando um ensaio láctico desidrogenase acoplado. A mistura de reação contendo bis-tris propano 100 mM (pH 9,0), piridoxal 5'-fosfato 10 μ M, acetoína 0-50 mM, L-alanina 0 a 5 mM, enzima purificada 0,14 ou

0,28 mg/mL, NADH 200 μ M e láctico desidrogenase 20 U/mL (Sigma; Louis, MO). A reação foi seguida pela medida da mudança em absorbância a 340 nm, que indica a oxidação de NADH. Sob estas condições, a k_{cat}/K_m para acetoína foi 10 $M^{-1} s^{-1}$ e aquela por L-alanina foi 400 $M^{-1} s^{-1}$.

5 A identidade do produto esperado de 3-amino-2-butanol foi confirmado pela comparação a um padrão sintético. Uma mistura de (*R,R*)- e (*S,S*)-3-amino-2-butanol foi sintetizada pelo método de Dickey *et al.* [*J Amer Chem Soc* 74:944 (1952)]: 5 g de *trans*-2,3-epoxibutano foi agitado vagorosamente em 150 mL NH_4OH de frio (4°C). A reação foi aquecida
10 vagorosamente para temperatura ambiente, lacrada e agitada à temperatura ambiente por mais 10 dias. Então, o excesso de amônia e água e epoxibutano residual foram removidos por evaporação rotativa sob vácuo a 40°C. O óleo transparente (2,9 g) resultante foi ressuspenso em água para uma concentração de 10% (peso por volume). A produção do produto desejado foi confirmado por
15 análise por RMN e comparação entre o espectro e aquele relatado por Levy *et al.* [*Org. Magnetic Resonance* 14:214 (1980)]. Uma mistura dos correspondentes isômeros (*2R,3S*)- e (*2S,3R*)- foi produzido usando o método idêntico com a exceção que o material de partida foi o *cis*- isômero de 2,3-epoxibutano.

Um método analítico para detecção de 3-amino-2-butanol foi
20 desenvolvido baseado no método de derivatização de *o*- ftalodialdeído para determinação de aminoácidos relatado por Roth [*Anal. Chem.* 43:880 (1971)]. Uma alíquota de 200 μ L de 3-amino-2-butanol 1 mM (mistura de isômeros) foi misturada com 200 μ L de uma solução de borato 50 mM (pH 9,5), ao qual foi adicionado 10 μ L de 2-mercaptoetanol 5 μ L/mL em etanol e 10 μ L de *o*-
25 ftalodialdeído 10 mg/mL em etanol. A solução foi incubada em temperatura ambiente por 10 min, no momento em que o derivativo foi extraído em 200 μ L de hexano. O hexano foi separado da solução aquosa por decantação, e 10 μ L foram injetados em uma coluna HPLC Chiracel OD (Daicel Chemical Industries;

Fort Lee, NJ). A coluna foi isocraticamente executada com uma fase móvel de hexano 90:10: isopropanol a uma taxa de 1 mL/min. Os isômeros derivatizados de 3-amino-2-butanol foram detectados por absorbância a 340 nm com tempos de retenção de cerca de 15,7 e 16,8 min [(2S, 3S) e (2R, 3R)], e 18,4 e 21,9 min [(2R, 3S) e (2S, 3R)]. e 18,4 e 21,9 min [(2R,3S) e (2S,3R)]. Para diferenciar os enantiômeros na primeira mistura, o isômero puro (2R, 3R) (Bridge Organics; Vicksburg, MI) foi também executado sob condições idênticas e descoberto para ser o pico a 16,8 min. Para diferenciar os enantiômeros 3091 na segunda mistura, a mistura foi primeiro cineticamente resolvida usando a alanina: acetoína aminotransferase 0,28 mg de enzima purificada foi incubada com piruvato 10 mM e 3-amino-2-butanol 10 mM [mistura 1:1 de isômeros (2R, 3S) e (2S, 3R)] em 1 ml de Tris propano-bis 100 mM (pH 9,0). Após 24 h à temperatura ambiente, uma alíquota foi retirada e analisada conforme descrito acima. A análise revelou que o pico foi a 18,4 min foi 95% esgotado, enquanto que o pico a 21,9 min foi de >90% retido. Uma alíquota de 100 µL do remanescente da mistura de reação foi misturada com 50 µL, de NADH20 mM e 10 µL do extrato da cepa TOP10/pTrc99a-BudC descrita no exemplo 9. A enzima BudC é conhecida por reduzir (R) - acetoína para *meso*- 2,3-butanodiol e (S) -acetoína para (S, S) - 2,3-butanodiol [Ui et al.(2004) *Letters in Applied Microbiology* 39:533-537]. Após 3 h, as amostras foram tomadas a partir da reação e analisadas conforme descrito acima para acetoína e butanodiol. A análise indicou que o principal produto da redução foi *meso* -2,3-butanodiol, indicando que o produto da reação de aminotransferase foi (R) -acetoína e, por conseguinte, o isômero de 3-amino-2-butanol consumido foi o isômero (2R, 3S). Assim, o tempo de retenção de 18,4 min pode ser atribuído a este isômero e de 21,9 para o isômero (2S, 3R).

Para confirmar que o produto da reação catalisada por APT - alanina:acetoína aminotransferase foi o 3-amino-2-butanol, 0,28 mg de enzima

purificada foi incubada com acetoína 10 mM, L-alanina 10 mM, 50 U desidrogenase láctica e NADH 200 µM em 1 ml de bis-tris propano 100 mM (pH 9,0). A mistura de reação foi incubada à temperatura ambiente por 20 h, após o qual uma alíquota de 200 µL foi retirada e derivatizada como descrito acima.

- 5 Os tempos de retenção dos produtos derivatizados foram 15,8 min (produto principal) e 18,5 min (produto secundário), que correspondem aos dos padrões de (2S, 3S) - e (2R, 3S) - 3-amino-2-butanol.

EXEMPLO 14

IDENTIFICAÇÃO E CLONAGEM DE *ERWINIA CAROTOVORA* SUBESPÉCIE *ATROSEPTICA*

10

AMINOÁLCOOL QUINASE E AMINOÁLCOOL O-FOSFATO LIASE

O objetivo deste exemplo é descrever a identificação e clonagem de seqüências codificando uma quinase aminoácido e aminoálcool O-fosfato liase da bactéria *Erwinia carotovora*. Essas duas enzimas são parte da via 1 para a conversão de 3-amino-2-butanol a 2-butanona através do intermediário 3-amino-2-butanol fosfato, conforme mostrado na figura 1.

15

PREVISÃO DA *ERWINIA* AMINOÁLCOOL QUINASE E DO AMINOÁLCOOL O-FOSFATO-LIASE

As atividades de aminoálcool quinase dependente de ATP e de aminoálcool O -fosfato-liase têm sido detectadas em várias espécies de *Pseudomonas* e *Erwinia*, incluindo *Pseudomonas* sp. P6 (NCIB10431), 20 *Pseudomonas putida* NCIB 10558 (Jones et al. (1973) *Biochem. J.* 134:167-182), *Erwinia carotovora*, *Erwinia amanas*, *Erwinia milletiae*, e *Erwinia atroseptica* (Jones et al. (1973) *Biochem. J.* 134:959-968). Nestes estudos, os extratos das espécies supracitadas foram mostrados por terem atividade para a conversão enzimática de aminopropanol através de aminopropanol O -fosfato a propionaldeído, e para a conversão de etanolamina através de etanolamina O- 25 fosfato a acetaldeído.

A seqüência do genoma da cepa *Erwinia atroseptica* em que estas atividades foram reportadas a existir (agora designada como de *Erwinia*

carotovora subespécie *atroseptica* cepa SCRI1043 (ATCC BAA-672)) tem sido determinada no Sanger Institute (Bell et al. *Proc. Natl. Acad. Sci. USA* 101(30): 11105-11110). Análise das quinases putativas na *Erwinia carotovora* subespécie *atroseptica* genoma revelou uma seqüência de operons (SEQ ID nº 275) que codificam uma proteína putativa (ECA2059; SEQ ID nº 124) que é 39% idêntica para uma *Rhizobium loti* homoserina quinase e uma aminotransferase (ECA2060; Nº SEQ ID: 126) dependente de piridoxal-fosfato (*PLP-pyridoxal phosphate*) classe III putativo que é 58% idêntica a uma aminotransferase putativa a partir de *Rhizobium meliloti*. Esperava-se que ECA2059 fosse uma aminoálcool quinase e que ECA2060 fosse uma aminoálcool O -fosfato-liase, que usasse PLP como cofator.

CLONAGEM DO AMINOÁLCOOL QUINASE PUTATIVA E AMINOÁLCOOL O-FOSFASE

LIASE PUTATIVA A PARTIR DE ERWINIA CAROTOVORA SUBESPÉCIE ATROSEPTICA

O DNA genômico de *Erwinia carotovora* subespécie *atroseptica* (ATCC #: BAA-672D) foi obtido a partir da American Type Culture Collection (ATCC). O operon que codifica a aminoálcool quinase (KA) putativa e aminoálcool O -fosfato-liase (AT), foi nomeado KA-TA (SEQ ID nº . 275. Este operon foi amplificado a partir do DNA genômico de *Erwinia* pela Phusion DNA polimerase (Finnzymes; via New England Biolabs; Ipswich, MA) usando iniciadores OT872 (SEQ. ID. Nº 127) e OT873 (SEQ. ID. Nº 128). Um fragmento de DNA de 2,4 kb foi obtido pela reação de PCR, o que corresponde ao tamanho do operon KA-AT. O produto de PCR foi digerido com as enzimas de restrição EcoRI e PstI, e clonado no vetor pKK223-3 (Amersham Biosciences; Piscataway, NJ) o qual foi digerido com as mesmas enzimas de restrição. Este produziu plasmídeo pKK223.KA-AT, que contém a *Erwinia* aminoálcool quinase putativa e operon liase sob controle do promotor *tac*. Semelhantemente, plasmídeos pKK223.KA e pKK223.AT foram feitos, que colocaram as regiões de codificação da putativa *Erwinia* quinase e da putativa

129
SA

Erwinia liase em vetores separados, cada sob o controle do promotor *tac*. Para a clonagem de PCR da região de codificação KA (SEQ ID nº 123), iniciadores OT872 (SEQ. ID. Nº 127) e OT879 (SEQ. ID. Nº 129) foi usada; e para a clonagem de PCR da região de codificação AT (SEQ ID nº 125), iniciadores OT873 (SEQ. ID. Nº 128) e OT880 (SEQ. ID. Nº 130) foram usados nas amplificações da PCR, que geraram produtos PCR de 1,1 kb e 1,3 kb, respectivamente. Os produtos de PCR foram cada digeridos com EcoRI e PstI, e ligados no vetor pKK223-3 para gerar pKK223.KA e pKK223.AT.

A ATIVIDADE IN VIVO DO AMINOÁLCOOL QUINASE PUTATIVA E AMINOÁLCOOL O-

FOSFATO LIASE PUTATIVA A PARTIR DE *ERWINIA CAROTOVORA* SUBESPÉCIE

ATROSEPTICA

Os plasmídeos pKK223.KA-AT, pKK223.KA, pKK223.AT e pKK223-3 foram transformados na cepa *E. coli* MG1655. Os transformantes foram reaplicados por esfregaço em uma placa com meio mínimo (*minimal media*) MOPS contendo 1% de glicose, 0,5% de aminopropanol como uma única fonte de nitrogênio, IPTG 1 mM e 100 µg/mL de ampicilina. A expressão de genes KA-AT, KA e AT foram induzidos pelo IPTG. Uma placa de controle não tinha IPTG incluído. As placas foram incubadas a 37°C por 7 dias. Na placa com IPTG, somente a cepa MG1655/pKK223.KA-AT cresceu, enquanto todas as outras três cepas não cresceram. Na placa sem IPTG adicionado, a cepa MG1655/pKK223.KA-AT cresceu, mas as colônias foram significativamente menores que aquelas na placa contendo IPTG, que corresponde aos níveis de expressão mais baixos de KA evAT nas células não induzidas. Nenhuma das outras três cepas cresceu sobre esta placa. Isso indica que a co-expressão dos genes putativos KA e AT de *Erwinia* forneceu suficiente atividade enzimática que permitiu a cepa *E. coli* MG1655/pKK223.KA-AT usar aminopropanol como uma única fonte de nitrogênio. Expressão da de cada enzima individual, quer KA ou AT não foi

suficiente para fornecer essa atividade enzimática in vivo.

EXEMPLO 15

ATIVIDADE IN VITRO DA *ERWINIA* PUTATIVA AMINOÁLCOOL QUINASE E

AMINOÁLCOOL O-FOSFATO LIASE

5 SUBCLONAGEM DO OPERON *ERWINIA* KA-AT NO VETOR pBAD.HisB E NA INDUÇÃO

DA EXPRESSÃO PROTÉICA

Os níveis de expressão de proteína de enzimas *Erwinia* putativas KA e AT expressas em células MG1655 do vetor pKK223.KA-AT foram analisadas por análise SDS-Page. O nível de expressão da enzima *Erwinia* AT foi relativamente
10 baixo, com uma nova proteína banda detectada no correto peso molecular de 7,8-20 g (46 kD) na fração solúvel de um extrato da célula, enquanto não era detectada nova proteína banda no tamanho previsto para a enzima KA.

Em um esforço para melhorar a expressão dos genes putativos de *Erwinia* KA e AT, o operon KA-AT foi subclonado nos sítios EcoRI e HindIII do
15 vetor pBAD.HisB-EcoRI. A pBAD.HisB-EcoRI foi derivada do vetor pBAD.HisB (Invitrogen), pela substituição do sítio NcoI no pBAD.HisB por um sítio EcoRI através de QuickChange mutagênese sítio-dirigida (Stratagene, La Jolla, CA) usando os iniciadores OT909 (SEQ ID. nº 131) & OT910 (SEQ ID. nº 132). No plasmídeo construído pBAD.KA-AT, o operon KA-AT foi colocado diretamente
20 sob controle do promotor *araB* (sem sua identificação).

O plasmídeo pBAD.KA-AT foi transformado na cepa *E. coli* TOP10. 50 mL da cultura da cepa TOP10/pBAD.KA-AT foi cultivada para meados da fase log ($DO_{600} = 0,6$) em LB, meio de ampicilina 100 µg/mL, a 37°C com agitação a 250 rpm. A cultura foi induzida através da adição de L-
25 arabinose para uma concentração final de 0,1% (peso por volume), e foi ainda incubada a 37°C durante 5 h antes da colheita por centrifugação. O pélete da célula foi ressuspenso em Tris-HCl 50 mM congelado, pH 8,0, e atingidos por sonicação sobre o gelo com um Fischer Sonic Model 300 Dismembrator

(Fischer, Pittsburgh, PA) a 50% de potência, repetindo quatro ciclos de sonicação por 30 segundos com 60 segundos de repouso entre cada ciclo. Cada amostra após a sonicação foi centrifugada ($15000 \times g$, 4 min, 4°C). Os extratos clarificados livres da célula foram analisados por nível de expressão de proteína e atividade de aminoálcool O- fosfato liase.

SÍNTESE QUÍMICA DE AMINOBTANOL O-FOSFATO E AMINOPROPANOL O-FOSFATO

O substrato de (*R, R*)-3-amino-2-butanol O -fosfato foi sintetizado por um método baseado no que foi reportado por Ferrari e Ferrari (patente US 2730542 [1956]) para fosfoetanolamina: 10 mmol de H_3PO_4 em uma solução aquosa 50% (peso por volume) foi misturado com uma solução 50% (peso por volume) de (*R,R*)-3-amino-2-butanol (Bridge Organics; Vicksburg, MI) com agitação em gelo. Depois de misturada, a solução foi lentamente aquecida à temperatura ambiente e, em seguida, agitada sob vácuo e aquecida a 70°C . Depois de 1 h a 70°C , a temperatura foi lentamente aumentada para 185°C e mantida ali por mais 2 h. Então, a reação foi resfriada à temperatura ambiente e liberada do vácuo. O restante do material foi dissolvido em água, e a análise por RMN indicou que 80% do material inicial foi convertido em produto com 20% permanecendo não-reagido. Produtos adicionais não foram observados.

Os substratos adicionais de (*2R, 3S*) -3-amino-2-butanol O -fosfato e (*2S, 3R*) -- 3-amino-2-butanol O -fosfato foram sintetizados pelo mesmo procedimento utilizando uma mistura 1:1 de (*2R, 3S*) -3-amino-2 -butanol e (*2S, 3R*) -3-amino-2-butanol (sintetizada como descrito em exemplo 13) como o material de partida. DL -1-amino-2-propanol O -fosfato, (*S*) - 2-amino-1-propanol O- fosfato, e (*R*) -2-amino-1-propanol O-fosfato foram sintetizados pelo mesmo procedimento utilizando DL -1-amino-2-propanol, (*R*) -2-amino-1-propanol ou (*S*) -2-amino-1-propanol como o material de partida.

ANÁLISE DA ATIVIDADE DE AMINOPROPANOL O-FOSFATO LIASE CODIFICADA PELO

OPERON PUTATIVO ERWINIA KA-AT

O ensaio aminopropanol O -fosfato-liase foi realizado conforme descrito por Jones et al. (1973, *Biochem. J.* 134:167-182) e Gori G. et al. (1995, *Chromatographia* 40:336). A formação de aminopropanol a partir de propionaldeído O -fosfato foi testada colorimetricamente com MBTH, que permite a detecção de formação de aldeído. A reação foi realizada da seguinte forma. Em um 1 mL da reação, 100 mg de extrato livre de células de *E. coli* TOP10/pBAD.KA-AT foi adicionado para DL -1-amino-2-propanol O -fosfato 10 mM em Tris-HCl 100 mM, pH 7,8, com PLP 0,1 mM. A reação foi incubada a 37°C por 10 min e 30 min, com uma alíquota de 100 µL de mistura da reação removida em cada ponto de tempo e misturada com 100 µL de MBTH-HCl 6 mg/mL em glicina 375 mM, pH 2,7. Essa mistura foi incubada a 100°C por 3 min, resfriada em gelo durante 15 a 30 s, e 1 mL de FeCl₃.6 H₂O 3,3 mg/mL (em HCl 10 mM) foi adicionado, seguido de incubação por 30 minutos em temperatura ambiente. A absorbância da mistura de reação que contém o aldeído-MBTH aduto, foi medida a 670 nm. Os resultados do ensaio são indicados na tabela 9. Na presença do substrato de fosfato aminopropanol PLP e extrato livre de células, formação de aldeído foi detectada, conforme indicado por um Abs₆₇₀ que era mais elevado do que o controle do fundo de até 0,3. Na ausência de um deles, do substrato ou do extrato livre de célula, a formação de aldeído não foi detectada. Na ausência do PLP adicionado, uma quantidade um pouco menor do aldeído foi detectada, provavelmente devido à presença de PLP no extrato livre de célula. O extrato livre de célula da cultura TOP10/pBAD.KA-AT não-induzida não produziu qualquer aldeído detectável na reação. Estes resultados indicaram que a *Erwinia* aminoálcool O- fosfato liase putativa catalisa a conversão de aminopropanol O -fosfato para propionaldeído.

133
A**TABELA 9.**

ENSAIO DE AMINOPROPANOL O-FOSFATO LIASE. A AMOSTRA 1 FOI A CÉLULA SEM O EXTRATO DO CONTROLE NÃO INDUZIDO DE *E. COLI* TOP10/pBAD.KA-AT. As AMOSTRA 2 A 5 CONTINHAM A CÉLULA LIVRE DE EXTRATO DA CULTURA INDUZIDA *E. COLI* TOP10/pBAD.KA-AT.

Número da amostra	Indução por 0,1% arabinose	Aminopropanol O-fosfato	PLP	Extrato de enzima (100 mg/mL)	DO ₆₇₀ , 10 min	DO ₆₇₀ , 30 min
1	Não induzido	(+)	(+)	(+)	0,262	0,255
2	induzido	(+)	(+)	(+)	1,229	2,264
3	induzido	(-)	(+)	(+)	0,303	0,223
4	induzido	(+)	(-)	(+)	0,855	1,454
5	induzido	(+)	(+)	(-)	0,156	0,065

ANÁLISE DA ATIVIDADE DE ERWINIA AMINOÁLCOOL O-FOSFATO LIASE COM RESPEITO AO SUBSTRATO DE AMINOBUTANOL O-FOSFATO

A atividade da aminoálcool O- fosfato liase rumo à aminobutanol O -fosfato de substratos foi estudada em condições idênticas às descritas acima. A reação foi conduzida a 37°C de um dia para o outro em 1 mL da reação que continha 100 microgramas de extrato livre da célula de *E. coli* TOP10/pBAD.KA-AT, aminobutanol O -fosfato 10 mM (quer seja a mistura de (R, R) + (S, S) ou a mistura de isômeros (R, S) + (S, I) descritas no exemplo 15), em Tris -HCl 100 mM, pH 7,8, com PLP 0,1 mM. As alíquotas de 100 µL da mistura de reação foram retiradas e o produto 2-butanona foi detectado usando o método de derivatização MBTH descrito em métodos gerais. Os dois picos representando os isômeros derivatizado de 2-butanona foram observados. Portanto, a *Erwinia* aminoálcool O -fosfato-liase é um aminobutanol fosfato

fosfo-liase além de uma aminopropanol fosfato fosfo-liase.

**A ANÁLISE DA ATIVIDADE DA ERWINIA O ÁLCOOL AMINO-FOSFATO-LIASE NO
SENTIDO DE ESTEREOISÔMEROS DE AMINOPROPANOL O-FOSFATO E
AMINOBTANOL O-FOSFATO.**

5 A atividade da *Erwinia* aminoálcool O -fosfato-liase para vários estereoisômeros de aminopropanol O -fosfato e aminobutanol O -fosfato foi estudada em condições idênticas às descritas acima. Na presença da *Erwinia* aminoálcool O-fosfato liase, ambos (R) e (S)-2-amino-1-propanol O-fosfato foram convertidos para propanona pela enzima, mas o produto produzido era
10 muito mais alto com o isômero (S). A enzima também produziu butanona a partir de ambas as misturas de isômeros de 3-amino) - 2-butanol O -fosfato, com um produto produzido mais elevado encontrado na reação contendo os isômeros dos substratos (R, S) e (S, I). Ambos os produtos butanona e propanona foram derivatizados por MBTH, e detectados por HPLC como
15 descrito em métodos gerais.

**OTIMIZAÇÃO DO NÍVEL DE EXPRESSÃO GÊNICA PARA ERWINIA AMINOÁLCOOL
QUINASE E AMINOÁLCOOL O-FOSFATO-LIASE**

 A fim de melhorar a expressão de níveis para a *Erwinia* aminoálcool quinase e aminoálcool O -fosfato-liase em *E. coli*, as regiões de
20 codificação códon otimizadas para ambas as enzimas (chamado EKA: SEQ ID nº 155 e EAT: SEQ ID nº 156 respectivamente) foram sintetizadas por DNA2.0 (Redwood City, CA). Cada região de codificação foi sintetizada com caudas de 5' e 3' incluindo sítios de restrição para clonagem: EKA tem sítios 5' BbsI e 3' EcoRI, HindIII; EAT tem sítios 5' EcoRI e 3' HindIII. As regiões de codificação
25 EKA e EAT foram fornecidas a partir de DNA2.0 como plasmídeos pEKA e pEAT, que estavam no vetor pJ51 de DNA2.0. A região de codificação da EKA otimizada foi subclonada pela união de um fragmento digerido BbsI e HindIII de pEKA no vetor pBAD.HisB entre os sítios NcoI e HindIII, para gerar plasmídeo

pBAD.EKA. No plasmídeo resultante a região codificante é 5' para a sua identificação, então uma região de codificação para uma amino-terminal His₆ identificação unida à *Erwinia* aminoálcool quinase foi construída através da execução de uma QuickChange reação mutagênese dirigida utilizando a um sítio
5 usando iniciadores N° SEQ ID: 157 e N° SEQ ID: 158 para gerar o vetor pBAD.His-EKA.

pBAD.His-EKA foi transformada em cepa *E. coli* BL21AI (*F^{ompT hsdSB (rB⁻ mB⁻) gal dcm araB::T7RNAP-tetA}*; Invitrogen) para produzir cepa BL21AI/pBAD.HisA-EKA. 50 mL de uma cultura BL21-AI/pBAD.HisA-
10 EKA foi cultivada para o meio da fase log (DO₆₀₀ = 0,6), induzida com arabinose 0,1%, e incubada adicionalmente a 30°C de um dia para o outro. Os extratos livres de célula foram preparados por sonicação. A proteína de fusão marcada com His₆ de *Erwinia* aminoálcool quinase foi purificada usando o Probond™ Sistema de Purificação (Invitrogen), sob condições de
15 purificação não-desnaturante seguindo as instruções do fabricante.

RESULTADO PROFILÁTICO

A atividade quinase da *Erwinia* aminoálcool quinase marcada com His₆ - é analisada pelo ensaio ADP Quest (DiscoverX, Fremont, CA), seguindo as instruções do fabricante. Este é um ensaio bioquímico que mede
20 o acúmulo de ADP, um produto da reação aminoálcool quinase usando quer aminopropanol ou aminobutanol como substrato. O substrato 10 mM é misturado com a *Erwinia* aminoálcool quinase marcada com His₆, em Tris-HCl 100 mM, pH 7,8, MgCl 10 mM₂, KCl 2 mM, ATP 0,1 mM, e incubado a 37°C por 1 h em 0,2 mL da reação. O reagente ADP A (100 µL) e reagente
25 ADP B (200 µL) são adicionados e a mistura é incubada em temperatura ambiente por 30 min. Um sinal fluorescente indicando atividade é medido com comprimento de onda de excitação de 530 nm e comprimento de onda de emissão de 590 nm.

EXEMPLO 16**EXPRESSÃO DA INTEIRA VIA 3****CONSTRUÇÃO DO VETOR PCL1925-KoDD-DDR:**

O vetor pTrc99a::BudABC (descrito no exemplo 9) é digerido com
5 EcoRI, e o DNA é tratado com polimerase Klenow DNA para embotar as
extremidades. O vetor embotado é, subseqüentemente, digerido com SpeI para
produzir um fragmento 2,5 kb contendo os genes *budA* e *budB*. O vetor
pCL1925-ter-T5chnA (descrito no exemplo 9) é digerido com HindIII, e o DNA foi
tratado com Klenow DNA polimerase para embotar as extremidades. O vetor
10 embotado é, subseqüentemente, digerido com XbaI para produzir um fragmento
4,6 kb que é então ligado para o fragmento *budAB* de pTrc99a::BudABC. O
plasmídeo resultante, chamado pCLBudAB-ter-T5chnA, é usado para
transformar células de *E. coli* Top10, e colônias únicas são então testadas para
estrutura de plasmídeo própria por PCR usando iniciadores pCL1925vecF (SEQ
15 ID nº 62) e N84seqR3 (SEQ ID nº 159). O plasmídeo é preparada a partir de
uma única colônia que produz um produto PCR do tamanho esperado de 1,4 kb.

CONSTRUÇÃO DO VETOR PKK223.KA-AT-APT

O gene APT é amplificado a partir do vetor pBAD.APT (descrito no
exemplo 12) por PCR usando iniciadores APTfor (SEQ ID nº 162; 5' incluindo
20 RBS e sítio SmaI) e APTrev (SEQ ID nº 163; 3' adicional sítio SmaI). O produto
do tamanho esperado de 1,7 kbp é purificado por gel e digerido com SmaI para
produzir extremidades embotadas. O vetor PKK223.KA-AT (descrito no exemplo
14) é digerido com PstI, e o DNA foi tratado com Klenow DNA polimerase para
embotar as extremidades. O fragmento de DNA resultante é ligado com o
25 produto PCR digerido por SmaI, e o produto de ligação é usado para transformar
células de *E. coli* Top10. As colônias Individuais resistentes à ampicilina são
testadas para PCR usando iniciadores OT872 (SEQ ID nº 127) e APTrev (SEQ
ID nº 163). A presença de um produto do PCR do tamanho esperado de 4,1 kbp

indica que o gene codificando APT está presente e orientado na mesma direção dos genes que codificam KA e AT. A sequência dos elementos de inserção é verificada usando os iniciadores APTseqRev (SEQ ID nº 160) e APTseqFor (SEQ ID nº 161). Este plasmídeo é chamado pKK223.KA-AT-APT. A expressão
5 própria de todos os três genes é verificada pelo crescimento de 5 mL de uma cultura de Top10/pKK223.KA-AT-APT em LB + 100 µg/mL de ampicilina a 37°C, com agitação. Quando a DO₆₀₀ atinge ~ 0,8, a expressão dos genes no plasmídeo é induzida pela adição de IPTG a 0,4 mM. A expressão é avaliada pela SDS-PAGE e ensaios de atividade conforme descrito acima.

10 CONSTRUÇÃO DE CEPAS DE PRODUÇÃO DE 2-BUTANOL E PRODUÇÃO DE 2-BUTANONA E 2-BUTANOL

E. A cepa de E. coli MG1655 é transformada com ambos pKK223.KA-AT-APT e pCLBudAB-ter-T5chnA, e transformantes selecionados pela resistência a ampicilina e espectinomicina, indicativos da presença dos
15 plasmídeos. As células são inoculadas nos frascos de agitação (aproximadamente 175 mL de volume total) contendo 50 ou 150 mL de TM3a/meio de glicose (com antibióticos adequados) para representar o meio e condições com pouco oxigênio, respectivamente. 0,4 mM de IPTG é adicionado para induzir expressão de genes a partir de pKK223.KA-AT-APT. Como um controle negativo, células de MG1655
20 são cultivadas no mesmo meio, mas sem antibióticos. Os frascos são inoculados a uma DO₆₀₀ inicial de ≤ 0,01 e incubados a 34°C com agitação a 300 rpm por 24 h. Os frascos contendo 50 mL do meio são tampados com tampas com respiro; os frascos contendo 150 mL, foram tampados com tampas sem respiro para minimizar troca de ar. A cepa 1655/pKK223.KA-AT-APT/pCLBudAB-ter-T5chnA
25 que compreende uma biossintética de 2-butanol produz ambos 2-butanona e 2-butanol sobre condições de oxigênio em baixas e médias concentrações enquanto a cepa de controle negativo não produz níveis detectáveis nem de 2-butanona nem de 2-butanol.

EXEMPLO 17CARACTERIZAÇÃO DE ATIVIDADE DE GLICEROL DESIDRATASE BUTANODIOLDESIDRATASE

Glicerol desidratase (E.C. 4.2.1.30) e diol desidratase (E.C. 4.2.1.28), enquanto relacionadas estruturalmente, são freqüentemente distinguidas na técnica à base de várias diferenças, que incluem especificidade de substrato. Este exemplo demonstra que glicerol desidratase converte *meso*-2,3-butanodiol para 2-butanona. A cepa recombinante *E. coli* KLP23/pSYCO12, que compreende genes de *Klebsiella pneumoniae* codifica as múltiplas subunidades da glicerol desidratase (alfa: SEQ ID nº 145 (região de codificação) e 146 (proteína); beta SEQ ID nº 147 (região de codificação) e 148 (proteína); e gama SEQ ID nº 149 (região de codificação) e 150 (proteína)) e genes de *Klebsiella pneumoniae* codificando as múltiplas subunidades de glicerol desidratase reativase (subunidade grande, SEQ ID nº 151 (região de codificação) e 152 (proteína); e pequena subunidade, SEQ ID nº 153 (região de codificação) e 154 (proteína)), é descrito em Emptage et al. US 6.514.733 e no WO 2003089621, os quais estão aqui incorporados, a título de referência. Uma célula bruta, sem extrato de KLP23/pSYCO12 foi preparada pelos métodos conhecidos pelo versado na técnica. O ensaio da enzima foi realizado na ausência de luz em tampão 80 mM HEPES, pH 8,2 a 37°C com 12 µM coenzima B₁₂ e 10 mM de *meso*-2,3-butanodiol. A formação de 2-butanona foi monitorada pela HPLC (coluna Shodex SH-1011 e SH-G coluna de proteção com índice de refração detecção; H₂SO₄ 0,01 M como a fase móvel a um fluxo de 0,5 mL/min e uma temperatura da coluna de 50°C; tempo de retenção de 2-butanona = 40,2 min). A taxa de formação de 2-butanona pela preparação de glicerol desidratase foi determinada para ser 0,4 nmol/min/mg de proteína bruta.

EXEMPLO 18**ANÁLISE ESTRUTURAL DE DIOL/GLICEROL DESIDRATASES ATRAVÉS DE GERAÇÃO E****VALIDAÇÃO DE UM PERFIL HMM PARA DIOL/GLICEROL DESIDRATASE****COMPROVADOS EXPERIMENTALMENTE**

5 As enzimas diol desidratase e glicerol desidratase pertencem às classes de enzima 4.2.1.28 e 4.2.1.30, respectivamente. Cada uma das enzimas em ambas as classes são um complexo de três subunidades: grande (também chamado de alfa), médio (também chamado de beta) e pequeno (também chamado de gama). Foi descoberto que em algumas glicerol
10 desidratases, as subunidades grande e média são unidas.

IDENTIFICANDO MEMBROS DA FAMÍLIA PELA SEQÜÊNCIA

A enzima butanodiol desidratase de *Klebsiella oxytoca* foi usada como uma enzima protótipo para identificar uma família de enzimas de diol e glicerol desidratases. As seqüências de aminoácidos da alfa (GenBank Nº: BAA08099; SEQ ID nº 8), beta (GenBank Nº: BAA08100; SEQ ID nº 10) e gama (GenBank Nº: BAA08101; SEQ ID nº 12) cada uma das subunidades foi executada conforme uma seqüência de consulta em uma pesquisa BLASTp na base de dados não-redundantes de proteína utilizando parâmetros pré-definidos. As seqüências com similaridades relevantes foram extraídas. A relevância foi
15 julgada pela pontuação E-valor, definição de proteína, detalhes incluídos no relatório GenBank para as proteínas combinadas e revisão da literatura sobre o assunto. Para a subunidade grande, a saída BLAST mostrou uma queda abrupta no E-Valor de e-20 para um E-Valor de 1,5. Todas as seqüências comparadas com um E-Valor de 1,5 ou maior tinham definições inconsistentes com elas
20 sendo desidratases. Muitas destas seqüências foram identificadas como subunidade beta de polimerase RNA dirigida a DNA. Existiram comparações compatíveis com E-valores cerca de e-20, que foram seqüências parciais. As seqüências sem anotação foram incluídas se o E-Valor foi menor que 1,5.

Usando a *Klebsiella oxytoca* butanodiol desidratase subunidade alfa como uma referência, 50 homólogos foram identificados como membros desta família de proteínas. Este grupo incluiu algumas seqüências que não são proteínas de comprimento total. As seqüências protótipo de comprimento total identificadas para a família de subunidade alfa de diol/glicerol desidratases são as SEQ ID nº 8 e SEQ ID nº 93, 99, 105, 135, 138, 141, 146, 164, 167, 170, 173, 176, 179, 182, 185, 188, 191, 194, 197, 200, 203, 206, 209, 212, 215, 218, 221, 224, 227, 130, 243, 254, 255, 256, 257, 258 e 259. SEQ ID nº 233, 235, 237, 239, 241, 246, 247 incluem ambas as subunidades alfa e beta, que são unidas nestes casos.

Usando a *Klebsiella oxytoca* butanodiol desidratase subunidade beta como uma referência, 51 homólogos foram identificados como membros desta família de proteínas. Este grupo incluiu algumas seqüências que não são proteínas de comprimento total. As seqüências protótipo de comprimento total identificadas para a família de subunidade beta de diol/glicerol desidratases são as SEQ ID nº 10 e SEQ ID nº 95, 101, 107, 136, 139, 142, 148, 165, 168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 222, 225, 228, 231, 244, 250, 252, 260, 261, 262, 263, 364, 265, 266 e 167.

Usando a *Klebsiella oxytoca* butanodiol desidratase subunidade gama como uma referência, 48 homólogos foram identificados como membros desta família de proteínas. Este grupo incluiu algumas seqüências que não são proteínas de comprimento total. As seqüências protótipo de comprimento total identificadas para a família de subunidade gama de diol/glicerol desidratases são as SEQ ID nº 12 e SEQ ID nº 97, 103, 109, 137, 140, 143, 150, 166, 169, 172, 175, 178, 181, 184, 187, 190, 193, 196, 199, 202, 205, 208, 211, 214, 217, 220, 223, 226, 229, 232, 234, 236, 238, 240, 242, 245, 248, 249, 251, 253, 268, 270, 271, 272, 273 e 274.

IDENTIFICAÇÃO DE MEMBROS DA FAMÍLIA COM FUNÇÃO EXPERIMENTALMENTE AVALIADA

Para cada sequência identificado através da análise descrita acima, uma busca por evidência experimental de sua função bioquímica foi conduzida nas bases de dados de BRENDA, UniProt e NCBI Entrez. BRENDA é uma base de dados revisados por humanos (*human-curated*) que contém informações detalhadas sobre propriedades cinéticas, físicas e bioquímicas da enzima, extraídas da literatura experimental e com links para as bases de dados relevantes (Cologne University Bioinformatics Center). A UniProt Knowledgebase [reservatório de conhecimento] é composta de uma parte revisada por humanos, a base de dados Swiss-Prot e de um machine annotated supplement [suplemento de anotações por máquina], a base de dados TrEMBL. A base de dados curada Swiss-Prot (Swiss Institute of Bioinformatics [Instituto Suíço de Bioinformática]) fornece um alto nível de anotação de proteína que inclui arquitetura de domínio, modificações pós-traducionais e variantes de sequência. A NCBI Entrez é o sistema integrado de recuperação, com pesquisa baseada em texto, usado no NCBI (National Center for Biotechnology Information, Bethesda, MD) pelos grandes bancos de dados, incluindo a PubMed, seqüências de nucleotídeo e proteínas, estruturas de proteína, genomas completos e Taxonomia.

Através da análise de informações e referências identificadas a partir destas bases de dados, oito diol/glicerol desidratases com função experimentalmente verificada como um diol ou glicerol desidratase foram identificados. Estes são mostrados na tabela 10.

TABELA 10

DIOL/GLICEROL DESIDRATASES COM FUNÇÃO EXPERIMENTALMENTE VERIFICADA

Organismo	Subunidade	GenBank nº	SEQ ID nº	Tipo	Referência
Klebsiella oxytoca	grande	gi 6980836	8	diol	Shibata et al., Structure 1999; 7:997-1008
K. oxytoca	média	gi 6980837	10	diol	

Organismo	Subunidade	GenBank nº	SEQ ID nº	Tipo	Referência
K. oxytoca	pequena	gi 6980838	12	diol	
Klebsiella pneumoniae	grande	gi 4063702	105	diol	Tobimatsu et al., 1998; Biosci. Biotechnol. Biochem. 62:1774-1777
K. pneumoniae	média	gi 94470233	107	diol	
K. pneumoniae	pequena	gi 4063704	109	diol	
Clostridium pasteurianum	grande	gi 3360389	135	glicerol	Macis et al., FEMS Microbiol Lett. 1998; 164(1):21-8
C.pasteurianum	média	gi 3360390	136	glicerol	
C.pasteurianum	pequena	gi 3360391	137	glicerol	
Escherichia blattae	grande	gi 60099613	138	glicerol	Sönke et al. J. of Mol. Micro. e Biotech. 2004; 8:150-168
E. blattae	média	gi 57340191	139	glicerol	
E. blattae	pequena	gi 57340192	140	glicerol	
Klebsiella pneumoniae	grande	gi 24158719	146	glicerol	Willard, Thesis (1994), U of Wisconsin-Madison
K. pneumoniae	média	gi 24158720	148	glicerol	
K. pneumoniae	pequena	gi 24158721	150	glicerol	
Citrobacter freundii	grande	gi 1169287	141	glicerol	Seyfried, Gottschalk; J. Bacteriol. 178:5793-5796 (1996)
C. freundii	média	gi 1229154	142	glicerol	
C. freundii	pequena	gi 1229155	143	Glicero l	
Lactobacillus brevis	grande	gi 116334196	164	diol	Schuetz e Radler 1984; Arch. Microbiol. 139, 366-370
L. brevis	média	gi 116334195	165	diol	
L. brevis	pequena	gi 116334194	166	diol	
Lactobacillus collinoides	grande	gi 18857678	99	diol	Sauvageot et al., 2002; Eur J Biochem. 269(22):5731-7.
L. collinoides	média	gi 18857679	101	diol	
L. collinoides	pequena	gi 18857680	103	diol	

O conjunto de 8 aminoácidos de cada subunidade de diol/glicerol

143
AP

desidratases com função experimentalmente determinada, registradas na tabela 10, foram comparadas através de um alinhamento múltiplo de sequência usando ClustalW com parâmetros pré-estabelecidos. A percentagem de identidade com a subunidade grande das seqüências variou de 97,6% a 58,4%. A percentagem de identidade com a subunidade média das seqüências variou de 89,5% a 41,7%. A percentagem de identidade com a subunidade pequena das seqüências variou de 83,3% a 36,4%. Portanto, a quantidade de identidade de seqüência entre algumas subunidades das seqüências foi muito baixa (como 36,4% e 41,7%), mesmo que estas subunidades fossem conhecidas como componentes de enzimas conhecidas através de dados experimentais para executar a mesma função. O baixo nível de percentagem de identidade das seqüências tornou inviável a utilização deste critério de correlações de estrutura/função.

RELAÇÃO DAS SEQÜÊNCIAS DE DIOL/GLICEROL DESIDRATASES VERIFICADOS

EXPERIMENTALMENTE PARA OUTROS DIOL/GLICEROL DESIDRATASES

Para realizar esta análise, seqüências altamente redundantes que são >95% idênticas foram retiradas do conjunto de seqüências para subunidades grandes, médias ou pequenas, com exceção de que todas as seqüências de função experimentalmente verificada foram mantidas. As seqüências de proteína parciais ou truncadas também foram removidas. Um alinhamento de seqüência múltiplo foi realizado nas seqüências restantes usando ClustalW com parâmetros pré-estabelecidos. A percentagem de identidade com as subunidades grandes das seqüências variou de 97,6% (a alta percentagem é de multiplas seqüências verificadas experimentalmente) a 42,8%. A percentagem de identidade com as subunidades pequenas das seqüências variou de 91,9% a 26,4%. A percentagem de identidade com as subunidades pequenas das seqüências variou de 85,2% a 20,5%. Estas percentagens de identidades têm faixas semelhantes às percentagens de identidades para as seqüências verificadas experimentalmente.

144

Com base nos múltiplos alinhamentos da sequência, as árvores filogenéticas foram construídas utilizando o algoritmo "neighbor-joining" (como implementado no MEGA pacote de software versão 3.1; Kumar et al., 2004 *Briefings in Bioinformatics* 5:150-163.) As árvores filogenéticas são mostrados nas figuras 2 (subunidade grande), 3 (subunidade média), e 4 (subunidade pequena), com as identidades das seqüências mapeadas listados em uma chave para cada figura. Conforme pode ser observado a partir das posições marcadas para as seqüências de função verificada experimentalmente (em círculos cinza claro e cinza para diol desidratase e glicerol desidratase, respectivamente), estas seqüências estão espalhadas pela maioria árvore. Entretanto, cada árvore não inclui uma ramificação sem membros verificados experimentalmente, mas que parecem pertencer à família diol/glicerol desidratase.

CONSTRUINDO UM PERFIL HIDDEN MARKOV MODEL (HMM) DA FAMÍLIA DE DIOL/GLICEROL-DESIDRATASE COM BASE NO CONJUNTOS DE OITO SEQÜÊNCIAS DE SUBUNIDADES

Uma caracterização alternativa da estrutura/função dos conjuntos de subunidades da família de diol/glicerol desidratase de enzimas foi realizada utilizando o pacote de software HMMER (A teoria que apoia o perfil HMMs é descrita em R. Durbin, S. Eddy, A. Krogh e G. Mitchison, *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, Cambridge University Press, 1998; Krogh et al., 1994; *J. Mol. Biol.* 235:1501–1531), seguindo o manual do usuário que está disponível através de HMMER (Janelia Farm Research Campus, Ashburn, VA).

Cada conjunto de sequência de 8 aminoácidos para as subunidades grandes, médias e pequenas da diol/glicerol desidratases caracterizada funcionalmente (na Tabela 10) foi analisada separadamente usando o programa de software HMMER. A saída do programa de software HMMER é um Perfil Hidden Markov Model (HMM) que caracteriza as seqüências

de entrada. Como indicado no manual do usuário, os perfis HMMs são modelos estatísticos de alinhamentos múltiplos de sequência. Eles capturam informações específicas de posição sobre como conservada cada coluna do alinhamento é, e quais resíduos de aminoácidos são mais prováveis de ocorrer em cada posição.

- 5 Assim HMMs têm uma base formal probabilística. Os Perfis HMMs para um grande número de famílias de proteínas estão disponíveis ao público na base de dados PFAM (Janelia Farm Research Campus, Ashburn, VA).

Cada perfil HMM foi construído como se segue:

ETAPA 1. CONSTRUIR UM ALINHAMENTO DE SEQÜÊNCIA

- 10 As oito seqüências para a subunidade grande da diol/glicerol desidratases verificada funcionalmente (SEQ ID N °: 8, 99, 105, 135, 138, 141, 146 e 164) foram alinhadas usando Clustal W com parâmetros pré-estabelecidos. O mesmo foi feito para o conjunto de seqüências de subunidade média (SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165) e o conjunto de seqüências de
- 15 pequena subunidade (SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166).

ETAPA 2. CONSTRUIR UM PERFIL HMM

- O programa Hmmbuild foi executado em cada conjunto de seqüências alinhadas usando os parâmetros pré-estabelecidos. O hmmbuild lê o arquivo de alinhamento múltiplo de seqüência, constrói um novo perfil
- 20 HMM e salva o arquivo de perfil HMM. Usando este programa um perfil não calibrado foi gerado a partir do alinhamento múltiplo para cada conjunto das seqüências de subunidades acima descritas.

- As informações seguintes, baseadas no guia do usuário do software HMMER, apresentam algumas descrições da forma que o programa hmmbuild
- 25 prepara um perfil HMM. Um perfil HMM é capaz de modelar alinhamentos com vãos, por exemplo, incluindo inserções e supressões, que permitem que o software descreva um domínio conservado completo (em vez de apenas um pequeno grupamento sem gaps). As inserções e supressões são modeladas

usando estados de inserção (I) e estados de supressão (D). Todas as colunas que contêm mais de uma determinada fração de caracteres de gap x serão atribuídas como uma coluna de inserção. Como um padrão, x é ajustado para 0,5. Cada estado de compatibilidade tem um estado I e um D associados a ele. O HMMER nomeia um grupo de três estados (M/D/I) na mesma posição consensual do alinhamento um "nó". Esses estados são interligados com setas chamadas *probabilidades do estado de transição*. Os estados M (compatibilidade[Match]) e I são emissores, enquanto que os estados D são silenciosos. As transições são dispostas de modo que em cada nó, quer o estado M é usado (e um resíduo seja alinhado e pontuado) ou o estado D é utilizado (e nenhum resíduo é alinhado, resultando num carácter de vão de supressão, '-'). As inserções ocorrem entre "nós", e os estados I têm uma auto-transição, permitindo que um ou mais resíduos inseridos ocorram entre as colunas de consenso.

A pontuação de resíduos em um estado de compatibilidade (ou seja, comparar as pontuações do estado de emissão), ou em um estado de inserção (ou seja, inserir as pontuações do estado de emissão) são proporcionais a $\text{Log}_2(p_x) / (\text{null}_x)$. Onde p_x é a probabilidade de um resíduo de aminoácidos, em uma posição particular no alinhamento de acordo com o perfil HMM e null_x é a probabilidade de acordo com o modelo NULL. O modelo Null é um modelo probabilístico simples de um estado com o conjunto de probabilidades de emissões pré-calculadas para cada um dos 20 aminoácidos derivados da distribuição de aminoácidos na SWISSPROT release 24.

As pontuações dos estados de transição são também calculadas como log de parâmetros ímpares e são proporcionais para $\text{Log}_2(t_x)$. Onde t_x é a probabilidade de transitar de um estado emissor para um não emissor.

ETAPA 3. CALIBRAR O PERFIL HMM

Cada Perfil HMM foi lido usando `hmmcalibrate` que pontua um grande número de seqüências aleatórias sintetizadas com o perfil (o número

147
SA

padrão de seqüências sintéticas usadas é 5000), ajusta uma distribuição de valores extremos (EVD-extreme value distribution) para o histograma destas pontuações e salva novamente o arquivo HMM incluindo agora os parâmetros EVD. Estes parâmetros EVD (μ e λ) são usados para calcular os E-values de pontuação "bit" quando o perfil é pesquisado em uma base de dados de seqüências de proteínas. O `hmmcalibrate` escreve dois parâmetros no arquivo HMM em uma linha marcada como "EVD": estes são os parâmetros μ (localização) e λ (escala), parâmetros de uma distribuição de valores extremos (EVD) que melhor ajusta um histograma das pontuações calculadas em seqüências geradas aleatoriamente com aproximadamente o mesmo comprimento e a mesma composição de resíduos de SWISS-PROT. Esta calibração foi feita uma vez para cada perfil HMM.

Os perfis HMM calibrado para as subunidades grandes, médias e pequenas dos conjuntos de seqüências são fornecidas no apêndice como as tabelas Excel Perfil HMM alfa, Perfil HMM beta e Perfil HMM gama. Cada Perfil HMM é fornecido em uma tabela que indica a probabilidade de cada aminoácido ocorrer em cada posição na seqüência de aminoácidos. A maior probabilidade é destacada para cada posição. A tabela 11 mostra algumas linhas do Perfil HMM preparado para as subunidades grandes de diol/glicerol desidratases com função verificada experimentalmente.

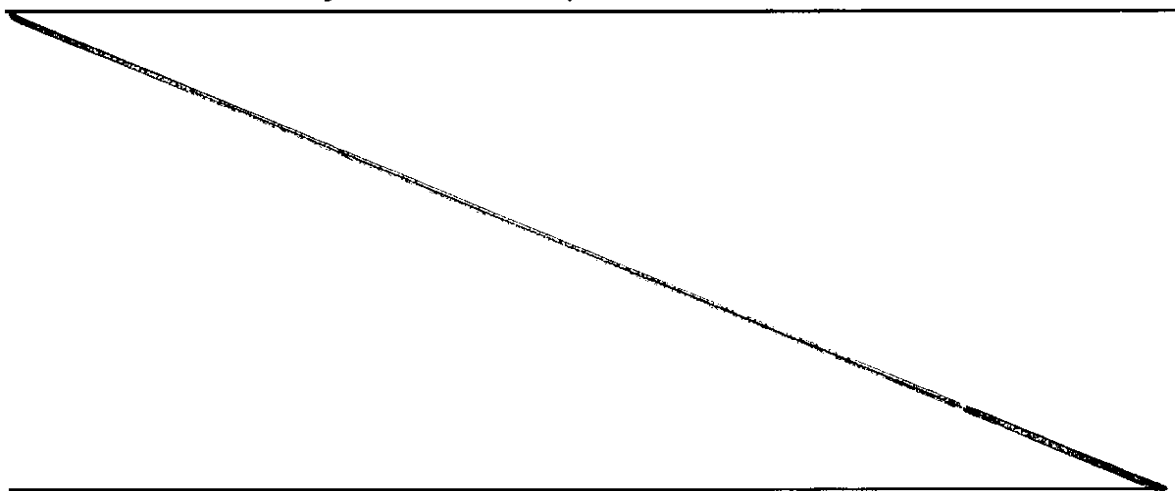


TABELA 11

UMA PORÇÃO DA SUBUNIDADE GRANDE DO PERFIL HMM.

HMM	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
	m>m	m>i	m>d	i>m	i>i	d>m	d>d	b>m	m>e											
	-585	*	-1585																	
1(M)	-965	-982	-2125	-1844	-276	-2022	-1302	373	-1390	610	4141	-1701	-2229	-1356	-1425	-1444	-1028	126	-1279	-880
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249
-	-33	-6045	-7087	-894	-1115	-701	-1378	-585	*											
2(K)	-603	-1732	-469	811	-2182	-1397	205	-1770	1954	-1654	-880	-59	-1465	629	1868	-495	-471	-1439	-1719	-1294
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249
-	-350	-6045	-2321	-894	-1115	-701	-1378	*	*											
3(R)	-781	-1268	-994	-512	-1638	-1313	-151	-1479	778	-1459	-913	-539	-1557	136	3077	-831	-760	-1263	-1437	-1152
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249
-	-41	-5736	-6778	-894	-1115	-535	-1690	*	*											

148

Os aminoácidos são representados por códigos de uma letra.

A primeira linha para cada posição relata as pontuações de emissão compatíveis: probabilidade para cada aminoácido estar naquele estado (pontuação mais elevada é realçada). A segunda linha relata as pontuações de emissão de inserção e a terceira linha relata sobre as pontuações em transição de estados: $M \rightarrow M$, $M \rightarrow I$, $M \rightarrow D$; $I \rightarrow M$, $I \rightarrow I$; $D \rightarrow M$, $D \rightarrow D$; $B \rightarrow M$; $M \rightarrow E$.

A tabela 11 mostra que, para as subunidades grandes, metionina tem uma probabilidade de 4141 de estar na primeira posição, a mais elevada probabilidade que é realçada. Na segunda posição lisina tem a maior probabilidade, que é 1954. Na terceira posição arginina tem a maior probabilidade, que é 3077.

ETAPA 4. TESTAR A ESPECIFICIDADE E SENSIBILIDADE DO PERFIL CONSTRUÍDO HMMs

Cada perfil HMM foi avaliado usando hmmsearch, que lê um perfil HMM de hmmfile e pesquisa o arquivo da seqüência em busca de seqüências compatíveis significativamente similares. O arquivo de seqüência pesquisado foi a base de dados de proteínas não redundante GenBank. O tamanho da base de dados (parâmetro Z) foi ajustado para 1 bilhão. Este ajuste de tamanho garante que os E-valores significativos na base de dados corrente continuarão a ser significativos no futuro previsível. O valor de corte de E-Valor foi ajustado para 10.

O perfil HMMs para as subunidades grandes, médias e pequenas de diol/glicerol desidratases com função verificada experimentalmente foi específica nas únicas subunidades de diol/glicerol desidratases que foram recuperadas, conforme indicado pela anotação das seqüências coincidentes e sensível nas seqüências parciais pares de subunidades de diol/glicerol desidratases que foram recuperadas. Cada uma das seqüências recuperadas tinha um E-valor de 0,01 ou menos.

Todas as seqüências nas árvores filogenéticas nas figuras 2, 3 e 4 foram recuperadas na correspondência de perfil HMM. Todas as seqüências

150
A

nas ramificações das árvores que não contenham seqüências com função experimentalmente verificada foram pareadas. Assim, todas as diol e glicerol desidratases estavam ligados à 8 diol e glicerol desidratases com função experimentalmente verificada através de correspondência com o perfil HMMs para as subunidades grandes, médias ou pequenas das enzimas. As subunidades de diol e glicerol desidratase de comprimento total que correspondem ao perfil HMMs têm os seguintes SEQ ID N °:

Subunidades (alfa) grandes: 8, 93, 99, 105, 135, 138, 141, 146, 164, 167, 170, 173, 176, 179, 182, 185, 188, 191, 194, 197, 200, 203, 206, 209, 212, 215, 218, 221, 224, 227, 130, 243, 254, 255, 256, 257, 258 e 259.

Subunidades grandes + médias unidas (a porção da subunidade grande e da subunidade média coincidem com o perfil grande e o perfil média, respectivamente): 233, 235, 237, 239, 241, 246 e 247.

Subunidades (beta) médias: 10, 95, 101, 107, 136, 139, 142, 148, 165, 168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 222, 225, 228, 231, 244, 250, 252, 260, 261, 262, 263, 364, 265, 266 e 167.

Subunidades (gama) pequenas: 12, 97, 103, 109, 137, 140, 143, 150, 166, 169, 172, 175, 178, 181, 184, 187, 190, 193, 196, 199, 202, 205, 208, 211, 214, 217, 220, 223, 226, 229, 232, 234, 236, 238, 240, 242, 245, 248, 249, 251, 253, 268, 270, 271, 272, 273 e 274.

Esta análise mostra que o perfil HMM para cada subunidade, que foi preparada usando seqüências com função verificada experimentalmente, fornece uma estrutura que está ligada à função de enzimas de diol/glicerol desidratase. Comparando todas as seqüências acima para os perfis HMM, por sua vez, fornece um link estrutura/função para estas seqüências.

TABELA 12

HMMER2.0 [2.3.2]	Nome e versão do programa																			
NOME alpha_exp_seqs	Nome do arquivo de entrada de alinhamento de sequência																			
LENG 557	Comprimento do alinhamento: Inclui indels																			
ALPH Amino	Tipo de resíduos																			
MAP sim	Mapa dos estados correspondentes às colunas do alinhamento																			
COM hmmbuild alpha.hmm alpha_exp_seqs.aln	Comandos usados para gerar o arquivo: este significa que hmmbuild (parâmetros pré-estabelecidos) foi aplicado para o arquivo de alinhamento																			
COM hmmscalibrate alpha.hmm	Comandos usados para gerar o arquivo: este significa que hmmscalibrate (parâmetros pré-estabelecidos) foi aplicado para o arquivo de alinhamento																			
NSEQ 8	Número de seqüências no arquivo de alinhamento																			
DATA Fri Mar 30 19:05:15 2007	Quando o arquivo foi gerado																			
XT -8455 -4 -1000 -1000 -8455 -4 -8455 -4																				
NULT -4 -8455																				
NULO 595 -1558 85 338 -294 453 -1158 197 249 902 -1085 -142 -21 -313 45 531 201 384 -1998 -644	A distribuição da probabilidade de transição para o modelo nulo (estado único G).																			
O símbolo para distribuição da probabilidade de emissão para o modelo nulo (estado G): (por exemplo 4 ou 20) números inteiros. A probabilidade de null usada para converter estas de volta às probabilidades do modelo é 1/K.																				
EVD -264.989197 0,112643	A distribuição de valores extremos dos parâmetros μ e λ , respectivamente, ambos valores de ponto flutuante. λ é positivo e não nulo. Estes valores são ajustados quando o modelo é calibrado com hmmscalibrate.																			
HMM	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
	m->m	m->i	m->d	i->m	i->i	d->m	d->d	b->m	m->e											
	-585	*	-1585																	
1(M)	-965	-982	-2125	-1844	-276	-2022	-1302	373	-1390	610	4141	-1701	-2229	-1356	-1425	-1444	-1028	126	-1279	-880
																				1

150

151

-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	46	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	-585	*												
2(K)	-603	-1732	-469	811	-2182	-1397	205	-1770	-1954	-1654	-860	-59	-1465	629	1868	-495	-471	-1439	-1719	-1294	2
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-350	-6045	-2321	-894	-1115	-701	-1378	*	*												
3(R)	-781	-1268	-994	-512	-1638	-1313	-151	-1479	778	-1459	-913	-539	-1557	136	3077	-831	-760	-1263	-1437	-1152	3
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-41	-5736	-6778	-894	-1115	-535	-1690	*	*												
4(Q)	-149	-1077	25	226	-1650	-945	-85	-1408	309	-1505	-735	33	-1280	-1816	-34	1644	-225	-1020	-1731	-1155	4
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
5(K)	-1028	-1686	-800	-442	-2171	-1521	-263	-1873	3103	-1859	-1251	-547	-1750	74	599	-1004	-972	-1620	-1812	-1528	5
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
6(R)	-1181	-1614	-1431	-903	-2068	-1632	-446	-1944	-559	-1867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	6
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
7(F)	-1602	-1340	-2512	-2455	3740	-2315	-379	-478	-2225	-175	-190	-1874	-2503	-1781	-2079	-1892	-1647	-683	210	1257	7

[illegible]

[illegible]

19(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	19
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
20(F)	-1300	-1041	-2794	-2396	2872	-2678	-676	440	-2076	1465	808	-1982	-2569	-1598	-1976	-1862	-1239	98	-109	791	20
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
21(V)	-922	-605	-3064	-2618	-752	-2841	-2139	2359	-2383	244	349	-2414	-2784	-2191	-2442	-2058	-911	2430	-1948	-1560	21
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
22(K)	-395	-1326	-130	117	-1960	1239	-86	-1633	2073	-1681	-915	-74	-1403	278	289	-375	-423	-1258	-1820	-1352	22
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
23(E)	-597	-1990	829	2406	-2276	-963	-75	-2069	89	-2078	-1318	1286	-1352	285	-438	-411	-615	-1663	-2255	-1512	23
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
24(W)	-1926	-1577	-2416	-2370	192	-2070	-795	-1602	-1882	-1347	-1210	-2018	-2383	-1854	-1777	-2150	-1992	-1643	5696	560	24
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

154
A

[illegible]

[illegible]

[illegible]

158

42(I)	-935	-625	-3052	-2609	-708	-2832	-2125	2677	-2364	307	391	-2408	-2777	-2168	-2419	-2052	-926	2060	-1920	-1540	42
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
43(K)	-489	-1476	-535	74	-1899	-1347	163	-1474	2257	-1467	-693	-94	-1441	566	1161	-436	566	-1175	-1619	-1191	43
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
44(I)	-935	-625	-3052	-2609	-708	-2832	-2125	2677	-2364	307	391	-2408	-2777	-2166	-2419	-2052	-926	2060	-1920	-1540	44
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
45(E)	-190	-1482	610	1617	-1728	-1033	193	-1411	949	-1462	-616	305	-1203	607	49	-87	-141	-15	-1713	-1064	45
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
46(N)	-508	-1893	708	1172	-2155	-1000	25	-1908	309	-1897	-1104	2413	-1320	976	-163	-338	-492	-1519	-2096	-1394	46
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
47(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	47
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

162
P

65(R)	-439	-1400	-571	103	-1664	-1368	1527	-1291	1414	-314	-534	-77	-1418	594	1819	-398	-324	-1021	-1489	-1036	65
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
66(F)	-1902	-1421	-2805	-2592	3214	-2763	351	-984	-2354	-657	-540	-1668	-2714	-1627	-2106	-1950	-1820	-1082	1026	3156	66
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
67(I)	-1016	-805	-2637	-2367	-591	-2470	-1853	-3039	-2040	329	346	-2180	-2589	-1943	-2070	-1879	-1051	1184	-1698	-1261	57
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
68(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	68
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
69(D)	-196	-1649	2118	1061	-1949	-1018	231	-1691	1047	-1647	-758	919	-1183	668	786	-60	-144	-1263	-1826	-1142	69
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
70(Y)	-1667	-1473	-2001	-1964	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	501	4149	70
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

166

88(I)	-1120	-818	-3064	-2567	-204	-2914	1888	2467	-2246	1556	866	-2411	-2743	-1931	-2224	-2108	-1072	1115	-1497	-1212	88
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
90(N)	-477	-1612	-233	247	-2000	-1284	183	-1637	1365	-1583	-780	1891	-1397	598	1526	-382	-381	-1301	-1701	-1216	90
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
91(M)	-965	-982	-2125	-1844	-276	-2022	-1302	373	-1390	610	4141	-1701	-2229	-1356	-1425	-1444	-1028	126	-1279	-880	91
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
92(L)	-1306	-1067	-2981	2443	125	-2920	-1717	814	-2062	2337	2190	-2353	-2682	-1686	-2014	-2092	-1231	368	-1248	-1063	92
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
93(C)	-206	3367	-2319	-1919	-689	-1494	-1270	951	-1603	-153	201	-1455	-1861	-1411	-1625	-746	-346	2000	-1422	-1027	93
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
94(D)	-1130	-2063	3441	243	-2502	-1157	-662	-2514	-843	-2578	-2029	-96	-1658	-421	-1434	970	-1239	-2138	-2388	-1915	94
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

170

111(P)	-934	-1220	-1352	-1451	-2187	-1314	-1456	-2160	-1488	-2268	-1827	-1359	3760	-1471	-1571	-1129	-1216	-1780	-2050	-2023	111
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
112(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	112
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
113(K)	-1028	-1686	-800	-442	-2171	-1521	-263	-1873	3103	-1859	-1251	-547	-1750	74	599	-1004	-972	-1620	-1812	-1528	113
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
114(I)	1297	-396	-1045	-533	-555	-1309	-316	1303	-228	-265	-313	-551	-1491	-201	106	-396	-102	374	-1004	-589	114
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
115(V)	433	-490	-819	416	-611	-1396	-259	391	-179	-21	288	-448	-1513	-95	-488	-437	-123	1568	-1055	-615	115
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
116(E)	-988	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	116
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

134(R)	-1181	-1614	-1431	-903	-2066	-1632	-446	-1944	559	-1867	-1326	-899	-1873	-156	3432	-1125	-1148	-1717	-1746	-1529	134
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
135(A)	2048	-367	-849	-680	-1659	-627	-739	-1190	-650	-1511	-785	-478	1606	-534	-899	74	-21	-698	-1911	-1515	135
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
136(R)	-1181	-1614	-1431	-903	-2066	-1632	-446	-1944	559	-1867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	136
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
137(R)	-957	-1769	-1130	-344	-2309	-1642	117	-1834	1694	-1695	-991	-391	-1689	518	2808	-887	-779	-1577	-1692	-1419	137
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
138(T)	-52	-555	-1173	-1112	-1571	-872	-1041	-862	-950	-1353	-866	-843	-1455	-937	-1084	-285	3069	-557	-1898	-1539	138
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
139(P)	-934	-1220	-1352	-1451	-2187	-1314	-1456	-2160	-1488	-2268	-1827	-1359	3760	-1471	-1571	-1129	-1216	-1780	-2050	-2023	139
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

124

[illegible]

[illegible]

[illegible]

178

157(A)	2806	444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	157
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
158(D)	-1130	-2063	3441	243	-2502	-1157	-662	-2514	-843	-2578	-2029	-96	-1658	-421	-1434	-970	-1239	-2138	-2388	-1915	158
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
159(A)	2806	444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	159
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
160(A)	2806	444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	160
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
161(E)	-704	-2274	1609	2494	-2517	-929	-118	-2327	-117	-2302	-1553	440	-1360	238	-777	-462	-729	-1891	-2491	-1671	161
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
162(A)	2429	-274	-996	-928	-1860	817	-856	-1435	-970	-1170	-1101	-575	-1171	-793	-1166	139	6	-817	-2094	-1749	162
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

174(V)	-177	-366	-1747	-1376	-835	-1408	-1080	842	-1132	-278	103	-1145	-1767	-1025	-1282	-633	1362	2195	-1521	-1119	174
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
175(A)	2031	-283	-945	-876	-1891	1659	-933	-1488	-939	-1802	-1022	-547	-1163	-758	-1150	148	11	-850	-2114	-1762	175
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
176(I)	-932	-615	-3069	-2624	-731	-2846	-2141	-2579	-2386	277	371	-2421	-2787	-2189	-2441	-2065	-921	2203	-1937	-1555	176
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
177(A)	2214	-291	-1447	-1169	-1154	-884	-959	74	-982	-736	-262	-833	-1414	-858	-1151	-156	-103	1000	-1694	-1298	177
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
178(R)	-1181	-1614	-1431	-903	-2066	-1632	-446	-1944	559	-1867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	178
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
179(Y)	-1667	-1473	-2001	-1967	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	-501	4149	179
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												

181

182

180(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721		
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		180
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*													
181(P)	-370	-1016	-600	-377	-1687	-1102	-432	-1453	72	-1586	-902	-444	2899	-166	707	-467	-495	-1110	-1727	-1330	181	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*													
182(F)	-1055	-829	-2811	-2263	2475	-2617	-1120	681	-1966	1473	2107	-2020	-2452	-1505	-1885	-1736	-975	308	-672	-125	182	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*													
183(N)	-637	-1349	-5	-148	-1773	-1078	-624	-1965	-499	-2112	-1514	3456	-1574	-435	-810	-661	-818	-1560	-1900	-1327	183	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*													
184(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	184	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*													
185(I)	-1143	-842	-3082	-2578	-159	-2929	-1882	2333	-2257	1719	917	-2424	-2743	-1919	-2225	-2119	-1091	1050	-1466	-1201	185	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		

[illegible]

[illegible]

197(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	197
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
198(V)	-700	-665	-2337	-2135	-873	-1965	-1727	1107	-1863	-105	22	-1851	-2286	-1794	-1920	-1368	-823	2927	-1816	-1362	198
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
199(L)	-1198	-902	-3098	-2586	-74	-2954	-1866	1998	-2259	2016	1009	-2444	-2745	1887	-2211	-2142	-1140	903	-1411	-1174	199
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
200(T)	-52	-555	-1173	-1112	-1571	-872	-1041	-862	-950	1353	-866	-843	-1455	-937	-1084	-285	3069	-557	-1898	-1539	200
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
201(Q)	-862	-1548	-321	-247	-1672	-1336	-419	-1721	130	-1660	-1161	-427	-1678	3540	-59	855	-904	-1495	-1728	-1237	201
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
202(C)	-608	5196	-2368	-2362	-1635	-1253	-1702	-987	-2034	-1540	-1191	-1745	-1831	-1953	-1922	-934	-927	-770	-1855	-1692	202
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												

135

186

203(S)	1057	-274	-916	-815	-1832	-508	-865	-1436	-820	-1752	-977	-502	-1145	-674	-1039	2412	41	-810	-2079	-1679	203
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
204(V)	-925	-616	-3000	-2497	-527	-2814	-1921	1840	-2252	949	560	-2316	-2703	-1995	-2284	-1988	-889	2430	-1695	-1364	204
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
205(E)	-988	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	205
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
206(E)	-988	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	206
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
207(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	207
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
208(T)	-323	-1596	-1187	-1153	-1884	-973	55	-1567	265	-1645	-837	335	-1257	437	-273	-205	1701	-1205	-1920	-1239	208
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
215(R)	-664	-1241	-1101	-424	-1173	-1559	-60	-804	818	181	-343	-483	-1657	246	2582	-741	-578	-719	1332	-877	215
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
216(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1175	-1590	-1790	-958	-1093	-1921	-2198	-2279	216
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
217(F)	-707	-668	-1832	-1360	2343	-1971	1711	-39	-1065	993	361	-1104	-1992	-805	-1140	-1047	-642	-73	294	1256	217
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
218(T)	-52	-555	-1173	-1112	-1571	-872	-1041	-862	-950	-1353	-866	-843	-1455	-937	-1084	-285	3069	-557	-1898	-1539	218
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
219(C)	1104	2197	-1323	-968	-1444	651	-781	-1064	-812	-1343	-561	-576	-1127	-634	-1023	1662	138	-550	-1740	-1381	219
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												

188

[illegible]

226(V)	-700	-665	-2337	-2135	-873	-1965	-1727	1107	-1863	-105	22	-1851	-2286	-1794	-1920	-1368	-823	2927	-1816	-1362	226
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
227(Y)	-1167	-1473	-2001	-1967	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	501	4149	227
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
228(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	228
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
229(T)	-52	-555	-1173	-1112	-1571	-872	-1041	-862	-950	-1353	-866	-843	-1455	-937	-1084	-285	3069	-557	-1898	-1539	229
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
230(E)	-717	-2337	2049	2259	-2571	-918	-117	-2393	-138	-2352	-1602	461	-1357	240	-821	-464	-744	-1944	-2542	-1698	230
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
231(K)	397	-1331	0	518	-1619	-1070	-250	-1306	1059	-1341	-481	786	38	674	793	224	-41	-946	-1579	-969	231
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

144

249(C)	1738	3050	-1597	-1286	-1304	-553	-920	-809	-1074	-1196	-460	-726	-1161	-870	-1191	1014	128	-366	-1662	-1327	249
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
250(Y)	-1667	-1473	-2001	-1964	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	501	4149	250
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
251(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	251
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
252(S)	116	-498	-840	-867	-1751	-671	-949	-1625	-916	-1896	-1218	-622	-1304	-824	-1097	2786	-220	-1038	-1992	-1530	252
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
253(R)	-1181	-1614	-1431	-903	-2066	-1632	446	-1944	559	-867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	253
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
254(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	254
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

272(A)	1700	-251	-1028	-731	-1556	-572	-693	-1113	-631	-1425	-647	-472	-1141	-490	-887	934	1645	-603	-1818	-1429	272
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
273(E)	-988	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2007	-1742	273
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
274(G)	307	-371	-641	-650	-1997	2231	-857	-1752	-837	-1971	-1163	-431	-1164	-654	-1088	1356	-32	-1032	-2185	-1778	274
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
275(K)	-1028	-1686	-800	-442	-2171	-1521	-263	-1873	3103	-1859	-1251	-547	-1750	74	599	-1004	-972	-1620	-1812	-1528	275
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
276(S)	116	-498	-840	-867	-1751	-671	-949	-1625	-916	-1896	-1218	-622	-1304	-824	-1097	2786	-220	-1038	-1992	-1530	276
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
277(M)	-965	-982	-2125	-1844	-276	-2022	-1302	373	-1390	610	4141	-1701	-2229	-1356	-1425	-1444	-1028	126	-1279	-880	277
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

295(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	295
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
296(L)	-1369	-1199	-2542	-2295	-129	-2453	-1603	488	-1867	2607	742	-2168	-2539	-1724	-1835	-1992	-1384	130	-1284	-897	296
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
297(Q)	-862	-1548	-321	-247	-1672	-1336	-419	-1721	130	-1660	-1161	-427	-1678	3540	-59	-855	-904	-1495	-1728	-1237	297
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
298(N)	-637	-1349	-5	-148	-1773	-1078	-624	-1965	-499	-2112	-1514	3456	-1574	-435	-810	-661	-818	-1560	-1900	-1327	298
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
299(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	299
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
300(S)	1375	-276	-871	-686	-1859	1354	-765	-1534	-710	-1758	-930	-428	-1109	-536	-985	1511	86	-862	-2062	-1669	300
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

	-688	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	320
318(L)	-1369	-1199	-2542	-2295	-129	-2453	-1603	488	-1867	2607	742	-2168	-2539	-1724	-1835	-1992	-1384	130	-1284	-897	318
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
319(G)	1867	-289	-929	-871	-1906	1873	-938	-1506	-944	-1818	-1039	-545	-1165	-761	-1155	146	5	-862	-2126	-1774	319
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
320(E)	-688	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	320
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
321(N)	-637	-1349	-5	-148	-1773	-1078	-624	-1965	-499	-2112	-1514	3456	-1574	-435	-810	-661	-818	-1560	-1900	-1327	321
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
322(L)	-1369	-1199	-2542	-2295	-129	-2453	-1603	488	-1867	2607	742	-2168	-2539	-1724	-1835	-1992	-1384	130	-1284	-897	322
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
323(I)	-1123	-822	-3067	-2569	-197	-2917	-1887	2447	-2248	1582	875	-2413	-2743	-1929	-2225	-2110	-1075	1105	-1492	-1210	323
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

219

341(H)	-1192	-1547	-812	-781	-520	-1557	4586	-1881	-457	-1789	-1357	-866	-1913	-669	-600	-1231	-1258	-1678	-918	-90	341
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
342(S)	116	-498	-840	-867	-1751	-671	-949	-1625	-916	-1896	-1218	-622	-1304	-824	-1097	2786	-220	-1038	-1992	-1530	342
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
343(D)	-1130	-2063	3441	243	-2502	-1157	-662	-2514	-843	-2578	-2029	-96	-1658	-421	-1434	-970	-1239	-2138	-2388	-1915	343
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
344(M)	-962	-721	-2832	-2307	-165	-2664	-1626	2366	-1987	903	2608	-2145	-2549	-1688	-1979	-1822	-914	1020	-1353	-1088	344
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
345(R)	-1181	-1614	-1431	-903	-2066	-1632	-446	-1944	559	-1867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	345
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
346(R)	-1181	-1614	-1431	-903	-2068	-1632	-446	-1944	559	-1867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	346
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

214

364(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	364
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
365(Y)	-1667	-1473	-2001	-1964	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	501	4149	365
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
366(S)	1123	-271	-926	-817	-1828	-508	-863	-1430	-818	-1746	-969	-502	-1143	-871	-1038	2382	45	-805	-2074	-1676	366
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
367(A)	2139	-236	-1024	-877	-1766	-505	-870	-1347	-846	-1673	-895	-530	-1137	-669	-1059	1515	68	-744	-2025	-1643	367
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
368(V)	-151	-569	-631	524	-788	-1293	-269	234	-134	-481	122	-371	-1478	-70	-451	-389	637	1708	-1201	-747	368
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
369(P)	-478	-1420	327	1138	-1929	-1032	-248	-1622	-3	-1747	-1039	36	2562	70	-402	-436	-543	-1278	-1965	-1426	369
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
370(N)	-637	-1349	-5	-148	-1773	-1078	-624	-1965	-499	-2112	-1514	3456	-1574	-435	-810	-661	-818	-1560	-1900	-1327	370
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
371(Y)	-806	-794	-1826	-1397	917	-2006	-90	-29	-1025	35	1798	-1154	-2057	-849	-1103	-1127	-760	-95	211	3271	371
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
372(D)	-1130	-2063	3441	243	-2502	-1157	-662	-2514	-843	-2578	-2029	-96	-1658	-421	-1434	-970	-1239	-2138	-2388	-1915	372
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
373(N)	-637	-1349	-5	-148	-1773	-1078	-624	-1965	-499	-2112	-1514	3456	-1574	-435	-810	-661	-818	-1560	-1900	-1327	373
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
374(M)	-204	-445	-1398	-949	-453	-1403	-637	369	-636	129	3007	-865	-1648	-574	-809	-563	1278	377	-1116	-731	374
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
375(F)	1602	-1340	-2512	-2455	3740	-2315	-379	-478	-2225	-175	-190	-1874	-2503	-1781	-2079	-1892	-1647	-683	210	1257	375

[illegible]

[illegible]

213
A

387(D)	-1130	-2063	3441	243	-2502	-1157	-662	-2514	-843	-2578	-2029	-96	-1658	-421	-1434	-970	-1239	-2138	-2388	-1915	387
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
388(Y)	-1667	-1473	-2001	-1964	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	501	4149	388
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
389(N)	-199	-701	-549	-93	-228	-1316	21	533	64	-536	109	1814	-1432	121	-268	-348	-154	-118	-681	1083	389
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
390(V)	-596	-427	-2444	-1942	-648	-2270	-1439	1955	-1702	78	364	-1768	-2320	-1523	-1801	-1418	532	2258	-1522	-1137	390
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
391(M)	-1151	-859	-3123	-2554	25	-2901	-1760	1894	-2267	1875	2113	-2393	-2661	-1798	-2177	-2046	-1074	735	-1288	-1136	391
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
392(Q)	-513	-1819	522	1560	-2087	-1069	43	-1821	462	-1801	-1025	265	-1348	2403	72	-362	-480	-1458	-1986	-1342	392
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

222

410(I)	-1016	-805	-2637	-2367	-591	-2470	-1853	3039	-2040	329	346	-2180	-2589	-1943	-2070	-1879	-1051	1184	-1698	-1261	410
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
411(K)	1191	-905	-299	140	-1432	-1011	-7	-980	1412	-1186	-412	-29	-1257	344	101	-104	626	-667	-1531	-1026	411
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
412(I)	442	-488	-2154	-1781	-680	-1947	-1383	2623	-1524	86	307	-1582	-2159	-1402	-1631	-1169	-575	1239	-1576	-1178	412
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
413(R)	-1181	-1614	-1431	-903	-2066	-1632	-446	-1944	559	-1867	-1326	-899	-1873	-156	3432	-1225	-1148	-1717	-1746	-1529	413
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
414(N)	-406	-1604	364	514	-1650	-1088	1550	-1590	512	-1604	-816	2266	-1324	905	133	-295	-367	-1249	-1692	-1010	414
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
415(K)	-1028	-1886	-800	-442	-2171	-1521	-263	-1873	3103	-1859	-1251	-547	-1750	74	599	-1004	-972	-1620	-1812	-1528	415
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

226
AP

433(T)	737	-257	-1006	-798	-1590	-576	-766	-1064	-703	-1448	-698	-511	-1166	-583	-923	595	2548	-574	-1871	-1487	433
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
434(D)	-1130	-2063	3441	243	-2502	-1157	-662	-2514	-843	-2578	-2029	-96	-1658	-421	-1434	-970	-1239	-2138	-2388	-1915	434
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
435(E)	-988	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	435
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
436(E)	-988	-1921	379	3052	-2309	-1196	-495	-2108	-350	-2194	-1612	-103	-1621	-210	-770	-862	-1042	-1799	-2207	-1742	436
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
437(V)	-700	-665	-2337	-2135	-873	-1965	-1727	1107	-1863	-105	22	-1851	-2286	-1794	-1920	-1368	-823	2927	-1816	-1362	437
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
438(E)	-462	-1877	1098	2281	-2164	-948	4	-1884	157	-1917	-1119	395	-1290	383	-418	-292	231	-1485	-2158	-1417	438
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

479(Y)	499	-364	-899	-371	787	-1296	-82	7	-211	-277	393	-419	-1419	-73	-488	927	-30	144	587	1023	479
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
480(K)	-84	-1454	654	613	-1738	-1039	285	-1458	1165	-1450	-558	861	-1146	994	487	248	-23	-503	-1655	-1002	480
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
481(H)	-138	-1185	181	434	-1561	117	2574	-1298	388	-1392	-582	681	-1230	426	-73	519	-152	-942	-1647	-1037	481
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
482(G)	-478	-1458	1461	350	-2273	2341	-418	-2092	-439	-2185	-1451	76	-1399	-120	-981	-425	-627	-1604	-2284	-1711	482
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
483(F)	-1075	-853	-2688	-2263	2906	-2522	-760	1338	-1963	741	748	-1891	-2464	-1545	-1898	-1690	-1030	450	-271	583	483
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
484(K)	-186	-1612	1105	1050	-1898	-1008	220	-1635	1350	-1612	-730	353	-372	652	63	-56	254	-1218	-1811	-1126	484
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

502(Y)	-1894	-1443	-2713	-2588	2338	-2734	353	-1021	-2261	-702	-584	-1632	-2696	-1582	-2053	-1924	-1818	-1110	1018	3836	502
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
503(L)	-1369	-1199	-2542	-2295	-129	-2453	-1603	488	-1867	2607	742	-2168	-2539	-1724	-1835	-1992	-1384	130	-1284	-897	503
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
504(Q)	-528	-1553	-18	249	-1355	-1246	1676	-1537	686	-1512	-787	17	-1438	2674	436	-447	-468	-1245	-1443	-765	504
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
505(T)	-52	-555	-1173	-1112	-1571	-872	-1041	-862	-950	-1353	-866	-843	-1455	-937	-1084	-285	3089	-557	-1898	-1539	505
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
506(S)	116	-498	-840	-867	-1751	-671	-949	-1625	-916	-1896	-1218	-622	-1304	-824	-1097	2786	-220	-1038	-1992	-1530	506
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
507(A)	2806	-444	-1270	-1265	-1732	-744	-1182	-1034	-1212	-1545	-1010	-869	-1380	-1102	-1325	-156	-259	-640	-2039	-1721	507
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

242

525(Y)	-1667	-1473	-2001	-1964	1181	-2210	-25	-1224	-1658	-1018	-871	-1505	-2415	-1423	-1618	-1747	-1692	-1248	501	4149	525
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
526(Q)	394	-1108	-203	338	-1300	-1159	184	-918	1042	-1043	968	94	-1255	1737	277	-133	-76	-656	-1380	-845	526
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
527(G)	-758	-1103	-1313	-1487	-2460	3331	-1562	-2502	-1720	-2644	-2081	-1324	-1775	-1590	-1790	-958	-1093	-1921	-2198	-2279	527
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
528(P)	-934	-1220	-1352	-1451	-2187	-1314	-1456	-2160	-1488	-2268	-1827	-1359	3760	-1471	-1571	-1129	-1216	-1780	-2050	-2023	528
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
529(G)	1163	-332	-849	-861	-1983	2434	-981	-1591	-996	-1909	-1142	-553	-1190	-813	-1198	105	-47	-932	-2188	-1842	529
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
530(T)	-52	-555	-1173	-1112	-1571	-872	-1041	-862	-950	-1353	-866	-843	-1455	-937	-1084	-285	3069	-557	-1898	-1539	530
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

246
PP

548(P)	1061	433	-770	-684	-1731	-657	-791	-1256	-689	-1580	-892	-499	2706	-594	-921	8	-100	-769	-1976	-1581	548
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
549(W)	-38	-416	-726	-186	608	365	-32	-47	-33	-324	375	490	-1365	81	-348	-269	377	100	1582	-231	549
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
550(A)	1982	-301	-1528	-1215	-1040	-1011	-971	334	-1023	-585	-132	-907	-1498	-891	-1190	-268	-147	1432	-1620	-1225	550
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
551(M)	-956	-675	-3000	-2437	-98	-2720	-1640	1822	-2169	1596	1921	-2230	-2553	-1752	-2099	-1856	-891	1429	-1284	-1077	551
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
552(D)	-409	-1868	1976	711	-2139	-994	95	-1893	949	-1858	-1023	375	-1268	1572	-112	-239	-379	-1478	-2048	-1334	552
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-33	-6045	-7087	-894	-1115	-701	-1378	*	*												
553(P)	-934	-1220	-1352	-1451	-2187	-1314	-1458	-2160	-1488	-2268	-1827	-1359	3760	-1471	-1571	-1129	-1216	-1780	-2050	-2023	553
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

TABELA 13

HMMER2.0 [2.3.2]
NOME beta_exp_seqs
LENG 238
ALPH Amino
RF no
CS no

MAP sim
 COM hmmbuild beta_hmm2 beta_exp_seqs.a
 COM hmmscalibrate beta_hmm2
 NSEQ 8
 DATA Fri Mar 30 18:30:25 2007
 CKSUM 9853

XT -8455 -4 -1000 -1000 -8455 -4 -8455 -4

NULT -4 -8455

NULO 595 -1558 85 338 -294 453 -1158 197 249 902 -1085 -142 -21 -313 45 531 201 384 -1998 -644

EVD -152.743744 0.132052

HMM	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	
	m->m	m->i	m->d	i->m	i->j	d->m	d->d	b->m	m->e												
	-561	*	-1634																		
1(M)	-117	-798	-402	82	-937	-1185	8	-544	307	-732	2122	-103	-1331	-1400	-38	722	-88	-348	-1198	-701	3
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-701	-1378	-561	*												
2(E)	-1082	-2012	315	3129	-2414	-1266	-581	-2231	-449	-2306	-1728	-179	-1697	-299	-873	-951	-1139	-1916	-2295	-1840	4
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-701	-1378	*	*												
3(I)	-1084	-859	-2738	-2477	-657	-2548	-1955	3104	-2151	274	286	-2280	-2665	-2051	-2174	-1970	-1120	1153	-1775	-1342	5

[illegible]

[illegible]

15(V)	748	-413	-2029	-1720	-969	-1515	-1368	951	-1502	-332	-2	-1391	-1912	-1365	-1614	-776	-423	2546	-1740	-1342	17
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-701	-1378	*	*												
16(L)	-1296	-988	-3217	-2703	-76	-3057	-1956	1866	-2372	2163	1015	-2560	-2822	-1967	-2303	-2251	-1234	794	-1447	-1233	18
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-701	-1378	*	*												
17(N)	1158	-703	-271	-42	-1756	-755	-333	-1434	-123	-1598	-780	1499	-1204	-10	-522	1429	-63	-929	-1892	-1374	19
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-701	-1378	*	*												
18(E)	-1082	-2012	315	3129	-2414	-1266	-581	-2231	-449	-2306	-1728	-179	-1697	-299	-873	-951	-1139	-1916	-2295	-1840	20
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-701	-1378	*	*												
19(M)	-510	-499	-2163	-1628	-454	-1946	-1027	664	-1344	132	3301	-1444	-2063	-1141	-1422	-1064	853	1357	-1169	-825	21
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-24	-6495	-7537	-894	-1115	-701	-1378	*	*												
20(Q)	-595	-1941	-47	1215	-2308	-1349	8	-1994	1563	-1906	-1064	12	-1502	2357	337	-457	-516	-1599	-2025	-1446	22
-	-149	-498	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

38(A)	1607	-764	-122	105	-1345	-864	-110	-927	240	-1134	-440	-28	-1219	1445	-70	-82	-102	-620	-1536	-1024	41
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-37	-5859	-6901	-894	-1115	-1577	-589	*	*												
39(A)	1803	-165	-972	-743	-1388	-531	-683	-645	-621	-1166	-492	-458	-1107	-509	-835	-189	1642	-256	-1723	-1338	42
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-262	-5859	-2747	-894	-1115	-1577	589	*	*												
40(T)	355	-282	-441	-217	-1238	-551	-308	-723	-120	-1049	-364	-157	1124	-76	-394	184	1668	-353	-1502	-1065	43
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-44	-5640	-6682	-894	-1115	1694	-534	*	*												
41(V)	1381	-162	-1227	-907	-788	-898	-715	613	-712	-293	125	-683	-1345	-611	-902	-143	7	1512	-1390	-978	44
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-44	-5640	-6682	-894	-1115	1694	-534	*	*												
42(T)	1348	-93	-776	-557	-1228	-464	-527	-425	-433	-977	-333	-323	-1020	-345	-654	258	1732	-77	-1578	-1179	45
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-44	-5640	-6682	-894	-1115	-1250	-787	*	*												
43(A)	1932	-257	-1412	-1102	-1001	-936	-884	327	-912	-552	-96	-813	-1424	-790	-1089	-191	-85	1321	-1569	-1170	46
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

253
P

61(F)	-1369	-1089	-2974	-2547	2668	-2815	-904	479	-2229	1773	892	-2159	-2660	-1714	-2105	-1996	-1299	90	-319	-513	65
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-1391	-693	*	*												
62(Q)	-318	-1420	-175	306	-1747	-1213	155	-1383	1463	-1416	-607	66	-1336	1585	492	-256	1223	-1062	-1632	-1103	66
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-31	-6123	-7165	-894	-1115	-393	-2068	*	*												
63(H)	-965	-2315	304	2557	-2345	-1402	2832	-2376	-96	-2333	-1572	-69	-1734	-25	-542	-787	-951	-1991	-2354	-1617	67
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
64(V)	-243	-624	-951	-402	-698	164	-280	507	700	-540	158	-512	-1577	-129	-534	129	-199	1501	-1065	-632	68
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
65(G)	-1103	2683	1617	1044	-3018	2466	-544	-2852	-622	-2806	-2065	78	-1745	-197	-1335	-856	-1156	-2377	-3010	-2151	69
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
66(I)	127	-601	-1729	700	-741	-2010	-897	1823	-1032	-346	165	-1212	-2069	-884	-1251	-1065	-513	1823	-1321	-915	70
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

84(F)	-2605	-2164	-3410	-3518	4186	-3034	-940	-1504	-3300	-1072	-1142	-2698	-3246	-2697	-2999	-2815	-2670	-1743	-274	806	88
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
85(D)	1368	-1571	1781	109	-2469	1557	-540	-2201	-430	-2276	-1455	-143	-1580	-184	-971	-465	-630	-1675	-2520	-1865	89
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
86(K)	-543	-1715	-308	1446	-1982	-1429	-43	-1616	1943	117	-835	-107	-1542	369	238	-460	-465	-1300	-1873	-1318	90
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
87(Y)	768	-433	-1427	-878	-230	-1635	1002	28	-656	-299	359	-841	-1719	-495	-842	-672	-251	1365	-705	2062	91
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
88(Q)	-713	-1108	-1266	-711	-658	-1906	-585	-189	-306	733	1857	-852	-1945	2725	-557	-934	-646	-321	-1349	-939	92
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
89(H)	-509	-1727	1105	324	-1970	-1263	2947	-1737	98	-1797	-978	19	-1505	215	-403	-416	1593	-1376	-2025	-1356	93
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*																
90(H)	257	-1526	-246	303	-1778	-1296	1995	-1458	1182	-1499	-646	24	-1387	1802	104	-236	-241	-136	-1741	-1141	94				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*																
91(N)	-205	-1061	-522	-287	-2096	-1074	-595	-1786	-360	-1930	-1104	2136	-1508	-259	-766	834	2022	-1276	-2209	-1676	95				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*																
92(M)	-1753	-1370	-3997	-3423	-278	-3613	-2472	2205	-3134	2048	2292	-3216	-3244	-2474	-2907	-2799	-1664	281	-1734	-1673	96				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*																
93(V)	-961	-720	-3067	-2588	-1001	-2736	-1942	2099	-2323	-277	18	-2312	-2758	-2110	-2361	-1900	1460	2254	-1899	-1513	97				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*																
94(G)	-798	-2252	1834	409	-2613	1891	-331	-2390	1009	-2358	-1542	66	-1618	55	-720	-619	-806	-1944	-2552	-1804	98				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*																
95(M)	-1293	-943	-3609	-3096	-758	-3225	-2328	2635	-2838	287	2654	-2820	-3076	-2478	-2765	-2408	-1249	1673	-1950	-1676	99				

[illegible]

[illegible]

267
A

107(G)	-1814	-2025	-2478	-2748	-3577	3650	-2700	-3847	-3052	-3864	-3354	-2492	-2699	-2874	-2984	-2040	-2187	-3148	-3081	-3439	111
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
108(I)	-1408	-993	-3860	-3463	-1315	-3565	-3081	2912	-3283	-218	-139	-3218	-3431	-3098	-3328	-2843	-1397	2338	-2695	-2286	112
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
109(E)	-1877	-2708	-271	3539	-3235	-1868	-1322	-3210	-1299	-3203	-2676	-849	-2342	-1076	-1736	-1723	-1961	-2857	-2976	-2640	113
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
110(E)	-1877	-2708	-271	3539	-3235	-1868	-1322	-3210	-1299	-3203	-2676	-849	-2342	-1076	-1736	-1723	-1961	-2857	-2976	-2640	114
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
111(E)	-1877	-2708	-271	3539	-3235	-1868	-1322	-3210	-1299	-3203	-2676	-849	-2342	-1076	-1736	-1723	-1961	-2857	-2976	-2640	115
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
112(G)	-1814	-2025	-2478	-2748	-3577	3650	-2700	-3847	-3052	-3864	-3354	-2492	-2699	-2874	-2984	-2040	-2187	-3148	-3081	-3439	118
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

24
A

130(A)	2576	-604	-1716	-1592	-2503	895	-1526	-2207	-1588	-2442	-1589	-1066	-1541	-1356	-1766	1089	-332	-1390	-2720	-2403	134
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
131(W)	1057	-423	-1396	-834	-306	-1603	1535	20	-598	-296	367	-808	-1690	-452	-791	-637	-225	1077	2854	-289	135
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
132(D)	-1299	-3082	2892	2152	-3290	-1299	-612	-3186	-788	-3089	-2397	95	-1806	-275	-1580	-987	-1355	-2689	-3263	-2321	136
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
133(A)	1980	-568	-1754	-1474	-1797	1737	-1271	-1167	-1349	-1645	-923	-1058	-1593	-1159	-1530	-274	-322	876	-2145	-1790	137
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
134(D)	1587	-2023	1916	376	-2494	-1219	-341	-2252	-201	-2265	-1443	1603	-1579	40	-787	-539	-712	-1798	-2489	-1753	138
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
135(H)	-380	-1637	-315	259	-1928	-1346	1589	-1609	1185	-1608	-749	1433	-1428	490	1250	-293	-304	142	-1806	-1220	139
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

153(I)	-1608	-1307	-3432	-3238	-1179	-3095	-2687	3474	-2941	-191	-206	-2987	-3216	-2818	-2915	-2635	-1653	-832	-2363	-1952	157
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
154(H)	-2323	-2419	-1868	-1956	-1368	-2396	5123	-3143	-1645	-2924	-2595	-1989	-2806	-1873	-1712	-2381	-2439	-2923	-1719	-935	158
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
155(Q)	-1762	-2287	-1132	-1144	-2520	-2040	-1289	-2773	-740	-2625	-2183	-1302	-2433	4141	-881	-1765	-1840	-2519	-2457	-2076	159
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
156(K)	-1383	-2308	-1643	-684	-2977	-2064	-157	-2423	2495	-2166	-1434	-724	-2050	1271	2417	-1265	-1155	-2139	-2093	-1856	160
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
157(D)	-1253	-2521	3255	144	-3228	1167	-878	-3157	-1066	-3127	-2440	-203	-1918	-574	-1777	-1067	-1387	-2610	-3163	-2448	161
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
158(L)	-1096	-1358	-1544	-1152	-644	-2175	-905	-263	-552	2080	225	-1234	-2259	2027	-699	-1367	-1047	-506	-1463	-1019	162
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

279
A

176(D)	-1277	-3051	2640	2426	-3262	-1297	-598	-3152	-756	-3057	-2356	97	-1797	-258	-1537	-970	-1330	-2657	-3231	-2297	180
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
177(T)	741	-572	-1785	-1337	-1017	-1495	-1012	1165	-1127	-673	-131	-1150	-1838	-980	-1309	-671	2444	264	-1540	-1155	181
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
178(Y)	-2952	-2203	-3627	-3744	2312	-3534	-52	-1936	-3354	-1436	-1429	-2308	-3446	-2368	-2918	-2769	-2850	-2068	674	4232	182
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
179(R)	-2176	-2415	-2504	-1923	-3054	-2403	-1269	-3053	-167	-2856	-2350	-1844	-2662	-997	3887	-2228	-2131	-2807	-2501	-2452	183
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
180(Q)	541	-1165	-629	-174	-1300	-1442	-271	-837	115	125	-362	-362	-1609	2716	-208	-502	-409	-670	-1574	-1069	184
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*												
181(I)	-1608	-1307	-3432	-3238	-1179	-3095	-2687	3474	-2941	-191	-206	-2987	-3216	-2818	-2915	-2635	-1653	832	-2363	-1952	185
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

199(T)	-410	-679	-2187	-1873	-1334	-1555	-1509	252	-1620	-819	-414	-1498	-2017	-1493	-1731	-819	2683	1729	-1987	-1601	203	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
200(V)	-640	-504	-2456	-1869	-431	-2217	-1096	1416	-1525	-927	467	-1651	-2226	-1310	527	-1301	-584	2026	-1123	-778	204	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
201(N)	-1465	-2069	-762	-997	-2634	-1768	-1516	-3059	-1508	-3127	-2580	4001	-2312	-1391	-1789	-1503	-1700	-2572	-2658	-2191	205	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
202(D)	-2059	-2867	3860	-532	-3452	-1877	-1516	-3665	-1847	-3616	-3145	-882	-2414	-1318	-2469	-1872	-2198	-3236	-3167	-2847	206	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
203(Q)	-1762	-2287	-1132	-1144	-2520	-2040	-1289	-2773	-740	-2625	-2183	-1302	-2433	4141	-881	-1765	-1840	-2519	-2457	-2076	207	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
204(M)	-1736	-1678	-2975	-2821	-903	-2619	-2137	-342	-2336	24	4777	-2581	-2886	-2281	-2280	-2241	-1843	-654	-1928	-1574	208	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		

-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*															
205(V)	2074	-651	-2390	-2110	-1415	-1553	-1686	337	-1896	-878	-474	-1620	-2049	-1709	-1970	-833	-621	2210	-2098	-1727				209
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249				
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*															
206(R)	-2176	-2415	-2504	-1923	-3054	-2403	-1269	-3053	-167	-2856	-2350	-1844	-2662	-997	3887	-2228	-2131	-2807	-2501	-2452				210
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249				
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*															
207(P)	-464	-829	-1870	-1666	-1484	-1471	-1433	-307	-1476	-1083	-664	-1375	2872	-1386	-1610	-788	-689	1790	-2036	-1633				211
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249				
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*															
208(K)	-1185	-2213	-949	-394	-2739	-1847	-198	-2317	2731	-2136	-1393	-533	-1934	2194	759	-1067	-1032	-2005	-2126	-1773				212
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249				
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*															
209(Y)	-2985	-2209	-3678	-3814	3177	-3561	-43	-1955	-3416	-1446	-1439	-2321	-3463	-2389	-2950	-2792	-2877	-2087	687	3771				213
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249				
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*															
210(Q)	-762	-1364	-955	-522	-1120	-1735	-477	-761	27	-599	2479	-672	-1881	2972	-193	-868	-726	-767	-1578	-1080				214

[illegible]

[illegible]

287

222(K)	-1892	-2397	-1701	-1300	-3050	-2216	-948	-2829	3583	-2696	-2116	-1348	-2450	-619	78	-1865	-1815	-2562	-2460	-2321		226
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
223(H)	-815	-1622	-721	-285	-154	-1730	3084	-1465	74	-1447	-801	-458	-1830	2002	-226	-776	-742	-1248	-656	1967	227	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
224(V)	-1294	-1167	-3176	-3053	-1510	-2597	-2599	776	-2808	-665	-568	-2655	-2928	-2721	-2804	-2093	-1429	3343	-2527	-2099	228	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
225(V)	-1294	-1167	-3176	-3053	-1510	-2597	-2599	776	-2808	-665	-568	-2665	-2928	-2721	-2804	-2093	-1429	3343	-2527	-2099	229	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
226(Q)	256	-702	-834	-278	-767	-1453	-225	-282	-122	-586	1554	-425	-1546	1737	-468	-439	956	664	-1117	-668	230	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-21	-6672	-7714	-894	-1115	-701	-1378	*	*													
227(G)	-983	-2484	2007	425	-2938	2044	-519	-2776	-557	-2732	-1965	1439	-1705	-165	-1244	-772	-1048	-2274	-2930	-2091	231	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		

[illegible]

[illegible]

29/

2(S)	-714	-1508	-1330	-1230	-3093	-1545	-1557	-2841	-1413	-2974	-2115	-2328	-2119	-1253	-1778	2548	824	-2105	-3194	-2709	2
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
3(E)	-1483	-3123	1915	2354	-3400	-1775	-923	-3218	-893	-3153	-2360	-393	-2183	-548	-1565	-1244	1945	-2744	-3349	-2503	3
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-1162	-7442	-869	-894	-1115	-701	-1378	*	*												
4(I)	-1121	-753	-3426	-2999	-988	-3155	-2539	2556	-2786	50	147	-2768	-3061	-259	-2831	-2397	-1109	2478	-2257	-1863	4
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-27	-6308	-7350	-894	-1115	-213	-2863	*	*												
5(D)	-683	-2011	2419	66	-2272	-1583	-322	411	745	-2004	-1135	766	-1725	104	-470	-581	551	-1598	-2239	-1583	5
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
6(D)	887	-1768	1716	30	-2010	-1563	-265	-85	855	-1760	-908	-259	-1671	151	-387	-505	1220	-1340	-2037	-1419	6
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
7(M)	-1176	-998	-3214	-2614	-619	-2754	-1611	1640	-2276	1581	2955	913	-2724	-1911	-2174	-1847	-1112	-23	-1437	-1150	7
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

295
A

25(T)	-290	-1158	-455	992	-1440	-1324	-182	-1023	92	-1236	-454	-204	-1490	162	-346	1073	304	618	-1621	-1082	25
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*												
26(A)	2848	-686	-2099	-2005	-2426	-1015	-1766	-1806	-1874	-2283	-1547	-1327	-1708	-1670	-1958	-344	1455	-1205	-2739	-2445	26
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*												
27(P)	-117	-814	-1355	-1082	-2287	-1034	-1151	-1960	-1028	-2169	-1343	-882	2175	-896	-1323	1590	1811	-1330	-2476	-2063	27
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*												
28(A)	1889	-735	-1594	-1300	-2159	-1036	-1254	-1753	-1199	-2034	-1241	-997	1840	-1048	-1449	-299	1828	-1186	-2406	-2028	28
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*												
29(T)	1146	-812	-1308	-911	-2113	-1049	-968	-1778	-813	-1976	-1149	-782	1200	-689	-1149	1530	1750	-1213	-2300	-1865	29
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*												
30(A)	2697	-693	-1984	-1936	-2645	-959	-1785	-2307	-1901	-2607	-1776	-1275	-1671	-1663	-2006	1734	-463	-1490	-2890	-2593	30
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*																
31(T)	1184	-788	-1393	-1051	-2205	1459	-1096	-1875	-979	-2081	-1251	-859	1250	-835	-1291	-265	1793	-1268	-2400	-1985	31				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*																
32(A)	2154	-729	-1615	-1319	-2152	-1035	-1263	-1746	-1215	-2029	-1236	-1006	1292	-1061	-1461	-298	1830	-1180	-2403	-2027	32				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*																
33(T)	1265	-698	-1787	-1520	-2328	-990	-1421	-1963	-1411	-2231	-1411	-1092	-1625	-1232	-1627	1653	2477	-1298	-2566	-2217	33				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*																
34(Q)	1091	-1611	-110	225	-2092	488	-167	-1806	180	-1836	-972	1294	-1480	-1910	-301	-326	-396	-1385	-2066	-1429	34				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*																
35(G)	-578	-1644	1459	30	-2499	1928	-604	-2224	-483	-2306	-1490	-223	-1659	-249	-1016	-545	1609	-1718	-2557	-1909	35				
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249					
-	-19	-6798	-7840	-894	-1115	-1832	-476	*	*																
36(Q)	807	-1248	-477	-5	-1851	-1220	-253	-1525	83	-1631	-790	-211	-1467	1893	-361	1196	1007	-1126	-1919	-1350	36				

[illegible]

[illegible]

48(A)	2566	-1673	-3241	-2895	1995	-2977	-836	-1285	-2608	-1350	-940	-2278	-3082	-2148	-2520	-2124	-1792	-1243	-329	2352	48
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*													
49(T)	556	-1927	-417	1122	-2324	-1539	-326	-2051	46	-2052	-1161	914	-1696	109	-456	1237	1507	-1625	-2264	-1607	49
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*													
50(K)	-1937	2062	-2451	-1377	-3415	-2629	-734	-2850	2929	-2646	-1935	-1362	-2625	-319	2284	-1860	-1722	-2588	-2595	-2389	50
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*													
51(H)	-1373	3432	-2086	-1398	-1984	-2275	-3997	-1962	-97	-2063	-1380	-1376	-2458	-680	1376	-1482	-1351	-1730	-2090	-1576	51
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*													
52(P)	-877	-1501	-2346	-2559	-3585	-1679	-2607	-3503	-2784	-3706	-2880	-1979	3765	-2531	-2839	1007	-1303	-2500	-3737	-3472	52
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*													
53(E)	-2083	-3956	1757	3314	-4156	-1930	-1331	-4122	-1636	-3984	-3863	-536	-2485	-1014	-2522	-1729	-2166	-3591	-4126	-3126	53
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

71(V)	-2174	-1982	-4186	-4170	-2484	-3406	-3713	42	-3993	-1584	-1517	-3697	-3771	-3891	-3924	-3057	-2329	3646	-3507	-3148	71
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
72(L)	-2729	-2229	-5089	-4576	-885	-4785	-3686	2209	-4314	2611	306	-4471	-4194	-3468	-3998	-4097	-2634	-271	-2572	-2589	72
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
73(N)	956	-2274	1213	-94	-2844	-1631	-685	-2609	-449	-2590	-1722	2259	-1939	-278	-1001	1526	-973	-2118	-2789	-2080	73
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
74(G)	-1785	-3613	2207	1292	-3820	2213	-1112	-3710	-1261	-3595	-2872	1221	-2322	-764	-2053	-1476	-1833	-3198	-3789	-2827	74
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
75(K)	-708	-2181	1118	746	-2508	-1590	-284	-2254	1976	-2187	-1279	1149	-1720	186	1307	-575	-648	-1808	-2344	-1667	75
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
76(V)	-2021	-1555	-4630	-4292	-2072	-4355	-4168	1922	-4183	-894	-829	-4054	-4177	-4066	-4278	-3696	-2019	3287	-3647	-3180	76
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

94(I)	-2036	-1562	-4657	-4316	-2070	-4401	-4214	3210	-4217	-878	-820	-4089	-4201	-4095	-4314	-3744	-2031	2315	-3666	-3206	94
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
95(A)	3527	-1872	-3386	-3637	-3817	-2114	-3249	-3478	-3634	-3846	-3228	-2714	-2847	-3362	-3481	-1630	-1813	-2725	-3849	-3859	95
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
96(E)	657	-2144	-303	1992	-2460	-1564	-271	-2208	916	-2153	-1249	1296	-1696	1197	-360	-543	-616	-1765	-2321	-1636	96
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
97(D)	-1428	-3090	2456	0	-3356	626	-868	-3185	-800	-3101	-2290	1589	-2147	2323	-1457	-1192	-1436	-2703	-3286	-2444	97
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
98(M)	2458	-1063	-2439	-2109	-2009	-1619	-1748	-1532	-1906	-1860	3077	-1662	-2171	-1718	-2036	877	-861	-1187	-2405	-2053	98
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
99(G)	-1362	-2477	-504	-475	-2812	2719	1818	-2783	-476	-2759	-1994	-738	-2262	1773	-798	-1267	-1389	-2373	-2786	-2136	99
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

311

117(P)	-877	-1501	-2346	-2559	-3585	-1679	-2607	-3503	-2784	-3706	-2880	-1979	3765	-2531	-2839	1007	-1303	-2500	-3737	-3472	117
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
118(D)	-3359	-3855	4091	-1760	-4660	-2944	-2744	-5137	-3226	-4926	-4562	-2110	-3493	-2619	-3804	-3177	-3522	-4662	-4114	-4101	118
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
119(D)	1000	-3403	2984	1734	-3674	-1814	-1070	-3527	-1163	-3445	-2700	-428	-2286	-718	-1910	-1407	-1729	-3031	-3645	-2731	119
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
120(R)	-1763	-2178	-2815	-2009	-2626	-2577	-1312	-1725	-215	-2141	-1691	-1848	-2804	-1008	3553	-1932	-1770	1073	-2627	-2315	120
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
121(I)	-2084	-1616	-4657	-4243	-1653	-4365	-3815	3132	-4093	890	-443	-4016	-4093	-3767	-4069	-3656	-2054	1752	-3174	-2884	121
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
122(L)	-3371	-2888	-4716	-4628	-1407	-4203	-3653	-947	-4306	3212	-381	-4472	-4191	-3767	-3993	-4264	-3364	-1600	-2791	-2682	122
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

140(A)	3050	-2169	1557	-918	-3547	-1815	-1738	-3268	-1907	-3465	-2733	-1101	-2422	-1474	-2446	-1331	-1572	-2621	-3648	-3069	140
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
141(I)	-1583	-1435	-3802	-3522	-1810	-3145	-2999	3393	-3226	-885	-741	-3089	-3390	-3083	-3224	-2453	1116	598	-2920	-2528	141
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
142(A)	3159	-1216	-2991	-3125	-3530	-1484	-2756	-3260	-3083	-3557	-2679	-2048	-2256	-2728	-3008	1226	-1046	-2226	-3770	-3578	142
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
143(D)	785	-2776	3038	-83	-3174	-1756	-838	-2947	1206	-2909	-2094	-427	-2118	-452	-1226	-1122	-1320	-2485	-3100	-2339	143
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
144(E)	-2080	-3966	2020	3228	-4159	-1924	-1325	-4125	-1633	-3985	-3364	-528	-2480	-1008	-2526	-1723	-2163	-3592	-4133	-3124	144
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
145(L)	-3371	-2888	-4716	-4628	-1407	-4203	-3653	-947	-4306	3212	-381	-4472	-4191	-3787	-3993	-4264	-3364	-1600	-2791	-2682	145
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[illegible]

[illegible]

[illegible]

163(E)	-1353	-2969	1558	2718	-3260	316	-827	-3072	-733	-3001	-2179	-361	-2110	-438	-1376	-1133	601	-2596	-3191	-2371	163
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
164(V)	-770	-686	-2514	-1945	-447	-2242	-1009	-25	-1638	907	143	1079	-2296	-1373	-1632	-1314	-714	2075	-994	1702	164
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
165(Y)	-4022	-3300	-4202	-4471	-162	-3952	-1522	-3554	-4217	-3007	-3045	-3501	-4179	-3602	-3853	-3928	-4069	-3631	-835	4786	165
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
166(E)	-538	-1578	-594	1619	-1749	-1591	-271	182	64	-466	-709	-313	-1679	970	667	560	-477	169	-1859	-1283	166
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
167(Q)	-583	-1987	-408	1304	-2289	-1544	-207	-2017	1086	-1990	-1087	-190	-1644	2012	-258	724	-522	445	-2178	-1519	167
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
168(R)	-2066	-2822	-1934	-1456	-3549	-2571	-1068	-3277	108	-3049	-2371	1548	-2762	-689	3583	-2003	-1971	-2938	-2872	-2605	168
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

31

[illegible]

-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												
175(N)	355	-2563	2286	-73	-3075	-1686	-794	-2864	-638	-2825	-1980	2330	-2039	-399	-1233	944	-1176	-2366	-3020	-2262	175
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-12	-7442	-8484	-894	-1115	-701	-1378	*	*												

321

LISTAGEM DE SEQUÊNCIAS

<110> E.I. du Pont de Nemours and Co.

<120> Produção fermentativa de álcoois de quatro carbonos

<130> CL3893 PCT

<160> 275

<170> PatentIn versão 3.4

<210> 1

<211> 780

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 1

```

atgaatcatt ctgctgaatg cacctgcgaa gagagtctat gcgaaaccct gcgggcgttt      60
tccgcgcagc atcccgagag cgtgctctat cagacatcgc tcatgagcgc cctgctgagc      120
ggggttttacg aaggcagcac caccatcgcg gacctgctga aacacggcga tttcggcctc      180
ggcaccttta atgagctgga cggggagctg atcgccttca gcagtcaggt ctatcagctg      240
cgcgccgacg gcagcgcgcg caaagcccag ccggagcaga aaacgccgtt cgcggtgatg      300
acctggttcc agccgcagta ccggaaaacc ttgaccatc cggtagaccg ccagcagctg      360
cacgaggtag tcgaccagca aatcccctct gacaacctgt tctgcgccct gcgcatcgac      420
ggccatttcc gccatgccc aaccgcacc gtgccgcgcc agacgccgcc gtaccgggcg      480
atgaccgacg tcctcgacga tcagccggtg ttccgcttta accagcgcga aggggtgctg      540
gtcggcttcc ggaccccgca gcatatgcag gggatcaacg tcgccgggta tcacgagcac      600
tttattaccg atgaccgcaa aggcggcggt cacctgctgg attaccagct cgaccatggg      660
gtgctgacct tcggcgaaat tcacaagctg atgatcgacc tgcccgccga cagcgcgttc      720
ctgcaggcta atctgcatcc cgataatctc gatgccgcca tccgttccgt agaaagttaa      780

```

<210> 2

<211> 259

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 2

```

Met Asn His Ser Ala Glu Cys Thr Cys Glu Glu Ser Leu Cys Glu Thr
1           5           10           15

```

Leu Arg Ala Phe Ser Ala Gln His Pro Glu Ser Val Leu Tyr Gln Thr
20 25 30

Ser Leu Met Ser Ala Leu Leu Ser Gly Val Tyr Glu Gly Ser Thr Thr
35 40 45

Ile Ala Asp Leu Leu Lys His Gly Asp Phe Gly Leu Gly Thr Phe Asn
50 55 60

Glu Leu Asp Gly Glu Leu Ile Ala Phe Ser Ser Gln Val Tyr Gln Leu
65 70 75 80

Arg Ala Asp Gly Ser Ala Arg Lys Ala Gln Pro Glu Gln Lys Thr Pro
85 90 95

Phe Ala Val Met Thr Trp Phe Gln Pro Gln Tyr Arg Lys Thr Phe Asp
100 105 110

His Pro Val Ser Arg Gln Gln Leu His Glu Val Ile Asp Gln Gln Ile
115 120 125

Pro Ser Asp Asn Leu Phe Cys Ala Leu Arg Ile Asp Gly His Phe Arg
130 135 140

His Ala His Thr Arg Thr Val Pro Arg Gln Thr Pro Pro Tyr Arg Ala
145 150 155 160

Met Thr Asp Val Leu Asp Asp Gln Pro Val Phe Arg Phe Asn Gln Arg
165 170 175

Glu Gly Val Leu Val Gly Phe Arg Thr Pro Gln His Met Gln Gly Ile
180 185 190

Asn Val Ala Gly Tyr His Glu His Phe Ile Thr Asp Asp Arg Lys Gly
195 200 205

Gly Gly His Leu Leu Asp Tyr Gln Leu Asp His Gly Val Leu Thr Phe
210 215 220

Gly Glu Ile His Lys Leu Met Ile Asp Leu Pro Ala Asp Ser Ala Phe
225 230 235 240

Leu Gln Ala Asn Leu His Pro Asp Asn Leu Asp Ala Ala Ile Arg Ser
245 250 255

Val Glu Ser

<210> 3
<211> 1680
<212> DNA
<213> *Klebsiella pneumoniae*

<400> 3
atggacaaac agtatccggt acgccagtgg gcgcacggcg ccgatctcgt cgtcagtcag 60
ctggaagctc agggagtagc ccagggtgttc ggcattccccg gcgccaaaat tgacaaggctc 120
ttcgactcac tgctggattc ctcgattcgc attattccgg tacgccacga agccaacgcc 180
gcgttttatgg ccgccgccgt cggacgcatt accggcaaaag cgggcgtggc gctggtcacc 240
tccggtccgg gctgttccaa cctgatcacc ggcattggcca ccgcgaacag cgaaggcgac 300
ccggtggtgg ccctgggagg cgcggtaaaa cgcgccgata aagcgaagca ggtccaccag 360
agtatggata cgggtggcgat gttcagcccc gtcaccaaata acgccgtcga ggtgacggcg 420
ccggatgcgc tggcggaagt ggtctccaac gccttccgcg ccgccgagca gggccggccg 480
ggcagcgcgt tcgttagcct gccgcaggat gtggtcgatg gcccggtcag cggcaaagtg 540
ctgccggcca gcggggcccc gcagatgggc gccgcgccgg atgatgccat cgaccagggtg 600
gcgaagctta tcgcccaggc gaagaacccg atcttctctg tcggcctgat ggccagccag 660
ccggaaaaca gcaaggcgct gcgccgtttg ctggagacca gccatattcc agtcaccagc 720
acctatcagg ccgccggagc ggtgaatcag gataacttct ctgcttcgc cggccgggtt 780
gggctgttta acaaccaggc cggggaccgt ctgctgcagc tcgccgacct ggtgatctgc 840
atcggtctaca gcccgggtga atacgaaccg gcgatgtgga acagcgga cgcgacgctg 900
gtgcacatcg acgtgctgcc cgcctatgaa gagcgcaact acaccccgga tgtcgagctg 960
gtgggcgata tcgccggcac tctcaacaag ctggcgcaaa atatcgatca tcggctgggtg 1020
ctctccccgc agggggcgga gatcctccgc gaccgccagc accagcgcgga gctgctggac 1080
cgccgcggcg cgcagctgaa ccagtttgcc ctgcatccgc tgcgcatcgt tcgcgccatg 1140
caggacatcg tcaacagcga cgtcacgttg accgtggaca tgggcagctt ccatatctgg 1200
attgcccgtt acctgtacag cttccgcgcc cgtcagggtg tgatctccaa cggccagcag 1260
accatgggag tcgccctgcc ctgggctatc ggcgcctggc tgggtcaatcc tgagcgaaaa 1320
gtgggtctccg tctccggcga cggcggttct ctgcagtcga gcatggagct ggagaccgcc 1380
gtccgcctga aagccaacgt actgcacctg atctgggtcg ataacggcta caacatgggtg 1440
gccattcagg aagagaaaaa ataccagcgc ctgtccggcg tcgagttcgg gccgatggat 1500
tttaaagcct atgccgaatc cttcggcgcg aaaggggttg ccgtggaaaag cgccgaggcg 1560
ctggagccga ccctgcacgc ggcgatggac gtcgacggcc cggcggtggg ggccattccg 1620

gtggattatc gcgataaccc gctgctgatg ggccagctgc atctgagtca gattctgtaa 1680

<210> 4

<211> 559

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 4

Met Asp Lys Gln Tyr Pro Val Arg Gln Trp Ala His Gly Ala Asp Leu
1 5 10 15

Val Val Ser Gln Leu Glu Ala Gln Gly Val Arg Gln Val Phe Gly Ile
20 25 30

Pro Gly Ala Lys Ile Asp Lys Val Phe Asp Ser Leu Leu Asp Ser Ser
35 40 45

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

Ala Ala Val Gly Arg Ile Thr Gly Lys Ala Gly Val Ala Leu Val Thr
65 70 75 80

Ser Gly Pro Gly Cys Ser Asn Leu Ile Thr Gly Met Ala Thr Ala Asn
85 90 95

Ser Glu Gly Asp Pro Val Val Ala Leu Gly Gly Ala Val Lys Arg Ala
100 105 110

Asp Lys Ala Lys Gln Val His Gln Ser Met Asp Thr Val Ala Met Phe
115 120 125

Ser Pro Val Thr Lys Tyr Ala Val Glu Val Thr Ala Pro Asp Ala Leu
130 135 140

Ala Glu Val Val Ser Asn Ala Phe Arg Ala Ala Glu Gln Gly Arg Pro
145 150 155 160

Gly Ser Ala Phe Val Ser Leu Pro Gln Asp Val Val Asp Gly Pro Val
165 170 175

Ser Gly Lys Val Leu Pro Ala Ser Gly Ala Pro Gln Met Gly Ala Ala
180 185 190

Pro Asp Asp Ala Ile Asp Gln Val Ala Lys Leu Ile Ala Gln Ala Lys
195 200 205

339

Asn Pro Ile Phe Leu Leu Gly Leu Met Ala Ser Gln Pro Glu Asn Ser
210 215 220

Lys Ala Leu Arg Arg Leu Leu Glu Thr Ser His Ile Pro Val Thr Ser
225 230 235 240

Thr Tyr Gln Ala Ala Gly Ala Val Asn Gln Asp Asn Phe Ser Arg Phe
245 250 255

Ala Gly Arg Val Gly Leu Phe Asn Asn Gln Ala Gly Asp Arg Leu Leu
260 265 270

Gln Leu Ala Asp Leu Val Ile Cys Ile Gly Tyr Ser Pro Val Glu Tyr
275 280 285

Glu Pro Ala Met Trp Asn Ser Gly Asn Ala Thr Leu Val His Ile Asp
290 295 300

Val Leu Pro Ala Tyr Glu Glu Arg Asn Tyr Thr Pro Asp Val Glu Leu
305 310 315 320

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

His Arg Leu Val Leu Ser Pro Gln Ala Ala Glu Ile Leu Arg Asp Arg
340 345 350

Gln His Gln Arg Glu Leu Leu Asp Arg Arg Gly Ala Gln Leu Asn Gln
355 360 365

Phe Ala Leu His Pro Leu Arg Ile Val Arg Ala Met Gln Asp Ile Val
370 375 380

Asn Ser Asp Val Thr Leu Thr Val Asp Met Gly Ser Phe His Ile Trp
385 390 395 400

Ile Ala Arg Tyr Leu Tyr Ser Phe Arg Ala Arg Gln Val Met Ile Ser
405 410 415

Asn Gly Gln Gln Thr Met Gly Val Ala Leu Pro Trp Ala Ile Gly Ala
420 425 430

Trp Leu Val Asn Pro Glu Arg Lys Val Val Ser Val Ser Gly Asp Gly
435 440 445

340

Gly Phe Leu Gln Ser Ser Met Glu Leu Glu Thr Ala Val Arg Leu Lys
450 455 460

Ala Asn Val Leu His Leu Ile Trp Val Asp Asn Gly Tyr Asn Met Val
465 470 475 480

Ala Ile Gln Glu Glu Lys Lys Tyr Gln Arg Leu Ser Gly Val Glu Phe
485 490 495

Gly Pro Met Asp Phe Lys Ala Tyr Ala Glu Ser Phe Gly Ala Lys Gly
500 505 510

Phe Ala Val Glu Ser Ala Glu Ala Leu Glu Pro Thr Leu His Ala Ala
515 520 525

Met Asp Val Asp Gly Pro Ala Val Val Ala Ile Pro Val Asp Tyr Arg
530 535 540

Asp Asn Pro Leu Leu Met Gly Gln Leu His Leu Ser Gln Ile Leu
545 550 555

<210> 5
<211> 771
<212> DNA
<213> *Klebsiella pneumoniae*

<400> 5
atgaaaaaag tcgcacttgt taccggcgcc ggccaggga ttggtaaagc tatcgccctt 60
cgtctggtga aggatggatt tgccgtggcc attgccgatt ataacgacgc caccgccaaa 120
gcggtcgctt cggaaatcaa ccaggccggc ggacacgccg tggcggtgaa agtggatgtc 180
tccgaccgag atcaggtatt tgccgccgtt gaacaggcgc gcaaaacgct gggcggtctt 240
gacgtcatcg tcaataacgc cgggtgtggca ccgtctacgc cgatcgagtc cattaccccg 300
gagattgtcg acaaagtcta caacatcaac gtcaaagggg tgatctgggg tattcaggcg 360
gcggtcgagg cctttaagaa agaggggcac ggcgggaaaa tcatcaacgc ctgttcccag 420
gccggccacg tcggcaaccc ggagctggcg gtgtatagct ccagtaaatt cgcggtacgc 480
ggcttaaccc agaccgccgc tcgcgacctc gcgccgctgg gcatcacggt caacggctac 540
tgcccgggga ttgtcaaaac gccaatgtgg gccgaaattg accgccaggt gtccgaagcc 600
gccggtaaac cgctgggcta cggtagcgcc gagttcgcca aacgcatcac tctcggtcgt 660
ctgtccgagc cggaagatgt cgccgcctgc gtctcctatc ttgccagccc ggattctgat 720
tacatgaccg gtcagtcggt gctgatcgac ggcgggatgg tatttaacta a 771

341
A

<210> 6
 <211> 256
 <212> PRT
 <213> *Klebsiella pneumoniae*

<400> 6

Met Lys Lys Val Ala Leu Val Thr Gly Ala Gly Gln Gly Ile Gly Lys
 1 5 10 15

Ala Ile Ala Leu Arg Leu Val Lys Asp Gly Phe Ala Val Ala Ile Ala
 20 25 30

Asp Tyr Asn Asp Ala Thr Ala Lys Ala Val Ala Ser Glu Ile Asn Gln
 35 40 45

Ala Gly Gly His Ala Val Ala Val Lys Val Asp Val Ser Asp Arg Asp
 50 55 60

Gln Val Phe Ala Ala Val Glu Gln Ala Arg Lys Thr Leu Gly Gly Phe
 65 70 75 80

Asp Val Ile Val Asn Asn Ala Gly Val Ala Pro Ser Thr Pro Ile Glu
 85 90 95

Ser Ile Thr Pro Glu Ile Val Asp Lys Val Tyr Asn Ile Asn Val Lys
 100 105 110

Gly Val Ile Trp Gly Ile Gln Ala Ala Val Glu Ala Phe Lys Lys Glu
 115 120 125

Gly His Gly Gly Lys Ile Ile Asn Ala Cys Ser Gln Ala Gly His Val
 130 135 140

Gly Asn Pro Glu Leu Ala Val Tyr Ser Ser Ser Lys Phe Ala Val Arg
 145 150 155 160

Gly Leu Thr Gln Thr Ala Ala Arg Asp Leu Ala Pro Leu Gly Ile Thr
 165 170 175

Val Asn Gly Tyr Cys Pro Gly Ile Val Lys Thr Pro Met Trp Ala Glu
 180 185 190

Ile Asp Arg Gln Val Ser Glu Ala Ala Gly Lys Pro Leu Gly Tyr Gly
 195 200 205

Thr Ala Glu Phe Ala Lys Arg Ile Thr Leu Gly Arg Leu Ser Glu Pro
 210 215 220

Glu Asp Val Ala Ala Cys Val Ser Tyr Leu Ala Ser Pro Asp Ser Asp
 225 230 235 240

Tyr Met Thr Gly Gln Ser Leu Leu Ile Asp Gly Gly Met Val Phe Asn
 245 250 255

<210> 7
 <211> 1665
 <212> DNA
 <213> *Klebsiella oxytoca*

<400> 7
 atgagatcga aaagatttga agcactggcg aaacgccctg tgaatcagga cggcttcggt 60
 aaggagtgga tcgaagaagg ctttatcgcg atggaaagcc cgaacgaccc aaaaccgtcg 120
 attaaaaatcg ttaacggcgc ggtgaccgag ctggacggga aaccggtaag cgattttgac 180
 ctgatcgacc actttatcgc ccgctacggt atcaacctga accgcgccga agaagtgatg 240
 gcgatggatt cgggtcaagct ggccaacatg ctgtgcgatc cgaacgttaa acgcagcgaa 300
 atcgtcccg c tgaccaccgc gatgacgccg gcgaaaattg tcgaagtggg ttcgcatatg 360
 aacgtcgtcg agatgatgat ggcgatgcag aaaatgcgcg cccgccgcac cccgtcccag 420
 caggcgcacg tcaccaacgt caaagataac ccggtacaga ttgccgccga cgccgccgaa 480
 ggggcatggc gcggatttga cgaacaggaa accaccgttg cggtagcgcg ctatgcgccg 540
 ttcaacgcca tcgcgctgct ggtgggctcg caggtaggcc gtccgggctg gctgacgcag 600
 tgctcgtgga aagaagccac cgagctgaag ctccggcatgc tgggccacac ctgctacgcc 660
 gaaaccatct ccgtctacgg caccgagccg gtctttaccg acggcgacga cacgccgtgg 720
 tcgaagggct tcctcgctc gtctacgcc tctcgcgggc tgaaaatgcg ctttacctcc 780
 ggctccggct cggaagtgca gatgggctac gccgaaggca aatccatgct ttatctggaa 840
 gcgcgctgca tctacatcac caaagccgcg ggcgtagcagg gtctgcaaaa cggttccgta 900
 agctgcatcg gcgtgccgtc tgcggtgcct tccggcattc gcgcggtgct ggcggaaaac 960
 ctgatctggt cgctcgtgga tctggagtgc gcctccagca acgaccagac cttcaccac 1020
 tccgatatgc gtcgtaccgc gcgcctgctg atgcagttcc tgccgggcac cgactttatc 1080
 tcctccgggt attccgcggg gccgaactac gacaacatgt tcgccggctc caacgaagat 1140
 gccgaagact ttgacgacta caacgtcatc cagcgcgacc tgaaggtgga cggcggtttg 1200
 cgtccggttc gcgaagagga cgtcatcgcc atccgtaaca aagccgcccg cgcgctgcag 1260
 gccgtgtttg ccggaatggg gctgccgccg attacgatg aagaagtga agccgcgacc 1320
 tacgcccacg gttcgaaaga tatgccggag cgcaacatcg tcgaagacat caagttcgcc 1380

caggaaatca tcaataaaaa ccgcaacggt ctggaagtgg tgaaagcgct ggcgcagggc 1440
 ggattcaccg acgtggccca ggacatgctc aacatccaga aagctaagct gaccggggac 1500
 tacctgcata cctccgcgat tatcgtcggc gacgggcagg tgctgtcagc cgtcaacgac 1560
 gtcaacgact atgccggtcc ggcaacgggc tatcgctgc agggcgaacg ctgggaagag 1620
 attaaaaaca tccctggcgc tcttgatccc aacgagattg attaa 1665

<210> 8
 <211> 554
 <212> PRT
 <213> *Klebsiella oxytoca*

<400> 8

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Gly Phe Ile Ala Met Glu
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Ala Val
 35 40 45

Thr Glu Leu Asp Gly Lys Pro Val Ser Asp Phe Asp Leu Ile Asp His
 50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Asn Arg Ala Glu Glu Val Met
 65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
 85 90 95

Lys Arg Ser Glu Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
 100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ala
 115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
 130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
 145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
 165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

345

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Gly Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 9
<211> 675
<212> DNA
<213> *Klebsiella oxytoca*

<400> 9
atggaaatta atgaaaaatt gctgcgccag ataattgaag acgtgctcag cgagatgaag 60
ggcagcgata aaccggtctc gtttaatgcg ccggcggcct ccgcggcgcc ccaggccacg 120
ccgcccgcg gcgacggctt cctgacggaa gtgggcgaag cgcgtcaggg aaccacgacg 180
gacgaagtga ttatcgccgt cggcccggct ttcggcctgg cgcagaccgt caatatcgtc 240
ggcatccgc ataagagcat ttgcgcgaa gtcattgccg gtattgaaga agaaggcatt 300
aaggcgcgcg tgattcgctg ctttaaattc tccgacgtgg ccttcgtcgc cgttgaaggt 360
aatcgctga gcggctccg catctctatc ggcatccagt cgaaaggcac cacggtgatc 420
caccagcagg ggctgccgcc gctctctaac ctggagctgt tcccgcaggc gccgctgctg 480
accctggaaa cctatcgcca gatcgga aaaacgccgcc gctatgcgaa acgcgaatcg 540

346
JL

ccgcagccgg tcccgcgct gaatgaccag atggcgcggc cgaagtacca ggcgaaatcg 600
 gccattttgc acattaaaga gaccaagtac gtggtgacgg gcaaaaaccc gcaggaactg 660
 cgcgtggcgc ttgga 675

<210> 10
 <211> 224
 <212> PRT
 <213> *Klebsiella oxytoca*

<400> 10

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
 1 5 10 15

Ser Glu Met Lys Gly Ser Asp Lys Pro Val Ser Phe Asn Ala Pro Ala
 20 25 30

Ala Ser Ala Ala Pro Gln Ala Thr Pro Pro Ala Gly Asp Gly Phe Leu
 35 40 45

Thr Glu Val Gly Glu Ala Arg Gln Gly Thr Gln Gln Asp Glu Val Ile
 50 55 60

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

Gly Ile Pro His Lys Ser Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
 85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
 115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
 130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
 145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
 165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
 180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 11
<211> 522
<212> DNA
<213> *Klebsiella oxytoca*

<400> 11
atgaataaccg acgcaattga atcgatggta cgcgacgtat tgagccgcat gaacagcctg 60
cagggcgagg cgctgcggc ggctccggcg gctggcggcg cgtcccgtag cgccagggtc 120
agcgactacc cgctggcgaa caagcaccgc gaatgggtga aaaccgccac caataaaacg 180
ctggacgact ttacgctgga aaacgtgctg agcaataaag tcaccgcca ggatatgcgt 240
attaccccg aaaccctgcg cttacaggct tctattgcc aagacgcggg ccgcgaccgg 300
ctggcgatga acttcgagcg cgccgccgag ctgaccgcgg taccggacga tcgcattctt 360
gaaatctaca acgcccctcg cccctatcgc tcgacgaaag aggagctgct ggcgatcgcc 420
gacgatctcg aaagccgcta tcaggcgaag atttgcgccg ctttcgttcg cgaagcggcc 480
acgctgtacg tcgagcgtaa aaaactcaaa ggcgacgatt aa 522

<210> 12
<211> 173
<212> PRT
<213> *Klebsiella oxytoca*

<400> 12

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
1 5 10 15

Met Asn Ser Leu Gln Gly Glu Ala Pro Ala Ala Ala Pro Ala Ala Gly
20 25 30

Gly Ala Ser Arg Ser Ala Arg Val Ser Asp Tyr Pro Leu Ala Asn Lys
35 40 45

His Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe
50 55 60

Thr Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg
65 70 75 80

348
A

Ile Thr Pro Glu Thr Leu Arg Leu Gln Ala Ser Ile Ala Lys Asp Ala
85 90 95

Gly Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr
100 105 110

Ala Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
115 120 125

Tyr Arg Ser Thr Lys Glu Glu Leu Leu Ala Ile Ala Asp Asp Leu Glu
130 135 140

Ser Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala
145 150 155 160

Thr Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 13
<211> 1041
<212> DNA
<213> *Rhodococcus ruber*

<400> 13
atgaaagccc tccagtacac cgagatcggc tccgagccgg tcgtcgtcga cgtccccacc 60
ccggcgcccc ggccgggtga gatcctgctg aaggtcaccg cggccggctt gtgccactcg 120
gacatcttcg tgatggacat gccggcagag cagtacatct acggtcttcc cctcaccctc 180
ggccacgagg gcgtcggcac cgtcgccgaa ctcggcgccg gcgtcaccgg attcgagacg 240
ggggacgccc tcgccgtgta cgggcccgtg ggggtgcggtg cgtgccacgc gtgcgcgcgc 300
ggccgggaga actactgcac ccgcgccgcc gagctgggca tcaccccgcc cgggtctcggc 360
tcgccgggt cgatggccga gtacatgata gtcgactcgg cgcgccacct cgtcccgatc 420
ggggacctcg accccgtcgc ggcggttccg ctcaccgacg cgggcctgac gccgtaccac 480
gcgatctcgc gggctctgcc cctgctggga cccggctcga ccgcggtcgt catcggggtc 540
ggcggactcg ggcacgtcgg catccagatc ctgcgcgccg tcagcgcggc ccgcgtgatc 600
gccgtcgatc tcgacgacga ccgactcgcg ctcgcccgcg aggtcggcgc cgacgcggcg 660
gtgaagtccg gcgccggggc ggccggacgc atccgggagc tgaccggcgg tgagggcgcg 720
acggcggtgt tcgacttcgt cggcgccag tcgacgatcg acacggcgca gcagggtggtc 780
gcgatcgacg ggcacatctc ggtggtcggc atccatgccg gcgcccacgc caaggctcggc 840
ttcttcatga tcccgttcgg cgcgtccgtc gtgacgccgt actggggcac gcggtccgag 900

ctgatggacg tcgtggacct ggcccgtgcc ggccggctcg acatccacac cgagacgttc 960
 accctcgacg agggaccac ggctaccgg cggctacgag agggcagcat ccgcggccgc 1020
 ggggtggtcg tcccgggctg a 1041

<210> 14
 <211> 346
 <212> PRT
 <213> *Klebsiella oxytoca*

<400> 14

Met Lys Ala Leu Gln Tyr Thr Glu Ile Gly Ser Glu Pro Val Val Val
 1 5 10 15

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

Thr Ala Ala Gly Leu Cys His Ser Asp Ile Phe Val Met Asp Met Pro
 35 40 45

Ala Glu Gln Tyr Ile Tyr Gly Leu Pro Leu Thr Leu Gly His Glu Gly
 50 55 60

Val Gly Thr Val Ala Glu Leu Gly Ala Gly Val Thr Gly Phe Glu Thr
 65 70 75 80

Gly Asp Ala Val Ala Val Tyr Gly Pro Trp Gly Cys Gly Ala Cys His
 85 90 95

Ala Cys Ala Arg Gly Arg Glu Asn Tyr Cys Thr Arg Ala Ala Glu Leu
 100 105 110

Gly Ile Thr Pro Pro Gly Leu Gly Ser Pro Gly Ser Met Ala Glu Tyr
 115 120 125

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

Pro Val Ala Ala Val Pro Leu Thr Asp Ala Gly Leu Thr Pro Tyr His
 145 150 155 160

Ala Ile Ser Arg Val Leu Pro Leu Leu Gly Pro Gly Ser Thr Ala Val
 165 170 175

Val Ile Gly Val Gly Gly Leu Gly His Val Gly Ile Gln Ile Leu Arg
 180 185 190

Ala Val Ser Ala Ala Arg Val Ile Ala Val Asp Leu Asp Asp Asp Arg
195 200 205

Leu Ala Leu Ala Arg Glu Val Gly Ala Asp Ala Ala Val Lys Ser Gly
210 215 220

Ala Gly Ala Ala Asp Ala Ile Arg Glu Leu Thr Gly Gly Glu Gly Ala
225 230 235 240

Thr Ala Val Phe Asp Phe Val Gly Ala Gln Ser Thr Ile Asp Thr Ala
245 250 255

Gln Gln Val Val Ala Ile Asp Gly His Ile Ser Val Val Gly Ile His
260 265 270

Ala Gly Ala His Ala Lys Val Gly Phe Phe Met Ile Pro Phe Gly Ala
275 280 285

Ser Val Val Thr Pro Tyr Trp Gly Thr Arg Ser Glu Leu Met Asp Val
290 295 300

Val Asp Leu Ala Arg Ala Gly Arg Leu Asp Ile His Thr Glu Thr Phe
305 310 315 320

Thr Leu Asp Glu Gly Pro Thr Ala Tyr Arg Arg Leu Arg Glu Gly Ser
325 330 335

Ile Arg Gly Arg Gly Val Val Val Pro Gly
340 345

<210> 15
<211> 29
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 15
caccatggac aaacagtatc cggtacgcc

29

<210> 16
<211> 25
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 16
cgaagggcga tagctttacc aatcc 25

<210> 17
<211> 32
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 17
caccatgaat cattctgctg aatgcacctg cg 32

<210> 18
<211> 22
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 18
gatactgttt gtccatgtga cc 22

<210> 19
<211> 28
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 19
caccatgaaa aaagtcgcac ttgttacc 28

<210> 20
<211> 15
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 20
ttagttaaat accat 15

<210> 21
<211> 23
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 21
caccatgaga tcgaaaagat ttg 23

<210> 22
<211> 22
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 22
cttagagaag ttaatcgtcg cc 22

— <210> 23
<211> 25
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 23
caccatgaaa gccctccagt acacc 25

<210> 24
<211> 18
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 24
cgtcgtgtca tgcccggg 18

<210> 25
<211> 36
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 25
gatcgaattc gtttaaactt agttttctac cgcacg 36

<210> 26
<211> 34
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 26
gatcgcatgc aagctttcat atagtcggaa ttcc 34

<210> 27
<211> 43
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 27
gatcgaattc gtttaaacaaggaggtctg attcatgaga tcg 43

<210> 28
<211> 22
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 28
gatcggattc ttaatcgtcg cc 22

<210> 29
<211> 36
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 29
gatcggatcc aaaggaggtc gggcgcgcatga aagccc 36

<210> 30
<211> 32
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 30
gatctctaga aagctttcag cccgggacga cc 32

<210> 31
<211> 21
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

354

<400> 31
acttttctttc gcctgtttca c 21

<210> 32
<211> 79
<212> DNA
<213> Seqüência artificial

<220>
<223> Iniciador

<400> 32
catgaagctt gtttaaactc ggtgaccttg aaaataatga aaacttatat tgttttgaaa 60
ataatgaaaa ctttatattg 79

<210> 33
<211> 39
<212> DNA
<213> Seqüência artificial

<220>
<223> Iniciador BABC F

<400> 33
gagctcgaat tcaaaggagg aagtgtatat gaatcattc 39

<210> 34
<211> 35
<212> DNA
<213> Seqüência artificial

<220>
<223> Iniciador BAB R

<400> 34
ggatcctcta gaattagtta aataccatcc cgccg 35

<210> 35
<211> 17
<212> DNA
<213> Seqüência artificial

<220>
<223> Iniciador M13 forward

<400> 35
gtaaaacgac ggccagt 17

<210> 36
<211> 16
<212> DNA
<213> Seqüência artificial

<220>

<223> Iniciador M13 Reverse

<400> 36
aacagctatg accatg

16

<210> 37
<211> 20
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador N83 SeqF2

<400> 37
gctggattac cagctcgacc

20

<210> 38
<211> 20
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador N83SeqF3

<400> 38
cggacgcatt accggcaaag

20

<210> 39
<211> 20
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador N84 SeqR4

<400> 39
cgaagcgaga gaagttatcc

20

<210> 40
<211> 38
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador BC Spe F

<400> 40
actagtaaag gaggaaagag tatgaagaag gtcgcact

38

<210> 41
<211> 26
<212> DNA
<213> Sequência artificial

<220>

<223> Iniciador BC Xba R

<400> 41

tctagaaagc aggggcaagc catgtc

26

<210> 42

<211> 20

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador Trc F

<400> 42

ttgacaatta atcatccggc

20

<210> 43

<211> 20

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador Trc R

<400> 43

cttctctcat ccgcaaaac

20

<210> 44

<211> 38

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador DDo For

<400> 44

aagcttaaag gaggctgatt catgagatcg aaaagatt

38

<210> 45

<211> 27

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador DDo Rev

<400> 45

tctagattat tcctcctgct gttctcc

27

<210> 46

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador DDko seq F2

<400> 46

gcatggcgcg gatttgacga ac

22

<210> 47

<211> 22

<212> DNA

<213> Seqüência artificial

<220>

<223> Iniciador DDko seq F5

<400> 47

cattaaagag accaagtacg tg

22

<210> 48

<211> 24

<212> DNA

<213> Seqüência artificial

<220>

<223> Iniciador DDko seq F7

<400> 48

atatcctggt ggtgtcgctcg gcgt

24

<210> 49

<211> 22

<212> DNA

<213> Seqüência artificial

<220>

<223> Iniciador DDko seq F9

<400> 49

tctttgtcac caacgccctg cg

22

<210> 50

<211> 22

<212> DNA

<213> Seqüência artificial

<220>

<223> Iniciador DDko seq R1

<400> 50

gcccaccgcg ctcgccgccg cg

22

<210> 51

<211> 22

<212> DNA

<213> Seqüência artificial

<220>

<223> Iniciador DDko seq R3

<400> 51

ccccaggat ggcggcttcg gc

22

<210> 52

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador DDko seq R7

<400> 52

gggccgacgg cgataatcac tt

22

<210> 53

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador DDko seq R10

<400> 53

ttcttcgatc cactccttaa cg

22

<210> 54

<211> 56

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador ChnA F

<400> 54

catcaattga ctacgtagtc gtacgtgtaa ggaggtttga aatggaaaaa attatg

56

<210> 55

<211> 40

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador ChnA R

<400> 55

catgctagcc ccgggtatct tctactcatt ttttatttcg

40

<210> 56

<211> 22

<212> DNA

<213> Artificial Squence

<220>

<223> Iniciador chnSeq F1

<400> 56

ctcaacaggg tgtaagtgtgta gt

22

<210> 57

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador chnSeq R1

<400> 57

cgttttgata tagccaggat gt

22

<210> 58

<211> 35

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador Top ter F1

<400> 58

ctagaagtca aaagcctccg accggaggct tttga

35

<210> 59

<211> 60

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador Top ter F2

<400> 59

ctgctcgagt tgctagcaag tttaaaca aaagcccg ctcattaggc gggctgagct

60

<210> 60

<211> 37

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador Bot ter R1

<400> 60

cagccgcct aatgagcggg ctttttttg tttaaac

37

<210> 61

<211> 50

<212> DNA

<213> Sequência artificial

<220>

360
SA

<223> Iniciador Bot ter R2

<400> 61

ttgctagcaa ctcgagcagt caaaagcctc cggtcggagg cttttgactt

50

<210> 62

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador pCL1925 vec F

<400> 62

cggtatcatc aacaggctta cc

22

<210> 63

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador pCL1925 vec R1

<400> 63

agggttttcc cagtcacgac gt

22

<210> 64

<211> 22

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador pCL1925 vec R2

<400> 64

cgcaatagtt ggcgaagtaa tc

22

<210> 65

<211> 20

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador N84 Seq R2

<400> 65

gcatcgagat tatcgggatg

20

<210> 66

<211> 208

<212> DNA

<213> *Escherichia coli*

<400> 66

361
A

atcgcccgca ttcttgccgc atcttcccc ggcgtcacac cgaagtaacg tttaaactca 60
 cggctgtgta ggctggagct gcttcgaagt tcctatactt tctagagaat aggaacttcg 120
 gaataggaac taaggaggat attcatatga ttacgttgga tgtcagccgc cgtatatacg 180
 aagccgcccc ctaagctttt tacgcctc 208

<210> 67
 <211> 42
 <212> DNA
 <213> Sequência Artificial
 <220>
 <223> Promotor 1.6GI Variante

<400> 67
 gcccttgaca atgccacatc ctgagcaaat aattcaacca ct 42

<210> 68
 <211> 42
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Promotor 1.5 GI

<400> 68
 gcccttgact atgccacatc ctgagcaaat aattcaacca ct 42

<210> 69
 <211> 3240
 <212> DNA
 <213> *Klebsiella oxytoca*

<400> 69
 ggcgcggtcc gccaggcggc cacctccgcg cgcgaaatcg gcaaaaccgt ccttgcgacc 60
 ctcggtgctg aaccgaaaaa cgatcgcccc tcctacatct gataccacag aggctgattc 120
 atgagatcga aaagatttga agcactggcg aaacgccctg tgaatcagga cggcttcggt 180
 aaggagtgga tcgaagaagg ctttatcgcg atggaaagcc cgaacgaccc aaaaccgctc 240
 attaaaatcg ttaacggcgc ggtgaccgag ctggacggga aaccggtaag cgattttgac 300
 ctgatcgacc actttatcgc ccgctacggt atcaacctga accgcgccga agaagtgatg 360
 gcgatggatt cgggtcaagct ggccaacatg ctgtgcatc cgaacgttaa acgcagcgaa 420
 atcgtccccg tgaccaccgc gatgacgccg gcgaaaattg tcgaagtggc ttcgcatatg 480
 aacgtcgctc agatgatgat ggcgatgcag aaaatgcgcg cccgccgcac cccgtcccag 540
 caggcgcacg tcaccaacgt caaagataac ccggtacaga ttgccgccga cgcgcgccga 600
 ggggcatggc gcggatttga cgaacaggaa accaccgttg cggtagcgcg ctatgcgccg 660

ttcaacgcca tcgcgctgct ggtgggctcg caggtaggcc gtccgggcgt gctgacgcag 720
 tgctcgctgg aagaagccac cgagctgaag ctccggcatgc tgggccacac ctgctacgcc 780
 gaaaccatct ccgctctacgg caccgagccg gtctttaccg acggcgacga cacgccgtgg 840
 tcgaagggct tcctcgctc gtctacgcc tctcgcgggc tgaaaatgcg ctttacctcc 900
 ggctccggct cggaagtgc gatgggctac gccgaaggca aatccatgct ttatctggaa 960
 gcgcgctgca tctacatcac caaagccgcg ggcgtacagg gtctgcaaaa cggttccgta 1020
 agctgcatcg gcgtgccgtc tgcggtgcct tccggcattc gcgcggtgct ggcggaaaac 1080
 ctgatctgtt cgtcgctgga tctggagtgc gcctccagca acgaccagac cttcaccac 1140
 tccgatatgc gtcgtaccgc gcgcctgctg atgcagttcc tgccgggcac cgactttatc 1200
 tcctccgggtt attccgcggg gccgaactac gacaacatgt tcgccggctc caacgaagat 1260
 gccgaagact ttgacgacta caacgtcatc cagcgcgacc tgaagggtgga cggcggtttg 1320
 cgtccgggtc gcgaagagga cgtcatcgcc atccgtaaca aagccgcccg cgcgctgcag 1380
 gccgtgtttg ccggaatggg gctgccgccg attaccgatg aagaagttga agccgcgacc 1440
 tacgccacg gttcgaaaga tatgccggag cgcaacatcg tcgaagacat caagttcgcc 1500
 caggaaatca tcaataaaaa ccgcaacggt ctggaagtgg tgaaagcgct ggcgcagggc 1560
 ggattcaccg acgtggccca ggacatgctc aacatccaga aagctaagct gaccggggac 1620
 tacctgcata cctccgcgat tatcgtcggc gacgggcagg tgctgtcagc cgtcaacgac 1680
 gtcaacgact atgccgggtc ggcaacgggc tatcgctgc agggcgaaacg ctgggaagag 1740
 attaaaaaca tccctggcgc tcttgatccc aacgagattg attaaggggt gagaaatgga 1800
 aattaatgaa aaattgctgc gccagataat tgaagacgtg ctacgcgaga tgaagggcag 1860
 cgataaaccg gtctcgttta atgcgccggc ggcctccgcg gcgccccagg ccacgccgcc 1920
 cgccggcgac ggcttcctga cggaagtggg cgaagcgcgt cagggaaccc agcaggacga 1980
 agtgattatc gccgtcggcc cggctttcgg cctggcgag accgtcaata tcgtcggcat 2040
 cccgcataag agcattttgc gcgaagtc attgccgtatt gaagaagaag gcattaaggc 2100
 gcgcgtgatt cgctgcttta aatcctccga cgtggccttc gtcgccgttg aaggtaatcg 2160
 cctgagcggc tccggcatct ctatcggcat ccagtcgaaa ggcaccacgg tgatccacca 2220
 gcaggggctg ccgccgtct ctaacctgga gctgttccc caggcgccgc tgctgaccct 2280
 ggaaacctat cgccagatcg gcaaaaacgc cggccgctat gcgaaacgcg aatcgccgca 2340
 gccggtcccg acgctgaatg accagatggc gcggccgaag taccaggcga aatcggccat 2400
 tttgcacatt aaagagacca agtacgtggt gacgggcaaa aaccgcgagg aactgcgcgt 2460

ggcgctttga taaaggataa ctccatgaat accgacgcaa ttgaatcgat ggtacgcgac 2520
 gtattgagcc gcatgaacag cctgcagggc gaggcgcctg cggcggctcc ggcggctggc 2580
 ggcgcgctccc gtagcgccag ggtcagcgac tacccgctgg cgaacaagca cccggaatgg 2640
 gtgaaaaccg ccaccaataa aacgctggac gactttacgc tggaaaacgt gctgagcaat 2700
 aaagtcaccg cccaggatat gcgtattacc ccggaaaccc tgcgcttaca ggcttctatt 2760
 gccaaagacg cgggcccgcga ccggctggcg atgaacttcg agcgcgccgc cgagctgacc 2820
 gcggtaccgg acgatcgcat tcttgaaatc tacaacgccc tccgccccta tcgctcgacg 2880
 aaagaggagc tgctggcgat cgccgacgat ctcgaaagcc gctatcaggc gaagatttgc 2940
 gccgctttcg ttcgcgaagc ggccacgctg tacgtcgagc gtaaaaaact caaaggcgac 3000
 gattaacttc tctaagtaat tcgagatgca ttgaggcggc aagtgagtga caaattcgtc 3060
 tggaacgaat ttgaacagcc ataggctggc tttagttagg gacagggatg tccctcataa 3120
 ccccgatgag cttactgtag taagtgattc gggtgaaaga acgcagccaa caaaaaggca 3180
 gtttgaagta cgacgagaaa aggggcatgt gatgcatat atagctggca ttgatatcgg 3240

<210> 70

<211> 2640

<212> DNA

<213> *Klebsiella oxytoca*

<400> 70

acgtcgagcg taaaaaactc aaaggcgacg attaaacttct ctaagtaatt cgagatgcat 60
 tgaggcggca agtgagtgc aaattcgtct ggaacgaatt tgaacagcca taggctggct 120
 ttagtgaggg acagggatgt ccctcataac cccgatgagc ttactgtagt aagtgattcg 180
 ggtgaaagaa cgacgccaac aaaaaggcag tttgaagtac gacgagaaaa ggggcatgtg 240
 atgcatata tagctggcat tgatatcggc aactcatcga cggaagtcgc cctggcgacc 300
 ctggatgagg ctggcgcgct gacgatcacc cacagcgcg tggcggaaac caccggaatc 360
 aaaggcacgt tgcgtaacgt gttcgggatt caggaggcgc tcgccctcgt cgccagaggc 420
 gccgggatcg ccgtcagcga tatttcgctc atccgcatca acgaagcgac gccggtgatt 480
 ggcgatgtgg cgatggaaac cattaccgaa accatcatca ccgaatcgac catgatcggc 540
 cataaccga aaacgcccgg cggcgcgggg cttggcacag gcatcaccat tacgccgacg 600
 gagctgctaa cccgcccggc ggacgcgccc tatatcctgg tgggtgctgc ggcgttcgat 660
 tttgccgata tcgccagcgt gattaacgct tccctgcgcg ccgggtatca gattaccggc 720
 gtcattttac agcgcgacga tggcgtgctg gtcagcaacc ggctggaaaa accgctgccg 780
 atcggtgacg aagtgtgtga catcgaccgc attccgctgg ggatgctggc ggcgattgag 840

gtcgccgttc cggggaaggt catcgaaacc ctctctaacc cttacggcat cgccaccgtc 900
 tttaacctca gccccgagga gacgaagaac atcgtcccga tggccccggc gctgattggc 960
 aaccgttccg ccgtggtggt caaaacgccca tccggcgacg tcaaagcgcg cgcgataccc 1020
 gccggtaatc ttgagctgct ggcccagggc cgtagcgtgc gcgtggatgt ggccgccggc 1080
 gccgaagcca tcatgaaagc ggtcgacggc tgcggcaggc tcgataacgt caccggcgaa 1140
 tccggcacca atatcggcgg catgctggaa cacgtgcgcc agaccatggc cgagctgacc 1200
 aacaagccga gcagcgaaat atttattcag gacctgctgg ccgttgatac ctcggtaccg 1260
 gtgagcgta cggcggtct ggccggggag ttctcgtgg agcaggccgt gggcatcgcc 1320
 tcgatggtga aatcggatcg cctgcagatg gcaatgatcg cccgcgaaat cgagcagaag 1380
 ctcaatatcg acgtgcagat cggcggcgca gaggccgaag ccgccatcct gggggcgctg 1440
 accacgccgg gcaccacccg accgctggcg atcctcgacc tcggcgcggg ctccaccgat 1500
 gcctccatca tcaaccccaa aggcgacatc atcgccaccc atctcgccgg cgcaggcgac 1560
 atggtgacga tgattattgc ccgcgagctg gggctggaag accgctatct ggcggaagag 1620
 atcaagaagt acccgctggc taagggtgaa agcctgttcc atttacgcca cgaggacggc 1680
 agcgtgcagt tcttctccac gccgctgccg cccgccgtgt tcgcccgcgt ctgctggtg 1740
 aaagcggacg aactggtgcc gctgcccggc gatttagcgc tggaaaaagt gcgcgccatt 1800
 cgccgcagcg ccaaagagcg ggtctttgtc accaacgccc tgcgcgcgct gcgtcaggtc 1860
 agccccaccg gcaacattcg cgatattccg ttcgtggtgc tggtcggcgg ttcgtcgtg 1920
 gatttcgaag tcccgcagct ggtcaccgat gcgctggcgc actaccgcct ggttgccgga 1980
 cggggaaata ttcgcggcag cgaggggccc cgaaacgcgg tggccaccgg cctgattctc 2040
 tcctggcata aggagtttgc gcatgaacgg taatcacagc gccccggcca tcgcgatcgc 2100
 cgatcatcgac ggctgcgacg gcctgtggcg cgaagtgcgt ctgggtatcg aagaggaagg 2160
 tatccctttc cggtccagc atcaccggc cgagaggtc gtggacagcg cctggcaggc 2220
 ggcgcgcagc tcgccgctgc tgggtggcat cgctgcgac cgccatatgc tggtcgtgca 2280
 ctacaagaat ttaccgcgat cggcgccgct ttttacgtg atgcatcatc aggacagtca 2340
 ggcccatcgc aacaccggta ataacgcggc acggctggtc aaggggatcc ctttccggga 2400
 tctgaatagc gaagcaacag gagaacagca ggatgaataa cgactggga ctggttga 2460
 caaaagggtt agtgggcgcc attgaggccg ccgatgcgat ggtgaaatcc gccaacgtgc 2520
 agctggtcgg ctacgaaaaa attggctcgg gcctcgtcac cgtgatggtg cgcggcgacg 2580
 tcggcgcggt caaagcggcg gtagacgcgg gcagcgggc ggcgagcgcg gtgggcgaag 2640

365
JP

<210> 71
 <211> 756
 <212> DNA
 <213> *Acinetobacter sp.*

<400> 71
 atggaaaaaa ttatgtcaaa taaattcaac aataaagtcg ctttaattac tggcgctggt 60
 tcagggtattg gtaaaagcac cgcactgctt ttggctcaac aggggtgtaag tgtagtggtt 120
 tcagatatta acctggaagc agcacagaaa gttgtggacg aaattgtcgc tttaggcggg 180
 aaagcggctg cgaataaggc caatactgct gagcctgaag acatgaaagc tgcagtcgag 240
 tttgcgggtca gcacttttgg tgcactgcat ttggccttca ataatgcggg aattctgggt 300
 gaagttaact ccaccgaaga attgagcatt gaaggatggc gtcgtgtgat tgatgtgaac 360
 ttgaatgcgg ttttctacag catgcattat gaagttcctg caatcttggc cgcagggggc 420
 ggagcgaattg tcaataccgc ttctattgca ggcttgatcg ggattcaaaa tatttcagggc 480
 tatgtcgtcg caaaacatgg cgtaacgggt ctaacgaaag cggcggcatt ggaatatgca 540
 gataaagggg ttcgcattaa ttcagtacat cctggctata tcaaaacgcc ttgattgca 600
 gaatttgaag aagcagaaat ggtaaaacta catccgattg gtcgtttggg acagccggaa 660
 gaagttgctc aggttggtgc cttcctactt tctgatgatg cttcatttgt gaccggtagt 720
 cagtatgtgg tcgatggtgc atatacctcg aaataa 756

<210> 72
 <211> 251
 <212> PRT
 <213> *Acinetobacter sp.*

<400> 72

Met Glu Lys Ile Met Ser Asn Lys Phe Asn Asn Lys Val Ala Leu Ile
 1 5 10 15

Thr Gly Ala Gly Ser Gly Ile Gly Lys Ser Thr Ala Leu Leu Leu Ala
 20 25 30

Gln Gln Gly Val Ser Val Val Val Ser Asp Ile Asn Leu Glu Ala Ala
 35 40 45

Gln Lys Val Val Asp Glu Ile Val Ala Leu Gly Gly Lys Ala Ala Ala
 50 55 60

Asn Lys Ala Asn Thr Ala Glu Pro Glu Asp Met Lys Ala Ala Val Glu
 65 70 75 80

366
AA

Phe Ala Val Ser Thr Phe Gly Ala Leu His Leu Ala Phe Asn Asn Ala
85 90 95

Gly Ile Leu Gly Glu Val Asn Ser Thr Glu Glu Leu Ser Ile Glu Gly
100 105 110

Trp Arg Arg Val Ile Asp Val Asn Leu Asn Ala Val Phe Tyr Ser Met
115 120 125

His Tyr Glu Val Pro Ala Ile Leu Ala Ala Gly Gly Gly Ala Ile Val
130 135 140

Asn Thr Ala Ser Ile Ala Gly Leu Ile Gly Ile Gln Asn Ile Ser Gly
145 150 155 160

Tyr Val Ala Ala Lys His Gly Val Thr Gly Leu Thr Lys Ala Ala Ala
165 170 175

Leu Glu Tyr Ala Asp Lys Gly Ile Arg Ile Asn Ser Val His Pro Gly
180 185 190

Tyr Ile Lys Thr Pro Leu Ile Ala Glu Phe Glu Glu Ala Glu Met Val
195 200 205

Lys Leu His Pro Ile Gly Arg Leu Gly Gln Pro Glu Glu Val Ala Gln
210 215 220

Val Val Ala Phe Leu Leu Ser Asp Asp Ala Ser Phe Val Thr Gly Ser
225 230 235 240

Gln Tyr Val Val Asp Gly Ala Tyr Thr Ser Lys
245 250

<210> 73

<211> 17417

<212> DNA

<213> *Acinetobacter* sp.

<400> 73

ctagcattta cgcgtgaggt aggtgggtag gtctgtaatg tgaagatcta cgaggaaatc 60

ggcgtcatga cgtgaggtcc agcgaaccgt cttgcgtaat ccgtcattca tggtagagtaa 120

cattgcccgt atttcgcgtt cagtatatag cagaccagca tgattaacga gatcctgggt 180

attttagttcc ggacacccaa agtcccatgc ggtcgccaga tccagtaagt cgactacgac 240

ttgctcatct gtagccaacc ccgcaatcac ttccacaatt ttcacagtg gaaccggatt 300

gaagaaatgg aaacctgcga tacggccctg atgctgacac gcagatgcaa ttgagggtcac 360
 agatagttag gatgtatttg aaaccagaat agtttcttca gccacaatcc tttcaagctg 420
 tttaaacaaa gtttgcttga tttccagatt ttcaataatt gcttctacga ccagatcaac 480
 gccagcaacc tcttcaatgc tttccaagat aatcaatcgg gctaagggtat ccacaagctg 540
 ctgttcggtt aactttcctt tagcagctag tttgtgcaag gttactttta atttttccaa 600
 gccttgctca gcagcgccgg gtttagcatc aaataaacgg acctcaacac ccgcctgtgc 660
 tgcaatttgc gcaataccca ttcccattac gcctgtgcca atcaaggcca ttttttgaat 720
 cgtcatgact tattttcctt gatattgagg gcttcgcttt tcgaaaagg cattgacgcc 780
 ttctttttga tcttgtgtat caaataaaat ttggaaggct ttacgctcta atgccaaagc 840
 accatcgagt ggcataattg cacctagtgt tgtgacttct ttgatctgtt caacggcaat 900
 cggtgagagt tgggcaatct gtgtcgcaat ttcaaccgct ttagcaaggg tttgatcatc 960
 ctcaaccact tcggaaacca accccatttt gtcagcttct tctgcagaaa agatctttcc 1020
 tgttaacact atttgcatgg ctttaaactt ccctaccgca cgcagtaagc gttgggtacc 1080
 accagcacct ggcatacagg ccaatttgac ttcaggctga ccaaactggg ctgattttcc 1140
 ggcaataatg atgtctgcat gcattgcaag ttcacacca ccaccaatg catatccatt 1200
 cacagcagcc acaatcggtt tagggcaatc aataatggcc cgccagtact gttccgtatg 1260
 gcgtaaatac atgtctacgg tttttgcagt ggtgaagtcc cgcatatccg cacctgctgc 1320
 aaatactttt tcaccaccag taatgacaat tgcgcggact gtatcagatg cagcgagctg 1380
 ctcaaacatt gctgcgagct gttggcgcag ttccagattc aatgcatttc tagtatctgg 1440
 acgatgtagt tcaacaatgg ccacaccatt actttgaata tctaaattca atatttcatt 1500
 ttccataaca acctacatgt ttgcgatagc ggtttattta aaccaaatac acctgttttt 1560
 ttgcaacaat aaagcccaca ggaacatagt tttaaattaa aaattggcta aaaatattta 1620
 aaaaacacaa ataaaatacc gcacagcggg atttgatata aatattattg catttatttt 1680
 tccattctgt catattattt tcattccaaa gcattagatc acccctgcat gaagcagaga 1740
 tggctaaatt tacctatcta atacaagggc ttaaaaatga ttcgcgatca agacacatta 1800
 aatcagctgg ttgacatgat ccgtcagttt gtcgatggcg ttcttattcc caatgaagaa 1860
 attgttgcgg aaaccgatga aattccagct gaaatcgtgc agcaaatgaa agaactgggt 1920
 ctttttgggtc tcaccattcc tgaggaatat gagggctctt gcctgaccat ggaggaagag 1980
 gtttacattg catttgaact gggacgtacc tctcctgctt tccgttcact gatcggcact 2040
 aacaatggga tcggttcac caggcttaatt attgatggct ccgaagagca gaaacagtat 2100

tttttgccac gtctggcaag tggtgaaatt attggttcat tctgtttaac tgaacctgat 2160
 tccggttcag atgtgcctc tttaaaaacc acagcgggta aagatgggta tcattacatt 2220
 ttaaattggca ctaagcgta catcaccaat gcaccgcatg cgggtgtctt tactgtcatg 2280
 gcacgtacca gtaccgaaat taaaggtaca ggtggaattt cagcctttat cgtggacagt 2340
 aaaactcctg gtatttcctt gggtaaactg gataagaaga tgggccaaaa aggtgcacat 2400
 acctgtgatg tgatttttga aaactgtcgt attcctgcat ctgcactcat tgggtgggtgtt 2460
 gaaggtgtag gtttttaaac tgcaatgaag gtacttgata aaggccgtat tcatattgct 2520
 gcattaagtg taggtgctgc tacgcgtatg ctggaagatt ccctacaata tgccgttgag 2580
 cgcaaacagt ttggtcaagc gattgcgaac ttccagtga ttcaagggtat gttagccgat 2640
 tctaaagctg aaattttacgc agcaaaatgt atggtattag atgtgtcccg acttcgtgat 2700
 gctggacaga atgtcagcac ggaagcatct tgtgccaaaga tgtttgccac tgaaatgtgt 2760
 ggccgtgtcg cagatcgtgg cgtacagatc catggtgggtg cgggttatat cagtgaatat 2820
 gctattgagc gttttttaccg tgatgtacgt ttattccgtt tgtatgaagg tacaacgcaa 2880
 atccaacagg tcattattgc ccgcaatatg atccgtgaag cgactcaata attgtataac 2940
 aggtattgag tgtatctaaa aggacgggat tagtgattta agctataact tgaataactaa 3000
 tcctgacttt ttgatggcaa ggctataaaa cctcctagct cattttatct ctaagctaata 3060
 cacagctgaa agatattttc agtcttcac cttaccagac agttcacaat acaaaattgg 3120
 attttatgaa tatgcaagaa caagaaatcg aacgcgaatc aatggagttt gacgtcgtga 3180
 ttgtcggcgc aggaccggcc ggtctttctg cagcgatcaa gatccgtcaa cttgcaattg 3240
 aaaacaacct gaacgatctg tcggtttgtg tgggtggaaaa aggctctgaa gtcggtgcgc 3300
 acatcttgtc cgggtgcggtg ctggaaccac gtgccatgaa tgagctgttc ccgaactgga 3360
 aggaagaagg tgcaccttta aatgtttccag tgaccgaaga caagacctat ttctgtctct 3420
 cggatgaaaa atcacaagaa gcgccacact ggatgggtgcc taaaacctg cataacgatg 3480
 gcaactatgt tatctcgtc ggcaacgtag tgcgctgggtt gggtcaaaaa gcggaagagc 3540
 tggaagtatc tattttcccg ggctttgccg ctgctgaaat tctgtaccat gcagatgggt 3600
 cggtgaaagg cattcaaacc ggtgacatgg gcattggcaa ggatggcgaa ccgaccata 3660
 actttactcc gggctatgaa ctgcatgcca aatacacct gtttgctgaa ggctgccgtg 3720
 gccacctcgg caagcgttta attgccaaat acaacctcga taaagattca gatccacaac 3780
 attacggtat cggatatcaa gagctgtggg aaatcgacct ggcgaaacac aagccagggtc 3840
 tgggtgatgca cgggtgccggc tggccattgt ctgaaaccgg ttcttcaggc ggctgggtgtt 3900

tgtatcatgc ggaaaacaat cagggtgactt tgggcatgat cgtcgatctg tcttacacca 3960
 acccgcatat gtatccgttt atggaaatgc agcgtggaa aacccatccg ctgatcaagc 4020
 agtatctgga aggtggcaaa cgtatttctt atggcgcgcg tgcggtaacc aaaggcggct 4080
 ttaactcgct accgaaatit accttcccgg gcggatcgct gattggtgac gatgccggct 4140
 tcctgaactt tgccaaaatc aagggtcac ataccgcgat gaaatccggc atgctctgcg 4200
 gtgaagcagt gtttgaagcc attgctgccg gtgtggaaaa aggtggtgac cttgcggttg 4260
 cgcgtgtgac ggaaggcgaa gacttgtttt ccaaaaaact gacttcttac accgacaagt 4320
 tcaataatag ctggctgaaa gaagagctgt acaactcgcg taactttggc ccggccatgc 4380
 acaagtttgg tcagtggctc ggtggtgcgt ttaactttat cgaccagaac gtgtttaagg 4440
 ● tgccgtttac cctgcatgac ctggtgacgg atttcggtgc gctgaaaacc gtcgatgcgg 4500
 tgaacttcaa gccgaattat caaaaccgg atggcaaaact gacctttgac cgtctgtctt 4560
 cgggtgttgt atccaacacg gtgcatgaag aaaaccagcc agcgcattta aaactgactg 4620
 acacttcgat tccggtgaat gtcaacctgc caaaatggga tgaaccggcg cagcgctact 4680
 gccccgcggg tgtatacgaa atcatggaaa atgatgacgg ttcgaaacgc ttccagatca 4740
 atgcagccaa ctgtgtgcac tgcaagacct gtgacatcaa ggatccttca cagaacatca 4800
 cctgggtaac accggaaggt ggtggtggtc caaactatcc gaatatgtaa gtctaatac 4860
 ttcaaggaag aggtttccca ttcccttctt ttctagcaga tgaagaagct tgcaactaaa 4920
 agagattggt tggatcagtt acccaaatc gttgaaaaga ttttaactct tcgattttta 4980
 ttttttaggt aatcctagcc ctctcggggg ctaggattaa aaattttaag ttattccaac 5040
 acgaatgaca aattgttcaa tgcaaaataa aaacatacaa tatataaata tatttttta 5100
 ttaaaacata agattacaat aaaataagaa tttttatttg gagtttgttt tttttctaca 5160
 atgatcatta tgtacaattt ttaggttcac cccatccaag ccttgtgatt gcattcctgc 5220
 gattctttat tcaatgaata agcaatgcta ttaatcagca atgaataacc agcactgcag 5280
 attttgaata aattcacatg tcgtaatgga gattatcatg tcacaaaaaa tggattttga 5340
 tgctatcgtg attggtggtg gttttggcgg actttatgca gtcaaaaaat taagagacga 5400
 gctcgaactt aagggttcagg cttttgataa agccacggat gtcgcaggta cttgggtactg 5460
 gaaccgttac ccagggtgat tgtcggatac agaaaccac ctctactgct attcttggga 5520
 taaagaatta ctacaatcgc tagaaatcaa gaaaaaatat gtgcaaggcc ctgatgtacg 5580
 caagtattta cagcaagtgg ctgaaaagca tgatttaaag aagagctatc aattcaatac 5640
 cgcggttcaa tcggctcatt acaacgaagc agatgccttg tgggaagtca ccaactgaata 5700

tgggtgataag tacacggcgc gtttcctcat cactgcttta ggcttattgt ctgcgcctaa 5760
 cttgccaaac atcaaaggca ttaatcagtt taaagggtgag ctgcatcata ccagccgctg 5820
 gccagatgac gtaagttttg aaggtaaacg tgtcggcgtg attggtacgg gttccaccgg 5880
 tgttcagggtt attacggctg tggcacctct ggctaaacac ctactgtct tccagcgttc 5940
 tgcacaatac agcgttccaa ttggcaatga tccactgtct gaagaagatg ttaaaaagat 6000
 caaagacaat tatgacaaaa tttgggatgg tgtatggaat tcagcccttg cctttggcct 6060
 gaatgaaagc acagtgccag caatgagcgt atcagctgaa gaacgcaagg cagtttttga 6120
 aaaggcatgg caaacagggtg gcggtttccg tttcatgttt gaaactttcg gtgatattgc 6180
 caccaatatg gaagccaata tcgaagcgca aaatttcatt aagggtaaaa ttgctgaaat 6240
 cgtcaagat ccagccattg cacagaagct tatgccacag gatttgtatg caaacgtcc 6300
 gttgtgtgac agtggttact acaacacctt taaccgtgac aatgtccgtt tagaagatgt 6360
 gaaagccaat ccgattgttg aaattaccga aaacgggtgtg aaactcgaaa atggcgattt 6420
 cgttgaatta gacatgctga tatgtgccac aggttttgat gccgtcgatg gcaactatgt 6480
 gcgcatggac attcaaggta aaaacggctt ggccatgaaa gactactgga aagaagggtcc 6540
 gtcgagctat atgggtgtca ccgtaaataa ctatccaaac atgttcatgg tgcttggacc 6600
 gaatggcccc tttaccaacc tgccgccatc aattgaatca cagggtggaat ggatcagtga 6660
 taccattcaa tacacggttg aaaacaatgt tgaatccatt gaagcgacaa aagaagcgga 6720
 agaacaatgg actcaaactt gcgccaatat tgcggaaatg accttattcc ctaaagcgca 6780
 atcctggatt tttggtgcga atatcccggg caagaaaaac acggtttact tctatctcgg 6840
 tggtttaaaa gaatatcgca gtgcgctagc caactgcaaa aaccatgcct atgaaggttt 6900
 tgatattcaa ttacaacgtt cagatatcaa gcaacctgcc aatgcctaaa tatatggggg 6960
 gcatcccca tattccattt tgtttaacat cagtcatatg ccagggatgt cttatcatga 7020
 actatccaaa tataccttta tatatcaacg gtgagtttct agatcatacc aatagagacg 7080
 tcaaagaagt ttttaatcca gtgaaccatg aatgtattgg actcatggcc tgtgcatcac 7140
 aagcagacct ggactacgca cttgaaagtt cacaacaggc ttttctaagg tggaaaaaaa 7200
 cttctcctat caccgtagt gaaatcctca gaacctttgc gaaactagcg cgtgaaaaag 7260
 cagcagaaat cgggcgcaat attacccttg atcaaggtaa gcccctgaaa gaagccattg 7320
 cagaagtcac tgtctgtgca gaacatgcag aatggcatgc agaagaatgc cgacgcattt 7380
 atggccgtgt tattccaccg cgtaacccaa atgtacagca actagtagtc agagaaccgc 7440
 tgggcgtatg tctggcattt tcaccgtgga atttcccgtt taatcaggca attcgtaaaa 7500

tttctgctgc aattgctgcc ggctgcacca tcattgtgaa aggttctggc gacacaccaa 7560
 gcgcggtata tgcgattgcc cagctatttc atgaggcggg tttgccgaat ggtgtgctga 7620
 atgtgatttg ggggtgactca aacttcattt ctgattacat gatcaaactg cccgatcatcc 7680
 aaaagatttc attcacaggc tcaaccccg tgggtaaaaa attagcctcg caagcgagtc 7740
 tgtatatgaa gccttgacc atggaattgg gtggtcatgc accggtcacg gtctgtgatg 7800
 atgctgatat tgatgccgct gttgaacatc tggtcgggta taaattccgt aatgcaggac 7860
 aggtctgtgt atcaccaacc cgtttttatg tgcaggaagg tatttataag gaattttctg 7920
 agaaagtggg gtttaagagcc aaacagatca aagtggggtg tggcttagac gcacccctcag 7980
 atatgggacc attggctcaa gctcgccgca tgcattgcaat gcaacaaatt gttgaagatg 8040
 cggttcataa aggcataaaa ttactgcttg gcggaaataa aatttctgac aaaggcaatt 8100
 tttttgaacc aacgggtactc ggtgacttgt gcaatgacac ccagtttatg aatgacgagc 8160
 catttggtcc gatcattggg ttgatacctt ttgacacaat agaccatgtc ctggaagaag 8220
 caaatcgatt accatttgga ttagcctctt acgcttttac cacatccagc aaaaatgcgc 8280
 atcaaacttc atacggactg gaggctggca tggtttcgat taaccacatg ggattggcgc 8340
 tcgctgaaac accttttggg ggtattaagg atagcgggtt tggtagtgaa ggggggtatcg 8400
 aaacctttga cggttacctc agaaccaaat ttattacgca actcaattag aaatggatct 8460
 tgggtgtcgt aggcacacca attctctttt gactttaagg atgaaagtta aatgagcaca 8520
 gacaaagcaa atacgtgat caaacccgaa gatgtcgtgt tatggattcc gggtaatgtc 8580
 acaattgaca gcatgaatgc cggttgggaa aacattgcaa tcagagggta cgaatatacc 8640
 aacctcgatg tgcattattc tgccatgcgt gactacatga tcgtcaacta taaaaaaagt 8700
 gcggcggaaa tgcgtagaaa aggcgatgcc tcttgggata cccaagtggg taagccgggt 8760
 tatgtctcct tgttgacctg tgggtgaagat tcccgcgtgg cgtggaatga ccatattgcc 8820
 gtcacccatg tctacatttc gcatgactcc atcacctcaa tggcgaataa ggtgtttgat 8880
 tatgatatcg cttcgatccg aatcagagac gaagtcggtg tggaagatca tgttttacct 8940
 gctctgactt cacttttaga actagaatta aagcaagggt gtttaggtgg aaacctgtat 9000
 ttagagagca ttaaaaacca gatcgccctg catttactcc gtcagtatgc caaattagat 9060
 ttaaggaag gacagtgccg ttctgggttt actcccctac aacgcagact gttattagaa 9120
 tttatcaatg aaaacatgag cattaaaatt accctcgaag atttagcggg attagtcaag 9180
 atgagcgtgc ctcatttaat gagaaaaatt aaagtcgatt ttggtaattc ccctgctgcc 9240
 tacatcatga atctcagggt gcaatttgct aaacgtttgc tcacttcaaa aaaagaaatt 9300

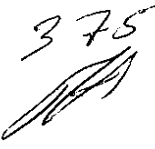
372

ccactgaaag tgattgccag tgaagccggt ttttgcgatc agagccatat gacccgagta 9360
 tttcaaaaat tttttgggaa aacacccatc gaaatcagac aggaacacac caatctcgtg 9420
 tctgaaaatt cagtctcctc tattgttttt tgagtactaa gagccacgca agaacctgat 9480
 tttcaataaa gcatccactg aaaaccagtg tggacttaca tgcattatct atgcaaaaata 9540
 acaaatgtca tgtgagtatc aagatatact ttctatcgct atcaagaact tgccagtaca 9600
 ggcaatatgg atgcactcat caaccagagt cgcagaactc caaatTTaaa aaaccgagtg 9660
 gatgagcaaa ctgaataagc tgttggtgat tttgcaatcc aatatccagc ttatgggtcag 9720
 catcggacca gtaatgagct acgtcagatt ggcatcttcg tatctggcag cgggtgtgcgc 9780
 tctatctggc ttagacacaa tcttgagaat ttcaaaaagc gattaaaggc acttgaaatt 9840
 aaagttgctc aagaaggcat tcagttgaat gatcagcaga ttgccgcatt agaacgtaaa 9900
 catgaagatg atgttgcttg tggtgaaatt gaaacacatc atccagggtta ccttgaggca 9960
 caagatactt tttatgtcgg aaatctaaaa ggtgttgggc atatttatca gcaaactttt 10020
 attgatactt atagcaaagt ggttcactgc aagctgtaca caaccaagac accaatcaca 10080
 gccgcagatt tattgaatga ccgcgtgtta ccattctatg agtcacaagg attgccaatg 10140
 cttcgcattt tgaccgacag aggcaccgaa tattgcggtg aagttgaaca tcacgattat 10200
 gagctttatt tggctctgaa tgatattgat cacactaaaa ctaaagcagc atcaccacaa 10260
 acaaatggga tctgtgagcg cttccataag acgatcttgc aggagtttta tcagattact 10320
 tttcgaaaga aactctatag ctcatagaa gagttacagc ttgatctaga cggttggctg 10380
 aaattctata atactgaacg aaccatcag ggtaagggtg gtaatggcag atgagcagca 10440
 ttgctgcgca agattgcaac attacttgat ggaaaacgta tttgggctga aaagaattta 10500
 gttcaaatTT aacctgacag tcttaagcaa atatcggtta ctatcagatc aggtttgaga 10560
 taccgtctga aacgtcaagt aaatgattga gaattcatgc tcaataatct gcttgataag 10620
 gctgttggtg tttgagcaca ccataacaaa gatgaatcaa cttcctcatc gcggctccaa 10680
 tcgctatcat cttggtttta ccattcgcca ataaacgttc attcattgcc ctgatgtgag 10740
 ggttatgccg agttgcgaca atggctgcca tatataaacc agcacgtatt ttggaagagc 10800
 ccgctttgga taaacggctt ctgccatgaa tggaactacc cgattgcttt tgaatgggga 10860
 ccaaaccgac aaaggcagcc gcttgactag ccctttcaaa agtatggctg cgcaagaaac 10920
 tgagcattaa taaactgggt cgatctgcaa tggctggaat actgctgagc agttctttat 10980
 catTTTTtaa atcaggattc tgattaatgt gatcatcaat ttgctggctg ataccctgaa 11040
 tgtgtttgtt taactgttca atactcttgt ggatagactg aagtacaggt tccatcgtga 11100

aggtcgactc tgctttttcc aaacgattct tttcacgttg taaatcttca caaagaatag 11160
 ctcttctatc cagcaaagca ttcagcaatt gaatatgttt aggtaaaggt tgccaaaaat 11220
 gtagatcggc agtcatcgca aatcgagcta ggacctcact atccaccttg tctgttttat 11280
 tcagcttaga catactctga gcaaaatata gagctctggc aggattgggt acacagactt 11340
 gatagcccg c atcaataaaa tatttaacca agagttcatg ataaatagat gttgcttcca 11400
 ttaaaataat ggtctgcgta gaagttgcag catgctgctt tagccagggt tgaagttgct 11460
 caaaaccttt tgggtgtattt gaaaaagttt tgggttttctt tttatttgca gaattttcta 11520
 aaattaaaca gcaatcaatt ttagcttttag caacatcaat accaagataa aacataatct 11580
 ttacctgctt tatttatcca attattgttt tagcataacc accgtctttt cttgtgaatg 11640
 cagcatcaaa gtgcttggtta ccgtccagag ttgtgcaagt ggtagggca aattacagggt 11700
 tttatctcaa actctaactt tatgttttgc tagtacacga aactctgcaa tttgcaatat 11760
 agtgatagct aatcactatg aatggtaaga tacaagctag tacacataag aagatattac 11820
 ttcttctcag gcagattcgc agcaaagaaa aattttccct tacaacaata gataaaagaa 11880
 aagaggggtat caccctctt tctcttttat atgggggtat cttctactca ttttttattt 11940
 cgagggtatat gcaccatcga ccacatactg actaccggtc acaaatgaag catcatcaga 12000
 aagtaggaag gcaacaacct gagcaacttc ttccggctgt cccaaacgac caatcggatg 12060
 tagttttacc atttctgctt cttcaaattc tgcaatcaaa ggcgttttga tatagccagg 12120
 atgtactgaa ttaatgcgaa tccctttatc tgcatattcc aatgccgccg ctttcgtag 12180
 acccgttacg ccatgttttg cagcgacata gcctgaaata ttttgaatcc cgatcaagcc 12240
 tgcaatagaa gcggtattga caatcgctcc gccccctgcg gccaagattg caggaacttc 12300
 ataatgcatg ctgtagaaaa ccgcattcaa gttcacatca atcacacgac gccatccttc 12360
 aatgctcaat tcttcggtgg agttaacttc acccagaatt cccgcattat tgaaggccaa 12420
 atgcagtgc ccaaaagtgc tgaccgcaa ctcgactgca gctttcatgt cttcaggctc 12480
 agcagtattg gccttattcg cagccgcttt cccgcctaaa gcgacaattt cgtccacaac 12540
 tttctgtgct gcttccagggt taatatctga aaccactaca cttacaccct gttgagccaa 12600
 aagcagtgcg gtgctttttac caatacctga accagcgcca gtaattaaag cgactttatt 12660
 gttgaattta tttgacataa ttttttccat ttcaaatttt aagcatcaaa gcttgtttca 12720
 tattttaaga ttcaagaaac cagatccggt agatgactcg tctgccaagc gacaacctgt 12780
 ctgatatcag gcttgcgatt caccctgtag acggttttca ttctaaatt ctgtatttcc 12840
 aagttatata aacaaaagtg ctaatctatg gggaattccc aggatccaaa caaatagaat 12900

374
A

gccatgaaag catcttttgc caagcgctgt gctgtatggt tcctagacaa accaccaacg 12960
ataactgcaa ctttttgaac tccttacaat ttccttattt tctttccct tcatcgcata 13020
aaaatagttt ttgcattcac aacaaaatca gcatgaatag tttttaaact cactgtacat 13080
atcttctata ttgatgacca agctggatat tgaattgcaa aattctatac agcctgttca 13140
acatgatcga tttagaaggc atacagtaaa cgtgactgaa gtccagaaat ttccaagcca 13200
ttttcaacat tcacatcttg tcgccattgt aataatagct gcagattcgg cttgatattg 13260
gtagaagcag aaacgacaaa ggtatctttt ctatcactgc cacgttcagt gacaccattc 13320
accttttctt taccgccatc ggtatgtctc cagggtgacag ccaaattgga tttatcggtc 13380
actttataga gtgcggagaa atctgtctgg aaaaaaacct ctttctcaat gttggtatat 13440
ttttgctcgc tataaagttc aaactgcccc accccctcaa gcgcaaattt atcagttaaa 13500
gcatggtaat aaccggcctg aacattatat tgatagcgat cattactgat ggcaaaaccc 13560
ttcgtttcat tactgccggt aggtacggtc aaaaaaccac cgaaaccaa atagcgccct 13620
ttttcagcat catgcaatgg ccaggcgata ccaccacaa ttaaatacc gacacccgag 13680
atatcatcag cgccattcat cttttgcttg gcaaaaggca agaggaattg aggatctaca 13740
atccaatccc ctacttcaat aaaacgaacg taacgcaata ttcccaaatac aatgcttaaa 13800
tcgagatcat cagcgacttt atcaccattt gcatacgctt tatccgcttc cgtatgctgg 13860
taataggcaa ccgctaagtt ggttccccct ggaagtgcct gataatcccc ggcatcagaa 13920
ctcaccctg cggttgagc gtccaaagcg gcagttaaag caaagaccaa agcagctatt 13980
ttttgatttg aacgatgata gaaatagttt ttcatttggt tcatttttaa ctctccggtg 14040
ttttgactca tttttttaa atgagtcttc ctagcacaaa gaccactcag gtctttgcgc 14100
aatttcttga ttttgatttg ggtattaaat atggaaaaac gttgggtgat cagttttcgt 14160
gcataagcac aatacgccc atgacgttgc catctttcaa gtctccaaat gcggaattga 14220
tctgcgaaat tggcagtttt ttcacgggaa tggctgacat gtgggtttct ttcaccagct 14280
ccaccagctc tcttaattcc tctaccgtcc ctacataact gccctggatt gtgagtggtc 14340
tcattggaat caccggaatg gaaagcttaa ttttccccc catcaatccg cagatcacia 14400
tatgcccacc acgtgcagca ctgcgaagg caaggctcaa tgttgattga ctgccaacca 14460
gatcaaggat cagacgtgca ccaccgtcag ttgcctgaat cagctgttga gcagatcct 14520
cacttcgggt attgatgacc gataatgcac cggcagcacg tgctgcttcc agtttgctgt 14580
catcaatata aactacgatt gcgcctttgg cttgcatagc tttgagcaac tcgagtcca 14640
tcagccctaa accaccggca ccaatgatca ccaccggctc gctttgaatc aaatcaccga 14700

375


atTTTTcag tgcactgtat gttgtcacgc ctgcacatgc caaagggtgca gcttcagcca 14760
gatccagacc tgcaatatcc accagatatc gtggatgcgg cacgatgata tattcggcaa 14820
aaccacccgg cttggcgatg cctaactgtt gcgggtttggc acacagggtt tcttcgccac 14880
gtttacagta gttgcattca ccgcaaccaa tccatggatg aaccaagctg accatgccga 14940
ccttgactga ttccgcatct ggaccgacag caaccacctg acctgtaatt tcatgactta 15000
aggttaaggg tggcttcagc ccacgatctg caagggataa acgcttgccc ccacctagat 15060
cataataacc ttcccataag tgtaaatccg tatggcatag acctgcggtt ttacatgga 15120
gtaaaacttc agtaccttcc ggttgcgga tttctttctc aacgtcttcg agtggttgtc 15180
catgatgcgt cacgcagtaa cagtgcata atctctcctt tgaaacaata aaatagacgg 15240
ccttgtagtg aacaaagtct ttatttact aagttttata cgccgtgtgg gcatgattt 15300
atgctttaaa ccactgcgca attttcgcta attcttgatc agcttcactt gcacgcccag 15360
ctaggaaagg aaaaacgtgc tgcattgtgt ccaccacaga taaagtcaca tcaacaccct 15420
ctttttttgc aatatcagca agacgtgttg cattgtctac aagtgattca actgatccgg 15480
cattgatata caaacgtggg aaaacctgat aattggcttt taacggattc gccaatggat 15540
ttgccggatc accatgttca cccaagaaca ttgttgacat gcctttaagc agatccactg 15600
taatcaaggc atcagtggca tcgttgctga tcagggtttc acctttgtgc tccatatcca 15660
gccaaggaga gaatgcaatc actgctctg gcaactcaat cctttcattt cgtagattga 15720
gtacggttga tategccaga ttcccccccg cagaatcccc tgcggtcagc atattttttg 15780
cagtaaagcc acgctggagt agttctttat atactgctgt cacgtcctga atttggtgccg 15840
ggaagacatg ttctggtgaa cgtcggtaat caaccacaaa tgcggatacc cctaaatact 15900
tggccaaatg ccccaccagc ttacggtgac tggccgaaga accgaccgca aatccaccgc 15960
catgggtata aatgatgact ttggataagt cagcatcttt cggataaatc caaagacctt 16020
ctacacctgc cacaacatcg aatttataag acacttcttc cggttccaat gtaggttgat 16080
gccattcatc aaacatactg cgaaagtctt caatggatcat attcggattt tcctgcatcc 16140
gtcttgacca gttcgcataa aaatcgaaaa gaaattgagt attgctttgt gtgctattca 16200
ttttaaaatc cttgatttga tatttaagga ataaatccta gttttattcc atgaagatat 16260
aaaaacttga gtgccatcac tcatggctag acactcagaa gatccaaatc taaagagtgg 16320
ctttgcatca ctggtttgat acaatttttt gcatgactaa gtaatctacg gataatctaa 16380
ccgtttcaaa ttagtatttt aaaatgtaaa aaatacatac cagcgaatgt tttctgcaaa 16440
atcgcatcct gttcaatata gcttttgatc ctacttattc tcttttctat tccagtccgt 16500

tataaaaaag ctttcattca ttttcatgca atcatgagct atgaatgttc ttaaacatta 16560
aacgattgtg tgtatggctg acttgtacat tcttgtactt atttttgtat aaaatgatca 16620
ggctcatcaa tttatgggaa aaattacaat tcgggtacaa tatctttcct gtttcatgaa 16680
tctattcaac tcattaaact tacgaccctc aactgcccaa aatcatagga tctgccgatc 16740
cacttgcaga attagcaatg ctaaaacatg aactccaaag agttactaaa aaaagagcat 16800
attaaaaaaa agccgtggca tatttcgcaa gccagttcaa gtcaggtagt tctttattca 16860
gtacctcagt taaacttttag attttcataa cgatgggttat tctgcatggc taaatacgct 16920
aatcagcaaa aaactctcca aaagataggc acagaaacac atatcaacca taaaaaccat 16980
ctcagacagt atattttacaa gcctctaatt caccgcactc acacttctct gcaagccttt 17040
ttaaataccc tgtacaaagt tctcagcctg atgaagcttc accttggact tagctttcag 17100
ttcagcctgt acttgggtcag tttctgaatt ttcatttgca taaaactcct ccaccacatc 17160
cataccctcc tcaatgtcag tttcaaaatg tgcatgtca tagccttgcc gtgccatttg 17220
aatggcttat tgaagattaa tggcatcacg taaagttaaa tccacgtaat acacagggtg 17280
tcgatagctt tgcgtcgtag actttctcga agagtcaatt gcagcggtag gcatgacagc 17340
aagccattca atgccgcatg gtaataactc agccgtgcgg ccaacgttcg tatgctgtta 17400
aaacccgggtt attctaa 17417

<210> 74
<211> 1164
<212> DNA
<213> *Escherichia coli*

<400> 74
atgaacaact ttaatctgca caccccaacc cgcattctgt ttggtaaagg cgcaatcgct 60
ggtttacgcg aacaaattcc tcacgatgct cgcgtattga ttacctacgg cggcggcagc 120
gtgaaaaaaaa cggcggttct cgatcaagtt ctggatgccc tgaaaggcat ggacgtgctg 180
gaatttggcg gtattgagcc aaacccgggt tatgaaacgc tgatgaacgc cgtgaaactg 240
gttcgcgaac agaaagtac tttcctgctg gcgggttggcg gcggttctgt actggacggc 300
accaaattta tcgccgcagc ggctaactat ccggaataa tcgatccgtg gcacattctg 360
caaacgggcy gtaaagagat taaaagcgcc atcccgatgg gctgtgtgct gacgctgcc 420
gcaaccgggt cagaatccaa cgcaggcgcg gtgatctccc gtaaaaccac aggcgacaag 480
caggcggtcc attctgcccc tgttcagccg gtatttgccg tgctcgatcc gggttatacc 540
tacaccctgc cgccgctca ggtggctaac ggcgtagtgg acgcctttgt acacaccgtg 600
gaacagtatg ttaccaaacc ggttgatgcc aaaattcagg accgtttcgc agaaggcatt 660

377
//

```

ttgctgacgc taatcgaaga tgggtccgaaa gccctgaaag agccagaaaa ctacgatgtg      720
cgcgccaacg tcatgtgggc ggcgactcag gcgctgaacg gtttgattgg cgctggcgta      780
ccgcaggact gggcaacgca tatgctgggc cacgaactga ctgcgatgca cggctctggat      840
cacgcgcaaa cactggctat cgtcctgcct gcactgtgga atgaaaaacg cgataccaag      900
cgcgctaagc tgctgcaata tgctgaacgc gtctggaaca tcaactgaagg ttccgatgat      960
gagcgtattg acgccgcgat tgccgcaacc cgcaatttct ttgagcaatt aggcgtgccg     1020
accacctct cgcactacgg tctggacggc agtccatcc cggttttgct gaaaaaactg     1080
gaagagcacg gcatgacca actgggcgaa aatcatgaca ttacgttgga tgtcagccgc     1140
cgatatatacg aagccgcccg ctaa                                           1164

```

<210> 75
 <211> 387
 <212> PRT
 <213> *Escherichia coli*

<400> 75

Met Asn Asn Phe Asn Leu His Thr Pro Thr Arg Ile Leu Phe Gly Lys
1 5 10 15

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

Leu Ile Thr Tyr Gly Gly Gly Ser Val Lys Lys Thr Gly Val Leu Asp
35 40 45

Gln Val Leu Asp Ala Leu Lys Gly Met Asp Val Leu Glu Phe Gly Gly
50 55 60

Ile Glu Pro Asn Pro Ala Tyr Glu Thr Leu Met Asn Ala Val Lys Leu
65 70 75 80

Val Arg Glu Gln Lys Val Thr Phe Leu Leu Ala Val Gly Gly Gly Ser
85 90 95

Val Leu Asp Gly Thr Lys Phe Ile Ala Ala Ala Ala Asn Tyr Pro Glu
100 105 110

Asn Ile Asp Pro Trp His Ile Leu Gln Thr Gly Gly Lys Glu Ile Lys
115 120 125

Ser Ala Ile Pro Met Gly Cys Val Leu Thr Leu Pro Ala Thr Gly Ser
130 135 140

378
A

Glu Ser Asn Ala Gly Ala Val Ile Ser Arg Lys Thr Thr Gly Asp Lys
145 150 155 160

Gln Ala Phe His Ser Ala His Val Gln Pro Val Phe Ala Val Leu Asp
165 170 175

Pro Val Tyr Thr Tyr Thr Leu Pro Pro Arg Gln Val Ala Asn Gly Val
180 185 190

Val Asp Ala Phe Val His Thr Val Glu Gln Tyr Val Thr Lys Pro Val
195 200 205

Asp Ala Lys Ile Gln Asp Arg Phe Ala Glu Gly Ile Leu Leu Thr Leu
210 215 220

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

Arg Ala Asn Val Met Trp Ala Ala Thr Gln Ala Leu Asn Gly Leu Ile
245 250 255

Gly Ala Gly Val Pro Gln Asp Trp Ala Thr His Met Leu Gly His Glu
260 265 270

Leu Thr Ala Met His Gly Leu Asp His Ala Gln Thr Leu Ala Ile Val
275 280 285

Leu Pro Ala Leu Trp Asn Glu Lys Arg Asp Thr Lys Arg Ala Lys Leu
290 295 300

Leu Gln Tyr Ala Glu Arg Val Trp Asn Ile Thr Glu Gly Ser Asp Asp
305 310 315 320

Glu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Glu Gln
325 330 335

Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser
340 345 350

Ile Pro Ala Leu Leu Lys Lys Leu Glu Glu His Gly Met Thr Gln Leu
355 360 365

Gly Glu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Glu
370 375 380

379
AAla Ala Arg
385

<210> 76
 <211> 1623
 <212> DNA
 <213> *Bacillus subtilis*

<400> 76
 atgtatttgg cattccaggt gcaaaaattg atgcggtatt tgacgcttta caagataaag 60
 gacctgaaat tatcgttgcc cggcacgaac aaaacgcagc aattcatggc ccaagcagtc 120
 ggccgtttta ctggaaaacc gggagtcgtg ttagtcacat caggaccggg tgcctctaac 180
 ttggcaacag gcctgctgac agcgaacact gaaggagacc ctgtcgttgc gcttgctgga 240
 aacgtgatcc gtgcatatcg tttaaaacgg acacatcaat ctttggataa tgcggcgcta 300
 ttccagccga ttacaaaata cagtgtagaa gttcaagatg taaaaaatat accggaagct 360
 gttacaaatg catttaggat agcgtcagca gggcaggctg gggccgcttt tgtgagcttt 420
 ccgcaagatg ttgtgaatga agtcacaaat acgaaaaacg tgcgtgctgt tgcagcgcca 480
 aaactcggtc ctgcagcaga tgatgcaatc agtgcggcca tagcaaaaat ccaaacagca 540
 aaacttcctg tcgttttggg cggcatgaaa ggcggaagac cggaagcaat taaagcgggt 600
 cgcaagcttt tgaaaaaggt tcagcttcca tttgttgaaa catatcaagc tgccggtacc 660
 ctttctagag atttagagga tcaatatatt ggccgtatcg gtttgttccg caaccagcct 720
 ggcgatttac tgctagagca ggcagatggt gttctgacga tcggctatga cccgattgaa 780
 tatgatccga aattctggaa tatcaatgga gaccggacaa ttatccattt agacgagatt 840
 atcgttgaca ttgatcatgc ttaccagcct gatcttgaat tgatcgggtga cattccgtcc 900
 acgatcaatc atatcgaaca cgatgctgtg aaagtggaaat ttgcagagcg tgagcagaaa 960
 atcctttctg atttaaaaca atatatgcat gaagggtgagc aggtgcctgc agattggaaa 1020
 tcagacagag cgcacctctt tgaaatcggt aaagagttgc gtaatgcagt cgatgatcat 1080
 gttacagtaa cttgcgatat cggttcgcac tccatttgga tgtcacgtta tttccgcagc 1140
 tacgagccgt taacattaat gatcagtaac ggtatgcaaa cactcggcgt tgcgcttcct 1200
 tgggcaatcg gcgcttcatt ggtgaaaccg ggagaaaaaag tggtttctgt ctctgggtgac 1260
 ggcggtttct tattctcagc aatggaatta gagacagcag ttcgactaaa agcaccaatt 1320
 gtacacattg tatggaacga cagcacatat gacatggtgc atttccagca attgaaaaaa 1380
 tataaccgta catctgcggt cgatttcgga aatatcgata tcgtgaaata tgcggaaagc 1440
 ttcgagacaa ctgcgttgcg cgtagaatca ccagaccagc tggcagatgt tctgcgtcaa 1500

380
AA

ggcatgaacg ctgaagggtcc tgtcatcatc gatgtcccggttgactacag tgataacatt 1560
 aatttagcaa gtgacaagct tccgaaagaa ttcggggaac tcatgaaaac gaaagctctc 1620
 tag 1623

<210> 77
 <211> 540
 <212> PRT
 <213> *Bacillus subtilis*
 <400> 77

Met Tyr Leu Ala Phe Gln Val Gln Lys Leu Met Arg Tyr Leu Thr Leu
 1 5 10 15

Tyr Lys Ile Lys Asp Leu Lys Leu Ser Leu Pro Gly Thr Asn Lys Thr
 20 25 30

Gln Gln Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly
 35 40 45

Val Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly
 50 55 60

Leu Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly
 65 70 75 80

Asn Val Ile Arg Ala Tyr Arg Leu Lys Arg Thr His Gln Ser Leu Asp
 85 90 95

Asn Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln
 100 105 110

Asp Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala
 115 120 125

Ser Ala Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val
 130 135 140

Val Asn Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro
 145 150 155 160

Lys Leu Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys
 165 170 175

Ile Gln Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly
 180 185 190

381
A

Arg Pro Glu Ala Ile Lys Ala Val Arg Lys Leu Leu Lys Lys Val Gln
195 200 205

Leu Pro Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp
210 215 220

Leu Glu Asp Gln Tyr Phe Gly Arg Ile Gly Leu Phe Arg Asn Gln Pro
225 230 235 240

Gly Asp Leu Leu Leu Glu Gln Ala Asp Val Val Leu Thr Ile Gly Tyr
245 250 255

Asp Pro Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg
260 265 270

Thr Ile Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr
275 280 285

Gln Pro Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His
290 295 300

Ile Glu His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys
305 310 315 320

Ile Leu Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro
325 330 335

Ala Asp Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu
340 345 350

Leu Arg Asn Ala Val Asp Asp His Val Thr Val Thr Cys Asp Ile Gly
355 360 365

Ser His Ser Ile Trp Met Ser Arg Tyr Phe Arg Ser Tyr Glu Pro Leu
370 375 380

Thr Leu Met Ile Ser Asn Gly Met Gln Thr Leu Gly Val Ala Leu Pro
385 390 395 400

Trp Ala Ile Gly Ala Ser Leu Val Lys Pro Gly Glu Lys Val Val Ser
405 410 415

Val Ser Gly Asp Gly Gly Phe Leu Phe Ser Ala Met Glu Leu Glu Thr
420 425 430

382
A

Ala Val Arg Leu Lys Ala Pro Ile Val His Ile Val Trp Asn Asp Ser
435 440 445

Thr Tyr Asp Met Val His Phe Gln Gln Leu Lys Lys Tyr Asn Arg Thr
450 455 460

Ser Ala Val Asp Phe Gly Asn Ile Asp Ile Val Lys Tyr Ala Glu Ser
465 470 475 480

Phe Gly Ala Thr Ala Leu Arg Val Glu Ser Pro Asp Gln Leu Ala Asp
485 490 495

Val Leu Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val
500 505 510

Pro Val Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro
515 520 525

Lys Glu Phe Gly Glu Leu Met Lys Thr Lys Ala Leu
530 535 540

<210> 78
<211> 1680
<212> DNA
<213> *Klebsiella terrigena*

<400> 78
atggacaaac cgcgtcacga acgtcaatgg gcccacgggtg ccgacttaat cgctcagccag 60
cttgaggccc agggcgtagc ccaggctctt gccatccccg gtgccaaaat cgacaagggtg 120
tttgattccc tctctgactc ctcaatccgc attattccgg tgcgccacga ggctaacgcc 180
gcctttatgg ccgcggcggt cgggcggatt accggtaaag cgggcgctgc gctggtgacc 240
tccggtcccc gctgctcaaa cctgattacc ggcatggcca ccgccaatag cgaaggcgac 300
ccggtggtgg cgctgggagg cgcggtgaag cgcgcggata aggccaagct gggtcaccaa 360
agcatggaca ccgtggcgat gttcagcccc gtcaccaaata acgccgtcga ggtgaccgcc 420
tccgacgcgc tggccgaggt ggtctccaac gcctttcgcg ccgccgaaca ggggcgtccg 480
gggagcgcgt ttgtcagcct gccgcaggat atcgttgacg gccccgccag cggcagcacg 540
ctgcccgcga gcagagcgcc gcagatgggc gccgcgccgg atggcgccgt tgacagcggt 600
gcgcaggcga tcgccgcggc gaagaaccct atcttctctg tcgggctgat ggccagccag 660
ccggaaaaca gccgcgccct gcaccgccat gctggaaaaa agccatattc cggtcaccag 720
cacctatcag gcgcgggggc ggtaaatacag gataacttcg cccgcttcgc cggccgggta 780

ggcctgttta ataaccaggc gggcgatcgc ctgctgcgtc aggcggacct gatcatctgc 840
 atcggctata gcccgggtga gtacgaaccg gcgatgtgga acagcggcac ggcaaccctg 900
 gtgcatatcg acgtgctgcc ggcctatgaa gagcgggaact acgtcccgga tatcgagctg 960
 gtgggcgaca tcgccgccac cctcgagaag ctggcccagc gcattgaaca tcggctggtg 1020
 ttaactccgc aggcggcgga catcctcgc gaccgccagc gccagcggga gctgcttgac 1080
 cgccgcgggg cgcagctgaa tcagtttgcg ctccaccgc tgcgcacgt gcgggcatg 1140
 caggatatcg tcaatagcga cgtcaccttg accgtcgata tgggcagttt ccatatctgg 1200
 attgcccgt acctctacag cttccgcgcc cgccagggtga tgatctcaa cggtcagcaa 1260
 acgatgggcg tcgcgctgcc gtgggcaatc ggcgcgtggc tggccaatcc gcagcgcaag 1320
 gtggtctcgg tatccggcga tggcggcttc ctgcagtcga gcatggagct ggagaccgcc 1380
 gtgcgcctgc acgccaatat tctgcacatc atctgggtcg ataacggcta caacatggtg 1440
 gcgattcagg aacagaagaa atatcagcgc ctctccggcg tggagttcgg cccggtcgat 1500
 ttcaaagtct acgccgaagc gttcggggcc tgcgggtttg cggtagagag cgccgaggcc 1560
 ctggagccga ccctgcgcgc ggcgatggat gtcgacggcc cggcgggtgg cgccattccg 1620
 gtcgattacc gcgataaccc tctgctgatg ggccagctcc atctcagcca aatactgtga 1680

<210> 79
 <211> 559
 <212> PRT
 <213> *Klebsiella terrigena*
 <400> 79

Met Asp Lys Pro Arg His Glu Arg Gln Trp Ala His Gly Ala Asp Leu
1 5 10 15

Ile Val Ser Gln Leu Glu Ala Gln Gly Val Arg Gln Val Phe Gly Ile
20 25 30

Pro Gly Ala Lys Ile Asp Lys Val Phe Asp Ser Leu Leu Asp Ser Ser
35 40 45

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

Ala Ala Val Gly Arg Ile Thr Gly Lys Ala Gly Val Ala Leu Val Thr
65 70 75 80

Ser Gly Pro Gly Cys Ser Asn Leu Ile Thr Gly Met Ala Thr Ala Asn
85 90 95

Ser Glu Gly Asp Pro Val Val Ala Leu Gly Gly Ala Val Lys Arg Ala
100 105 110

Asp Lys Ala Lys Leu Val His Gln Ser Met Asp Thr Val Ala Met Phe
115 120 125

Ser Pro Val Thr Lys Tyr Ala Val Glu Val Thr Ala Ser Asp Ala Leu
130 135 140

Ala Glu Val Val Ser Asn Ala Phe Arg Ala Ala Glu Gln Gly Arg Pro
145 150 155 160

Gly Ser Ala Phe Val Ser Leu Pro Gln Asp Ile Val Asp Gly Pro Ala
165 170 175

Ser Gly Ser Thr Leu Pro Ala Ser Arg Ala Pro Gln Met Gly Ala Ala
180 185 190

Pro Asp Gly Ala Val Asp Ser Val Ala Gln Ala Ile Ala Ala Ala Lys
195 200 205

Asn Pro Ile Phe Leu Leu Gly Leu Met Ala Ser Gln Pro Glu Asn Ser
210 215 220

Arg Ala Leu His Arg His Ala Gly Lys Lys Pro Tyr Ser Gly His Gln
225 230 235 240

His Leu Ser Gly Ala Gly Ala Val Asn Gln Asp Asn Phe Ala Arg Phe
245 250 255

Ala Gly Arg Val Gly Leu Phe Asn Asn Gln Ala Gly Asp Arg Leu Leu
260 265 270

Arg Gln Ala Asp Leu Ile Ile Cys Ile Gly Tyr Ser Pro Val Glu Tyr
275 280 285

Glu Pro Ala Met Trp Asn Ser Gly Thr Ala Thr Leu Val His Ile Asp
290 295 300

Val Leu Pro Ala Tyr Glu Glu Arg Asn Tyr Val Pro Asp Ile Glu Leu
305 310 315 320

Val Gly Asp Ile Ala Ala Thr Leu Glu Lys Leu Ala Gln Arg Ile Glu
325 330 335

His Arg Leu Val Leu Thr Pro Gln Ala Ala Asp Ile Leu Ala Asp Arg
340 345 350

Gln Arg Gln Arg Glu Leu Leu Asp Arg Arg Gly Ala Gln Leu Asn Gln
355 360 365

Phe Ala Leu His Pro Leu Arg Ile Val Arg Ala Met Gln Asp Ile Val
370 375 380

Asn Ser Asp Val Thr Leu Thr Val Asp Met Gly Ser Phe His Ile Trp
385 390 395 400

Ile Ala Arg Tyr Leu Tyr Ser Phe Arg Ala Arg Gln Val Met Ile Ser
405 410 415

Asn Gly Gln Gln Thr Met Gly Val Ala Leu Pro Trp Ala Ile Gly Ala
420 425 430

Trp Leu Val Asn Pro Gln Arg Lys Val Val Ser Val Ser Gly Asp Gly
435 440 445

Gly Phe Leu Gln Ser Ser Met Glu Leu Glu Thr Ala Val Arg Leu His
450 455 460

Ala Asn Ile Leu His Ile Ile Trp Val Asp Asn Gly Tyr Asn Met Val
465 470 475 480

Ala Ile Gln Glu Gln Lys Lys Tyr Gln Arg Leu Ser Gly Val Glu Phe
485 490 495

Gly Pro Val Asp Phe Lys Val Tyr Ala Glu Ala Phe Gly Ala Cys Gly
500 505 510

Phe Ala Val Glu Ser Ala Glu Ala Leu Glu Pro Thr Leu Arg Ala Ala
515 520 525

Met Asp Val Asp Gly Pro Ala Val Val Ala Ile Pro Val Asp Tyr Arg
530 535 540

Asp Asn Pro Leu Leu Met Gly Gln Leu His Leu Ser Gln Ile Leu
545 550 555

<210> 80
<211> 768
<212> DNA
<213> *Bacillus subtilis*

<400> 80

```

atgaaacgag aaagcaacat tcaagtgtc agccgtggc aaaaagatca gcctgtgagc      60
cagatttatc aagtatcaac aatgacttct ctattagacg gagtatatga cggagatttt      120
gaactgtcag agattccgaa atatggagac ttcggtatcg gaacctttaa caagcttgac      180
ggagagctga ttgggtttga cggcgaattt taccgtcttc gctcagacgg aaccgcgaca      240
ccggtccaaa atggagaccg ttcaccgttc tggtcattta cgttctttac accggacatg      300
acgcacaaaa ttgatgcgaa aatgacacgc gaagactttg aaaaagagat caacagcatg      360
ctgccaaagca gaaacttatt ttatgcaatt cgcattgacg gattgtttta aaaggtgcag      420
acaagaacag tagaacttca agaaaaacct tacgtgccaa tgggtgaagc ggtcaaaaca      480
cagccgattt tcaacttcga caacgtgaga ggaacgattg taggtttctt gacaccagct      540
tatgcaaacg gaatcgccgt ttctggctat cacctgcact tcattgacga aggacgcaat      600
tcaggcggac acgtttttga ctatgtgctt gaggattgca cggttacgat ttctcaaaaa      660
atgaacatga atctcagact tccgaacaca gcggatttct ttaatgcgaa tctggataac      720
cctgattttg cgaaagatat cgaaacaact gaaggaagcc ctgaataa      768

```

<210> 81

<211> 255

<212> PRT

<213> *Bacillus subtilis*

<400> 81

```

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

```

```

Gln Pro Val Ser Gln Ile Tyr Gln Val Ser Thr Met Thr Ser Leu Leu
          20           25           30

```

```

Asp Gly Val Tyr Asp Gly Asp Phe Glu Leu Ser Glu Ile Pro Lys Tyr
          35           40           45

```

```

Gly Asp Phe Gly Ile Gly Thr Phe Asn Lys Leu Asp Gly Glu Leu Ile
          50           55           60

```

```

Gly Phe Asp Gly Glu Phe Tyr Arg Leu Arg Ser Asp Gly Thr Ala Thr
65           70           75           80

```

```

Pro Val Gln Asn Gly Asp Arg Ser Pro Phe Cys Ser Phe Thr Phe Phe
          85           90           95

```

387
H

Thr Pro Asp Met Thr His Lys Ile Asp Ala Lys Met Thr Arg Glu Asp
100 105 110

Phe Glu Lys Glu Ile Asn Ser Met Leu Pro Ser Arg Asn Leu Phe Tyr
115 120 125

Ala Ile Arg Ile Asp Gly Leu Phe Lys Lys Val Gln Thr Arg Thr Val
130 135 140

Glu Leu Gln Glu Lys Pro Tyr Val Pro Met Val Glu Ala Val Lys Thr
145 150 155 160

Gln Pro Ile Phe Asn Phe Asp Asn Val Arg Gly Thr Ile Val Gly Phe
165 170 175

Leu Thr Pro Ala Tyr Ala Asn Gly Ile Ala Val Ser Gly Tyr His Leu
180 185 190

His Phe Ile Asp Glu Gly Arg Asn Ser Gly Gly His Val Phe Asp Tyr
195 200 205

Val Leu Glu Asp Cys Thr Val Thr Ile Ser Gln Lys Met Asn Met Asn
210 215 220

Leu Arg Leu Pro Asn Thr Ala Asp Phe Phe Asn Ala Asn Leu Asp Asn
225 230 235 240

Pro Asp Phe Ala Lys Asp Ile Glu Thr Thr Glu Gly Ser Pro Glu
245 250 255

<210> 82

<211> 780

<212> DNA

<213> *Klebsiella terrigena*

<400> 82

gtgaatcatt atcctgaatg cacctgccag gagagcctgt gcgaaaccgt acgcggcttc 60

tccgcccacc accctgatag cgttatctat cagacctctc tgatgagcgc gctgctgagc 120

ggggtctatg agggtagcac caccatcgcc gacctgctga cccacggcga cttcgggtctc 180

ggcaccttta acgaactcga tggcgaactg attgccttta gcagcgaggt ctaccagctg 240

cgcgctgacg gcagcgcgcg taaagcccg gcggatcaaa aaacgccctt cgcggtgatg 300

acctggttca gaccgcagta ccgtaaaacc ttgaccacc cggtcagccg ccagcagctg 360

cacgacgtta tcgaccagca aatccctcc gataacctgt tctgcgcctt gcatattgat 420

388

gggtcactttc gccacgcccac caccgcgacc gtgccgcggc agacgccgcc ctatcgggcg 480
 atgaccgacg tgctcgatga ccagccggtt ttccgcttca accagcgcaa ggggacgctg 540
 gtcggctttc gcaccccgca gcatatgcag ggccttaacg ttgccggcta ccacgagcac 600
 tttattaccg acgatcgcca gggcggcggc catctgctgg actaccagct cgatagcggc 660
 gtgctgacct tcggcgagat ccacaagctg atgattgacc tcccggccga cagcgctttc 720
 ctgcaggccg acctgcatcc tgacaatctc gatgccgcta ttcgtgcggt agaaaactaa 780

<210> 83
 <211> 259
 <212> PRT
 <213> *Klebsiella terrigena*

<400> 83

Met Asn His Tyr Pro Glu Cys Thr Cys Gln Glu Ser Leu Cys Glu Thr
1 5 10 15

Val Arg Gly Phe Ser Ala His His Pro Asp Ser Val Ile Tyr Gln Thr
20 25 30

Ser Leu Met Ser Ala Leu Leu Ser Gly Val Tyr Glu Gly Ser Thr Thr
35 40 45

Ile Ala Asp Leu Leu Thr His Gly Asp Phe Gly Leu Gly Thr Phe Asn
50 55 60

Glu Leu Asp Gly Glu Leu Ile Ala Phe Ser Ser Glu Val Tyr Gln Leu
65 70 75 80

Arg Ala Asp Gly Ser Ala Arg Lys Ala Arg Ala Asp Gln Lys Thr Pro
85 90 95

Phe Ala Val Met Thr Trp Phe Arg Pro Gln Tyr Arg Lys Thr Phe Asp
100 105 110

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

Pro Ser Asp Asn Leu Phe Cys Ala Leu His Ile Asp Gly His Phe Arg
130 135 140

His Ala His Thr Arg Thr Val Pro Arg Gln Thr Pro Pro Tyr Arg Ala
145 150 155 160

Met Thr Asp Val Leu Asp Asp Gln Pro Val Phe Arg Phe Asn Gln Arg
165 170 175

Lys Gly Thr Leu Val Gly Phe Arg Thr Pro Gln His Met Gln Gly Leu
180 185 190

Asn Val Ala Gly Tyr His Glu His Phe Ile Thr Asp Asp Arg Gln Gly
195 200 205

Gly Gly His Leu Leu Asp Tyr Gln Leu Asp Ser Gly Val Leu Thr Phe
210 215 220

Gly Glu Ile His Lys Leu Met Ile Asp Leu Pro Ala Asp Ser Ala Phe
225 230 235 240

Leu Gln Ala Asp Leu His Pro Asp Asn Leu Asp Ala Ala Ile Arg Ala
245 250 255

Val Glu Asn

<210> 84
<211> 1053
<212> DNA
<213> *Bacillus cereus*

<400> 84
atgaaagcac tactttggca taatcaacgt gatgtacgag tagaagaagt accagaacca 60
acagtaaaac caggaacagt gaaaatcaaa gttaaattgt gtggtatttg tgggacagac 120
ttgcatgaat atttagcagg gcctatTTTT attccaacag aagaacatcc attaacacat 180
gtgaaagcac ctgtttttt aggtcatgag tttagtgggt aggtaataga gattgggtgaa 240
ggagttacat ctcataaagt gggagaccgc gttgtttag agccaattta ttcttgtggt 300
aaatgtgaag cttgtaaaca tggacattac aatgtttgtg aacaacttgt tttccacggt 360
cttggcggag aaggcggcgg tttctctgaa tatacagtag taccagaaga tatggttcat 420
cacattccag atgaaatgac gtatgaacaa ggtgcgcttg tagaaccagc agcagtagca 480
gttcatgcag tacgtcaaag taaattaaaa gaaggggaag ctgtagcggg atttggttgc 540
ggtccaattg gacttcttgt tatccaagca gctaaagcag caggagcaac tcctgttatt 600
gcagttgaac tttctaaaga acgtcaagag ttagcgaaat tagcaggtgc ggattatgta 660
ttaaattccag caactcaaga tgtgttagct gaaattcgta acttaacaaa tggtttaggt 720
gtaaatgtta gctttgaagt aacagggtgt gaagttgtac tacgccaagc gattgaaagt 780

acaagcttcg aaggacaaac tgtaattggt agtgtatggg aaaaagacgc aacaattact 840
 ccaaataact tagtattaaa agaaaaagaa gttattggta ttttaggata ccgtcacatc 900
 ttcccagctg ttattaaatt gattagctcc ggtcaaattc aagcagagaa attaattacg 960
 aaaaaaatta cagtggatca agttgttgaa gaaggatttg aagcacttgt aaaagataaa 1020
 acacaagtga aaattcttgt ttcacctaaa taa 1053

<210> 85
 <211> 350
 <212> PRT
 <213> *Bacillus cereus*

<400> 85

Met Lys Ala Leu Leu Trp His Asn Gln Arg Asp Val Arg Val Glu Glu
 1 5 10 15

Val Pro Glu Pro Thr Val Lys Pro Gly Thr Val Lys Ile Lys Val Lys
 20 25 30

Trp Cys Gly Ile Cys Gly Thr Asp Leu His Glu Tyr Leu Ala Gly Pro
 35 40 45

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

Val Ile Leu Gly His Glu Phe Ser Gly Glu Val Ile Glu Ile Gly Glu
 65 70 75 80

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

Tyr Ser Cys Gly Lys Cys Glu Ala Cys Lys His Gly His Tyr Asn Val
 100 105 110

Cys Glu Gln Leu Val Phe His Gly Leu Gly Gly Glu Gly Gly Gly Phe
 115 120 125

Ser Glu Tyr Thr Val Val Pro Glu Asp Met Val His His Ile Pro Asp
 130 135 140

Glu Met Thr Tyr Glu Gln Gly Ala Leu Val Glu Pro Ala Ala Val Ala
 145 150 155 160

Val His Ala Val Arg Gln Ser Lys Leu Lys Glu Gly Glu Ala Val Ala
 165 170 175

391
AA

Val Phe Gly Cys Gly Pro Ile Gly Leu Leu Val Ile Gln Ala Ala Lys
180 185 190

Ala Ala Gly Ala Thr Pro Val Ile Ala Val Glu Leu Ser Lys Glu Arg
195 200 205

Gln Glu Leu Ala Lys Leu Ala Gly Ala Asp Tyr Val Leu Asn Pro Ala
210 215 220

Thr Gln Asp Val Leu Ala Glu Ile Arg Asn Leu Thr Asn Gly Leu Gly
225 230 235 240

Val Asn Val Ser Phe Glu Val Thr Gly Val Glu Val Val Leu Arg Gln
245 250 255

Ala Ile Glu Ser Thr Ser Phe Glu Gly Gln Thr Val Ile Val Ser Val
260 265 270

Trp Glu Lys Asp Ala Thr Ile Thr Pro Asn Asn Leu Val Leu Lys Glu
275 280 285

Lys Glu Val Ile Gly Ile Leu Gly Tyr Arg His Ile Phe Pro Ala Val
290 295 300

Ile Lys Leu Ile Ser Ser Gly Gln Ile Gln Ala Glu Lys Leu Ile Thr
305 310 315 320

Lys Lys Ile Thr Val Asp Gln Val Val Glu Glu Gly Phe Glu Ala Leu
325 330 335

Val Lys Asp Lys Thr Gln Val Lys Ile Leu Val Ser Pro Lys
340 345 350

<210> 86

<211> 1053

<212> DNA

<213> *Bacillus cereus*

<400> 86

atgaaagcac tacttttgga taatcaacgt gatgtacgag tagaagaagt accagaacca 60

acagtaaaac caggaacagt gaaaatcaaa gttaaattggt gtggtatttg tgggacagac 120

ttgcatgaat atttagcagg gcctatTTTT attccaacag aagaacatcc attaacacat 180

gtgaaagcac ctgttatttt aggtcatgag tttagtgggt aggtaataga gattggtgaa 240

ggagttacat ctcataaagt gggagaccgc gttgtttag agccaattta ttcttgtggt 300

392
A

```

aaatgtgaag cttgtaaaca tggacattac aatgtttgtg aacaacttgt tttccacggt 360
cttggcggag aaggcggcgg tttctctgaa tatacagtag taccagaaga tatggttcat 420
cacattccag atgaaatgac gtatgaacaa ggtgcgcttg tagaaccagc agcagtagca 480
gttcatgcag tacgtcaaag taaattaaaa gaaggggaag ctgtagcggg atttggttgc 540
ggtccaattg gacttcttgt tatccaagca gctaaagcag caggagcaac tcctgttatt 600
gcagttgaac tttctaaaga acgtcaagag ttagcgaaat tagcagggtgc ggattatgta 660
ttaaatccag caactcaaga tgtgttagct gaaattcgta acttaacaaa tggtttaggt 720
gtaaatgtta gctttgaagt aacagggtgtt gaagttgtac tacgccaagc gattgaaagt 780
acaagcttcg aaggacaaac tgtaattgtt agtgtatggg aaaaagacgc aacaattact 840
ccaaataact tagtattaaa agaaaaagaa gttattggta ttttaggata ccgtcacatc 900
ttcccagctg ttattaaatt gattagctcc ggtcaaattc aagcagagaa attaattacg 960
aaaaaaatta cagtggatca agttgttgaa gaaggatttg aagcacttgt aaaagataaa 1020
acacaagtga aaattcttgt ttcacctaaa taa 1053

```

```

<210> 87
<211> 350
<212> PRT
<213> Bacillus cereus

```

```

<400> 87

```

```

Met Lys Ala Leu Leu Trp His Asn Gln Arg Asp Val Arg Val Glu Glu
1           5           10           15

```

```

Val Pro Glu Pro Thr Val Lys Pro Gly Thr Val Lys Ile Lys Val Lys
          20           25           30

```

```

Trp Cys Gly Ile Cys Gly Thr Asp Leu His Glu Tyr Leu Ala Gly Pro
35           40           45

```

```

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

```

```

Val Ile Leu Gly His Glu Phe Ser Gly Glu Val Ile Glu Ile Gly Glu
65           70           75           80

```

```

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

```

```

Tyr Ser Cys Gly Lys Cys Glu Ala Cys Lys His Gly His Tyr Asn Val
100           105           110

```

Cys Glu Gln Leu Val Phe His Gly Leu Gly Gly Glu Gly Gly Gly Phe
115 120 125

Ser Glu Tyr Thr Val Val Pro Glu Asp Met Val His His Ile Pro Asp
130 135 140

Glu Met Thr Tyr Glu Gln Gly Ala Leu Val Glu Pro Ala Ala Val Ala
145 150 155 160

Val His Ala Val Arg Gln Ser Lys Leu Lys Glu Gly Glu Ala Val Ala
165 170 175

Val Phe Gly Cys Gly Pro Ile Gly Leu Leu Val Ile Gln Ala Ala Lys
180 185 190

Ala Ala Gly Ala Thr Pro Val Ile Ala Val Glu Leu Ser Lys Glu Arg
195 200 205

Gln Glu Leu Ala Lys Leu Ala Gly Ala Asp Tyr Val Leu Asn Pro Ala
210 215 220

Thr Gln Asp Val Leu Ala Glu Ile Arg Asn Leu Thr Asn Gly Leu Gly
225 230 235 240

Val Asn Val Ser Phe Glu Val Thr Gly Val Glu Val Val Leu Arg Gln
245 250 255

Ala Ile Glu Ser Thr Ser Phe Glu Gly Gln Thr Val Ile Val Ser Val
260 265 270

Trp Glu Lys Asp Ala Thr Ile Thr Pro Asn Asn Leu Val Leu Lys Glu
275 280 285

Lys Glu Val Ile Gly Ile Leu Gly Tyr Arg His Ile Phe Pro Ala Val
290 295 300

Ile Lys Leu Ile Ser Ser Gly Gln Ile Gln Ala Glu Lys Leu Ile Thr
305 310 315 320

Lys Lys Ile Thr Val Asp Gln Val Val Glu Glu Gly Phe Glu Ala Leu
325 330 335

Val Lys Asp Lys Thr Gln Val Lys Ile Leu Val Ser Pro Lys
340 345 350

394
AS

<210> 88
 <211> 1113
 <212> DNA
 <213> *Lactococcus lactis*

<400> 88
 ttgcctgaaa cgacaaccat cctatataga ggaggcggtt ttatgcgcgc agcacgtttt 60
 tacgaccgcg gggatatccg cattgatgaa attaatgaac caatagtaaa agctggccaa 120
 gttggcattg atgtggcttg gtgtggaatt tgtggaacag atctccatga attttttagat 180
 ggcccaattt tttgtccgtc agcagaacat cctaatacaa ttactggaga agtaccacca 240
 gtcactcttg gacatgaaat gtctgggggt gtaaatttta taggtgaagg agtaagcgga 300
 cttaaagtag gtgaccatgt cgttgtcgaa ccttatatcg ttcccgaagg gactgataca 360
 agtgaaactg gacattataa cctctcagaa ggctcaaact ttattgggtt gggcggaagt 420
 ggtggagggt tggctgaaaa aatttctggt gatgaacggt gggttcaca aattcctgat 480
 aacttaccat tggatgaagc tgctctaatt gagccactat cagtcggcta tcacgctggt 540
 gaacgagcaa atttaagtga aaagagtacg gtattagttg ttggtgctgg accaattgga 600
 ctattaactg ctgccgttgc aaaagcgcaa ggacatactg ttatcatcag tgaacctagt 660
 ggacttcgtc gtaaaaaagc acaagaagca caagttgctg attatttctt caatccaatt 720
 gaagatgaca ttcaagctaa agttcatgaa attaatgaaa aaggagtggg cgcagccttt 780
 gaatgtacct ctgtccaacc gggatttgac gcttgtctag atgcgattcg tatgggtgga 840
 acagttgtca ttgtcgcaat ttggggcaag cctgctagtg ttgatatggc aaaattagta 900
 atcaaagaag ctaacctttt aggaacgatt gcttataata aactcatcc aaaaacaatt 960
 gatttagtat caacaggtaa aataaaattg gaccaattca tcacagctaa aatcggtttg 1020
 gatgatttga ttgataaagg attcgatagc ctgattcatc ataataaac agctgttaaa 1080
 attttagttt caccaactgg taaaggtcta taa 1113

<210> 89
 <211> 370
 <212> PRT
 <213> *Lactococcus lactis*

<400> 89

Met Pro Glu Thr Thr Thr Ile Leu Tyr Arg Gly Gly Val Phe Met Arg
 1 5 10 15

Ala Ala Arg Phe Tyr Asp Arg Gly Asp Ile Arg Ile Asp Glu Ile Asn
 20 25 30

395

Glu Pro Ile Val Lys Ala Gly Gln Val Gly Ile Asp Val Ala Trp Cys
 35 40 45

Gly Ile Cys Gly Thr Asp Leu His Glu Phe Leu Asp Gly Pro Ile Phe
 50 55 60

Cys Pro Ser Ala Glu His Pro Asn Pro Ile Thr Gly Glu Val Pro Pro
 65 70 75 80

Val Thr Leu Gly His Glu Met Ser Gly Val Val Asn Phe Ile Gly Glu
 85 90 95

Gly Val Ser Gly Leu Lys Val Gly Asp His Val Val Val Glu Pro Tyr
 100 105 110

Ile Val Pro Glu Gly Thr Asp Thr Ser Glu Thr Gly His Tyr Asn Leu
 115 120 125

Ser Glu Gly Ser Asn Phe Ile Gly Leu Gly Gly Asn Gly Gly Gly Leu
 130 135 140

Ala Glu Lys Ile Ser Val Asp Glu Arg Trp Val His Lys Ile Pro Asp
 145 150 155 160

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

Tyr His Ala Val Glu Arg Ala Asn Leu Ser Glu Lys Ser Thr Val Leu
 180 185 190

Val Val Gly Ala Gly Pro Ile Gly Leu Leu Thr Ala Ala Val Ala Lys
 195 200 205

Ala Gln Gly His Thr Val Ile Ile Ser Glu Pro Ser Gly Leu Arg Arg
 210 215 220

Lys Lys Ala Gln Glu Ala Gln Val Ala Asp Tyr Phe Phe Asn Pro Ile
 225 230 235 240

Glu Asp Asp Ile Gln Ala Lys Val His Glu Ile Asn Glu Lys Gly Val
 245 250 255

Asp Ala Ala Phe Glu Cys Thr Ser Val Gln Pro Gly Phe Asp Ala Cys
 260 265 270

396

Leu Asp Ala Ile Arg Met Gly Gly Thr Val Val Ile Val Ala Ile Trp
 275 280 285

Gly Lys Pro Ala Ser Val Asp Met Ala Lys Leu Val Ile Lys Glu Ala
 290 295 300

Asn Leu Leu Gly Thr Ile Ala Tyr Asn Asn Thr His Pro Lys Thr Ile
 305 310 315 320

Asp Leu Val Ser Thr Gly Lys Ile Lys Leu Asp Gln Phe Ile Thr Ala
 325 330 335

Lys Ile Gly Leu Asp Asp Leu Ile Asp Lys Gly Phe Asp Thr Leu Ile
 340 345 350

His His Asn Glu Thr Ala Val Lys Ile Leu Val Ser Pro Thr Gly Lys
 355 360 365

Gly Leu
 370

<210> 90
 <211> 705
 <212> DNA
 <213> *Pyrococcus furiosus*

<400> 90
 atgaagggttg ccgtaattac tggggcatcc cgtggaatcg gggaagctat agcaaaggcc 60
 cttgctgaag atggatattc ccttgcctta ggggctagaa gtgttgatag gttagagaag 120
 attgccaagg aactcagcga aaaacatggg gtggagggtat tttacgacta cctcgatgta 180
 tcaaaaccag aaagcgttga agagtttgca aggaaaacgc tagctcactt tggagatgtg 240
 gacgttggtt tggccaatgc ggggcttggg tactttggta ggcttgaaga gcttacagaa 300
 gagcagttcc acgaaatgat tgaagtaaac cttttgggag tttggagaac aataaaagct 360
 ttcttaaaact ccttaaagcg gactggagga gtggctattg ttgttacttc agatgtttct 420
 gcaaggctac ttccatacgg tggagggttat gtggcaacta aatgggctgc aagagcattg 480
 gtaaggacct tccagattga gaatccagat gtgagggttct tgcagctaag acctggagca 540
 gtagatacat attttgaggg gagcaaagct gggaagccaa aggagcaagg gtattttaaa 600
 cctgaggaag ttgctgaggg agtaaaatac ctctaagac ttccaaagga tgttagggtt 660
 gaggaattaa tgttgcgctc aatttatcaa aaacctgagt attga 705

<210> 91

<211> 234

<212> PRT

<213> *Pyrococcus furiosus*

<400> 91

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

Ile Ala Lys Ala Leu Ala Glu Asp Gly Tyr Ser Leu Ala Leu Gly Ala
20 25 30

Arg Ser Val Asp Arg Leu Glu Lys Ile Ala Lys Glu Leu Ser Glu Lys
35 40 45

His Gly Val Glu Val Phe Tyr Asp Tyr Leu Asp Val Ser Lys Pro Glu
50 55 60

Ser Val Glu Glu Phe Ala Arg Lys Thr Leu Ala His Phe Gly Asp Val
65 70 75 80

Asp Val Val Val Ala Asn Ala Gly Leu Gly Tyr Phe Gly Arg Leu Glu
85 90 95

Glu Leu Thr Glu Glu Gln Phe His Glu Met Ile Glu Val Asn Leu Leu
100 105 110

Gly Val Trp Arg Thr Ile Lys Ala Phe Leu Asn Ser Leu Lys Arg Thr
115 120 125

Gly Gly Val Ala Ile Val Val Thr Ser Asp Val Ser Ala Arg Leu Leu
130 135 140

Pro Tyr Gly Gly Gly Tyr Val Ala Thr Lys Trp Ala Ala Arg Ala Leu
145 150 155 160

Val Arg Thr Phe Gln Ile Glu Asn Pro Asp Val Arg Phe Phe Glu Leu
165 170 175

Arg Pro Gly Ala Val Asp Thr Tyr Phe Gly Gly Ser Lys Ala Gly Lys
180 185 190

Pro Lys Glu Gln Gly Tyr Leu Lys Pro Glu Glu Val Ala Glu Ala Val
195 200 205

Lys Tyr Leu Leu Arg Leu Pro Lys Asp Val Arg Val Glu Glu Leu Met
210 215 220

398
AS

Leu Arg Ser Ile Tyr Gln Lys Pro Glu Tyr
225 230

<210> 92
<211> 1665
<212> DNA
<213> *Salmonella typhimurium*

<400> 92
atgagatcga aaagatttga agcactggcg aaacgccctg tgaatcagga cggctttgtt 60
aaggagtgga tcgaagaagg ctttatcgcg atggaaagcc cgaacgaccc aaaaccgtcg 120
ataaaaatcg ttaacggcgc ggtaaccgag ctggacggaa aaccggttag cgaattcgac 180
ctgatcgacc actttatcgc ccgctacggc atcaacctga accgcgccga agaagtgatg 240
gcgatggatt cgggtcaagct ggctaacatg ctgtgcgatc cgaacgtcaa gcgcagcgaa 300
atcgtttccgc taaccaccgc gatgacccca gcgaaaattg tcgaagtggg ttcgcatatg 360
aacgtgggtg agatgatgat ggcgatgcag aaaatgcgcg cccgccgtac tccatctcaa 420
caggcgcacg tcaccaacgt taaagacaac ccggtgcaaa ttgccgccga tgccgccgaa 480
ggcgcatggc gcggggttga cgaacaagag acgacgggtg cggtagcgcg ctatgcgccg 540
ttcaacgcca tcgcgctgct gggttggttct caggtaggtc gtccgggggt actgactcaa 600
tgctcgctgg aagaagccac cgagctgaag ctccggcatgc tgggccacac ctgctacgcc 660
gaaaccatct ccgtttacgg caccgagccg gtcttcaccg acggtgacga taccatcg 720
tcgaagggtt tcttagcctc ttcctacgcc tctcgcggcc tgaaaatgcg cttcacctcc 780
ggctccggct ccgaagtgca gatgggctac gccgaaggca aatccatgct gtatctggaa 840
gcgcgctgca tctatatcac caaagccgcg ggcgttcagg ggctgcaaaa cggctccgta 900
agcagcatcg gcgtaccgtc tgccgtgccg tcaggcattc gtgccgtgct ggcggaaaac 960
ctgatctgct cttcgctgga tctggaatgc gcctccagta acgaccagac cttcacccac 1020
tccgatatgc gtcgtaccgc tcgcctgctg atgcagttcc tgccgggtac cgactttatc 1080
tcctccgggt attccgcggt gccgaactac gacaacatgt tcgccgggtc caacgaagat 1140
gcggaagact ttgacgacta caacgttata cagcgtgacc tgaaagtgga cggcgggtctg 1200
cgcccgggtc gcgaagagga cgttatcgcc atccgtaaca aagccgcccg cgcgctgcag 1260
gccgtgtttg ccggaatggg actgccgccg attaccgatg aagaagtga agccgcgacc 1320
tatgcccacg gttcgaaaga tatgccggag cgcaacatcg tcgaagacat caagtgcgcc 1380
caggaaatca tcaataaaaa ccgcaacggg ctggaagttg tgaaagcgct ggctcagggc 1440
gggtttaccg acgtggccca ggacatgctc aacatccaga aagccaagct aaccggcgac 1500

399
AS

tatttgcaca cctccgccat tatcgtcggc gacggacaag tgctctctgc ggттаатgac 1560
 gtcaatgact atgccgggtcc ggcaacaggt tatcgcctgc agggagaacg ctgggaagag 1620
 attaaaaaca tccctggcgc tcttgatccc aacgagattg attaa 1665

<210> 93
 <211> 554
 <212> PRT
 <213> *Salmonella typhimurium*

<400> 93

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Ala Val
 35 40 45

Thr Glu Leu Asp Gly Lys Pro Val Ser Glu Phe Asp Leu Ile Asp His
 50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Asn Arg Ala Glu Glu Val Met
 65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
 85 90 95

Lys Arg Ser Glu Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
 100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ala
 115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
 130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
 145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
 165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
 180 185 190

400
A

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Ser Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

4/21
AS

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Gly Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 94

<211> 675

<212> DNA

<213> *Salmonella typhimurium*

<400> 94

atggaaatta atgaaaaatt gctgcgccag ataattgaag acgtactccg cgatatgaag	60
ggcagcgata aaccggtctc gtttaatgcg cctgcggcat ccacagcacc acagaccgct	120
gcgcctgcgg gcgacggctt tctgaccgaa gtgggcgaag cgcgccaggg cactcagcag	180
gacgaagtca ttatcgccgt cggcccggca tttggcctgg cgcaaaccgt caatatcgtc	240
ggcttaccgc ataagagcat tctgcgcgaa gtcattgccg gtattgaaga agaaggcatc	300
aaggcgcgcg tgattcgctg ctttaaatct tccgacgtgg cgttcgctgc cgttgaaggt	360
aaccgcctga gcggatccgg catctccatc ggcatccagt cgaaaggtac tacggttatc	420
caccagcagg ggctaccgcc gctctccaac ctggagctgt tcccgcaggc accgctgctg	480
acgctggaaa cctaccgtca gattggtaaa aacgccgccc gctatgcgaa acgagaatca	540
ccgcagccgg tccctacgct caatgaccag atggcacgcc cgaagtacca ggcaaagtcg	600
gccattttgc atattaaaga gaccaagtac gtcgtgacgg gcaaaaaccc gcaggaactg	660

402

cgcggtggcgc tttga

675

<210> 95

<211> 224

<212> PRT

<213> *Salmonella typhimurium*

<400> 95

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
 1 5 10 15

Arg Asp Met Lys Gly Ser Asp Lys Pro Val Ser Phe Asn Ala Pro Ala
 20 25 30

Ala Ser Thr Ala Pro Gln Thr Ala Ala Pro Ala Gly Asp Gly Phe Leu
 35 40 45

Thr Glu Val Gly Glu Ala Arg Gln Gly Thr Gln Gln Asp Glu Val Ile
 50 55 60

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

Gly Leu Pro His Lys Ser Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
 85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
 115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
 130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
 145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
 165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
 180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
 195 200 205

403

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
 210 215 220

<210> 96
 <211> 522
 <212> DNA
 <213> *Salmonella typhimurium*

<400> 96
 atgaataaccg acgcaattga atcgatgggc cgggacgtat tgagccgcat gaacagcctg 60
 cagggcgatg cgccagcagc ggctcctgcg gcaggcggca cgtcccgcag cgcaaagggtc 120
 agcgactacc cgctggcgaa caaacacccg gaatgggtga aaaccgccac caataaaacg 180
 ctggacgact ttacgctgga aaacgtgctg agcaataaag tcaccgctca ggatatgcgt 240
 attaccccg aaaccctgcg cttacaggcc tctatcgcca aagatgcggg tcgcgaccgg 300
 ctggcgatga acttcgaacg cgccgccgaa ctgaccgcgg taccggacga tcgcattctt 360
 gaaatctaca acgccccttcg tccgtatcgt tcaacgaaag aagagctgct cgctatcgcc 420
 gacgatctcg aaaaccgtta tcaggcaaag atttgcgag ctttcgttcg tgaagcggca 480
 gggctgtacg ttgagcgtaa aaaactcaaa ggcgacgatt aa 522

<210> 97
 <211> 173
 <212> PRT
 <213> *Salmonella typhimurium*

<400> 97

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
 1 5 10 15

Met Asn Ser Leu Gln Gly Asp Ala Pro Ala Ala Ala Pro Ala Ala Gly
 20 25 30

Gly Thr Ser Arg Ser Ala Lys Val Ser Asp Tyr Pro Leu Ala Asn Lys
 35 40 45

His Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe
 50 55 60

Thr Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg
 65 70 75 80

Ile Thr Pro Glu Thr Leu Arg Leu Gln Ala Ser Ile Ala Lys Asp Ala
 85 90 95

Gly Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr
 100 105 110

Ala Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
 115 120 125

Tyr Arg Ser Thr Lys Glu Glu Leu Leu Ala Ile Ala Asp Asp Leu Glu
 130 135 140

Asn Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala
 145 150 155 160

Gly Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
 165 170

<210> 98

<211> 1677

<212> DNA

<213> *Lactobacillus collinoides*

<400> 98

ttggaacgtc aaaaaagatt tgaaaaatta gagaaacgtc cagtgcattt agatgggttc	60
gttaagaact gggacgacga aggttttagtt gcccttaacg gtaagaacga tccaaagcca	120
agcattacga tcgaaaacgg tgttgttact gaaatggatg gtaagaagaa ggcagacttc	180
gaccttatcg acaagtacat cgctgaatac gggatcaact tggacaatgc tgaaaagact	240
ttaaacacag attcagttaa gatcgccaac atgatgtgtg atcctaacgt ctcccgtgct	300
gaaattattg aatatacaac tgctatgaca ccagccaagg ctgctgaagt tatcagccag	360
ttaaacttcg ctgaaatgat catggcaact caaaagatgc ggccacgctg gaccctatg	420
actcaagtcc acgctaccaa cactttggat aaccagttg aaatcgctgc tgatgctgcc	480
gaagctgcat tacgtggggt tcctgaagaa gaaaccacca ctgccattgc tcggtatgcg	540
ccaatgaacg ctatttcaat catggttggg gcccaagcag gccgtcctgg tgttatcacc	600
caatgttcag ttgaagaagc tgacgaattg agtttgggga tgcgtgggtt tactgcctat	660
gctgaaacca tttcagttta tgggactgac cgggtcttca ctgatggtga tgataccct	720
tggtcaaaag gtttcttagc ttcttgctac gtttcacgtg gtttgaagat gcggtttact	780
tcaggtgccg gttcagaagc tatgatgggc tacactgaag gtaaataaat gctttacctt	840
gaagctcggt gtatctacat taccaaggcg tcaggtgttc aaggtctgca aaacgggtggt	900
gttagttgta tcgggatgcc aggtgccgtc gttggtggta tccgtgaagt cttaggtgaa	960
aacttactat gtatgtcact tgatgttgaa tgtgcttctg gttgtgacca agccttctct	1020

cactctgaca ttctgtcggac tggccggatg attggccaat tcatcgctgg tactgattac 1080
 ctgtcatcag gttacgctgc cgaagaaaac atggataaca ctttcgctgg ttcaaaccatg 1140
 gatgttcttg actacgatga ttacatcact ttggaacgtg atatggctat taacgggtggt 1200
 atcatgccaa ttaccgaaga ggaatctatt aagattcgtc acaaggctgc ggttgctatc 1260
 caagctgtct ttgatggctt aggcctacca cagatcactg atgaagaagt tgaagccgca 1320
 acttatggca gcaattcaaa cgacatgccaa aaacgtgaca tggttcaaga tatgaaagct 1380
 gctcaaggctc tgatgactcg tggcattact gttgttgacg ttatcaaggc cttatatgac 1440
 catgatatta aagacgtcgc tgaggctgtg cttaagttag cgcaacaaaa ggtttgtggt 1500
 gattacctgc aaacatctgc tgtcttcttg gatggttgga agtgtacttc agctattaac 1560
 aacgctaacg attacaaagg cccaggtact ggttaccgtc tatgggaaga caaagacaaa 1620
 tgggatcgtc tagaaaacgt tccgtgggct ttggatcctc agaagttgga attctaa 1677

<210> 99
 <211> 558
 <212> PRT
 <213> *Lactobacillus collinoides*

<400> 99

Met Glu Arg Gln Lys Arg Phe Glu Lys Leu Glu Lys Arg Pro Val His
1 5 10 15

Leu Asp Gly Phe Val Lys Asn Trp Asp Asp Glu Gly Leu Val Ala Leu
20 25 30

Asn Gly Lys Asn Asp Pro Lys Pro Ser Ile Thr Ile Glu Asn Gly Val
35 40 45

Val Thr Glu Met Asp Gly Lys Lys Lys Ala Asp Phe Asp Leu Ile Asp
50 55 60

Lys Tyr Ile Ala Glu Tyr Gly Ile Asn Leu Asp Asn Ala Glu Lys Thr
65 70 75 80

Leu Asn Thr Asp Ser Val Lys Ile Ala Asn Met Met Cys Asp Pro Asn
85 90 95

Val Ser Arg Ala Glu Ile Ile Glu Tyr Thr Thr Ala Met Thr Pro Ala
100 105 110

Lys Ala Ala Glu Val Ile Ser Gln Leu Asn Phe Ala Glu Met Ile Met
115 120 125

Ala Thr Gln Lys Met Arg Pro Arg Arg Thr Pro Met Thr Gln Val His
130 135 140

Ala Thr Asn Thr Leu Asp Asn Pro Val Glu Ile Ala Ala Asp Ala Ala
145 150 155 160

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

Ala Arg Tyr Ala Pro Met Asn Ala Ile Ser Ile Met Val Gly Ala Gln
180 185 190

Ala Gly Arg Pro Gly Val Ile Thr Gln Cys Ser Val Glu Glu Ala Asp
195 200 205

Glu Leu Ser Leu Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile
210 215 220

Ser Val Tyr Gly Thr Asp Arg Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Phe Thr Ser Gly Ala Gly Ser Glu Ala Met Met Gly Tyr Thr
260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr
275 280 285

Lys Ala Ser Gly Val Gln Gly Leu Gln Asn Gly Gly Val Ser Cys Ile
290 295 300

Gly Met Pro Gly Ala Val Val Gly Gly Ile Arg Glu Val Leu Gly Glu
305 310 315 320

Asn Leu Leu Cys Met Ser Leu Asp Val Glu Cys Ala Ser Gly Cys Asp
325 330 335

Gln Ala Phe Ser His Ser Asp Ile Arg Arg Thr Gly Arg Met Ile Gly
340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Leu Ser Ser Gly Tyr Ala Ala Glu
355 360 365

407
A

Glu Asn Met Asp Asn Thr Phe Ala Gly Ser Asn Met Asp Val Leu Asp
370 375 380

Tyr Asp Asp Tyr Ile Thr Leu Glu Arg Asp Met Ala Ile Asn Gly Gly
385 390 395 400

Ile Met Pro Ile Thr Glu Glu Glu Ser Ile Lys Ile Arg His Lys Ala
405 410 415

Ala Val Ala Ile Gln Ala Val Phe Asp Gly Leu Gly Leu Pro Gln Ile
420 425 430

Thr Asp Glu Glu Val Glu Ala Ala Thr Tyr Gly Ser Asn Ser Asn Asp
435 440 445

Met Pro Lys Arg Asp Met Val Gln Asp Met Lys Ala Ala Gln Gly Leu
450 455 460

Met Thr Arg Gly Ile Thr Val Val Asp Val Ile Lys Ala Leu Tyr Asp
465 470 475 480

His Asp Ile Lys Asp Val Ala Glu Ala Val Leu Lys Leu Ala Gln Gln
485 490 495

Lys Val Cys Gly Asp Tyr Leu Gln Thr Ser Ala Val Phe Leu Asp Gly
500 505 510

Trp Lys Cys Thr Ser Ala Ile Asn Asn Ala Asn Asp Tyr Lys Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Leu Trp Glu Asp Lys Asp Lys Trp Asp Arg Leu
530 535 540

Glu Asn Val Pro Trp Ala Leu Asp Pro Gln Lys Leu Glu Phe
545 550 555

<210> 100

<211> 693

<212> DNA

<213> *Lactobacillus collinoides*

<400> 100

gtgagttcag aaatcgatga aacattgctt agaaatatca ttaaaggcgt tttaaatgaa 60

gttcaaaact ctgatacgcc aatttccttt ggtggccaag atgcagcccc agttgccggt 120

gccaaggaag gtgccgcacc agaaaagaag ttggattggt tccaacacgt tggaatcgcc 180

aaaccaggtt tgtcaaagga tgaagttgta attggtgttg ccccagcatt tgctgaagtg 240
 ttgacgcaaa ctatgacgaa gatccaacac aaagacatcc tgcgtcaaatt cattgccgga 300
 gttgaagaag aagggtctcaa ggcccgtgtc gttaagggtt atcggacttc agacgtttcc 360
 ttcgtttccg ctgatgttga caagttgtca gggttcaggaa tttcagttgc cgttcaatca 420
 aaggggacaa cgattattca ccaaaaggat caagcacctg tgtcaaacct tgaattgttc 480
 ccacaggctc cagttttgac attggacgct taccgtcaaa tcggtaagaa cgctgcccag 540
 tatgctaagg gtatgtcacc aaccccagtg ccaacaatta acgaccagat ggcacgtgtg 600
 caatatcaag cactttctgc tttgatgcac atcaaggaaa caaaacaggt tgttggtggg 660
 aagcctgctg aagaaattaa ggtaaccttt tag 693

<210> 101
 <211> 230
 <212> PRT
 <213> *Lactobacillus collinoides*

<400> 101

Met Ser Ser Glu Ile Asp Glu Thr Leu Leu Arg Asn Ile Ile Lys Gly
1 5 10 15

Val Leu Asn Glu Val Gln Asn Ser Asp Thr Pro Ile Ser Phe Gly Gly
20 25 30

Gln Asp Ala Ala Pro Val Ala Gly Ala Lys Glu Gly Ala Ala Pro Glu
35 40 45

Lys Lys Leu Asp Trp Phe Gln His Val Gly Ile Ala Lys Pro Gly Leu
50 55 60

Ser Lys Asp Glu Val Val Ile Gly Val Ala Pro Ala Phe Ala Glu Val
65 70 75 80

Leu Thr Gln Thr Met Thr Lys Ile Gln His Lys Asp Ile Leu Arg Gln
85 90 95

Ile Ile Ala Gly Val Glu Glu Glu Gly Leu Lys Ala Arg Val Val Lys
100 105 110

Val Tyr Arg Thr Ser Asp Val Ser Phe Val Ser Ala Asp Val Asp Lys
115 120 125

Leu Ser Gly Ser Gly Ile Ser Val Ala Val Gln Ser Lys Gly Thr Thr
130 135 140

409
AA

Ile Ile His Gln Lys Asp Gln Ala Pro Leu Ser Asn Leu Glu Leu Phe
145 150 155 160

Pro Gln Ala Pro Val Leu Thr Leu Asp Ala Tyr Arg Gln Ile Gly Lys
165 170 175

Asn Ala Ala Gln Tyr Ala Lys Gly Met Ser Pro Thr Pro Val Pro Thr
180 185 190

Ile Asn Asp Gln Met Ala Arg Val Gln Tyr Gln Ala Leu Ser Ala Leu
195 200 205

Met His Ile Lys Glu Thr Lys Gln Val Val Val Gly Lys Pro Ala Glu
210 215 220

Glu Ile Lys Val Thr Phe
225 230

<210> 102

<211> 522

<212> DNA

<213> *Lactobacillus collinoides*

<400> 102

atgagtgaag tagatgactt agtagctaga attgctgctc agctacaaca aagtggaaac 60
gcttctagtg cctcaactag tgccggtact tctgctggtt ccgagaaaga attagggcgca 120
gcagattacc cactatttga aaagcaccca gatcaaatca agacgccatc aggtaaaaat 180
gttgaagaaa tcaccttgga aaatgttatt aacggcaagg tagacgcaaa ggatatgcgg 240
attacgcccc caaccctgaa gttacaaggt gaaattgctg ccaacgcagg tcggccagca 300
atccaacgga acttccagcg ggcttctgaa ttaacttcag ttcccgatga tggtgttttg 360
gacttatata attcattacg gccattccgt tcaaccaagc aagaattatt ggataccgcc 420
aaggagcttc gtgacaagta tcacgcacct atctgtgccg gctggttcga agaagcagcc 480
gaaaactacg aagtcaacaa gaagttgaag ggcgataact ag 522

<210> 103

<211> 173

<212> PRT

<213> *Lactobacillus collinoides*

<400> 103

Met Ser Glu Val Asp Asp Leu Val Ala Arg Ile Ala Ala Gln Leu Gln
1 5 10 15

4/10
H

Gln Ser Gly Asn Ala Ser Ser Ala Ser Thr Ser Ala Gly Thr Ser Ala
20 25 30

Gly Ser Glu Lys Glu Leu Gly Ala Ala Asp Tyr Pro Leu Phe Glu Lys
35 40 45

His Pro Asp Gln Ile Lys Thr Pro Ser Gly Lys Asn Val Glu Glu Ile
50 55 60

Thr Leu Glu Asn Val Ile Asn Gly Lys Val Asp Ala Lys Asp Met Arg
65 70 75 80

Ile Thr Pro Ala Thr Leu Lys Leu Gln Gly Glu Ile Ala Ala Asn Ala
85 90 95

Gly Arg Pro Ala Ile Gln Arg Asn Phe Gln Arg Ala Ser Glu Leu Thr
100 105 110

Ser Val Pro Asp Asp Val Val Leu Asp Leu Tyr Asn Ser Leu Arg Pro
115 120 125

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

Asp Lys Tyr His Ala Pro Ile Cys Ala Gly Trp Phe Glu Glu Ala Ala
145 150 155 160

Glu Asn Tyr Glu Val Asn Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 104

<211> 1665

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 104

atgagatcga aaagatttga agcactggcg aaacgccctg tgaatcagga tggtttcggt 60
aaggagtgga ttgaagaggg ctttatcgcg atggaaagtc ctaacgatcc caaaccttct 120
atccgcacgc tcaacggcgc ggtgaccgaa ctgcacggta aaccggttga cgagttcgac 180
ctgattgacc actttatcgcg gcgctacggc attaatctcg cccgggcccga agaagtgatg 240
gccatggatt cggttaagct cgccaacatg ctctgcgacc cgaacgttaa acgcagcgac 300
atcgtgccgc tcactaccgc gatgaccccg gcgaaaatcg tggaagtggg gtcgcatatg 360
aacgtggtcg agatgatgat ggcgatgcaa aaaatgcgcg cccgccgcac gccgtcccag 420

411
AS

caggcgcatg tcactaatat caaagataat ccggtacaga ttgccgccga cgccgctgaa 480
 ggcgcatggc gcggccttga cgaacaggag accaccgtcg ccgtggcgcg ctacgcgcgg 540
 ttcaacgcca tcgccctgct ggtgggttca caggttggcc gccccggcgt cctcaccag 600
 tgttcgctgg aagaagccac cgagctgaaa ctgggcatgc tgggccacac ctgctatgcc 660
 gaaaccattt cgggtatacgg tacggaaccg gtgtttaccg atggcgatga cactccatgg 720
 tcgaaaggct tcctgcctc ctctacgcc tcgcgcggcc tgaaaatgcg ctttacctcc 780
 gggtccgggt ctgaagtaca gatgggctat gccgaaggca aatcgatgct ttatctcgaa 840
 gcgcgctgca tctacatcac caaagccgcc ggggtgcaag gcctgcagaa tggctccgtc 900
 agctgtatcg gcgtaccgtc cgccgtgccg tccgggatcc gcgccgtact ggcggaaaac 960
 ctgatctgct cagcgctgga tctggagtgc gcctccagca acgatcaaac ctttaccac 1020
 tcggatatgc ggcgtaccgc gcgtctgctg atgcagttcc tgccaggcac cgacttcac 1080
 tcctccgggt actcggcggg gcccaactac gacaacatgt tcgccgggtc caacgaagat 1140
 gccgaagact tcgatgacta caacgtgatc cagcgcgacc tgaaggtcga tgggggtctg 1200
 cggccgggtg gtgaagagga cgtgatcgcc attcgcaaca aagccgcccg cgcgctgcag 1260
 gcggtatttg ccggcatggg ttgcccgcct attacggatg aagaggtaga agccgccacc 1320
 tacgccacg gttcaaaaga tatgcctgag cgcaatatcg tcgaggacat caagtttgct 1380
 caggagatca tcaacaagaa ccgcaacggc ctggagggtg tgaaagccct ggcgaaaggc 1440
 ggcttccccg atgtcgccca ggacatgctc aatattcaga aagccaagct caccggcgac 1500
 tacctgcata cctccgccat cattgttggc gagggccagg tgctctcggc cgtgaatgac 1560
 gtgaacgatt atgccgggtc ggcaacaggc taccgcctgc aaggcgagcg ctgggaagag 1620
 attaaaaata tcccgggcgc gctcgatccc aatgaacttg gctaa 1665

<210> 105

<211> 554

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 105

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Arg Ile Val Asn Gly Ala Val
 35 40 45

412
AA

Thr Glu Leu Asp Gly Lys Pro Val Asp Glu Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Ala Arg Ala Glu Glu Val Met
65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Lys Arg Ser Asp Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ala
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Ile Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Arg Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

413
As

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ala Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Gly Leu Glu Val Val Lys Ala Leu Ala Lys Gly
465 470 475 480

Gly Phe Pro Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Gly Glu Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

4/4
AS

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Leu Gly
545 550

<210> 106

<211> 687

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 106

atggaaatta acgaaacgct gctgcgccag attatcgaag aggtgctgtc ggagatgaaa 60
tcaggcgcag ataagccggt ctcttttagc gcgccggcgt ctgtcgccctc tgccgcgccg 120
gtcgccggttg cgctgtgtc cggcgacagc ttcctgacgg aaatcggcga agccaaaccc 180
ggcacgcagc aggatgaagt cattattgcc gtcgggccag cgtttgggtct ggcgcaaacc 240
gccaatatcg tcggcattcc gcataaaaat attctgcgcg aagtgatcgc cggcattgag 300
gaagaaggca tcaaagcccc ggtgatccgc tgctttaagt catctgacgt cgccttcgtg 360
gcagtggaag gcaaccgcct gagcggtcc gccatctcga tcggtattca gtcgaaaggc 420
accaccgtca tccaccagcg cggcctgccg ccgctttcca atctggaact cttcccgag 480
gcgccgtgt taacgctgga aacctaccgt cagattggca aaaacgccgc gcgctacgcc 540
aaacgcgagt cgccgcagcc ggtgccgacg cttaacgatc agatggctcg tcccaaatac 600
caggcgaagt cggccatttt gcacattaaa gagaccaa atcgtggtgac gggcaaaaac 660
ccgcaggaac tgcgcgtggc gctttaa 687

<210> 107

<211> 228

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 107

Met Glu Ile Asn Glu Thr Leu Leu Arg Gln Ile Ile Glu Glu Val Leu
1 5 10 15

Ser Glu Met Lys Ser Gly Ala Asp Lys Pro Val Ser Phe Ser Ala Pro
20 25 30

Ala Ser Val Ala Ser Ala Ala Pro Val Ala Val Ala Pro Val Ser Gly
35 40 45

Asp Ser Phe Leu Thr Glu Ile Gly Glu Ala Lys Pro Gly Thr Gln Gln
50 55 60

415

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

Ala Asn Ile Val Gly Ile Pro His Lys Asn Ile Leu Arg Glu Val Ile
85 90 95

Ala Gly Ile Glu Glu Glu Gly Ile Lys Ala Arg Val Ile Arg Cys Phe
100 105 110

Lys Ser Ser Asp Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser
115 120 125

Gly Ser Gly Ile Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile
130 135 140

His Gln Arg Gly Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln
145 150 155 160

Ala Pro Leu Leu Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala
165 170 175

Ala Arg Tyr Ala Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn
180 185 190

Asp Gln Met Ala Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His
195 200 205

Ile Lys Glu Thr Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu
210 215 220

Arg Val Ala Leu
225

<210> 108

<211> 525

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 108

atgaataccg acgcaattga atccatggta cgcgacgtgc tgagccggat gaacagccta 60

caggacgggg taacgcccgc gccagccgcg ccgacaaacg acaccgttcg ccagccaaaa 120

gtagcgact acccgttagc gacctgccat ccggagtggg tcaaaaccgc taccaataaa 180

acgctcgatg acctgacgct ggagaacgta ttaagcgatc gcgttacggc gcaggacatg 240

cgcatactc cggaacgct gcgtatgcag gcggcgatcg ccaggatgc cggacgcgat 300

416

cggctggcga tgaactttga gcgggccgca gagctcaccg cggttcccga cgaccgaatc 360
 cttgagatct acaacgccct gcgccatac cgttccaccc aggcggagct actggcgatc 420
 gctgatgacc tcgagcatcg ctaccaggca cgactctgtg ccgcctttgt tcgggaagcg 480
 gccgggctgt acatcgagcg taagaagctg aaaggcgacg attaa 525

<210> 109

<211> 174

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 109

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
 1 5 10 15

Met Asn Ser Leu Gln Asp Gly Val Thr Pro Ala Pro Ala Ala Pro Thr
 20 25 30

Asn Asp Thr Val Arg Gln Pro Lys Val Ser Asp Tyr Pro Leu Ala Thr
 35 40 45

Cys His Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp
 50 55 60

Leu Thr Leu Glu Asn Val Leu Ser Asp Arg Val Thr Ala Gln Asp Met
 65 70 75 80

Arg Ile Thr Pro Glu Thr Leu Arg Met Gln Ala Ala Ile Ala Gln Asp
 85 90 95

Ala Gly Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu
 100 105 110

Thr Ala Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg
 115 120 125

Pro Tyr Arg Ser Thr Gln Ala Glu Leu Leu Ala Ile Ala Asp Asp Leu
 130 135 140

Glu His Arg Tyr Gln Ala Arg Leu Cys Ala Ala Phe Val Arg Glu Ala
 145 150 155 160

Ala Gly Leu Tyr Ile Glu Arg Lys Lys Leu Lys Gly Asp Asp
 165 170

418


<210> 110
 <211> 1833
 <212> DNA
 <213> *Klebsiella oxytoca*

<400> 110
 atgcatata tagctggcat tgatatcggc aactcatcga cggaagtcgc cctggcgacc 60
 ctggatgagg ctggcgcgct gacgatcacc cacagcgcgc tggcggaac caccggaatc 120
 aaaggcacgt tgcgtaacgt gtccgggatt caggaggcgc tcgccctcgt cgccagaggc 180
 gccgggatcg ccgtcagcga tatttcgctc atccgcatca acgaagcgac gccggtgatt 240
 ggcgatgtgg cgatggaaac cattaccgaa accatcatca ccgaatcgac catgatcggc 300
 cataacccga aaacgcccgg cggcgcgggg cttggcacag gcatcaccat tacgcccag 360
 gagctgctaa cccgcccggc ggacgcgccc tatatcctgg tgggtgctgc ggcgttcgat 420
 tttgccgata tcgccagcgt gattaacgct tccctgcgcy ccgggtatca gattaccggc 480
 gtcattttac agcgcgacga tggcgtgctg gtcagcaacc ggctggaaaa accgctgccg 540
 atcggtgacg aagtgtgta catcgaccgc attccgctgg ggatgctggc ggcgattgag 600
 gtcgccgttc cggggaaggt catcgaaacc ctctctaacc cttacggcat cgccaccgtc 660
 tttaacctca gccccgagga gacgaagaac atcgtcccga tggccccggc gctgattggc 720
 aaccgttccg ccgtggtggt caaaacgcca tccggcgacg tcaaagcgcy cgcgataccc 780
 gccggtaatc ttgagctgct ggcccagggc cgtagcgtgc gcgtggatgt ggccgcccgc 840
 gccgaagcca tcatgaaagc ggtcgacggc tgcggcaggc tcgataacgt caccggcgaa 900
 tccggcacca atatcggcgg catgctggaa cacgtgcgcc agaccatggc cgagctgacc 960
 aacaagccga gcagcgaaat atttattcag gacctgctgg ccgttgatac ctcggtaccg 1020
 gtgagcgta cggcggtct ggccggggag ttctcgtgag agcaggccgt gggcatcgcc 1080
 tcgatggtga aatcggatcg cctgcagatg gcaatgatcg cccgcgaaat cgagcagaag 1140
 ctcaatatcg acgtgcagat cggcggcgca gaggccgaag ccgccatcct gggggcgctg 1200
 accacgccgg gcaccaccg accgctggcg atcctcgacc tcggcgcggg ctccaccgat 1260
 gcctccatca tcaaccccaa aggcgacatc atcgccacc atctcgccgg cgcaggcgac 1320
 atggtgacga tgattattgc ccgcgagctg gggctggaag accgctatct ggcggaagag 1380
 atcaagaagt acccgctggc taagggtgaa agcctgttcc atttacgcca cgaggacggc 1440
 agcgtgcagt tcttctccac gccgctgccg cccgccgtgt tcgcccgcgt ctgctgggtg 1500
 aaagcggacg aactggtgcc gctgcccggc gatttagcgc tggaaaaagt gcgcgccatt 1560
 cgccgcagcy ccaaagagcy ggtctttgtc accaacgccc tgcgcgcgct gcgtcaggtc 1620

413
A

agccccaccg gcaacattcg cgatattccg ttcgtggtgc tggtcggcgg ttcgtcgctg 1680
 gatttcgaag tcccgcagct ggtcaccgat gcgctggcgc actaccgcct ggttgccgga 1740
 cggggaaata ttcgcggcag cgagggcccc cgaaacgcgg tggccaccgg cctgattctc 1800
 tcctggcata aggagtttgc gcatgaacgg taa 1833

<210> 111
 <211> 610
 <212> PRT
 <213> *Klebsiella oxytoca*

<400> 111

Met Arg Tyr Ile Ala Gly Ile Asp Ile Gly Asn Ser Ser Thr Glu Val
 1 5 10 15

Ala Leu Ala Thr Leu Asp Glu Ala Gly Ala Leu Thr Ile Thr His Ser
 20 25 30

Ala Leu Ala Glu Thr Thr Gly Ile Lys Gly Thr Leu Arg Asn Val Phe
 35 40 45

Gly Ile Gln Glu Ala Leu Ala Leu Val Ala Arg Gly Ala Gly Ile Ala
 50 55 60

Val Ser Asp Ile Ser Leu Ile Arg Ile Asn Glu Ala Thr Pro Val Ile
 65 70 75 80

Gly Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser
 85 90 95

Thr Met Ile Gly His Asn Pro Lys Thr Pro Gly Gly Ala Gly Leu Gly
 100 105 110

Thr Gly Ile Thr Ile Thr Pro Gln Glu Leu Leu Thr Arg Pro Ala Asp
 115 120 125

Ala Pro Tyr Ile Leu Val Val Ser Ser Ala Phe Asp Phe Ala Asp Ile
 130 135 140

Ala Ser Val Ile Asn Ala Ser Leu Arg Ala Gly Tyr Gln Ile Thr Gly
 145 150 155 160

Val Ile Leu Gln Arg Asp Asp Gly Val Leu Val Ser Asn Arg Leu Glu
 165 170 175

419
AS

Lys Pro Leu Pro Ile Val Asp Glu Val Leu Tyr Ile Asp Arg Ile Pro
180 185 190

Leu Gly Met Leu Ala Ala Ile Glu Val Ala Val Pro Gly Lys Val Ile
195 200 205

Glu Thr Leu Ser Asn Pro Tyr Gly Ile Ala Thr Val Phe Asn Leu Ser
210 215 220

Pro Glu Glu Thr Lys Asn Ile Val Pro Met Ala Arg Ala Leu Ile Gly
225 230 235 240

Asn Arg Ser Ala Val Val Val Lys Thr Pro Ser Gly Asp Val Lys Ala
245 250 255

Arg Ala Ile Pro Ala Gly Asn Leu Glu Leu Leu Ala Gln Gly Arg Ser
260 265 270

Val Arg Val Asp Val Ala Ala Gly Ala Glu Ala Ile Met Lys Ala Val
275 280 285

Asp Gly Cys Gly Arg Leu Asp Asn Val Thr Gly Glu Ser Gly Thr Asn
290 295 300

Ile Gly Gly Met Leu Glu His Val Arg Gln Thr Met Ala Glu Leu Thr
305 310 315 320

Asn Lys Pro Ser Ser Glu Ile Phe Ile Gln Asp Leu Leu Ala Val Asp
325 330 335

Thr Ser Val Pro Val Ser Val Thr Gly Gly Leu Ala Gly Glu Phe Ser
340 345 350

Leu Glu Gln Ala Val Gly Ile Ala Ser Met Val Lys Ser Asp Arg Leu
355 360 365

Gln Met Ala Met Ile Ala Arg Glu Ile Glu Gln Lys Leu Asn Ile Asp
370 375 380

Val Gln Ile Gly Gly Ala Glu Ala Glu Ala Ala Ile Leu Gly Ala Leu
385 390 395 400

Thr Thr Pro Gly Thr Thr Arg Pro Leu Ala Ile Leu Asp Leu Gly Ala
405 410 415

420
AA

Gly Ser Thr Asp Ala Ser Ile Ile Asn Pro Lys Gly Asp Ile Ile Ala
420 425 430

Thr His Leu Ala Gly Ala Gly Asp Met Val Thr Met Ile Ile Ala Arg
435 440 445

Glu Leu Gly Leu Glu Asp Arg Tyr Leu Ala Glu Glu Ile Lys Lys Tyr
450 455 460

Pro Leu Ala Lys Val Glu Ser Leu Phe His Leu Arg His Glu Asp Gly
465 470 475 480

Ser Val Gln Phe Phe Ser Thr Pro Leu Pro Pro Ala Val Phe Ala Arg
485 490 495

Val Cys Val Val Lys Ala Asp Glu Leu Val Pro Leu Pro Gly Asp Leu
500 505 510

Ala Leu Glu Lys Val Arg Ala Ile Arg Arg Ser Ala Lys Glu Arg Val
515 520 525

Phe Val Thr Asn Ala Leu Arg Ala Leu Arg Gln Val Ser Pro Thr Gly
530 535 540

Asn Ile Arg Asp Ile Pro Phe Val Val Leu Val Gly Gly Ser Ser Leu
545 550 555 560

Asp Phe Glu Val Pro Gln Leu Val Thr Asp Ala Leu Ala His Tyr Arg
565 570 575

Leu Val Ala Gly Arg Gly Asn Ile Arg Gly Ser Glu Gly Pro Arg Asn
580 585 590

Ala Val Ala Thr Gly Leu Ile Leu Ser Trp His Lys Glu Phe Ala His
595 600 605

Glu Arg
610

<210> 112

<211> 378

<212> DNA

<213> *Klebsiella oxytoca*

<400> 112

atgaacggta atcacagcgc cccggccatc gcgatcgccg tcatcgacgg ctgcgacggc

421
AS

ctgtggcgcg aagtgtgct gggatcgaa gaggaaggta tccctttccg gctccagcat 120
 caccggcgcg gagaggtcgt ggacagcgcc tggcaggcgg cgcgcagctc gccgctgctg 180
 gtgggcatcg cctgcgaccg ccatatgctg gtcgtgcact acaagaattt acccgcatcg 240
 gcgcccgttt ttacgtgat gcatcatcag gacagtcagg cccatcgcaa caccggtaat 300
 aacgcggcac ggctggtcaa ggggatccct ttccgggatc tgaatagcga agcaacagga 360
 gaacagcagg atgaataa 378

<210> 113
 <211> 125
 <212> PRT
 <213> *Klebsiella oxytoca*

<400> 113

Met Asn Gly Asn His Ser Ala Pro Ala Ile Ala Ile Ala Val Ile Asp
 1 5 10 15

Gly Cys Asp Gly Leu Trp Arg Glu Val Leu Leu Gly Ile Glu Glu Glu
 20 25 30

Gly Ile Pro Phe Arg Leu Gln His His Pro Ala Gly Glu Val Val Asp
 35 40 45

Ser Ala Trp Gln Ala Ala Arg Ser Ser Pro Leu Leu Val Gly Ile Ala
 50 55 60

Cys Asp Arg His Met Leu Val Val His Tyr Lys Asn Leu Pro Ala Ser
 65 70 75 80

Ala Pro Leu Phe Thr Leu Met His His Gln Asp Ser Gln Ala His Arg
 85 90 95

Asn Thr Gly Asn Asn Ala Ala Arg Leu Val Lys Gly Ile Pro Phe Arg
 100 105 110

Asp Leu Asn Ser Glu Ala Thr Gly Glu Gln Gln Asp Glu
 115 120 125

<210> 114
 <211> 1833
 <212> DNA
 <213> *Salmonella typhimurium*

<400> 114
 atgcgatata tagctggcat tgacatcggg aactcatcaa cggaagtcgc actggcgcg 60

caagatgaga ctggcgcaact gacgattaca cacagcgcg c tggcggaaac caccgggatac 120
 aaaggcacgt tgcgtaacgt gttcggcatt caggaagcgc tcgccctcgt cgcaaagcgc 180
 gcggggatca atgtcagaga tatttcgctc atccgcatta acgaagccac gccggtgatt 240
 ggcgatgtgg cgatggaaac cattaccgaa accatcatca ccgaatcgac aatgatcggc 300
 cataacccaa aaacgccggg cggagcaggc cttggtgtgg gtatcacgat tacgccggag 360
 gagctgttaa cccgcccggc ggactcgtcc tatattctgg tggatcgtc agcctttgat 420
 ttgctgata tcgccaatgt tatcaacgcc tcaatgcgcg ccggatacca gattaccggc 480
 gtcattttgc agcgcgacga tggcgactg gtcagcaacc ggctggaaaa atcgtaccg 540
 attgtcgatg aagttctgta catcgaccgc attccgctgg ggatgctggc ggcgattgaa 600
 gtcgccgtgc cgggaaaggt tatcgaaacc ctctctaacc cttacggcat cgccaccgta 660
 tttaatctca acgccgatga gacaaaaaac atcgtcccga tggcgcgcg cgtgattggc 720
 aaccgttccg ccgtggtggg taaaacgcc tccggcgacg tcaaagcgcg cgcaataccc 780
 gccggttaacc tggagctgca ggctcagggt cgtaccgtgc gcgtggatgt tgccgccggg 840
 gccgaagcca tcatgaaagc ggtggacggg tgcggcaagc tcgacaacgt caccggcgag 900
 gccgggacca atatcggcgg catgctggag cacgtgcgcc agaccatggc cgaactgacc 960
 aacaagccga gcagtgagat tttcattcag gatctactgg ccgttgacac ctcggttccg 1020
 gtgagcgtca ccggcgggtct ggccggggag ttctcgtgg agcaggccgt cggcatcgcc 1080
 tcgatggtga aatcagaccg tctgcagatg gcgatgattg cccgtgaaat tgagcagaag 1140
 cttaatatcg acgtgcagat cggcggcgt gaggtgaag ccgccattct gggcgcgctg 1200
 accacgccgg gtaccacccg accgctggcg atcctcgacc tcggcgcggg ctccaccgat 1260
 gcctccatca tcaaccctaa aggcgaaatc atcgccaccc atctcgccgg ggcaggcgac 1320
 atggtcacga tgattattgc ccgcgaactg gggctggaag accgctatct ggcggaagag 1380
 atcaaaaaat acccgctggc taaggctgaa agcctgttcc acttacgcca cgaggacggc 1440
 agcgtccagt tcttcccgac gccgctgcct cccgccgtgt tcgcccgcgt ctgctggtg 1500
 aaaccggacg aactggtgcc gcttcccggc gacttagcgc tggaaaaagt gcgcgccatt 1560
 cgccgcagcg ctaaagaacg cgtctttgtc accaacgccc tgcgcgcgt gcgccaggtc 1620
 agtccaaccg gcaacattcg cgatattccg ttcgtggtgc tggctggcgg ctcgtcgctg 1680
 gatttcgaag ttccgcagct ggtcaccgat gcgctggcgc actaccgcct ggtcgccggg 1740
 cgaggaaata ttcgcggcag cgaaggccca agaaacgcgg tggccaccgg cctgattctc 1800
 tcctggcata aggagtttgc gcatggacag taa 1833

423
P

<210> 115
 <211> 610
 <212> PRT
 <213> *Salmonella typhimurium*

<400> 115

Met Arg Tyr Ile Ala Gly Ile Asp Ile Gly Asn Ser Ser Thr Glu Val
 1 5 10 15

Ala Leu Ala Arg Gln Asp Glu Thr Gly Ala Leu Thr Ile Thr His Ser
 20 25 30

Ala Leu Ala Glu Thr Thr Gly Ile Lys Gly Thr Leu Arg Asn Val Phe
 35 40 45

Gly Ile Gln Glu Ala Leu Ala Leu Val Ala Lys Arg Ala Gly Ile Asn
 50 55 60

Val Arg Asp Ile Ser Leu Ile Arg Ile Asn Glu Ala Thr Pro Val Ile
 65 70 75 80

Gly Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser
 85 90 95

Thr Met Ile Gly His Asn Pro Lys Thr Pro Gly Gly Ala Gly Leu Gly
 100 105 110

Val Gly Ile Thr Ile Thr Pro Glu Glu Leu Leu Thr Arg Pro Ala Asp
 115 120 125

Ser Ser Tyr Ile Leu Val Val Ser Ser Ala Phe Asp Phe Ala Asp Ile
 130 135 140

Ala Asn Val Ile Asn Ala Ser Met Arg Ala Gly Tyr Gln Ile Thr Gly
 145 150 155 160

Val Ile Leu Gln Arg Asp Asp Gly Val Leu Val Ser Asn Arg Leu Glu
 165 170 175

Lys Ser Leu Pro Ile Val Asp Glu Val Leu Tyr Ile Asp Arg Ile Pro
 180 185 190

Leu Gly Met Leu Ala Ala Ile Glu Val Ala Val Pro Gly Lys Val Ile
 195 200 205

Glu Thr Leu Ser Asn Pro Tyr Gly Ile Ala Thr Val Phe Asn Leu Asn
 210 215 220

424
NA

Ala Asp Glu Thr Lys Asn Ile Val Pro Met Ala Arg Ala Leu Ile Gly
225 230 235 240

Asn Arg Ser Ala Val Val Val Lys Thr Pro Ser Gly Asp Val Lys Ala
245 250 255

Arg Ala Ile Pro Ala Gly Asn Leu Glu Leu Gln Ala Gln Gly Arg Thr
260 265 270

Val Arg Val Asp Val Ala Ala Gly Ala Glu Ala Ile Met Lys Ala Val
275 280 285

Asp Gly Cys Gly Lys Leu Asp Asn Val Thr Gly Glu Ala Gly Thr Asn
290 295 300

Ile Gly Gly Met Leu Glu His Val Arg Gln Thr Met Ala Glu Leu Thr
305 310 315 320

Asn Lys Pro Ser Ser Glu Ile Phe Ile Gln Asp Leu Leu Ala Val Asp
325 330 335

Thr Ser Val Pro Val Ser Val Thr Gly Gly Leu Ala Gly Glu Phe Ser
340 345 350

Leu Glu Gln Ala Val Gly Ile Ala Ser Met Val Lys Ser Asp Arg Leu
355 360 365

Gln Met Ala Met Ile Ala Arg Glu Ile Glu Gln Lys Leu Asn Ile Asp
370 375 380

Val Gln Ile Gly Gly Ala Glu Ala Glu Ala Ala Ile Leu Gly Ala Leu
385 390 395 400

Thr Thr Pro Gly Thr Thr Arg Pro Leu Ala Ile Leu Asp Leu Gly Ala
405 410 415

Gly Ser Thr Asp Ala Ser Ile Ile Asn Pro Lys Gly Glu Ile Ile Ala
420 425 430

Thr His Leu Ala Gly Ala Gly Asp Met Val Thr Met Ile Ile Ala Arg
435 440 445

Glu Leu Gly Leu Glu Asp Arg Tyr Leu Ala Glu Glu Ile Lys Lys Tyr
450 455 460

425

Pro Leu Ala Lys Val Glu Ser Leu Phe His Leu Arg His Glu Asp Gly
465 470 475 480

Ser Val Gln Phe Phe Pro Thr Pro Leu Pro Pro Ala Val Phe Ala Arg
485 490 495

Val Cys Val Val Lys Pro Asp Glu Leu Val Pro Leu Pro Gly Asp Leu
500 505 510

Ala Leu Glu Lys Val Arg Ala Ile Arg Arg Ser Ala Lys Glu Arg Val
515 520 525

Phe Val Thr Asn Ala Leu Arg Ala Leu Arg Gln Val Ser Pro Thr Gly
530 535 540

Asn Ile Arg Asp Ile Pro Phe Val Val Leu Val Gly Gly Ser Ser Leu
545 550 555 560

Asp Phe Glu Val Pro Gln Leu Val Thr Asp Ala Leu Ala His Tyr Arg
565 570 575

Leu Val Ala Gly Arg Gly Asn Ile Arg Gly Ser Glu Gly Pro Arg Asn
580 585 590

Ala Val Ala Thr Gly Leu Ile Leu Ser Trp His Lys Glu Phe Ala His
595 600 605

Gly Gln
610

<210> 116
<211> 372
<212> DNA
<213> *Salmonella typhimurium*

<400> 116
atggacagta atcacagcgc cccggctatc gtcattaccg ttatcaacga ctgcgccagc 60
ctctggcacg aagtgtgtgct gggcattgaa gaggaaggca tccctttcct gcttcagcat 120
caccgggctg gagatatcgt tgacagcgcc tggcaggcgg cgcgagctc gccgctgctg 180
gtcggcattg cctgcgatcg acactcgtcg gtcgtgcatt acaagaattt acccgcatcg 240
gcgccgcttt ttacgctgat gcatcatcag gacagtcagg cccaacgcaa caccggtaat 300
aacgcggcac ggctggtcaa agggatccct ttcgggatct ccatgcttaa tcacaggaga 360
acggcagtat ga 372

<210> 117
 <211> 123
 <212> PRT
 <213> *Salmonella typhimurium*

<400> 117

Met Asp Ser Asn His Ser Ala Pro Ala Ile Val Ile Thr Val Ile Asn
 1 5 10 15

Asp Cys Ala Ser Leu Trp His Glu Val Leu Leu Gly Ile Glu Glu Glu
 20 25 30

Gly Ile Pro Phe Leu Leu Gln His His Pro Ala Gly Asp Ile Val Asp
 35 40 45

Ser Ala Trp Gln Ala Ala Arg Ser Ser Pro Leu Leu Val Gly Ile Ala
 50 55 60

Cys Asp Arg His Ser Leu Val Val His Tyr Lys Asn Leu Pro Ala Ser
 65 70 75 80

Ala Pro Leu Phe Thr Leu Met His His Gln Asp Ser Gln Ala Gln Arg
 85 90 95

Asn Thr Gly Asn Asn Ala Ala Arg Leu Val Lys Gly Ile Pro Phe Gly
 100 105 110

Ile Ser Met Leu Asn His Arg Arg Thr Ala Val
 115 120

<210> 118
 <211> 1833
 <212> DNA
 <213> *Lactobacillus collinoides*

<400> 118
 atgacacgtg taattggtgt tgatatcggg aattcctcta cagaagttgc gcttgctgat 60
 gtgtctgaca gtggtgaagt aaatttcatt aattctggaa tttccgatac aactggcatt 120
 aaaggacta aacaaaattt gatcgggggtg cgtaaatacca tccagatcgt tttgaaaaag 180
 tcgaatatgc aaatttccga tgttgacctg attcggatca acgaagcaac gcccggttatc 240
 ggtgatgttg ccatggagac catcaccgaa acggtgatta ctgaatcgac gatgatcggc 300
 cacaaccag ggactcctgg ggggtgtcggg actgggttctg gttacacggg gaatttgctt 360
 gatttggtga gccaaacgga taaggatcgt ccttatatcg ttatcatctc gaaagaaatc 420
 gattttgctg acgcagctaa gctgatcaac gcttatgtgg cttctgggta taatattacc 480

gctgccattc tgcaaagtga tgatgggggtg ctgatcaata atcggttgac ccataagatt 540
 cccatcgtag atgaagtctc acagatcgac aagggtaccgt tgaacatgct tgccgcagtg 600
 gaagttgcac cgcctggcaa agtaattgct caactttcca acccgtagtg cattgccaca 660
 ctgttcgaac tttcctctga agaaaccaag aacattgtgc cagttgcccg agccttaatc 720
 ggaaaccggt cagcgggttggt tattaaaacc cctgccgggtg atgttaaagc tcgtgttata 780
 ccagccggga aaatcttgat caatggccaa ccgaatgggtc atgggtgaagt taacgttgcg 840
 gctggtgccg atgccatcat gaaaaagggtg aacgagttcg atagtgtcga tgacattacc 900
 ggtgaatcgg gcactaacgt tgggtgggatg cttgaaaaag ttcgtcaaac aatggctgag 960
 ttgaccgaca agcaaaatag cgacattgcc attcaagatt tattagctgt caatacgctc 1020
 gttccagtaa cgggtgcgtg tgggtctgggt ggtgaattct caatggaaca agccgttggtg 1080
 attgctgcta tgggtcaaatac tgatcacttg caaatgcaag cgattgcaga cctgatgaaa 1140
 gatgaatttc acgttcaagt cgaaatcggc ggtgctgaag ctgaatcagc catcctcgggt 1200
 gcgctaacia cgccagggtg gacaaaacca attgccatcc ttgatttggg ggctgggttca 1260
 acggatgcat caattatcaa ccaaaaggac gaaaagggtc ctattcactt ggctgggtgcc 1320
 ggtgatattg ttacatgat catcaattct gaacttgggt tgggaagacc atatttagct 1380
 gaggatatta agaaatatcc gctggctaaa gttgataatc tattccagct acggcatgaa 1440
 gatggtgccg ttcaattctt tgaagatcca ttacctgctg atttatttgc cagagttgtg 1500
 gctgttaaac cagatgggtta cgaaccactt cctggtaatt tgagtatcga gaaagttaaa 1560
 atcgtccgtc aaactgctaa gaagcgggtg ttcgtaacga acgcaattcg tgccttacac 1620
 cacgttagcc caacaggtaa tatccgagat atcccatctg tggtcattgt cggcgggtca 1680
 gccctcgatt ttgaaattcc acaattgggt accgatgaat tatcacactt taacttagtt 1740
 gcaggtcgtg gtaatatctg gggaattgaa ggtccacgga acgccgtggc aactgggttg 1800
 attctttcat acgcgagtga gaagagggga tag 1833

<210> 119
 <211> 610
 <212> PRT
 <213> *Lactobacillus collinoides*

<400> 119

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

Ala Leu Ala Asp Val Ser Asp Ser Gly Glu Val Asn Phe Ile Asn Ser
20 25 30

Gly Ile Ser Asp Thr Thr Gly Ile Lys Gly Thr Lys Gln Asn Leu Ile
35 40 45

Gly Val Arg Lys Ser Ile Gln Ile Val Leu Lys Lys Ser Asn Met Gln
50 55 60

Ile Ser Asp Val Asp Leu Ile Arg Ile Asn Glu Ala Thr Pro Val Ile
65 70 75 80

Gly Asp Val Ala Met Glu Thr Ile Thr Glu Thr Val Ile Thr Glu Ser
85 90 95

Thr Met Ile Gly His Asn Pro Gly Thr Pro Gly Gly Val Gly Thr Gly
100 105 110

Ser Gly Tyr Thr Val Asn Leu Leu Asp Leu Leu Ser Gln Thr Asp Lys
115 120 125

Asp Arg Pro Tyr Ile Val Ile Ile Ser Lys Glu Ile Asp Phe Ala Asp
130 135 140

Ala Ala Lys Leu Ile Asn Ala Tyr Val Ala Ser Gly Tyr Asn Ile Thr
145 150 155 160

Ala Ala Ile Leu Gln Ser Asp Asp Gly Val Leu Ile Asn Asn Arg Leu
165 170 175

Thr His Lys Ile Pro Ile Val Asp Glu Val Ser Gln Ile Asp Lys Val
180 185 190

Pro Leu Asn Met Leu Ala Ala Val Glu Val Ala Pro Pro Gly Lys Val
195 200 205

Ile Ala Gln Leu Ser Asn Pro Tyr Gly Ile Ala Thr Leu Phe Glu Leu
210 215 220

Ser Ser Glu Glu Thr Lys Asn Ile Val Pro Val Ala Arg Ala Leu Ile
225 230 235 240

Gly Asn Arg Ser Ala Val Val Ile Lys Thr Pro Ala Gly Asp Val Lys
245 250 255

Ala Arg Val Ile Pro Ala Gly Lys Ile Leu Ile Asn Gly Gln Pro Asn
260 265 270

429
[Signature]

Gly His Gly Glu Val Asn Val Ala Ala Gly Ala Asp Ala Ile Met Lys
275 280 285

Lys Val Asn Glu Phe Asp Ser Val Asp Asp Ile Thr Gly Glu Ser Gly
290 295 300

Thr Asn Val Gly Gly Met Leu Glu Lys Val Arg Gln Thr Met Ala Glu
305 310 315 320

Leu Thr Asp Lys Gln Asn Ser Asp Ile Ala Ile Gln Asp Leu Leu Ala
325 330 335

Val Asn Thr Ser Val Pro Val Thr Val Arg Gly Gly Leu Ala Gly Glu
340 345 350

Phe Ser Met Glu Gln Ala Val Gly Ile Ala Ala Met Val Lys Ser Asp
355 360 365

His Leu Gln Met Gln Ala Ile Ala Asp Leu Met Lys Asp Glu Phe His
370 375 380

Val Gln Val Glu Ile Gly Gly Ala Glu Ala Glu Ser Ala Ile Leu Gly
385 390 395 400

Ala Leu Thr Thr Pro Gly Thr Thr Lys Pro Ile Ala Ile Leu Asp Leu
405 410 415

Gly Ala Gly Ser Thr Asp Ala Ser Ile Ile Asn Gln Lys Asp Glu Lys
420 425 430

Val Ala Ile His Leu Ala Gly Ala Gly Asp Met Val Thr Met Ile Ile
435 440 445

Asn Ser Glu Leu Gly Leu Glu Asp Pro Tyr Leu Ala Glu Asp Ile Lys
450 455 460

Lys Tyr Pro Leu Ala Lys Val Asp Asn Leu Phe Gln Leu Arg His Glu
465 470 475 480

Asp Gly Ala Val Gln Phe Phe Glu Asp Pro Leu Pro Ala Asp Leu Phe
485 490 495

Ala Arg Val Val Ala Val Lys Pro Asp Gly Tyr Glu Pro Leu Pro Gly
500 505 510

430
A

Asn Leu Ser Ile Glu Lys Val Lys Ile Val Arg Gln Thr Ala Lys Lys
515 520 525

Arg Val Phe Val Thr Asn Ala Ile Arg Ala Leu His His Val Ser Pro
530 535 540

Thr Gly Asn Ile Arg Asp Ile Pro Phe Val Val Ile Val Gly Gly Ser
545 550 555 560

Ala Leu Asp Phe Glu Ile Pro Gln Leu Val Thr Asp Glu Leu Ser His
565 570 575

Phe Asn Leu Val Ala Gly Arg Gly Asn Ile Arg Gly Ile Glu Gly Pro
580 585 590

Arg Asn Ala Val Ala Thr Gly Leu Ile Leu Ser Tyr Ala Ser Glu Lys
595 600 605

Arg Gly
610

<210> 120

<211> 351

<212> DNA

<213> *Lactobacillus collinoides*

<400> 120

atggcatttg attctgaacg tccgtcaatt ctattggcga caccaacggg ttctaattggc 60
caacttccag aagttctaaa accaatgctc aatgggtattg aagaagaaca gattcctttt 120
cagattctcg atatggaagg cggttcagca gttgagcggg cttataacgc gtcagttgct 180
tcacgattat cagtgggcgt tgggtttgat gatgcacata tcattgtgca ttataaaaac 240
ttgaaaccag aaaaaccgct gtttgatggt gccatcactg atgcagcatc cattcgtaaa 300
gttggcgcaa acgccgctcg acttgtaaag ggagttccat tcaagaagta a 351

<210> 121

<211> 116

<212> PRT

<213> *Lactobacillus collinoides*

<400> 121

Met Ala Phe Asp Ser Glu Arg Pro Ser Ile Leu Leu Ala Thr Pro Thr
1 5 10 15

Gly Ser Asn Gly Gln Leu Pro Glu Val Leu Lys Pro Met Leu Asn Gly
20 25 30

431
A

Ile Glu Glu Glu Gln Ile Pro Phe Gln Ile Leu Asp Met Glu Gly Gly
35 40 45

Ser Ala Val Glu Arg Ala Tyr Asn Ala Ser Val Ala Ser Arg Leu Ser
50 55 60

Val Gly Val Gly Phe Asp Asp Ala His Ile Ile Val His Tyr Lys Asn
65 70 75 80

Leu Lys Pro Glu Lys Pro Leu Phe Asp Val Ala Ile Thr Asp Ala Ala
85 90 95

Ser Ile Arg Lys Val Gly Ala Asn Ala Ala Arg Leu Val Lys Gly Val
100 105 110

Pro Phe Lys Lys
115

<210> 122
<211> 453
<212> PRT
<213> *Vibrio fluvialis*

<400> 122

Met Asn Lys Pro Gln Ser Trp Glu Ala Arg Ala Glu Thr Tyr Ser Leu
1 5 10 15

Tyr Gly Phe Thr Asp Met Pro Ser Leu His Gln Arg Gly Thr Val Val
20 25 30

Val Thr His Gly Glu Gly Pro Tyr Ile Val Asp Val Asn Gly Arg Arg
35 40 45

Tyr Leu Asp Ala Asn Ser Gly Leu Trp Asn Met Val Ala Gly Phe Asp
50 55 60

His Lys Gly Leu Ile Asp Ala Ala Lys Ala Gln Tyr Glu Arg Phe Pro
65 70 75 80

Gly Tyr His Ala Phe Phe Gly Arg Met Ser Asp Gln Thr Val Met Leu
85 90 95

Ser Glu Lys Leu Val Glu Val Ser Pro Phe Asp Ser Gly Arg Val Phe
100 105 110

Tyr Thr Asn Ser Gly Ser Glu Ala Asn Asp Thr Met Val Lys Met Leu
115 120 125

Trp Phe Leu His Ala Ala Glu Gly Lys Pro Gln Lys Arg Lys Ile Leu
130 135 140

Thr Arg Trp Asn Ala Tyr His Gly Val Thr Ala Val Ser Ala Ser Met
145 150 155 160

Thr Gly Lys Pro Tyr Asn Ser Val Phe Gly Leu Pro Leu Pro Gly Phe
165 170 175

Val His Leu Thr Cys Pro His Tyr Trp Arg Tyr Gly Glu Glu Gly Glu
180 185 190

Thr Glu Glu Gln Phe Val Ala Arg Leu Ala Arg Glu Leu Glu Glu Thr
195 200 205

Ile Gln Arg Glu Gly Ala Asp Thr Ile Ala Gly Phe Phe Ala Glu Pro
210 215 220

Val Met Gly Ala Gly Gly Val Ile Pro Pro Ala Lys Gly Tyr Phe Gln
225 230 235 240

Ala Ile Leu Pro Ile Leu Arg Lys Tyr Asp Ile Pro Val Ile Ser Asp
245 250 255

Glu Val Ile Cys Gly Phe Gly Arg Thr Gly Asn Thr Trp Gly Cys Val
260 265 270

Thr Tyr Asp Phe Thr Pro Asp Ala Ile Ile Ser Ser Lys Asn Leu Thr
275 280 285

Ala Gly Phe Phe Pro Met Gly Ala Val Ile Leu Gly Pro Glu Leu Ser
290 295 300

Lys Arg Leu Glu Thr Ala Ile Glu Ala Ile Glu Glu Phe Pro His Gly
305 310 315 320

Phe Thr Ala Ser Gly His Pro Val Gly Cys Ala Ile Ala Leu Lys Ala
325 330 335

Ile Asp Val Val Met Asn Glu Gly Leu Ala Glu Asn Val Arg Arg Leu
340 345 350

Ala Pro Arg Phe Glu Glu Arg Leu Lys His Ile Ala Glu Arg Pro Asn
355 360 365

433

Ile Gly Glu Tyr Arg Gly Ile Gly Phe Met Trp Ala Leu Glu Ala Val
370 375 380

Lys Asp Lys Ala Ser Lys Thr Pro Phe Asp Gly Asn Leu Ser Val Ser
385 390 395 400

Glu Arg Ile Ala Asn Thr Cys Thr Asp Leu Gly Leu Ile Cys Arg Pro
405 410 415

Leu Gly Gln Ser Val Val Leu Cys Pro Pro Phe Ile Leu Thr Glu Ala
420 425 430

Gln Met Asp Glu Met Phe Asp Lys Leu Glu Lys Ala Leu Asp Lys Val
435 440 445

Phe Ala Glu Val Ala
450

<210> 123
<211> 1122

<212> DNA

<213> *Erwinia caratovora subespécie atroseptica*

<400> 123

atgtctgacg gacgactcac cgcacttttt cctgcattcc cacacccggc gtccaatcag	60
cccgtatttg ccgaggcttc accgcacgac gacgagttaa tgacgcaggc cgtaccgcag	120
gtttcctgtc agcaggcggtt ggcgattgag cagcaagaat atggcttgtc tgggcagatg	180
tcgctgcttc agggcgagcg tgatgtgaat ttctgtctga cgggtgacgcc agatgaacgc	240
tacatgctga aagtcacaa tgcggcagaa cctgccgacg tcagcaattt ccaaaccgcg	300
ctgctgctgc atcttgcccg tcaggcacct gaactgcccg taccgcgtat caggctgaca	360
aaagcgggtc agtcggaaac aggcgttgag atcgatggtg tactgctgag tgtgcggctt	420
gtgagctatc tggcaggaat gccgcagtat ctggcctcac cgtcaacggc gctgatgccg	480
cagttggggg gaacgctggc gcagttggat aacgcgcttc acagctttac gcatccggcg	540
gcaaaccgtg cgctgctgtg ggatatcagc cgggcagagc aggtgcgtcc ttacctcgat	600
ttcgtttctg aaccgcagca gtatcagcat cttcagcgta tttttgaccg ttatgacagt	660
aacgttgctc ctctgttgac gacgctacgt cgtcaggica ttcataacga tctgaatccg	720
cataacgtgc tgggtggatgg atcgtcgccg acgcggtta ctggcattat cgattttggc	780
gatgccgtat ttgccccgtt aatttgcgaa gtcgcgacgg cactggcgta tcagatcggc	840
gatggaaccg atttgttggg gcatgttgtg ccgtttgttg cggcctatca ccaacgcatt	900

ccgtagcac cggaggagat tgcgctgtta cccgatctga tagcgacccg tatggcgctg 960
 accctgacca ttgcgcagtg gcgagcatcg cgttatcccg acaatcgga gtatctgctg 1020
 cgtaacgtgc cgcgctgttg gcacagtttg cagcgattg cgacctattc ccatgcgcaa 1080
 tttttgactc gcctacagca ggtttgcccg gagaatgcgc ga 1122

<210> 124

<211> 374

<212> PRT

<213> *Erwinia caratovora subespécie atroseptica*

<400> 124

Met Ser Asp Gly Arg Leu Thr Ala Leu Phe Pro Ala Phe Pro His Pro
 1 5 10 15

Ala Ser Asn Gln Pro Val Phe Ala Glu Ala Ser Pro His Asp Asp Glu
 20 25 30

Leu Met Thr Gln Ala Val Pro Gln Val Ser Cys Gln Gln Ala Leu Ala
 35 40 45

Ile Ala Gln Gln Glu Tyr Gly Leu Ser Gly Gln Met Ser Leu Leu Gln
 50 55 60

Gly Glu Arg Asp Val Asn Phe Cys Leu Thr Val Thr Pro Asp Glu Arg
 65 70 75 80

Tyr Met Leu Lys Val Ile Asn Ala Ala Glu Pro Ala Asp Val Ser Asn
 85 90 95

Phe Gln Thr Ala Leu Leu Leu His Leu Ala Arg Gln Ala Pro Glu Leu
 100 105 110

Pro Val Pro Arg Ile Arg Ser Thr Lys Ala Gly Gln Ser Glu Thr Gly
 115 120 125

Val Glu Ile Asp Gly Val Leu Leu Arg Val Arg Leu Val Ser Tyr Leu
 130 135 140

Ala Gly Met Pro Gln Tyr Leu Ala Ser Pro Ser Thr Ala Leu Met Pro
 145 150 155 160

Gln Leu Gly Gly Thr Leu Ala Gln Leu Asp Asn Ala Leu His Ser Phe
 165 170 175

435

Thr His Pro Ala Ala Asn Arg Ala Leu Leu Trp Asp Ile Ser Arg Ala
180 185 190

Glu Gln Val Arg Pro Tyr Leu Asp Phe Val Ser Glu Pro Gln Gln Tyr
195 200 205

Gln His Leu Gln Arg Ile Phe Asp Arg Tyr Asp Ser Asn Val Ala Pro
210 215 220

Leu Leu Thr Thr Leu Arg Arg Gln Val Ile His Asn Asp Leu Asn Pro
225 230 235 240

His Asn Val Leu Val Asp Gly Ser Ser Pro Thr Arg Val Thr Gly Ile
245 250 255

Ile Asp Phe Gly Asp Ala Val Phe Ala Pro Leu Ile Cys Glu Val Ala
260 265 270

Thr Ala Leu Ala Tyr Gln Ile Gly Asp Gly Thr Asp Leu Leu Glu His
275 280 285

Val Val Pro Phe Val Ala Ala Tyr His Gln Arg Ile Pro Leu Ala Pro
290 295 300

Glu Glu Ile Ala Leu Leu Pro Asp Leu Ile Ala Thr Arg Met Ala Leu
305 310 315 320

Thr Leu Thr Ile Ala Gln Trp Arg Ala Ser Arg Tyr Pro Asp Asn Arg
325 330 335

Glu Tyr Leu Leu Arg Asn Val Pro Arg Cys Trp His Ser Leu Gln Arg
340 345 350

Ile Ala Thr Tyr Ser His Ala Gln Phe Leu Thr Arg Leu Gln Gln Val
355 360 365

Cys Pro Glu Asn Ala Arg
370

<210> 125

<211> 1272

<212> DNA

<213> *Erwinia caratovora subespécie atroseptica*

<400> 125

atgacagcga cagaagcttt gctggcgcgc cgtcagcgag tggtgggagg cggttatcgc

ctgttttatg aagagccgct gcatgtcgcg cgcggcgagg gcgtgtggct gttcgatcac 120
 caagggaac gttatctgga tgtctacaat aatgtggctt cggtcggaca ttgccacccc 180
 gcggtggttg aagccgtggc gcgacagagc gcacaactca ataccacac gcgctatttg 240
 caccacgcga ttgtcgattt tgcggaagat ttgctgagcg aatttcccgc cgaattgaac 300
 aatgtaatgc tgacctgtac cggcagttag gctaacgatc tggcgctgcg tatcgcccga 360
 catgtcacgg gcgggacggg gatgttgggt acgcgctggg cgtatcacgg cgtgaccagc 420
 gcgctggcgg aactgtctcc gtcgctgggg gatggcggtt tgcgcggtag ccatgtgaag 480
 ctgatcgacg cgccagacac ttatcgtcag cccggtgcat ttcttaccag cattcgtgaa 540
 gcgctggcgc agatgcaacg ggaaggattt cgtcctgcgg cgctgctggt agataccatt 600
 ttttccagcg atggcgtgtt ctgtgcgccg gaaggcgaaa tggcacaggc ggcggcggtt 660
 atccgtcagg cgggcgggct gtttattgcy gatgaagtgc agccgggctt cgggcgcacc 720
 ggggaatcac tgtggggctt tgcgcgccac aatgtcgtcc ctgatttggg gagtctaggg 780
 aaaccgatgg gcaacggaca tcccatcgtt ggattggtgg ggcgttccgc tctgttcgac 840
 gcatttgggc gcgatgtgcg ctatttcaat accttggcg gcaatccggt ttctgtcag 900
 gcggcgacg cgggtgctgc ggtgattcgg gaagagcagt tgcagcagaa tgcccagcgg 960
 gtcggtgatt atctgcggca agggttgcag caactggcgc agcatttccc gctgattggg 1020
 gatattcggg cttacggcct gtttattggg gcggagctgg tcagcgatcg cgaaagtaaa 1080
 acgccggcaa gtgaatccgc gttgcagggt gtgaatgcga tgcgccaacg tgggtgtgctc 1140
 atcagcgca cggggccagc ggcgaacata ctgaaaattc gcccgcgct ggtgtttctg 1200
 gaagaacacg ccgatgtgtt cttaccacg ctgagtgacg ttttagcgct catcggcact 1260
 cgtgcacaga ga 1272

<210> 126

<211> 424

<212> PRT

<213> *Erwinia caratovora subspécie atroseptica*

<400> 126

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

Gly Gly Tyr Arg Leu Phe Tyr Glu Glu Pro Leu His Val Ala Arg Gly
20 25 30

Glu Gly Val Trp Leu Phe Asp His Gln Gly Lys Arg Tyr Leu Asp Val
35 40 45

437
A

Tyr Asn Asn Val Ala Ser Val Gly His Cys His Pro Ala Val Val Glu
50 55 60

Ala Val Ala Arg Gln Ser Ala Gln Leu Asn Thr His Thr Arg Tyr Leu
65 70 75 80

His His Ala Ile Val Asp Phe Ala Glu Asp Leu Leu Ser Glu Phe Pro
85 90 95

Ala Glu Leu Asn Asn Val Met Leu Thr Cys Thr Gly Ser Glu Ala Asn
100 105 110

Asp Leu Ala Leu Arg Ile Ala Arg His Val Thr Gly Gly Thr Gly Met
115 120 125

Leu Val Thr Arg Trp Ala Tyr His Gly Val Thr Ser Ala Leu Ala Glu
130 135 140

Leu Ser Pro Ser Leu Gly Asp Gly Val Val Arg Gly Ser His Val Lys
145 150 155 160

Leu Ile Asp Ala Pro Asp Thr Tyr Arg Gln Pro Gly Ala Phe Leu Thr
165 170 175

Ser Ile Arg Glu Ala Leu Ala Gln Met Gln Arg Glu Gly Ile Arg Pro
180 185 190

Ala Ala Leu Leu Val Asp Thr Ile Phe Ser Ser Asp Gly Val Phe Cys
195 200 205

Ala Pro Glu Gly Glu Met Ala Gln Ala Ala Ala Leu Ile Arg Gln Ala
210 215 220

Gly Gly Leu Phe Ile Ala Asp Glu Val Gln Pro Gly Phe Gly Arg Thr
225 230 235 240

Gly Glu Ser Leu Trp Gly Phe Ala Arg His Asn Val Val Pro Asp Leu
245 250 255

Val Ser Leu Gly Lys Pro Met Gly Asn Gly His Pro Ile Ala Gly Leu
260 265 270

Val Gly Arg Ser Ala Leu Phe Asp Ala Phe Gly Arg Asp Val Arg Tyr
275 280 285

438
H

Phe Asn Thr Phe Gly Gly Asn Pro Val Ser Cys Gln Ala Ala His Ala
290 295 300

Val Leu Arg Val Ile Arg Glu Glu Gln Leu Gln Gln Asn Ala Gln Arg
305 310 315 320

Val Gly Asp Tyr Leu Arg Gln Gly Leu Gln Gln Leu Ala Gln His Phe
325 330 335

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

Leu Val Ser Asp Arg Glu Ser Lys Thr Pro Ala Ser Glu Ser Ala Leu
355 360 365

Gln Val Val Asn Ala Met Arg Gln Arg Gly Val Leu Ile Ser Ala Thr
370 375 380

Gly Pro Ala Ala Asn Ile Leu Lys Ile Arg Pro Pro Leu Val Phe Leu
385 390 395 400

Glu Glu His Ala Asp Val Phe Leu Thr Thr Leu Ser Asp Val Leu Ala
405 410 415

Leu Ile Gly Thr Arg Ala Gln Arg
420

<210> 127
<211> 35
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 127
ctccggaatt catgtctgac ggacgactca ccgca

35

<210> 128
<211> 46
<212> DNA
<213> Sequência artificial

<220>
<223> Iniciador

<400> 128
ttccaatgca ttggctgcag ttatctctgt gcacgagtgc cgatga

46

439
A

<210> 129
 <211> 40
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 129
 aacagccaag cttggctgca gtcacgcgc attctccggg 40

<210> 130
 <211> 40
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 130
 tctccggaat tcatgacgtc tgaaatgaca gcgacagaag 40

<210> 131
 <211> 33
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 131
 gctaacagga ggaagaattc atgggggggtt ctc 33

<210> 132
 <211> 33
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 132
 gagaaccccc catgaattct tcctcctggt agc 33

<210> 133
 <211> 723
 <212> DNA
 <213> *Klebsiella terrigena*

<400> 133
 atgcaaaaag tcgcacttgt caccggcgcc ggtcagggca tcggtaaagc tatcgccctg 60

cgtctggtga aggatggatt tgccgtggca atcgccgatt acaacgacgc tacggccaca 120

gcggtagccg ctgaaatcaa ccaggccggc ggccgcgcgg tggccattaa ggtcgacgtc 180

440
AS

tcgcgccggg accaggtttt cgccgccgtt gagcaggcgc gtaaagccct gggcggattc 240
aacgttatcg tcaacaacgc cggcatcgcg ccgtcaacgc cgatcgagtc catcaccgag 300
gagatcgtcg accgggtcta taacatcaac gttaagggcg tcatctgggg gatgcaggcg 360
gcggtggagg ccttcaaaaa agagggggcac ggcgggaaga tcgtcaacgc ctgctcccag 420
gccggccacg tcggcaaccc ggagctggcg gtctacagtt cgagtaaatt cgccgtgcgc 480
ggcctgacgc aaaccgccgc ccgcgatctg gcgccgctgg gcatcaccgt taacggcttc 540
tgcccaggga tcgttaagac gccaatgtgg gcggagattg accgtcagtg tcggaagcgg 600
cgggcaaacc gctgggctac ggcacggctg aatttgccaa acgcatcacc cttggccgcc 660
tgtcggagcc tgaagacgtc gccgcctgcg tgtcgttcct cgccagcccg gattccgact 720
ata 723

<210> 134
<211> 241
<212> PRT
<213> *Klebsiella terrigena*

<400> 134

Met Gln Lys Val Ala Leu Val Thr Gly Ala Gly Gln Gly Ile Gly Lys
1 5 10 15

Ala Ile Ala Leu Arg Leu Val Lys Asp Gly Phe Ala Val Ala Ile Ala
20 25 30

Asp Tyr Asn Asp Ala Thr Ala Thr Ala Val Ala Ala Glu Ile Asn Gln
35 40 45

Ala Gly Gly Arg Ala Val Ala Ile Lys Val Asp Val Ser Arg Arg Asp
50 55 60

Gln Val Phe Ala Ala Val Glu Gln Ala Arg Lys Ala Leu Gly Gly Phe
65 70 75 80

Asn Val Ile Val Asn Asn Ala Gly Ile Ala Pro Ser Thr Pro Ile Glu
85 90 95

Ser Ile Thr Glu Glu Ile Val Asp Arg Val Tyr Asn Ile Asn Val Lys
100 105 110

Gly Val Ile Trp Gly Met Gln Ala Ala Val Glu Ala Phe Lys Lys Glu
115 120 125

4/4/1


Gly His Gly Gly Lys Ile Val Asn Ala Cys Ser Gln Ala Gly His Val
 130 135 140

Gly Asn Pro Glu Leu Ala Val Tyr Ser Ser Ser Lys Phe Ala Val Arg
 145 150 155 160

Gly Leu Thr Gln Thr Ala Ala Arg Asp Leu Ala Pro Leu Gly Ile Thr
 165 170 175

Val Asn Gly Phe Cys Pro Gly Ile Val Lys Thr Pro Met Trp Ala Glu
 180 185 190

Ile Asp Arg Gln Cys Arg Lys Arg Arg Ala Asn Arg Trp Ala Thr Ala
 195 200 205

Arg Leu Asn Leu Pro Asn Ala Ser Pro Leu Ala Ala Cys Arg Ser Leu
 210 215 220

Lys Thr Ser Pro Pro Ala Cys Arg Ser Ser Pro Ala Arg Ile Pro Thr
 225 230 235 240

Ile

<210> 135

<211> 554

<212> PRT

<213> *Clostridium pasteurianum*

<400> 135

Met Lys Ser Lys Arg Phe Gln Val Leu Ser Glu Arg Pro Val Asn Lys
 1 5 10 15

Asp Gly Phe Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met Ser
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Lys Glu Gly Lys Val
 35 40 45

Ile Glu Leu Asp Gly Lys Asn Arg Glu Asp Phe Asp Met Ile Asp Arg
 50 55 60

Phe Ile Ala Asn Tyr Gly Ile Asn Leu Asn Arg Ala Glu Asp Val Ile
 65 70 75 80

Lys Met Asp Ser Val Lys Leu Ala Lys Met Leu Val Asp Ile Asn Val
85 90 95

Asp Arg Lys Thr Ile Val Glu Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Gly Asn Met Asn Val Val Glu Met Met Met Ala
115 120 125

Leu Gln Lys Met Arg Ala Arg Lys Thr Pro Ser Asn Gln Cys His Val
130 135 140

Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Ala Ala Ile Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Gly Ile Val
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ala Gln Val
180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Cys Ala Ile Glu Glu Ala Thr Glu
195 200 205

Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val Ser
210 215 220

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

Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Ala Leu Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Met Thr Gly Ala Leu Pro Ser Gly Ile Arg Ala Val Leu Gly Glu Asn
305 310 315 320

443
AS

Leu Ile Thr Thr Met Leu Asp Ile Glu Val Ala Ser Ala Asn Asp Gln
325 330 335

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Met Leu Met Gln
340 345 350

Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ser Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Met Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Ser Glu Glu Glu Val Ile Thr Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Ile Gln Ala Val Phe Glu Gly Leu Lys Leu Pro Ala Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Val Thr Tyr Ser His Gly Ser Lys Asp Val
435 440 445

Pro Glu Arg Asn Val Val Glu Asp Leu Lys Ala Ala Glu Glu Met Ile
450 455 460

Asn Arg Gly Ile Thr Gly Ile Asp Val Val Lys Ala Leu Ser Lys His
465 470 475 480

Gly Phe Asp Asp Ile Ala Glu Asn Ile Leu Asn Met Leu Lys Gln Arg
485 490 495

Ile Ser Gly Asp Tyr Leu Gln Thr Ser Ala Ile Ile Asp Lys Asn Phe
500 505 510

Asn Val Val Ser Ala Val Asn Asp Cys Asn Asp Tyr Met Gly Pro Gly
515 520 525

Thr Gly Tyr Arg Leu Ser Lys Glu Arg Trp Asp Glu Ile Lys Asn Ile
530 535 540

Pro Asn Ala Met Lys Pro Glu Asp Ile Lys
545 550

444
SA

<210> 136
 <211> 179
 <212> PRT
 <213> *Clostridium pasteurianum*

<400> 136

Met Glu Leu Lys Glu Lys Asp Ile Ala Leu Ser Gly Asn Gln Ser Asn
 1 5 10 15

Glu Val Val Ile Gly Ile Ala Pro Ala Phe Gly Lys Tyr Gln His Gln
 20 25 30

Ser Ile Val Gly Val Pro His Asp Lys Ile Leu Arg Glu Leu Ile Ala
 35 40 45

Gly Ile Glu Glu Glu Gly Leu Lys Ser Arg Val Val Arg Ile Ile Arg
 50 55 60

Thr Ser Asp Val Ser Phe Ile Ala His Asp Ala Ala Val Leu Ser Gly
 65 70 75 80

Ser Gly Ile Gly Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His
 85 90 95

Gln Lys Asp Leu Leu Pro Leu Asn Asn Leu Glu Leu Phe Pro Gln Ala
 100 105 110

Pro Leu Leu Asp Leu Asp Ile Phe Arg Leu Ile Gly Lys Asn Ala Ala
 115 120 125

Lys Tyr Ala Lys Gly Glu Ser Pro Asn Pro Val Pro Thr Arg Asn Asp
 130 135 140

Gln Met Val Arg Pro Lys Phe Gln Ala Lys Ala Ala Leu Leu His Ile
 145 150 155 160

Lys Glu Thr Lys His Val Val Gln Asn Ala Lys Pro Ile Glu Leu Glu
 165 170 175

Ile Ile Ser

<210> 137
 <211> 146
 <212> PRT
 <213> *Clostridium pasteurianum*

2/45


<400> 137

Met Ser Asp Ile Thr Asn Asn Ile Lys Val Asp Tyr Glu Asn Asp Tyr
 1 5 10 15

Pro Leu Ala Ala Lys Arg Ser Glu Trp Ile Lys Thr Pro Thr Gly Lys
 20 25 30

Asn Leu Lys Asp Ile Thr Leu Glu Ala Val Ile Asp Glu Asn Val Lys
 35 40 45

Ala Glu Asp Val Arg Ile Ser Arg Asp Thr Leu Glu Leu Gln Ala Gln
 50 55 60

Val Ala Glu Gly Ser Gly Arg Cys Ala Ile Ala Arg Asn Phe Arg Arg
 65 70 75 80

Ala Ala Glu Leu Ile Ser Ile Ser Asp Glu Arg Ile Leu Glu Ile Tyr
 85 90 95

Asn Ala Leu Arg Pro Tyr Arg Ser Thr Lys Asn Glu Leu Leu Ala Ile
 100 105 110

Ala Asp Glu Leu Glu Glu Lys Tyr Asp Ala Lys Val Asn Ala Asp Phe
 115 120 125

Ile Arg Glu Ala Ala Glu Val Tyr Ser Lys Arg Asn Lys Val Arg Ile
 130 135 140

Glu Asp
 145

<210> 138

<211> 555

<212> PRT

<213> *Escherichia blattae*

<400> 138

Met Arg Arg Ser Lys Arg Phe Glu Val Leu Glu Lys Arg Pro Val Asn
 1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
 20 25 30

Gly Ser Pro Trp Asp Pro Pro Ser Ser Val Lys Val Glu Gln Gly Arg
 35 40 45

446
H

Ile Val Glu Leu Asp Gly Lys Ala Arg Ala Asp Phe Asp Met Ile Asp
50 55 60

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Ile Glu Glu Thr Glu His Ala
65 70 75 80

Met Gly Leu Asp Ala Leu Thr Ile Ala Arg Met Leu Val Asp Ile Asn
85 90 95

Val Ser Arg Ala Glu Ile Ile Lys Val Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Val Glu Val Met Ser His Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Ile Gly Ser Gln
180 185 190

Ser Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Phe Thr Ser Tyr Ala Glu Thr Val
210 215 220

Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ala
260 265 270

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

448

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

Leu Arg Pro Val Ser Glu Glu Glu Thr Ile Ala Ile Arg Asn Lys Ala
405 410 415

Ala Arg Ala Val Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Pro Val
420 425 430

Thr Asp Glu Glu Val Thr Ala Ala Thr Tyr Ala His Gly Ser Lys Asp
435 440 445

Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Arg Ala Leu Ser Val
465 470 475 480

Asn Gly Phe Asp Asp Val Ala Asn Asn Ile Leu Asn Met Leu Arg Gln
485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Glu
500 505 510

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

448
J

Gly Thr Gly Tyr Arg Ile Ser Pro Gln Arg Trp Glu Glu Ile Lys Asn
530 535 540

Ile Ala Thr Val Ile Gln Pro Asp Ser Ile Glu
545 550 555

<210> 139

<211> 196

<212> PRT

<213> *Escherichia blattae*

<400> 139

Met Glu Thr Thr Gln Lys Lys Ala Pro Val Phe Thr Leu Asn Leu Val
1 5 10 15

Glu Ser Gly Val Ala Lys Pro Gly Glu Arg Ser Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys Tyr Gln His Lys Thr Leu Ile Asp
35 40 45

Met Pro His Lys Ala Ile Ile Lys Glu Leu Val Ala Gly Val Glu Glu
50 55 60

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

His Val Val Ala Asp Ala Lys Pro Val Thr Leu Asn Ile Glu Ile Thr
180 185 190

c/49

Arg Glu Glu Ala
195

<210> 140
<211> 141
<212> PRT
<213> *Escherichia blattae*

<400> 140

Met Thr Thr Thr Lys Met Ser Ala Ala Asp Tyr Pro Leu Ala Ser Arg
1 5 10 15

Cys Pro Glu Arg Ile Gln Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
20 25 30

Thr Leu Glu Asn Val Leu Ala Gly Lys Val Gly Pro Gln Asp Val Arg
35 40 45

Ile Ser Arg Glu Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
50 55 60

His Arg His Ala Ile Ala Arg Asn Leu Arg Arg Ala Gly Glu Leu Ile
65 70 75 80

Ala Ile Pro Asp Ala Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
85 90 95

Tyr Arg Ser Ser Val Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu
100 105 110

Thr Arg Tyr Gln Ala Thr Val Asn Ala Ala Phe Ile Arg Glu Ala Ala
115 120 125

Glu Val Tyr Arg Gln Arg Asp Lys Leu Arg Lys Glu Ala
130 135 140

<210> 141
<211> 555
<212> PRT
<213> *Citrobacter freundii*

<400> 141

Met Arg Arg Ser Lys Arg Phe Glu Val Leu Ala Gln Arg Pro Val Asn
1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

4/50
S

Glu Ser Pro Tyr Asp Pro Ala Ser Ser Val Lys Val Glu Asn Gly Arg
35 40 45

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

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Pro Glu Ala Glu Arg Ala
65 70 75 80

Met Gln Leu Asp Ala Leu Glu Ile Ala Arg Met Leu Val Asp Ile His
85 90 95

Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Arg Leu Glu Val Met Ala Gln Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln
180 185 190

Cys Gly Ala Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val
210 215 220

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

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser
260 265 270

451
A

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

Leu Arg Pro Val Thr Glu Glu Glu Thr Ile Ala Ile Arg Asn Lys Ala
405 410 415

Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Leu Ile
420 425 430

Ser Asp Glu Glu Val Asp Ala Ala Thr Tyr Ala His Gly Ser Lys Asp
435 440 445

Met Pro Ala Arg Asn Val Val Glu Asp Leu Ala Ala Val Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Ser Ser
465 470 475 480

Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Arg Gln
485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln
500 505 510

452
A

Phe Asp Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540

Ile Ala Gly Val Val Gln Pro Gly Ser Ile Glu
545 550 555

<210> 142

<211> 194

<212> PRT

<213> *Citrobacter freundii*

● <400> 142

Met Glu Cys Thr Thr Glu Arg Lys Pro Val Phe Thr Leu Gln Val Ser
1 5 10 15

Glu Gly Glu Ala Ala Lys Ala Asp Glu Arg Val Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys Tyr Gln His Lys Thr Leu Ile Asp
35 40 45

Met Pro His Lys Ala Ile Leu Lys Glu Leu Val Ala Gly Ile Glu Glu
50 55 60

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

4/53
J.

His Val Val Gln Asp Arg Ala Pro Val Thr Leu His Ile Ala Leu Val
 180 185 190

Arg Glu

<210> 143
 <211> 142
 <212> PRT
 <213> *Citrobacter freundii*

<400> 143

Met Asn Asp Asn Ile Met Thr Ala Gln Asp Tyr Pro Leu Ala Thr Arg
 1 5 10 15

Cys Pro Glu Lys Ile Gln Thr Pro Thr Gly Lys Pro Leu Thr Glu Ile
 20 25 30

Thr Leu Glu Asn Val Leu Ala Gly Arg Val Gly Pro Gln Asp Val Arg
 35 40 45

Ile Ser Gln Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
 50 55 60

Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
 65 70 75 80

Ala Ile Pro Asp Ala Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
 85 90 95

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

His Thr Trp His Ala Thr Val Asn Ala Gly Phe Val Arg Glu Ser Ala
 115 120 125

Glu Val Tyr Leu Gln Arg Asn Lys Leu Arg Lys Gly Ser Gln
 130 135 140

<210> 144
 <211> 1359
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Amino:piruvato transaminase códon otimizada de *Vibrio fluvialis*

<400> 144

458
P

atgaacaaac cacagtcttg ggaagctcgt gcagaaacct attctctgta cggcttcact 60
 gacatgccgt ccctgcacca gcgtggtact gttgttgtca cgcacggcga aggtccgtac 120
 attgttgacg tcaatggctg ccgttatctg gacgctaatt ctggcctgtg gaatatgggt 180
 gcaggttttg accataaggg tctgatcgac gcagctaagg ctgagtacga gcgttttccg 240
 ggctaccatg cgttcttcgg tcgtatgagc gatcagacgg tgatgctgtc cgaaaaactg 300
 gtagaagtct ctccgttcga cagcggccgt gtgttctata cgaacagcgg tagcgaagca 360
 aacgacacta tggttaagat gctgtggttc ctgcatgcgg cggaaggtaa gccacaaaag 420
 cgcaaaattc tgaccggtt gaacgcgtat cacggcggtta ctgcagttag cgcctccatg 480
 accggtaaac cgtacaacag cgttttcggt ctgccgctgc caggtttcgt tcacctgact 540
 tgccctcact actggcggtta cgggtgaagaa ggcgagacgg aagaacaatt cgttgcacgc 600
 ctggcacgcg aactggaaga gactatccag cgtgagggtg ctgacactat cgctggcttc 660
 tttgctgagc cggttatggg tgcagggtgt gttattccgc ctgctaaagg ttattttcag 720
 gctattctgc caatcctgcg taaatatgac atcccgggtta tctctgacga agttatctgt 780
 ggttttggtc gactggcaa cacctgggggt tgcgtaactt atgattttac tccggatgct 840
 atcatctcta gcaaaaacct gaccgccggt ttcttcccga tgggcgcagt gatcctgggt 900
 ccagaactga gcaagcgcct ggaaaccgca attgaagcaa tcgaggaatt tccgcacggc 960
 tttaccgcgt cgggccatcc ggtaggctgt gcaatcgcg cgtgaaagcgat cgatgttggt 1020
 atgaacgaag gcctggcggg aaacgttcgc cgtctggcac cgcgcttcga agaacgtctg 1080
 aaacatatcg cggaacgtcc gaacattggt gaatatcggt gtatcggttt tatgtgggct 1140
 ctggaggcag tcaaagacaa agcgtctaaa actccgttcg atggcaatct gagcgtgagc 1200
 gaacgtatcg ccaacacttg caccgacctg ggtctgatct gccgtccact gggccaaagc 1260
 gtagtgctgt gtccgccgtt tatectgacc gaagcgcaaa tggacgaaat gttcgacaaa 1320
 ctggagaaag cactggataa agtggttcga gaggtggca 1359

<210> 145
 <211> 1668
 <212> DNA
 <213> *Klebsiella pneumoniae*

<400> 145
 atgaaaagat caaaacgatt tgcagtactg gcccagcgcc ccgtcaatca ggacgggctg 60
 attggcgagt ggcctgaaga ggggctgatc gccatggaca gcccttttga cccggtctct 120
 tcagtaaaag tggacaacgg tctgatcgtc gaactggacg gcaaacgccg ggaccagttt 180
 gacatgatcg accgatttat cgccgattac gcgatcaacg ttgagcgcac agagcaggca 240

atgcgcctgg aggcgggtgga aatagcccggt atgctggtgg atattcacgt cagccgggag 300
 gagatcattg ccatcactac cgccatcacg ccggccaaag cggtcgaggt gatggcgag 360
 atgaacgtgg tggagatgat gatggcgctg cagaagatgc gtgcccgcg gacccccctcc 420
 aaccagtgcc acgtcaccaa tctcaaagat aatccgggtgc agattgccgc tgacgccgcc 480
 gaggccggga tccgcggctt ctcagaacag gagaccacgg tcggtatcgc gcgctacgcg 540
 ccgtttaacg ccctggcgct gttggtcggg tcgcagtgcg gccgccccgg cgtgttgacg 600
 cagtgctcgg tgggaagaggc caccgagctg gagctgggca tgcgtggctt aaccagctac 660
 gccgagacgg tgtcgggtcta cggcaccgaa gcggtattta ccgacggcga tgatacgccg 720
 tgggtcaaagg cgttcctcgc ctccggcctac gcctcccgcg ggttgaaaat gcgctacacc 780
 tccggcaccg gatccgaagc gctgatgggc tattcggaga gcaagtcgat gctctacctc 840
 gaatcgcgct gcattctcat tactaaaggc gccgggggttc agggactgca aaacggcgcg 900
 gtgagctgta tcggcatgac cggcgctgtg ccgtcgggca ttcgggcggg gctggcgga 960
 aacctgatcg cctctatgct cgacctcgaa gtggcggtccg ccaacgacca gactttctcc 1020
 cactcggata ttcgccgcac cgcgcgcacc ctgatgcaga tgctgccggg caccgacttt 1080
 attttctccg gctacagcgc ggtgccgaac tacgacaaca tgttcgccgg ctcgaaacttc 1140
 gatgcggaag attttgatga ttacaacatc ctgcagcgtg acctgatggg tgacggcggc 1200
 ctgcgtccgg tgaccgaggc ggaaaccatt gccattcgcc agaaagcggc gcgggcgatc 1260
 caggcggttt tccgcgagct ggggctgccg ccaatcgccg acgaggaggt ggaggccgcc 1320
 acctacgcgc acggcagcaa cgagatgccg ccgcgtaacg tgggtggagga tctgagtgcg 1380
 gtggaagaga tgatgaagcg caacatcacc ggcctcgata ttgtcggcgc gctgagccgc 1440
 agcggctttg aggatatcgc cagcaatatt ctcaatatgc tgcgccagcg ggtcaccggc 1500
 gattacctgc agacctcggc cattctcgat cggcagttcg aggtggtgag tgcggtcaac 1560
 gacatcaatg actatcaggg gccgggcacc ggctatcgca tctctgccga acgctggggc 1620
 gagatcaaaa atattccggg cgtggttcag cccgacacca ttgaataa 1668

<210> 146

<211> 555

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 146

Met Lys Arg Ser Lys Arg Phe Ala Val Leu Ala Gln Arg Pro Val Asn
 1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

Asp Ser Pro Phe Asp Pro Val Ser Ser Val Lys Val Asp Asn Gly Leu
35 40 45

Ile Val Glu Leu Asp Gly Lys Arg Arg Asp Gln Phe Asp Met Ile Asp
50 55 60

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Glu Arg Thr Glu Gln Ala
65 70 75 80

Met Arg Leu Glu Ala Val Glu Ile Ala Arg Met Leu Val Asp Ile His
85 90 95

Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Val Glu Val Met Ala Gln Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln
180 185 190

Cys Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val
210 215 220

Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

457

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser
260 265 270

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

Leu Arg Pro Val Thr Glu Ala Glu Thr Ile Ala Ile Arg Gln Lys Ala
405 410 415

Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Pro Ile
420 425 430

Ala Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Asn Glu
435 440 445

Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Ser Arg
465 470 475 480

Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Arg Gln
485 490 495

458

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln
500 505 510

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540

Ile Pro Gly Val Val Gln Pro Asp Thr Ile Glu
545 550 555

<210> 147

<211> 585

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 147

gtgcaacaga caaccctaat tcagccctct ttaccctga aaaccgcga gggcggggta 60
gcttctgccg atgaacgcgc cgatgaagtg gtgatcggcg tcggccctgc cttcgataaa 120
caccagcatc acactctgat cgatatgcc catggcgca tcctcaaaga gctgattgcc 180
ggggtggaag aagaggggct tcacgcccgg gtggtgcgca ttctgcgcac gtccgacgtc 240
tcctttatgg cctgggatgc ggccaacctg agcggctcgg ggatcggcat cggtatccag 300
tcgaagggga ccacggtcat ccatcagcgc gatctgctgc cgctcagcaa cctggagctg 360
ttctcccagg cgccgtgct gacgctggag acctaccggc agattggcaa aaacgctgcg 420
cgctatgcgc gcaaagagtc accttcgccg gtgccggtgg tgaacgatca gatggtgcgg 480
ccgaaattta tggccaaagc cgcgctattt catatcaaag agaccaaaca tgtggtgcag 540
gacgccgagc ccgtcacct gcacatcgac ttagtaaggg agtga 585

<210> 148

<211> 194

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 148

Met Gln Gln Thr Thr Gln Ile Gln Pro Ser Phe Thr Leu Lys Thr Arg
1 5 10 15

Glu Gly Gly Val Ala Ser Ala Asp Glu Arg Ala Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys His Gln His His Thr Leu Ile Asp
35 40 45

6/59
JH

Met Pro His Gly Ala Ile Leu Lys Glu Leu Ile Ala Gly Val Glu Glu
50 55 60

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

His Val Val Gln Asp Ala Glu Pro Val Thr Leu His Ile Asp Leu Val
180 185 190

Arg Glu

<210> 149

<211> 426

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 149

atgagcgaga aaaccatgcg cgtgcaggat tatccgtag ccacccgctg cccggagcat 60

atcctgacgc ctaccggcaa accattgacc gatattaccc tcgagaaggt gctctctggc 120

gaggtgggcc cgcaggatgt gcggatctcc cgccagaccc ttgagtacca ggcgagatt 180

gccgagcaga tgcagcgcca tgcggtggcg cgcaatttcc gccgcgcggc ggagcttatc 240

gccattcctg acgagcgcat tctggctatc tataacgcgc tgcgcccgtt ccgctcctcg 300

caggcggagc tgctggcgat cgccgacgag ctggagcaca cctggcatgc gacagtgaat 360

gccgcctttg tccgggagtc ggcggaagtg tatcagcagc ggcataagct gcgtaaagga 420

469

agctaa

426

<210> 150

<211> 141

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 150

Met Ser Glu Lys Thr Met Arg Val Gln Asp Tyr Pro Leu Ala Thr Arg
 1 5 10 15

Cys Pro Glu His Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
 20 25 30

Thr Leu Glu Lys Val Leu Ser Gly Glu Val Gly Pro Gln Asp Val Arg
 35 40 45

Ile Ser Arg Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
 50 55 60

Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
 65 70 75 80

Ala Ile Pro Asp Glu Arg Ile Leu Ala Ile Tyr Asn Ala Leu Arg Pro
 85 90 95

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

His Thr Trp His Ala Thr Val Asn Ala Ala Phe Val Arg Glu Ser Ala
 115 120 125

Glu Val Tyr Gln Gln Arg His Lys Leu Arg Lys Gly Ser
 130 135 140

<210> 151

<211> 1824

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 151

atgccgttaa tagccgggat tgatatcggc aacgccacca ccgaggtggc gctggcgctcc 60

gactaccgc aggcgagggc gtttggtgcc agcgggatcg tcgcgacgac gggcatgaaa 120

gggacgcggg acaatatcgc cgggaccctc gccgcgctgg agcaggccct ggcgaaaaca 180

ccgtggtcga tgagcgatgt ctctgcacatc tatcttaacg aagccgcgcc ggtgattggc 240

4/6/

gatgtggcga tggagaccat caccgagacc attatcaccg aatcgaccat gatcgggtcat 300
 aacccgcaga cgccgggcgg ggtgggcgtt ggcgtgggga cgactatcgc cctcgggcgg 360
 ctggcgacgc tgccggcggc gcagtatgcc gaggggtgga tcgtactgat tgacgacgcc 420
 gtcgatttcc ttgacgccgt gtggtggctc aatgaggcgc tcgaccgggg gatcaacgtg 480
 gtggcggcga tcctcaaaaa ggacgacggc gtgctggtga acaaccgcct gcgtaaaacc 540
 ctgccggtgg tggatgaagt gacgctgctg gagcaggtcc ccgagggggg aatggcggcg 600
 gtggaagtgg ccgcgcgggg ccaggtggtg cggatcctgt cgaatcccta cgggatcgcc 660
 accttcttcg ggctaagccc ggaagagacc caggccatcg tccccatcgc ccgcgccttg 720
 attggcaacc gttccgcggt ggtgctcaag accccgcagg gggatgtgca gtcgcgggtg 780
 atccccggcg gcaacctcta cattagcggc gaaaagcgcc gcggagaggc cgatgtcgcc 840
 gagggcgcgg aagccatcat gcaggcgatg agcgcctgcg ctccggtacg cgacatccgc 900
 ggcgaaccgg gcacccacgc cggcggcatg cttgagcggg tgcgcaaggt aatggcgctc 960
 ctgaccggcc atgagatgag cgcgatatac atccaggatc tgctggcggg ggatacgttt 1020
 attccgcgca aggtgcaggg cgggatggcc ggcgagtgcg ccatggagaa tgccgtcggg 1080
 atggcggcga tgggtgaaagc ggatcgtctg caaatgcagg ttatcgcccg cgaactgagc 1140
 gcccgactgc agaccgaggt ggtggtgggc ggcgtggagg ccaacatggc catcgccggg 1200
 gcgttaacca ctcccggtg tgcggcgccg ctggcgatcc tcgacctcgg cgccggctcg 1260
 acggatgcgg cgatcgtcaa cgcggagggg cagataacgg cgggtccatct cgccggggcg 1320
 gggaatatgg tcagcctggt gattaaaacc gagctgggcc tcgaggatct ttcgctggcg 1380
 gaagcgataa aaaaataccc gctggccaaa gtggaaagcc tgttcagtat tcgtcacgag 1440
 aatggcgcgg tggagtcttt tcgggaagcc ctcagcccgg cgggtgttcg caaagtgggtg 1500
 tacatcaagg agggcgaaact ggtgccgatc gataacgcca gcccgctgga aaaaattcgt 1560
 ctcgtgcgcc ggcaggcgaa agagaaagtg tttgtcacca actgcctgcg cgcgctgcgc 1620
 caggtctcac ccggcggttc cattcgcgat atcgctttg tgggtgctggg gggcggtca 1680
 tcgctggact ttgagatccc gcagcttatac acggaagcct tgtcgcacta tggcgtggtc 1740
 gccgggcagg gcaatattcg gggaacagaa gggccgcgca atgcggtcgc caccgggctg 1800
 ctactggccg gtcaggcgaa ttaa 1824

<210> 152

<211> 607

<212> PRT

<213> *Klebsiella pneumoniae*

464

<400> 152

Met Pro Leu Ile Ala Gly Ile Asp Ile Gly Asn Ala Thr Thr Glu Val
1 5 10 15

Ala Leu Ala Ser Asp Tyr Pro Gln Ala Arg Ala Phe Val Ala Ser Gly
20 25 30

Ile Val Ala Thr Thr Gly Met Lys Gly Thr Arg Asp Asn Ile Ala Gly
35 40 45

Thr Leu Ala Ala Leu Glu Gln Ala Leu Ala Lys Thr Pro Trp Ser Met
50 55 60

Ser Asp Val Ser Arg Ile Tyr Leu Asn Glu Ala Ala Pro Val Ile Gly
65 70 75 80

Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser Thr
85 90 95

Met Ile Gly His Asn Pro Gln Thr Pro Gly Gly Val Gly Val Gly Val
100 105 110

Gly Thr Thr Ile Ala Leu Gly Arg Leu Ala Thr Leu Pro Ala Ala Gln
115 120 125

Tyr Ala Glu Gly Trp Ile Val Leu Ile Asp Asp Ala Val Asp Phe Leu
130 135 140

Asp Ala Val Trp Trp Leu Asn Glu Ala Leu Asp Arg Gly Ile Asn Val
145 150 155 160

Val Ala Ala Ile Leu Lys Lys Asp Asp Gly Val Leu Val Asn Asn Arg
165 170 175

Leu Arg Lys Thr Leu Pro Val Val Asp Glu Val Thr Leu Leu Glu Gln
180 185 190

Val Pro Glu Gly Val Met Ala Ala Val Glu Val Ala Ala Pro Gly Gln
195 200 205

Val Val Arg Ile Leu Ser Asn Pro Tyr Gly Ile Ala Thr Phe Phe Gly
210 215 220

Leu Ser Pro Glu Glu Thr Gln Ala Ile Val Pro Ile Ala Arg Ala Leu
225 230 235 240

Ile Gly Asn Arg Ser Ala Val Val Leu Lys Thr Pro Gln Gly Asp Val
245 250 255

Gln Ser Arg Val Ile Pro Ala Gly Asn Leu Tyr Ile Ser Gly Glu Lys
260 265 270

Arg Arg Gly Glu Ala Asp Val Ala Glu Gly Ala Glu Ala Ile Met Gln
275 280 285

Ala Met Ser Ala Cys Ala Pro Val Arg Asp Ile Arg Gly Glu Pro Gly
290 295 300

Thr His Ala Gly Gly Met Leu Glu Arg Val Arg Lys Val Met Ala Ser
305 310 315 320

Leu Thr Gly His Glu Met Ser Ala Ile Tyr Ile Gln Asp Leu Leu Ala
325 330 335

Val Asp Thr Phe Ile Pro Arg Lys Val Gln Gly Gly Met Ala Gly Glu
340 345 350

Cys Ala Met Glu Asn Ala Val Gly Met Ala Ala Met Val Lys Ala Asp
355 360 365

Arg Leu Gln Met Gln Val Ile Ala Arg Glu Leu Ser Ala Arg Leu Gln
370 375 380

Thr Glu Val Val Val Gly Gly Val Glu Ala Asn Met Ala Ile Ala Gly
385 390 395 400

Ala Leu Thr Thr Pro Gly Cys Ala Ala Pro Leu Ala Ile Leu Asp Leu
405 410 415

Gly Ala Gly Ser Thr Asp Ala Ala Ile Val Asn Ala Glu Gly Gln Ile
420 425 430

Thr Ala Val His Leu Ala Gly Ala Gly Asn Met Val Ser Leu Leu Ile
435 440 445

Lys Thr Glu Leu Gly Leu Glu Asp Leu Ser Leu Ala Glu Ala Ile Lys
450 455 460

Lys Tyr Pro Leu Ala Lys Val Glu Ser Leu Phe Ser Ile Arg His Glu
465 470 475 480

464
A

Asn Gly Ala Val Glu Phe Phe Arg Glu Ala Leu Ser Pro Ala Val Phe
485 490 495

Ala Lys Val Val Tyr Ile Lys Glu Gly Glu Leu Val Pro Ile Asp Asn
500 505 510

Ala Ser Pro Leu Glu Lys Ile Arg Leu Val Arg Arg Gln Ala Lys Glu
515 520 525

Lys Val Phe Val Thr Asn Cys Leu Arg Ala Leu Arg Gln Val Ser Pro
530 535 540

Gly Gly Ser Ile Arg Asp Ile Ala Phe Val Val Leu Val Gly Gly Ser
545 550 555 560

Ser Leu Asp Phe Glu Ile Pro Gln Leu Ile Thr Glu Ala Leu Ser His
565 570 575

Tyr Gly Val Val Ala Gly Gln Gly Asn Ile Arg Gly Thr Glu Gly Pro
580 585 590

Arg Asn Ala Val Ala Thr Gly Leu Leu Leu Ala Gly Gln Ala Asn
595 600 605

<210> 153
<211> 354
<212> DNA
<213> *Klebsiella pneumoniae*

<400> 153
atgtcgtttt caccgccagg cgtacgcctg ttttacgatc cgcgcgggca ccatgccggc 60
gccatcaatg agctgtgctg ggggctggag gagcaggggg tcccctgccca gaccataacc 120
tatgacggag gcggtgacgc cgctgcgctg ggcgccttg cggccagaag ctcgcccctg 180
cgggtgggta tcgggctcag cgcgtccggc gagatagccc tcactcatgc ccagctgccg 240
gcggacgcgc cgctggctac cggacacgtc accgatagcg acgatcaact gcgtacgctc 300
ggcgccaacg ccgggcagct gggttaaagtc ctgccgttaa gtgagagaaa ctga 354

<210> 154
<211> 117
<212> PRT
<213> *Klebsiella pneumoniae*

<400> 154

Met Ser Leu Ser Pro Pro Gly Val Arg Leu Phe Tyr Asp Pro Arg Gly
1 5 10 15

465

His His Ala Gly Ala Ile Asn Glu Leu Cys Trp Gly Leu Glu Glu Gln
20 25 30

Gly Val Pro Cys Gln Thr Ile Thr Tyr Asp Gly Gly Gly Asp Ala Ala
35 40 45

Ala Leu Gly Ala Leu Ala Ala Arg Ser Ser Pro Leu Arg Val Gly Ile
50 55 60

Gly Leu Ser Ala Ser Gly Glu Ile Ala Leu Thr His Ala Gln Leu Pro
65 70 75 80

Ala Asp Ala Pro Leu Ala Thr Gly His Val Thr Asp Ser Asp Asp Gln
85 90 95

Leu Arg Thr Leu Gly Ala Asn Ala Gly Gln Leu Val Lys Val Leu Pro
100 105 110

Leu Ser Glu Arg Asn
115

<210> 155

<211> 1125

<212> DNA

<213> Sequência artificial

<220>

<223> aminoálcool quinase códon otimizada de *Erwinia caratovora*
subespécie atroseptica

<400> 155

atgagcgatg gccgtctgac cgcactgttt cctgcatttc cacatccggc atccaaccag	60
ccagtgtttg cggaggcttc cccgcacgac gatgaactga tgacgcaggc ggtgccgcag	120
gtttcctgcc agcaagccct ggcaattgcc cagcaggaat atggcctgag cggtcagatg	180
agcctgctgc agggcgaacg tgacgttaat ttctgtctga ccgtaacgcc agatgaacgc	240
tatatgctga aagtcacaa cgctgctgaa ccggcagatg tgagcaactt tcagactgcg	300
ctgctgctgc acctggcacg tcaggcgcca gaactgccag tccctcgtat ccgctccacg	360
aaggctggtc agtctgaaac gggcgtcgaa attgatggtg ttctgctgcg tgtgcgtctg	420
gtttcctacc tggctggcat gccgcagtac ctggcgtctc cgagcacggc actgatgcc	480
cagctgggcg gtactctggc gcagctggac aacgctctgc actctttcac ccatccggcg	540
gctaaccgtg ctctgctgtg ggacatctcc cgcgcagagc aggtccgccc gtacctggac	600
ttcgtagcg agccgcagca gtatcagcac ctgcagcgca tctttgatcg ctatgactct	660

466/AA

aacgtggcac cgctgctgac gacgctgcgc cgccagggtta tccacaacga cctgaacccg 720
 cataacgtcc tggctgatgg ttccagcccg acgcgcgtca cgggtatcat cgacttcggc 780
 gatgcagtgt tcgcgccgct gatctgtgag gttgcgaccg ctctggcgta ccaaattggc 840
 gacggcacgg atctgctgga acatgtggta ccgtttgtcg cagcgtatca ccagcgtatt 900
 ccgctggcgc cggaggaaat cgccctgctg ccagatctga tcgcgacccg catggcactg 960
 actctgacca tcgctcagtg gcgtgcgtct cgctaccagc ataaccgcga atacctgctg 1020
 cgcaacgtgc cgcgctgctg gcaactccctg cagcgtatcg caacttacag ccacgcacaa 1080
 tttctgacgc gcctgcagca ggtttgccca gaaaacgctc gttga 1125

<210> 156
 <211> 1275
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Aminoálcool O-fosfato liase códon otimizada de *Erwinia caratovora*
subespécie atroseptica

<400> 156
 atgactgcaa ctgaagctct gctggcacgt cgtcagcgcg ttctgggcgg tggctaccgt 60
 ctgttctacg aagaaccgct gcatgttgca cgcggcgaag gtgtatggct gttcgatcat 120
 cagggtaaac gttacctgga cgtatataac aacgtagcta gcgtaggcta ctgtcaccgc 180
 gccgtttaga aagcggtcgc gcgtcaatct gcgcaactga acaccatac gcgctacctg 240
 catcacgcga tcgtagattt tgctgaagat ctgctgtctg agttcccggc agaactgaac 300
 aacgtcatgc tgacctgtac tggctccgaa gcgaacgacc tggccctgcg cattgcgcgt 360
 cacgttacgg gtggtaccgg catgctgggtg acccgttggg cctaccatgg tgttacgtcc 420
 gctctggcgg agctgtcccc gtccctgggc gacggcgtag tacgcggttc ccacgtaaag 480
 ctgatcgatg ctccggatac ctaccgtcag ccgggtgctt tcctgacctc tatccgcgaa 540
 gcgctggcac agatgcagcg tgaaggattt cgtccggcgg ctctgctggt tgatactatc 600
 ttctcctccg acggtgtatt ctgtgcgccg gaagggtgaga tggcccaggc agccgcactg 660
 atccgtcagg ccggtggcct gttcattgcy gacgaagtgc agccgggctt tggctcgatc 720
 ggtgaatccc tgtgggggtt cgcacgtcat aacgtggttc cagatctggt ttctctgggc 780
 aaaccgatgg gtaacggcca tccgattgct ggtctggtag gtcgctccgc actgttcgac 840
 gcttttggtc gtgatgttcg ctactttaat actttcggcg gtaaccagc atcctgccag 900
 gcggcacatg ctgttctgcy cgttatccgt gaagaacagc tgcagcagaa cgcgcagcgt 960
 gttggtgatt atctgcgcca aggtctgcag cagctggcac aacacttccc gctgatcgg 1020

467
JA

gacattcgtg catatggtct gtttatcggt gctgaactgg tttccgaccg tgaatccaaa 1080
 accccagcga gcgagtctgc actgcagggt gttaacgcga tgcgtcagcg tgggtgactg 1140
 atctccgcaa ccggcccggc ggcgaacatt ctgaagatcc gtcctccgct ggtattcctg 1200
 gaggaacacg cggacgtgtt cctgactacc ctgtccgacg tgctggcgct gatcggtact 1260
 cgtgcacagc gttaa 1275

<210> 157
 <211> 77
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 157
 caggaggaat taaccatggg gggttctcat catcatcatc atcatggtga cgatgacgat 60
 aagatgagcg atggccg 77

<210> 158
 <211> 77
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 158
 cgccatcgc tcatttattc gtcacgtca ccatgatgat gatgatgatg agaaccccc 60
 atgggttaatt cctcctg 77

<210> 159
 <211> 20
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 159
 ggacctgctt cgctttatcg 20

<210> 160
 <211> 15
 <212> DNA
 <213> Sequência artificial

<220>
 <223> Iniciador

<400> 160

468
JA

gctagagatg atagc

15

<210> 161

<211> 18

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador

<400> 161

ggaagagact atccagcg

18

<210> 162

<211> 50

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador

<400> 162

gcgcgcccgg gaagaaggag ctcttcacca tgaacaaacc acagtcttgg

50

<210> 163

<211> 28

<212> DNA

<213> Sequência artificial

<220>

<223> Iniciador

<400> 163

gcgcgcccgg gttcatgccca cctctgcg

28

<210> 164

<211> 558

<212> PRT

<213> *Lactobacillus brevis*

<400> 164

Met	Lys	Arg	Gln	Lys	Arg	Phe	Glu	Glu	Leu	Glu	Lys	Arg	Pro	Ile	His
1				5					10					15	

Leu	Asp	Gly	Phe	Val	Lys	Glu	Trp	Pro	Glu	Glu	Gly	Phe	Val	Ala	Met
			20					25					30		

Met	Gly	Pro	Asn	Asp	Pro	Lys	Pro	Ser	Ile	Lys	Ile	Glu	Asn	Gly	Lys
		35					40					45			

Val	Thr	Glu	Met	Asp	Ser	Lys	Pro	Ala	Ala	Asp	Phe	Asp	Leu	Ile	Asp
	50					55					60				

469
JA

Leu Tyr Ile Ala Lys Tyr Gly Ile Lys Leu Glu Asn Ala Glu Lys Val
65 70 75 80

Met Ala Met Asp Ser Thr Lys Ile Ala Asn Met Leu Cys Asp Pro Asn
85 90 95

Val Pro Arg Lys Asp Ile Ile Glu Ile Thr Thr Ala Met Thr Pro Ala
100 105 110

Lys Ala Glu Glu Val Ile Ser Lys Leu Asn Phe Ala Glu Met Ile Met
115 120 125

Ala Thr Gln Lys Met Arg Pro Arg Arg Thr Pro Ala Thr Gln Cys His
130 135 140

Val Thr Asn Ile Arg Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Asp Ala Ala Leu Arg Gly Phe Pro Glu Gln Glu Thr Thr Thr Ala Val
165 170 175

Ala Arg Tyr Ala Pro Leu Asn Ala Ile Ser Leu Met Val Gly Ala Gln
180 185 190

Thr Gly Arg Pro Gly Val Ile Thr Gln Cys Ser Val Glu Glu Ala Glu
195 200 205

Glu Leu Ser Leu Gly Met Arg Gly Phe Thr Gly Tyr Ala Glu Thr Ile
210 215 220

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

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Met Met Gly Tyr Thr
260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Ala Ser Gly Val Gln Gly Leu Gln Asn Gly Gly Val Ser Cys Ile
290 295 300

470
A

Gly Ile Pro Gly Ser Val Pro Ser Gly Ile Arg Ser Val Leu Gly Glu
305 310 315 320

Asn Leu Leu Cys Met Met Leu Asp Leu Glu Cys Ala Ser Ala Asn Asp
325 330 335

Gln Ala Phe Ser His Ser Asp Met Arg Arg Thr Glu Arg Leu Leu Gly
340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ser Ser Thr
355 360 365

Pro Asn Tyr Asp Asn Thr Phe Ala Gly Ser Asn Thr Asp Gly Leu Asp
370 375 380

Tyr Asp Asp Tyr Tyr Val Met Glu Arg Asp Leu Ala Ile Asn Gly Gly
385 390 395 400

Ile His Pro Val Asp Glu Gln Thr Ile Ile Lys Ala Arg Asn Lys Ala
405 410 415

Ala Arg Ala Leu Gln Gly Val Phe Glu Asp Leu Gly Leu Pro Lys Ile
420 425 430

Thr Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala Asn Thr Ser Lys Asp
435 440 445

Met Pro Glu Arg Asn Met Val Glu Asp Met Lys Ala Ala Gln Asp Leu
450 455 460

Met Asp Arg Gly Ile Thr Gly Val Asp Ile Val Lys Ala Leu Phe Asn
465 470 475 480

His Gly Phe Lys Asp Val Ala Gln Ala Val Leu Asp Leu Gln Lys Gln
485 490 495

Lys Val Cys Gly Asp Phe Leu Gln Thr Ser Ala Ile Phe Asp Ser Lys
500 505 510

Trp His Val Ile Ser Ala Val Asn Asp Ala Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Leu Glu Glu Asp Thr Glu Glu Trp Glu Arg Ile
530 535 540

481
SP

Lys Asn Leu Pro Phe Ala Ile Asp Pro Gln Asn Met Gln Leu
545 550 555

<210> 165
<211> 239
<212> PRT
<213> *Lactobacillus brevis*

<400> 165

Met Ala Gln Glu Ile Asp Glu Asn Leu Leu Arg Asn Ile Ile Arg Asp
1 5 10 15

Val Ile Ala Glu Thr Gln Thr Gly Asp Thr Pro Ile Ser Phe Lys Ala
20 25 30

Asp Ala Pro Ala Ala Ser Ser Ala Thr Thr Ala Thr Ala Ala Pro Val
35 40 45

Asn Gly Asp Gly Pro Glu Pro Glu Lys Pro Val Asp Trp Phe Lys His
50 55 60

Val Gly Val Ala Lys Pro Gly Tyr Ser Arg Asp Glu Val Val Ile Ala
65 70 75 80

Val Ala Pro Ala Phe Ala Glu Val Met Asp His Asn Leu Thr Gly Ile
85 90 95

Ser His Lys Glu Ile Leu Arg Gln Met Val Ala Gly Ile Glu Glu Glu
100 105 110

Gly Leu Lys Ala Arg Ile Val Lys Val Tyr Arg Thr Ser Asp Val Ser
115 120 125

Phe Cys Gly Ala Glu Gly Asp His Leu Ser Gly Ser Gly Ile Ala Ile
130 135 140

Ala Ile Gln Ser Lys Gly Thr Thr Ile Ile His Gln Lys Asp Gln Glu
145 150 155 160

Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Val Leu Asp Gly
165 170 175

Asp Thr Tyr Arg Ala Ile Gly Lys Asn Ala Ala Glu Tyr Ala Lys Gly
180 185 190

Met Ser Pro Ser Pro Val Pro Thr Val Asn Asp Gln Met Ala Arg Val
195 200 205

472
JA

Gln Tyr Gln Ala Leu Ser Ala Leu Met His Ile Lys Glu Thr Lys Gln
210 215 220

Val Val Met Gly Lys Pro Ala Glu Gln Ile Glu Val Asn Phe Asn
225 230 235

<210> 166

<211> 175

<212> PRT

<213> *Lactobacillus brevis*

<400> 166

Met Ser Glu Ile Asp Asp Leu Val Ala Lys Ile Val Gln Gln Ile Gly
1 5 10 15

Gly Thr Glu Ala Ala Asp Gln Thr Thr Ala Thr Pro Thr Ser Thr Ala
20 25 30

Thr Gln Thr Gln His Ala Ala Leu Ser Lys Gln Asp Tyr Pro Leu Tyr
35 40 45

Ser Lys His Pro Glu Leu Val His Ser Pro Ser Gly Lys Ala Leu Asn
50 55 60

Asp Ile Thr Leu Asp Asn Val Leu Asn Asp Asp Ile Lys Ala Asn Asp
65 70 75 80

Leu Arg Ile Thr Pro Asp Thr Leu Arg Met Gln Gly Glu Val Ala Asn
85 90 95

Asp Ala Gly Arg Asp Ala Val Gln Arg Asn Phe Gln Arg Ala Ser Glu
100 105 110

Leu Thr Ser Ile Pro Asp Asp Arg Leu Leu Glu Met Tyr Asn Ala Leu
115 120 125

Arg Pro Tyr Arg Ser Thr Lys Ala Glu Leu Leu Ala Ile Ser Ala Glu
130 135 140

Leu Lys Asp Lys Tyr His Ala Pro Val Asn Ala Gly Trp Phe Ala Glu
145 150 155 160

Ala Ala Asp Tyr Tyr Glu Ser Arg Lys Lys Leu Lys Gly Asp Asn
165 170 175

473

<210> 167

<211> 554

<212> PRT

<213> *Salmonella enterica subspécie enterica sorovar Choleraesuis*

<400> 167

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Ala Val
 35 40 45

Thr Glu Leu Asp Gly Lys Pro Val Ser Glu Phe Asp Leu Ile Asp His
 50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Asn Arg Ala Glu Glu Val Met
 65 70 75 80

Ala Met Asp Ser Ile Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
 85 90 95

Lys Arg Ser Glu Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
 100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ala
 115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
 130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
 145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
 165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
 180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
 195 200 205

474

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Thr Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Gly Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 168

<211> 224

<212> PRT

<213> *Salmonella enterica subspécie enterica sorovar Choleraesuis cepa SC-B67*

<400> 168

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
1 5 10 15

Arg Asp Met Lys Gly Ser Asp Lys Pro Val Ser Phe Asn Ala Pro Ala
20 25 30

Ala Ser Thr Ala Pro Gln Thr Ala Ala Pro Ala Gly Asp Gly Phe Leu
35 40 45

Thr Glu Val Gly Glu Ala Arg Gln Gly Thr Gln Gln Asp Glu Val Ile
50 55 60

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

Gly Leu Pro His Lys Ser Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 169

<211> 173

<212> PRT

<213> *Salmonella enterica subspécie enterica sorovar Choleraesuis cepa*
SC-B67

<400> 169

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
1 5 10 15

Met Asn Ser Leu Gln Gly Asp Ala Pro Ala Ala Ala Ser Ala Ala Gly
20 25 30

Gly Thr Ser Arg Ser Ala Lys Val Ser Asp Tyr Pro Leu Ala Asn Lys
35 40 45

His Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe
50 55 60

Thr Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg
65 70 75 80

487
H

Ile Thr Pro Glu Thr Leu Arg Leu Gln Ala Ser Ile Ala Lys Asp Ala
85 90 95

Gly Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr
100 105 110

Ala Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
115 120 125

Tyr Arg Ser Thr Lys Glu Glu Leu Met Ala Ile Ala Asp Asp Leu Glu
130 135 140

Asn Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala
145 150 155 160

Thr Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 170
<211> 554
<212> PRT
<213> *Escherichia coli*

<400> 170

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val
35 40 45

Thr Glu Leu Asp Gly Lys Pro Ala Ser Gln Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Ala Arg Ala Glu Glu Val Ile
65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Lys Arg Ser Asp Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ala
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Phe Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Phe Ser Ala Val Pro
355 360 365

479
A

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Ser Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Met Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Asp Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 171

<211> 222

<212> PRT

<213> *Escherichia coli*

<400> 171

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
1 5 10 15

Ala Glu Met Gln Pro Ser Asp Lys Ser Val Ser Phe Arg Ala Pro Val
20 25 30

480
AS

Ser Ala Thr Val Ser Ser Ala Pro Asp Thr Gly Asn Phe Leu Thr Glu
35 40 45

Ile Gly Glu Ala Gln Gln Gly Thr Gln Gln Asp Glu Val Ile Ile Ala
50 55 60

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

Pro His Lys Asn Ile Leu Arg Glu Val Ile Ala Gly Ile Glu Glu Glu
85 90 95

Gly Ile Lys Ala Arg Val Ile Arg Cys Phe Lys Ser Ser Asp Val Ala
100 105 110

Phe Val Ala Val Glu Gly Asp Arg Leu Ser Gly Ser Gly Ile Ala Ile
115 120 125

Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly Leu Pro
130 135 140

Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu Thr Leu
145 150 155 160

Asp Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Lys Arg
165 170 175

Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala Arg Pro
180 185 190

Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr Lys Tyr
195 200 205

Val Val Thr Gly Lys Lys Pro Gln Glu Leu Arg Val Thr Phe
210 215 220

<210> 172

<211> 172

<212> PRT

<213> *Escherichia coli*

<400> 172

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Asn Arg
1 5 10 15

481
A

Met Asn Ser Leu Gln Asp Ala Ala Pro Val Ser Ala Val Pro Asn Ala
20 25 30

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

Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe Thr
50 55 60

Leu Glu Asn Val Leu Ser Asp Asn Val Thr Ala Leu Asp Met Arg Ile
65 70 75 80

Thr Pro Glu Thr Leu Arg Ile Gln Ala Ala Ile Ala Arg Asp Ala Gly
85 90 95

Cys Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr Ser
100 105 110

Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro Tyr
115 120 125

Arg Ser Thr Lys Gln Glu Leu Ile Ala Ile Ala Asp Asp Leu Glu Gln
130 135 140

Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala Glu
145 150 155 160

Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 173

<211> 554

<212> PRT

<213> *Shigella sonnei*

<400> 173

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val
35 40 45

y82
A

Thr Glu Leu Asp Gly Lys Pro Ala Ser Gln Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Ala Arg Ala Glu Glu Val Ile
65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Lys Arg Ser Asp Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ala
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

483
AS

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Phe Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Phe Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Ser Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Ser Phe Thr Asp Val Ala Gln Asp Met Leu Asn Met Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Asp Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

484
A

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 174
<211> 222
<212> PRT
<213> *Shigella sonnei*

<400> 174

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
1 5 10 15

Ala Glu Met Gln Pro Ser Asp Lys Ser Val Ser Phe Arg Ala Pro Val
20 25 30

Ser Ala Thr Val Pro Ser Ala Pro Asp Thr Gly Asn Phe Leu Thr Glu
35 40 45

Ile Gly Glu Ala Gln Gln Gly Thr Gln Gln Asp Glu Val Ile Ile Ala
50 55 60

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

Pro His Lys Asn Ile Leu Arg Glu Val Ile Ala Gly Ile Glu Glu Glu
85 90 95

Gly Ile Lys Ala Arg Val Ile Arg Cys Phe Lys Ser Ser Asp Val Ala
100 105 110

Phe Val Ala Val Glu Gly Asp Arg Leu Ser Gly Ser Gly Ile Ala Ile
115 120 125

Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly Leu Pro
130 135 140

Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu Thr Leu
145 150 155 160

Asp Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Lys Arg
165 170 175

485
AP

Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala Arg Pro
180 185 190

Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr Lys Tyr
195 200 205

Val Val Thr Gly Lys Lys Pro Gln Glu Leu Arg Val Thr Phe
210 215 220

<210> 175

<211> 172

<212> PRT

<213> *Shigella sonnei*

● <400> 175

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Asn Arg
1 5 10 15

Met Asn Ser Leu Gln Asp Ala Ala Pro Val Ser Ala Val Pro Asn Ala
20 25 30

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

Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe Thr
50 55 60

Leu Glu Asn Val Leu Ser Asp Asn Val Thr Ala Leu Asp Met Arg Ile
65 70 75 80

Thr Pro Glu Thr Leu Arg Ile Gln Ala Ala Ile Ala Arg Asp Ala Gly
85 90 95

Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr Ser
100 105 110

Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro Tyr
115 120 125

Arg Ser Thr Lys Gln Glu Leu Ile Ala Ile Ala Asp Asp Leu Glu Gln
130 135 140

Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala Glu
145 150 155 160

Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 176

<211> 554

<212> PRT

<213> *Yersinia bercovieri*

<400> 176

Met Lys Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val
35 40 45

Thr Glu Leu Asp Gly Lys Pro Gln Ser Ser Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Asp His Ala Glu Glu Val Met
65 70 75 80

Lys Met Asp Ser Ile Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Ser Arg Arg Thr Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ser
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Phe Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

487
A

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Val Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Ile Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ala Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

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

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

438

Asp Glu Glu Val Ile Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Ser Lys Asn Arg Thr Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Glu Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Ile Ala Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Lys Gly Asp Asn
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Glu Gly Ala Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Asn Ala Leu Asp Pro Asn Glu Leu Gly
545 550

<210> 177

<211> 221

<212> PRT

<213> *Yersinia bercovieri*

<400> 177

Met Val Asp Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Gly Val
1 5 10 15

Leu Gln Glu Met Gln Gly Glu Lys Asn Ser Val Ser Phe Lys Gln Glu
20 25 30

Ser Gln Pro Ala Thr Ala Val Ala Ser Gly Asp Phe Leu Thr Glu Val
35 40 45

Gly Glu Ala Arg Pro Gly Ser Asn Gln Asp Glu Val Ile Ile Ala Val
50 55 60

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

His Lys Asn Ile Leu Arg Glu Leu Ile Ala Gly Ile Glu Glu Glu Gly
85 90 95

Ile Lys Ala Arg Val Ile Arg Cys Phe Lys Ser Ser Asp Val Ala Phe
100 105 110

Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile Ser Ile Gly
115 120 125

Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly Leu Pro Pro
130 135 140

Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu Thr Leu Glu
145 150 155 160

Thr Tyr Arg Leu Ile Gly Lys Asn Ala Ala Arg Tyr Ala Lys Arg Glu
165 170 175

Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala Arg Pro Lys
180 185 190

Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr Lys Tyr Val
195 200 205

Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 178

<211> 174

<212> PRT

<213> *Yersinia bercovieri*

<400> 178

Met Asn Ser Glu Ala Ile Glu Ser Met Val Arg Asp Val Leu Asn Lys
1 5 10 15

Met Asn Ser Leu Gln Gly Gln Ala Pro Ala Ala Cys Pro Ala Pro Ala
20 25 30

Ala Ser Ser Arg Ser Asp Ala Lys Val Ser Asp Tyr Pro Leu Ala Asn
35 40 45

Lys His Pro Asp Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp
50 55 60

Leu Thr Leu Ala Asn Val Leu Asn Gly Ser Val Thr Ser Gln Asp Leu
65 70 75 80

490
A

Arg Ile Thr Pro Glu Ile Leu Arg Ile Gln Ala Ser Ile Ala Lys Asp
85 90 95

Ala Gly Arg Pro Leu Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu
100 105 110

Thr Ala Val Pro Asp Asp Lys Val Leu Asp Ile Tyr Asn Ala Leu Arg
115 120 125

Pro Phe Arg Ser Ser Lys Glu Glu Leu Asn Ala Ile Ala Asp Asp Leu
130 135 140

Glu Lys Thr Tyr Gln Ala Thr Ile Cys Ala Ala Phe Val Arg Glu Ala
145 150 155 160

Ala Val Leu Tyr Val Gln Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 179

<211> 551

<212> PRT

<213> *Yersinia mollaretii*

<400> 179

Met Lys Ser Lys Arg Phe Glu Ala Arg Pro Val Asn Gln Asp Gly Phe
1 5 10 15

Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu Ser Pro Asn
20 25 30

Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val Thr Glu Leu
35 40 45

Asp Gly Lys Pro Gln Ser Ser Phe Asp Leu Ile Asp His Phe Ile Ala
50 55 60

Arg Tyr Gly Ile Asn Leu Asp His Ala Glu Glu Val Met Lys Met Asp
65 70 75 80

Ser Ile Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val Ser Arg Arg
85 90 95

Thr Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys Ile Val Glu
100 105 110

491
A

Val Val Ser His Met Asn Val Val Glu Met Met Met Ser Met Gln Lys
115 120 125

Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val Thr Asn Val
130 135 140

Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu Gly Ala Phe
145 150 155 160

Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala Arg Tyr Ala
165 170 175

Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val Gly Arg Pro
180 185 190

Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu Leu Lys Leu
195 200 205

Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser Val Tyr Gly
210 215 220

Thr Glu Pro Val Phe Thr Asp Gly Asp Asp Thr Pro Trp Ser Lys Gly
225 230 235 240

Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met Arg Phe Thr
245 250 255

Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu Gly Lys Ser
260 265 270

Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys Ala Ala Gly
275 280 285

Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Val Gly Val Pro Ser
290 295 300

Ala Val Pro Ser Gly Ile Arg Ala Ile Leu Ala Glu Asn Leu Ile Cys
305 310 315 320

Ser Ala Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln Thr Phe Thr
325 330 335

His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln Phe Leu Pro
340 345 350

492
A

Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro Asn Tyr Asp
355 360 365

Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe Asp Asp Tyr
370 375 380

Asn Val Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu Arg Pro Val
385 390 395 400

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

Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr Asp Glu Glu
420 425 430

Val Ile Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met Pro Asp Arg
435 440 445

Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile Ser Lys Asn
450 455 460

Arg Thr Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly Gly Phe Glu
465 470 475 480

Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys Ile Ala Gly
485 490 495

Asp Tyr Leu His Thr Ser Ala Ile Ile Lys Gly Asp Asn Gln Val Leu
500 505 510

Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala Thr Gly Tyr
515 520 525

Arg Leu Glu Gly Ala Arg Trp Glu Glu Ile Lys Asn Ile Pro Asn Ala
530 535 540

Leu Asp Pro Asn Glu Leu Gly
545 550

<210> 180

<211> 224

<212> PRT

<213> *Yersinia mollaretii*

<400> 180

493
AA

Met Val Asp Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Gly Val
1 5 10 15

Leu Gln Glu Met Gln Gly Asp Asn Asn Thr Val Ser Phe Lys Pro Val
20 25 30

Ser Gln Pro Ala Thr Ala Ala Thr Ala Val Ala Ala Gly Asp Phe Leu
35 40 45

Thr Glu Val Gly Glu Ala Arg Pro Gly Ser Asn Gln Asp Glu Val Ile
50 55 60

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

Gly Ile Pro His Lys Asn Ile Leu Arg Glu Leu Ile Ala Gly Ile Glu
85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Ile Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Asp Thr Tyr Arg Leu Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 181
<211> 175
<212> PRT

494
SA<213> *Yersinia mollaretii*

<400> 181

Met Asn Ser Glu Ala Ile Glu Ser Met Val Arg Asp Val Leu Asn Lys
1 5 10 15

Met Asn Ser Leu Gln Gly Gln Thr Pro Ala Ala Ala Cys Ala Ala Pro
20 25 30

Ala Ala Ser Ser Arg Ser Asn Ala Lys Val Ser Asp Tyr Pro Leu Ala
35 40 45

Asn Lys His Pro Asp Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp
50 55 60

Asp Leu Thr Leu Ala Asn Val Leu Asn Gly Ser Val Thr Ser Gln Asp
65 70 75 80

Leu Arg Ile Thr Pro Glu Ile Leu Arg Ile Gln Ala Ser Ile Ala Lys
85 90 95

Asp Ala Gly Arg Pro Leu Leu Ala Met Asn Phe Glu Arg Ala Ala Glu
100 105 110

Leu Thr Ala Val Pro Asp Asp Lys Val Leu Asp Ile Tyr Asn Ala Leu
115 120 125

Arg Pro Phe Arg Ser Ser Lys Glu Glu Leu Asn Ala Ile Ala Asp Asp
130 135 140

Leu Glu Lys Thr Tyr Lys Ala Thr Ile Cys Ala Ala Phe Val Arg Glu
145 150 155 160

Ala Ala Val Leu Tyr Val Gln Arg Lys Lys Leu Lys Gly Asp Asp
165 170 175

<210> 182

<211> 554

<212> PRT

<213> *Yersinia enterocolitica subspécie enterocolitica*

<400> 182

Met Lys Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val
35 40 45

Thr Glu Leu Asp Gly Lys Pro Gln Ser Ser Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Ala His Ala Glu Glu Val Met
65 70 75 80

Lys Met Asp Ser Ile Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Ser Arg Arg Thr Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ser
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Phe Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

496
A

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Val Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Ile Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ala Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Met Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

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

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Ile Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Ser Lys Asn Arg Thr Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Glu Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

497
AA

Ile Ala Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Lys Gly Asp Asn
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Glu Gly Ala Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Asn Ala Leu Asp Pro Asn Glu Leu Gly
545 550

<210> 183

<211> 223

<212> PRT

<213> *Yersinia enterocolitica subespécie enterocolitica*

<400> 183

Met Val Asp Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Gly Val
1 5 10 15

Leu Gln Glu Met Gln Gly Asp Ser Asn Thr Val Ser Phe Lys Gln Glu
20 25 30

Thr Gln Pro Thr Ala Thr Ala Ala Val Ala Ser Gly Asp Phe Leu Thr
35 40 45

Glu Val Gly Glu Ala Arg Pro Gly Thr His Gln Asp Glu Val Ile Ile
50 55 60

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

Ile Pro His Lys Asn Ile Leu Arg Glu Leu Ile Ala Gly Ile Glu Glu
85 90 95

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

Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile Ser
115 120 125

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly Leu
130 135 140

Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu Thr
145 150 155 160

498
AA

Leu Glu Thr Tyr Arg Leu Ile Gly Lys Asn Ala Ala Arg Tyr Ala Lys
165 170 175

Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala Arg
180 185 190

Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr Lys
195 200 205

Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 184

<211> 174

<212> PRT

<213> *Yersinia enterocolitica subespécie enterocolitica*

<400> 184

Met Asn Ser Glu Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Lys
1 5 10 15

Met Asn Ser Leu Gln Gly Gln Thr Pro Thr Thr Ala Cys Ala Pro Ala
20 25 30

Ala Ser Ser Arg Ser Asp Ala Lys Val Ser Asp Tyr Pro Leu Ala Asn
35 40 45

Lys His Pro Asp Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp
50 55 60

Leu Thr Leu Ala Asn Val Leu Asn Gly Ser Val Thr Ser Gln Asp Leu
65 70 75 80

Arg Ile Thr Pro Glu Ile Leu Arg Ile Gln Ala Ser Ile Ala Lys Asp
85 90 95

Ala Gly Arg Pro Leu Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu
100 105 110

Thr Ala Val Pro Asp Asp Lys Val Leu Asp Ile Tyr Asn Ala Leu Arg
115 120 125

Pro Phe Arg Ser Ser Lys Glu Glu Leu Asn Ala Ile Ala Asp Asp Leu
130 135 140

Glu Gln Thr Tyr Lys Ala Thr Ile Cys Ala Ala Phe Val Arg Glu Ala
145 150 155 160

499
A

Ia Val Leu Tyr Val Gln Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 185
<211> 554
<212> PRT
<213> *Yersinia intermedia*

<400> 185

Met Lys Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val
35 40 45

Thr Glu Leu Asp Gly Lys Pro Gln Ser Ser Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Ala His Ala Glu Glu Val Met
65 70 75 80

Lys Met Asp Ser Ile Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Ser Arg Arg Thr Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ser
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Tyr Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

500
AA

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Val Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Ile Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ala Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Phe Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

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

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Ile Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Ser Lys Asn Arg Thr Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Glu Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Ile Ala Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Lys Gly Asp Asn
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Glu Gly Ala Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Asn Ala Leu Asp Pro Asn Glu Leu Gly
545 550

<210> 186

<211> 224

<212> PRT

<213> *Yersinia intermedia*

<400> 186

Met Val Asp Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Gly Val
1 5 10 15

Leu Gln Glu Met Gln Gly Asp Lys Asn Thr Val Ser Phe Lys Gln Asp
20 25 30

Thr Gln Pro Ala Ala Ala Ala Ala Thr Val Ala Glu Gly Asn Phe Leu
35 40 45

Thr Glu Val Gly Glu Ala Arg Pro Gly Ser Asn Gln Asp Glu Val Ile
50 55 60

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

Gly Ile Pro His Lys Asn Ile Leu Arg Glu Leu Ile Ala Gly Ile Glu
85 90 95

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

502

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 187

<211> 175

<212> PRT

<213> *Yersinia intermedia*

<400> 187

Met Asn Ser Glu Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Lys
1 5 10 15

Met Asn Ser Leu Gln Ser Gln Thr Pro Ala Ala Ala Cys Ala Ala Pro
20 25 30

Ala Thr Ser Ser Arg Asn Asp Ala Lys Val Ser Asp Tyr Pro Leu Ala
35 40 45

Asn Lys His Pro Asp Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp
50 55 60

Asp Leu Thr Leu Ala Asn Val Leu Asn Gly Ser Val Thr Ser Gln Asp
65 70 75 80

Leu Arg Ile Thr Pro Glu Ile Leu Arg Ile Gln Ala Ser Ile Ala Arg
85 90 95

Asp Ala Gly Arg Pro Leu Leu Ala Ile Asn Phe Glu Arg Ala Ala Glu
100 105 110

503

Leu Thr Ala Val Pro Asp Asp Lys Val Leu Asp Ile Tyr Asn Ala Leu
115 120 125

Arg Pro Phe Arg Ser Ser Lys Glu Glu Leu Asn Ala Ile Ala Asp Asp
130 135 140

Leu Glu Gln Thr Tyr Lys Ala Thr Ile Cys Ala Ala Phe Val Arg Glu
145 150 155 160

Ala Ala Val Leu Tyr Val Gln Arg Lys Lys Leu Lys Gly Asp Asp
165 170 175

<210> 188

<211> 554

<212> PRT

<213> *Listeria welshimeri*

<400> 188

Met Lys Ser Lys Arg Phe Glu Glu Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Leu Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly Lys Val
35 40 45

Val Glu Met Asp Ser Lys Lys Leu Ala Asp Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Lys Tyr Gly Val Asp Leu Ser Arg Ala Glu Glu Val Met
65 70 75 80

Gln Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Pro Arg Glu Lys Ile Val Leu Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ser
115 120 125

Met Gln Lys Met Arg Ser Arg Arg Thr Pro Thr Thr Gln Ala His Val
130 135 140

504

Thr Asn Leu Arg Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Ala Ala Ile Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Val
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ser Leu Leu Val Gly Ser Gln Thr
180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Glu Leu Gly Met Arg Gly Leu Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Ile Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Thr Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys Ile Gly
290 295 300

Ile Pro Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Ala Val Met Leu Asp Leu Glu Val Ala Ser Gly Asn Asp Gln
325 330 335

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ser Ala Thr Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Asp Asp Phe
370 375 380

Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Thr Pro Val Thr Glu Glu Glu Val Val Ala Val Arg Asn Lys Ala Ala
405 410 415

Arg Val Ile Gln Val Val Phe Glu Lys Leu Gly Leu Pro Thr Val Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala Arg Gly Ser Asn Asp Met
435 440 445

Pro Glu Arg Asn Met Val Glu Asp Ile Lys Ala Ala Ala Glu Met Met
450 455 460

Asp Arg Gly Val Thr Gly Leu Asp Val Val Lys Ala Leu Ser Ala Gly
465 470 475 480

Gly Phe Asp Asp Val Ala Glu Ser Val Leu Asn Met Leu Lys Gln Arg
485 490 495

Val Ser Gly Asp Phe Leu His Thr Ser Ala Ile Ile Asp Lys Asp Trp
500 505 510

Asn Val Ile Ser Ser Val Asn Asp Leu Asn Asp Tyr Ala Gly Pro Gly
515 520 525

Thr Gly Tyr Arg Leu Glu Gly Glu Arg Trp Glu Lys Leu Lys Asn Ile
530 535 540

Ala Val Ala Val Asp Ala Asn Glu Leu Asp
545 550

<210> 189

<211> 219

<212> PRT

<213> *Listeria welshimeri*

<400> 189

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

Leu Asp Glu Leu Gln Leu Lys Glu Asp Lys Val Ser Phe Gln Lys Glu
20 25 30

Pro Ser Asn Pro Val Val Ser Asp Glu Ser Phe Leu Thr Glu Val Gly
35 40 45

506
AA

Asp Ala Lys Pro Gly Arg Gln Lys Asp Glu Val Val Ile Ala Val Ala
50 55 60

Pro Ala Phe Gly Lys Tyr Gln Thr Lys Asn Ile Val Gly Val Pro His
65 70 75 80

Lys Gln Ile Leu Arg Glu Val Ile Ala Gly Ile Glu Glu Glu Gly Leu
85 90 95

Lys Ala Arg Val Val Arg Val Phe Arg Ser Ser Asp Val Ala Phe Val
100 105 110

Ala Val Glu Gly Asp Arg Leu Ser Gly Ser Gly Ile Cys Ile Gly Ile
115 120 125

Gln Ser Arg Gly Thr Ala Leu Ile His Gln Lys Asp Leu Gln Pro Leu
130 135 140

Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Ile Thr Leu Glu Thr
145 150 155 160

Tyr Arg Ala Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys Gly Glu Ser
165 170 175

Pro Asn Pro Val Pro Met Val Asn Asp Gln Met Ala Arg Pro Lys Phe
180 185 190

Gln Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His Val Val
195 200 205

Gln Gly Lys Asn Ala Val Glu Leu Gln Val Asn
210 215

<210> 190

<211> 170

<212> PRT

<213> *Listeria welshimeri*

<400> 190

Met Asn Gln Glu Ala Leu Glu Asn Met Val Arg Asn Ile Leu Gln Glu
1 5 10 15

Val Asn Ser Gly Ala Val Ser Thr Thr Thr Ser Gln Lys Val Ser Gly
20 25 30

Asp Thr Leu Thr Val Arg Asp Tyr Pro Leu Gly Lys Lys Arg Pro Glu
35 40 45

507

Leu Val Lys Thr Ser Thr Ser Lys Ser Leu Asp Asp Ile Thr Leu Lys
50 55 60

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

Glu Thr Leu Lys Met Gln Ala Gln Val Ala Arg Asp Ala Gly Arg Ala
85 90 95

Thr Leu Ala Asn Asn Phe Glu Arg Ala Ala Glu Leu Thr Ile Val Pro
100 105 110

Asp Glu Arg Ile Leu Glu Ile Tyr Asn Ala Met Arg Pro Tyr Arg Ser
115 120 125

Ser Lys Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu Ser Val Tyr
130 135 140

His Ala Thr Ile Cys Ser Asn Tyr Val Arg Glu Ala Ala Gln Leu Tyr
145 150 155 160

Gln Glu Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 191

<211> 554

<212> PRT

<213> *Listeria innocua*

<400> 191

Met Lys Ser Lys Arg Phe Glu Glu Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Leu Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly Lys Val
35 40 45

Val Glu Met Asp Ser Lys Lys Leu Ala Asp Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Lys Tyr Gly Val Asp Leu Ser Arg Ala Glu Glu Val Met
65 70 75 80

gob

Lys Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Pro Arg Glu Lys Ile Val Leu Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ser
115 120 125

Met Gln Lys Met Arg Ser Arg Arg Thr Pro Thr Thr Gln Ala His Val
130 135 140

Thr Asn Leu Arg Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Ala Ala Ile Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Val
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ser Leu Leu Val Gly Ser Gln Thr
180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Glu Leu Gly Met Arg Gly Leu Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Ile Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Thr Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys Ile Gly
290 295 300

Ile Pro Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

509
P

Leu Ile Ala Val Met Leu Asp Leu Glu Val Ala Ser Gly Asn Asp Gln
325 330 335

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ser Ala Thr Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Asp Asp Phe
370 375 380

Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Thr Pro Val Thr Glu Glu Glu Val Val Ala Val Arg Asn Lys Ala Ala
405 410 415

Arg Val Ile Gln Val Val Phe Asp Lys Leu Gly Leu Pro Ala Val Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala Arg Gly Ser Asn Asp Met
435 440 445

Pro Glu Arg Asn Met Val Glu Asp Ile Lys Ala Ala Ala Glu Met Met
450 455 460

Asp Arg Gly Val Thr Gly Leu Asp Val Val Lys Ala Leu Ala Ala Gly
465 470 475 480

Gly Phe Asp Asp Val Ala Glu Ser Val Leu Asn Met Leu Lys Gln Arg
485 490 495

Val Ser Gly Asp Phe Leu His Thr Ser Ala Ile Ile Asp Lys Asp Trp
500 505 510

Asn Val Ile Ser Ser Val Asn Asp Leu Asn Asp Tyr Ala Gly Pro Gly
515 520 525

Thr Gly Tyr Arg Leu Glu Gly Glu Arg Trp Glu Lys Leu Lys Asp Ile
530 535 540

Ala Val Ala Val Asp Ala Asn Glu Leu Asp
545 550

510
AA

<210> 192
 <211> 219
 <212> PRT
 <213> *Listeria innocua*

<400> 192

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

Leu Asp Glu Leu Gln Leu Lys Glu Asp Lys Val Ser Phe Gln Lys Glu
 20 25 30

Gln Pro Ser Val Ala Val Ser Asp Glu Ser Phe Leu Thr Glu Val Gly
 35 40 45

Asp Ala Lys Pro Gly Arg Gln Lys Asp Glu Val Val Ile Ala Val Ala
 50 55 60

Pro Ala Phe Gly Lys Tyr Gln Thr Lys Asn Ile Val Gly Val Pro His
 65 70 75 80

Lys Gln Ile Leu Arg Glu Val Ile Ala Gly Ile Glu Glu Glu Gly Leu
 85 90 95

Lys Ala Arg Val Val Arg Val Phe Arg Ser Ser Asp Val Ala Phe Val
 100 105 110

Ala Val Glu Gly Asp Lys Leu Ser Gly Ser Gly Ile Cys Ile Gly Ile
 115 120 125

Gln Ser Arg Gly Thr Ala Leu Ile His Gln Lys Asp Leu Gln Pro Leu
 130 135 140

Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Ile Thr Leu Glu Thr
 145 150 155 160

Tyr Arg Ala Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys Gly Glu Ser
 165 170 175

Pro Asn Pro Val Pro Met Val Asn Asp Gln Met Ala Arg Pro Lys Phe
 180 185 190

Gln Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His Val Val
 195 200 205

511
A

Gln Gly Lys Asn Ala Val Glu Leu Gln Val Asn
210 215

<210> 193
<211> 170
<212> PRT
<213> *Listeria innocua*

<400> 193

Met Asn Gln Glu Ala Leu Glu Asn Met Val Arg Asn Ile Leu Gln Glu
1 5 10 15

Val Asn Ser Gly Ala Val Thr Thr Thr Thr Ser Gln Lys Ala Ser Gly
20 25 30

Asp Ser Leu Thr Val Arg Asp Tyr Pro Leu Gly Thr Lys Arg Pro Glu
35 40 45

Leu Val Lys Thr Ala Ser Ser Lys Ser Leu Asp Asp Ile Thr Leu Lys
50 55 60

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

Glu Thr Leu Lys Met Gln Ala Gln Val Ala Arg Asp Ala Gly Arg Ala
85 90 95

Thr Leu Ala Asn Asn Phe Glu Arg Ala Ala Glu Leu Thr Val Val Pro
100 105 110

Asp Glu Arg Ile Leu Glu Ile Tyr Asn Ala Met Arg Pro Tyr Arg Ser
115 120 125

Ser Arg Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu Asn Val Tyr
130 135 140

Gln Ala Thr Ile Cys Ser Asn Tyr Val Arg Glu Ala Ala Gln Leu Tyr
145 150 155 160

Gln Glu Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 194
<211> 554
<212> PRT
<213> *Listeria monocytogenes*

G12

<400> 194

Met Lys Ser Lys Arg Phe Glu Glu Leu Ala Lys Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Leu Ile Ala Met Glu
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly Lys Val
 35 40 45

Val Glu Met Asp Ser Lys Lys Leu Ala Asp Phe Asp Leu Ile Asp His
 50 55 60

Phe Ile Ala Lys Tyr Gly Val Asp Leu Ser Arg Ala Glu Glu Val Met
 65 70 75 80

Gln Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
 85 90 95

Pro Arg Glu Lys Ile Val Leu Leu Thr Thr Ala Met Thr Pro Ala Lys
 100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ser
 115 120 125

Met Gln Lys Met Arg Ser Arg Arg Thr Pro Thr Thr Gln Ala His Val
 130 135 140

Thr Asn Leu Arg Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
 145 150 155 160

Ala Ala Ile Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Val
 165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ser Leu Leu Val Gly Ser Gln Thr
 180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
 195 200 205

Leu Glu Leu Gly Met Arg Gly Leu Thr Cys Tyr Ala Glu Thr Ile Ser
 210 215 220

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

513

Ser Lys Gly Ile Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Thr Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys Ile Gly
290 295 300

Ile Pro Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Ala Val Met Leu Asp Leu Glu Val Ala Ser Gly Asn Asp Gln
325 330 335

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ser Ala Thr Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Asp Asp Phe
370 375 380

Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Thr Pro Val Thr Glu Glu Glu Val Val Ala Val Arg Asn Lys Ala Ala
405 410 415

Arg Val Ile Gln Ala Val Phe Asp Lys Leu Gly Leu Pro Glu Val Thr
420 425 430

Asp Ala Glu Val Glu Ala Ala Thr Tyr Ala Arg Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Met Val Glu Asp Ile Lys Ala Ala Ala Glu Met Met
450 455 460

Asp Arg Gly Val Thr Gly Leu Asp Val Val Lys Ala Leu Ser Ala Gly
465 470 475 480

514
AA

Gly Phe Asp Asp Val Ala Glu Ser Val Leu Asn Met Leu Lys Gln Arg
485 490 495

Val Ser Gly Asp Phe Leu His Thr Ser Ala Ile Ile Asp Lys Asp Trp
500 505 510

Asn Val Ile Ser Ser Val Asn Asp Leu Asn Asp Tyr Ala Gly Pro Gly
515 520 525

Thr Gly Tyr Arg Leu Glu Gly Glu Arg Trp Glu Lys Leu Lys Asp Ile
530 535 540

Ala Val Ala Val Asp Ala Asn Glu Leu Asp
545 550

<210> 195

<211> 219

<212> PRT

<213> *Listeria monocytogenes*

<400> 195

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

Leu Asp Glu Leu Gln Leu Lys Glu Asp Lys Val Ser Phe Gln Lys Glu
20 25 30

Gln Pro Ser Val Ala Val Ser Asp Glu Ser Phe Leu Thr Glu Val Gly
35 40 45

Asp Ala Glu Pro Gly Arg Gln Lys Asp Glu Val Val Ile Ala Val Ala
50 55 60

Pro Ala Phe Gly Lys Tyr Gln Thr Lys Asn Ile Val Gly Val Pro His
65 70 75 80

Lys Gln Ile Leu Arg Glu Leu Ile Ala Gly Ile Glu Glu Glu Gly Leu
85 90 95

Lys Ala Arg Val Val Arg Val Phe Arg Ser Ser Asp Val Ala Phe Val
100 105 110

Ala Val Glu Gly Asp Lys Leu Ser Gly Ser Gly Ile Cys Ile Gly Ile
115 120 125

Gln Ser Arg Gly Thr Ala Leu Ile His Gln Lys Asp Leu Gln Pro Leu
130 135 140

515


Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Ile Thr Leu Glu Thr
 145 150 155 160

Tyr Arg Ala Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys Gly Glu Ser
 165 170 175

Pro Asn Pro Val Pro Met Val Asn Asp Gln Met Ala Arg Pro Lys Phe
 180 185 190

Gln Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His Val Val
 195 200 205

Gln Gly Lys Asn Ala Val Glu Leu Gln Val Asn
 210 215

<210> 196
 <211> 170
 <212> PRT
 <213> *Listeria monocytogenes*

<400> 196

Met Asn Gln Glu Ala Leu Glu Asn Met Val Arg Asn Ile Leu Gln Glu
 1 5 10 15

Val Asn Ser Gly Gly Val Ser Thr Thr Thr Ser Gln Lys Val Ser Gly
 20 25 30

Asp Thr Leu Thr Val Arg Asp Tyr Pro Leu Gly Thr Lys Arg Pro Glu
 35 40 45

Leu Val Lys Thr Ser Thr Ser Lys Ser Leu Asp Asp Ile Thr Leu Lys
 50 55 60

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

Glu Thr Leu Lys Met Gln Ala Gln Val Ala Arg Asp Ala Gly Arg Ala
 85 90 95

Thr Leu Ala Asn Asn Phe Glu Arg Ala Ala Glu Leu Thr Ile Val Pro
 100 105 110

Asp Glu Arg Ile Leu Glu Ile Tyr Asn Ala Met Arg Pro Tyr Arg Ser
 115 120 125

Ser Arg Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu Ser Val Tyr
 130 135 140

5/16

His Ala Thr Ile Cys Ser Asn Tyr Val Arg Glu Ala Ala Gln Leu Tyr
 145 150 155 160

Gln Glu Arg Lys Lys Leu Lys Gly Asp Asn
 165 170

<210> 197

<211> 554

<212> PRT

<213> *Salmonella enterica subespécie enterica sorovar Typhi*

<400> 197

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
 20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Ala Val
 35 40 45

Thr Glu Leu Asp Gly Lys Pro Val Ser Glu Phe Asp Leu Ile Asp His
 50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Asn Arg Ala Glu Glu Val Met
 65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
 85 90 95

Lys Arg Ser Glu Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
 100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ala
 115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
 130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
 145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
 165 170 175

517

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ser Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

S-18
PA

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Gly Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 198

<211> 224

<212> PRT

<213> *Salmonella enterica subspécie enterica sorovar Typhi*

<400> 198

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
1 5 10 15

Arg Asp Met Lys Gly Ser Asp Lys Pro Val Ser Phe Asn Thr Pro Ala
20 25 30

Ala Ser Thr Ala Pro Gln Thr Ala Ala Pro Ala Gly Asp Gly Phe Leu
35 40 45

Thr Glu Val Gly Glu Ala Arg Gln Gly Thr Gln Gln Asp Glu Val Ile
50 55 60

519
AA

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

Gly Leu Pro His Lys Ser Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Asp Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Thr Leu
210 215 220

<210> 199

<211> 173

<212> PRT

<213> *Salmonella enterica subspécie enterica sorovar Typhi*

<400> 199

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
1 5 10 15

Met Asn Ser Leu Gln Gly Asp Ala Pro Ala Ala Ala Pro Ala Ala Gly
20 25 30

Gly Thr Ser Arg Ser Ala Lys Val Ser Asp Tyr Pro Leu Ala Asn Lys
35 40 45

520
A

His Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe
50 55 60

Thr Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg
65 70 75 80

Ile Thr Pro Lys Thr Leu Arg Leu Gln Ala Ser Ile Ala Lys Asp Ala
85 90 95

Gly Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr
100 105 110

Ala Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
115 120 125

Tyr Arg Ser Thr Lys Glu Glu Leu Leu Ala Ile Ala Asp Asp Leu Glu
130 135 140

Asn Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala
145 150 155 160

Gly Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 200

<211> 554

<212> PRT

<213> *Escherichia coli*

<400> 200

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Val Val
35 40 45

Thr Glu Leu Asp Gly Lys Pro Ala Ser Gln Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Ala Arg Ala Glu Glu Val Ile
65 70 75 80

521
SA

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Lys Arg Ser Asp Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ala
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Ile Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

522
AP

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Phe Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Phe Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Ser Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Met Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Asp Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

523
A

<210> 201
 <211> 222
 <212> PRT
 <213> *Escherichia coli*

<400> 201

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
 1 5 10 15

Ala Glu Met Gln Thr Ser Asp Lys Ser Val Ser Phe Arg Ala Pro Val
 20 25 30

Ser Ala Thr Val Ser Ser Val Pro Asp Thr Glu Asn Phe Leu Thr Glu
 35 40 45

Ile Gly Glu Ala Gln Gln Gly Thr Gln Gln Asp Glu Val Ile Ile Ala
 50 55 60

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

Pro His Lys Asn Ile Leu Arg Glu Val Ile Ala Gly Ile Glu Glu Glu
 85 90 95

Gly Ile Lys Ala Arg Val Ile Arg Cys Phe Lys Ser Ser Asp Val Ala
 100 105 110

Phe Val Ala Val Glu Gly Asp Arg Leu Ser Gly Ser Gly Ile Ala Ile
 115 120 125

Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly Leu Pro
 130 135 140

Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu Thr Leu
 145 150 155 160

Asp Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Lys Arg
 165 170 175

Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala Arg Pro
 180 185 190

Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr Lys Tyr
 195 200 205

524
AP

Val Val Thr Gly Lys Lys Pro Gln Glu Leu Arg Val Thr Phe
210 215 220

<210> 202
<211> 172
<212> PRT
<213> *Escherichia coli*

<400> 202

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Asn Arg
1 5 10 15

Met Asn Ser Leu Gln Asp Ala Ala Pro Val Ser Ala Val Pro Ser Ala
20 25 30

Ser Ile Gln Ser Ala Lys Val Thr Asp Tyr Pro Leu Ala Asn Lys His
35 40 45

Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe Thr
50 55 60

Leu Glu Asn Val Leu Ser Asp Asn Val Thr Ala Gln Asp Met Arg Ile
65 70 75 80

Thr Pro Glu Thr Leu Arg Ile Gln Ala Ala Ile Ala Arg Asp Ala Gly
85 90 95

Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr Ser
100 105 110

Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro Tyr
115 120 125

Arg Ser Thr Lys Gln Glu Leu Ile Ala Ile Ala Asp Asp Leu Glu Arg
130 135 140

Arg Tyr Gln Ala Gln Ile Cys Ala Ala Phe Val Arg Glu Ala Ala Glu
145 150 155 160

Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 203
<211> 554
<212> PRT
<213> *Listeria monocytogenes*

525
A

<400> 203

Met Lys Ser Lys Arg Phe Glu Glu Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Leu Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly Lys Val
35 40 45

Val Glu Met Asp Ser Lys Lys Leu Ala Asp Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Lys Tyr Gly Val Asp Leu Ser Arg Ala Glu Glu Val Met
65 70 75 80

Gln Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Pro Arg Glu Lys Ile Val Leu Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met Met Ser
115 120 125

Met Gln Lys Met Arg Ser Arg Arg Thr Pro Thr Thr Gln Ala His Val
130 135 140

Thr Asn Leu Arg Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Ala Ala Ile Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Val
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ser Leu Leu Val Gly Ser Gln Thr
180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Glu Leu Gly Met Arg Gly Leu Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

52.6
JP

Ser Lys Gly Ile Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Thr Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys Ile Gly
290 295 300

Ile Pro Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Ala Val Met Leu Asp Leu Glu Val Ala Ser Gly Asn Asp Gln
325 330 335

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ser Ala Thr Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Asp Asp Phe
370 375 380

Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Thr Pro Val Thr Glu Glu Val Val Ala Val Arg Asn Lys Ala Ala
405 410 415

Arg Val Ile Gln Ala Val Phe Asp Lys Leu Gly Leu Pro Glu Val Thr
420 425 430

Asp Ala Glu Val Glu Ala Ala Thr Tyr Ala Arg Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Met Val Glu Asp Ile Lys Ala Ala Ala Glu Met Met
450 455 460

Asp Arg Gly Val Thr Gly Leu Asp Val Val Lys Ala Leu Ser Ala Gly
465 470 475 480

527

Gly Phe Asp Asp Val Ala Glu Ser Val Leu Asn Met Leu Lys Gln Arg
 485 490 495

Val Ser Gly Asp Phe Leu His Thr Ser Ala Ile Ile Asp Lys Asp Trp
 500 505 510

Asn Val Ile Ser Ser Val Asn Asp Leu Asn Asp Tyr Ala Gly Pro Gly
 515 520 525

Thr Gly Tyr Arg Leu Glu Gly Glu Arg Trp Glu Lys Leu Lys Asp Ile
 530 535 540

Ala Val Ala Val Asp Ala Asn Glu Leu Glu
 545 550

<210> 204

<211> 219

<212> PRT

<213> *Listeria monocytogenes*

<400> 204

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

Leu Asp Glu Leu Gln Leu Lys Glu Asp Lys Val Ser Phe Gln Lys Glu
 20 25 30

Gln Pro Ser Val Ala Val Ser Asp Glu Ser Phe Leu Thr Glu Val Gly
 35 40 45

Asp Ala Glu Pro Gly Arg Gln Lys Asp Glu Val Val Ile Ala Val Ala
 50 55 60

Pro Ala Phe Gly Lys Tyr Gln Thr Lys Asn Ile Val Gly Val Pro His
 65 70 75 80

Lys Gln Ile Leu Arg Glu Val Ile Ala Gly Ile Glu Glu Glu Gly Leu
 85 90 95

Lys Ala Arg Val Val Arg Val Phe Arg Ser Ser Asp Val Ala Phe Val
 100 105 110

Ala Val Glu Gly Asp Lys Leu Ser Gly Ser Gly Ile Cys Ile Gly Ile
 115 120 125

528
JA

Gln Ser Arg Gly Thr Ala Leu Ile His Gln Lys Asp Leu Gln Pro Leu
130 135 140

Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Ile Thr Leu Glu Thr
145 150 155 160

Tyr Arg Ala Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys Gly Glu Ser
165 170 175

Pro Asn Pro Val Pro Met Val Asn Asp Gln Met Ala Arg Pro Lys Phe
180 185 190

Gln Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His Val Val
195 200 205

Gln Gly Lys Asn Ala Val Glu Leu Gln Val Asn
210 215

<210> 205

<211> 170

<212> PRT

<213> *Listeria monocytogenes*

<400> 205

Met Asn Gln Glu Ala Leu Glu Asn Met Val Arg Asn Ile Leu Gln Glu
1 5 10 15

Val Asn Ser Gly Ala Val Ser Thr Thr Thr Ser Gln Lys Val Ser Gly
20 25 30

Asp Thr Leu Thr Val Arg Asp Tyr Pro Leu Gly Thr Lys Arg Pro Glu
35 40 45

Leu Val Lys Thr Ser Thr Ser Lys Ser Leu Asp Asp Ile Thr Leu Lys
50 55 60

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

Glu Thr Leu Lys Met Gln Ala Gln Val Ala Arg Asp Ala Gly Arg Ala
85 90 95

Thr Leu Ala Asn Asn Phe Glu Arg Ala Ala Glu Leu Thr Ile Val Pro
100 105 110

529
AA

Asp Glu Arg Ile Leu Glu Ile Tyr Asn Ala Met Arg Pro Tyr Arg Ser
115 120 125

Ser Arg Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu Ser Val Tyr
130 135 140

His Ala Thr Ile Cys Ser Asn Tyr Val Arg Glu Ala Ala Gln Leu Tyr
145 150 155 160

Gln Glu Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 206

<211> 553

<212> PRT

<213> *Streptococcus sanguinis*

<400> 206

Met Arg Ser Lys Arg Phe Glu Val Leu Ser Glu Arg Pro Ile Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Val Glu Glu Gly Leu Val Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Gln Asn Gly Lys Val
35 40 45

Val Glu Leu Asp Gly Lys Pro Lys Glu Glu Phe Asp Leu Ile Asp Gln
50 55 60

Phe Ile Ala Asn Tyr Gly Ile Asp Leu Ser Leu Ala Glu Glu Val Ile
65 70 75 80

Gln Met Asp Ser Arg Glu Ile Ala Asn Lys Ile Leu Thr Pro Ser Val
85 90 95

Pro Arg Thr Glu Ile Ile Lys Leu Thr Lys Ala Met Thr Pro Ala Lys
100 105 110

Ile Ile Glu Val Val Asn Gln Met Asn Val Val Glu Ile Met Met Cys
115 120 125

Leu Gln Lys Met Arg Thr Arg Lys Gln Thr Ala Thr Gln Ala His Val
130 135 140

930

Thr Asn Val Asn Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Leu Arg Gly Phe Ala Glu Gln Glu Thr Thr Val Ala Val Val
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ser Leu Met Ile Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Glu Phe Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Thr Gly Ser Glu Val Gln Met Gly Gln Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Leu Tyr Val Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Thr Gln Asn Gly Ser Val Ser Cys Ile Gly
290 295 300

Ile Pro Ala Ala Val Pro Ser Gly Ile Arg Ala Val Val Ala Glu Asn
305 310 315 320

Leu Ile Ala Ser Met Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Leu Arg Arg Ser Val Arg Thr Leu Met Gln
340 345 350

Phe Ala Pro Gly Thr Asp Phe Ile Asn Ser Gly Tyr Ser Ala Thr Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Trp Asp Ala Glu Asp Phe
370 375 380

530
A

Asp Asp Tyr Asn Val Leu Gln Arg Asp Leu Arg Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Glu Val Val Lys Val Arg Asn Lys Ala Ala
405 410 415

Arg Val Met Gln Ala Leu Phe Lys Gly Leu Gly Leu Pro Gln Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asp Lys Val Glu Asp Ile Lys Ala Ala Gln Gly Ile Leu
450 455 460

Glu Arg Gly Val Gln Gly Ala Asp Leu Ile Lys Ala Leu Ala Asn Asn
465 470 475 480

Gly Phe Pro Glu Val Ala Asn Glu Leu Leu Asn Leu Phe Lys Gln Arg
485 490 495

Val Ala Gly Asp Phe Leu Gln Thr Ser Ala Ile Phe Asp Arg Asp Trp
500 505 510

Asn Val Ile Ser Ala Val Asn Ser Pro Asn Asp Tyr Val Gly Val Gly
515 520 525

Thr Gly His Arg Leu Val Gly Glu Glu Trp Glu Lys Val Lys Asp Ile
530 535 540

Pro Lys Ala Ile Asp Pro Arg Asp Val
545 550

<210> 207

<211> 222

<212> PRT

<213> *Streptococcus sanguinis*

<400> 207

Met Thr Glu Ile Asn Glu Thr Leu Leu Arg Ser Ile Ile Ala Glu Val
1 5 10 15

Met Lys Glu Met Ser Ala Asn Thr Lys Glu Thr Ala Ala Glu Thr Ser
20 25 30

Glu Lys Pro Val Thr Lys Pro Val Ser Asn Glu Lys Ala Val Ile Arg
35 40 45

532
A

Thr Val Gly Val Ala Lys Pro Ser Gln Ser Thr Asp Glu Val Val Ile
50 55 60

Ala Val Gly Pro Ala Phe Gly Glu Gln Gln Val Lys Thr Met Val Asp
65 70 75 80

Ile Pro His Thr Glu Val Leu Arg Gln Leu Val Ala Gly Ile Glu Glu
85 90 95

Glu Gly Leu Lys Ala Arg Ile Val Lys Val Tyr Arg Ser Ser Asp Val
100 105 110

Ala Phe Val Ala Val Glu Gly Asp His Leu Ser Gly Ser Gly Ile Ser
115 120 125

Ile Gly Val Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
130 135 140

Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu Thr
145 150 155 160

Pro Glu Thr Tyr Arg Leu Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys
165 170 175

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

Pro Lys Tyr Gln Ala Tyr Ser Ala Leu Leu His Ile Lys Glu Thr Lys
195 200 205

Leu Val Lys Arg Gly Lys Pro Ala Asp Glu Cys Gln Val Ile
210 215 220

<210> 208

<211> 171

<212> PRT

<213> *Streptococcus sanguinis*

<400> 208

Met Ser Glu Ser Val Glu Thr Leu Val Lys Gln Ile Leu Ala Glu Leu
1 5 10 15

Ser Asp Ser Gly Ser Ala Ser Gln Gly Ala Val Asn Arg Pro Val Ser
20 25 30

933
AA

Ser Asp Glu Ala Thr Ala Ala Asp Tyr Pro Ile Ser Lys Lys His Pro
35 40 45

Asp Trp Ile Lys Val Gly Gln Asp Lys Lys Phe Glu Asp Ile Thr Leu
50 55 60

Glu Asn Ile Leu Ser Gly Tyr Val Lys Ala Glu Asp Leu Arg Ile Lys
65 70 75 80

Pro Glu Ile Leu Ile Lys Gln Gly Glu Ile Ala Lys Asn Ala Gly Arg
85 90 95

Glu Ala Ile Gln Tyr Asn Phe Ser Arg Ala Ala Glu Leu Thr Lys Val
100 105 110

Pro Asp Ala Arg Val Leu Glu Ile Tyr Asn Ala Leu Arg Pro Tyr Arg
115 120 125

Ser Ser Lys Gln Glu Leu Leu Asp Ile Ala Asn Glu Leu Glu Asn Gln
130 135 140

Tyr Gly Ala Val Ile Cys Ala Gly Phe Val Arg Glu Ala Ala Glu Asn
145 150 155 160

Tyr Glu Arg Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 209

<211> 555

<212> PRT

<213> *Escherichia blattae*

<400> 209

Met Arg Arg Ser Lys Arg Phe Glu Val Leu Glu Lys Arg Pro Val Asn
1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

Gly Ser Pro Trp Asp Pro Pro Ser Ser Val Lys Val Glu Gln Gly Arg
35 40 45

Ile Val Glu Leu Asp Gly Lys Ala Arg Ala Asp Phe Asp Met Ile Asp
50 55 60

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Ile Asp Glu Thr Glu His Ala
65 70 75 80

Met Gly Leu Asp Ala Leu Thr Ile Ala Arg Met Leu Val Asp Ile Asn
85 90 95

Val Ser Arg Ala Glu Ile Ile Lys Val Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Val Glu Val Met Ser His Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Ile Gly Ser Gln
180 185 190

Ser Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Phe Thr Ser Tyr Ala Glu Thr Val
210 215 220

Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ala
260 265 270

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Gly Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Pro Gly Leu Ala Glu
305 310 315 320

535

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
 325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
 340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
 355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
 370 375 380

Phe Asp Gly Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
 385 390 395 400

Leu Arg Pro Val Ser Glu Glu Glu Thr Ile Ala Ile Arg Asn Lys Ala
 405 410 415

Ala Arg Ala Val Gln Ala Val Phe Arg Gly Leu Gly Leu Pro Pro Val
 420 425 430

Thr Asp Glu Glu Val Thr Ala Ala Thr Tyr Ala His Gly Ser Lys Asp
 435 440 445

Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met
 450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Arg Ala Leu Ser Val
 465 470 475 480

Asn Gly Phe Asp Asp Val Ala Asn Asn Ile Leu Asn Met Leu Arg Gln
 485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Glu
 500 505 510

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
 515 520 525

Gly Thr Gly Tyr Arg Ile Ser Pro Gln Arg Trp Glu Glu Ile Lys Asn
 530 535 540

Ile Ala Thr Val Ile Gln Pro Asp Ser Ile Glu
 545 550 555

536
A

<210> 210
 <211> 196
 <212> PRT
 <213> *Escherichia blattae*

<400> 210

Met Glu Thr Thr Gln Lys Lys Ala Pro Val Phe Thr Leu Asn Leu Val
 1 5 10 15

Glu Ser Gly Val Ala Lys Pro Gly Glu Arg Ser Asp Glu Val Val Ile
 20 25 30

Gly Val Gly Pro Ala Phe Asp Lys Tyr Gln His Lys Thr Leu Ile Asp
 35 40 45

Met Pro His Lys Ala Ile Ile Lys Glu Met Val Ala Gly Val Glu Glu
 50 55 60

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
 65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
 85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
 100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
 115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
 130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
 145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
 165 170 175

His Val Val Gly Asp Ala Lys Pro Val Thr Leu Asn Ile Glu Ile Thr
 180 185 190

Arg Glu Glu Ala
 195

<210> 211

537
AP

<211> 140
 <212> PRT
 <213> *Escherichia blattae*

<400> 211

Met Thr Thr Thr Lys Met Ser Ala Ala Asp Tyr Pro Leu Ala Ser Arg
 1 5 10 15

Cys Pro Glu Arg Ile Gln Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
 20 25 30

Thr Leu Glu Asn Val Leu Ala Gly Lys Val Gly Thr Gln Asp Glu Arg
 35 40 45

Ile Ser Arg Glu Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
 50 55 60

His Arg His Ala Ile Ala Arg Asn Leu Arg Arg Ala Gly Glu Leu Ile
 65 70 75 80

Ala Ile Pro Asp Ala Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
 85 90 95

Tyr Arg Ser Ser Val Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu
 100 105 110

Thr Arg Tyr Gln Ala Thr Val Asn Ala Ala Phe Ile Arg Glu Ala Ala
 115 120 125

Glu Val Tyr Arg Gln Arg Asp Lys Leu Arg Lys Glu
 130 135 140

<210> 212
 <211> 554
 <212> PRT
 <213> *Clostridium perfringens*

<400> 212

Met Lys Ser Lys Arg Phe Gln Val Leu Ser Glu Arg Pro Val Asn Gln
 1 5 10 15

Asp Gly Leu Ile Gly Glu Trp Ala Asp Glu Gly Leu Ile Ala Leu Asp
 20 25 30

Ser Pro Asn Asp Pro Lys Ser Ser Ile Lys Ile Glu Asn Gly Ile Ile
 35 40 45

538

Thr Glu Leu Asp Gly Arg Ser Arg Asp Glu Phe Asp Met Ile Asp Lys
 50 55 60

Phe Ile Ala Glu Tyr Ala Ile Asn Ile Glu Asp Ala Glu Ala Ser Met
 65 70 75 80

Lys Leu Ser Ser Lys Glu Ile Ala Arg Arg Leu Val Asp Ile Asn Val
 85 90 95

Ser Arg Asp Glu Ile Val Lys Ile Thr Thr Ser Ile Thr Pro Met Lys
 100 105 110

Ala Val Glu Val Ile Gln Glu Met Asn Val Val Glu Met Met Met Ala
 115 120 125

Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ala Asn Gln Cys His Val
 130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
 145 150 155 160

Ala Ala Leu Arg Gly Phe Ala Glu Gln Glu Thr Thr Val Gly Ile Val
 165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ala Ile Leu Val Gly Ser Gln Val
 180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Cys Ala Val Glu Glu Ala Thr Glu
 195 200 205

Leu Asp Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val Ser
 210 215 220

Val Tyr Gly Thr Glu Ser Val Phe Thr Asp Gly Asp Asp Thr Pro Trp
 225 230 235 240

Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
 245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Ala Leu Met Gly Tyr Ser Glu
 260 265 270

Gly Arg Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Tyr Ile Thr Lys
 275 280 285

539

Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile Gly
290 295 300

Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Gly Glu Asn
305 310 315 320

Leu Ile Ala Ala Met Leu Asp Ile Glu Val Ala Ser Ala Asn Asp Gln
325 330 335

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Met Leu Met Gln
340 345 350

Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Thr Glu Glu Glu Thr Ile Lys Val Arg Asn Lys Ala Ala
405 410 415

Lys Cys Ile Gln Ile Ile Phe Arg Glu Leu Gly Phe Pro Glu Val Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Cys His Gly Ser Lys Glu Met
435 440 445

Pro Asn Arg Asn Val Val Glu Asp Leu Lys Ala Ala Glu Glu Met Leu
450 455 460

Glu Arg Arg Ile Thr Gly Leu Asp Ile Ile Lys Ala Leu Ser Lys Asn
465 470 475 480

Gly Met Glu Asp Ile Ala Asn Asn Leu Leu Asn Met Leu Lys Gln Arg
485 490 495

Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Lys Asp Phe
500 505 510

Asn Val Ile Ser Ala Val Asn Asp Val Asn Asp Tyr Met Gly Pro Gly
515 520 525

540
AS

Thr Gly Tyr Arg Leu Asp Gly Gln Arg Trp Glu Glu Ile Lys Lys Val
530 535 540

Pro Thr Val Met Arg Pro Glu Asp Ile Glu
545 550

<210> 213

<211> 190

<212> PRT

<213> *Clostridium perfringens*

<400> 213

Met Thr Met Glu Glu Arg Thr Phe Ile Pro Glu Ile Thr Val Glu Glu
1 5 10 15

Val Gly Glu Ala Lys Val Gly Leu Arg Ser Asp Glu Val Val Ile Gly
20 25 30

Leu Ala Pro Ala Phe Leu Lys Tyr Gln Asn Lys Thr Ile Val Asp Val
35 40 45

Pro His Thr Glu Thr Leu Leu Glu Ile Ile Ala Gly Ile Glu Glu Glu
50 55 60

Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val Ser
65 70 75 80

Phe Ile Ala His Asp Ala Ala Cys Leu Ser Gly Ser Gly Ile Gly Ile
85 90 95

Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Lys Asp Leu Leu
100 105 110

Pro Leu Asn Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr Thr
115 120 125

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

Glu Ser Pro Thr Pro Val Pro Val Lys Asn Asp Gln Met Val Arg Pro
145 150 155 160

Lys Phe Met Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His
165 170 175

541
AS

Val Glu Pro Gly Lys Lys Pro Val Gln Leu Glu Val Lys Phe
180 185 190

<210> 214

<211> 141

<212> PRT

<213> *Clostridium perfringens*

<400> 214

Met Glu Asn Lys Arg Met Thr Ala Ala Asp Tyr Pro Leu Thr Ser Lys
1 5 10 15

Arg Lys Gly Asp Ile Lys Thr Pro Thr Gly Lys Ala Leu Glu Asp Ile
20 25 30

Thr Leu Glu Lys Val Leu Ser Gly Glu Ile Asn Ala Asp Asp Ile Arg
35 40 45

Ile Ser Pro Glu Thr Leu Glu Met Gln Ala Gln Ile Ala Glu Ser Met
50 55 60

Asn Arg Asp Ala Ile Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80

Arg Val Pro Asp Asp Arg Ile Leu Glu Met Tyr Asn Ala Leu Arg Pro
85 90 95

Tyr Arg Ser Thr Lys Glu Asp Leu Phe Lys Ile Ala Asp Glu Leu Glu
100 105 110

Thr Lys Tyr Asp Ala Lys Val Asn Ala Asp Phe Val Arg Glu Ala Ala
115 120 125

Glu Val Tyr Glu Thr Arg Asn Lys Leu Arg Ile Glu Glu
130 135 140

<210> 215

<211> 408

<212> PRT

<213> *Yersinia frederiksenii*

<400> 215

Met Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu Gly Ala
1 5 10 15

Phe Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala Arg Tyr
20 25 30

542

Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val Gly Arg
35 40 45

Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu Leu Lys
50 55 60

Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser Val Tyr
65 70 75 80

Gly Thr Glu Pro Val Phe Thr Asp Gly Asp Asp Thr Pro Trp Ser Lys
85 90 95

Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met Arg Phe
100 105 110

Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu Gly Lys
115 120 125

Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys Ala Ala
130 135 140

Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Cys Val Gly Val Pro
145 150 155 160

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

Cys Ser Ala Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln Thr Phe
180 185 190

Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln Phe Leu
195 200 205

Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro Asn Tyr
210 215 220

Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe Asp Asp
225 230 235 240

Tyr Asn Val Met Gln Arg Asp Leu Lys Val Asp Gly Gly Leu Arg Pro
245 250 255

Val Leu Glu Ala Asp Val Val Ala Ile Arg Asn Lys Ala Ala Arg Ala
260 265 270

943
113

Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr Asp Glu
275 280 285

Glu Val Ile Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met Pro Glu
290 295 300

Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile Ser Lys
305 310 315 320

Asn Arg Thr Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly Gly Phe
325 330 335

Glu Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys Ile Ala
340 345 350

Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Lys Gly Asp Asn Gln Val
355 360 365

Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala Thr Gly
370 375 380

Tyr Arg Leu Glu Gly Ala Arg Trp Glu Glu Ile Lys Asn Ile Pro Asn
385 390 395 400

Ala Leu Asp Pro Asn Glu Leu Gly
405

<210> 216

<211> 224

<212> PRT

<213> *Yersinia frederiksenii*

<400> 216

Met Val Asp Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Gly Val
1 5 10 15

Leu Gln Glu Met Gln Gly Asp Gln Asn Thr Val Ser Phe Lys Gln Glu
20 25 30

Thr Gln Pro Ala Ala Ala Val Asn Thr Thr Ala Ser Gly Asp Phe Leu
35 40 45

Thr Glu Val Gly Glu Ala Arg Pro Gly Thr His Gln Asp Glu Val Ile
50 55 60

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

544

Gly Ile Pro His Lys Asn Ile Leu Arg Glu Leu Ile Ala Gly Ile Glu
85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Leu Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 217
<211> 175
<212> PRT
<213> *Yersinia frederiksenii*

<400> 217

Met Asn Ser Glu Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Lys
1 5 10 15

Met Asn Ser Leu Gln Gly Gln Ser Pro Ala Ala Ala Cys Pro Ala Pro
20 25 30

Ala Ala Ser Ser Arg Ser Asp Ala Lys Val Ser Asp Tyr Pro Leu Ala
35 40 45

Asn Lys His Pro Asp Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp
50 55 60

545
AB

Asp Leu Thr Leu Ala Asn Val Leu Asn Gly Ser Val Thr Ser Gln Asp
65 70 75 80

Leu Arg Ile Thr Pro Glu Ile Leu Arg Ile Gln Ala Ser Ile Ala Lys
85 90 95

Asp Ala Gly Arg Pro Leu Leu Ala Met Asn Phe Glu Arg Ala Ala Glu
100 105 110

Leu Thr Ala Val Pro Asp Asp Lys Val Leu Asp Ile Tyr Asn Ala Leu
115 120 125

Arg Pro Phe Arg Ser Ser Lys Asp Glu Leu Asn Ala Ile Ala Asp Asp
130 135 140

Leu Glu Lys Thr Tyr Lys Ala Thr Ile Cys Ala Ala Phe Val Arg Glu
145 150 155 160

Ala Ala Val Leu Tyr Val Gln Arg Lys Lys Leu Lys Gly Asp Asp
165 170 175

<210> 218

<211> 554

<212> PRT

<213> *Thermoanaerobacter ethanolicus*

<400> 218

Met Arg Ser Lys Arg Phe Glu Val Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Ile Ile Ala Glu Trp Pro Glu Val Gly Leu Ile Ala Val Asn
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly Lys Ile
35 40 45

Val Glu Met Asp Gly Lys Lys Arg Glu Glu Phe Asp Met Ile Glu Gln
50 55 60

Phe Ile Ala Asp Tyr Ala Ile Asn Ile Glu Met Ala Glu Lys Ala Met
65 70 75 80

Gly Met Asp Ser Leu Glu Ile Ala Arg Met Leu Val Asp Ile Asn Val
85 90 95

Pro Arg Glu Glu Ile Val Lys Ile Val Ser Gly Leu Thr Pro Ala Lys
100 105 110

546
JA

Val Val Glu Val Val Asn His Leu Asn Val Val Glu Ile Met Met Ala
115 120 125

Ile Gln Lys Met Arg Ala Arg Lys Arg Pro Gly Asn Gln Ala His Val
130 135 140

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

Ala Ala Leu Arg Gly Phe Asp Glu Leu Glu Thr Thr Val Gly Ile Thr
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln Thr
180 185 190

Gly Arg Gly Gly Val Leu Thr Gln Asp Ala Leu Glu Glu Ser Phe Glu
195 200 205

Leu Thr Ile Gly Met Arg Gly Phe Thr Ser Tyr Ala Glu Thr Ile Ser
210 215 220

Val Tyr Gly Thr Glu Gln Val Phe Val Asp Gly Asp Asp Thr Pro Trp
225 230 235 240

Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Thr Gly Ser Glu Cys Gln Met Gly Phe Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Val Met Ile Ala Lys
275 280 285

Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys Ile Gly
290 295 300

Val Pro Gly Ala Val Pro Gly Gly Ile Arg Ala Val Ala Ala Glu Asn
305 310 315 320

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

Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Leu Leu
340 345 350

547
A

Met Leu Pro Gly Thr Asp Leu Ile Phe Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Leu Met Leu Gln Arg Asp Leu Met Val Glu Gly Gly Val
385 390 395 400

Ser Pro Val Thr Glu Asp Glu Val Ile Ala Val Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ser Val Phe Lys Lys Leu Asn Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Pro Arg Asp Val Asn Ala Asp Leu Lys Ala Ala Thr Glu Leu Met
450 455 460

Glu Lys Gly Leu Thr Gly Leu Asp Val Val Lys Ala Leu Ala Glu Ser
465 470 475 480

Gly Phe Thr Asp Val Ala Glu Asn Val Leu Asn Met Leu Lys Gln Arg
485 490 495

Ile Ala Gly Asp Tyr Leu Gln Thr Ser Ala Val Leu Asp Lys Asp Phe
500 505 510

Asn Ile Asp Ser Ala Val Asn Asn Ser Asn Asp Tyr Lys Gly Pro Gly
515 520 525

Thr Gly Phe Arg Leu Ser Lys Glu Arg Trp Glu Arg Ile Lys Asn Val
530 535 540

Pro Gln Ala Leu Lys Pro Glu Asp Leu Ala
545 550

<210> 219

<211> 229

<212> PRT

<213> *Thermoanaerobacter ethanolicus*

<400> 219

Met Val Lys Thr Glu Ser Leu Val Glu Gln Ile Val Lys Glu Val Leu
1 5 10 15

548
A

Lys Lys Leu Glu Asn Val Glu Ile Ala Ala Pro Ala Thr Gln Ser Ser
20 25 30

Asp Asp Ala Asn Gln Glu Trp Glu Met Ile Ile Glu Glu Ile Gly Glu
35 40 45

Ala Lys Gln Gly Val Asn Val Asp Glu Val Val Ile Gly Val Ser Pro
50 55 60

Gly Phe Tyr Ile Lys Phe Lys Lys Asn Ile Ile Gly Ile Pro Leu Gly
65 70 75 80

Asn Ile Leu Arg Glu Ile Ile Ser Gly Ile Thr Glu Gln Gly Leu Lys
85 90 95

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

His Thr Ala Ala Lys Leu Ser Gly Ser Gly Ile Gly Ile Gly Ile Gln
115 120 125

Ser Arg Gly Thr Val Val Ile His Gln Lys Asp Leu Gln Pro Leu Asn
130 135 140

Asn Leu Glu Leu Phe Pro Gln Cys Pro Val Leu Thr Leu Glu Thr Tyr
145 150 155 160

Arg Ala Ile Gly Arg Asn Ala Ala Leu Tyr Ala Lys Gly Glu Ser Pro
165 170 175

Thr Pro Val Pro Val Gln Asn Asp Gln Met Ala Arg Pro Lys Tyr Gln
180 185 190

Ala Ile Ala Ala Val Met His Asn Phe Glu Thr Lys Tyr Val Gln Thr
195 200 205

Gly Ala Lys Pro Val Glu Leu Lys Val Ser Phe Ala Arg Lys Gly Gly
210 215 220

Asn Lys Ser Asp Arg
225

<210> 220

<211> 182

<212> PRT

<213> *Thermoanaerobacter ethanolicus*

549
SP

<400> 220

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

Ile Ile Val Arg Glu Val Leu Thr Asn Leu Thr Ser Asp Lys Gly Thr
20 25 30

Gln Asn Gln Gln Lys Thr Ala Ser Ser Ser Leu Pro Lys Leu Asp Pro
35 40 45

Lys Arg Asp Tyr Pro Leu Ala Lys Asn Lys Pro Glu Leu Ala Lys Ser
50 55 60

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

Gly Lys Val Leu Pro Asp Asp Leu Lys Ile Ser Pro Glu Thr Leu Leu
85 90 95

Ala Gln Ala Glu Ile Ala Glu Ala Ala Gly Arg Lys Gln Leu Ala Asn
100 105 110

Asn Phe Arg Arg Ala Ala Glu Leu Thr Lys Val Pro Asp Lys Arg Ile
115 120 125

Leu Glu Ile Tyr Asn Ala Leu Arg Pro Tyr Arg Ser Thr Lys Glu Glu
130 135 140

Leu Leu Ala Ile Ala Asp Glu Leu Asp Asn Ala Tyr Gly Ala Lys Val
145 150 155 160

Cys Ala Ala Phe Val Arg Glu Ala Ala Glu Val Tyr Glu Arg Arg Gly
165 170 175

Arg Leu Lys Gly Met Glu
180

<210> 221

<211> 558

<212> PRT

<213> *Lactobacillus hilgardii*

<400> 221

Met Lys Arg Gln Lys Arg Phe Glu Lys Leu Glu Lys Arg Pro Val His
1 5 10 15

Leu Asp Gly Phe Val Lys Glu Trp Asp Asp Glu Gly Leu Val Ala Met
20 25 30

Glu Gly Lys Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly Val
35 40 45

Val Thr Glu Leu Asp Gly Lys Lys Lys Asp Asp Phe Asp Leu Ile Asp
50 55 60

Gln Tyr Ile Ala Glu Tyr Gly Ile Asn Leu Asn Asn Ala Glu Lys Val
65 70 75 80

Met Lys Met Asp Ser Leu Lys Ile Ala Lys Met Leu Val Asp Pro Asn
85 90 95

Asp Ser Arg Ser Glu Ile Ile Gln Leu Thr Thr Ala Met Thr Pro Ala
100 105 110

Lys Ala Glu Glu Val Ile Ser Lys Leu Asn Phe Ala Glu Met Ile Met
115 120 125

Ala Ala Gln Lys Met Arg Pro Arg Arg Thr Pro Ala Thr Gln Cys His
130 135 140

Ile Thr Asn Thr Leu Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Ala Leu Arg Gly Phe Pro Glu Gln Glu Thr Thr Thr Ala Val
165 170 175

Ala Arg Tyr Ala Pro Met Asn Ala Ile Ser Ile Met Val Gly Ser Gln
180 185 190

Thr Gly Arg Pro Gly Val Ile Thr Gln Cys Ser Val Glu Glu Ser Glu
195 200 205

Glu Leu Ser Leu Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile
210 215 220

Ser Val Tyr Gly Thr Asp Arg Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
245 250 255

551
AA

Met Arg Phe Thr Ser Gly Ala Gly Ser Glu Ala Met Met Gly Tyr Thr
260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Phe Ile Thr
275 280 285

Lys Ala Ser Gly Val Gln Gly Leu Gln Asn Gly Gly Val Ser Cys Ile
290 295 300

Gly Ile Pro Gly Ala Val Pro Gly Gly Ile Arg Glu Val Leu Gly Glu
305 310 315 320

Asn Leu Leu Cys Met Met Leu Asp Ile Glu Cys Ala Ser Gly Asn Asp
325 330 335

Gln Ala Phe Ser His Ser Asp Ile Arg Arg Thr Glu Arg Met Ile Gly
340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ala Ala Glu
355 360 365

Glu Asn Arg Asp Asn Thr Phe Ala Gly Ser Asn Met Asp Val Leu Asp
370 375 380

Tyr Asp Asp Tyr Cys Ser Met Glu Arg Asp Leu Ala Val Asn Gly Ser
385 390 395 400

Ile Leu Pro Ile His Glu Glu Asp Ala Ile Lys Ile Arg Asn Arg Ala
405 410 415

Ala Lys Ala Ile Gln Ala Val Phe Asp Gly Leu Gly Leu Pro Pro Ile
420 425 430

Ser Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala Ser Asn Ser Gly Asp
435 440 445

Met Pro Lys Arg Asp Met Val Gln Asp Ile Lys Ala Ala Gln Asp Leu
450 455 460

Met Asn Arg Asp Ile Thr Ile Ser Asp Ile Ile Lys Ala Leu Tyr Asn
465 470 475 480

His Gly Phe Lys Asp Val Ala Gln Ala Val Leu Thr Leu Ala Gln Gln
485 490 495

Lys Val Cys Gly Asp Tyr Leu Gln Thr Ser Ala Ile Phe Asn Gly Lys
500 505 510

Trp His Cys Val Ser Ala Ile Asn Asp Ala Asn Asp Tyr Glu Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Val Trp Glu Asp Lys Glu Gln Trp Lys Lys Leu
530 535 540

Lys Asp Ile Pro Trp Ala Val Asp Pro Gln His Met Asn Phe
545 550 555

<210> 222

<211> 217

<212> PRT

<213> *Lactobacillus hilgardii*

<400> 222

Met Asn Asp Thr Asp Gln Ala Ile Ser Phe Lys Ala Asp Ser Thr Thr
1 5 10 15

Ala Thr Asp Thr Ala Thr Ala Ala Lys Pro Ser Thr Ser Gln Glu Ser
20 25 30

Asp Ile Lys Pro Val Asp Trp Phe Lys His Val Gly Val Ala Lys Pro
35 40 45

Gly Phe Ser Lys Asp Glu Val Val Ile Ala Val Ala Pro Ala Phe Ala
50 55 60

Glu Val Leu Thr Lys Thr Met Thr Lys Ile Ser His Arg Asp Val Leu
65 70 75 80

Arg Gln Val Ile Ala Gly Ile Glu Glu Glu Gly Ile Lys Ala Arg Val
85 90 95

Ile Lys Val Tyr Arg Thr Ser Asp Val Ser Phe Cys Ser Val Glu Ala
100 105 110

Asp Lys Leu Ser Gly Ser Gly Ile Ala Ile Ala Ile Gln Ser Lys Gly
115 120 125

Thr Thr Ile Ile His Gln Arg Asp Gln Glu Pro Leu Asn Asn Leu Glu
130 135 140

Leu Phe Pro Gln Ala Pro Val Leu Thr Leu Asp Thr Tyr Arg Ala Ile
145 150 155 160

553
P

Gly Lys Asn Ala Ala Gln Tyr Ala Arg Gly Met Ser Pro Asn Pro Val
165 170 175

Pro Thr Val Asn Asp Gln Met Ala Arg Val Gln Tyr Gln Ala Leu Ser
180 185 190

Thr Leu Met His Ile Gln Glu Thr Lys Gln Val Val Met Gly Lys Pro
195 200 205

Ala Asp Glu Ile Gln Val Ser Val Glu
210 215

<210> 223

<211> 176

<212> PRT

<213> *Lactobacillus hilgardii*

<400> 223

Met Ser Glu Val Asp Glu Leu Val Ser Lys Ile Met Asn Glu Leu Lys
1 5 10 15

Asp Gly Gln Ser Ser Ser Thr Pro Asn Ser Thr Pro Glu Ser Ala Pro
20 25 30

Ser Gln Val Pro Thr Gly Lys Lys Leu Gly Lys Ser Asp Tyr Pro Leu
35 40 45

Tyr Glu Lys His Pro Lys Asp Val Lys Ser Pro Thr Gly Lys Ser Leu
50 55 60

Asn Glu Ile Thr Leu Glu Asn Ile Ile Asn Gly Asn Val Thr Ser Lys
65 70 75 80

Asp Leu Arg Ile Thr Pro Asp Thr Leu Arg Met Gln Gly Gln Ile Ala
85 90 95

Ala Ser Ala Gly Arg Val Ala Ile Gln Arg Asn Phe Gln Arg Ala Ala
100 105 110

Gly Leu Thr Ala Ile Pro Asp Asp Arg Val Leu Ala Leu Tyr Asn Ser
115 120 125

Leu Arg Ala Tyr Arg Ser Phe Lys Gln Glu Leu Leu Asp Thr Ala Asn
130 135 140

Glu Leu Arg Thr Lys Tyr His Ala Pro Val Cys Ala Gly Trp Phe Glu
145 150 155 160

554
A

Glu Ala Ala Glu Asn Tyr Glu Lys Asp Arg Lys Leu Lys Gly Asp Asn
165 170 175

<210> 224

<211> 558

<212> PRT

<213> *Lactobacillus reuteri*

<400> 224

Met Lys Arg Gln Lys Arg Phe Glu Glu Leu Glu Lys Arg Pro Ile His
1 5 10 15

Gln Asp Thr Phe Val Lys Glu Trp Pro Glu Glu Gly Phe Val Ala Met
20 25 30

Met Gly Pro Asn Asp Pro Lys Pro Ser Val Lys Val Glu Asn Gly Lys
35 40 45

Ile Val Glu Met Asp Gly Lys Lys Leu Glu Asp Phe Asp Leu Ile Asp
50 55 60

Leu Tyr Ile Ala Lys Tyr Gly Ile Asn Ile Asp Asn Val Glu Lys Val
65 70 75 80

Met Asn Met Asp Ser Thr Lys Ile Ala Arg Met Leu Val Asp Pro Asn
85 90 95

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

Lys Ala Glu Glu Ile Ile Ser Lys Leu Asp Phe Gly Glu Met Ile Met
115 120 125

Ala Val Lys Lys Met Arg Pro Arg Arg Lys Pro Asp Asn Gln Cys His
130 135 140

Val Thr Asn Thr Val Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Asp Ala Ala Leu Arg Gly Phe Pro Glu Gln Glu Thr Thr Thr Ala Val
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Ile Ser Ile Leu Ile Gly Ala Gln
180 185 190

Thr Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

555
 AP

Glu Leu Gln Leu Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile
 210 215 220

Ser Val Tyr Gly Thr Asp Arg Val Phe Thr Asp Gly Asp Asp Thr Pro
 225 230 235 240

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
 245 250 255

Met Arg Phe Thr Ser Gly Ala Gly Ser Glu Val Leu Met Gly Tyr Pro
 260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Leu Leu Thr
 275 280 285

Lys Ala Ser Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
 290 295 300

Glu Ile Pro Gly Ala Val Pro Asn Gly Ile Arg Glu Val Leu Gly Glu
 305 310 315 320

Asn Leu Leu Cys Met Met Cys Asp Ile Glu Cys Ala Ser Gly Cys Asp
 325 330 335

Gln Ala Tyr Ser His Ser Asp Met Arg Arg Thr Glu Arg Phe Ile Gly
 340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Ile Asn Ser Gly Tyr Ser Ser Thr
 355 360 365

Pro Asn Tyr Asp Asn Thr Phe Ala Gly Ser Asn Thr Asp Ala Met Asp
 370 375 380

Tyr Asp Asp Met Tyr Val Met Glu Arg Asp Leu Gly Gln Tyr Tyr Gly
 385 390 395 400

Ile His Pro Val Lys Glu Glu Thr Ile Ile Lys Ala Arg Asn Lys Ala
 405 410 415

Ala Lys Ala Leu Gln Ala Val Phe Glu Asp Leu Gly Leu Pro Lys Ile
 420 425 430

Thr Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala Asn Thr His Asp Asp
 435 440 445

556

Met Pro Lys Arg Asp Met Val Ala Asp Met Lys Ala Ala Gln Asp Met
450 455 460

Met Asp Arg Gly Ile Thr Ala Ile Asp Ile Ile Lys Ala Leu Tyr Asn
465 470 475 480

His Gly Phe Lys Asp Val Ala Glu Ala Ile Leu Asn Leu Gln Lys Gln
485 490 495

Lys Val Val Gly Asp Tyr Leu Gln Thr Ser Ser Ile Phe Asp Lys Asp
500 505 510

Trp Asn Val Thr Ser Ala Val Asn Asp Gly Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Leu Tyr Glu Asp Lys Glu Glu Trp Asp Arg Ile
530 535 540

Lys Asp Leu Pro Phe Ala Leu Asp Pro Glu His Leu Glu Leu
545 550 555

<210> 225

<211> 236

<212> PRT

<213> *Lactobacillus reuteri*

<400> 225

Met Ala Asp Ile Asp Glu Asn Leu Leu Arg Lys Ile Val Lys Glu Val
1 5 10 15

Leu Ser Glu Thr Asn Gln Ile Asp Thr Lys Ile Asp Phe Asp Lys Ser
20 25 30

Asn Asp Ser Thr Ala Thr Ala Thr Gln Glu Val Gln Gln Pro Asn Ser
35 40 45

Lys Ala Val Pro Glu Lys Lys Leu Asp Trp Phe Gln Pro Val Gly Glu
50 55 60

Ala Lys Pro Gly Tyr Ser Lys Asp Glu Val Val Ile Ala Val Gly Pro
65 70 75 80

Ala Phe Ala Thr Val Leu Asp Lys Thr Glu Thr Gly Ile Pro His Lys
85 90 95

Glu Val Leu Arg Gln Val Ile Ala Gly Ile Glu Glu Glu Gly Leu Lys
100 105 110

557

Ala Arg Val Val Lys Val Tyr Arg Ser Ser Asp Val Ala Phe Cys Ala
115 120 125

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

Ser Lys Gly Thr Thr Val Ile His Gln Lys Asp Gln Asp Pro Leu Gly
145 150 155 160

Asn Leu Glu Leu Phe Pro Gln Ala Pro Val Leu Thr Pro Glu Thr Tyr
165 170 175

Arg Ala Ile Gly Lys Asn Ala Ala Met Tyr Ala Lys Gly Glu Ser Pro
180 185 190

Glu Pro Val Pro Ala Lys Asn Asp Gln Leu Ala Arg Ile His Tyr Gln
195 200 205

Ala Ile Ser Ala Ile Met His Ile Arg Glu Thr His Gln Val Val Val
210 215 220

Gly Lys Pro Glu Glu Glu Ile Lys Val Thr Phe Asp
225 230 235

<210> 226

<211> 172

<212> PRT

<213> *Lactobacillus reuteri*

<400> 226

Met Met Ser Glu Val Asp Asp Leu Val Ala Lys Ile Met Ala Gln Met
1 5 10 15

Gly Asn Ser Ser Ser Ala Asn Ser Ser Thr Gly Thr Ser Thr Ala Ser
20 25 30

Thr Ser Lys Glu Met Thr Ala Asp Asp Tyr Pro Leu Tyr Gln Lys His
35 40 45

Arg Asp Leu Val Lys Thr Pro Lys Gly His Asn Leu Asp Asp Ile Asn
50 55 60

Leu Gln Lys Val Val Asn Asn Gln Val Asp Pro Lys Glu Leu Arg Ile
65 70 75 80

Thr Pro Glu Ala Leu Lys Leu Gln Gly Glu Ile Ala Ala Asn Ala Gly
85 90 95

558

Arg Pro Ala Ile Gln Lys Asn Leu Gln Arg Ala Ala Glu Leu Thr Arg
100 105 110

Val Pro Asp Glu Arg Val Leu Glu Met Tyr Asp Ala Leu Arg Pro Phe
115 120 125

Arg Ser Thr Lys Gln Glu Leu Leu Asn Ile Ala Lys Glu Leu Arg Asp
130 135 140

Lys Tyr Asp Ala Asn Val Cys Ala Ala Trp Phe Glu Glu Ala Ala Asp
145 150 155 160

Tyr Tyr Glu Ser Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 227

<211> 558

<212> PRT

<213> *Lactobacillus diolivorans*

<400> 227

Met Lys Arg Gln Lys Arg Phe Glu Lys Leu Glu Lys Arg Pro Val His
1 5 10 15

Leu Asp Gly Phe Val Lys Glu Trp Asp Asp Glu Gly Leu Val Ala Met
20 25 30

Glu Gly Lys Asn Asp Pro Lys Pro Ser Val Lys Val Glu Asn Gly Val
35 40 45

Val Thr Glu Leu Asp Gly Lys Lys Lys Ala Asp Phe Asp Leu Ile Asp
50 55 60

Lys Tyr Ile Ala Glu Tyr Gly Leu Asn Ile Asp Lys Ala Glu Glu Val
65 70 75 80

Met Ala Met Asp Ser Thr Lys Ile Ala Lys Met Leu Val Asp Pro Asn
85 90 95

Val Pro Arg Ser Glu Ile Val Lys Leu Thr Thr Ala Met Thr Pro Ala
100 105 110

Lys Ala Glu Glu Val Ile Ser Lys Leu Asn Phe Gly Glu Met Ile Met
115 120 125

599

Ala Val Gln Lys Met Arg Pro Arg Arg Thr Pro Ala Thr Gln Cys His
130 135 140

Ile Thr Asn Thr Arg Asp Asn Pro Val Glu Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Ser Leu Arg Gly Phe Pro Glu Gln Glu Thr Thr Thr Ala Thr
165 170 175

Ala Arg Tyr Ala Pro Leu Asn Ala Ile Ser Ile Met Val Gly Ser Gln
180 185 190

Thr Gly Arg Pro Gly Val Ile Ser Gln Cys Ser Val Glu Gly Ser Glu
195 200 205

Glu Leu Ser Leu Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile
210 215 220

Ser Val Tyr Gly Thr Asp Arg Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Phe Thr Ser Gly Ala Gly Ser Glu Val Met Met Gly Tyr Thr
260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ser Gly Val Gln Gly Leu Gln Asn Gly Gly Val Ser Cys Ile
290 295 300

Gly Ile Pro Gly Ala Val Pro Ser Gly Leu Arg Gly Val Leu Gly Glu
305 310 315 320

Asn Leu Leu Cys Met Met Ile Asp Val Glu Cys Ala Ser Gly Cys Asp
325 330 335

Gln Ala Phe Ser His Ser Asp Ile Arg Arg Thr Glu Arg Met Ile Gly
340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ala Ala Glu
355 360 365

560

Glu Asn Arg Asp Asn Thr Phe Ala Gly Ser Asn Met Asp Val Leu Asp
370 375 380

Tyr Asp Asp Tyr Cys Ser Met Glu Arg Asp Leu Ser Ile Asn Gly Gly
385 390 395 400

Ile Val Pro Ile His Glu Glu Asp Ala Ile Lys Ile Arg Asn Lys Gly
405 410 415

Ala Lys Ala Ile Gln Ala Val Phe Asp Gly Leu Gly Leu Pro Lys Ile
420 425 430

Thr Asp Glu Glu Val Glu Ala Ala Thr Tyr Gly Ser Asn Ser Asp Asp
435 440 445

Met Pro Lys Arg Asp Met Val Gln Asp Met Lys Ala Ala Gln Asp Leu
450 455 460

Met Asp Arg Gly Thr Thr Ile Ser Asp Val Ile Lys Ala Leu Tyr Asp
465 470 475 480

His Asp Phe Lys Asp Val Ala Glu Ala Val Leu Lys Leu Ala Gln Gln
485 490 495

Lys Val Cys Gly Asp Tyr Leu Gln Thr Ser Ala Ile Phe Asp Gly Asp
500 505 510

Trp Asn Cys Ile Ser Ala Val Asn Asp Thr Asn Asp Tyr Met Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Leu Trp Glu Asp Lys Pro Gln Asp Gln Arg Tyr
530 535 540

Ser Met Gly Asn Gly Ser Thr Thr Tyr Gly Leu Leu Ile Asp
545 550 555

<210> 228

<211> 245

<212> PRT

<213> *Lactobacillus diolivorans*

<400> 228

Met Lys Asp Thr Asp Thr Pro Ile Ser Phe Gly Lys Lys Asp Ser Ser
1 5 10 15

Ala Asn His Glu Ser Asn Gln Leu Ala Leu His Gly Asp Ser Trp Phe
20 25 30

561
AA

Lys Ser Gly Asp Asp Pro Val Ser Asn Ser Ser Ser Ala Thr Ala Thr
35 40 45

Ala Ser Ala Pro Ser Gln Ser Thr Ser Gly Asn Gly Asp Ile Lys Ser
50 55 60

Leu Asp Trp Phe Lys His Val Gly Val Ala Lys Pro Gly Leu Ser Lys
65 70 75 80

Asp Glu Val Val Ile Gly Val Ala Pro Ala Phe Ala Glu Val Leu Thr
85 90 95

Lys Thr Met Thr Lys Ile Pro His Lys Asp Val Leu Arg Gln Ile Ile
100 105 110

Ala Gly Ile Glu Glu Glu Gly Leu Lys Ala Arg Val Val Lys Val Tyr
115 120 125

Arg Thr Ser Asp Val Ser Phe Ile Ala Asp Glu Val Asp Lys Leu Ser
130 135 140

Gly Ser Gly Ile Ala Val Ala Val Gln Ser Lys Gly Thr Ala Ile Ile
145 150 155 160

His Gln Lys Asp Gln Glu Pro Leu Asn Asn Leu Glu Leu Phe Pro Gln
165 170 175

Ala Pro Val Ile Asp Leu Pro Thr Tyr Arg Ala Ile Gly Lys Asn Ala
180 185 190

Ala Gln Tyr Ala Lys Gly Leu Ser Pro Asp Pro Val Pro Thr Val Asn
195 200 205

Asp Gln Met Ala Arg Val Gln Tyr Thr Ala Leu Ser Ser Leu Met His
210 215 220

Ile Ala Glu Thr Lys Arg Val Val Val Gly Lys Pro Ala Glu Glu Ile
225 230 235 240

Gln Val Thr Phe Asn
245

<210> 229

<211> 180

<212> PRT

<213> *Lactobacillus diolivorans*

562
/s

<400> 229

Met Ser Glu Val Asp Glu Leu Val Ser Lys Ile Leu Ser Gln Leu Asn
1 5 10 15

Asn Gly Asp Ser Ser Gln Ser Thr Gly Thr Thr Thr Ala Ser Ser Val
20 25 30

Thr Pro Thr Ser Ser Pro Ala Asn Ser Gly Lys Thr Phe Asp Lys Gly
35 40 45

Asp Tyr Pro Leu Phe Arg Lys His Pro Asp Asp Val Lys Thr Pro Thr
50 55 60

Gly Lys Ala Val Ser Asp Ile Thr Leu Asp Asn Val Val Ser Gly Lys
65 70 75 80

Val Asp Ser Lys Asp Leu Arg Ile Thr Ala Asn Thr Leu Arg Arg Gln
85 90 95

Gly Glu Ile Ala Ala Ser Ala Gly Arg Pro Ala Ile Gln Arg Asn Phe
100 105 110

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

Phe Tyr His Ala Leu Arg Pro Phe Arg Ser Ser Lys Gln Asp Leu Leu
130 135 140

Asp Ile Ala Lys Gln Leu Arg Asp Thr Tyr His Ala Pro Val Cys Ala
145 150 155 160

Asn Trp Phe Glu Glu Ala Ala Gly Asn Tyr Glu Ile Ser Lys Lys Leu
165 170 175

Lys Gly Asp Asn
180

<210> 230
<211> 558
<212> PRT
<213> *Lactobacillus reuteri*

<400> 230

Met Lys Arg Gln Lys Arg Phe Glu Glu Leu Glu Lys Arg Pro Ile His
1 5 10 15

563

Gln Asp Thr Phe Val Lys Glu Trp Pro Glu Glu Gly Phe Val Ala Met
20 25 30

Met Gly Pro Asn Asp Pro Lys Pro Ser Val Lys Val Glu Asn Gly Lys
35 40 45

Ile Val Glu Met Asp Gly Lys Lys Leu Glu Asp Phe Asp Leu Ile Asp
50 55 60

Leu Tyr Ile Ala Lys Tyr Gly Ile Asn Ile Asp Asn Val Glu Lys Val
65 70 75 80

Met Asn Met Asp Ser Thr Lys Ile Ala Arg Met Leu Val Asp Pro Asn
85 90 95

Val Ser Arg Glu Ser Ile Ile Glu Ile Thr Ser Ala Leu Thr Pro Ala
100 105 110

Lys Ala Glu Glu Ile Ile Ser Lys Leu Asp Phe Gly Glu Met Ile Met
115 120 125

Ala Ile Lys Lys Met Arg Pro Arg Arg Lys Pro Asp Asn Gln Cys His
130 135 140

Val Thr Asn Thr Val Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Asp Ala Ala Leu Arg Gly Phe Pro Glu Gln Glu Thr Thr Thr Ala Val
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Ile Ser Ile Leu Ile Gly Ala Gln
180 185 190

Thr Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Gln Leu Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile
210 215 220

Ser Val Tyr Gly Thr Asp Arg Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
245 250 255

564
A

Met Arg Phe Thr Ser Gly Ala Gly Ser Glu Val Leu Met Gly Tyr Pro
260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Leu Leu Thr
275 280 285

Lys Ala Ser Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Glu Ile Pro Gly Ala Val Pro Asn Gly Ile Arg Glu Val Leu Gly Glu
305 310 315 320

Asn Leu Leu Cys Met Met Cys Asp Ile Glu Cys Ala Ser Gly Cys Asp
325 330 335

Gln Ala Tyr Ser His Ser Asp Met Arg Arg Thr Glu Arg Phe Ile Gly
340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Ile Asn Ser Gly Tyr Ser Ser Thr
355 360 365

Pro Asn Tyr Asp Asn Thr Phe Ala Gly Ser Asn Thr Asp Ala Met Asp
370 375 380

Tyr Asp Asp Met Tyr Val Met Glu Arg Asp Leu Gly Gln Tyr Tyr Gly
385 390 395 400

Ile His Pro Val Gln Glu Glu Thr Ile Ile Lys Ala Arg Asn Lys Ala
405 410 415

Ala Lys Ala Leu Gln Ala Val Phe Glu Asp Leu Gly Leu Pro Lys Ile
420 425 430

Thr Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala Asn Thr His Asp Asp
435 440 445

Met Pro Lys Arg Asp Met Val Ala Asp Met Lys Ala Ala Gln Asp Met
450 455 460

Met Asp Arg Gly Ile Thr Ala Val Asp Ile Ile Lys Ala Leu Tyr Asn
465 470 475 480

His Gly Phe Lys Asp Val Ala Glu Ala Val Leu Asn Leu Gln Lys Gln
485 490 495

565
AS

Lys Val Val Gly Asp Tyr Leu Gln Thr Ser Ser Ile Phe Asp Lys Asp
500 505 510

Trp Asn Ile Thr Ser Ala Val Asn Asp Gly Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Leu Tyr Glu Asp Lys Glu Glu Trp Asp Arg Ile
530 535 540

Lys Asp Leu Pro Phe Ala Leu Asp Pro Glu His Leu Glu Leu
545 550 555

<210> 231

<211> 236

<212> PRT

<213> *Lactobacillus reuteri*

<400> 231

Met Ala Asp Ile Asp Glu Asn Leu Leu Arg Lys Ile Val Lys Glu Val
1 5 10 15

Leu Asn Glu Thr Asn Gln Ile Asp Thr Lys Ile Asn Phe Asp Lys Glu
20 25 30

Asn Asn Ser Thr Ala Thr Ala Thr Glu Glu Val Gln Gln Pro Asn Ser
35 40 45

Lys Ala Val Pro Glu Lys Lys Leu Asp Trp Phe Gln Pro Ile Gly Glu
50 55 60

Ala Lys Pro Gly Tyr Ser Lys Asp Glu Val Val Ile Ala Val Gly Pro
65 70 75 80

Ala Phe Ala Thr Val Leu Asp Lys Thr Glu Thr Gly Ile Pro His Lys
85 90 95

Glu Val Leu Arg Gln Val Ile Ala Gly Ile Glu Glu Glu Gly Leu Lys
100 105 110

Ala Arg Val Val Lys Val Tyr Arg Ser Ser Asp Val Ala Phe Cys Ala
115 120 125

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

566
JP

Ser Lys Gly Thr Thr Val Ile His Gln Lys Asp Gln Asp Pro Leu Gly
145 150 155 160

Asn Leu Glu Leu Phe Pro Gln Ala Pro Val Leu Thr Pro Glu Thr Phe
165 170 175

Arg Ala Ile Gly Lys Asn Ala Ala Met Tyr Ala Lys Gly Glu Ser Pro
180 185 190

Glu Pro Val Pro Ala Lys Asn Asp Gln Leu Ala Arg Ile His Tyr Gln
195 200 205

Ala Ile Ser Ala Ile Met His Ile Arg Glu Thr His Gln Val Val Val
210 215 220

Gly Lys Pro Glu Glu Glu Ile Lys Val Thr Phe Asp
225 230 235

<210> 232

<211> 171

<212> PRT

<213> *Lactobacillus reuteri*

<400> 232

Met Ser Glu Val Asp Asp Leu Val Ala Lys Ile Met Ala Gln Met Gly
1 5 10 15

Asn Ser Ser Ser Ser Asp Ser Ser Thr Ser Ala Thr Ser Thr Asn Asn
20 25 30

Gly Lys Glu Met Thr Ala Asp Asp Tyr Pro Leu Tyr Gln Lys His Arg
35 40 45

Asp Leu Val Lys Thr Pro Ser Gly Lys Lys Leu Asp Asp Ile Thr Leu
50 55 60

Gln Lys Val Val Asn Asp Gln Val Asn Pro Lys Glu Leu Arg Ile Thr
65 70 75 80

Pro Glu Ala Leu Lys Leu Gln Gly Glu Ile Ala Ala Asn Ala Gly Arg
85 90 95

Pro Ala Ile Gln Lys Asn Leu Gln Arg Ala Ala Glu Leu Thr Arg Val
100 105 110

Pro Asp Glu Arg Val Leu Gln Met Tyr Asp Ala Leu Arg Pro Phe Arg
115 120 125

567
AA

Ser Thr Lys Gln Glu Leu Leu Asp Ile Ala Asn Glu Leu Arg Asp Lys
130 135 140

Tyr His Ala Glu Val Cys Ala Ala Trp Phe Glu Glu Ala Ala Asp Tyr
145 150 155 160

Tyr Glu Ser Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 233

<211> 756

<212> PRT

<213> *Mesorhizobium loti*

● <400> 233

Met Thr Pro Lys Leu Asn Arg Trp Lys Arg Phe Ala Asp Trp Asp Glu
1 5 10 15

Arg Pro Leu Arg Leu Asp Lys Phe Ala Ala Glu Asp Pro Ala Asn Gly
20 25 30

Phe Ser Ala Phe Ser Ser Pro Ala Asp Pro Lys Pro Gly Ile Gly Ile
35 40 45

Lys Asp Gly Arg Val Ile Ser Leu Asp Gly Val Leu Glu His Asp Tyr
50 55 60

Asp Met Ile Asp Arg Phe Ile Ala Arg His His Ile Asp Pro Glu Val
65 70 75 80

Ala Pro Glu Ala Met Ala Leu Asp Ser Ala Thr Val Ala Arg Trp Leu
85 90 95

Val Asp Met Asn Val Pro Arg Glu Thr Leu Val Arg Leu Ala His Gly
100 105 110

Met Thr Pro Ala Lys Leu Ala Glu Val Val Ser Gln Leu Asn Ala Leu
115 120 125

Glu Ile Ala Phe Ala Tyr Ser Lys Met Arg Ala Arg Lys Thr Pro Gly
130 135 140

Asn Gln Ala His Val Thr Asn Ala Lys Asp Asp Pro Leu Gln Leu Ala
145 150 155 160

Ala Asp Ala Ala Thr Ala Val Ala Phe Gly Phe Asp Glu Ile Glu Thr
165 170 175

568
A

Thr Met Arg Val Ser Arg Asn Ala Trp Ser Asn Ala Val Ala Cys Ala
180 185 190

Val Gly Gly Ala Val Gly Arg Trp Gly Thr Leu Phe Gln Cys Ser Ser
195 200 205

Glu Glu Ala Glu Glu Leu Arg Ile Ala Met Ala Gly Phe Thr Ser Tyr
210 215 220

Ala Glu Thr Val Ser Val Tyr Gly Thr Glu Lys Ser Phe Thr Asp Gly
225 230 235 240

Asp Asp Thr Pro Trp Ser Lys Ala Phe Leu Ala Ala Ala Tyr Ala Ser
245 250 255

Arg Gly Val Lys Met Arg Cys Thr Ser Gly Ala Gly Ser Glu Leu Leu
260 265 270

Met Gly Phe His Glu Ala Lys Ser Leu Leu Tyr Leu Glu Ala Arg Cys
275 280 285

Leu Cys Leu Gln Arg Gly Met Gly Val Gln Gly Thr Gln Asn Gly Gly
290 295 300

Ile Asp Gly Ala Pro Leu Thr Ala Thr Ile Pro Gly Gly Val Arg Glu
305 310 315 320

Leu Met Ala Glu Asn Leu Ile Ala Val Trp Leu Asp Leu Glu Cys Ala
325 330 335

Ser Gly Asn Asp Ala Arg Ser Thr Glu Ser Glu Ile Arg Val Gly Ala
340 345 350

Lys Ile Leu Pro Tyr Leu Ile Ala Gly Ser Asp Leu Ile Cys Ser Gly
355 360 365

Met Gly Ser Ile Leu Lys Tyr Asp Asn Ser Phe Asn Pro Ser Leu Ile
370 375 380

Asn Gly Glu Glu Leu Glu Asp Tyr Leu Val Leu Gln Arg Asp Phe Glu
385 390 395 400

Ala Asp Gly Gly Leu Thr Pro Leu Pro Glu Ser Arg Ala Ile Glu Leu
405 410 415

569
H

Arg Glu Arg Ala Val Glu Ala Ile Ala Ala Val Phe Glu Glu Leu Gly
420 425 430

Leu Ser Ser Pro Thr Glu Asp Met Lys Ser Ser Val Val Tyr Ala Ser
435 440 445

Gly Ser Asp Asp Thr Arg Ser Leu Met Pro Arg Asp Val Ser Phe Ile
450 455 460

Ser Glu Ala Ile Lys Glu Arg Gly Ile Thr Val Ile Asp Val Val Lys
465 470 475 480

Ala Leu Ala Lys Arg Gly Phe Arg Glu Glu Ala Glu Asn Leu Leu Asp
485 490 495

Val Val Lys Leu Arg Leu Ser Gly Asp Tyr Leu Gln Thr Ser Ala Met
500 505 510

Ile Arg Asn Gly Arg Ile Val Ser Ala Val Asn Asp Pro Asn Asp Tyr
515 520 525

Leu Gly Pro Gly Ser Gly Tyr Arg Val Ser Glu Glu Arg Arg Leu Gln
530 535 540

Leu Asn Asp Ile Arg Asp Val Leu Asp Gln Lys Glu Val Leu Arg Ser
545 550 555 560

Glu Ala Leu His Glu Lys Asp Glu Ala Arg His Ile Arg Tyr Arg Asn
565 570 575

Leu Gly Pro Ala Ala Asn Gly Ser Ala Lys Asp Asp Val Val Ile Gly
580 585 590

Ile Ser Pro Ala Phe Gly Leu Lys Leu Tyr Arg Thr Thr Ala Gly His
595 600 605

Arg Leu Ser Glu Val Leu Gly Ala Met Leu Asp Ala Ile His Ala Arg
610 615 620

Gly Leu Lys Ala Arg Val Val Arg Phe Arg His Thr Ala Asp Thr Ser
625 630 635 640

Phe Leu Gly Leu Ser Ala Ala Arg Leu Ala Gly Ser Gly Ile Gly Ile
645 650 655

570
A

Gly Ile Gln Ala Lys Gly Thr Ala Val Ile His Gln Arg Asp Arg Gln
660 665 670

Pro His Asn Asn Leu Glu Leu Phe Ser Asn Ala Pro Ile Thr Arg Leu
675 680 685

Glu His Tyr Arg Ala Leu Gly Ala Asn Ala Ala Ala Tyr Ala Leu Gly
690 695 700

Glu Met Pro Glu Pro Val Val Val Pro Gln Arg Gly Glu Ala Met Gly
705 710 715 720

Ser Arg Tyr His Ala Arg Val Ala Leu Ile Tyr Ala Ile Glu Thr Gly
725 730 735

Leu Thr Glu Ala Gly Ala Ala Pro Glu Glu Val Asp Val Ile Leu Thr
740 745 750

Gly Val Lys Ser
755

<210> 234

<211> 139

<212> PRT

<213> *Mesorhizobium loti*

<400> 234

Met Thr His Thr Arg Ala Asp Tyr Pro Leu Ala Glu Thr Gln Pro Gly
1 5 10 15

Ala Val Lys Gly Lys Arg Gly Lys Ser Leu Ala Glu Ile Thr Leu Asp
20 25 30

Ser Val Leu Ala Gly Asp Val Thr Met Glu Asp Leu Arg Ile Thr Pro
35 40 45

Gln Ala Leu Gln Ala Gln Ala Asp Val Ala Arg Asp Val Gly Arg Pro
50 55 60

Thr Leu Ala Leu Asn Phe Glu Arg Gly Ala Glu Leu Val Glu Val Pro
65 70 75 80

Gln Asp Phe Ile Met Gln Val Tyr Glu Leu Leu Arg Pro Gly Arg Ala
85 90 95

571
A

Lys Ser Lys Glu Glu Leu Leu Gln Ala Ala Thr Thr Met Arg Asp Thr
100 105 110

Tyr Gln Ala Glu Arg Ile Ala Arg Phe Ile Glu Glu Ala Ala Glu Thr
115 120 125

Tyr Ala Ala Arg Gly Leu Phe Thr Phe Arg Phe
130 135

<210> 235

<211> 746

<212> PRT

<213> *Mycobacterium vanbaalenii*

<400> 235

Met Ala Asp Glu Leu Gly Arg Phe Arg Val Leu Asn Ser Lys Pro Val
1 5 10 15

Asn Leu Asp Gly Phe Ser Val Pro Asp Ala Gly Leu Gly Leu Val Ala
20 25 30

Met Ser Ser Pro His Asp Pro Ala Pro Ser Leu Lys Ile Arg Gly Gly
35 40 45

Glu Val Val Glu Leu Asp Gly Lys Gly Ala Gly Glu Phe Asp Val Ile
50 55 60

Asp Glu Phe Ile Ala Arg Tyr Gly Ile Asp Leu Thr Val Ala Glu Glu
65 70 75 80

Ala Met Ala Leu Gly Asp Glu Thr Leu Ala Arg Met Val Val Asp Ile
85 90 95

Asn Val Pro Arg Ala Glu Val Val Arg Leu Ile Gly Gly Thr Thr Pro
100 105 110

Ala Lys Leu Ala Arg Val Val Ala Leu Leu Ser Pro Val Glu Met Gln
115 120 125

Met Ala Met Ala Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Ala
130 135 140

His Val Thr Asn Gln Leu Asp Asp Pro Leu Leu Ile Ala Ala Asp Ala
145 150 155 160

572
A

Ala Ser Ala Val Ala Tyr Gly Phe Arg Glu Val Glu Thr Thr Val Pro
165 170 175

Val Leu Gly Asp Ala Pro Ser Asn Ala Val Ala Leu Leu Ile Gly Ser
180 185 190

Gln Val Gly Ser Pro Gly Ala Met Ala Gln Cys Ser Ile Glu Glu Ala
195 200 205

Leu Glu Leu Arg Leu Gly Leu Arg Gly Leu Thr Ser Tyr Ala Glu Thr
210 215 220

Ile Ser Ile Tyr Gly Thr Glu Gln Val Phe Val Asp Gly Asp Asp Thr
225 230 235 240

Pro Phe Ser Lys Ala Ile Leu Thr Ser Ala Tyr Ala Ser Arg Gly Leu
245 250 255

Lys Met Arg Val Thr Ser Gly Gly Gly Ala Glu Val Leu Met Gly Ala
260 265 270

Ala Glu Lys Cys Ser Ile Leu Tyr Leu Glu Ser Arg Cys Val Ser Leu
275 280 285

Ala Arg Ala Leu Gly Ser Gln Gly Val Gln Asn Gly Gly Ile Asp Gly
290 295 300

Val Gly Val Val Ala Ser Val Pro Glu Gly Met Lys Glu Leu Leu Ala
305 310 315 320

Glu Asn Leu Met Val Met Met Arg Asp Leu Glu Ser Cys Ala Gly Asn
325 330 335

Asp Asn Leu Ile Ser Glu Ser Asp Ile Arg Arg Ser Ala His Thr Leu
340 345 350

Pro Val Leu Leu Ala Gly Ala Asp Phe Val Phe Ser Gly Phe Gly Ser
355 360 365

Ile Pro Arg Tyr Asp Asn Ala Phe Ala Leu Ser Asn Phe Asn Ser Asp
370 375 380

Asp Met Asp Asp Phe Leu Val Leu Gln Arg Asp Trp Gly Ala Asp Gly
385 390 395 400

573/11

Gly Leu Arg Thr Val Ser Pro Glu His Leu Glu Ala Val Arg Arg Arg
 405 410 415

Ala Ala Lys Ala Val Gln Ala Val Tyr Arg Asp Leu Gly Leu Ala Asp
 420 425 430

Tyr Glu Asp Ala Arg Val Glu Glu Val Val Ala Ala Asn Gly Ser Arg
 435 440 445

Asp Leu Pro Ala Gly His Pro Lys Met Val Ala Glu Ala Ala Ala Ser
 450 455 460

Ile Glu Ala Arg Gln Leu Thr Val Phe Asp Val Ile Ala Ser Leu His
 465 470 475 480

Arg Thr Gly Phe Thr Asp Glu Ala Glu Ala Ile Thr Thr Leu Thr Arg
 485 490 495

Glu Arg Leu Arg Gly Asp Gln Leu Gln Thr Ser Ala Ile Phe Asp Glu
 500 505 510

Lys Phe Arg Val Leu Ser Lys Leu Thr Asp Pro Asn Asp Tyr Thr Gly
 515 520 525

Pro Ala Thr Gly Tyr Ala Leu Thr Asp Arg Arg Arg Ala Glu Ile Asp
 530 535 540

Ala Ile Arg Gln Ala Arg Ser Ser Ala Glu Leu Thr Ala Asp Gln Glu
 545 550 555 560

Ser Tyr Arg Gly His Val Leu Val Thr Asp Val Glu Pro Ala Gln Gln
 565 570 575

Gly Ser Asp Pro Arg Glu Val Cys Ile Gly Leu Ser Pro Ala Trp Gly
 580 585 590

Arg Ser Val Trp Leu Thr Leu Cys Gly Leu Thr Ile Gly Glu Val Leu
 595 600 605

Arg Gln Ile Ser Ala Gly Leu Glu Glu Glu Gly Cys Ile Ala Arg Pro
 610 615 620

Val Arg Val Arg Ser Thr Ile Asp Val Gly Leu Ile Gly Leu Thr Ala
 625 630 635 640

574

Ala Arg Leu Ser Gly Ser Gly Ile Gly Ile Gly Leu Gln Gly Lys Gly
645 650 655

Thr Ala Leu Ile His Arg Arg Asp Leu Ala Pro Leu Ala Asn Leu Glu
660 665 670

Leu Phe Ser Val Ala Pro Leu Leu Thr Ala Lys Met Tyr Arg Glu Leu
675 680 685

Gly Lys Asn Ala Ala Arg His Ala Lys Gly Met Ala Pro Val Pro Ile
690 695 700

Phe Thr Gly Gly Thr Asp Glu Ser Ile Ser Ala Arg Tyr His Ala Arg
705 710 715 720

Ala Val Ala Leu Val Ala Leu Glu Arg Glu Ser Cys Glu Pro Gly Gln
725 730 735

Pro Pro Val Thr Val Lys Val Glu Trp Pro
740 745

<210> 236

<211> 119

<212> PRT

<213> *Mycobacterium vanbaalenii*

<400> 236

Met Ala Met Ser Glu Ile Ser Ala Ala Ser Arg Glu Asn Ile Thr Val
1 5 10 15

Gly Asn Ala Val Asp Gly Lys Leu Gly Leu Gly Asp Leu Arg Met Asp
20 25 30

Pro Ala Thr Leu Ala His Gln Ala Ala Val Ala Glu Ala Gly Gly Asn
35 40 45

Pro Gln Leu Ala Glu Asn Phe Arg Arg Ala Ala Glu Leu Ala Thr Ile
50 55 60

Glu Asp Glu Gln Val Met Ala Leu Tyr Glu Ala Leu Arg Pro His Arg
65 70 75 80

Ser Thr Ala Ala Glu Leu Glu Glu Leu Arg Ala Ser Leu Leu Ala Arg
85 90 95

Gly Ala Pro Arg Cys Ala Ala Leu Val Glu Gln Ala Ala Ala Val Tyr
100 105 110

Ala Arg Arg Gly Leu Leu Arg
115

<210> 237

<211> 739

<212> PRT

<213> *Mycobacterium sp.*

<400> 237

Met Arg Ile Leu Asp Ala Lys Pro Val Asn Leu Asp Gly Phe Ser Val
1 5 10 15

Thr Asp Pro Ala Leu Gly Leu Val Ala Met His Ser Pro His Asp Pro
20 25 30

Gln Pro Ser Leu Val Val Arg Asp Gly Arg Val Val Glu Leu Asp Gly
35 40 45

Arg Pro Ala Ala Asp Phe Asp Val Ile Asp Glu Phe Ile Ala Arg Tyr
50 55 60

Gly Ile Asp Leu Thr Val Ala Glu Glu Ala Met Ala Leu Asp Asp Ala
65 70 75 80

Val Leu Ala Arg Met Ala Val Asp Val Asn Val Pro Arg Ala Glu Val
85 90 95

Val Arg Leu Ile Gly Gly Thr Thr Pro Ala Lys Leu Ala Arg Val Met
100 105 110

Ala Val Met Thr Pro Val Glu Met Gln Met Ala Met His Lys Met Arg
115 120 125

Ala Arg Arg Thr Pro Ser Asn Gln Ala His Val Thr Asn Gln Leu Asp
130 135 140

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

Phe Arg Glu Val Glu Thr Thr Val Pro Val Phe Gly Asp Ala Pro Ser
165 170 175

576
AA

Asn Ala Ile Ala Leu Leu Ile Gly Ser Gln Val Gly Val Pro Gly Ala
180 185 190

Met Ala Gln Cys Ser Ile Glu Glu Ala Met Glu Leu Arg Leu Gly Leu
195 200 205

Arg Gly Leu Thr Ser Tyr Ala Glu Thr Ile Ser Ile Tyr Gly Thr Glu
210 215 220

Gln Val Phe Val Asp Gly Asp Asp Thr Pro Phe Ser Lys Ala Ile Leu
225 230 235 240

Thr Ala Ala Tyr Ala Ser Arg Gly Leu Lys Met Arg Val Thr Ser Gly
245 250 255

Gly Gly Ala Glu Val Leu Met Gly Ala Ala Glu Lys Cys Ser Ile Leu
260 265 270

Tyr Leu Glu Ser Arg Cys Val Ser Leu Ala Arg Ala Leu Gly Ser Gln
275 280 285

Gly Val Gln Asn Gly Gly Ile Asp Gly Val Gly Val Val Ala Ser Val
290 295 300

Pro Glu Gly Met Lys Glu Leu Leu Ala Glu Asn Leu Met Val Met Met
305 310 315 320

Arg Asp Leu Glu Ser Cys Ala Gly Asn Asp Asn Leu Ile Ser Glu Ser
325 330 335

Asp Ile Arg Arg Ser Ala His Thr Leu Pro Val Leu Leu Ala Gly Ala
340 345 350

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

Phe Ala Leu Ser Asn Phe Asn Ala Asp Asp Met Asp Asp Phe Leu Val
370 375 380

Leu Gln Arg Asp Trp Gly Ala Asp Gly Gly Leu Arg Thr Val Ser Arg
385 390 395 400

Glu His Leu Ala Arg Val Arg Arg Arg Ala Ala Thr Ala Val Gln Ala
405 410 415

587
AP

Val Tyr Arg Asp Leu Gly Leu Ala Asp Phe Asp Asp Thr Arg Ile Asp
420 425 430

Ala Val Val Val Ala Asn Asp Ser Arg Asp Leu Pro Ala Gly Asp Pro
435 440 445

Lys Ala Val Ala Glu Ala Ala Thr Ala Ile Glu Ala Arg Gln Leu Thr
450 455 460

Val Phe Asp Val Val Ala Ala Leu His Arg Thr Gly Tyr Ala Pro Glu
465 470 475 480

Ala Glu Ala Ile Met Arg Leu Thr Arg Glu Arg Leu Arg Gly Asp Gln
485 490 495

Leu Gln Thr Ser Ala Ile Phe Asp Asp Gln Phe Gln Val Leu Ser Lys
500 505 510

Ile Thr Asp Pro Asn Asp Tyr Ala Gly Pro Gly Ser Gly Tyr Thr Pro
515 520 525

Thr Glu Lys Arg Arg Ala Glu Ile Asp Gly Ile Arg Gln Ala Arg Thr
530 535 540

Ser Ala Glu Leu Thr Ala Asp Gln Ala Glu His Arg Gly His Val Val
545 550 555 560

Phe Ser Asp Val Glu Pro Ala His Gln Gly Ser Asp Pro Arg Glu Val
565 570 575

Cys Ile Gly Leu Ser Pro Ala Leu Gly Arg Ser Val Trp Leu Thr Leu
580 585 590

Cys Gly Leu Thr Val Gly Glu Val Leu Arg Gln Leu Ser Ala Gly Leu
595 600 605

Glu Glu Glu Gly Cys Val Pro Arg Leu Val Arg Val Arg Ser Thr Ile
610 615 620

Asp Val Gly Leu Ile Gly Leu Thr Ala Ala Arg Leu Ser Gly Ser Gly
625 630 635 640

Ile Gly Ile Gly Leu Gln Gly Lys Gly Thr Ala Leu Ile His Arg Arg
645 650 655

578
P

Asp Leu Ala Pro Leu Ala Asn Leu Glu Leu Phe Ser Val Ala Pro Leu
660 665 670

Leu Thr Ala Lys Met Tyr Arg Glu Leu Gly Arg Asn Ala Ala Arg His
675 680 685

Ala Lys Gly Met Ala Pro Leu Pro Ile Leu Ala Gly Gly Thr Asp Glu
690 695 700

Ser Ile Ser Ala Arg Tyr His Ala Arg Ala Val Ala Leu Val Ala Leu
705 710 715 720

Glu Arg Gln Ala Cys Glu Pro Gly Gln Ala Pro Ile Thr Val Glu Ala
725 730 735

Lys Arg Val

<210> 238
<211> 110
<212> PRT
<213> *Mycobacterium sp.*

<400> 238

Met Thr Glu Lys Phe Thr Val Ala Ala Ala Val Asp Gly Lys Leu Thr
1 5 10 15

Leu Ser Asp Leu Arg Met Asp Pro Ala Thr Leu Ala Tyr Gln Ala Val
20 25 30

Val Ala Glu Gln Asp Gly Asn Pro Gln Leu Ala Glu Asn Phe Leu Arg
35 40 45

Ala Ala Glu Leu Ala Val Ile Asp Asp Glu Ala Val Met Lys Phe Tyr
50 55 60

Glu Ala Leu Arg Pro His Arg Ser Thr Ala Ala Glu Leu Glu Glu Leu
65 70 75 80

Arg Val Ser Leu Glu Thr Gly Gly Ala Ser Arg Cys Ala Glu Leu Val
85 90 95

Arg Gln Ala Ala Glu Val Tyr Ala Arg Arg Gly Leu Leu Arg
100 105 110

579
P

<210> 239

<211> 746

<212> PRT

<213> *Mycobacterium flavescens*

<400> 239

Met Ala Asp Glu Leu Gly Arg Phe Arg Val Leu Asn Ser Lys Pro Val
 1 5 10 15

Asn Leu Asp Gly Phe Ser Val Pro Asp Ala Ala Leu Gly Leu Val Ala
 20 25 30

Met Ser Ser Pro His Asp Pro Ala Pro Ser Leu Val Ile Arg Asp Gly
 35 40 45

Ala Val Val Glu Leu Asp Ser Lys Asp Val Ala Glu Phe Asp Val Ile
 50 55 60

Asp Glu Phe Ile Ala Arg Tyr Gly Ile Asp Leu Ser Val Ala Glu Glu
 65 70 75 80

Ala Met Ala Leu Asp Asp Glu Thr Leu Ala Arg Met Val Val Asp Ile
 85 90 95

Asn Val Pro Arg Ala Gly Val Val Arg Leu Ile Gly Gly Thr Thr Pro
 100 105 110

Ala Lys Leu Ala Arg Val Val Ala Leu Leu Ser Pro Val Glu Met Gln
 115 120 125

Met Ala Met Val Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Ala
 130 135 140

His Val Thr Asn Gln Leu Asp Asp Pro Leu Leu Ile Ala Ala Asp Ala
 145 150 155 160

Ala Ser Ala Val Ala Tyr Gly Phe Arg Glu Val Glu Thr Thr Val Pro
 165 170 175

Val Leu Gly Asp Ala Pro Ser Asn Ala Val Ala Leu Leu Ile Gly Ser
 180 185 190

Gln Val Gly Thr Pro Gly Ala Met Ala Gln Cys Ser Ile Glu Glu Ala
 195 200 205

580
PA

Leu Glu Leu Arg Leu Gly Leu Arg Gly Leu Thr Ser Tyr Ala Glu Thr
210 215 220

Ile Ser Ile Tyr Gly Thr Glu Gln Val Phe Ile Asp Gly Asp Asp Thr
225 230 235 240

Pro Phe Ser Lys Ala Ile Leu Thr Ser Ala Tyr Ala Ser Arg Gly Leu
245 250 255

Lys Met Arg Val Thr Ser Gly Gly Gly Ala Glu Val Leu Met Gly Ala
260 265 270

Ala Gln Lys Cys Ser Ile Leu Tyr Leu Glu Ser Arg Cys Val Ser Leu
275 280 285

Ala Arg Ala Leu Gly Ser Gln Gly Val Gln Asn Gly Gly Ile Asp Gly
290 295 300

Val Gly Val Val Ala Ser Val Pro Glu Gly Met Lys Glu Leu Leu Ala
305 310 315 320

Glu Asn Leu Met Val Met Met Arg Asp Leu Glu Ser Cys Ala Gly Asn
325 330 335

Asp Asn Leu Ile Ser Glu Ser Asp Ile Arg Arg Ser Ala His Thr Leu
340 345 350

Pro Val Leu Leu Ala Gly Ala Asp Phe Ile Phe Ser Gly Phe Gly Ser
355 360 365

Ile Pro Arg Tyr Asp Asn Ala Phe Ala Leu Ser Asn Phe Asn Ser Asp
370 375 380

Asp Met Asp Asp Phe Leu Val Leu Gln Arg Asp Trp Gly Ala Asp Gly
385 390 395 400

Gly Leu Arg Thr Val Ser Pro Glu His Leu Glu Ala Val Arg Arg Arg
405 410 415

Ala Ala Gln Ala Val Gln Ala Val Tyr Arg Asp Leu Gly Leu Ala Asp
420 425 430

Tyr Asp Asp Ala Arg Val Glu Asp Val Val Ala Ala Asn Gly Ser Arg
435 440 445

581

Asp Leu Pro Ala Gly His Pro Lys Met Val Ala Glu Ala Ala Ala Ser
 450 455 460

Ile Glu Ala Arg Gln Leu Thr Val Phe Asp Val Ile Ala Ala Leu His
 465 470 475 480

Arg Thr Gly Phe Thr Glu Glu Ala Glu Ala Ile Thr Thr Leu Thr Arg
 485 490 495

Glu Arg Leu Arg Gly Asp Gln Leu Gln Thr Ser Ala Ile Phe Asp Glu
 500 505 510

Gln Phe Arg Val Phe Ser Lys Leu Thr Asp Pro Asn Asp Tyr Arg Gly
 515 520 525

Pro Ala Thr Gly Tyr Ala Leu Thr Glu Gln Arg Arg Ala Glu Ile Asp
 530 535 540

Ala Ile Arg Gln Ala Arg Ser Gly Ala Glu Leu Thr Val Asp Gln Glu
 545 550 555 560

Glu His Arg Gly His Val Val Val Thr Asp Val Glu Pro Ala His Gln
 565 570 575

Gly Ser Asp Pro Arg Glu Val Cys Ile Gly Leu Ser Pro Ala Trp Gly
 580 585 590

Arg Ser Val Trp Leu Thr Leu Cys Gly Leu Thr Val Gly Glu Val Leu
 595 600 605

Arg Gln Ile Ala Ala Gly Leu Glu Glu Glu Gly Cys Ile Ala Arg Thr
 610 615 620

Val Arg Val Cys Ser Thr Ile Asp Thr Gly Leu Ile Gly Leu Thr Ala
 625 630 635 640

Ala Arg Leu Ser Gly Ser Gly Ile Gly Ile Gly Leu Gln Gly Lys Gly
 645 650 655

Thr Ala Leu Ile His Arg Arg Asp Leu Ala Pro Leu Ala Asn Leu Glu
 660 665 670

Leu Phe Ser Val Ala Pro Leu Leu Thr Ala Lys Met Tyr Arg Glu Leu
 675 680 685

582
AA

Gly Lys Asn Ala Ala Arg His Ala Lys Gly Met Ala Pro Val Pro Ile
690 695 700

Phe Thr Gly Gly Thr Asp Glu Ser Ile Ser Ala Arg Tyr His Ala Arg
705 710 715 720

Ala Val Ala Leu Val Ala Leu Glu Arg Asp Ala Cys Glu Pro Gly Arg
725 730 735

Ala Pro Val Thr Val Lys Val Glu Ala Arg
740 745

<210> 240

<211> 109

<212> PRT

<213> *Mycobacterium flavescens*

<400> 240

Met Ser Glu Ile Thr Thr Gln Asn Ala Val Asp Gly Lys Leu Gly Leu
1 5 10 15

Gly Asp Leu Arg Met Asp Pro Ala Val Leu Glu His Gln Ala Ala Val
20 25 30

Ala Glu Glu Gly Gly Asn Pro Gln Leu Ala Glu Asn Phe Arg Arg Ala
35 40 45

Ala Glu Leu Ala Thr Ile Asp Asp Asp Thr Val Met Ala Leu Tyr Glu
50 55 60

Ala Leu Arg Pro His Arg Ser Thr Ala Ala Glu Leu Asp Glu Leu His
65 70 75 80

Ser Ser Leu Ile Ala Arg Gly Ala Pro Arg Cys Ala Ala Leu Val Glu
85 90 95

Gln Ala Ala Ala Val Tyr Ala Arg Arg Gly Leu Leu Arg
100 105

<210> 241

<211> 739

<212> PRT

<213> *Mycobacterium sp.*

<400> 241

Met Arg Ile Leu Asp Ala Lys Pro Val Asn Leu Asp Gly Phe Ser Val
1 5 10 15

583

Thr Asp Pro Ala Leu Gly Leu Val Ala Met His Ser Pro His Asp Pro
20 25 30

Gln Pro Ser Leu Val Val Arg Asp Gly Arg Val Val Glu Leu Asp Gly
35 40 45

Arg Pro Ala Ala Asp Phe Asp Val Ile Asp Glu Phe Ile Ala Arg Tyr
50 55 60

Gly Ile Asp Leu Thr Val Ala Glu Glu Ala Met Ala Leu Asp Asp Ala
65 70 75 80

Val Leu Ala Arg Met Ala Val Asp Val Asn Val Pro Arg Ala Glu Val
85 90 95

Val Arg Leu Ile Gly Gly Thr Thr Pro Ala Lys Leu Ala Arg Val Met
100 105 110

Ala Val Met Thr Pro Val Glu Met Gln Met Ala Met His Lys Met Arg
115 120 125

Ala Arg Arg Thr Pro Ser Asn Gln Ala His Val Thr Asn Gln Leu Asp
130 135 140

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

Phe Arg Glu Val Glu Thr Thr Val Pro Val Phe Gly Asp Ala Pro Ser
165 170 175

Asn Ala Ile Ala Leu Leu Ile Gly Ser Gln Val Gly Val Pro Gly Ala
180 185 190

Met Ala Gln Cys Ser Ile Glu Glu Ala Met Glu Leu Arg Leu Gly Leu
195 200 205

Arg Gly Leu Thr Ser Tyr Ala Glu Thr Ile Ser Ile Tyr Gly Thr Glu
210 215 220

Gln Val Phe Val Asp Gly Asp Asp Thr Pro Phe Ser Lys Ala Ile Leu
225 230 235 240

Thr Ala Ala Tyr Ala Ser Arg Gly Leu Lys Met Arg Val Thr Ser Gly
245 250 255

584

Gly Gly Ala Glu Val Leu Met Gly Ala Ala Glu Lys Cys Ser Ile Leu
260 265 270

Tyr Leu Glu Ser Arg Cys Val Ser Leu Ala Arg Ala Leu Gly Ser Gln
275 280 285

Gly Val Gln Asn Gly Gly Ile Asp Gly Val Gly Val Val Ala Ser Val
290 295 300

Pro Glu Gly Met Lys Glu Leu Leu Ala Glu Asn Leu Met Val Met Met
305 310 315 320

Arg Asp Leu Glu Ser Cys Ala Gly Asn Asp Asn Leu Ile Ser Glu Ser
325 330 335

Asp Ile Arg Arg Ser Ala His Thr Leu Pro Val Leu Leu Ala Gly Ala
340 345 350

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

Phe Ala Leu Ser Asn Phe Asn Ala Asp Asp Met Asp Asp Phe Leu Val
370 375 380

Leu Gln Arg Asp Trp Gly Ala Asp Gly Gly Leu Arg Thr Val Ser Arg
385 390 395 400

Glu His Leu Ala Arg Val Arg Arg Arg Ala Ala Thr Ala Val Gln Ala
405 410 415

Val Tyr Arg Asp Leu Gly Leu Ala Asp Phe Asp Asp Thr Arg Ile Asp
420 425 430

Ala Val Val Val Ala Asn Asp Ser Arg Asp Leu Pro Ala Gly Asp Pro
435 440 445

Lys Ala Val Ala Glu Ala Ala Thr Ala Ile Glu Ala Arg Gln Leu Thr
450 455 460

Val Phe Asp Val Val Ala Ala Leu His Arg Thr Gly Tyr Ala Pro Glu
465 470 475 480

Ala Glu Ala Ile Met Arg Leu Thr Arg Glu Arg Leu Arg Gly Asp Gln
485 490 495

585
AA

Leu Gln Thr Ser Ala Ile Phe Asp Asp Gln Phe Gln Val Leu Ser Lys
500 505 510

Ile Thr Asp Pro Asn Asp Tyr Ala Gly Pro Gly Ser Gly Tyr Thr Pro
515 520 525

Thr Glu Lys Arg Arg Ala Glu Ile Asp Gly Ile Arg Gln Ala Arg Thr
530 535 540

Ser Ala Glu Leu Thr Ala Asp Gln Ala Glu His Arg Gly His Val Val
545 550 555 560

Phe Ser Asp Val Glu Pro Ala Arg Gln Gly Ser Asp Pro Arg Glu Val
565 570 575

Cys Ile Gly Leu Ser Pro Ala Leu Gly Arg Ser Val Trp Leu Thr Leu
580 585 590

Cys Gly Leu Thr Val Gly Glu Val Leu Arg Gln Leu Ser Ala Gly Leu
595 600 605

Glu Glu Glu Gly Cys Val Pro Arg Leu Val Arg Val Arg Ser Thr Ile
610 615 620

Asp Val Gly Leu Ile Gly Leu Thr Ala Ala Arg Leu Ser Gly Ser Gly
625 630 635 640

Ile Gly Ile Gly Leu Gln Gly Lys Gly Thr Ala Leu Ile His Arg Arg
645 650 655

Asp Leu Ala Pro Leu Ala Asn Leu Glu Leu Phe Ser Val Ala Pro Leu
660 665 670

Leu Thr Ala Lys Met Tyr Arg Glu Leu Gly Arg Asn Ala Ala Arg His
675 680 685

Ala Lys Gly Met Ala Pro Leu Pro Ile Leu Ala Gly Gly Thr Asp Glu
690 695 700

Ser Ile Ser Ala Arg Tyr His Ala Arg Ala Val Ala Leu Val Ala Leu
705 710 715 720

Glu Arg Gln Ala Cys Glu Pro Gly Gln Ala Pro Ile Thr Val Glu Ala
725 730 735

586
/

Lys Arg Val

<210> 242
 <211> 110
 <212> PRT
 <213> *Mycobacterium sp.*

<400> 242

Met Thr Glu Lys Phe Thr Val Ala Ala Ala Val Asp Gly Lys Leu Thr
 1 5 10 15

Leu Ser Asp Leu Arg Met Asp Pro Ala Thr Leu Ala Tyr Gln Ala Val
 20 25 30

Val Ala Glu Gln Asp Gly Asn Pro Gln Leu Ala Glu Asn Phe Leu Arg
 35 40 45

Ala Ala Glu Leu Ala Val Ile Asp Asp Glu Ala Val Met Lys Leu Tyr
 50 55 60

Glu Ala Leu Arg Pro His Arg Ser Thr Ala Ala Glu Leu Glu Glu Leu
 65 70 75 80

Arg Val Ser Leu Glu Thr Gly Gly Ala Ser Arg Cys Ala Glu Leu Val
 85 90 95

Arg Gln Ala Ala Glu Val Tyr Ala Arg Arg Gly Leu Leu Arg
 100 105 110

<210> 243
 <211> 570
 <212> PRT
 <213> *Mycobacterium smegmatis*

<400> 243

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

Arg Thr Lys Leu Leu Glu Glu Arg Pro Val Asn Leu Asp Gly Phe Val
 20 25 30

Gln Glu Trp Pro Glu Val Gly Met Val Ala Met Asp Ser Ala Phe Asp
 35 40 45

Pro Glu Pro Ser Val Arg Val Glu Asn Gly Val Ile Val Glu Met Asp
 50 55 60

588

Gly Arg Ala Arg Ala Asp Phe Asp Phe Ile Asp Gln Phe Ile Ala Asp
65 70 75 80

His Ala Ile Asp Val Ala Thr Thr Glu Gln Ser Met Ala Leu Pro Ala
85 90 95

Val Glu Ile Ala Arg Met Leu Val Asp Pro Arg Val Thr Arg Asp Glu
100 105 110

Val Ile Ala Val Thr Gly Gly Leu Thr Pro Ala Lys Leu Leu Glu Val
115 120 125

Ala Lys Asn Leu Asn Ile Val Glu Ile Met Met Gly Ile Gln Lys Met
130 135 140

Arg Ala Arg Arg Thr Pro Ala Asn Gln Ala His Cys Thr Ser Ala Arg
145 150 155 160

Asp Asn Pro Leu Gln Val Ala Cys Glu Ala Ala Glu Ala Ser Leu Arg
165 170 175

Gly Phe Ser Glu Val Glu Thr Thr Leu Gly Val Val Arg Tyr Ala Pro
180 185 190

Leu Val Ala Met Ala Leu Gln Ile Gly Ser Gln Val Gly Ser Gly Gly
195 200 205

Leu Leu Thr Gln Cys Ala Leu Glu Glu Ala Thr Glu Leu Glu Leu Gly
210 215 220

Met Arg Gly Ile Thr Ala Tyr Ala Glu Thr Ile Ser Val Tyr Gly Thr
225 230 235 240

Glu Ser Val Phe Val Asp Gly Asp Asp Thr Pro Trp Ser Lys Ala Phe
245 250 255

Leu Ala Ala Ala Tyr Ala Ser Arg Gly Ile Lys Met Arg Phe Thr Ser
260 265 270

Gly Thr Gly Ser Glu Val Gln Met Gly Asn Ala Glu Gly Arg Ser Met
275 280 285

Leu Tyr Leu Glu Ile Arg Cys Ile Leu Val Ala Lys Gly Ala Gly Val
290 295 300

588
A

Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys Ile Gly Val Pro Gly Ala
305 310 315 320

Val Pro Ala Gly Ile Arg Ala Val Ala Ala Glu Asn Leu Ile Ala Ser
325 330 335

Ala Val Asp Leu Glu Cys Ala Ser Gly Asn Asp Gln Ser Phe Ser His
340 345 350

Ser Pro Met Arg Arg Thr Ala Arg Leu Leu Pro Gln Met Leu Pro Gly
355 360 365

Thr Asp Phe Val Cys Ser Gly Tyr Ser Ala Val Pro Asn Tyr Asp Asn
370 375 380

Met Phe Ala Gly Ser Asn Leu Asp Ala Glu Asp Phe Asp Asp Phe Asn
385 390 395 400

Thr Ile Gln Arg Asp Leu Gln Val Asp Gly Gly Leu Arg His Val Asn
405 410 415

Glu Ala Glu Ile Val Ala Ala Arg Arg Arg Ala Ala Gln Ala Leu Gln
420 425 430

Ala Val Phe Arg Tyr Leu Asp Leu Pro Ala Ile Thr Asp Ala Glu Ile
435 440 445

Glu Ala Ala Val Tyr Ala His Gly Ser Arg Glu Leu Ile Pro Arg Asp
450 455 460

Val Leu Glu Asp Leu Lys Gly Ala Gln Gln Val Met Asp Arg Asn Val
465 470 475 480

Thr Gly Leu Asp Leu Val Lys Ala Leu Glu Ser Thr Gly Phe Ala Asp
485 490 495

Ile Ala Glu Asn Leu Leu Ala Val Leu Arg Gln Arg Val Ser Gly Asp
500 505 510

Leu Leu Gln Thr Ser Ala Ile Met Thr Arg Asp Leu Lys Pro Leu Ser
515 520 525

Ala Val Asn Asp Arg Asn Asp Tyr Ala Gly Pro Gly Thr Gly Tyr Arg
530 535 540

589
A

Pro Ser Gly Ala Arg Trp Glu Glu Met Lys Arg Leu Arg His Val Thr
 545 550 555 560

Ser Ala Glu Asn Pro Glu Leu Glu Val Glu
 565 570

<210> 244

<211> 200

<212> PRT

<213> *Mycobacterium smegmatis*

<400> 244

Met Ile Met Ser Ala Gln Ser Thr Gln Ala Gln Arg Thr Leu Glu Leu
 1 5 10 15

Val Gly Asp Ser Pro Ala Glu Pro Gly Lys Arg Ser Asp Glu Val Val
 20 25 30

Leu Ala Val Ser Pro Ala Phe Ala Asp Phe Phe Ser Gln Thr Ile Ile
 35 40 45

Gly Leu Ser His Ala Asp Val Ile Arg Glu Ile Leu Ala Gly Ile Glu
 50 55 60

Glu Gln Glu Val His Ala Arg Cys Ile Arg Val Arg His Ser Ser Asp
 65 70 75 80

Leu Ala Val Leu Ala His Thr Ala Ala Lys Leu Ser Gly Ser Gly Ile
 85 90 95

Gly Ile Gly Ile Leu Ser Arg Gly Thr Ser Met Ile His Gln Arg Asp
 100 105 110

Leu Pro Arg Leu Ser Ser Leu Glu Leu Phe Pro Gln Ser Pro Leu Met
 115 120 125

Thr Leu Glu Thr Tyr Arg Ser Ile Gly Ser Asn Ala Ala Gln Tyr Ala
 130 135 140

Lys Gly Glu Ser Pro Glu Pro Val Pro Thr Leu Asn Asp Gln Met Ala
 145 150 155 160

Arg Pro Arg Trp Gln Ala Lys Ala Ala Leu Leu His Leu Lys Glu Thr
 165 170 175

590
11A

Glu Gln Val Val Lys Gln Ala Lys Pro Val Glu Val Val Pro Gln Phe
180 185 190

Gly Val Ala Glu Ala Leu Gly Thr
195 200

<210> 245

<211> 114

<212> PRT

<213> *Mycobacterium smegmatis*

<400> 245

Met Asn Asp Lys Phe Thr Asp Lys Tyr Thr Val Ala Ala Ala Val Asp
1 5 10 15

Gly Lys Leu Asp Leu Ser Asp Leu Arg Met Asp Pro Ala Val Leu Ala
20 25 30

His Gln Ala Val Ile Ala Glu Glu Asn Gly Asn Pro Gln Leu Ala Glu
35 40 45

Asn Phe Leu Arg Ala Ala Glu Leu Ala Thr Ile Asp Asp Glu Asp Val
50 55 60

Met Arg Leu Tyr Glu Ala Leu Arg Pro Tyr Arg Ser Ser Ala Glu Asp
65 70 75 80

Leu Asp Ala Leu Gln Ala Ser Leu Glu Ser Arg Gly Ala Ala Arg Cys
85 90 95

Ala Glu Leu Val Arg Gln Ala Ala Glu Ala Tyr Ala Arg Arg Gly Leu
100 105 110

Leu Arg

<210> 246

<211> 754

<212> PRT

<213> *Mycobacterium smegmatis*

<400> 246

Met His Leu Asn Ser Thr Glu Val Asp Ser Arg Leu Gly Arg Ile Arg
1 5 10 15

Leu Leu Asp Arg Gln Arg Val Asn Leu Asp Gly Phe Ala Asp Val Asp
20 25 30

591
AS

Ala Glu Leu Gly Met Ile Ser His Leu Ser Pro Asn Asp Pro Glu Pro
35 40 45

Ser Trp Val Val Ala Asp Asp Gly Thr Val Leu Glu Met Asp Ser Lys
50 55 60

Pro Ala Glu Asp Phe Asp Thr Ile Asp Glu Phe Ile Val Lys Tyr Ala
65 70 75 80

Ile Asp His Glu Gln Ala Pro Arg Ser Met Ala Met Thr Asp Leu Asp
85 90 95

Leu Ala Arg Met Ile Val Asp Pro Gly Arg Pro Arg Glu Glu Ile Leu
100 105 110

Arg Val Cys Ser Gly Leu Thr Pro Ala Lys Met Ala Arg Val Val Ala
115 120 125

Ser Leu Gln Pro Val Glu Ile Gln Met Ala Met Met Lys Met Arg Ala
130 135 140

Arg Arg Thr Pro Ala Asn Gln Ala His Val Thr Asn Arg Leu Asp Asp
145 150 155 160

Pro Leu Leu Ile Ala Ala Asp Ala Ala Thr Ala Val Val Tyr Gly Phe
165 170 175

Arg Glu Leu Glu Ala Thr Val Pro Val Leu Asp Asp Ala Pro Ala Val
180 185 190

Ala Val Gly Leu Leu Ile Gly Ser Gln Val Pro Ala Pro Gly Ala Leu
195 200 205

Thr Gln Cys Ser Val Glu Glu Ala Arg Glu Leu Glu Leu Gly Val Arg
210 215 220

Gly Leu Val Ser Tyr Ala Glu Thr Val Ser Val Tyr Gly Thr Glu Gln
225 230 235 240

Val Phe Thr Asp Gly Asp Asp Thr Pro Trp Ser Lys Ala Phe Leu Thr
245 250 255

Ser Cys Tyr Ala Ser Arg Gly Ile Lys Met Arg Leu Ser Ser Gly Ala
260 265 270

592
AP

Gly Ser Glu Val Leu Met Gly Gln Ala Glu Gly Lys Ser Met Asn Tyr
275 280 285

Leu Glu Ala Arg Cys Val Ala Leu Ala Arg Gly Ile Gly Ala Gln Gly
290 295 300

Val Gln Asn Gly Gly Val Asp Gly Ala Ala Ile Thr Ala Ser Val Pro
305 310 315 320

Gly Gly Val Lys Glu Leu His Ala Glu Asn Leu Met Val Met Leu Arg
325 330 335

Gly Leu Glu Ser Cys Ser Gly Asn Asp Ser Leu Met Ser Glu Ser Thr
340 345 350

Met Arg Arg Thr Ser Arg Thr Leu Pro Thr Leu Leu Ser Gly Ser Asp
355 360 365

Phe Ile Phe Ser Gly Phe Gly Ser Val Val Ser Tyr Asp Asn Met Phe
370 375 380

Gly Pro Ser Asn Phe Asn Ala Ala Asp Leu Asp Asp Tyr Leu Val Leu
385 390 395 400

Gln Arg Asp Trp Gly Val Asp Gly Gly Leu Arg Ser Val Asp Pro Thr
405 410 415

Thr Leu Glu Ser Met Arg Arg Glu Ala Ala Glu Ala Thr Arg Ala Val
420 425 430

Phe Glu Tyr Leu Gly Leu Ala Asp Phe Asp Asp Asp His Val Glu Ala
435 440 445

Val Val Gly Ala Glu Gly Ser Lys Asp Leu Pro Gln Asp Asp Gly Val
450 455 460

Lys Val Leu Ser Ala Ala Arg Met Ile Asp Gln Ser Gly Leu Thr Val
465 470 475 480

Leu Asp Ile Val Ser Ala Leu Ala Glu Thr Gly Phe Thr His Ile Ala
485 490 495

Asp Arg Val Leu Gly Met Ala Arg Ala Arg Val Thr Gly Asp Tyr Leu
500 505 510

593
H

Gln Thr Ala Ala Ile Phe Asp Glu Glu Met Asn Val Leu Ser Ala Leu
515 520 525

His Asp Pro Asn Asp Tyr Arg Gly Pro Gly Thr Gly Tyr Arg Pro Thr
530 535 540

Pro Glu Arg Gln Ala Gln Ile Asp Ala Val Arg Gln Ala Arg Ser Val
545 550 555 560

Ala Asp Leu Val Lys Glu Gln Ala Thr Ser Ala Gln Pro Asp Arg Leu
565 570 575

Arg Val Leu Gly Ala Ala Thr Val Gly Glu Asp Pro Arg Glu Val Val
580 585 590

Ile Gly Val Ser Pro Ala Phe Gly Thr Lys Leu Phe Arg Thr Leu Ser
595 600 605

Gly Met Thr Ile Tyr Asp Val Leu Glu Gln Ile Leu Ala Gly Leu Glu
610 615 620

Glu Glu His Cys Val Pro Arg Leu Val Arg Ile Thr Asp Ser Ile Asp
625 630 635 640

Leu Gly Ala Ile Gly Lys Ser Ala Ala Gln Leu Ser Gly Ser Gly Ile
645 650 655

Gly Val Gly Leu Gln Ala Lys Gly Thr Thr Leu Ile His Arg Arg Asp
660 665 670

Leu Pro Pro Leu Ala Asn Leu Glu Leu Leu Ser Val Ala Pro Leu Ile
675 680 685

Thr Pro Glu Met Tyr Arg Leu Ile Gly Ile Asn Ala Gly Arg His Ala
690 695 700

Lys Gly Ala Thr Pro Ser Pro Met Arg Asn Ala Tyr Thr Asp Glu Ala
705 710 715 720

Ile Thr Ala Arg Tyr His Thr Lys Val Val Ser Met Val Ala Ile Glu
725 730 735

Arg Glu Glu Ser Glu Arg Arg Glu Thr Gly Asn Val Glu Leu Glu Ile
740 745 750

594
A

Thr Arg

<210> 247
 <211> 757
 <212> PRT
 <213> *Mycobacterium smegmatis*

<400> 247

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

Gly Arg Met Arg Ile Leu Asp Ala Lys Pro Val Asn Leu Asp Gly Phe
 20 25 30

Ser Val Pro Asp Pro Asp Leu Gly Leu Ala Ala Met Ser Ser Pro His
 35 40 45

Asp Pro Gln Pro Ser Leu Val Ile Arg Asp Gly Arg Val Val Glu Met
 50 55 60

Asp Gly Lys Ala Ala Glu Asp Phe Asp Val Ile Asp Glu Phe Ile Ala
 65 70 75 80

Arg Tyr Gly Leu Asp Leu Asp Val Ala Pro Glu Ala Met Ala Met Ser
 85 90 95

Asp Ile Asp Leu Ala Arg Met Ala Val Asp Ile Asn Val Pro Arg Ala
 100 105 110

Glu Val Val Arg Leu Ile Ala Gly Thr Thr Pro Ala Lys Leu Ala Lys
 115 120 125

Val Ile Ala Val Leu Thr Pro Val Glu Met Gln Ala Ala Met Ala Lys
 130 135 140

Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Ala His Val Thr Asn Gln
 145 150 155 160

Leu Asp Asp Pro Leu Leu Ile Ala Ala Asp Ala Ala Ser Ala Val Ala
 165 170 175

Tyr Gly Phe Arg Glu Val Glu Thr Thr Val Pro Val Leu Ala Asp Ala
 180 185 190

Pro Ser Asn Ala Val Ala Leu Leu Ile Gly Ser Gln Val Gly Val Pro
 195 200 205

595
AA

Gly Ala Met Ala Gln Cys Ser Ile Glu Glu Ala Leu Glu Leu Arg Leu
210 215 220

Gly Leu Arg Gly Leu Thr Ser Tyr Ala Glu Thr Ile Ser Ile Tyr Gly
225 230 235 240

Thr Glu Gln Val Phe Val Asp Gly Asp Asp Thr Pro Phe Ser Lys Ala
245 250 255

Ile Leu Thr Ser Ala Tyr Ala Ser Arg Gly Leu Lys Met Arg Val Thr
260 265 270

Ser Gly Gly Gly Ala Glu Val Leu Met Gly Ala Ala Glu Lys Cys Ser
275 280 285

Ile Leu Tyr Leu Glu Ser Arg Cys Val Ser Leu Ala Arg Ala Leu Gly
290 295 300

Ser Gln Gly Val Gln Asn Gly Gly Ile Asp Gly Val Gly Val Val Ala
305 310 315 320

Ser Val Pro Asp Gly Met Lys Glu Leu Leu Ala Glu Asn Leu Met Val
325 330 335

Met Met Arg Asp Leu Glu Ser Cys Ala Gly Asn Asp Asn Leu Ile Ser
340 345 350

Glu Ser Asp Ile Arg Arg Ser Ala His Thr Leu Pro Val Leu Leu Ala
355 360 365

Gly Ala Asp Phe Ile Phe Ser Gly Phe Gly Ser Ile Pro Arg Tyr Asp
370 375 380

Asn Ala Phe Ala Leu Ser Asn Phe Asn Ser Asp Asp Met Asp Asp Phe
385 390 395 400

Leu Val Leu Gln Arg Asp Trp Gly Ala Asp Gly Gly Leu Arg Thr Val
405 410 415

Pro Ala Asp Gln Leu Ala Ala Val Arg Arg Arg Ala Ala Arg Ala Val
420 425 430

Gln Ala Val Tyr Arg Asp Leu Gly Leu Ala Asp Phe Asp Asp Gln His
435 440 445

596
A

Ile Glu Asn Val Val Ala Ala Asn Gly Ser Arg Asp Leu Pro Pro Gly
450 455 460

Asp Pro Lys Ala Val Leu Glu Ala Ala Asn Ala Ile Glu Ala Lys Gln
465 470 475 480

Leu Thr Val Phe Asp Val Val Ala Ser Leu Lys Arg Thr Gly Phe Asp
485 490 495

Pro Glu Ala Glu Ala Ile Met Arg Leu Thr Ala Glu Arg Met Arg Gly
500 505 510

Asp Gln Leu Gln Thr Ser Ala Ile Phe Asp Glu Gln Phe Arg Val Leu
515 520 525

Ser Lys Ile Thr Asp Pro Asn Asp Tyr Ala Gly Pro Gly Thr Gly Tyr
530 535 540

Thr Leu Ser Glu Gln Arg Arg Ala Glu Ile Asp Asn Ile Arg Gln Gln
545 550 555 560

Arg Ser Ala Ala Glu Leu Thr Ala Asp Gln Ala Glu His Ala Gly His
565 570 575

Ile Thr Val Thr Glu Ile Glu Pro Ala Arg Gln Gly Ser Asp Pro Arg
580 585 590

Glu Val Cys Ile Gly Leu Ser Pro Ala Leu Gly Arg Ser Val Trp Leu
595 600 605

Ser Leu Cys Gly Leu Pro Ile Gly Glu Val Ile Arg Gln Ile Ser Ala
610 615 620

Gly Leu Glu Glu Glu Gly Cys Val Pro Arg Phe Val Arg Val Arg Ser
625 630 635 640

Thr Ile Asp Val Gly Leu Ile Gly Leu Thr Ala Ala Lys Leu Ala Gly
645 650 655

Ser Gly Ile Gly Ile Gly Leu Gln Gly Lys Gly Thr Ala Leu Ile His
660 665 670

Arg Arg Asp Leu Ala Pro Leu Ala Asn Leu Glu Leu Phe Ser Val Ala
675 680 685

597
P

Pro Leu Leu Thr Ala Arg Asn Tyr Arg Glu Leu Gly Arg Asn Ala Ala
690 695 700

Arg His Ala Lys Gly Met Ala Pro Val Pro Ile Leu Thr Gly Gly Thr
705 710 715 720

Asp Glu Ser Ile Ser Ala Arg Tyr His Ala Arg Ala Val Ala Leu Val
725 730 735

Ala Leu Glu Arg Gln Ala Ser Glu Pro Gly Glu Ala Pro Val Thr Val
740 745 750

Glu Val Arg Arg Pro
755

<210> 248

<211> 142

<212> PRT

<213> *Mycobacterium smegmatis*

<400> 248

Met Ser Glu Pro Val Leu Asp Pro Ala Val Asp Tyr Pro Leu Ser Leu
1 5 10 15

Asn Arg Lys Asp Leu Leu Thr Thr Pro Asn Gly Lys Pro Ile Asp Ala
20 25 30

Ile Thr Met Asp Ala Val Met Ser Gly Glu Val Ser Ala Ser Asp Leu
35 40 45

Arg Ile Thr Pro Glu Thr Leu Arg Leu Gln Ala Gln Ile Ala Glu Gly
50 55 60

Val Gly Arg Lys Gln Leu Gly Ala Asn Leu Arg Arg Ala Ala Glu Met
65 70 75 80

Thr Ala Ile Ser Asp Glu Arg Val Leu Gln Ile Tyr Asn Ala Leu Arg
85 90 95

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

Glu Thr Gln Tyr Gly Ala Thr Met Leu Ala Gly Leu Val Arg Glu Ala
115 120 125

Ala Asp Val Tyr Glu Arg Arg Asp Ile Leu Ala Thr Ser Glu
130 135 140

598
14

<210> 249
 <211> 119
 <212> PRT
 <213> *Mycobacterium smegmatis*

<400> 249

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

Glu Ala Ala Arg Gly Gly Asp Leu Thr Leu Asp Asp Ile Arg Ile Ser
 20 25 30

Arg Asp Thr Leu Ile Ser Gln Ala Glu Ala Ala Glu Arg Thr Gly Ser
 35 40 45

Glu Gln Leu Gly Leu Asn Leu Arg Arg Ala Ala Glu Leu Thr Ala Leu
 50 55 60

Ser Ser Asp Asp Met Leu Ala Ala Tyr Glu Ala Leu Arg Pro Gly Arg
 65 70 75 80

Ser Thr Phe Ser Glu Leu Glu Ala Leu Ala Gln Arg Leu Ala Ala Gln
 85 90 95

Glu Ala His Thr Cys Ala Gln Leu Val Arg Glu Ala Ala Ala Ala Tyr
 100 105 110

Arg Arg Arg Gly Leu Leu Arg
 115

<210> 250
 <211> 224
 <212> PRT
 <213> *Salmonella enterica subspécie enterica sorovar Paratyphi*

<400> 250

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
 1 5 10 15

Arg Asp Met Lys Gly Ser Asp Lys Pro Val Ser Phe Asn Thr Pro Ala
 20 25 30

Ala Ser Thr Ala Pro Gln Thr Ala Ala Pro Ala Gly Asp Gly Phe Leu
 35 40 45

Thr Glu Val Gly Glu Ala Arg Gln Gly Thr Gln Gln Asp Glu Val Ile
 50 55 60

599
A

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

Gly Leu Pro His Lys Ser Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Asp Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 251

<211> 173

<212> PRT

<213> *Salmonella enterica subspécie enterica sorovar Paratyphi*

<400> 251

Met Asn Ile Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
1 5 10 15

Met Asn Ser Leu Gln Gly Asp Ala Pro Ala Ala Ala Pro Ala Ala Gly
20 25 30

Gly Thr Ser Arg Ser Ala Lys Val Ser Asp Tyr Pro Leu Ala Asn Lys
35 40 45

His Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe
50 55 60

600
H

Thr Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg
65 70 75 80

Ile Thr Pro Glu Thr Leu Arg Leu Gln Ala Ser Ile Ala Lys Asp Ala
85 90 95

Gly Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr
100 105 110

Ala Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
115 120 125

Tyr Arg Ser Thr Lys Glu Glu Leu Leu Ala Ile Ala Asp Asp Leu Glu
130 135 140

Asn Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala
145 150 155 160

Gly Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
165 170

<210> 252

<211> 188

<212> PRT

<213> *Clostridium perfringens*

<400> 252

Met Glu Glu Arg Thr Phe Ile Pro Glu Ile Thr Val Glu Glu Val Gly
1 5 10 15

Glu Ala Lys Val Gly Leu Arg Ser Asp Glu Val Val Ile Gly Leu Ala
20 25 30

Pro Ala Phe Leu Lys Tyr Gln Asn Lys Thr Ile Val Asp Val Pro His
35 40 45

Thr Glu Thr Leu Leu Glu Ile Ile Ala Gly Ile Glu Glu Glu Gly Leu
50 55 60

His Ala Arg Val Val Arg Val Leu Arg Thr Ser Asp Val Ser Phe Ile
65 70 75 80

Ala His Asp Ala Ala Cys Leu Ser Gly Ser Gly Ile Gly Ile Gly Ile
85 90 95

Gln Ser Lys Gly Thr Thr Val Ile His Gln Lys Asp Leu Leu Pro Leu
100 105 110

601
AA

Asn Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr Pro Glu Thr
115 120 125

Tyr Arg Leu Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys Gly Glu Ser
130 135 140

Pro Thr Pro Val Pro Val Lys Asn Asp Gln Met Val Arg Pro Lys Phe
145 150 155 160

Met Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His Val Glu
165 170 175

Pro Gly Lys Lys Pro Val Gln Leu Glu Val Lys Phe
180 185

<210> 253

<211> 141

<212> PRT

<213> *Clostridium perfringens*

<400> 253

Met Glu Asn Lys Arg Met Thr Ala Ala Asp Tyr Pro Leu Thr Ser Lys
1 5 10 15

Arg Lys Gly Asp Ile Lys Thr Pro Thr Gly Lys Ala Leu Glu Asp Ile
20 25 30

Thr Leu Glu Lys Val Leu Ser Gly Glu Ile Asn Ala Asp Asp Ile Arg
35 40 45

Ile Ser Pro Glu Thr Leu Glu Met Gln Ala Gln Ile Ala Glu Ser Met
50 55 60

Asn Arg Asp Ala Ile Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80

Arg Val Pro Asp Asp Arg Ile Leu Glu Met Tyr Asn Ala Leu Arg Pro
85 90 95

Tyr Arg Ser Thr Lys Glu Asp Leu Phe Lys Ile Ala Asp Glu Leu Glu
100 105 110

Thr Lys Tyr Asp Ala Lys Val Asn Ala Asn Phe Val Arg Glu Ala Ala
115 120 125

602
H

Glu Val Tyr Glu Thr Arg Asn Lys Leu Arg Ile Glu Glu
130 135 140

<210> 254

<211> 554

<212> PRT

<213> *Salmonella enterica subespécie enterica sorovar Typhimurium*

<400> 254

Met Arg Ser Lys Arg Phe Glu Ala Leu Ala Lys Arg Pro Val Asn Gln
1 5 10 15

Asp Gly Phe Val Lys Glu Trp Ile Glu Glu Gly Phe Ile Ala Met Glu
20 25 30

Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Val Asn Gly Ala Val
35 40 45

Thr Glu Leu Asp Gly Lys Pro Val Ser Glu Phe Asp Leu Ile Asp His
50 55 60

Phe Ile Ala Arg Tyr Gly Ile Asn Leu Asn Arg Ala Glu Glu Val Met
65 70 75 80

Ala Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro Asn Val
85 90 95

Lys Arg Ser Glu Ile Val Pro Leu Thr Thr Ala Met Thr Pro Ala Lys
100 105 110

Ile Val Glu Val Val Ser His Met Asn Val Val Glu Met Met Met Ala
115 120 125

Met Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Gln Gln Ala His Val
130 135 140

Thr Asn Val Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala Glu
145 150 155 160

Gly Ala Trp Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala Val Ala
165 170 175

Arg Tyr Ala Pro Phe Asn Ala Ile Ala Leu Leu Val Gly Ser Gln Val
180 185 190

Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala Thr Glu
195 200 205

Leu Lys Leu Gly Met Leu Gly His Thr Cys Tyr Ala Glu Thr Ile Ser
210 215 220

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

Ser Lys Gly Phe Leu Ala Ser Ser Tyr Ala Ser Arg Gly Leu Lys Met
245 250 255

Arg Phe Thr Ser Gly Ser Gly Ser Glu Val Gln Met Gly Tyr Ala Glu
260 265 270

Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr Lys
275 280 285

Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Val Ser Ser Ile Gly
290 295 300

Val Pro Ser Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu Asn
305 310 315 320

Leu Ile Cys Ser Ser Leu Asp Leu Glu Cys Ala Ser Ser Asn Asp Gln
325 330 335

Thr Phe Thr His Ser Asp Met Arg Arg Thr Ala Arg Leu Leu Met Gln
340 345 350

Phe Leu Pro Gly Thr Asp Phe Ile Ser Ser Gly Tyr Ser Ala Val Pro
355 360 365

Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Glu Asp Ala Glu Asp Phe
370 375 380

Asp Asp Tyr Asn Val Ile Gln Arg Asp Leu Lys Val Asp Gly Gly Leu
385 390 395 400

Arg Pro Val Arg Glu Glu Asp Val Ile Ala Ile Arg Asn Lys Ala Ala
405 410 415

Arg Ala Leu Gln Ala Val Phe Ala Gly Met Gly Leu Pro Pro Ile Thr
420 425 430

Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp Met
435 440 445

Pro Glu Arg Asn Ile Val Glu Asp Ile Lys Phe Ala Gln Glu Ile Ile
450 455 460

Asn Lys Asn Arg Asn Gly Leu Glu Val Val Lys Ala Leu Ala Gln Gly
465 470 475 480

Gly Phe Thr Asp Val Ala Gln Asp Met Leu Asn Ile Gln Lys Ala Lys
485 490 495

Leu Thr Gly Asp Tyr Leu His Thr Ser Ala Ile Ile Val Gly Asp Gly
500 505 510

Gln Val Leu Ser Ala Val Asn Asp Val Asn Asp Tyr Ala Gly Pro Ala
515 520 525

Thr Gly Tyr Arg Leu Gln Gly Glu Arg Trp Glu Glu Ile Lys Asn Ile
530 535 540

Pro Gly Ala Leu Asp Pro Asn Glu Ile Asp
545 550

<210> 255

<211> 524

<212> PRT

<213> *Listeria monocytogenes*

<400> 255

Met Glu Ser Pro Asn Asp Pro Lys Pro Ser Ile Lys Ile Glu Asn Gly
1 5 10 15

Lys Val Val Glu Met Asp Ser Lys Lys Leu Ala Asp Phe Asp Leu Ile
20 25 30

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

Val Met Gln Met Asp Ser Val Lys Leu Ala Asn Met Leu Cys Asp Pro
50 55 60

Asn Val Pro Arg Glu Lys Ile Val Leu Leu Thr Thr Ala Met Thr Pro
65 70 75 80

Ala Lys Ile Val Glu Val Val Ser Gln Met Asn Val Val Glu Met Met
85 90 95

605

Met Ser Met Gln Lys Met Arg Ser Arg Arg Thr Pro Thr Thr Gln Ala
 100 105 110

His Val Thr Asn Leu Arg Asp Asn Pro Val Gln Ile Ala Ala Asp Ala
 115 120 125

Ala Glu Ala Ala Ile Arg Gly Phe Asp Glu Gln Glu Thr Thr Val Ala
 130 135 140

Val Val Arg Tyr Ala Pro Phe Asn Ala Leu Ser Leu Leu Val Gly Ser
 145 150 155 160

Gln Thr Gly Arg Gly Gly Val Leu Thr Gln Cys Ser Leu Glu Glu Ala
 165 170 175

Thr Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Cys Tyr Ala Glu Thr
 180 185 190

Ile Ser Val Tyr Gly Thr Glu Pro Val Phe Thr Asp Gly Asp Asp Thr
 195 200 205

Pro Trp Ser Lys Gly Ile Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu
 210 215 220

Lys Met Arg Phe Thr Ser Gly Thr Gly Ser Glu Val Gln Met Gly Tyr
 225 230 235 240

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

Thr Lys Ala Ala Gly Val Gln Gly Leu Gln Asn Gly Ser Ile Ser Cys
 260 265 270

Ile Gly Ile Pro Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala
 275 280 285

Glu Asn Leu Ile Ala Val Met Leu Asp Leu Glu Val Ala Ser Gly Asn
 290 295 300

Asp Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Leu Leu
 305 310 315 320

Met Gln Phe Leu Pro Gly Thr Asp Tyr Ile Ser Ser Gly Tyr Ser Ala
 325 330 335

606
A

Thr Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Asp
340 345 350

Asp Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Lys Val Asp Gly
355 360 365

Gly Leu Thr Pro Val Thr Glu Glu Glu Val Val Ala Val Arg Asn Lys
370 375 380

Ala Ala Arg Val Ile Gln Ala Val Phe Asp Lys Leu Gly Leu Pro Glu
385 390 395 400

Val Thr Asp Ala Glu Val Glu Ala Ala Thr Tyr Ala Arg Gly Ser Lys
405 410 415

Asp Met Pro Glu Arg Asn Met Val Glu Asp Ile Lys Ala Ala Ala Glu
420 425 430

Met Met Asp Arg Gly Val Thr Gly Leu Asp Val Val Lys Ala Leu Ser
435 440 445

Ala Gly Gly Phe Asp Asp Val Ala Glu Ser Val Leu Asn Met Leu Lys
450 455 460

Gln Arg Val Ser Gly Asp Phe Leu His Thr Ser Ala Ile Ile Asp Lys
465 470 475 480

Asp Trp Asn Val Ile Ser Ser Val Asn Asp Leu Asn Asp Tyr Ala Gly
485 490 495

Pro Gly Thr Gly Tyr Arg Leu Glu Gly Glu Arg Trp Glu Lys Leu Lys
500 505 510

Asp Ile Ala Val Ala Val Asp Ala Asn Glu Leu Glu
515 520

<210> 256

<211> 555

<212> PRT

<213> *Escherichia blattae*

<400> 256

Met Arg Arg Ser Lys Arg Phe Glu Val Leu Glu Lys Arg Pro Val Asn
1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

607
A

Gly Ser Pro Trp Asp Pro Pro Ser Ser Val Lys Val Glu Gln Gly Arg
35 40 45

Ile Val Glu Leu Asp Gly Lys Ala Arg Ala Asp Phe Asp Met Ile Asp
50 55 60

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Ile Glu Glu Thr Glu His Ala
65 70 75 80

Met Gly Leu Asp Ala Leu Thr Ile Ala Arg Met Leu Val Asp Ile Asn
85 90 95

Val Ser Arg Ala Glu Ile Ile Lys Val Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Glu Glu Val Met Ser His Met Asn Val Val Glu Met Met Val
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Ile Gly Ser Gln
180 185 190

Ser Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Phe Thr Ser Tyr Ala Glu Thr Val
210 215 220

Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ala
260 265 270

608
A

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

Leu Arg Pro Val Ser Glu Glu Glu Thr Ile Ala Ile Arg Asn Lys Ala
405 410 415

Ala Arg Ala Val Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Pro Val
420 425 430

Thr Asp Glu Glu Val Thr Ala Ala Thr Tyr Ala His Gly Ser Lys Asp
435 440 445

Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Arg Ala Leu Ser Val
465 470 475 480

Asn Gly Phe Asp Asp Val Ala Asn Asn Ile Leu Asn Met Leu Arg Gln
485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Glu
500 505 510

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Ile Ser Pro Gln Arg Trp Glu Glu Ile Lys Asn
530 535 540

Ile Ala Thr Val Ile Gln Pro Asp Ser Ile Glu
545 550 555

<210> 257

<211> 555

<212> PRT

<213> não identificada

● <400> 257

Met Arg Arg Ser Lys Arg Phe Glu Val Leu Ala Gln Arg Pro Val Asn
1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

Glu Ser Pro Tyr Asp Pro Ala Ser Ser Val Lys Val Glu Asn Gly Arg
35 40 45

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

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Pro Glu Ala Glu Arg Ala
65 70 75 80

Met Gln Leu Asp Ala Leu Glu Ile Ala Arg Met Leu Val Asp Ile His
85 90 95

Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Val Glu Val Met Ala Gln Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

610
AA

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln
180 185 190

Cys Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val
210 215 220

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

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser
260 265 270

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

64
M

Leu Arg Pro Val Thr Glu Glu Glu Thr Ile Ala Ile Arg Asn Lys Ala
405 410 415

Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Leu Ile
420 425 430

Ser Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Lys Asp
435 440 445

Met Pro Ala Arg Asn Val Val Glu Asp Leu Ala Ala Val Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Ser Ser
465 470 475 480

Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Arg Gln
485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln
500 505 510

Phe Asp Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540

Ile Ala Gly Val Val Gln Pro Gly Ser Ile Glu
545 550 555

<210> 258

<211> 555

<212> PRT

<213> não identificada

<400> 258

Met Arg Arg Ser Lys Arg Phe Glu Val Leu Ala Gln Arg Pro Val Asn
1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

Glu Ser Pro Tyr Asp Pro Ala Ser Ser Val Lys Val Glu Asn Gly Arg
35 40 45

Ile Val Glu Leu Asp Gly Lys Asp Arg Ala Gln Phe Asp Met Ile Asp
50 55 60

612
AA

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Ala Glu Thr Glu Arg Ala
65 70 75 80

Met Gln Leu Asp Ala Leu Glu Ile Ala Arg Met Leu Val Asp Ile His
85 90 95

Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Val Glu Val Met Ala Lys Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln
180 185 190

Cys Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val
210 215 220

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

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser
260 265 270

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

613

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val
355 360 365

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

Leu Arg Pro Val Thr Glu Glu Glu Thr Ile Ala Ile Arg Asn Lys Ala
405 410 415

Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Leu Ile
420 425 430

Ser Asp Glu Glu Val Asp Ala Ala Thr Tyr Ala His Gly Ser Lys Asp
435 440 445

Met Pro Ala Arg Asn Val Val Glu Asp Leu Ala Ala Glu Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Asn Ser
465 470 475 480

Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Lys Gln
485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln
500 505 510

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540

614
A

Ile Ala Gly Val Val Gln Pro Ser Ala Ile Glu
545 550 555

<210> 259
<211> 556
<212> PRT
<213> *Lactobacillus collinoides*

<400> 259

Met Glu Arg Gln Lys Arg Phe Glu Lys Leu Glu Glu Arg Pro Val His
1 5 10 15

Leu Asp Gly Phe Val Lys Asn Trp Asp Asp Glu Gly Leu Val Ala Leu
20 25 30

Asn Gly Lys Asn Asp Pro Lys Pro Ser Ile Thr Ile Glu Asn Gly Val
35 40 45

Val Thr Glu Met Asp Gly Lys Lys Lys Ala Asp Phe Asp Leu Ile Asp
50 55 60

Lys Tyr Ile Ala Glu Tyr Gly Ile Asn Leu Asp Asn Ala Glu Lys Thr
65 70 75 80

Leu Asn Thr Asp Ser Val Lys Ile Ala Asn Met Met Cys Asp Pro Asn
85 90 95

Val Ser Arg Ala Glu Ile Ile Glu Tyr Thr Thr Ala Met Thr Pro Ala
100 105 110

Lys Ala Ala Glu Val Ile Ser Gln Leu Asn Phe Ala Glu Met Ile Met
115 120 125

Ala Thr Gln Lys Met Arg Pro Arg Arg Thr Pro Met Ala Gln Val His
130 135 140

Ala Thr Asn Thr Leu Asp Asn Pro Val Glu Ile Ala Ala Asp Ala Ala
145 150 155 160

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

Ala Arg Tyr Ala Pro Met Asn Ala Ile Ser Ile Met Val Gly Ala Gln
180 185 190

615
M

Ala Gly Arg Pro Gly Val Ile Thr Gln Cys Ser Val Glu Glu Ala Asp
195 200 205

Glu Leu Ser Leu Gly Met Arg Gly Phe Thr Ala Tyr Ala Glu Thr Ile
210 215 220

Ser Val Tyr Gly Thr Asp Arg Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Gly Phe Leu Ala Ser Cys Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Phe Thr Ser Gly Ala Gly Ser Glu Ala Met Met Gly Tyr Thr
260 265 270

Glu Gly Lys Ser Met Leu Tyr Leu Glu Ala Arg Cys Ile Tyr Ile Thr
275 280 285

Lys Ala Ser Gly Val Gln Gly Leu Gln Asn Gly Gly Val Ser Cys Ile
290 295 300

Gly Met Pro Gly Ala Val Val Gly Gly Ile Arg Glu Val Leu Gly Glu
305 310 315 320

Asn Leu Leu Cys Met Ser Leu Asp Val Glu Cys Ala Ser Gly Cys Asp
325 330 335

Gln Ala Phe Ser His Ser Asp Ile Arg Arg Thr Gly Arg Met Ile Gly
340 345 350

Gln Phe Ile Ala Gly Thr Asp Tyr Leu Ser Ser Ser Tyr Ala Ala Glu
355 360 365

Glu Asn Met Asp Asn Thr Phe Ala Gly Ser Asn Met Asp Val Leu Asp
370 375 380

Cys Asp Asp Tyr Ile Thr Leu Glu Arg Asp Met Ala Ile Asn Gly Gly
385 390 395 400

Ile Met Pro Ile Thr Glu Glu Glu Ser Ile Lys Ile Arg His Lys Ala
405 410 415

Ala Val Ala Ile Gln Ala Val Phe Asp Gly Leu Gly Leu Pro Gln Ile
420 425 430

616
AB

Thr Asp Glu Glu Val Glu Ala Ala Thr Tyr Gly Ser Asn Ser Asn Asp
435 440 445

Met Pro Lys Arg Asp Met Val Gln Asp Met Lys Ala Ala Gln Gly Leu
450 455 460

Met Thr Arg Gly Ile Thr Val Val Asp Val Ile Lys Ala Leu Tyr Asp
465 470 475 480

His Asp Ile Lys Asp Val Ala Glu Ala Val Leu Lys Leu Ala Gln Gln
485 490 495

Lys Val Cys Gly Asp Tyr Leu Gln Thr Ser Ala Val Phe Leu Asp Gly
500 505 510

Trp Lys Cys Thr Ser Ala Ile Asn Asn Ala Asn Asp Tyr Lys Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Leu Trp Glu Asp Lys Asp Lys Trp Asp Arg Leu
530 535 540

Glu Asn Val Pro Trp Ala Leu Asp Pro Gln Lys Leu
545 550 555

<210> 260

<211> 224

<212> PRT

<213> não identificada

<400> 260

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
1 5 10 15

Ser Glu Met Gln Thr Ser Asp Lys Pro Val Ser Phe Arg Ser Ser Ala
20 25 30

Ala Ala Thr Ala Pro Gln Met Ala Thr Ala Pro Gly Asp Ser Phe Leu
35 40 45

Thr Glu Ile Gly Glu Ala Lys Gln Gly Gln Gln Gln Asp Glu Val Ile
50 55 60

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

Gly Ile Pro His Lys Asn Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
85 90 95

617

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 261

<211> 224

<212> PRT

<213> não identificada

<400> 261

Met Glu Ile Asn Glu Lys Leu Leu Arg Gln Ile Ile Glu Asp Val Leu
1 5 10 15

Ser Glu Met Gln Thr Ser Asp Lys Pro Val Ser Phe Arg Ser Ser Thr
20 25 30

Ala Ala Ser Ala Pro Gln Ala Ala Ala Ala Gln Gly Asp Ser Phe Leu
35 40 45

Thr Glu Ile Gly Glu Ala Lys Gln Gly Gln Gln Gln Asp Glu Val Ile
50 55 60

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

6.18
A

Gly Ile Pro His Lys Ser Ile Leu Arg Glu Val Ile Ala Gly Ile Glu
85 90 95

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

Val Ala Phe Val Ala Val Glu Gly Asn Arg Leu Ser Gly Ser Gly Ile
115 120 125

Ser Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Gln Gly
130 135 140

Leu Pro Pro Leu Ser Asn Leu Glu Leu Phe Pro Gln Ala Pro Leu Leu
145 150 155 160

Thr Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala
165 170 175

Lys Arg Glu Ser Pro Gln Pro Val Pro Thr Leu Asn Asp Gln Met Ala
180 185 190

Arg Pro Lys Tyr Gln Ala Lys Ser Ala Ile Leu His Ile Lys Glu Thr
195 200 205

Lys Tyr Val Val Thr Gly Lys Asn Pro Gln Glu Leu Arg Val Ala Leu
210 215 220

<210> 262

<211> 194

<212> PRT

<213> não identificada

<400> 262

Met Glu Cys Thr Ile Glu Arg Lys Pro Val Phe Ile Leu Gln Val Ser
1 5 10 15

Glu Gly Glu Ala Ala Lys Ala Gly Asp Arg Ala Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys Tyr Gln His Lys Thr Leu Ile Asp
35 40 45

Met Pro His Lys Ala Ile Leu Lys Glu Leu Val Ala Gly Ile Glu Glu
50 55 60

619
AA

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

His Val Val Gln Asp Ala Ala Pro Val Thr Leu His Ile Ala Leu Val
180 185 190

Lys Glu

<210> 263

<211> 194

<212> PRT

<213> não identificada

<400> 263

Met Glu Cys Thr Thr Glu Arg Lys Pro Val Phe Thr Leu Gln Val Ser
1 5 10 15

Glu Gly Glu Ala Ala Lys Ala Asp Asp Arg Ala Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys Tyr Gln His Lys Thr Leu Ile Asp
35 40 45

Met Pro His Lys Ala Ile Leu Lys Glu Leu Val Ala Gly Ile Glu Glu
50 55 60

620

Glu Gly Leu His Ala Arg Val Val Arg Val Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

His Val Val Gln Asp Ala Ala Pro Val Thr Leu His Ile Ala Leu Val
180 185 190

Arg Glu

<210> 264

<211> 194

<212> PRT

<213> não identificada

<400> 264

Met Glu Cys Thr Thr Glu Arg Lys Pro Val Phe Thr Leu Gln Val Ser
1 5 10 15

Glu Gly Glu Ala Ala Lys Ala Asp Glu Arg Val Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys Tyr Gln His Lys Thr Leu Ile Asp
35 40 45

Met Pro His Asn Ala Ile Leu Lys Glu Leu Val Ala Gly Ile Glu Glu
50 55 60

621
A

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

His Val Val Gln Asp Ala Ala Pro Val Thr Leu His Ile Ala Leu Val
180 185 190

Arg Glu

<210> 265

<211> 188

<212> PRT

<213> *Clostridium perfringens*

<400> 265

Met Glu Glu Arg Thr Phe Ile Pro Glu Ile Thr Val Glu Glu Val Gly
1 5 10 15

Glu Ala Lys Val Gly Leu Arg Ser Asp Glu Val Val Ile Gly Leu Ala
20 25 30

Pro Ala Phe Leu Lys Tyr Gln Asn Lys Thr Ile Val Asp Val Pro His
35 40 45

Thr Glu Thr Leu Leu Glu Ile Ile Ala Gly Ile Glu Glu Glu Gly Leu
50 55 60

622

His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val Ser Phe Ile
65 70 75 80

Ala His Asp Ala Ala Cys Leu Ser Gly Ser Gly Ile Gly Ile Gly Ile
85 90 95

Gln Ser Lys Gly Thr Thr Val Ile His Gln Lys Asp Leu Leu Pro Leu
100 105 110

Asn Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr Pro Glu Thr
115 120 125

Tyr Arg Leu Ile Gly Lys Asn Ala Ala Lys Tyr Ala Lys Gly Glu Ser
130 135 140

Pro Thr Pro Val Pro Val Lys Asn Asp Gln Met Val Arg Pro Lys Phe
145 150 155 160

Met Ala Lys Ala Ala Leu Leu His Ile Lys Glu Thr Lys His Val Glu
165 170 175

Pro Gly Lys Lys Pro Val Gln Leu Glu Val Lys Phe
180 185

<210> 266

<211> 146

<212> PRT

<213> não identificada

<400> 266

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

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
20 25 30

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
35 40 45

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
50 55 60

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
65 70 75 80

623
A

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
85 90 95

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
100 105 110

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
115 120 125

His Val Val Gln Asp Ala Glu Pro Val Thr Leu His Ile Asp Leu Val
130 135 140

Arg Glu
145

<210> 267

<211> 148

<212> PRT

<213> *Escherichia blattae*

<400> 267

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

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
20 25 30

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
35 40 45

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
50 55 60

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
65 70 75 80

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
85 90 95

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
100 105 110

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
115 120 125

624

His Val Val Ala Asp Ala Lys Pro Ala Thr Leu Asn Ile Glu Ile Thr
 130 135 140

Arg Glu Glu Ala
 145

<210> 268
 <211> 172
 <212> PRT
 <213> não identificada

<400> 268

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
 1 5 10 15

Met Asn Ser Leu Gln Gly Glu Thr Val Thr Pro Ser Ala Thr Val Ser
 20 25 30

Ser Thr His Thr Ala Lys Val Thr Asp Tyr Pro Leu Ala Asn Lys His
 35 40 45

Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe Thr
 50 55 60

Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg Ile
 65 70 75 80

Thr Pro Glu Thr Leu Arg Leu Gln Ala Glu Ile Ala Lys Asp Ala Gly
 85 90 95

Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr Ala
 100 105 110

Val Pro Asp Asp Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro Tyr
 115 120 125

Arg Ser Thr Lys Asp Glu Leu Met Ala Ile Ala Asp Asp Leu Glu Asn
 130 135 140

Arg Tyr Gln Ala Lys Ile Cys Ala Ala Phe Val Arg Glu Ala Ala Val
 145 150 155 160

Leu Tyr Val Glu Arg Lys Lys Leu Lys Gly Asp Asp
 165 170

<210> 269

625
H

<211> 121
 <212> PRT
 <213> não identificada

<400> 269

Met Asn Thr Asp Ala Ile Glu Ser Met Val Arg Asp Val Leu Ser Arg
 1 5 10 15

Met Asn Ser Leu Gln Gly Glu Thr Val Thr Pro Ala Ala Thr Gly Ser
 20 25 30

Ser Ala His Thr Ala Lys Val Thr Asp Tyr Pro Leu Ala Asn Lys His
 35 40 45

Pro Glu Trp Val Lys Thr Ala Thr Asn Lys Thr Leu Asp Asp Phe Thr
 50 55 60

Leu Glu Asn Val Leu Ser Asn Lys Val Thr Ala Gln Asp Met Arg Ile
 65 70 75 80

Thr Pro Glu Thr Leu Arg Ile Gln Ala Asp Ile Ala Lys Asp Ala Gly
 85 90 95

Arg Asp Arg Leu Ala Met Asn Phe Glu Arg Ala Ala Glu Leu Thr Ala
 100 105 110

Val Pro Asp Asp Arg Ile Leu Glu Val
 115 120

<210> 270
 <211> 170
 <212> PRT
 <213> *Listeria monocytogenes*

<400> 270

Met Asn Gln Glu Ala Leu Glu Asn Met Val Arg Asn Ile Leu Gln Glu
 1 5 10 15

Val Asn Ser Gly Ala Val Ser Thr Thr Thr Ser Gln Lys Val Ser Gly
 20 25 30

Asp Thr Leu Thr Val Arg Asp Tyr Pro Leu Gly Thr Lys Arg Pro Glu
 35 40 45

Leu Val Lys Thr Ser Thr Ser Lys Ser Leu Asp Asp Ile Thr Leu Lys
 50 55 60

62.6
AA

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

Glu Thr Leu Lys Met Gln Ala Gln Val Ala Arg Asp Ala Gly Arg Ala
85 90 95

Thr Leu Ala Asn Asn Phe Glu Arg Ala Ala Glu Leu Thr Ile Val Pro
100 105 110

Asp Glu Arg Ile Leu Glu Ile Tyr Asn Ala Met Arg Pro Tyr Arg Ser
115 120 125

Ser Arg Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu Ser Val Tyr
130 135 140

His Ala Thr Ile Cys Ser Asn Tyr Val Arg Glu Ala Thr Gln Leu Tyr
145 150 155 160

Gln Glu Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 271

<211> 142

<212> PRT

<213> não identificada

<400> 271

Met Asn Asp Asn Ile Met Thr Ala Gln Asp Tyr Pro Leu Ala Thr Arg
1 5 10 15

Cys Pro Glu Lys Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
20 25 30

Thr Leu Glu Asn Val Leu Ala Gly Arg Val Gly Pro Gln Asp Val Arg
35 40 45

Ile Ser Gln Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
50 55 60

Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80

Ala Ile Pro Asp Ala Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
85 90 95

Phe Arg Ser Ser Leu Ala Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu
100 105 110

626
A

His Thr Trp His Ala Thr Val Asn Ala Gly Phe Val Arg Glu Ser Ala
115 120 125

Glu Val Tyr Gln Gln Arg Asn Lys Leu Arg Lys Gly Ser Gln
130 135 140

<210> 272
<211> 142
<212> PRT
<213> não identificada

<400> 272

Met Asn Glu Asn Ile Met Thr Ala Gln Asp Tyr Pro Leu Ala Thr Arg
1 5 10 15

Cys Pro Glu Lys Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
20 25 30

Thr Leu Glu Asn Val Leu Ala Gly Arg Val Gly Pro Gln Asp Val Arg
35 40 45

Ile Ser Gln Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
50 55 60

Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80

Ala Ile Pro Asp Ala Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
85 90 95

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

His Thr Trp His Ala Thr Val Asn Ala Gly Phe Val Arg Glu Ser Ala
115 120 125

Asp Val Tyr Gln Gln Arg Asn Lys Leu Arg Lys Gly Ser Gln
130 135 140

<210> 273
<211> 142
<212> PRT
<213> não identificada

<400> 273

Met Asn Asp Asn Ile Met Thr Ala Gln Asp Tyr Pro Leu Ala Thr Arg
1 5 10 15

623
A

Cys Pro Glu Lys Ile Gln Thr Pro Thr Gly Lys Pro Leu Thr Glu Ile
20 25 30

Thr Leu Glu Asn Val Leu Ala Gly Arg Val Gly Pro Gln Asp Val Arg
35 40 45

Ile Ser Gln Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
50 55 60

Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80

Ala Ile Pro Asp Ala Arg Ile Leu Glu Ile Tyr Asn Ala Leu Arg Pro
85 90 95

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

His Thr Trp His Ala Thr Val Asn Ala Gly Phe Val Arg Glu Ser Ala
115 120 125

Glu Val Tyr Gln Gln Arg Asn Lys Leu Arg Lys Gly Ser Gln
130 135 140

<210> 274

<211> 170

<212> PRT

<213> *Listeria monocytogenes*

<400> 274

Met Asn Gln Glu Ala Leu Glu Asn Met Val Arg Asn Ile Leu Gln Glu
1 5 10 15

Val Asn Ser Gly Ala Val Ser Thr Thr Thr Ser Gln Lys Val Ser Gly
20 25 30

Asp Thr Leu Thr Val Arg Asp Tyr Pro Leu Gly Thr Lys Arg Pro Glu
35 40 45

Leu Val Lys Thr Ser Thr Ser Lys Ser Leu Asp Asp Ile Thr Leu Lys
50 55 60

Ser Val Leu Asp Gly Thr Ile Lys Pro Glu Asp Val Arg Val Thr Val
65 70 75 80

Glu Thr Leu Lys Met Gln Ala Gln Val Ala Arg Asp Ala Gly Arg Ala
85 90 95

629
AS

Thr Leu Ala Asn Asn Phe Glu Arg Ala Ala Glu Leu Thr Ile Val Pro
100 105 110

Asp Glu Arg Ile Leu Glu Ile Tyr Asn Ala Met Arg Pro Tyr Arg Ser
115 120 125

Ser Arg Glu Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu Ser Val Tyr
130 135 140

His Ala Thr Ile Cys Ser Asn Tyr Val Arg Glu Ala Ala Gln Leu Tyr
145 150 155 160

Gln Glu Arg Lys Lys Leu Lys Gly Asp Asn
165 170

<210> 275

<211> 2432

<212> DNA

<213> *Erwinia caratovora subespécie atroseptica*

<400> 275

atgtctgacg gacgactcac cgcacttttt cctgcattcc cacacccggc gtccaatcag 60
cccgtatttg ccgaggcttc accgcacgac gacgagttaa tgacgcaggc cgtaccgcag 120
gtttcctgtc agcaggcggtt ggcgattgcy cagcaagaat atggcttgctc tgggcagatg 180
tcgctgcttc agggcgagcg tgatgtgaat ttctgtctga cggtgacgcc agatgaacgc 240
tacatgctga aagtcaccaa tgcggcagaa cctgccgacg tcagcaattt ccaaaccgcg 300
ctgctgctgc atcttgcccg tcaggcacct gaactgcccg taccgcgtat caggctgaca 360
aaagcgggtc agtcggaaac aggcgttgag atcgatggtg tactgctgcy tgtgcggctt 420
gtgagctatc tggcaggaat gccgcagtat ctggcctcac cgtcaacggc gctgatgccg 480
cagttggggg gaacgctggc gcagttggat aacgcgcttc acagctttac gcatccggcg 540
gcaaaccgtg cgctgctgtg ggatatcagc cgggcagagc aggtgcgtcc ttacctcgat 600
ttcgtttctg aaccgcagca gtatcagcat cttcagcgta tttttgaccg ttatgacagt 660
aacgttgctc ctctgttgac gacgctacgt cgtcaggctc ttcataacga tctgaatccg 720
cataacgtgc tgggtgatgg atcgctgccg acgcgggtta ctggcattat cgattttggc 780
gatgccgtat ttgccccgtt aatttgcgaa gtcgcgacgg cactggcgta tcagatcggc 840
gatggaaccg atttggttga gcatgttggt cgtttgttg cggcctatca ccaacgcatt 900
ccgttagcac cggaggagat tgcgctgtta cccgatctga tagcgacccg tatggcgctg 960
accctgacca ttgcgcagtg gcgagcatcg cgttatcccg acaatcggga gtatctgctg 1020

cgtaacgtgc cgcgctgttg gcacagtttg cagcgcattg cgacctattc ccatgcgcaa 1080
 tttttgactc gcctacagca ggtttgcccc gagaatgcmc gatgaaccag aaaggaatga 1140
 cgtctatgac gtctgaaatg acagcgacag aagctttgct ggcmgcmccgt cagcmgagtgt 1200
 tgggcmggcmg ttatcmccctg ttttatgaag agccgctgca tctcmgcmcmg ggcmgagggcmg 1260
 tgtggctgtt cgatcaccaa gggaaacgtt atctggatgt ctacaataat gtggcttcmg 1320
 tcmgacattg ccccccmg gtggttgaag ccmgmgcmg acagagcmga caactcaata 1380
 cccacaccmg ctatttgcac ccmgcmgtt tcmgttttgc ggaagatttg ctgagcmgaat 1440
 tcccccmga attgaacaat gtaatgctga cctgtaccmg cagtgaggtt aacgatctgg 1500
 cmgtcmgtat cccccgat gtcacggcmg ggacggggat gttggtgacm cmgtggcmgt 1560
 atcacggcmgt gaccagcmg ctggcmgaac tgtctccgtc gctgggggat ggcmgttgm 1620
 gcmgtagcca tgtgaagctg atcmgcmcm cagacactta tcmgtcmccc ggtgcatttc 1680
 ttaccagcat tcmgtgaagcm ctggcmcmga tgcaacggga aggtattcmgt cctgcmgmcm 1740
 tgmgtgtaga taccattttt tccagcmgt gcmgttctg tcmgcmggaa ggcgaatgm 1800
 cacagggcmg ggcmgttgatc cmgtcmggcm gcmggctgtt tattgcmgat gaagtgcagc 1860
 cgggcttcmg gcmgaccmgg gaatcmctgt ggggcttgc gcmccacaat gcmgtccctg 1920
 atttgmgtgag tctagggaaa ccmgtggcm acggacatcc cmgtcmgtga ttgmgtggggc 1980
 gttccmgtct gttcmgcmga tttgggmcm atgtgcmcta tttcaatacc tttggcmga 2040
 atccgmgttc cmgtcmggcm gcmgaccmg tgmgtcmgggt gattcmggaa gagcmgttgc 2100
 agcagaatgc ccagcmggtc ggtgattatc tcmggcaagg gttcmgcmga ctggcmgcmg 2160
 atttcccmgt gattgmgtgat attcmggctt acggcmctgtt tattgmgtgcm gagctgmgtca 2220
 gcmgatcmga aagtaaaacm ccmgcaagt gattcmcmgt gcmggtgmgt aatgcmgatgc 2280
 gccaacmgtg tgmgtcmatc agcmgcmcm ggccagcmg gaacatactg aaaattcmgc 2340
 cmcmgtgmgt gtttctgmga gaacaccmcm atgtgttctt aaccacmgt agtgacmgtt 2400
 tagcmgtcat cmgcmctcmgt gcacagagat aa 2432

REIVINDICAÇÕES

1. CÉLULA HOSPEDEIRA MICROBIANA RECOMBINANTE, caracterizada pelo fato de que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol;
- iv) 2,3-butanodiol para 2-butanona; e
- v) 2-butanona para 2-butanol;

sendo que pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana e sendo que a dita célula hospedeira microbiana produz 2-butanol.

2. CÉLULA HOSPEDEIRA MICROBIANA RECOMBINANTE, caracterizada pelo fato de que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol; e
- iv) 2,3-butanodiol para 2-butanona;

sendo que pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana e sendo que a dita célula hospedeira microbiana produz 2-butanona.

3. CÉLULA HOSPEDEIRA, de acordo com as reivindicações 1 ou 2, caracterizada pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de piruvato em alfa-acetolactato é acetolactato sintase.

4. CÉLULA HOSPEDEIRA, de acordo com as reivindicações 1

ou 2, caracterizada pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de alfa-acetolactato em acetoína é acetolactato descarboxilase.

5 5. CÉLULA HOSPEDEIRA, de acordo com as reivindicações 1 ou 2, caracterizada pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de acetoína em 2,3-butanodiol é butanodiol desidrogenase.

6. CÉLULA HOSPEDEIRA, de acordo com as reivindicações 1 ou 2, caracterizada pelo fato de que o polipeptídeo que catalisa um substrato para 10 produzir a conversão de 2,3-butanodiol em 2-butanona é diol desidratase ou glicerol desidratase.

7. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 1, caracterizada pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de 2-butanona em 2-butanol é butanol desidrogenase.

15 8. CÉLULA HOSPEDEIRA, de acordo com as reivindicações 1 ou 2, caracterizada pelo fato de ser selecionada do grupo que consiste em: uma bactéria, uma cianobactéria, um fungo filamentoso e uma levedura.

9. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 8, caracterizada pelo fato de ser membro de um gênero selecionado do grupo 20 consistindo em *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Pediococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* e *Saccharomyces*.

10. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 3, 25 caracterizada pelo fato de que a acetolactato sintase tem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 4, SEQ ID nº 77 e SEQ ID nº 79 com base no método de alinhamento Clustal W usando os

parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

11. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 4, caracterizada pelo fato de que a acetolactato descarboxilase tem uma
5 seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 2, SEQ ID nº 81, e SEQ ID nº 83 com base no método de alinhamento Clustal W usando os parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH
● PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

10 12. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 5, caracterizada pelo fato de que a butanodiol desidrogenase tem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 6, SEQ ID nº 85, SEQ ID nº 87 e SEQ ID nº 89 com base no método de alinhamento Clustal W
15 usando os parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

13. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6, caracterizada pelo fato de que a diol desidratase ou glicerol desidratase compreendem subunidades grandes, médias e pequenas de comprimento
20 total e cada uma tem um parâmetro E-Valor de 0,01 ou menos quando pesquisada usando um perfil Hidden Markov Model, preparado com o uso das subunidades grandes das SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164; das subunidades média das SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165; e das subunidades pequenas das SEQ ID nº 12, 103, 109, 137, 140,
25 143, 150 e 166, sendo que cada pesquisa é executada usando-se o algoritmo hmmsearch com o parâmetro Z ajustado em 1 bilhão.

14. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6, caracterizada pelo fato de que a diol desidratase ou a glicerol desidratase é

identificada por meio de um processo que compreende as etapas de:

a) gerar um perfil "Hidden Markov Model" a partir do alinhamento da seqüência de aminoácidos correspondendo às subunidades grandes, médias e pequenas das enzimas de diol e glicerol desidratase, sendo que:

i) A subunidade grande compreende uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164;

ii) A subunidade média compreende uma seqüência de aminoácidos selecionado do grupo consistindo em SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165; e

iii) A subunidade pequena compreende uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166;

b) pesquisar pelo menos uma base de dados pública de seqüências de proteína contendo seqüências de diol e glicerol desidratases com o perfil "Hidden Markov Model" de (a) usando o algoritmo hmmsearch no qual o parâmetro Z é ajustado em 1 bilhão e o parâmetro E-Valor é ajustado em 0,01, para identificar um primeiro conjunto de dados de seqüências de aminoácidos de diol e glicerol desidratase; e

c) remover quaisquer seqüências parciais do primeiro conjunto de dados de (b) para gerar um segundo conjunto de dados de seqüências de aminoácidos de diol e glicerol desidratases, no qual as enzimas de diol desidratase e de glicerol desidratase são identificadas.

15. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6, caracterizada pelo fato de que a diol desidratase ou glicerol desidratase compreendem uma subunidade grande que compreende uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de

aminoácidos selecionada do grupo consistindo em SEQ ID nº 8, 93, 99, 105, 135, 138, 141, 146, 164, 167, 170, 173, 176, 179, 182, 185, 188, 191, 194, 197, 200, 203, 206, 209, 212, 215, 218, 221, 224, 227, 130, 243, 254, 255, 256, 257, 258 e 259, com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e séries de matriz de peso de proteína Gonnet 250.

16. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6, caracterizada pelo fato de que a diol desidratase ou glicerol desidratase compreendem uma subunidade média que compreende uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 10, 95, 101, 107, 136, 139, 142, 148, 165, 168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 222, 225, 228, 231, 244, 250, 252, 260, 261, 262, 263, 264, 265, 266 e 167 com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e as séries de matriz de peso de proteína Gonnet 250.

17. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6, caracterizada pelo fato de que a diol desidratase ou glicerol desidratase compreendem uma subunidade pequena que compreende uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 12, 97, 103, 109, 137, 140, 143, 150, 166, 169, 172, 175, 178, 181, 184, 187, 190, 193, 196, 199, 202, 205, 208, 211, 214, 217, 220, 223, 226, 229, 232, 234, 236, 238, 240, 242, 245, 248, 249, 251, 253, 268, 270, 271, 272, 273 e 274, com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e nas séries de matriz de peso de proteína Gonnet 250.

18. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6,

caracterizada pelo fato de que a diol desidratase ou glicerol desidratase compreendem subunidades grande, média e pequena unidas, que compreendem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que
5 consiste nas SEQ ID nº 233, 235, 237, 239, 241, 246 e 247, com base no método de alinhamento Clustal W usando os parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e nas séries de matriz de peso de proteína Gonnet 250.

19. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 6,
10 caracterizada pelo fato de que a diol desidratase ou glicerol desidratase compreendem subunidades grande, média e pequena unidas e têm pelo menos 95% de identidade com uma seqüência de aminoácidos que compreende todas as três seqüências de aminoácidos que codificam subunidades grandes, pequenas e médias, selecionadas do grupo que consiste
15 em:

- a) SEQ ID nº 8, SEQ ID nº 10 e SEQ ID nº 12;
- b) SEQ ID nº 93, SEQ ID nº 95 e SEQ ID nº 97;
- c) SEQ ID nº 99, SEQ ID nº 101, e SEQ ID nº 103;
- d) SEQ ID nº 105, SEQ ID nº 107 e SEQ ID nº 109;
- 20 e) SEQ ID nº 135, SEQ ID nº 136 e SEQ ID nº 137;
- f) SEQ ID nº 138, SEQ ID nº 139 e SEQ ID nº 140;
- g) SEQ ID nº 146, SEQ ID nº 148 e SEQ ID nº 150;
- h) SEQ ID nº 141, SEQ ID nº 142 e SEQ ID nº 143; e
- i) SEQ ID nº 164, SEQ ID nº 165 e SEQ ID nº 166;

25 com base no método de alinhamento Clustal W e com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

20. CÉLULA HOSPEDEIRA, de acordo com a reivindicação 7

caracterizada pelo fato de que a butanol desidrogenase tem uma sequência de aminoácidos tendo pelo menos 95% de identidade com uma sequência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 14, SEQ ID nº 72, SEQ ID nº 75 e SEQ ID nº 91 com base no método de alinhamento Clustal W e com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

21. MÉTODO PARA PRODUÇÃO DE 2-BUTANOL, caracterizado pelo fato de que compreende:

1) fornecer uma célula hospedeira microbiana recombinante que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol;
- iv) 2,3-butanodiol para 2-butanona; e
- v) 2-butanona para 2-butanol;

sendo que a pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana;

2) colocar a célula hospedeira de (1) em contato com um substrato fermentável de carbono em um meio de fermentação nas condições em que o 2-butanol é produzido.

22. MÉTODO PARA PRODUÇÃO DE 2-BUTANONA, caracterizado pelo fato de que compreende:

1) fornecer uma célula hospedeira microbiana recombinante que compreende pelo menos uma molécula de DNA que codifica um polipeptídeo que catalisa um substrato para produzir a conversão selecionada

342

do grupo consistindo em:

- i) piruvato para alfa-acetolactato;
- ii) alfa-acetolactato para acetoína;
- iii) acetoína para 2,3-butanodiol; e
- iv) 2,3-butanodiol para 2-butanona:

sendo que a pelo menos uma molécula de DNA é heteróloga à dita célula hospedeira microbiana;

2) colocar a célula hospedeira de (1) em contato com um substrato fermentável de carbono em um meio de fermentação nas condições em que a 2-butanona é produzida.

23. MÉTODO, de acordo com as reivindicações 21 ou 22, caracterizado pelo fato de que o substrato de carbono fermentável é selecionado do grupo que consiste em monossacarídeos, oligossacarídeos e polissacarídeos.

24. MÉTODO, de acordo com as reivindicações 21 ou 22, caracterizado pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de piruvato em alfa-acetolactato é acetolactato sintase.

25. MÉTODO, de acordo com as reivindicações 21 ou 22, caracterizado pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de alfa-acetolactato em acetoína é acetolactato descarboxilase.

26. MÉTODO, de acordo com as reivindicações 21 ou 22, caracterizado pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de acetoína em 2,3-butanodiol é butanodiol desidrogenase.

27. MÉTODO, de acordo com as reivindicações 21 ou 22, caracterizado pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de 2,3-butanodiol em 2-butanona é diol desidratase ou glicerol desidratase.

28. MÉTODO, de acordo com a reivindicação 21, caracterizado pelo fato de que o polipeptídeo que catalisa um substrato para produzir a conversão de 2-butanona em 2-butanol é butanol desidrogenase.

29. MÉTODO, de acordo com as reivindicações 21 ou 22, caracterizado pelo fato de que a célula é selecionada do grupo que consiste em: uma bactéria, uma cianobactéria, um fungo filamentoso e uma levedura.

30. MÉTODO, de acordo com a reivindicação 29, caracterizado pelo fato de que a célula é membro de um gênero selecionado do grupo consistindo em *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Pediococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* e *Saccharomyces*.

31. MÉTODO, de acordo com a reivindicação 24, caracterizado pelo fato de que a acetolactato sintase tem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 4, SEQ ID nº 77 e SEQ ID nº 79 com base no método de alinhamento Clustal W com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

32. MÉTODO, de acordo com a reivindicação 25, caracterizado pelo fato de que a acetolactato descarboxilase tem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 2, SEQ ID nº 81, e SEQ ID nº 83 com base no método de alinhamento Clustal W e com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

33. MÉTODO, de acordo com a reivindicação 26, caracterizado

pelo fato de que a butanodiol desidrogenase tem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 6, SEQ ID nº 85, SEQ ID nº 87 e SEQ ID nº 89 com base no método de alinhamento Clustal W e com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

34. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou glicerol desidratase compreendem subunidades grandes, médias e pequenas de comprimento total e cada uma tem um parâmetro E-Valor de 0,01 ou menos quando pesquisada usando-se um perfil Hidden Markov Model, preparado com o uso das subunidades grandes das SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164; das subunidades médias das SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165, e das subunidades pequenas das SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166, cada pesquisa sendo executada usando-se o algoritmo hmmsearch com o parâmetro Z ajustado em 1 bilhão.

35. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou a glicerol desidratase é identificada por meio de um processo que compreende as etapas de:

a) gerar um perfil "Hidden Markov Model" a partir do alinhamento da seqüência de aminoácidos correspondendo às subunidades grande, média e pequena das enzimas diol e glicerol desidratase, sendo que:

i) A subunidade grande compreende uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 8, 99, 105, 135, 138, 141, 146 e 164;

ii) A subunidade média compreende uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 10, 101, 107, 136, 139, 142, 148 e 165; e

iii) A subunidade pequena compreende uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 12, 103, 109, 137, 140, 143, 150 e 166;

b) pesquisar pelo menos uma base de dados pública de seqüências de proteína contendo seqüências de diol e glicerol desidratases com o perfil "Hidden Markov Model" de (a) usando o algoritmo hmmsearch com o parâmetro Z ajustado em um 1 bilhão e o parâmetro E-Valor ajustado em 0,01, para identificar um primeiro conjunto de dados de seqüências de aminoácidos de diol e glicerol desidratase; e

c) remover quaisquer seqüências parciais do primeiro conjunto de dados de (b) para gerar um segundo conjunto de dados de seqüências de aminoácidos de diol e glicerol desidratases no qual as enzimas de diol desidratase e de glicerol desidratase são identificadas.

36. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou glicerol desidratase compreendem uma subunidade grande que compreende uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 8, 93, 99, 105, 135, 138, 141, 146, 164, 167, 170, 173, 176, 179, 182, 185, 188, 191, 194, 197, 200, 203, 206, 209, 212, 215, 218, 221, 224, 227, 130, 243, 254, 255, 256, 257, 258 e 259, com base no método de alinhamento Clustal W e com o uso dos parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

37. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou glicerol desidratase compreendem uma subunidade média que compreende uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 10, 95, 101, 107, 136, 139, 142, 148, 165,

168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 222, 225, 228, 231, 244, 250, 252, 260, 261, 262, 263, 264 265, 266 e 167 com base no método de alinhamento Clustal W e com o uso dos parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

38. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou glicerol desidratase compreendem uma subunidade pequena que compreende uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo consistindo em SEQ ID nº 12, 97, 103, 109, 137, 140, 143, 150, 166, 169, 172, 175, 178, 181, 184, 187, 190, 193, 196, 199, 202, 205, 208, 211, 214, 217, 220, 223, 226, 229, 232, 234, 236, 238, 240, 242, 245, 248, 249, 251, 253, 268, 270, 271, 272, 273 e 274, com base no método de alinhamento Clustal W e com o uso dos parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

39. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou glicerol desidratase compreendem subunidades grande, média e pequena unidas, que compreendem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que consiste nas SEQ ID nº 233, 235, 237, 239, 241, 246 e 247, com base no método de alinhamento Clustal W e com o uso dos parâmetros pré-estabelecidos de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

40. MÉTODO, de acordo com a reivindicação 27, caracterizado pelo fato de que a diol desidratase ou glicerol desidratase compreendem subunidades grande, média e pequena unidas e têm pelo menos 95% de identidade com uma seqüência de aminoácidos que compreende todas as três seqüências de aminoácidos que codificam subunidades grandes, pequenas e

médias, selecionadas do grupo que consiste em:

- a) SEQ ID nº 8, SEQ ID nº 10 e SEQ ID nº 12;
- b) SEQ ID nº 93, SEQ ID nº 95 e SEQ ID nº 97;
- c) SEQ ID nº 99, SEQ ID nº 101, e SEQ ID nº 103;
- 5 d) SEQ ID nº 105, SEQ ID nº 107 e SEQ ID nº 109;
- e) SEQ ID nº 135, SEQ ID nº 136 e SEQ ID nº 137;
- f) SEQ ID nº 138, SEQ ID nº 139 e SEQ ID nº 140;
- g) SEQ ID nº 146, SEQ ID nº 148 e SEQ ID nº 150;
- h) SEQ ID nº 141, SEQ ID nº 142 e SEQ ID nº 143; e
- 10 i) SEQ ID nº 164, SEQ ID nº 165 e SEQ ID nº 166;

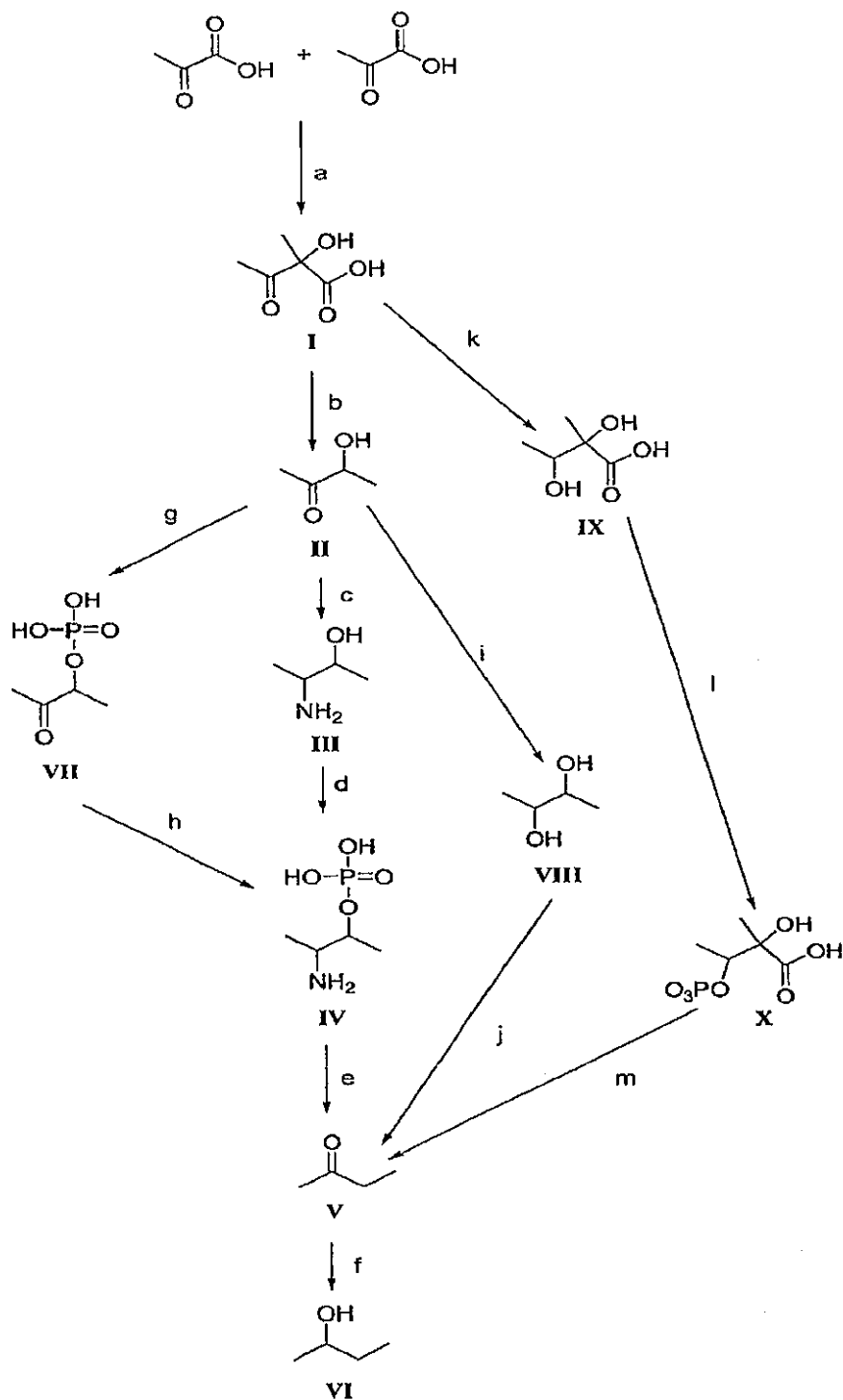
com base no método de alinhamento Clustal W e com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

41. MÉTODO, de acordo com a reivindicação 28, caracterizado
15 pelo fato de que a butanol desidrogenase tem uma seqüência de aminoácidos tendo pelo menos 95% de identidade com uma seqüência de aminoácidos selecionada do grupo que consiste em SEQ ID nº 14, SEQ ID nº 72, SEQ ID nº 75 e SEQ ID nº 91 com base no método de alinhamento Clustal W e com o uso dos parâmetros predeterminados de GAP PENALTY=10, GAP LENGTH
20 PENALTY=0,1 e das séries de matriz de peso de proteína Gonnet 250.

42. 2-BUTANOL, caracterizado pelo fato de que contém meio de fermentação de produto produzido pelo método conforme definido na reivindicação 21.

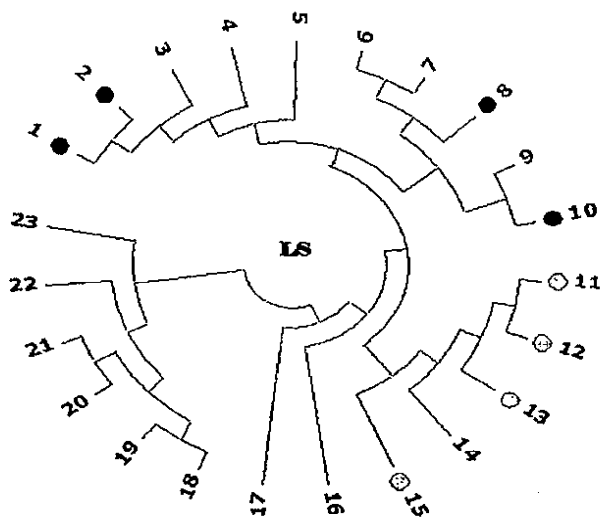
43. 2-BUTANONA, caracterizada pelo fato de que contém meio de
25 fermentação de produto produzido pelo método conforme definido na reivindicação 22.

Figura 1



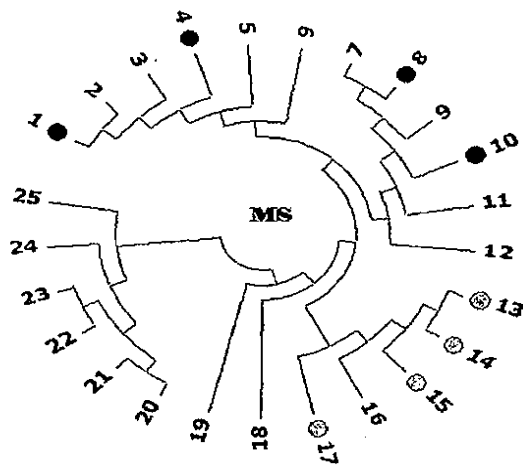
632
PA

Figura 2



1	gi 4063702 DD <i>Klebsiella pneumoniae</i>
2	gi 6980836 DD <i>Klebsiella Oxytoca</i> [pdbj]1D
3	gi 123442946 <i>Yersinia enterocolitica</i>
4	gi 116872527 <i>Listeria welshimeri</i>
5	gi 125717397 <i>Streptococcus sanguinis</i>
6	gi 37542647 <i>Lactobacillus diolivorans</i>
7	gi 37542653 <i>Lactobacillus hilgardii</i>
8	gi 18857678 DD <i>Lactobacillus collinoides</i>
9	gi 86604356 <i>Lactobacillus reuteri</i>
10	gi 116334196 DD <i>Lactobacillus brevis</i>
11	gi 1169287 GD <i>Citrobacter freundii</i>
12	gi 24158719 GD <i>Klebsiella pneumoniae</i>
13	gi 60099613 GD <i>Escherichia blattae</i>
14	gi 18309911 <i>Clostridium perfringens</i>
15	gi 3360389 GD <i>Clostridium pasteurianum</i>
16	gi 114844379 <i>Thermoanaerobacter ethanolicus</i>
17	gi 118472362 <i>Mycobacterium smegmatis</i>
18	gi 108801917 <i>Mycobacterium</i> sp. MCS
19	gi 118472557 <i>Mycobacterium smegmatis</i> str. MC2 155
20	gi 120406505 <i>Mycobacterium vanbaalenii</i> PYR-1
21	gi 89342701 <i>Mycobacterium flavescens</i>
22	gi 118470012 <i>Mycobacterium smegmatis</i> str. MC2 155
23	gi 13475607 <i>Mesorhizobium loti</i> MAFF303099

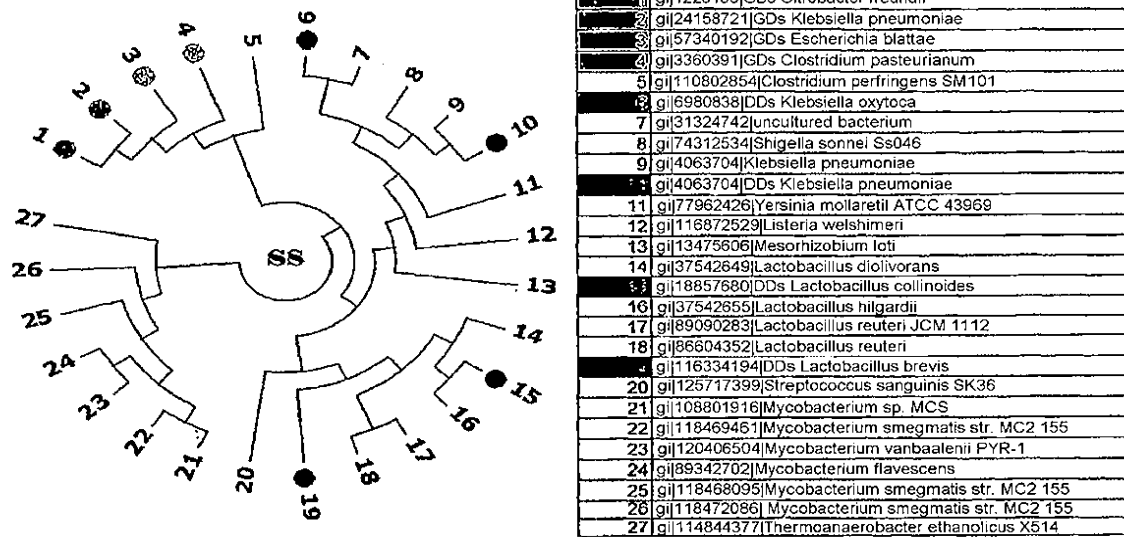
Figura 3

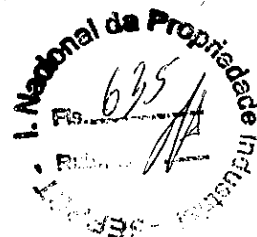


1	gi 6980837 DDm Klebsiella Oxytoca
2	gi 74312533 Shigella sonnei Ss046
3	gi 31324737 uncultured bacterium
4	gi 94470233 DDm Klebsiella pneumoniae
5	gi 77962425 Yersinia mollaretii ATCC 43969
6	gi 116872528 Listeria welshimeri
7	gi 37542648 Lactobacillus diolivorans
8	gi 18857679 DDm Lactobacillus collinoides
9	gi 37542654 Lactobacillus hilgardii
10	gi 116334195 DDm Lactobacillus brevis
11	gi 86604354 Lactobacillus reuteri
12	gi 125717398 Streptococcus sanguinis SK36
13	gi 1229154 GDm Citrobacter freundii
14	gi 24158720 GDm Klebsiella pneumoniae
15	gi 57340191 GDm Escherichia blattae
16	gi 18309912 Clostridium perfringens
17	gi 3360390 GDm Clostridium pasteurianum
18	gi 114844378 Thermoanaerobacter ethanolicus X514
19	gi 118467606 Mycobacterium smegmatis str. MC2 155
20	gi 108801917 Mycobacterium sp. MCS
21	gi 118472557 Mycobacterium smegmatis str. MC2 155
22	gi 120406505 Mycobacterium vanbaalenii PYR-1
23	gi 89342701 Mycobacterium flavescens
24	gi 118470012 Mycobacterium smegmatis str. MC2 155
25	gi 13475607 Mesorhizobium loti MAFF303099

634
AA

Figura 4





RESUMO

"CÉLULAS HOSPEDEIRAS MICROBIANAS RECOMBINANTES, MÉTODO PARA PRODUÇÃO DE 2-BUTANOL, MÉTODO PARA PRODUÇÃO DE 2-BUTANONA, 2-BUTANOL E 2-BUTANONA"

5 A presente invenção apresenta métodos para a produção fermentativa de álcoois de quatro carbonos. Especificamente butanol, de preferência 2-butanol, é produzido pelo crescimento fermentativo de bactérias recombinantes que expressam 2-butanol por via biossintética. Os
10 microorganismos recombinantes e os métodos da invenção podem também ser adaptados para produzir 2-butanona, um intermediário nas vias biossintéticas do 2-butanol aqui apresentadas.