



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

(11) PI 0314543-3 B1

(22) Data do Depósito: 08/10/2003

(45) Data de Concessão: 14/06/2016



(54) Título: CÉLULAS HOSPEDEIRAS, DNA RECOMBINANTE, UTILIZAÇÃO DOS MESMOS, PROCESSOS DE PRODUÇÃO DE UM POLIPEPTÍDEO IMPLICADO NA BIOSÍNTESE DAS ESPIRAMICINAS, DE PRODUÇÃO DE MACROLÍDEOS E DE PRODUÇÃO DE ESPIRAMICINA(S), E MICRO-ORGANISMO TRANSGÊNICO MUTANTE PRODUTOR DE MACROLÍDEO

(51) Int.Cl.: C12N 15/52; C07K 14/36; C12P 19/62; A61K 31/70

(30) Prioridade Unionista: 08/10/2002 FR 02 12489, 27/02/2003 FR 03 02439

(73) Titular(es): CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (C.N.R.S.). AVENTIS PHARMA S.A.

(72) Inventor(es): MARIE-HÉLÈNE BLONDELET-ROUAULT, HÉLÈNE DOMINGUEZ, EMMANUELLE DARBON-RONGERE, CLAUDE GERBAUD, ANNE GONDRAN, FATMA KARRAY, PATRICIA LACROIX, NATHALIE OESTREICHER-MERMET-BOUVIER, JEAN-LUC PERNODET, KARINE TUPHILE

Relatório Descritivo da Patente de Invenção para **"CÉLULAS HOSPEDEIRAS, DNA RECOMBINANTE, UTILIZAÇÃO DOS MESMOS, PROCESSOS DE PRODUÇÃO DE UM POLIPEPTÍDEO IMPLICADO NA BIOSÍNTESE DAS ESPIRAMICINAS, DE PRODUÇÃO DE MACROLÍDEOS E DE PRODUÇÃO DE ESPIRAMICINA(S), E MICRO-ORGANISMO TRANSGÊNICO MUTANTE PRODUTOR DE MACROLÍDEO"**.

A presente invenção refere-se ao isolamento e à identificação de novos genes da via de biossíntese das espiramicinas e de novos polipeptídeos implicados nesse biossíntese. Ela é também relativa à utilização desses genes com a finalidade de aumentar as taxas de produção e a pureza da espiramicina produzida.

A invenção refere-se também à utilização desses genes para a construção de mutantes que podem levar à síntese de novos antibióticos ou a formas derivadas de espiramicinas. A invenção refere-se ainda às moléculas produzidas graças à expressão desses genes e enfim composições farmacologicamente ativas de uma molécula produzida, graças à expressão desses genes.

As *Streptomyces* são bactérias filamentosas Gram positivo do solo. Elas exercem um papel importante na decomposição e a mineralização das matérias orgânicas graças à grande diversidade das enzimas degradativas que elas secretam. Elas apresentam fenômenos de diferenciações morfológicas únicas nos procariotes acompanhando-se de uma diferenciação metabólica caracterizada pelo fato de produção de metabólitos secundários apresentando uma extraordinária diversidade de estruturas químicas e de atividades biológicas. Dentre esses metabólitos, encontram-se as piramicinas produzidas no estado natural pela bactéria *Streptomyces ambofaciens*.

A espiramicina é um antibiótico da família dos macrolídeos útil tanto em medicina veterinária, quanto em medicina humana. Os macrolídeos são caracterizados pela presença de um ciclolactona sobre o qual são enxertados um açúcar ou vários. *Streptomyces ambofaciens* produzido no estado natural a espiramicina I, II e III (conforme Figura 1), todavia a atividade antibiótica da espiramicina I é nitidamente superior àquela das espiramicina II e III (Liu et al., 1999). A molécula de espiramicina I é constituída de um macrociclolactônico, denominado platenolídeo e de três açúcares: a

forosamina, a micaminose e a micarose (cf/ Figura 1). A atividade antibiótica das espiramicinas é devido a uma inibição da síntese protéica nos procariotes por um mecanismo que implica a ligação do antibiótico ao ribossoma bacteriano.

5 Certos compostos membros da família do macrolídeos e possuindo também um ciclolactona deram lugar a utilizações fora do domínio dos antibióticos. Assim, o produto FK506 tem efeitos imunossuppressores e oferece perspectivas de aplicação terapêutica no domínio do transplante de órgão, da artrite reumatóide e mais geralmente em patologias ligadas a reações autoimunes. Outros macrolídeos como a avermectina têm atividades inseticidas e anti-helmínticas.

 Numerosas vias de biossíntese, referentes aos antibióticos pertencentes a classes variadas, assim como outros metabólitos secundários (para uma revista, Chater K. 1990), foram até hoje estudadas nos
15 *Streptomyces*. Todavia, o conhecimento das vias de biossíntese das espiramicinas é apenas muito parcial até hoje.

 A biossíntese das espiramicinas é um processo complexo que comporta numerosas etapas e implicando numerosas enzimas (Omura et al., 1979a. Omura et al., 1979b). As espiramicinas pertencem à ampla classe
20 dos poliquetídeos que agrupa moléculas complexas particularmente abundantes nos microorganismos do solo. Essas moléculas são agrupadas não por analogia de estrutura, mas por uma certa similaridade ao nível de etapas de sua via de biossíntese. Com efeito, os poliquetídeos são produzidos por uma série complexa de reações, mas possuem em comum ter em sua via de
25 biossíntese uma série de reações catalisadas por uma enzima ou várias denominadas "poliketides synthases"(PKS). Nas *streptomyces ambofaciens*, a biossíntese do macrociclolactônico das espiramicinas (platenolídeo) é realizada por uma série de oito módulos codificados por cinco genes PKS diferentes (Kuhstoss S., 1996, patente 5.945.320). As espiramicinas são obtidas
30 a partir desse ciclolactona. Todavia, as diversas etapas e enzimas implicadas na síntese dos açúcares, assim como sua ligação ao ciclolactona e a modificação desse ciclo para a obtenção das espiramicinas continuam ainda

desconhecidas até hoje.

A patente US 5.514.544 descreve a clonagem de uma seqüência denominada *srmR* nos *streptomyces ambofaciens*. Ele representa a hipótese nessa patente que o gene *srmR* codifica uma proteína reguladora da trans-
5 crição dos genes implicados na biossíntese dos macrolídeos.

Em 1987, Richardson e colaboradores (Richardson et al., 1987) mostraram que a resistência a espiramicina de *S. ambofaciens* é conferida por pelo menos três genes, estes últimos foram denominados *srmA*, *srmB* e *srmC*. A patente US 4.886.757 descreve mais particularmente a caracteriza-
10 ção de um fragmento de DNA de *S. ambofaciens* que contém o gene *srmC*. Entretanto, a seqüência desse gene não foi divulgada. Em 1990, Richardson e colaboradores (Richardson et al., 1990) estabeleceu a hipótese da existência de três genes de biossíntese da espiramicina próxima de *srmB*. A patente US 5.098.837 refere-se à clonagem de cinco genes potencialmente
15 implicados na biossíntese da espiramicina. Esses genes foram batizados *srmD*, *srmE*, *srmF*, *srmG* e *srmH*.

Uma das dificuldades maiores na produção de compostos, tais como as espiramicinas reside no fato de volumes muito grandes de fermentação são necessários à produção de uma quantidade relativamente pequena de produto. Portanto, é desejável poder aumentar a eficácia de produção
20 dessas moléculas, a fim de diminuir-lhe os custos de produção.

A via de biossíntese das espiramicinas é um processo complexo e seria desejável identificar e suprimir as reações parasitas que poderiam existir durante esse processo. A finalidade dessa manipulação é de obter um
25 antibiótico mais puro e/ou uma melhoria da produtividade. A esse respeito, *streptomyces ambofaciens* produzido no estado natural a espiramicinas I, II e III (conforme Figura 1), todavia a atividade do antibiótico da espiramicina I é nitidamente superior àquela das espiramicina II e III (Iiu et al., 1999). Seria, portanto, desejável poder dispor de cepas que produzem apenas espiramici-
30 nas I.

Em virtude do interesse comercial pelos antibióticos macrolídeos, a obtenção de novos derivados, notadamente de análogos de espira-

- micinas que têm propriedades vantajosas é intensamente buscada. Seria desejável poder gerar em quantidade suficiente os intermediários de biossintese da via de biossintese das espiramicinas ou dos derivados das espiramicinas, notadamente com o objetivo de fabricar antibióticos híbridos derivados das espiramicinas.

Descrição geral da invenção

- A presente invenção resulta da clonagem de genes cujo produto intervém na biossintese das espiramicinas. A invenção refere-se, inicialmente, a novos genes da via de biossintese das espiramicinas e de novos polipeptídeos implicados nessa biossintese.

Os genes da via de biossintese e as seqüências codificantes associadas foram clonados e a seqüência de DNA de cada um foi determinada. As seqüências codificantes clonadas serão designadas pela seqüência

*orf1*c, orf2*c, orf3*c, orf4*c, orf5*, orf6*, orf7*c, orf8*, orf9*, orf10*, orf1, orf2, orf3, orf4, orf5, orf6, orf7, orf8, orf9c, orf10, orf11c, orf12, orf13c, orf14, orf15c, orf16, orf17, orf18, orf19, orf20, orf21c, orf22c, orf23c, orf24c, orf25c, orf26, orf27, orf28c, orf29, orf30c, orf31, orf32c, orf33 et orf34c.*

- A função das proteínas codificadas por essas seqüências na via de biossintese das espiramicinas é desenvolvida na discussão a seguir que é ilustrada pelas Figuras 5, 6 e 8.

- 1) Um primeiro objeto da invenção refere-se a um polinucleotídeos codificando um polipeptídeo implicado na biossintese das espiramicinas caracterizado pelo fato de a seqüência desse polinucleotídeo ser:

(a) um das seqüências

SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149,

(b) uma das seqüências constituídas pelas variantes das seqüências (a);

- (c) uma das seqüências derivadas das seqüências (a) e (b) em razão da degenerescência do código genético.

2) A presente invenção tem também por objeto um polinucleotídeo que se hibrida em condições de hibridização de forte estrigência em pelo menos um dos polinucleotídeos, segundo o parágrafo 1) acima.

3) A invenção refere-se também a um polinucleotídeo que apresenta pelo menos 70%, 80%, 85%, 90%, 95% ou 98% de identidade em nucleotídeos compreendendo pelo menos

10, 12, 15, 18, 20 a 25, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850 ou 1900 nucleotídeos consecutivos de um polinucleotídeo, segundo o parágrafo 1) acima.

4) A invenção refere-se também a um polinucleotídeo segundo o parágrafo 2) ou 3) acima, caracterizado pelo fato de ser isolado a partir de uma bactéria do gênero *streptomyces*.

5) A invenção refere-se também a um polinucleotídeo, segundo o parágrafo 2), 3) ou 4) acima, caracterizado pelo fato de codificar uma proteína implicada na biossíntese de um macrolídeo.

6) A invenção refere-se também a um polinucleotídeo, segundo o parágrafo 2), 3) ou 4) acima, caracterizado pelo fato de codificar uma proteína que tem uma atividade similar à proteína codificada pelo polinucleotídeo como qual se hibrida ou apresenta a identidade.

7) A invenção refere-se também a um polipeptídeo resultante da expressão de um polinucleotídeo, segundo o parágrafo 1), 2), 3), 4), 5) ou 6) acima.

8) um outro aspecto da invenção refere-se a um polipeptídeo implicado na biossíntese das espiamicinas, caracterizada pelo fato de a sequência desse polipeptídeo ser:

a) uma das seqüências

SEQ ID N° 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 27, 29, 31, 32, 33, 35, 37, 38, 39, 41, 42, 44, 46, 48, 50, 51, 52, 54, 55, 56, 57, 58, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 108, 110, 112, 114, 116, 117, 119, 121,

142, 144, 146, 148 e 150,

(b) uma das seqüências, tais como definidas em (a) exceto que ao longo dessa seqüência um ou vários aminoácidos foram substituídos, inseridos ou deletados, sem afetar as propriedades funcionais;

(c) uma das seqüências constituídas pelas variantes das seqüências (a) e (b).

9) Um outro objeto da invenção refere-se a um polipeptídeo que apresenta pelo menos 70%, 80%, 85%, 90%, 95% ou 98% de identidade em aminoácidos com um polipeptídeo, compreendendo pelo menos

10, 15, 20, 30 a

40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 480, 500, 520, 540, 560, 580, 600, 620 ou 640

aminoácidos consecutivos de um polipeptídeo segundo o parágrafo 8) acima.

10) Um outro aspecto da invenção refere-se a um polipeptídeo, segundo o parágrafo 9) acima, caracterizado pelo fato de ser isolado a partir de uma bactéria do gênero *streptomyces*.

11) um outro aspecto da invenção refere-se a um polipeptídeo, segundo o parágrafo 9) ou 10) acima, caracterizado pelo fato de codificar uma proteína implicada na biossíntese de um macrolídeo.

12) Um outro aspecto da invenção refere-se a um polipeptídeo, segundo o parágrafo 10) ou 11) acima caracterizado pelo fato de ter uma atividade similar àquela do polipeptídeo como qual ele divide a identidade.

13) um outro aspecto da invenção refere-se a um DNA recombinante caracterizado pelo fato de compreender pelo menos um polinucleotídeo, segundo um dos parágrafos 1), 2), 3), 4), 5) ou 6) acima.

14) um outro aspecto da invenção refere-se a um DNA recombinante segundo o parágrafo 13) acima, caracterizado pelo fato de esse DNA recombinante estar compreendido em um vetor.

15) Um outro aspecto da invenção refere-se a um DNA recombinante, segundo o parágrafo 14) acima, caracterizado pelo fato de esse vetor ser escolhido dentre os bacteriófagos, os plasmídeos, os fagemidas, os vetores integrativos, os fosmidas, os cosmidas, os vetores navettes, os BAC ou os PAC.

16) Um outro aspecto da invenção refere-se a um DNA recombinante, segundo o parágrafo 15) acima, caracterizado pelo fato de ser escolhido dentre pOS49,1, pOS49,11, pOSC49,12, pOS49,14, pOS49,16, pOS49,28, pOS44,1, pOS44,2, pOS44,4, pSPM5, pSPM7, pOS49,67, pOS49,88, pOS49,106, pOS49,120, pOS49,107, pOS49,32, pOS49,43, pOS49,44, pOS49,50, pOS49,99, pSPM17, pSPM21, pSPM502, pSPM504, pSPM507, pSPM508, pSPM509, pSPM1, pBXL1111, pBXL1112, pBXL1113, pSPM520, pSPM521, pSPM522, pSPM523, pSPM524, pSPM525, pSPM527, pSPM528, pSPM34, pSPM35, pSPM36, pSPM37, pSPM38, pSPM39, pSPM40, pSPM41, pSPM42, pSPM43, pSPM44, pSPM45, pSPM47, pSPM48, pSPM50, pSPM51, pSPM52, pSPM53, pSPM55, pSPM56, pSPM58, pSPM72, pSPM73, pSPM515, pSPM519, pSPM74, pSPM75, pSPM79, pSPM83, pSPM107, pSPM543 e pSPM106.

17) Um outro aspecto da invenção refere-se a um vetor de expressão, caracterizado pelo fato de compreender pelo menos uma sequência de ácido nucléico, codificando um polipeptídeo, segundo o parágrafo 7), 8), 9), 10, 11) ou 12) acima.

18) A invenção é também relativa a um sistema de expressão que compreende um vetor de expressão apropriado e uma célula hospedeira, permitindo a expressão de um ou vários polipeptídeos, segundo parágrafo 7), 8), 9), 10), 11) ou 12) acima.

19) A invenção é também relativa a um sistema de expressão, segundo o parágrafo 18 acima, caracterizado pelo fato de ser escolhido dentre os sistemas de expressão procariotes ou os sistemas de expressão eucariotes.

20) A invenção é também relativa a um sistema de expressão, segundo o parágrafo 19) acima caracterizado pelo fato de ser escolhido dentre os sistemas de expressão na bactéria *E. coli*, os sistemas de expressão baculovírus, permitindo uma expressão em células de levedos, os sistemas de expressão, permitindo uma expressão em células de mamíferos.

21) A invenção é também relativa a uma células hospedeira na qual foi introduzido pelo menos um polinucleotídeo e/ou pelo menos um DNA recombinante e/ou pelo menos um vetor de expressão, segundo um dos pa-

rágrafos 1), 2), 3), 4), 5), 6), 13), 14), 15), 16 ou 17) acima.

22) A invenção é também relativa a um processo de produção de um polipeptídeo, segundo o parágrafo 7), 8), 9), 10), 11) ou 12) acima caracterizado pelo fato de esse processo compreender as seguintes etapas:

- 5 a) inserir pelo menos um ácido nucléico codificando esse polipeptídeo em um vetor apropriado;
- b) cultivar, em um meio de cultura apropriado, uma célula hospedeira previamente transformada ou transfectada como vetor da etapa a);
- c) recuperar o meio de cultura acondicionado ou um extrato celular;
- 10 d) separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa c), esse polipeptídeo;
- e) se for o caso, caracterizar o polipeptídeo que recombina o produto.

15 23) Um outro aspecto da invenção refere-se a um microorganismo bloqueado em uma etapa da via de biossíntese de pelo menos um macrolídeo.

24) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 23) acima, caracterizado pelo fato de ser obtido, inativando a função de pelo menos uma proteína implicada na biossíntese desse ou desses macrolídeos em um microorganismo produtor desse ou desses macrolídeos.

25) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 24) acima, caracterizado pelo fato de a inativação dessa(s) proteína(s) ser feita por mutagênese no(s) gene(s) codificando essa(s) proteína(s) ou por expressão de um ou de vários RNA anti-sentido complementares do RNA ou RNA mensageiros, codificando essa(s) proteína(s).

26) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 25) acima, caracterizado pelo fato de a inativação dessa(s) proteína(s) ser feita por mutagênese por irradiação, por ação de agente químico mutagene, por mutagênese dirigida ou por substituição de gene.

27) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 25) ou 26) acima, caracterizado pelo fato de a(s) mutagênese(s) ser(em) feita(s) *in vitro* ou *in situ*, por supressão, substituição, deleção e/ou adição de uma ou várias bases no(s) gene(s) considerado(s) ou por inativação gênica.

28) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 23), 24), 25), 26) ou 27) acima caracterizado pelo fato de esse microorganismo ser uma bactéria do gênero *streptomyces*.

29) Um outro aspecto da invenção refere-se ao microorganismo, segundo o parágrafo 23), 24), 25), 26), 27) ou 28) acima caracterizado pelo fato de o macrolídeo ser a espiramicina.

30) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 23), 24), 25), 26), 27), 28) ou 29) acima, caracterizado pelo fato de esse microorganismo ser uma cepa de *S. ambofaciens*.

31) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 23), 24), 25), 26), 27), 28), 29) ou 30) acima, caracterizado pelo fato de a mutagênese ser feita em pelo menos um gene, compreendendo uma seqüência, de acordo com um dos parágrafos 1), 2), 3), 4), 5) ou 6) acima.

32) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 25), 26), 27), 28), 29) ou 30) ou 31) acima, caracterizado pelo fato de a mutagênese ser feita em pelo menos um ou vários genes, compreendendo uma das seqüências correspondentes a uma ou várias das seqüências SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149.

33) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 25), 26), 27), 28), 29), 30, 31 ou 32) acima, caracterizado pelo fato de a mutagênese consistir na inativação gênica de um gene, compreendendo uma seqüência correspondente à seqüência SEQ ID N° 13.

34) Um outro aspecto da invenção refere-se a uma cepa de *streptomyces ambofaciens*, caracterizada pelo fato de se tratar de uma cepa

escolhida dentre uma das cepas depositadas junto à Coleção Nacional de Culturas de Microorganismos (CNCM), em 10 de julho de 2002 sob o número de registro I-2909, I-2911, I-2912, I-2913, I-2914, I-2915, I-2916 ou I-2917.

35) Um outro aspecto da invenção refere-se a um processo de
5 preparação de um intermediário de biossíntese de macrolídeo, caracterizado pelo fato de, caracterizado pelo fato de compreender as seguintes etapas:

a) cultivar, em um meio de cultura apropriado, um microorganismo, segundo um dos parágrafos 23), 24), 25), 26), 27), 28), 29), 30), 31), 32), 33) ou 34) acima;

10 b) recuperar o meio de cultura acondicionado ou um extrato celular;

c) separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa b), esse intermediário de biossíntese.

15 36) Um outro aspecto da invenção refere-se a um processo de preparação de uma molécula derivada de um macrolídeo, caracterizado pelo fato de se preparar um intermediário de biossíntese segundo o processo do parágrafo 35) acima e de se modificar o intermediário assim produzido.

20 37) Um outro aspecto da invenção refere-se a um processo de preparação, segundo o parágrafo 36) acima, caracterizado pelo fato de se modificar esse intermediário por via química, bioquímica, enzimática e/ou microbiológica.

25 38) Um outro aspecto da invenção refere-se a um processo de preparação, segundo o parágrafo 36) ou 37) acima, caracterizado pelo fato de se introduzir nesse microorganismo um ou vários genes codificando proteínas capazes de modificar o intermediário servindo-se dele como substrato.

30 39) Um outro aspecto da invenção refere-se a um processo de preparação, segundo o parágrafo 36), 37), ou 38) acima, caracterizado pelo fato de o macrolídeo ser a espiramicina.

40) Um outro aspecto da invenção refere-se a um processo de preparação, segundo o parágrafo 36), 37), 38), ou 39) acima, caracterizado

pelo fato de o microorganismo utilizado ser uma cepa de *S. ambofaciens*.

41) Um outro aspecto da invenção refere-se a um microorganismo, que produz a espiramicina I, mas que não produz espiramicina II e III.

42) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41) acima, caracterizado pelo fato de compreender o conjunto dos genes necessários à biossíntese da espiramicina I, mas que o gene que compreende a sequência SEQ ID N° 13 ou uma de suas variantes ou uma das sequências derivadas destas, em virtude da degenerescência do código genético e codificando um polipeptídeo de sequência SEQ ID N° 14 ou uma de suas variantes não é expressa ou foi tornada inativa.

43) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 42) acima, caracterizado pelo fato de essa inativação ser realizada por mutagênese no gene que codifica essa proteína ou pela expressão de um RNA anti-sentido complementar do RNA mensageiro codificando essa proteína.

44) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 43) acima, caracterizado pelo fato de essa mutagênese ser feita no promotor desse gene, na sequência codificante ou em uma sequência não codificante, de maneira a tornar inativa a proteína codificada ou a impedir sua expressão ou sua tradução.

45) Um outro aspecto da invenção refere-se a um microorganismo, segundo a reivindicação 43) ou 44) acima, caracterizado pelo fato de a mutagênese ser feita por irradiação, por ação de agente químico mutagene, por mutagênese dirigida ou por substituição de gene.

46) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 43), 44), ou 45), acima, caracterizado pelo fato de a mutagênese ser feita *in vitro* ou *in situ*, por supressão, substituição, deleção e/ou adição de uma ou várias bases no gene considerado ou por inativação gênica.

47) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41) ou 42) acima, caracterizado pelo fato de esse microorganismo ser obtido expressando-se os genes da via de biossíntese

da espiramicinas, sem que estes compreendam o gene que compreende a sequência correspondente à SEQ ID N° 13 ou uma de suas variantes ou uma das seqüências derivadas desta, em virtude da degeneração do código genético e codificando um polipeptídeo da sequência SEQ ID N° 14 ou uma
5 de suas variantes.

48) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41), 42), 43), 44), 45), 46), ou 47) acima, caracterizado pelo fato de o microorganismo ser uma bactéria do gênero *streptomyces*.

10 49) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41), 42), 43), 44), 45), 46), 47) ou 48) acima, caracterizado pelo fato de o microorganismo ser obtido a partir de uma cepa de partida, produzindo as espiras I, II e III.

50) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41), 42), 43), 44), 45), 46), 47), 48) ou 49) acima, caracterizado pelo fato de ser obtido por mutagênese em um gene que compreende a sequência correspondente à SEQ ID N° 13 ou uma de suas variantes ou uma das seqüências derivadas destas em razão da degenerescência do código genético e codificando um polipeptídeo de sequência SEQ ID
15 N° 14 ou uma de suas variantes que têm a mesma função.

51) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41), 42), 43), 44), 45), 46), 47), 48), 49) ou 30) acima, caracterizado pelo fato de esse microorganismo ser obtido a partir de uma cepa de *S. ambofaciens* que produzindo as espiramicinas I, II e III, na
25 qual é feita uma inativação gênica do gene que compreende a sequência correspondente à SEQ ID N° 13 ou uma das seqüências derivadas desta em razão da degenerescência do código genético.

52) Um outro aspecto da invenção refere-se a uma cepa de *S. ambofaciens*, caracterizada pelo fato de se tratar da cepa depositada junto à
30 Coleção Nacional de Culturas de Microorganismos (CNCM) em 10 de julho de 2002, sob o número de registro I-2910.

53) Um outro aspecto da invenção refere-se a um processo de

produção de espiramicina I, caracterizado pelo fato de compreender as seguintes etapas:

(a) cultivar, em um meio de cultura apropriado, um microorganismo, segundo um dos parágrafos 41), 42), 43), 44), 45), 46), 47), 48), 49),
5 50), 51) ou 52) acima;

(b) recuperar o meio de cultura acondicionado ou um extrato celular;

(c) separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa b) a espiramicina I.

10 54) Um outro aspecto da invenção refere-se à utilização de um polinucleotídeo, segundo um dos parágrafo 1), 2), 3), 4), 5), ou 6) acima, para melhorar a produção em macrolídeos de um microorganismo.

55) Um outro aspecto da invenção refere-se a um microorganismo mutante produtor de macrolídeo, caracterizado pelo fato de possuir uma
15 modificação genética em pelo menos um gene que compreende uma sequência, segundo um dos parágrafos 1), 2), 3), 4), 5) ou 6) acima e/ou que superexpressa pelo menos um gene que compreende uma sequência, segundo um de seus parágrafos 1), 2), 3), 4), 5), ou 6) acima.

56) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55) acima, caracterizado pelo fato de a
20 modificação genética consistir em uma supressão, uma substituição, uma deleção e/ou uma adição de uma ou várias bases no(s) gene(s) considerado(s) com a finalidade de expressar uma ou proteínas que têm uma atividade superior ou de expressar um nível mais elevado dessa(s) proteína(s).

25 57) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55) acima, caracterizado pelo fato de a superexpressão do gene considerado ser obtida, aumentando-se o número de cópia desse gene e/ou colocando-se um promotor mais ativo do que o promotor selvagem.

30 58) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 55), ou 57) acima, caracterizado pelo fato de a superexpressão do gene considerado ser obtida, transformando-se um micro-

organismo produtor de macrolídeo por uma construção de DNA recombinante segundo o parágrafo 13, 14 ou 17 acima, permitindo a superexpressão desse gene.

5 59) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55), 56), 57), ou 58) acima, caracterizado pelo fato de a modificação genética ser feita em um ou vários genes, compreendendo uma de suas seqüências correspondentes a uma ou várias das seqüências SEQ ID Nº 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 10 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149 ou uma de suas variantes ou uma das seqüências derivadas destas em razão da degenerescência do código genético.

60) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55), 56), 57), 58) ou 59) acima, caracterizado pelo fato de o microorganismo superexpressar um ou vários genes 15 compreendendo uma das seqüências correspondentes a uma ou várias das seqüência

SEQ ID Nº 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147

e 149 ou uma de suas variantes ou uma das seqüências derivadas destas em razão da degenerescência do código genético.

20 61) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55), 56), 57), 58), 59) ou 60) acima, caracterizado pelo fato de esse microorganismo ser uma bactéria do gênero *streptomyces*.

25 62) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55), 56), 57), 58), 59), 60) ou 61) acima, caracterizado pelo fato de o macrolídeo ser a espiramicina.

63) Um outro aspecto da invenção refere-se a um microorganismo mutante, segundo o parágrafo 55), 56), 57), 58), 59) 60, 61) ou 62) acima, caracterizado pelo fato de esse microorganismo ser uma cepa de S.

ambofaciens.

64) Um outro aspecto da invenção refere-se a um processo de produção de macrolídeos, caracterizado pelo fato de compreender as seguintes etapas:

5 a) cultivar, em um meio de cultura apropriado, um microorganismo, segundo um dos parágrafos 55), 56), 57), 58), 59), 60), 61), 62), 63) ou 64) acima;

b) recuperar o meio de cultura acondicionado ou um extrato celular;

10 c) separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa b), esse(s) macrolídeo(s) produzido(s).

65) Um outro aspecto da invenção refere-se a utilização de uma seqüência e/ou de um DNA recombinante e/ou de um vetor, segundo um dos parágrafos 1), 2), 3), 4), 5), 6), 7), 8), 9), 10) 11), 12), 13), 14), 15), 16) ou 17) acima, para a preparação de antibióticos híbridos.

66) Um outro aspecto da invenção refere-se à utilização de pelo menos um polinucleotídeo e/ou pelo menos um DNA recombinante e/ou pelo menos um vetor de expressão e/ou pelo menos um polipeptídeo e/ou pelo menos uma célula hospedeira, segundo um dos parágrafos 1), 2), 3), 4), 5), 6), 7), 8), 9), 10), 11), 12), 13), 14), 15), 16), 17) ou 21) acima, para realizar uma ou várias bioconversões.

67) Um outro aspecto da invenção refere-se a um polinucleotídeo, caracterizado pelo fato de se tratar de um polinucleotídeo complementar de um dos polinucleotídeos, segundo o parágrafo 1), 2), 3), 4), 5) ou 6) acima.

68) Um outro aspecto da invenção refere-se a um microorganismo produtor de pelo menos uma espiamicinas, caracterizado pelo fato de superexpressar:

30 - um gene capaz de ser obtido por amplificação em cadeia por polimerase (PCR), utilizando o par de atrações de seqüência apresentada a seguir:

AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e 5' GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces ambofaciens*;

- ou um gene derivado deste em razão da degenerescência do código genético.

5 69) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 68 ou 90 acima, caracterizado pelo fato de uma bactéria do gênero *streptomyces*.

70) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 68, 69, ou 90, caracterizado pelo fato de tratar-se
10 de uma bactéria da espécie *Streptomyces ambofaciens*.

71) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 68, 69, 70 ou 90, caracterizado pelo fato de a superexpressão desse gene ser obtida por transferência desse microorganismo por um vetor de expressão.

15 72) Um outro aspecto da invenção refere-se a uma cepa de *Streptomyces ambofaciens*, caracterizada pelo fato de se tratar da cepa OSC2/pSPM75(1) ou da cepa OSC2/pSPM75(2) depositada junto à Coleção Nacional de Culturas de Microorganismos (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, 6 de outubro de 2003, sob
20 o número de registro I-3101.

73) Um outro aspecto da invenção refere-se a um DNA recombinante, caracterizado pelo fato de compreender:

- um polinucleotídeo capaz de ser obtido por amplificação em cadeia por polimerase, utilizando o par de atrações com a seguinte seqüência:
25 5' AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e 5' GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces ambofaciens*;

- ou um fragmento de pelo menos

10, 12, 15, 18, 20 à 25, 30, 40, 50, 60, 70, 80,
90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850,
900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1460, 1470,
1480, 1490 ou 1500

nucleotídeos consecutivos desse polinucleotídeo.

74) Um outro aspecto da invenção refere-se a um DNA recombinante, segundo o parágrafo 73 ou 91, caracterizado pelo fato de se tratar de um vetor.

5 75) Um outro aspecto da invenção refere-se a um DNA recombinante, segundo o parágrafo 73, 74 ou 91, caracterizado pelo fato de se tratar de um vetor de expressão.

76. Um outro aspecto da invenção refere-se a uma célula hospedeira, na qual foi introduzido pelo menos um DNA recombinante, segundo
10 um dos parágrafos 73, 74, 75 ou 91.

77) Um outro aspecto da invenção refere-se a um processo de produção de um polipeptídeo, caracterizado pelo fato de esse processo compreender as seguintes etapas:

a) transformar uma célula hospedeira por pelo menos um vetor
15 de expressão, segundo o parágrafo 75;

b) cultivar, em um meio de cultura apropriado, essa célula hospedeira;

c) recuperar o meio de cultura acondicionado ou um extrato celular;

20 d) separar e purificar, a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa c), desse polipeptídeo;

e) se for o caso, caracterizar o polipeptídeo recombinante produzido.

78) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 51, caracterizado pelo fato de a inativação gênica
25 ser feita por deleção em fase do gene ou de uma parte do gene que compreende a sequência correspondente à SEQ ID N° 13 ou uma das sequências derivadas desta, em razão da degenerescência do código genético.

79) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51 ou 78, caracterizado pelo fato de superexpressar além disso:

- um gene capaz de ser obtido por amplificação em cadeia por polimerase, utilizando o par de atrações da seguinte seqüência:

5' AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e

5' GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces ambofaciens*,

- ou um gene derivado destes, em razão da degenerescência do código genético.

80) Um outro aspecto da invenção refere-se a um vetor de expressão, caracterizado pelo fato de o polinucleotídeo de seqüência SEQ ID N° 47 ou o polinucleotídeo derivado deste, em razão da degenerescência do código genético ser colocado sob o controle de um promotor, permitindo a expressão da proteína codificada por esse polinucleotídeo no *streptomyces ambofaciens*.

81) Um outro aspecto da invenção refere-se a um vetor de expressão, segundo o parágrafo 80, caracterizado pelo fato de se tratar do plasmídeo pSPM524 ou pSPM525.

82) Um outro aspecto da invenção refere-se a uma cepa de *streptomyces ambofaciens* transformada por um vetor, segundo um parágrafo 80 ou 81.

83) Um outro aspecto da invenção refere-se a um microorganismo, segundo um dos parágrafos 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 78, 79 ou 92, caracterizado pelo fato de superexpressar, além disso, o gene de seqüência codificante SEQ ID N° 47 ou de uma seqüência codificante derivada desta em razão da degenerescência do código genético.

84) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 83, caracterizado pelo fato de se tratar da cepa SPM502 pSPM525 depositada junto à Coleção Nacional de Culturas de Mi-

croorganismos (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 26 de fevereiro de 2003, sob o número de registro I-2977.

85) Um outro aspecto da invenção refere-se a um processo de produção de espiramicina(s), caracterizado pelo fato de compreender as seguintes etapas:

a) cultivar, em um meio de cultura apropriado, um microorganismo, segundo um dos parágrafos 68, 69, 70, 71, 72, 78, 79, 82, 83, 84, 90 ou 92;

b) recuperar o meio de cultura acondicionado ou um extrato celular;

c) separar e purificar a partir do meio de cultura ou ainda a partir do extrato celular obtido na etapa b), as espiramicinas.

86) Um outro aspecto da invenção refere-se a um polipeptídeo, caracterizado pelo fato de sua sequência compreender a sequência SEQ ID N° 112 ou a sequência SEQ ID N° 142.

87) Um outro aspecto da invenção refere-se a um polipeptídeo, caracterizado pelo fato de sua sequência corresponder à sequência traduzida da sequência codificante:

- de um gene capaz de ser obtido por amplificação em cadeia por polimerase (PCR), utilizando o par de atrações da seguinte sequência:

5'AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e
5'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces ambofaciens*;

- ou de um gene derivado deste, em razão da degenerescência do código genético.

88) Um outro aspecto da invenção refere-se a um vetor de expressão de um polipeptídeo, segundo o parágrafo 86, 87 ou 93 no *streptomyces ambofaciens*.

89) Um outro aspecto da invenção refere-se a um vetor de expressão, segundo o parágrafo 88, caracterizado pelo fato de que é do plas-

mídeo pSPM75.

90) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 68, caracterizado pelo fato de o gene de sequência codificante SEQ ID N° 141 ou um gene derivado deste em razão da degenerescência do código genético.

91) Um outro aspecto da invenção refere-se a um DNA recombinado, segundo o parágrafo 73, caracterizado pelo fato de o polinucleotídeo capaz de ser obtido por amplificação em cadeia por polimerase ser um polinucleotídeo de sequência SEQ ID N° 141.

92) Um outro aspecto da invenção refere-se a um microorganismo, segundo o parágrafo 79, caracterizado pelo fato de o gene capaz de ser obtido por amplificação em cadeia por polimerase ser o gene de sequência codificante SEQ ID N° 141 ou um gene derivado deste em razão da degenerescência do código genético.

93) Um outro aspecto da invenção refere-se a um polipeptídeo, caracterizado pelo fato de sua sequência ser SEQ ID N° 142.

Definições gerais

O termo "isolado" no sentido da presente invenção designa um material biológico (ácido nucléico ou proteína) que foi subtraído em seu meio ambiente original (o meio ambiente no qual é localizado naturalmente).

Por exemplo, um polinucleotídeo presente no estado natural em uma planta ou um animal não é isolado. O mesmo polinucleotídeo separado dos ácidos nucléicos adjacentes no meio dos quais é naturalmente inserido no genoma da planta ou o animal é considerado como isolado.

Esse polinucleotídeo pode ser incluído em um vetor e/ou esse polinucleotídeo pode ser incluído em uma composição e permanecer, todavia, no estado isolado devido ao fato de o vetor ou a composição não constituir seu meio ambiente natural.

O termo *purificado* não necessita de que o material esteja presente sob uma forma de pureza absoluta, exclusiva da presença de outros compostos. Trata-se antes de tudo de uma definição relativa.

Um polinucleotídeo está no estado "purificado", após purificação

do material de partida ou do material natural de pelo menos uma ordem de grandeza, de preferência 2 ou 3, e, preferencialmente 4 ou 5 ordens de grandeza.

5 Com a finalidade da presente invenção o termo ORF (Open Reading Frame, isto é, quadro aberto de leitura) foi empregado para designar em particular a seqüência codificante de um gene.

Com a finalidade da presente descrição, a expressão "seqüência nucleotídica" pode ser empregada para designar indiferentemente um polinucleotídeo ou um ácido nucléico. A expressão "seqüência nucleotídica" engloba o próprio material genético e não está, portanto, restrita à informação
10 referente à sua seqüência.

Os termos "ácido nucléico", polinucleotídeo, oligonucleotídeo ou ainda seqüência nucleotídica englobam as seqüências de RNA, de DNA, de ADNc ou ainda seqüências híbridas RNA/DNA além disso, de um nucleotídeo, indiferentemente sob a forma simples cadeia ou sob a forma de duplex.
15

O termo nucleotídeo designa ao mesmo tempo os nucleotídeos naturais (A, T, G, C), assim como nucleotídeos modificados que compreendem pelo menos uma modificação tal como (1), um análogo de uma purina, (2), um análogo de uma pirimidina, ou (3) um açúcar análogo, exemplos
20 desses nucleotídeos modificados sendo descritos, por exemplo, no pedido PCT N° EO 95/04 064.

Com a finalidade da presente invenção, um primeiro polinucleotídeo é considerado como sendo "complementar" de um segundo polinucleotídeo, quando cada base do primeiro polinucleotídeo é emparelhado à base
25 complementar do segundo polinucleotídeo, cuja orientação é inversa. As bases complementares são A e T (ou A e U), ou C e G.

O termo "genes da via de biossíntese das espiamicinas" compreende igualmente os genes reguladores, e os genes conferem a resistência aos microorganismos produtores.

30 Será entendido por "fragmento" de um ácido nucléico de referência, de acordo com a invenção, uma seqüência nucleotídica de comprimento reduzido em relação ao ácido nucléico de referência e compreenden-

do, sobre a parte comum, uma seqüência em nucleotídeos idêntica ao ácido nucléico de referência.

Esse "fragmento" de ácido nucléico, de acordo com a invenção, pode estar, se for o caso, compreendido em um polinucleotídeo maior, do qual ele é constitutivo.

Desses fragmentos, compreendem ou alternativamente consistem em polinucleotídeos de comprimento que vai de 8, 10, 12, 15, 18, 20 a 25, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850 ou 1900 nucleotídeos consecutivos de um ácido nucléico, de acordo com a invenção.

Por "fragmento" de um polipeptídeo, de acordo com a invenção, entender-se-á um polipeptídeo cuja seqüência de aminoácidos é mais curta do que a do polipeptídeo de referência, e que compreende sobre toda a parte comum com esses polipeptídeos de referência, uma seqüência em aminoácidos idêntica.

Esses fragmentos podem, se for o caso, estar compreendidos no meio de um polipeptídeo maior, do qual eles fazem parte.

Esses fragmentos de um polipeptídeo, de acordo com a invenção, podem ter um comprimento de 10, 15, 20, 30 a 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 480, 500, 520, 540, 560, 580, 600, 620 ou 640 aminoácidos.

Por "condições de hibridização de forte estringência" no sentido da presente invenção, serão entendidas condições de hibridização que desfavorecem a hibridização de filamentos de ácidos nucléicos não homólogos. Condições de hibridizações de forte estringência podem ser, por exemplo, descritas como condições de hibridização no tampão descrito por Church & Gilbert (Church & Gilbert, 1984) a uma temperatura compreendida entre 55°C e 65°C, de maneira preferida a temperatura de hibridização é de 55°C, de maneira ainda mais preferida a temperatura de hibridização é de 60°C e

de maneira inteiramente preferida a temperatura de hibridização é de 65°C, seguida de uma ou várias lavagens feita em tampão 2X SSC (o tampão 1X SSC corresponde a uma solução aquosa 0,15M NaCl, 15 mM de citrato de sódio) a uma temperatura compreendida entre 55°C e 65°C, de maneira
5 preferida essa temperatura é de 55°C, de maneira ainda mais preferida essa temperatura é de 60°C e de maneira inteiramente preferida essa temperatura é de 65°C, seguida de uma ou várias lavagens em tampão 0,5X SSC a uma temperatura compreendida entre 55°C e 65°C, de maneira preferida essa temperatura é de 55°C, de maneira ainda mais preferida essa temperatura é de 60°C e de maneira inteiramente preferida essa temperatura é de
10 65°C.

É evidente que as condições de hibridização acima descritas podem ser adaptadas em função do comprimento do ácido nucléico, cuja hibridização é buscada ou do tipo marcação escolhida, segundo técnicas
15 conhecidas do técnico. As condições convenientes de hibridização podem, por exemplo, ser adaptadas, segundo a obra de F. Ausubel et al. (2002).

Por "variante" de um ácido nucléico, de acordo com a invenção, entender-se-á um ácido nucléico que difere de uma ou várias bases em relação ao polinucleotídeo de referência. Um ácido nucléico variante pode ser de
20 origem natural, tal como um variante alélico encontrado naturalmente, ou pode ser também um variante não natural obtido por exemplo por técnicas de mutagênese.

Em geral, as diferenças entre o ácido nucléico de referência e o ácido nucléico variante são reduzidas de tal modo que as seqüências nucleotídicas do ácido nucléico de referência e do ácido nucléico variante são
25 muito próximas e, em numerosas regiões, idênticas. As modificações de nucleotídeos presentes em um ácido nucléico variante podem ser silenciosas, o que significa que elas não alteram as seqüências de aminoácidos codificadas por esse ácido nucléico variante.

30 Todavia, as mudanças de nucleotídeos em um ácido nucléico variante podem também resultar de substituições, adições, deleções no polipeptídeo codificado pelo ácido nucléico variante em relação aos peptídeos

codificados pelo ácido nucléico de referência. Além disso, modificações de nucleotídeos nas regiões codificantes podem produzir substituições, conservadoras ou não conservadoras na seqüência de aminoácidos.

De preferência, os ácidos nucléicos variantes, de acordo com a invenção, codificam polipeptídeos que conservam sensivelmente a mesma função ou atividade biológica que o polipeptídeo do ácido nucléico de referência ou ainda a capacidade de serem reconhecidos por anticorpos dirigidos contra os polipeptídeos codificados pelo ácido nucléico inicial.

Certos ácidos nucléicos variantes codificarão assim formas mutadas dos polipeptídeos cujo estudo sistemático permitirá deduzir relações estrutura atividade das proteínas em questão.

Por variante de um polipeptídeo, de acordo com a invenção, entender-se-á principalmente um polipeptídeo cuja seqüência de aminoácidos contém uma ou várias substituições, adições ou deleções de pelo menos um resíduo de aminoácido, em relação à seqüência de aminoácidos do polipeptídeo de referência, naturalmente que as substituições de aminoácidos podem ser indiferentemente conservadoras ou não conservadoras.

De preferência, os polipeptídeos variantes, de acordo com a invenção, conservam sensivelmente a mesma função ou atividade biológica que o polipeptídeo de referência ou ainda a capacidade de serem reconhecidos por anticorpos dirigidos contra os polipeptídeos iniciais.

Por polipeptídeo tendo uma atividade similar a um polipeptídeo de referência no sentido da invenção, entende-se um polipeptídeo que tem uma atividade biológica próxima, mas não necessariamente idêntica, daquela do polipeptídeo de referência, tal como medido em um teste biológico conveniente à medida da atividade biológica do polipeptídeo de referência.

Por antibiótico híbrido no sentido da invenção entende-se um composto, gerado pela construção de uma via de biossíntese artificial utilizando a tecnologia do DNA recombinante.

30 Descrição Detalhada da Invenção

A presente invenção tem mais particularmente por objeto novos genes da via de biossíntese das espiamicinas e novos polipeptídeos impli-

cados nessa biossíntese, tais como apresentados na descrição detalhada abaixo.

Os genes da via de biossíntese foram clonados e a sequência de DNA desses genes foi determinada. As sequências obtidas foram analisadas graças ao programa FramePlot (Ishikawa J & Hotta K. 1999). Identificaram-se, dentre as fases abertas de leitura, as fases abertas de leitura que apresentam um uso dos códons típicos de *streptomyces*. Essa análise mostrou que essa região comporta 44 ORFs, localizadas de ambos lados de cinco genes codificando a enzima "polyketide synthase"(PKS), e apresentando um uso dos códons típico de *streptomyces*. De ambos lados desses cinco genes codificando a PKS, respectivamente 10 e 34 ORFs foram identificadas a jusante e a montante (a jusante e a montante sendo definidos pela orientação dos 5 genes de PKS todos orientados no mesmo sentido) (conforme Figura 3 e 37). Assim, as 34 fases abertas de leitura desse tipo, ocupando uma região de aproximadamente 41,7 kb (cf SEQ ID N° 1 que apresenta uma primeira região de 31 kb contendo 25 ORFs e SEQ ID N° 140 que apresenta uma região de aproximadamente 12,1 kb, dos quais 1,4 kb se sobrepõem a sequência precedente (SEQ ID N° 1) e aproximadamente 10,7 kb correspondem à sequência da sequência, esta última parte de aproximadamente 10,7 kb contendo 9 ORFs suplementares (dos quais um ORF de sequência parcial), conforme também Figura 3 e 37 e abaixo) foram identificadas a montante dos 5 genes codificando a PKS e 10 ocupando uma região de aproximadamente 11,1 kb (SEQ ID N° 2 e Figura 3), foram identificadas a jusante dos genes da PKS. Os 10 genes situados a jusante dos 5 genes PKS foram assim denominados *orf1*c*, *orf2*c*, *orf3*c*, *orf4*c*, *orf5*c*, *orf6*c*, *orf7*c*, *orf8*c*, *orf9*c*, *orf10*c* (SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19 e 21). O "c" acrescentado no nome do gene significa para ORF em questão que a sequência codificante é na orientação inversa (o filamento codificando é, portanto, o filamento complementar da sequência determinada em SEQ ID N° 2 para esses genes). Utilizando a mesma nomenclatura, as 34 ORFs a montante dos genes da PKS foram chamados *orf1*, *orf2*, *orf3*, *orf4*, *orf5*, *orf6*, *orf7*, *orf8*, *orf9c*, *orf10*, *orf11c*, *orf12*, *orf13c*, *orf14*, *orf15c*, *orf16*, *orf17*, *orf18*,

orf19, *orf20*, *orf21c*, *orf22c*, *orf23c*, *orf24c*, *orf25c*, *orf26*, *orf27*, *orf28c*, *orf29*, *orf30c*, *orf31*, *orf32c* e *orf34c* (SEQ ID Nº 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149) (conforme Figuras 3 e 37).

- 5 As seqüências protéicas deduzidas dessas fases abertas de leitura foram comparadas com aquelas presentes em diferentes bases de dados graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et al., 1997), CD-Search, COGs (Cluster of Orthologous Groups) (esses três programas são acessíveis notadamente junto ao National Center for Biotechnology Information (NCBI) (Bethesda, Maryland, USA)), FASTA ((Pearson W. R & D.J. Lipman, 1988) e (Pearson W. R., 1990), BEAUTY (Worley K. C. et al., 1995)), (esses dois programas são acessíveis notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França). Essas comparações permitiram formular hipóteses sobre a função dos produtos desses genes e
- 10
- 15 identificar aqueles capazes de serem implicados na biossíntese das espiramicinas.

Genes situados a jusante dos genes codificando os PKS

- Uma representação esquemática da organização da região é apresentada na Figura 3. Assim conforme será demonstrado abaixo, sobre
- 20 os 10 genes identificados a jusante dos genes codificando os PKS 9 parecem implicados na biossíntese ou a resistência às espiramicinas. Trata-se dos seguintes 9 genes: *orf1*c*, *orf2*c*, *orf3*c*, *orf4*c*, *orf5**, *orf6**, *orf7*c*, *orf8** e *orf9**.

- Na tabela 1 a seguir são apresentadas as referências à seqüência em DNA e em aminoácidos dos 10 genes identificados a jusante dos 5 genes PKS
- 25

TABELA 1

Gene	Posição na seqüência SEQ ID N° 2	Seqüência DNA	Seqüências poli- peptídicas
<i>Orf1*c</i>	10882 a 10172	SEQ ID N 3	SEQ ID N 4
<i>Orf2*c</i>	10052 a 8781	SEQ ID N 5	SEQ ID N 6
<i>Orf3*c</i>	8741 a 7476	SEQ ID N 7	SEQ ID N 8
<i>Orf4*c</i>	7459 a 6100	SEQ ID N 9	SEQ ID N 10
<i>Orf5*</i>	5302 a 5976	SEQ ID N 11	SEQ ID N 12
<i>Orf6*</i>	4061 a 5305	SEQ ID N 13	SEQ ID N 14
<i>Orf7*c</i>	3665 a 2817	SEQ ID N 15	SEQ ID N 16
<i>Orf8*</i>	1925 a 2755	SEQ ID N 17	SEQ ID N 18
<i>Orf9*</i>	1007 a 1888	SEQ ID N 19	SEQ ID N 20
<i>Orf10*</i>	710 a 937	SEQ ID N 21	SEQ ID N 22

O "c" acrescentado ao nome do gene indica que a seqüência codificante está na orientação inversa (o filamento codificante é, portanto, o filamento complementar da seqüência determinada em SEQ ID N° 2 para esses genes).

Com a finalidade de determinar a função dos polipeptídeos identificados três tipos de experiências foram feitas: a comparação das seqüências identificadas com seqüências de funções conhecidas, experiências de inativação de genes, conduzindo à produção de cepas mutantes e análises da produção em espiramicinas e em intermediários de biossíntese das espiramicinas por essas cepas mutantes.

As seqüências protéicas deduzidas dessas fases abertas de leitura foram inicialmente comparadas com aquelas presentes em diferentes bases de dados, graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et al., 1997), CD-search, COGs (Cluster of Orthologous Groups), FASTA (Pearson W.R & D.J. Lipman, 1988) e (Pearson W.R., 1990), BEAUTY (Worley K.C. et al., 1955). Essas comparações permitiram formular hipóteses a respeito da função dos produtos desses genes e identificar aqueles capazes de serem implicados na biossíntese de espiramicina. A tabela 2 mostra as proteínas que apresentam uma elevada similitude com

os produtos 10 genes situados a jusante dos 5 genes PKS.

Tabela 2

Produto de gene	Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
<i>orf1*c</i>	Ty1MI(<i>Orf3*</i>)(<i>S. fradiae</i>)	CAA57473	287	N-Metiltransferase
<i>orf2*c</i>	DnrQ produto gene (<i>Streptomyces peucetius</i>)	AAD15266	153	Desconhecida
<i>orf3*c</i>	Ty1MII(<i>orf2*</i>) (<i>S. fradiae</i>)	CAA57472	448	Glicosil transferase
<i>orf4*c</i>	Crotonil-CoA redutase (<i>S. coelicolor</i>)	NP_630556	772	Crotonil-CoA redutase
<i>orf5*</i>	MdmC (<i>S. mycarofaciens</i>)	B42719	355	O-metiltransferase
<i>orf6*</i>	3-O-aciltransferase (<i>S. mycarofaciens</i>)	Q00718	494	Aciltransferase
<i>orf7*c</i>	MdmA (<i>S. mycarofaciens</i>)	A60725	380	Proteína implicada na resistência à midecamicina
<i>orf8*</i>	ABC-transportador (<i>S. griseus</i>)	CAC22119	191	ABC-transportador
<i>orf9*</i>	ABC-transportador (<i>S. griseus</i>)	CAC22118	269	ABC-transportador
<i>orf10*</i>	Proteína putativa pequena conservada hipotética (<i>S. coelicolor</i>)	NP_627432	109	Desconhecida

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

- 5 Experiências de inativação de genes foram feitas para confirmar esses resultados. Os métodos aplicados consistem em efetuar uma substituição de gene. O gene alvo a interromper é substituído por uma cópia desse gene interrompida por um cassette, conferindo à resistência a um antibiótico (por exemplo a apramicina, a geneticina ou a higromicina). Os cassettes utilizados são contornados de ambos os lados por códons de terminação da
- 10 tradução em todas as fases de leitura e por terminadores de transcrição ativos em *streptomyces*. A inserção do cassette no gene alvo pode ser acom-

panhada ou não de uma deleção nesse gene alvo. O tamanho das regiões que costeam o cassette pode ir de algumas centenas a alguns milhares de pares de bases. Um segundo tipo de cassettes pode ser utilizado para a inativação dos genes: cassettes ditos "cassettes excisáveis". Esses cassettes

5 apresentam a vantagem de poderem ser excisados em *streptomyces* por uma recombinação de sítio específico após terem sido introduzidos no genoma de *S. ambofaciens*. A finalidade é de inativar certos genes em cepas de *streptomyces*, sem deixar na cepa final marcadores de seleção ou de grandes seqüências de DNA que não pertencem à cepa. Após excisão sub-

10 siste unicamente uma curta seqüência de uns trinta pares de bases (denominado sítio "cicatricial") no genoma da cepa (conforme Figura 10). A utilização desse sistema consiste, em uma primeira etapa, na substituição da cópia selvagem do gene alvo (graças a dois acontecimentos de recombinação homóloga, conforme Figura 9) por uma construção na qual um cassette

15 excisável foi inserida nesse gene alvo. A inserção desse cassette é acompanhada de uma deleção no gene alvo (conforme Figura 9). Em uma segunda etapa, a excisão do cassette excisável do genoma da cepa é provocada. O cassette excisável funciona graças a um sistema de recombinação sítio específico e tem por vantagem permitir a obtenção de mutantes de

20 *streptomyces* não portando finalmente gene de resistência. Ficou-se livre também de eventuais efeitos polares sobre a expressão dos genes situados a jusante do ou dos genes inativados (conforme Figura 10). As cepas assim construídas foram testadas para sua produção em espiramicinas.

A inativação dos genes *orf1*c*, *orf2*c*, *orf3*c* e *orf4*c* não foi feita,

25 pois as experiências de comparação de seqüência permitiram determinar que esses genes tinham uma similitude relativamente importante com genes implicados na biossíntese de um antibiótico relativamente próxima. Assim o gene *orf1*c* codifica uma proteína que apresenta uma identidade de 66% (determinada graças ao programa BLAST) com a proteína codificada por

30 gene *tylMI* que codifica uma N-metiltransferase implicada na biossíntese de tilosina e catalisando a 3-N-metilação durante a produção da micaminose nas *streptomyces fradiae* (Gandecha, A.R. et al., 1997; Número de acesso

GenBank: CAAA57473; Escore BLAST: 287). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico relativamente próxima e mais particularmente na biossíntese da micaminose sugere que o gene orf1*c codifica uma N-metiltransferase responsável por uma N-metilação, quando da biossíntese da forosamina ou da micaminose (conforme Figura 5 e 6). Essa hipótese é reforçada pelo fato de a proteína codificada pelo gene orf1*c apresenta uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 3).

TABELA 3

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Score BLAST*	Função complementar
Metiltransferase (<i>S. antibioticus</i>)	CAA05643	277	metiltransferase
N,N-dimetiltransferase (<i>S. venezuelae</i>)	AAC68678	268	N,N-dimetiltransferase
N-metilase provável snogX (<i>S. nogalater</i>)	T46679	243	N-metiltransferase

* maior similitude de sequência é associada a um escore BLASDT mais elevado (Altschul et al., 1990).

O gene orf2*c codifica uma proteína que apresenta uma relativa similitude (35% de identidade) com uma proteína codificada pelo gene tyIMIII codificando uma NDP hexose 3,4 isomerase implicada na biossíntese da tilosina nas *streptomyces fradiae* (Gandecha, A.R., et al., 1997; Número de acesso GenBank: CAAS57471; Escore BMAST; 130). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo e mais particularmente na biossíntese da micaminose sugere muito que o gene orf2c codifica uma NDP hexose 3,4 isomerase responsável de isomerização, quando da biossíntese de um dos açúcares da espiramicina, eventualmente a micaminose (conforme Figura 5 e 6).

O gene orf3* codifica uma proteína que apresenta uma relativa forte similitude (59% de identidade) com uma proteína codificada pelo gene

- tylMII codificando uma glicosiltransferase implicada na biossíntese da tilosina nas *streptomyces fradiae* (Gandecha, A.R. et al., 1997; Número de acesso GenBank: CAA57472; Escore BLAST: 448). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere muito que o gene orf*3c codifica uma glicosiltransferase. Essa hipótese é reforçada pelo fato de a proteína codificada pelo gene orf3*c apresenta uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 4).

TABELA 4

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Glicosil transferase (S. venezuelae)	AAC68677	426	Glicosil transferase
Glicosil transferase (S. antibioticus)	CAA05642	425	Glicosil transferase
Glicosil transferase (Saccharopolyspora erythraea)	CAA74710	395	Glicosil transferase
Glicosil transferase (S. antibioticus)	CAA05641	394	Glicosil transferase

- * Maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

- O gene orf4*c codifica uma proteína que apresenta uma relativa forte similitude com vários crotonil-CoA redutases. Notadamente, a proteína codificada por orf4*c possui uma similitude importante com uma crotonil CoA redutase de *streptomyces coelicolor* (Redenbach, M et al., 1996; Número de acesso GenBank: NP 630556; Escore BLAST: 772). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere muito que o gene orf4*c codifique também uma crotonil-CoA redutase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf4* apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 5)

TABELA 5

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Trans-2-enoil-CoA redutase (EC1.3.1.28) (<i>S. collinus</i>)	S72400	764	Trans-2-enoil-CoA redutase
Crotonil CoA redutase (<i>S. fradiae</i>)	CAA57474	757	Crotonil CoA redutase
Crotonil-CoA redutase (<i>S. cinnamomensis</i>)	AAD53915	747	Crotonil-CoA redutase

* maior similitude de seqüência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

- O gene orf6* apresenta uma certa similitude como gene mdmB presente nas *streptomyces mycarofaciens* (Hara et Hutchinsnson, 1992; número de acesso GenBank: A42719; Escore BLAST: 489) produtor de antibiótico macrolídeo. Nesse produtor, o gene é implicado na acilação do ciclo-lactona. O gene orf6* codificaria, portanto, uma aciltransferase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf6* apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 6).

Tabela 6

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
AcyA (<i>Streptomyces thermotolerans</i>)	J4001	450	Macrolídeo 3-O-aciltransferase
Midecamicina 4"-O-propionil transferase (<i>S. mycarofaciens</i>)	BAA09815	234	Midecamicina 4"-O-propioniltransferase
Micarose O-aciltransferase (<i>Micromonospora megalomicea subsp. Nigra</i>)	AAG13909	189	Micarose O-aciltransferase

* Maior similitude de seqüência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

- A inativação do gene orf6* foi realizada por uma dele-

ção/inserção em fase e foi mostrado que a cepa resultante não produz mais espiramicina II e III, mas unicamente a espiramicina I (Conforme Figura 1). Isto confirma que o gene *orf6** está bem implicado na síntese de espiramicina II e III. A enzima codificada por esse gene é responsável pela formação de espiramicina II e III por fixação de um grupo acetila ou butirila sobre o carbono 3. As cepas que não expressam mais a proteína codificada pelo gene *orf6** são particularmente interessantes, já que elas não produzem mais espiramicina II e III, mas somente a espiramicinas I. Conforme foi precisado acima, a atividade antibiótica da espiramicina I é nitidamente superior àquela das espiramicina II e III (Liu et al., 1999).

O gene *orf5** codifica uma proteína que apresenta uma relativa forte similitude com vários O-metiltransferase. Notadamente, a proteína codificada por *orf5** possui uma similitude importante com uma O-metiltransferase (EC 2.1.1.) MdmC de *streptomyces mycarofaciens* (Hara & Hutchinson, 1992; Número de acesso GenBank: B42719; Escore BLAST: 355). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico sugere muito que o gene *orf5** codifique também uma O-metiltransferase. O gene *orf5** seria implicado na formação de precursores incorporados no ciclolactona. Com efeito, a partir das comparações de seqüências, o produto gene *orf5** é também relativamente próximo de FkbG que é responsável pela metilação do hidroximalonil-ACP a partir (Wu et al., 2000; Hoffmeister et al., 2000; Número de acesso GenBank: AAF86386; Escore BLAST: 247 (Conforme Figura 8). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene *orf5** apresenta uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 7).

Tabela 7

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Provável O-metiltransferase (EC 2.1.1.)safC(<i>Myxococcus xanthus</i>)	T18553	223	O-metiltransferase
4-O-metiltransferase (EC 2.1.1.)-(<i>Streptomyces</i> sp.)	JC4004	222	O-metiltransferase

* maior similitude de seqüência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

Graças ao efeito polar da inserção de um cassette não excisado no gene *orf6**, pôde ser determinado que o gene *orf5** é essencial à via de biossíntese das espiramicinas. Com efeito, a inserção do cassette excisável na parte codificante do gene *orf6** acarreta uma parada total da produção em espiramicinas. Todavia, uma vez que o cassette inserido tenha sido excisado (e, portanto, quando só o gene *orf6** é inativado, conforme exemplos 14 e 15), encontra-se uma produção em espiramicinas I. Isto mostra que o gene *orf5** é essencial à via de biossíntese das espiramicinas, já que sua inativação acarreta uma parada total da produção em espiramicinas.

O gene *orf5** codifica uma proteína que apresenta uma relativa forte similitude com vários O-metiltransferase. O gene *orf5** seria uma O-metil transferase implicada seja diretamente na síntese do platenolídeo, seja na síntese de um precursor metilado (metoximalonil, ver a Figura 8) incorporado no platenolídeo pela PKS. Para verificar essa hipótese, experiências de análise CL/SM e de RMN foram feitas sobre uma cepa de *S. ambofaciens* de genótipo: *orf6*::att1Ωhyg+*. Nessa cepa o gene *orf5** não é expresso por causa do efeito polar de inserção, no gene *orf6**, do cassette que contém terminadores de transcrição (conforme exemplo 27). Foi mostrado que essa cepa produz uma molécula cujo espectro UV tem um comportamento similar àquele da espiramicina I, mas o espectro de massa mostra um íon molecular em 829. A diferença de massa de 14 em relação à massa da espiramicina pode ser explicada pela ausência de metil sobre o oxigênio levada pelo carbono N° 4 do ciclolactona (a estrutura desse composto é apresentada na

Figura 39). Os resultados obtidos por RMN são compatíveis com essa hipótese. A presença de um composto em 829 permite validar a hipótese do papel de orf5* na biossíntese das espiramicinas. Além disso, o produto correspondente à espiramicinas sem o grupo metila apresenta uma atividade microbiológica muito fraca (menor de um fator 10) em relação à espiramicina não modificada, quando testada sobre o microorganismo *Micrococcus luteus*.

O gene orf7*c codifica uma proteína que apresenta uma relativa forte similitude com uma proteína codificada pelo gene mdmA de *streptomyces mycarofaciens*, este codificando uma proteína implicada na resistência à midecamicina nesse organismo produtor (Hara et al., 1990; Número de acesso GenBank: A60725; Escore BLAST: 380). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico sugere fortemente que o gene orf7*c codifique também uma proteína implicada na resistência a espiramicina. Mais particularmente, a enzima codificada pelo gene orf7*c possui uma atividade metiltransferase e é implicada na resistência a espiramicinas em *streptomyces ambofaciens*. Foi demonstrado que esse gene confere uma resistência de tipo MLS I, resistência da qual se sabe que é devido à monometilação em posição A2058 do RNA ribossomal 23S (Pernodet et al., 1996). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf7*c apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 8).

TABELA 8

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Macrolídeo-lincosamida-estreptogramin B resistência determinante (<i>S. fradiae</i>)	JC5319	238	23S rRNA metiltransferase
23S ribossomal RNA metiltransferase ErmML (<i>Micrococcus luteus</i>)	AAL68827	119	23S rRNA metiltransferase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

O gene orf8* codifica uma proteína que apresenta uma relativa forte similitude com uma proteína de tipo transportador ABC em *streptomyces* griseus (Campelo, 2002, Número de acesso GenBank: CAC22119; Escore BLAST: 191). Essa similitude com uma proteína de tipo transportador ABC sugere fortemente que o gene orf8* codifique também uma proteína de tipo transportador ABC que pode ser implicada na resistência a espiromicinas. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf8* apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 9).

10

TABELA 9

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
AcrW (<i>Streptomyces galilaeus</i>)	BAB72060	94	Transportador ABC
Daunorubicina resistência transmembrana proteína (<i>Streptomyces peucetius</i>)	P32011	89	Resistência à daunorubicina
Probable ABC-transportador, transmembrana componente (<i>Streptomyces coelicolor</i>)	NP 626506	89	Transportador ABC

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990)

O gene orf9* codifica uma proteína que apresenta uma relativa forte similitude com uma proteína de tipo transportador ABC em *streptomyces* griseus (Campelo, 2002, Número de acesso GenBank: CAC22118; Escore BLAST: 269). Essa similitude com uma proteína de tipo transportador ABC sugere fortemente que o gene orf9* codifique também uma proteína de tipo transportador ABC que pode ser implicada na resistência a espiromicina. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf8* apresentar uma forte similitude com, outras proteínas de função similar em outros organismos (conforme tabela 10).

TABELA 10

Proteína apresentando uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Provável componente de ligação de ATP de proteína de transporte do tipo ABC (S. coelicolor)	NP-626505	231	Transportador ABC
Componente de ligação de ATP de transportador ABC putativo (Streptomyces Coelicolor)	NP-627624	228	Transportador ABC

* maior similitude de seqüência é associada a um escore BLAST mais elevado (Altschul et al. , 1990).

O gene orf10* codifica uma proteína que apresenta uma relativa forte similitude com uma proteína de função desconhecida. Todavia, genes similares a orf10* são encontrados no meio de vários grupos de genes implicados na biossíntese de antibióticos. Assim, um gene próximo de orf10* é encontrado em S. coelicolor (Redenbach et al., 1996, número de acesso GenBank: NP-627432, escore BLAST: 109). Um gene próximo (CouY) é também encontrado em S. rishiriensis (Wang et al., 2000, número de acesso GenBank"AAG29779, escore BLAST 97).

Genes situados a montante dos genes codificando os PKS

Na seqüência de DNA situada a montante dos genes codificando os PKS (a jusante e a montante sendo definido pela orientação dos 5 genes de PKS todos orientados no mesmo sentido) (conforme Figura 3), 34 ORFs foram identificados (conforme acima). Assim, as 34 fases abertas de leitura desse tipo, ocupando uma região de aproximadamente 41,7 kb (conforme SEQ ID Nº 1 que apresenta uma primeira região de aproximadamente 31 kb contendo 25 ORFs e a SEQ ID Nº 140 apresentando uma região de aproximadamente 12,1 kb, dos quais 1,4 kb se sobrepõem a seqüência precedente (SEQ ID Nº 1) e aproximadamente 10,7 kb correspondem ao sítio da seqüência), esta parte de aproximadamente 10,7 kb, contendo 9 ORFs suplementares (das quais um ORF de seqüência parcial), conforme também Figura 3 e 37 e abaixo). Uma representação esquemática da organização da re-

gião está presente na Figura 3 e 37. Os 34 genes identificados foram chamados: *orf1*, *orf2*, *orf3*, *orf4*, *orf5*, *orf6*, *orf7*, *orf8*, *orf9c*, *orf10*, *orf11c*, *orf12*, *orf13c*, *orf15*, *orf15c*, *orf16*, *orf17*, *orf18*, *orf19*, *orf20*, *orf21c*, *orf22c*, *orf23c*, *orf24c*, *orf25c*, *orf26*, *orf27*, *orf28c*, *orf29*, *orf30c*, *orf31*, *orf32c*, *orf33* e *orf34c*.

Na tabela 11 a seguir são apresentadas as referências à sequência em DNA e em aminoácidos dos 34 genes identificados a montante dos 5 genes PKS.

TABELA 11

Gene ¹	Posição na sequência SEQ ID N 1	Seqüência DNA	Seqüência(s) polipeptídica(s) ²
<i>orf1</i>	658 a 1869	SEQ ID N 23	SEQ ID N 24
<i>orf2</i>	1866 a 2405	SEQ ID N 25	SEQ ID N 26 e 27
<i>orf3</i>	2402 a 3568	SEQ ID N 28	SEQ ID N 29
<i>orf4</i>	3565 a 4473	SEQ ID N 30	SEQ ID N 31, 32 e 33
<i>orf5</i>	4457 a 5494	SEQ ID N 34	SEQ ID N 35
<i>orf6</i>	5491 a 6294	SEQ ID N 36	SEQ ID N 37, 38 e 39
<i>orf7</i>	6296 a 7705	SEQ ID N 40	SEQ ID N 41 e 42
<i>orf8</i>	8011 a 9258	SEQ ID N 43	SEQ ID N 44
<i>orf9</i>	10081 a 9362	SEQ ID N 45	SEQ ID N 46
<i>orf10</i>	10656 a 12623	SEQ ID N 47	SEQ ID N 48
<i>orf11c</i>	14482 a 12734	SEQ ID N 49	SEQ ID N 50, 51 e 52
<i>orf12</i>	14601 a 16031	SEQ ID N 53	SEQ ID N 54, 55, 56, 57, 58 e 59
<i>orf13c</i>	17489 a 16092	SEQ ID N 60	SEQ ID N 61
<i>orf14</i>	17809 a 18852	SEQ ID N 62	SEQ ID N 63
<i>orf15</i>	20001 a 18961	SEQ ID N 64	SEQ ID N 65
<i>orf16</i>	20314 a 21552	SEQ ID N 66	SEQ ID N 67
<i>orf17</i>	21609 a 22879	SEQ ID N 68	SEQ ID N 69
<i>orf18</i>	22997 a 24175	SEQ ID N 70	SEQ ID N 71
<i>orf19</i>	24177 a 25169	SEQ ID N 72	SEQ ID N 73
<i>orf20</i>	25166 a 26173	SEQ ID N 74	SEQ ID N 75

Gene ¹	Posição na seqüência SEQ ID N 1	Seqüência DNA	Seqüência(s) polipeptídica(s) ²
<i>orf21c</i>	27448 a 26216	SEQ ID N 76	SEQ ID N 77
<i>orf22c</i>	28560 a 27445	SEQ ID N 78	SEQ ID N 79
<i>orf23c</i>	29770 a 28649	SEQ ID N 80	SEQ ID N 81
<i>orf24c</i>	30074 a 29763	SEQ ID N 82	SEQ ID N 83
<i>orf25c</i>	30937 a 30071	SEQ ID N 83	SEQ ID N 84
Gene ¹	Posição na seqüência SEQ ID N 140	Seqüência DNA	Seqüência(s) polipeptídica(s) ²
<i>orf26</i>	1647 a 2864	SEQ ID N 107	SEQ ID N 108
<i>orf27</i>	2914 a 3534	SEQ ID N 109	SEQ ID N 110
<i>orf28c</i>	4967 a 3804	SEQ ID N 141	SEQ ID N 142
<i>orf29</i>	5656 a 6663	SEQ ID N 113	SEQ ID N 114
<i>orf30c</i>	7723 a 6686	SEQ ID N 115	SEQ ID N 116 e 117
	7534 a 6686	SEQ ID N 143	SEQ ID N 144
<i>orf31</i>	7754 a 8728	SEQ ID N 118	SEQ ID N 119
<i>orf32c</i>	19488 a 8977	SEQ ID N 145	SEQ ID N 146
<i>orf33</i>	19562 a 10837	SEQ ID N 147	SEQ ID N 148
<i>orf34c</i>	12134 a 10899	SEQ ID N 149	SEQ ID N 150

1. O "c" acrescentado no nome do gene indica que a seqüência codificante está na orientação inversa (o filamento codificante é, portanto, o filamento complementar da seqüência determinada em SEQ ID N° 1 ou SEQ ID N° 140 para esses genes).

- 5 2. Quando várias seqüências protéicas são indicadas para uma única orf, as proteínas correspondentes são provenientes de vários sítios possíveis de acionamento da tradução.

10 Com a finalidade de determinar a função dos polipeptídeos identificados na tabela 11 acima, três tipos de experiências foram feitas: a comparação da identidade das seqüências identificadas com seqüências de funções conhecidas, experiências de inativação de genes e análises da produção em espiramicinas dessas cepas mutantes.

As seqüências protéicas deduzidas dessas fases abertas de lei-

5 tura foram inicialmente comparadas com aquelas presentes em diferentes bases de dados graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et al., 1997), CD-search, COGs (Cluster of Orthologous Groups), FASTA (Pearson W. R & D.J. Lipman, 1988) e (Pearson W.R., 1990), BEAUTY (Worley K.C. et al., 1995)), (conforme acima). Essas comparações permitiram formular hipóteses sobre a função dos produtos desses genes e de identificar aqueles capazes de serem implicados na biossíntese de espiamicinas. A tabela 12 mostra as proteínas que apresentam uma forte similitude com os 34 genes situados a montante dos 5 genes PKS.

10

TABELA 12

Gene	Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
<i>orf1</i>	Citocromo P450 Ty1I (<i>Streptomyces fradiae</i>)	S49051	530	Citocromo P45
<i>orf2</i>	ORF15x4 (<i>Listonella anguillarum</i>)	AAB81630	113	Desconhecida
<i>orf3</i>	Proteína Aminotransferase-similar (<i>Streptomyces antibioticus</i>)	AAF59939	431	Aminotransferase
<i>orf4</i>	Alfa-D-glicose-1-fosfato timidilil-transferase (<i>Streptomyces venezuelae</i>)	AAC68682	404	Alfa-D-glicose-1-fosfato timidilil-transferase
<i>orf5</i>	AprE (<i>Streptomyces tenebrarius</i>)	AAG18457	476	dTDP-glicose 4,6-desidratase
<i>orf6</i>	Tioesterase (<i>Streptomyces avermitilis</i>)	BAB69315	234	Tioesterase
<i>orf7</i>	TyICVI (<i>Streptomyces fradiae</i>)	AAF29379	461	dNTP hexose 2,3-desidratase
<i>orf8</i>	provável aminotransferase	AAG23279	465	Aminotransferase
<i>orf9c</i>	SrmX (<i>Streptomyces ambofaciens</i>)	S25204	445	Metiltransferase
<i>orf10</i>	SrmR (<i>Streptomyces ambofaciens</i>)	S25203	1074	Proteína de regulação
<i>orf11c</i>	SrmB (<i>Streptomyces ambofaciens</i>)	S25202	955	Resistência à

Gene	Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
				espiramicina
orf12	UrdQ (<i>Streptomyces fradiae</i>)	AAF72550	634	NDP-hexose 3,4 desidratase
orf13c	SC4H2,17 (<i>Streptomyces coelicolor</i>)	T35116	619	Desconhecida
orf14	Putativa redutase (<i>Streptomyces coelicolor</i>)	CAB90862	147	Redutase
orf15c	Provável 3-cetorredutase (<i>Streptomyces antibioticus</i>)	T51102	285	3-cetorredutase
orf16	Hipotetical NDP hexose 3,4 isomerase (<i>Streptomyces fradiae</i>)	CAA57471	209	NDP hexose 3,4 isomerase
orf17	Glicosil transferase (<i>Streptomyces venezuelae</i>)	AAC68677	400	Glicosil transferase
orf18	Glicosil transferase (<i>Streptomyces rishiriensis</i>)	AAG29785	185	Glicosil transferase
orf19	NDP-hexose-4-cetorredutase TylCIV (<i>Streptomyces fradiae</i>)	AAD41822	266	NDP-hexose-4-cetorredutase
orf20	EryBII (<i>Saccharopolyspora erythraea</i>)	AAB84068	491	Aldocéteo redutase
orf21c	TylCIII (<i>Streptomyces fradiae</i>)	AAD41823	669	NDP-hexose 3-C metiltransferase
orf22c	FkbH (<i>Streptomyces hygroscopicus</i>)	AAF86387	463	Implicado na biossíntese da metoximalonila
orf23c	Fkbl (<i>Streptomyces hygroscopicus</i>)	AAF86388	387	Acyl CoA desidrogenase
orf24c	FkbJ (<i>Streptomyces hygroscopicus</i>)	AAF86389	87	Implicado na biossíntese da metoximalonila
orf25c	FkbK (<i>Streptomyces hygroscopicus</i>)	AAF86390	268	Acyl CoA desidrogenase

Gene	Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
<i>orf26</i>	TylCV (<i>Streptomyces fradiae</i>)	AAD41824	471	Micarosil transferase
<i>orf27</i>	TylCVII (<i>Streptomyces fradiae</i>)	AAD41825	243	NDP-hexose-3,5-(ou 5-)epimerase
<i>orf28c</i>	AcyB2 (<i>Streptomyces thermotolerans</i>)	JC2032	451	Proteína reguladora
<i>orf29</i>	Beta-manannase (<i>Sorangium cellulosum</i>)	AAK19890	139	Glicosil hidrolase
<i>orf30c</i>	Epimerase de açúcar nucleosídeo difosfato (<i>Corynebacterium glutamicum</i>)	NP_600590	89	Epimerase de açúcar nucleosídeo difosfato
<i>orf31</i>	Oxidoredutase (<i>Streptomyces coelicolor</i>)	NP_631148	261	Oxidoredutase
<i>orf32c</i>	Proteína reguladora da família GntR (<i>Streptomyces avermitilis</i>)	NP_824604	282	Proteína reguladora
<i>orf33</i>	Proteína hipotética (<i>Xanthomonas campestris</i>)	NP_635564	54	Desconhecida
<i>orf34c</i>	Arabinofuranosidase (<i>Streptomyces coelicolor</i>)	NP_630049	654	Arabinofuranosidase

* maior similitude de seqüência é associada a um açúcar BLAST mais elevado (Altschul et al., 1990).

Experiências de inativação de genes foram feitas para confirmar esses resultados. Os métodos aplicados consistem em fazer uma substituição de gene. O gene alvo a interromper é substituído por uma cópia desse gene interrompido por um cassette, conferindo a resistência a um antibiótico (por exemplo, a apramicina ou higromicina). Os cassettes utilizados são contornados de ambos os lados por códons de terminação da tradução em todas as fases de leitura e por terminadores de transcrição ativos em *streptomyces*. A inserção do cassette no gene alvo pode ser acompanhada

ou não de uma deleção nesse gene alvo. O tamanho das regiões que tornam o cassette pode ir de algumas centenas a vários milhares de pares de bases. Um segundo tipo de cassettes pode ser utilizado para a inativação de genes: cassettes ditos "cassettes excisáveis"(conforme acima). As cepas assim construídas foram testadas para sua produção em espiramicinas

O gene orf1 codifica uma proteína que apresenta uma relativa forte similitude com vários citocromo P450. Notadamente, a proteína codificada por orf1 possui uma similitude importante com a proteína codificada pelo gene tyll implicada na biossíntese de tilosina em *streptomyces fradiae* (Merson-D994; Davies, L.A. et al., 1994; Número de acesso GenBank: S49051; Escore BLAST: 530). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere fortemente que o gene orf1 codifica também um citocromo P450. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf1 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 13).

TABELA 13

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Citocromo Putativo P450 YJIB (Bacillus subtilis)	034374	248	Citocromo P450
Citocromo P450 113A1 (Saccharopolyspora erythraea)	P48635	237	Citocromo P450
Citocromo P450 Hidroxilase homóloga (Streptomyces caelestis)	AAC46023	208	Citocromo P450 hidroxilase
Citocromo P450 Monooxygenase (Streptomyces venezuelae)	AAC64105	206	Citocromo P450 monooxygenase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

O gene orf2 codifica uma proteína que apresenta uma relativa forte similitude com uma dTDP-6-deóxi-3,4-ceto-hexulose isomerase de Aneurinibacillus thermoaerophilus (Pfoestl, A. et al., 2003, Número de aces-

so GenBank: AAO06351; Escore BLAST: 118). Essa similitude sugere fortemente que o gene orf2 codifica uma isomerase responsável pela reação de isomerização necessária à biossíntese de um dos açúcares presentes na molécula de espiramicina, esse açúcar podendo ser a micarose (conforme Figura 5). A inativação do gene orf2 foi feita. Pôde ser mostrado que a cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf2 é bem implicado na biossíntese das espiramicinas.

O gene orf3 codifica uma proteína que apresenta uma relativa forte similitude com várias aminotransferases. Notadamente, a proteína codificada por orf3 possui uma similitude importante com uma aminotransferase de *Streptomyces antibioticus* implicada na biossíntese da oleandomicina (Draeger, G. et al. 1999; Número de acesso GenBank: AAF59939; Escore BLAST: 431). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere fortemente que o gene orf3 codifica uma 3 amino-transferase responsável pela reação de transaminação necessária à biossíntese de um dos açúcares aminados das espiramicinas (conforme Figura 5). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf3 apresentar uma forte similitude com outras proteínas de função similitude em outros organismos (conforme tabela 15).

20

TABELA 15

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Aminotransferase (<i>Streptomyces antibioticus</i>)	T51111	429	Aminotransferase
Transaminase (<i>Streptomyces venezuelae</i>)	AAC68680	419	Transaminase

* Maior similitude de seqüência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

A inativação do gene orf3 foi realizada. Foi possível assim mostrar que a cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf3 é bem implicado na biossíntese das espiramicinas. Isto confirma

25

que o gene orf3 é bem implicado na biossíntese das espiramicinas. A enzima codificada por esse gene é, portanto, bem responsável por uma etapa de bioconversão essencial à biossíntese das espiramicinas. A produção de espiramicinas pode ser complementada pela expressão da proteína TylB de *S. fradiae* (cf exemplo 23). Isto demonstra que o gene orf3 codifica uma 3 amino-transferase responsável pela reação de transaminação necessária à biossíntese da micaminose (conforme Figura 5). Como a micaminose é o primeiro açúcar a ser fixado sobre o platenolídeo, é esperado que a cepa interrompida em orf3 (OS49.67) acumule o platenolídeo.

Os intermediários de biossíntese da cepa interrompida no gene orf3 foram estudados (conforme exemplo 20). Essas experiências permitiram demonstrar que essa cepa produz duas formas de platenolídeo: o platenolídeo A e o platenolídeo B, a estrutura deduzida dessas duas moléculas foi apresentada na Figura 36. Essa cepa produz também o platenolídeo A + micarose e o platenolídeo B + micarose (conforme exemplo 20 e Figura 40). Esses compostos comportam um açúcar, mas não comportam micaminose. Além disso, caso sejam comparados a espiramicinas (conforme Figura 1), esses compostos comportam micarose no lugar da micaminose. Esses resultados estão em acordo com a implicação do produto do gene orf3 na biossíntese da micaminose e seu papel como 3 amino-transferase da reação de transaminação necessária à biossíntese da micaminose (conforme Figura 5). Pode-se observar que a especificidade de glicosilação não parece absoluta, já que se acham moléculas com a micarose fixada na posição normalmente ocupada pela micaminose (conforme Figura 40).

O gene orf4 codifica uma proteína que apresenta uma relativa forte similitude com várias NDP-glicose sintetases. Notadamente, a proteína codificada por orf4 possui uma similitude importante com uma alfa-D-glicose-1-fosfato timidililtransferase de *streptomyces venezuelae* (Xue Y et al., 1998; Número de acesso GenBank: AAC68682; Escore BLAST 404). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere fortemente que o gene orf4 codifica uma NDP-glicose sintetase responsável pela síntese do NDP-glicose necessária à biossíntese

dos três açúcares atípicos incorporados na molécula de espiramicina (conforme Figuras 4, 5 e 6). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf4 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 16).

5

TABELA 16

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Glicose-1-fosfato timidiltransferase (<i>Streptomyces avermitilis</i>)	BAA84594	402	Glicose-1-fosfato timidiltransferase
Ac1Y(<i>Streptomyces galilaeus</i>)	BAB72036	400	dTDP-1-glicose sintetase
Putativa glicose-1-fosfato timidiltransferase(<i>Saccharopolyspora spinosa</i>)	AAK83289	399	Glicose-1-fosfato timidiltransferase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990).

O gene orf5 codifica uma proteína que apresenta uma relativa forte similitude com glicose desidratases. Notadamente, a proteína codificada por orf5 possui uma similitude importante com uma dTDP-glicose 4,6-desidratase de *streptomyces tenebrarius* (Li,T.B. et al., 2001; Número de acesso GenBank: AAG18457, Escore BLAST: 476). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere fortemente que o gene orf5 codifique uma NDP glicose desidratase necessária à biossíntese dos três açúcares atípicos incorporados na molécula de espiramicinas (conforme Figura 4, 5 e 6). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf5 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 17)

10

15

TABELA 17

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
dTDP-glicose 4,6-desidratase (<i>Saccharopolyspora spinosa</i>)	AAK83290	464	dTDP-glicose 4,6-desidratase
Timidina difosfoglicose 4,6-desidratase (<i>Saccharopolyspora erythraea</i>)	AAA68211	445	Timidina difosfoglicose 4,6-desidratase
dTDP-glicose 4,6-desidratase (EC 4.2.1.46)-(<i>Streptomyces fradiae</i>)	S49054	443	dTDP-glicose 4,6-desidratase
TDP-glicose 4,6-desidratase (<i>Streptomyces venezuelae</i>)	AAC68681	421	TDP-glicose 4,6-desidratase
SgcA(<i>Streptomyces globisporus</i>)	AAF13998	418	dNDP-glicose-4,6-desidratase

* maior similitude de seqüência é associada a um escore BLAST mais elevado (Altschul et al., 1990).

- O gene orf6 codifica uma proteína que apresenta uma relativa
- 5 forte similitude com várias tioesterases. Notadamente, a proteína codificada por orf6 possui uma similitude importante com uma tioesterase de *streptomyces avermitilis* (Omura, S. et al., 2001; Número de acesso GenBank: BAB69315; Escore BLAST: 234). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere
- 10 fortemente que o gene orf6 codifique também uma tioesterase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf6 apresentar uma forte similitude com outras proteínas de função similitude em outros organismos (conforme tabela 18).

TABELA 18

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
RifR (<i>Amycolatopsis mediterranei</i>)	AAG52991	216	Tioesterase
Tioesterase -(<i>Streptomyces fradiae</i>)	S49055	215	Tioesterase
Tioesterase (<i>Streptomyces avermitilis</i>)	BAB69188	213	Tioesterase
Tioesterase II (EC 3.1.2.)-(<i>Streptomyces venezuelae</i>)	T17413	203	Tioesterase
Proteína PimI (<i>Streptomyces natalensis</i>)	CAC20922	201	Tioesterase
Tioesterase (<i>Streptomyces griseus</i>)	CAC22116	200	Tioesterase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990).

- O gene orf7 codifica uma proteína que apresenta uma relativa forte similitude com várias hexoses desidratases. Notadamente, a proteína codificada por orf7 possui uma similitude importante com uma dNTP hexose 2,3-desidratase (codificada pelo gene TylCVI) de *streptomyces fradiae* implicada na biossíntese da tilosina (Merson-Davies, L.A. et al., 1994; Número de acesso GenBank: AAF29379; Escore BLAST: 461). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere fortemente que o gene orf7 codifique também uma hexose 2-3 desidratase necessária à biossíntese de dois açúcares atípicos incorporados na molécula de espiamicinas (conforme Figura 4 e 6). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf7 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 19).

TABELA 19

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Rif18 (<i>Amycolatopsis mediterranei</i>)	AAG52988	459	Hexose desidratase
SimB3 (<i>Streptomyces antibioticus</i>)	AAK06810	444	dNDP-4-ceto-6-deóxi-glicose-2,3-desidratase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990).

O gene orf8 codifica uma proteína que apresenta uma relativa forte similitude com várias aminotransferases. Notadamente, a proteína codificada por orf8 possui uma similitude importante com uma aminotransferase provavelmente implicada na biossíntese da forosamina em *Saccharopolyspora spinosa* (Waldron, C. et al., 2001; Número de acesso GenBank: AAG23279; escore BLAST: 465). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo, sugere fortemente que o gene orf8 codifique uma 4 amino-transferase responsável pela reação de transaminação necessária à biossíntese da forosamina (conforme Figura 6). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf8 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 20).

TABELA 20

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Proteína putativa na biossíntese de amino-açúcar (<i>Bordetella bronchiseptica</i>)	CAA07666	213	Proteína implicada na biossíntese de amino-açúcar

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990).

A inativação do gene orf8 foi realizada. Pôde assim ser mostrado que a cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf8 é bem implicado na biossíntese das espiramicinas. A enzima codificada é, portanto, bem responsável por uma etapa de biossíntese essencial

à biossíntese das espiramicinas. A validação da hipótese do papel exercido pelo produto do gene orf8 na biossíntese da forosamina é fornecida pelo fato de um mutante inativado pelo gene orf8 produto da forocidina, esse mutante sendo, portanto, bloqueado no estágio da forocidina e não produz neo-
 5 espiramicina (conforme Figura 7 e Exemplo 25). Esses resultados estão em acordo com uma implicação do produto do gene orf8 na biossíntese de forosamina (conforme Figura 6).

O gene orf9c foi identificado nas *streptomyces ambofaciens* e foi designado srmX por Geistlich et al. (Geistlich, M. et al., 1992). A similitude da
 10 proteína codificada por esse gene com vários metiltransferases implicadas na via de biossíntese de outros antibióticos próximos sugere fortemente que o gene orf9c codifica uma metiltransferase responsável pela reação de metilação necessária à biossíntese da micaminose ou da forosamina (conforme Figuras 5 e 6). Essa hipótese é confirmada pelo fato de a proteína codificada
 15 pelo gene orf9c apresentar uma forte similitude com outras proteínas de função semelhante em outros organismos (conforme tabela 21).

TABELA 21

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
N,N-dimetiltransferase (<i>Streptomyces venezuelae</i>)	AAC68678	240	N,N-dimetiltransferase
Metiltransferase (<i>Streptomyces antibioticus</i>)	CAA05643	232	Metiltransferase
Putativo amino metilase (<i>Streptomyces nogalater</i>)	AAF01819	219	Aminometilase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

20 O gene orf10 foi identificado nas *streptomyces ambofaciens* e foi designado srmR por Geistlich et al. (Geistlich, M. et al., 1992). A proteína codificada por esse gene é implicada na regulação da via de biossíntese das espiramicinas nas *streptomyces ambofaciens*. A inativação do gene

orf10 foi realizada. Foi possível assim mostrar que a cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf10 é bem implicado na biossíntese das espiramicinas. A proteína codificada por esse gene é, portanto, bem essencial à biossíntese das espiramicinas.

5 Por outro lado, o ponto de acionamento da tradução de orf10 foi determinado e pôde ser mostrado que a superexpressão desse gene acarreta uma melhoria da produção em espiramicinas. O sítio de acionamento da tradução corresponde a um ATG situado a montante do ATG proposto por Geistlich et al. (Geistlich, M. et al., 1992). Foi além disso demonstrado que
10 essa extremidade 5' é essencial à função de Orf10, já que um mensageiro truncado em 5' é inativo (conforme exemplo 17). Para se obter o efeito buscado sobre a produção em espiramicinas, é, portanto, essencial que a superexpressão de orf10 seja realizada tomando-se o cuidado de não expressar um mensageiro truncado em 5' de orf10.

15 O gene orf11c foi identificado nas *streptomyces* ambofaciens e foi designado srmB por Geistlich et al. (Geistlich, M. et al., 1992) e Schoner et al. (Schoner B et al., 1992). A proteína codificada por esse gene é implicada na resistência a espiramicina nas *streptomyces* ambofaciens e é um transportador de tipo ABC.

20 O gene orf12 codifica uma proteína que apresenta uma relativa forte similitude com várias hexoses desidratases. Notadamente, a proteína codificada por orf12 possui uma similitude importante com uma NDP-hexose 3,4-desidratase codificada pelo gene UrdQ de *streptomyces* fradiae e implicada na biossíntese da urdamicina (Hoffmeister, D. et al., 2000; Número de
25 acesso GenBank: AAF72550; Escore BLAST: 634). Essa similitude com uma proteína implicada na via de biossíntese de um outro antibiótico próximo sugere fortemente que o gene orf12 codifica uma 3,4 desidratase responsável pela reação de desidratação necessária à biossíntese da forosamina (conforme Figura 6). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf12 apresentar uma forte similitude com outras proteínas de
30 função similar em outros organismos (conforme tabela 22).

TABELA 22

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Aknp (<i>Streptomyces galilaeus</i>)	AAF73452	625	3-desidratase
NDP-hexose 3,4-desidratase homólogo (<i>Streptomyces cyanogenus</i>)	AAD13547	624	NDP-Hexose 3,4-desidratase
Rdml (<i>Streptomyces purpurascens</i>)	AAL24451	608	Hexose-C-3-desidratase
Provável CDP-4-ceto-6-deoxiglicose-3-desidratase (E1) (<i>Streptomyces violaceoruber</i>)	T46528	602	CDP-4-ceto-6-deoxiglicose-3-desidratase
Provável NDP-hexose-3,4-desidratase (<i>Saccharopolyspora spinosa</i>)	AAG23278	582	NDP-hexose-3,4-desidratase
dNTP-hexose desidratase (<i>Amycolatopsis mediterranei</i>)	AAC01730	576	dNTP-hexose desidratase

* maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990)

- A inativação do gene orf12 foi feita. Pôde ser mostrado que a
- 5 cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf12 é bem implicado na biossíntese da espiramicina. A enzima codificada por esse gene é, portanto, bem responsável por uma etapa de bioconversão essencial à biossíntese da espiramicina. A validação da hipótese do papel exercido por Orf12 na biossíntese da forosamina é fornecida pelo fato de um
- 10 mutante inativado para o gene orf12 não produzir mais forosamina. Ele produz, todavia, uma pequena quantidade de forocidina. Esse mutante é, portanto, bloqueado no estágio da forocidina e não produz neo-espiramicina (conforme Figura 7 e exemplo 26). Esse mutante produz, além disso, um composto de estrutura apresentada na Figura 38. Este último composto
- 15 comporta dois açúcares, a micaminose e a micarose, mas não comporta fo-

rosamina. Além disso, caso se compare à estrutura da espiramicina (conforme Figura 1), esse composto comporta o açúcar micarose no lugar esperado da forosamina. Esses resultados estão em acordo com a implicação do produto do gene orf12 na biossíntese da forosamina (conforme Figura 6).

- 5 Pode-se observar que a especificidade de glicosilação não é absoluta, já que se observam moléculas nas quais a micarose é fixada na posição normalmente ocupada pela forosamina (ver a Figura 38).

O gene orf13c codifica uma proteína que apresenta uma relativa forte similitude com uma proteína de função desconhecida nas *streptomyces* coelicolor. Essa proteína foi chamada SC4H2.17 (número de acesso GeneBank: T35116; Escore BLAST: 619). A proteína codificada pelo gene orf13c apresenta também uma forte similitude com outras proteínas de outros organismos (conforme tabela 23)

TABELA 23

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Proteína hflX (<i>Mycobacterien leprae</i>)	S72938	473	Desconhecida
Possível ATP/GTP-proteína de ligação (<i>Mycobacterium leprae</i>)	NP-301739	470	Proteína que fixa o ATP/GTP
GTP-proteína de ligação (<i>Mycobacterium tuberculosis</i> CDC1551)	AAK47114	468	Proteína que fixa o GTP
ATP/GTP-proteína de ligação (<i>Streptomyces fradiae</i>)	T44592	388	Proteína que fixa o ATP/GTP

- 15 * maior grande similitude de seqüência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

Nenhuma função precisa foi assinada nas proteínas próxima daquela codificada por orf13c. A inativação do gene orf13c foi realizada com a finalidade de estudar a função desse gene na via de biossíntese das espira-

micinas nas *streptomyces* ambofaciens. Pôde ser mostrado que a cepa resultante produz espiramicinas. Isto indica que o gene orf13c não é essencial à biossíntese das espiramicinas e que não é essencial à sobrevivência da bactéria. A enzima codificada por esse gene não é, portanto, responsável por uma
5 etapa de bioconversão essencial à biossíntese das espiramicinas.

O gene orf14 codifica uma proteína que apresenta uma relativa forte similitude com uma redutase putativa (Redenbach, M. et al., 1996; Bentley et al., 2002; número de acesso GenBank: CAB90862; Escore BLAST: 147).

10 A inativação do gene orf14 foi feita. Pôde assim ser mostrado que a cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf14 é bem implicado na biossíntese das espiramicinas. A enzima codificada por esse gene é, portanto, bem responsável por uma etapa de bioconversão essencial à biossíntese das espiramicinas. Os intermediários de
15 biossíntese da cepa interrompida no gene orf14 foram estudados (conforme exemplo 20). Essas experiências permitiram demonstrar que essa cepa produz platenolídeo A, mas não produz platenolídeo B (conforme Figura 36).

O gene orf15c codifica uma proteína que apresenta uma relativa forte similitude com várias ceto-redutases. Notadamente, a proteína codificada por orf15c possui uma similitude importante com uma 3-ceto-redutase nas *streptomyces* antibioticus (Número de acesso GenBak: T51102, Escore BLAST: 285). Essa similitude sugere fortemente que o gene orf15c codifica uma 3 ceto-redutase responsável pela reação de redução necessária à biossíntese da forosamina (conforme Figura 6).Essa hipótese é confirmada pelo
20 fato de a proteína codificada pelo gene orf15c apresentar uma forte similitude com outras proteínas de função similitude em outros organismos (conforme tabela 24).
25

TABELA 24

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função comple- mentar
Óxi-redutase homóloga (<i>Streptomyces cytanogenus</i>)	AAD13550	272	Óxi-redutase
D-oliose 4-ceto-redutase (<i>Streptomyces galilaeus</i>)	CAB96550	265	D-oliose 4-ceto- redutase
AknQ (<i>Streptomyces galilaeus</i>)	AAF73453	263	3-Ceto-redutase putativa
Provável NDP-hexose-3-keto- redutase (<i>Saccharopolyspora spinosa</i>)	AAG23275	253	NDP-hexose-3- ceto-redutase.

* maior similitude de seqüência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

- O gene orf16 codifica uma proteína que apresenta uma relativa
- 5 forte similitude com várias isomerases. Notadamente, a proteína codificada por orf16 possui uma similitude importante com uma NDP hexose 3,4 isomerase nas *streptomyces fradiae* (Gandecha et al., 1997; Número de acesso GenBak: CAA57471, Escore Blast: 209). Essa similitude sugere fortemente que o gene orf16 codifique uma proteína implicada na biossíntese de um dos
- 10 açúcares da espiramicina (conforme Figuras 5 e 6). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf16 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 25).

TABELA 25

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Tautomerase putativa (Streptomyces venezuelae)	AAC68676	145	Tautomerase
TDP-4-ceto-6-deóxi-hexose 3,4-isomerase (Micromonospora megalomicea subsp. Nigra)	AAG13907	112	TDP-4-ceto-6-deóxi-hexose 3,4-isomerase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990)

- O gene orf17 codifica uma proteína que apresenta uma relativa
- 5 forte similitude com várias glicosil transferases. Notadamente, a proteína codificada por orf17 possui uma similitude importante com uma glicosil transferase de *streptomyces venezuelae* (Xue,Y. et al., 1998; Número de acesso GenBank: AAC68677; Escore BLAST: 400). A similitude da proteína codificada pelo gene orf17 com várias glicosil transferases implicadas na via
- 10 de biossíntese de outros antibióticos próximos sugere fortemente que esse gene codifique também uma glicosil transferase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf17 apresentar uma forte similitude com outras proteínas de função similitude em outros organismos (conforme tabela 26).

TABELA 26

Proteína que apresenta um similitude significativa	Número de acesso GenBank	Escore BLAST*	Função comple- mentar
Glicosiltransferase (<i>Streptomyces fradiae</i>)	CAA57472	399	Glicosiltransferase
Glicosiltransferase (<i>Streptomyces antibioticus</i>)	CAA05642	378	Glicosiltransferase
Glicosiltransferase (<i>Streptomyces antibioticus</i>)	CAA05641	360	Glicosiltransferase
Glicosiltransferase (<i>Saccha- ropolyspora erythrae</i>)	CAA74710	344	Glicosiltransferase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990)

- O gene orf18 codifica uma proteína que apresenta uma relativa forte similitude com várias glicosil transferases. Notadamente, a proteína codificada por orf18 possui uma similitude importante com uma glicosil transferase de *streptomyces rishiriensis* (Wang et al., 2000; Número de acesso GenBank: AAG29785; Escore BLAST: 185). A similitude da proteína codificada pelo gene orf18 com várias glicosil transferases implicadas na via de biossíntese de outros antibióticos próximos sugere fortemente que esse gene codifique também uma glicosil transferase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf18 apresenta uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 27).

TABELA 27

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
NOvM (<i>Streptomyces spheroides</i>)	AAF67506	184	Glicosiltransferase
Provável glicosil transferase (<i>Streptomyces violaceoruber</i>)	T46519	169	Glicosiltransferase
Glicosil transferase homóloga (<i>Streptomyces cyanogenus</i>)	AAD13553	167	Glicosiltransferase
Glicosil transferase homóloga (<i>Streptomyces cyanogenus</i>)	AAD13555	163	Glicosiltransferase
dNTP-hexose glicosil transferase (<i>Amycolatopsis mediterranei</i>)	AAC01731	160	Glicosiltransferase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

O gene orf19 codifica uma proteína que apresenta uma relativa forte similitude com várias ceto-redutases. Notadamente, a proteína codificada por orf19 possui uma similitude importante com um NDP-hexose 4-ceto-redutase (TylCIV) de *streptomyces fradiae* (Bate et al., 2000; Número de acesso GenBank: AAD41822; Escore BLAST: 266). A similitude da proteína codificada pelo gene orf19 com essa ceto-redutase implicada na via de biossíntese de um antibiótico próximo sugere fortemente que esse gene codifique também uma 4-ceto-redutase responsável pela reação de redução necessária à biossíntese da micarose (conforme Figura 4). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf19 apresenta uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 28).

TABELA 28

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
NDP-4-ceto-6-deóxi-hexose 4-ceto-redutase (Streptomyces venezuelae)	AAL14256	251	NDP-4-ceto-6-deóxi-hexose 4-ceto-redutase
EryBIV(Saccharopolyspora erythraea)	AAB84071	249	Óxido redutase
TDP-4-ceto-6-deóxi-hexose 4-ceto-redutase (Micromonospora megalomicea subsp. Nigra)	AAG13916	218	dTDP-4-ceto-6-deóxi-hexose 4-ceto-redutase
dTDP-4-ceto-6-deóxi-L-hexose 4-redutase (Streptomyces avermitilis)	BAA84595	212	dTDP-4-ceto-6-deóxi-L-hexose 4-redutase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

- O gene orf20 codifica uma proteína que apresenta uma relativa forte similitude com várias hexose redutases. Notadamente, a proteína codificada por orf20 possui uma similitude importante como gene EryBII da Saccharopolyspora erythraea que codifica uma dTDP-4-ceto-L-6-deóxi-hexose 2,3-redutase (Summers, R.G., et al. , 1997), Número de acesso GenBank: AAB84068; Escore BLAST: 491). A similitude da proteína codificada pelo gene orf20 com várias hexose redutases implicadas na via de biossíntese de outros antibióticos próximos sugere fortemente que esse gene codifica uma 2-3 redutase responsável pela reação de redução necessária à biossíntese da micarose (conforme Figura 4). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf20 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 29).

TABELA 29

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
TyICII (<i>Streptomyces fradiae</i>)	AAD41821	464	NDP-hexose 2,3 enoil redutase
TDP-4-ceto-6-deóxi-hexose 2,3-redutase (<i>Micromonospora megalomicea</i> subsp. <i>Nigra</i>)	AAG13914	446	TDP-4-ceto-6-deóxi-hexose 2,3-ceto-redutase
dTDP-4-ceto-6-deóxi-L-hexose 2,3-redutase (<i>Streptomyces avermitilis</i>)	BAA84599	377	dTDP-4-ceto-6-deóxi-L-hexose 2,3- redutase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

O gene orf21c codifica uma proteína que apresenta uma relativa forte similitude com várias hexose metiltransferases. Notadamente, a proteína codificada por orf21c possui uma similitude importante como gene TyICIII de *streptomyces fradiae* que codifica uma NDP-hexose 3-C-metiltransferase (Bate,N et al., 2000; Número de acesso GenBank: AAD41823; Escore BLAST: 669). A similitude da proteína codificada pelo gene orf21c com várias hexose metiltransferases implicadas na via de biossíntese de outros antibióticos próximos sugere fortemente que esse gene codifique uma hexose metiltransferase responsável pela reação de metilação necessária à biossíntese da micarose (conforme Figura 4). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf21c apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 30).

TABELA 30

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
EryH (<i>Saccharopolyspora erythraea</i>)	228448	592	Gene de biossíntese da eritromicina
S-adenosil-dependente metil transferase (<i>Coxiella burnetii</i>)	AAK71270	358	Metiltransferase
NovU (<i>Streptomyces sphaeroides</i>)	AAF67514	184	C-metiltransferase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

O gene orf22c codifica uma proteína que apresenta uma relativa forte similitude com a proteína codificada pelo gene fkbH de *streptomyces hygroscopicus* var. *ascomyceticus* que codifica uma enzima implicada na biossíntese da metoximalonila (Wu, K. et al., 2000; Número de acesso GenBank: AAF86387; Escore BLAST: 463). A similitude da proteína codificada pelo gene orf22c com essa proteína implicada na via de biossíntese de um outro macrolídeo próximo sugere fortemente que esse gene codifique também uma enzima implicada na biossíntese da metoximalonila nas *streptomyces ambofaciens* (conforme Figura 8).

O gene orf23c codifica uma proteína que apresenta uma relativa forte similitude com a proteína codificada pelo gene fkbI de *streptomyces hygroscopicus* var. *ascomyceticus* que codifica um acil CoA desidrogenase implicada na biossíntese da metoximalonila (Wu, K. et al., 2000; Número de acesso GenBank: AAF86388; Escore BLAST: 387). A similitude da proteína codificada pelo gene orf23c com várias acil CoA desidrogenases implicadas na via de biossíntese de outros antibióticos próximos sugere fortemente que esse gene codifique uma acil CoA desidrogenase implicada na biossíntese da metoximalonila (conforme Figura 8). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf23c apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 31).

TABELA 31

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Acil-CoA desidrogenase (Polysangium cellulosum)	AAK19892	171	Acil-CoA desidrogenase
Provável acil-CoA desidrogenase - (Streptomyces coelicolor)	T36802	160	Acil-CoA desidrogenase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990)

O gene orf24c codifica uma proteína que apresenta uma relativa forte similitude com a proteína codificada pelo gene fkbJ de *streptomyces hygroscopicus* var. *ascomyceticus* que codifica a proteína de ligação do grupo acila (Acil Carrier Protein (ACP)) implicada na biossíntese da metoximalonila (Wu, K. et al., 2000; Número de acesso GenBank: AAF86389; Escore BLAST: 87). Similitude da proteína codificada pelo gene orf24c com essa proteína implicada na via de biossíntese de um outro macrolídeo próximo sugere fortemente que esse gene codifique uma proteína implicada na biossíntese da metoximalonila nas *streptomyces ambofaciens* (conforme Figura 8).

O gene orf25c codifica uma proteína que apresenta uma relativa forte similitude com a proteína codificada pelo gene fkbK de *streptomyces hygroscopicus* var. *ascomyceticus* que codifica uma acil CoA desidrogenase implicada na biossíntese da metoximalonila (Wu, K. et al., 2000; Número de acesso GenBank: AAF86390; Escore BLAST: 268). A similitude da proteína codificada pelo gene orf25c com vários acil CoA desidrogenases implicadas na via de biossíntese de outros antibióticos próximos sugere fortemente que esse gene codifica uma acil CoA desidrogenase implicada na biossíntese da metoximalonila (conforme Figura 8). Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf25c apresenta uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 32).

TABELA 32

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Provável 3-hidroxi-butiril-CoA Desidrogenase (<i>Bacillus subtilis</i>)	P45856	177	3-hidroxi-butiril-CoA Desidrogenase
Proteína 3-hidroxi-butiril-CoA desidrogenase (<i>Bacillus thuringiensis</i> serovar <i>kurstaki</i>)	AAL32270	174	3-hidroxi-butiril-CoA Desidrogenase
3-hidroxi-butiril-CoA desidrogenase (<i>Deinococcus radiodurans</i>)	NP-294792	167	3-hidroxi-butiril-CoA Desidrogenase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990)

O gene orf26 codifica uma proteína que apresenta uma identidade de 65% (determinada graças ao programa BLAST) com a proteína codificada pelo gene *tylCV* que codifica uma micarosil transferase implicada na biossíntese de tilosina nas *streptomyces fradiae* (Bate, N. et al., 2000; Número de acesso GenBank: AAD41824, Escore BLAST: 471). Mais particularmente, *TylCV* é uma glicosil-transferase que liga a molécula de micarose, quando da síntese da tilosina. Essa similitude com uma proteína na via de biossíntese de um outro antibiótico relativamente próximo e mais particularmente na transferência da micarose, sugere que o gene orf26 codifica uma glicosil transferase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf26 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 33).

TABELA 33

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Glicosil transferase (<i>Streptomyces avermitilis</i>)	BAA84592	218	Glicosil transferase
CalG4 (<i>Micromonospora echinospora</i>)	AAM70365	217	Glicosil transferase
CalG2 (<i>Micromonospora echinospora</i>)	AAM70348	197	Glicosil transferase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

O gene orf27 codifica uma proteína que apresenta uma identidade de 70% (determinada graças ao programa BLAST) com a proteína codificada pelo gene tylCVII que codifica uma NDP-hexose 3,5- (ou 5-)epimerase implicada na biossíntese de tilosina nas *streptomyces fradiae* (Bate,N. et al., 2000; Número de acesso GenBank: AAD41825, Escore BLAST: 243). Mais particularmente, TylCVII é uma hexose 3,5- (ou 5-) epimerase implicada na biossíntese da micarose. Essa similitude com uma proteína implicada na via de biossíntese da micarose, sugere que o gene orf27 codifica uma epimerase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf27 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 34). Uma análise das sequências próximas obtidas graças ao programa BLAST sugere fortemente que o gene orf27 codifique uma 5 epimerase responsável da reação de epimerização necessária à biossíntese da micarose (conforme Figura 4).

TABELA 34

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
LanZ1 (<i>Streptomyces cyano-genus</i>)	AAS13558	172	NDP-hexose 3,5 epimerase
Epi (<i>Saccharopolyspora spinosa</i>)	AAK83288	169	TDP-4-ceto-6-deoxiglicose 3,5 epimerase
dNTP-hexose 3,5 epimerase (<i>Amycolatopsis mediterranei</i>)	AAC01732	166	dNTP-hexose 3,5-epimerase

* maior similitude de sequência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

- A sequência de orf28c foi em uma primeira etapa determinada de maneira parcial, já que a sequência de uma região de aproximadamente 450 pares de bases só foi determinada após nova seqüenciação (essa região é simbolizada por "N" na sequência incompleta SEQ ID N° 106). A sequência parcial dessa ORF (SEQ ID N° 111) foi todavia utilizada para a análise com os diferentes programas informáticos conforme explicado acima.
- Assim, pôde ser determinado que o gene orf28c codifique uma proteína que apresenta uma identidade de 64% sobre a sequência determinada (SEQ ID N° 112, que é a sequência parcial da proteína Orf28c) (determinada graças ao programa BLAST) com a proteína codificada pelo gene *acyB2* que codifica uma proteína reguladora implicada na biossíntese de carbomicina nas *streptomyces thermotolerans* (Arisawa,A., et al. 1993; Número de acesso GenBank: JC2032, Escore BLAST: 329). Essa similitude com uma proteína implicada na via de biossíntese de um antibiótico relativamente próximo sugere que o gene orf28c codifique uma proteína reguladora implicada próxima sugere que o gene orf28c codifique uma proteína reguladora implicada na biossíntese das espiamicinas. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf28c apresentar também uma forte similitude com a proteína TylR que é uma proteína reguladora implicada na biossíntese de tilosina nas *streptomyces fradiae* (Bate,N. et al., 1999); Número de aces-

so GenBank: AAF29380, Escore BLAST; 167).

O gene orf28c pôde ser amplificado graças a oligonucleotídeos situados de ambos os lados da seqüência não determinada e subclonado em um vetor de expressão. Pôde assim ser demonstrado que a superexpressão do gene orf28c aumenta significativamente a produção em espiramicinas da cepa OSC2 (conforme exemplo 24). Isto demonstra que a superexpressão de orf 28c leva a um aumento da produção em espiramicinas e confirma seu papel como regulador da via de biossíntese das espiramicinas.

A seqüência parcial de orf28c foi completada na seqüência e a região que falta de aproximadamente 450 pares de bases foi determinada (conforme SEQ ID N° 140 e SEQ ID N° 141). A seqüência completa dessa ORF (SEQ ID N° 141) foi utilizada para a análise com os diferentes programas informáticos conforme explicado acima. Pôde assim ser determinado que o gene orf28c codifique uma proteína que apresenta uma identidade de 69% na seqüência determinada (SEQ ID N° 142, que é a seqüência completa da proteína Orf28c) (determinada graças ao programa BLAST) com a proteína codificada pelo gene *acyB2* que codifica uma proteína reguladora implicada na biossíntese de carbomicina nas *streptomyces* *thermotolerans* (Arsawa, A., et al., 1993; Número de acesso GenBank: JC2032, Escore BLAST: 451). Essa similitude com uma proteína implicada na regulação da biossíntese de um antibiótico relativamente próxima sugere que o gene orf28c codifique uma proteína reguladora implicada na biossíntese das espiramicinas. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf28c apresenta também uma forte similitude com a proteína TyIR que é uma proteína reguladora implicada na regulação da biossíntese de tilosina nas *streptomyces* *fradiae* (Bate, N. et al., 1999; Número de acesso GenBank: AAF29380, Escore BLAST: 224). Os resultados da superexpressão desse gene (conforme exemplo 24) confirmam seu papel como regulador da via de biossíntese das espiramicinas.

A inativação do gene orf28c foi feita. Pôde assim ser mostrado que a cepa resultante não produz mais espiramicinas. Isto confirma que o gene orf28c está bem implicado na biossíntese das espiramicinas e é es-

sencial à biossíntese das espiramicinas. Esses resultados associados aos resultados da superexpressão desse gene (Conforme exemplo 24) estão em acordo com um papel de ativador essencial à biossíntese das espiramicinas de Orf28c.

- 5 O gene orf29 codifica uma proteína que apresenta uma identidade de 31% (determinada graças ao programa BLAST) com uma provável glicosil hidrolase localizada no grupo de genes implicado na biossíntese de sorafen A (um antifúngico da classe dos poliquetídeos) no *Sorangium cellulosum* (Ligon, J. et al., 2002; Número de acesso GenBank: AAK19890, Escore BLAST: 139). Essa similitude com uma proteína implicada na via de biossíntese de uma molécula relativamente próxima sugere que o gene orf29 codifique uma proteína que tem uma atividade glicosil hidrolase. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf29 apresentar uma forte similitude com outras proteínas de função similar em outros
- 10
- 15 organismos (conforme tabela 35). Uma análise da sequência da proteína codificada por orf29, graças ao programa CD-search (conforme acima) sugere também que o gene orf29 codifique uma glicosil hidrolase.

TABELA 35

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
ManA (<i>Caldicellulosiruptor saccharolyticus</i>)	AAC44232	136	Beta-1,4-mannanase
ManA (<i>Dictyoglomus thermophilum</i>)	AAB82454	129	Beta-mannanase

- * maior similitude de sequência é associada a um escore BLAST mais elevado (Altschul et al., 1990).
- 20

- A análise da sequência protéica deduzida do orf29 pelo programa SignalP (<http://www.cbs.dtu.dk/services/SignalP/>) (Nielsen, H. et al., 1997) mostra que essa proteína possui uma sequência sinal C-terminal com um sítio de corte predito entre as posições 30 e 31 (QSA/QA). Pode-se pre-
- 25 dizer que essa proteína é extracelular. Ela poderia, enquanto glicosil-

hidrolase, ter um papel na reativação da espiramicina inativada por glicosilação pelas glicosil-transferases GimA e/ou GimB (Gourmelen et al., 1998).

O gene orf30c codifica uma proteína que apresenta uma identidade de 31% (determinada graças ao programa BLAST) com uma epimerase de açúcar-nucleosídeo-difosfato de *Corynebacterium glutamicum* (Número de acesso GenBank: NP-600590, Escore BLAST: 89). Essa similitude sugere que o gene orf30c codifica uma epimerase. Essa hipótese é confirmada pelo fato de uma análise da sequência graças ao programa CD-search (conforme acima) sugere também que o gene orf30c codifique uma epimerase.

O orf30c apresenta dois códons de iniciações possíveis (conforme SEQ ID N° 115) que dão duas proteínas possíveis de 345 e 282 aminoácidos respectivamente (SEQ ID N° 116 e 117). Todavia, o uso de códons é típico de *streptomyces* somente a partir do segundo ATG, além disso, a sequência protéica deduzida da sequência entre o primeiro ATG e o segundo não se alinha com as sequências próximas identificadas, enquanto que a sequência protéica a mais curta (a partir do segundo ATG: SEQ ID N° 144) se alinha bem com o começo dessas proteínas. Pode-se, portanto, deduzir daí que o segundo ATG seja o bom códon de iniciação e que a sequência desse orf seja, portanto, aquele apresentado em SEQ ID N° 143 que uma vez traduzido corresponde à proteína de sequência SEQ ID N° 144.

O gene orf31 codifica uma proteína que apresenta uma identidade de 52% (determinada graças ao programa BLAST) com uma óxido-redutase nas *streptomyces coelicolor* (Número de acesso GenBank: NP_631148, Escore BLAST: 261). Essa similitude sugere que o gene orf31 codifique uma redutase. Essa hipótese é confirmada pelo fato de uma análise da sequência, graças ao programa CD-search (conforme acima) sugerir também que o gene orf31 codifique uma redutase. Essa hipótese é também confirmada pelo fato de a proteína codificada pelo gene orf31 apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 36).

TABELA 36

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Oxido-redutase putativa (<i>Streptomyces griseus</i>)	BAB79295	173	Óxido-redutase
MocA (<i>Xanthomonas axonopodis</i>)	NP-640644	171	Óxido-redutase

* maior similitude de seqüência é associada a um Escore BLAST mais elevado (Altschul et al., 1990).

- 5 A inativação do gene orf3 foi realizada. Pôde assim ser mostrado que a cepa resultante não produz espiramicinas. Isto confirma que o gene orf31 é bem implicado na biossíntese das espiramicinas. A enzima codificada por esse gene é, portanto, responsável de uma etapa de bioconversão essencial à biossíntese das espiramicinas.

- 10 A seqüência de orf32c foi inicialmente determinada de maneira parcial (cf exemplo 19), já que a seqüência codificante em 5' foi determinada apenas em uma segunda etapa. A seqüência parcial dessa orf (SEQ ID N° 120) foi todavia utilizada para análise com os diferentes programas informáticos conforme explicado acima. Pôde assim ser determinado que o gene orf32c codifica uma proteína que apresenta uma identidade de 47% sobre a
- 15 seqüência determinada (SEQ ID N° 121, que é a seqüência parcial da proteína Orf32c) (determinada graças ao programa BLAST) com uma proteína reguladora da família GntR nas *streptomyces coelicolor* (Número de acesso GenBank: NP_625576, Escore BLAST: 229). Essa similitude sugere que o gene orf32c codifique um regulador transcricional da família GntR. Essa
- 20 hipótese é confirmada pelo fato de a proteína codificada pelo gene orf32c apresentar uma forte similitude com outras proteínas de função similar em outros organismos.

- A seqüência parcial de orf32c foi completada pela seqüência e a região que falta foi determinada (conforme SEQ ID N° 140 e SEQ ID N° 145).
- 25 A seqüência completa dessa orf codifica uma proteína que apresenta uma

- identidade de 44% (determinada graças ao programa BLAST) com uma proteína reguladora da família GntR nas *streptomyces avermitilis* (Número de acesso GenBank: NP_824604, Escore BLAST: 282). Essa similitude sugere que o gene orf32c codifica um regulador transcricional da família GntR. Essa hipótese é confirmada pelo fato de a proteína codificada pelo gene orf32c apresentar uma forte similitude com outras proteínas de função similar em outros organismos (conforme tabela 37).

TABELA 37

Proteína que apresenta uma similitude significativa	Número de acesso GenBank	Escore BLAST*	Função complementar
Proteína reguladora da família GntR (<i>Streptomyces avermitilis</i>)	NP_828241	270	Proteína reguladora da família GntR
Proteína reguladora da família GntR (<i>Streptomyces coelicolor</i>)	NP_625576	266	Proteína reguladora da família GntR
SC5G8.04 (<i>Streptomyces coelicolor</i>)	NP_628991	258	Proteína reguladora da família GntR
Regulador transcricional (<i>Streptomyces venezuelae</i>)	AAF01064	224	Regulador transcricional
Proteína reguladora da família GntR (<i>Streptomyces avermitilis</i>)	NP_827432	239	Proteína reguladora da família GntR

- * maior similitude de sequência está associada a um Escore BLAST mais elevado (Altschul et al., 1990).

- A inativação do gene orf32c foi feita com a finalidade de estudar a função desse gene na via de biossíntese das espiramicinas em *streptomyces ambofaciens*. Pôde ser mostrado que a cepa resultante produz espiramicinas. Isto indica que o gene orf32c não é essencial à biossíntese das espiramicinas e que não é essencial à sobrevivência da bactéria.

- O gene orf33 codifica uma proteína que apresenta uma identidade de 49% (determinada graças ao programa BLAST) com uma proteína hipotética de *Xanthomonas campestris* (Número de acesso GenBank: NP_635564, Escore BLAST: 54).

- A sequência de orf34c é parcial. Com efeito, as comparações

feitas entre o produto dessa orf e os bancos de dados sugerem que a parte C-terminal dessa proteína não está no produto deduzido da sequência nucleotídica e, portanto, que essa orf seria mais longa e continuaria fora da região seqüenciada. A seqüência parcial dessa ORF foi, todavia, utilizada para a análise com os diferentes programas informáticos conforme explicado acima. Pôde ser determinado que o gene orf34c codifica uma proteína que apresenta uma identidade de 91% sobre a seqüência determinada (SEQ ID N° 150, que é a seqüência parcial da proteína Orf34c) (determinada graças ao programa BLAST) com uma arabinofuranosidase em *streptomyces coelicolor* (Bentley et al., 2002; Número de acesso GenBank: NP_630049, Escore BLAST: 654). Em *S. coelicolor* o gene que codifica essa arabinofuranosidase não parece implicado na biossíntese de metabólito secundário. Em *S. ambofaciens*, esse gene não é, portanto, provavelmente implicado na biossíntese de espiramicina.

A presente invenção tem também por objeto dos polinucleotídeos que se hibridam em condições de hibridização de forte estringência a pelo menos um dos polinucleotídeos de seqüência SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149 ou uma de suas variantes ou uma das seqüências derivadas desta, em virtude da degenerescência do código genético. Preferencialmente, esses polinucleotídeos são isolados a partir de uma bactéria do gênero *streptomyces*, mais preferencialmente, esses polinucleotídeos codificam proteínas implicadas na biossíntese de um macrolídeo e ainda mais preferencialmente esses polinucleotídeos codificam uma proteína que tem uma atividade similar à proteína codificada pelos polinucleotídeos com os quais eles se hibridam. As condições de hibridização de forte estringência podem ser definidas como condições de hibridização, desfavorecendo a hibridização de filamentos de ácidos nucléicos não homólogos. Condições de hibridizações de forte estringência podem ser, por exemplo, descritas como condições de hibridização no tampão descrito por Church & Gilbert (Church & Gilbert, 1984) a uma temperatura compreendida entre 55°C e 65°C, de

maneira preferida a temperatura de hibridização é de 55°C, de maneira ainda mais preferida a temperatura de hibridização é de 60°C e de maneira inteiramente preferida a temperatura de hibridização é de 65°C, seguida de uma ou várias lavagens feitas em tampão 2X SSC a uma temperatura compreendida entre 55°C e 65°C, de maneira preferida essa temperatura é de 55°C, de maneira ainda mais preferida essa temperatura é de 60°C e de maneira inteiramente preferida essa temperatura é de 65°C, seguida de uma ou várias lavagens em tampão 0.5X SSC a uma temperatura compreendida entre 55°C e 65°C, de maneira preferida essa temperatura é de 55°C, de maneira ainda mais preferida essa temperatura é de 60°C e de maneira inteiramente preferida essa temperatura é de 65°C. As condições de hibridização acima descritas podem ser adaptadas em função da extensão do ácido nucléico, cuja hibridização é buscada ou do tipo de marcação escolhido, segundo técnicas conhecidas do técnico. As condições convenientes de hibridização podem, por exemplo, ser adaptadas, segundo a obra de F. Ausubel et al., 2002.

A invenção refere-se também a um polinucleotídeo que tem pelo menos 70%, de forma mais preferida 80%, de forma mais preferida 85%, de forma ainda mais preferida 90%, de forma ainda mais preferida 95%, e de maneira inteiramente preferida 98% de identidade em nucleotídeos com um polinucleotídeo que compreende pelo menos 10, 12, 15, 18, 20 a 25, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850 ou 1900 nucleotídeos consecutivos de um polinucleotídeo escolhido no grupo constituído de seqüências nucleotídicas SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149 ou de uma de suas variantes ou uma das seqüências derivadas destas em razão da degenerescência do código genético, ou um polinucleotídeo de seqüência complementar. Preferencialmente, esses polinucleotídeos são isolados a partir de uma bactéria do gênero *streptomyces*, mais

preferencialmente, esses polinucleotídeos codificam proteínas implicadas na biossíntese de um macrolídeo e ainda mais preferencialmente esses polinucleotídeos codificam proteínas que têm atividades similares às proteínas codificadas pelos polinucleotídeos com os quais eles apresentam a identidade.

- 5 De forma inteiramente preferida, um polinucleotídeo, de acordo com a invenção, é escolhido no grupo constituído das seqüências nucleotídicas SEQ ID Nº 3, 5, 7, 9, 11, 13, 15,

17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147

e 149 ou um polinucleotídeo de seqüência complementar.

- 10 O alinhamento ótimo das seqüências para a comparação pode ser realizado de maneira informática com o auxílio de algoritmos conhecidos, por exemplo, os do pacote FASTA (Pearson W. R. & D. J. Lipman, 1988) e (Pearson W. R., 1990), acessível notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França. A título de ilustração, a percentagem de identidade de seqüência poderá ser determinada com o auxílio do software
- 15 LFASTA (Chao K. -M. et al., 1992) ou LALIGN (Huang X. e Miller W., 1991). Os programas LFASTA e LALIGN fazem parte do pacote FASTA. LALIGN volta aos alinhamentos sítios ótimos: esse programa é mais rigoroso, mas também mais lento do que LFASTA.

- 20 Um outro aspecto da invenção refere-se a um polipeptídeo resultante da expressão de uma seqüência de ácido nucléico, tal como definida acima. Preferencialmente, os polipeptídeos, de acordo com a invenção, apresentam pelo menos 70%, de forma mais preferida 80%, de forma mais preferida 85%, de forma ainda mais preferida 90%, de forma ainda mais preferida 95% e de maneira inteiramente preferida 98% de identidade em
- 25 aminoácidos com um polipeptídeo, compreendendo pelo menos 10, 15, 20, 30 a 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 480, 500, 520, 540, 560, 580, 600, 620 08 640 aminoácidos consecutivos de um polipeptídeo escolhido
- 30 dentre

SEQ ID N° 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 27, 29, 31, 32, 33, 35, 37, 38, 39, 41, 42, 44, 46, 48, 50, 51, 52, 54, 55, 56, 57, 58, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 108, 110, 112, 114, 116, 117, 119, 121, 142, 144, 146, 148

e 150 ou uma dessas seqüências exceto ao longo dessa seqüência um ou vários aminoácidos foram substituídos, inseridos ou deletados, sem afetar as propriedades funcionais ou uma das variantes dessas seqüências. Preferencialmente, os polipeptídeos, de acordo com a invenção, são expressos no estado natural por uma bactéria do gênero *streptomyces*, mais preferencialmente, esses polipeptídeos são implicados na biossíntese de um macrolídeo e ainda mais preferencialmente esses polipeptídeos têm uma atividade similar àquela do polipeptídeo como qual ele divide a identidade. De forma preferida, um polipeptídeo, de acordo com a invenção, é escolhido no grupo constituído das seqüências polipeptídicas

SEQ ID N° 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 27, 29, 31, 32, 33, 35, 37, 38, 39, 41, 42, 44, 46, 48, 50, 51, 52, 54, 55, 56, 57, 58, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 108, 110, 112, 114, 116, 117, 119, 121, 142, 144, 146, 148 e 150 ou uma dessas seqüências exceto que ao longo dessa seqüência um ou vários aminoácidos foram substituídos, inseridos ou deletados sem afetar as propriedades funcionais ou umas variantes dessas seqüências. De forma inteiramente preferida, um polipeptídeo, de acordo com a invenção, é escolhido no grupo constituído das seqüências polipeptídicas

SEQ ID N° 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 27, 29, 31, 32, 33, 35, 37, 38, 39, 41, 42, 44, 46, 48, 50, 51, 52, 54, 55, 56, 57, 58, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 108, 110, 112, 114, 116, 117, 119, 121, 142, 144, 146, 148 e 150.

O alinhamento ótimo das seqüências para a comparação pode ser realizado de maneira informática como auxílio de algoritmos conhecidos, por exemplo aqueles do pacote FASTA (Pearson W.R. & D.J. Lipman, 1988) e (Pearson W.R., 1990), acessível notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França. A título de ilustração, a percentagem de identidade de seqüência poderá ser determinada com o auxílio do software LFASTA (Chao K.M. et al., 1992) ou LALIGN (Huang X. e Miller W., 1991), utilizando os parâmetros por defeito, tal como definido pelo centro de pes-

quisas INFOBIOGEN, Evry, França. Os programas LFASTA E LALIGN fazem parte do *pacote* FASTA. LALIGN retorna os alinhamentos sítios ótimos: esse programa é mais rigoroso, mas também mais lento do que LFASTA.

Um outro aspecto da invenção refere-se a um DNA recombinante que compreende pelo menos um polinucleotídeo, tal como descrito acima. Preferencialmente, esse DNA recombinante é um vetor. Ainda mais preferencialmente, o vetor é escolhido dentre os bacteriófagos, os plasmídeos, os fagemídeos, os vetores integrativos, os foscídeos, os cosmídeos, os vetores navettes, os BAC (Bacterial Artificial Chromosome) ou os PAC (P1-derived Artificial Chromosome). A título ilustrativo, podem-se citar como bacteriófagos os fagos lambda e M13. Como plasmídeos, poder-se citar os plasmídeos que se replicam em *E.Coli* por exemplo pBR322 e seus derivados, pUC18 e seus derivados, pUC19 e seu derivados, pGB2 e seus derivados (Churchward G. et al., 1984), pACYC177 (número de acesso GenBank: X06402) e seus derivados, paCYC184 (número de acesso GenBank: X06403) e seus derivados. Pode-se também citar os plasmídeos que se replicam em *streptomyces* como, por exemplo, pIJ101 e seus derivados, pSG5 e seus derivados, SLP1 e seus derivados, SCP2* e seus derivados (Kieser et al. 2000). Como fagenídeos, podem-se citar a título ilustrativo pBluescript II e seus derivados (comercializados notadamente pela sociedade Stratagene (LaJolla, Califórnia, USA)), pGEM-T e seus derivados (comercializados pela sociedade PROMEGA (Madison Wisconsin, USA)), λ ZAPII e seu derivados (comercializados notadamente pela sociedade Stratagene (LaJolla, Califórnia, USA)). Como vetores integrativos, podem-se citar a título ilustrativo, os vetores integrativos nas *streptomyces* como, por exemplo, aqueles derivados de SLP1 (Kieser et al, 2000), aqueles derivados de pSAM2 (Kieser et al., 2000), os vetores que utilizam os sistemas de integração dos fagos PhiC31 (Kieser et al., 2000) (Por exemplo pSET152 (Bierman et al., 1992) ou de VWB (Van Mellaert L, et al., 1998), assim como os vetores que utilizam o sistema de integração de IS117 (Kieser et al. 2000). Como foscídeos, podem-se citar, a título ilustrativo, o foscídeo pFOS1 (comercializado pela sociedade New England Biolabs Inc., Beverly, Massachussets, USA) e seu de-

rivados. Como cosmídeos, podem-se citar a título ilustrativo o cosmídeo SuperCos e seus derivados (comercializados notadamente pela sociedade Stratagene (LaJolla, Califórnia, USA)), o cosmídeo pWED15 (Whahl et al., 1987) e seu derivados. Como vetores navettes, podem-se citar, a título ilustrativo, os plasmídeos navettes *E.Coli/ streptomyces* como, por exemplo plJ903 e seu derivados, a série dos plasmídeos pUWL, pCAO106, pWHM3, pOJ446 e seus derivados (Kieser et al., 2000), os BAC navettes *E.coli / streptomyces* como, por exemplo, aqueles descritos no pedido de patente WO 01/40497. Como BAC (Bacterial Artificial Chromosome) pode-se citar a título ilustrativo o BAC pBeloBAC11 (número de acesso GenBank: U51113). Como PAC (P1-derived Artificial Chromosome), pode-se citar a título ilustrativo o vetor pCYPAC6 (número de acesso GenBank: AF133437). De forma inteiramente preferida, um vetor, de acordo com a invenção, é escolhido dentre

pOS49.1, pOS49.11, pOSC49.12, pOS49.14, pOS49.16, pOS49.28, pOS44.1, pOS44.2, pOS44.4, pSPM5, pSPM7, pOS49.67, pOS49.88, pOS49.106, pOS49.120, pOS49.107, pOS49.32, pOS49.43, pOS49.44, pOS49.50, pOS49.99, pSPM17, pSPM21, pSPM502, pSPM504, pSPM507, pSPM508, pSPM509, pSPM1, pBXL1111, pBXL1112, pBXL1113, pSPM520, pSPM521, pSPM522, pSPM523, pSPM524, pSPM525, pSPM527, pSPM528, pSPM34, pSPM35, pSPM36, pSPM37, pSPM38, pSPM39, pSPM40, pSPM41, pSPM42, pSPM43, pSPM44, pSPM45, pSPM47, pSPM48, pSPM50, pSPM51, pSPM52, pSPM53, pSPM55, pSPM56, pSPM58, pSPM72, pSPM73, pSPM515, pSPM519, pSPM74, pSPM75, pSPM79, pSPM83, pSPM107, pSPM543 e pSPM106.

Um outro aspecto da invenção refere-se a um sistema de expressão que compreende um vetor de expressão apropriado e uma célula hospedeira que permite a expressão de um ou vários polipeptídeos, tais como descritos acima em um sistema biológico. Os vetores de expressão, de acordo com a invenção, compreendem uma sequência de ácido nucléico codificando um ou vários polipeptídeos, tais como descritos acima e podem ser destinados à expressão dos diferentes polipeptídeos, de acordo com a invenção em diversas células hospedeiras bem conhecidas do técnico. A

título de exemplo, podem-se citar os sistemas de expressão procariota, como os sistemas de expressão na bactéria *E. coli*, os sistemas de expressão eucariotas, como os sistemas de expressão baculovírus, que permitem uma expressão em células de insetos, os sistemas de expressão que permitem uma expressão em células de levedos, os sistemas de expressão que permitem uma expressão em células de mamíferos, notadamente células humanas. Os vetores de expressão utilizáveis nesses sistemas são bem conhecidos do técnico, no que refere-se às células procariotas, podem-se citar, a título ilustrativo, os vetores de expressão em *E. coli*, por exemplo, da família pET, comercializados pela sociedade Stratagene (LaJolla, Califórnia, USA), os vetores da família GATEWAY, comercializados pela sociedade Invitrogen (Carlsbad, Califórnia, USA), os vetores da família pBAD, comercializados pela sociedade Invitrogen (Carlsbad, Califórnia, USA), os vetores da família pMAL, comercializados pela sociedade New England Biolabs Inc. (Beverly, Massachusetts, USA), os vetores de expressão indutíveis pela ramnose citados na publicação Wilms B et al., 2001 e seus derivados, podem-se também citar os vetores de expressão nos *streptomyces*, como, por exemplo os vetores pIJ4123, pIJ6021, pPM927, pANT849, pANT850, pANT851, pANT1200, pANT1201, pANT1202 e seus derivados (Kieser et al., 2000). No que refere-se às células de levedos, pode-se citar a título ilustrativo o vetor pESC comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA). No que refere-se ao sistema de expressão baculovírus, permitindo a expressão em células de insetos, pode-se citar a título ilustrativo o vetor BacPAK6 comercializado pela sociedade BD Biosciences Clontech, (Palo Alto, Califórnia USA). No que refere-se às células de mamíferos, podem-se citar a título ilustrativo vetores que compreendem o promotor dos genes precoces do vírus CMV (citomegalovírus) (por exemplo, o vetor pCMV e seus derivados comercializados pela sociedade Stratagene (LaJolla, Califórnia, USA) ou o promotor dos genes imediatos (promotor imediato SV40) do vírus simiano vacuolisante SV40 (por exemplo, o vetor pSG5 comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA).

A invenção é também relativa a um processo de produção de um

polipeptídeo, tal como descrito acima, esse processo compreendendo as seguintes etapas:

- a) inserção de um ácido nucléico codificando esse polipeptídeo em um vetor apropriado;
- 5 b) cultivo, em um meio de cultura apropriado, de uma célula hospedeira previamente transformada ou transfeita com o vetor da etapa a);
- c) recuperação do meio de cultura condicionado ou um extrato celular, por exemplo, por sonicação ou por choque osmótico;
- d) separação e purificação, a partir desse meio de cultura, ou
10 ainda a partir do extrato celular obtido na etapa c), desse polipeptídeo;
- e) se for o caso, caracterização do polipeptídeo recombinante produzido.

Um polipeptídeo recombinante, de acordo com a invenção, pode ser purificado por passagem sobre uma série apropriada de colunas de cromatografia, segundo os métodos conhecidos do técnico, e descritos, por
15 exemplo, em F. Ausubel et al.(2002). Pode-se citar, a título ilustrativo, a técnica do "Tag-Histidine", que consiste em acrescentar uma curta seqüência polihistadina ao polipeptídeo a ser produzido, este podendo, em seguida, ser purificado sobre coluna de níquel. Um polipeptídeo, de acordo com a inven-
20 ção, pode ser igualmente preparado pelas técnicas de síntese *in vitro*. A título ilustrativo dessas técnicas, um polipeptídeo, de acordo com a invenção, poderá ser preparado graças ao sistema "rapid translation system (RTS)", comercializado notadamente pela sociedade Roche Diagnostics France S.A., Meyland, França.

25 Um outro aspecto da invenção refere-se a células hospedeiras, nas quais foi introduzido pelo menos um polinucleotídeo e/ou pelo menos um DNA recombinante e/ou pelo menos um vetor de expressão, de acordo com a invenção.

Um outro aspecto da invenção refere-se a microorganismos blo-
30 queados em uma etapa da via de biossíntese de pelo menos um macrolídeo. O interesse reside, por um lado, no estudo da função das proteínas mutadas e, por outro lado, na realização de microorganismos que produzem interme-

diários de biossíntese. Esses intermediários podem ser modificados, eventualmente após separação, seja por acréscimos de componentes particulares nos meios de produção, seja por introdução nos microorganismos assim mutados de outros genes codificando proteínas capazes de modificar o intermediário que serve, aí, como substrato. Esses intermediários podem assim ser modificados por via química, bioquímica, enzimática e/ou microbiológica. Os microorganismos bloqueados em uma etapa da via de biossíntese de macrolídeos podem ser obtidos ativando-se a função de uma ou várias proteínas implicadas na biossíntese desse(s) macrolídeo(s) em microorganismos produtores desse(s) macrolídeo(s). Conforme as proteínas inativadas, pode-se ainda obter microorganismos bloqueados em diferentes etapas da via de biossíntese dos macrolídeos. A inativação dessa(s) proteína(s) pode ser feita por qualquer método conhecido do técnico, por exemplo por mutagênese no(s) gene(s) codificando essa(s) proteína(s) ou pela expressão de um ou vários RNA anti-sentido complementares do RNA ou dos RNA mensageiros codificando essa(s) proteína(s). A mutagênese pode, por exemplo, ser feita por irradiação, por ação de agente químico mutágeno, por mutagênese dirigida, por substituição de gene ou qualquer outro método conhecido do técnico. As condições convenientes dessa mutagênese podem, por exemplo, ser adaptadas segundo o ensinamento contido nas obras de Kieser, T et al (2000) e Ausubel et al., (2002). A mutagênese pode ser feita *in vitro*, *in situ*, por supressão, substituição, deleção e/ou adição de uma ou várias bases no gene considerado ou por inativação gênica. Essa mutagênese pode ser feita em um gene que compreende uma sequência tal como descrita acima. Preferencialmente, os microorganismos bloqueados em uma etapa da via de biossíntese de macrolídeos são bactérias do gênero *streptomyces*. Mais preferencialmente, a inativação da função de uma ou várias proteínas implicadas na biossíntese do(s) macrolídeo(s) em questão é feita por mutagênese. Ainda mais preferencialmente, o macrolídeo em questão é a espiramicina e os microorganismos nos quais são feitas a ou as mutagêneses são cepas de *S. ambofaciens*. Mais preferencialmente, a mutagênese é feita em um ou vários genes que compreendem uma das se-

qüências correspondentes a uma ou várias das seqüências

SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147 e 149. De forma preferida, a ou as mutagêneses são feitas por inativação gênica. De forma inteiramente preferida, a mutagênese consiste na inativação gênica de um gene que compreende uma seqüência correspondente à seqüência SEQ ID N° 13.

A título ilustrativo, os seguintes microorganismos podem ser citados a título de exemplo desses microorganismos:

OS49.16 (*orf3::Ωhyg*, cf. exemple 2), OS49.67 (*orf3délétion en phase*, cf. exemple 6), OS49.107 (*orf8::Ωhyg*, cf. exemple 7), OS49.50 (*orf10::Ωhyg*, cf. exemple 8), SPM21 (*orf2::att3Ωaac-*, cf. exemple 10), SPM22 (*orf2::att3 délétion en phase*, cf. exemple 10), SPM501 (*orf6*::att1Ωhyg+*, cf. exemple 14), SPM502 (*orf6*::att1 délétion en phase*, cf. exemple 14), SPM507 (*orf12::att3Ωaac-*, cf. exemple 11), SPM508 (*orf13c::att3Ωaac-*, cf. exemple 12), SPM509 (*orf14::att3Ωaac-*, cf. exemple 13), SPM107 (*orf28c::att3aac+*, cf. exemple 29), SPM543 (*orf31::att3aac+*, cf. exemple 30), SPM106 (*orf32c::att3aac+*, cf. exemple 31).

exemple = exemplo

5 délétion en phase = deleção em fase aquosa

Um outro aspecto da invenção refere-se a um processo de preparação de um intermediário de biossíntese de macrolídeo utilizando microorganismos bloqueados em uma etapa da via de biossíntese de macrolídeos, tais como descritos acima. O processo consiste em cultivar, em um meio de cultura apropriado, um microorganismo bloqueado em uma etapa da via de biossíntese de macrolídeos, tais como descritos acima, recuperar o meio de cultura acondicionado ou um extrato celular, por exemplo por sonicação ou por choque osmótico, separar e purificar a partir do meio de cultura ou ainda a partir do extrato celular obtido na etapa precedente, esse intermediário de biossíntese. As condições de cultura desses microorganismos poderão ser determinadas, segundo técnicas bem conhecidas do técnico. O meio de cultura poderá, por exemplo, ser o meio MP5 ou o meio SL11 para

as *streptomyces* e notadamente para *streptomyces* ambofaciens (Pernodet et al., 1993). O técnico poderá notadamente se referir à obra de Kieser et al. (2000) no que refere-se à cultura das *streptomyces*. O intermediário produzido pode ser recuperado por quaisquer técnicas conhecidas do técnico. O

5 técnico poderá se referir, por exemplo, às técnicas ensinadas na patente americana US 3.000.785 e, mais particularmente, aos métodos de extração das espiramicinas descritas nessa patente.

Um outro objeto da invenção refere-se a um processo de preparação de uma molécula derivada de um macrolídeo utilizando os microorga-

10 nismos bloqueados em uma etapa da via de biossíntese desse macrolídeo, tais como descritos acima. O processo consiste em obter um intermediário de biossíntese, segundo o processo acima e em modificar o intermediário assim produzido, eventualmente após separação do meio de cultura. As condições de cultura desses microorganismos poderão ser determinadas

15 segundo técnicas bem conhecidas do técnico. O meio de cultura poderá, por exemplo, ser o meio MP5 ou o meio SL11 para as *streptomyces* e notadamente para *streptomyces* ambofaciens (Pernodet et al., 1993). O técnico poderá notadamente se referir à obra de Kieser et al. 2000, no que refere-se à cultura das *streptomyces*. Os intermediários produzidos podem ser modifi-

20 cados, eventualmente após separação, seja por acréscimo de componentes apropriados nos meios de produção, seja por introdução nos microorganismos de outros genes codificando proteínas capazes de modificar nos microorganismos de outros genes codificando proteínas capazes de modificar o intermediário servindo-se dele como substrato. Esses intermediários podem

25 assim ser modificados por via química, bioquímica, enzimática e/ou microbiológica. Mais preferencialmente, o macrolídeo em questão é a espiramicina e os microorganismos nos quais são feitas as mutagêneses são cepas de *S. ambofaciens*.

A invenção é também relativa a um microorganismo que produz

30 a espiramicinas I, mas não produz espiramicina II e III. Esse microorganismo compreende o conjunto dos genes necessários à biossíntese da espiramicina I, mas não produz espiramicina II e III, pois o gene que compreende a

seqüência correspondente à SEQ ID N° 13 ou uma de suas variantes ou uma das seqüências derivadas daquelas em razão da degenerescência do código genético e codificando um polipeptídeo de seqüência SEQ ID N° 14 ou uma de suas variantes não é expressa ou foi tornado inativo. A inativação

5 dessa proteína pode ser feita por qualquer método conhecido do técnico, por exemplo por mutagênese no gene que codifica essa proteína ou pela expressão RNA anti-sentido complementar do RNA mensageiro codificando essa proteína, naturalmente que se a expressão de orf5* for afetada por essa manipulação será necessário proceder a uma outra modificação para

10 que o gene orf5* seja corretamente expresso. A mutagênese pode ser feita na seqüência codificante ou em uma seqüência não codificante, de maneira a tornar inativa a proteína codificada ou a impedir sua expressão ou sua tradução. A mutagênese pode ser feita por mutagênese dirigida, por substituição de gene ou qualquer outro método conhecido do técnico. As condições

15 convenientes dessa mutagênese podem, por exemplo, ser adaptadas, segundo o ensinamento contido nas obras de Kieser, T et al (2000) e Ausubel et al., 2002. A mutagênese pode ser feita *in vitro* ou *in situ*, por supressão, substituição, deleção e/ou adição de uma ou várias bases no gene considerado ou por inativação Gênica. O microorganismo pode ser também obtido,

20 expressando-se os genes da via de biossíntese da espiramicina, sem que estes compreendam o gene que compreende a seqüência SEQ ID N° 13 ou uma de suas variantes ou uma das seqüências derivadas destas, em razão da degenerescência do código genético e codificando um polipeptídeo de seqüência SEQ ID N° 14 ou uma de suas variantes. Preferencialmente, o

25 microorganismo é uma bactéria do gênero *streptomyces*. Mais preferencialmente, o microorganismo que produz a espiramicinas I, mas que não produz a espiramicina II e III é obtido a partir de um microorganismo de partida, produzindo espiramicinas I, II e III. Ainda mais preferencialmente, o microorganismo é obtido por mutagêneses em um gene que compreende a seqüência

30 correspondente à SEQ ID N° 13 ou uma de suas variantes ou uma das seqüências derivadas destas, em razão da degenerescência do código genético e codificando um polipeptídeo de SEQ ID N° 14 ou uma de suas variantes

que tem a mesma função. Ainda mais preferencialmente, essa mutagênese é feita por inativação gênica. De maneira preferida, o microorganismo é obtido a partir de uma cepa de *S.ambofaciens* que produz espiramicinas I, II e III na qual é feita uma inativação gênica do gene que compreende a seqüência correspondente à SEQ ID N° 13 ou uma das seqüências derivadas desta, em razão da degenerescência do código genético. De maneira inteiramente preferida, a invenção gênica é feita por deleção em fase do gene ou de uma parte do gene que compreende a seqüência correspondente à SEQ ID N°13 ou uma das seqüências derivadas desta em virtude da degenerescência do código genético. A título ilustrativo, a cepa SPM502 (orf6*::attI, conforme exemplo 14) pode ser citada como um microorganismo que produz a espiramicinas I, mas que não produz espiramicina II e III.

A invenção é também relativa a um microorganismo que produz a espiramicinas I, mas que não produz espiramicina II e III, tal como descrito acima, caracterizado pelo fato de superexpressar, além disso:

- um gene capaz de ser obtido por amplificação em cadeia por polimerase, utilizando o par de atrações da seguinte seqüência:

5'AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e
5'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces ambofaciens*, mais preferencialmente trata-se do gene de seqüência codificante SEQ ID N° 141;

- ou um gene derivado destes em razão da degenerescência do código genético.

Um exemplo de seqüência desse gene é dado em SEQ ID N° 111 (DNA), todavia essa seqüência é parcial, já que ela não compreende a parte 3' da seqüência codificante correspondente. A tradução em proteína dessa parte de seqüência codificante é dada em SEQ ID N° 112. O técnico poderá facilmente completá-la, notadamente utilizando o ensinamento apresentado no exemplo 24. A seqüência indeterminada em SEQ ID N° 111 foi determinada em uma segunda etapa e a seqüência completa dessa orf

(orf28c) é dada em SEQ ID N° 141. A tradução em proteína dessa sequência codificante é dada em SEQ ID N° 142. É apresentado no exemplo 24 um método para clonar o gene orf28c e para fabricar um vetor de expressão que permite a expressão de orf28c. É também mostrado nesse exemplo que a
5 superexpressão do gene orf28c na cepa OSC2 leva à melhoria da produção em espiramicinas dessa cepa. A superexpressão do gene orf28c pode ser obtida aumentando-se o número de cópias desse gene e/ou colocando-se um promotor mais ativo que o promotor primitivo. De forma preferida, a superexpressão desse gene é obtida, fornecendo no microorganismo uma
10 construção de DNA recombinante, permitindo a superexpressão desse gene. De forma preferida, essa construção de DNA recombinante aumenta o número de cópia desse gene e permite obter uma superexpressão desse gene. Nessa construção de DNA recombinante, a sequência codificante do gene pode ser colocada sob o controle de um promotor mais ativo que o promotor
15 primitivo. Pode-se citar a título de ilustração o promotor ptrc ativo em *streptomyces ambofaciens* (Amann, E et al., 1988), assim como o promotor ermE*.

Assim, de forma preferida, a ou as cópias do gene orf28c (é) são colocadas sob o controle do promotor ermE*, como é o caso na construção pSPM75
20 (conforme exemplo 24).

A invenção é também relativa a um microorganismo que produz espiramicina I, mas que não produz espiramicina II e III, tal como descrito acima, caracterizado pelo fato de superexpressar, além disso, o gene de sequência codificante SEQ ID N° 47 ou de uma sequência codificante derivada desta em razão da degenerescência do código genético. Preferencialmente, esse microorganismo é caracterizado pelo fato de se tratar a cepa SPM502 pSPM525 depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, aos 26 de fevereiro de 2003, sob o número de registro I-2977.
25
30

A invenção é também relativa a um processo de produção de espiramicinas I, o processo consiste em cultivar em um meio de cultura

apropriado um microorganismo que produz a espiramicinas I, mas que não produz espiramicina II e III, tal como descrito acima, recuperar o meio de cultura acondicionado ou um extrato celular, separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa precedente a espiramicina I. As condições de cultura desse microorganismo poderão ser determinadas segundo técnicas bem conhecidas do técnico. O meio de cultura poderá, por exemplo, ser o meio MP5 ou o meio SL11 para as *streptomyces* e notadamente para *streptomyces ambofaciens* (Pernodet et al., 1993). O técnico poderá notadamente se referir à obra de Kieser et al. 2000 no que refere-se à cultura das *streptomyces*. A espiramicina I produzida pode ser recuperada por quaisquer técnicas conhecidas do técnico. O técnico poderá se referir, por exemplo, às técnicas ensinadas na patente americana US 3.000.785 e, mais particularmente, aos métodos de extração das espiramicinas descritas nessa patente.

Um outro aspecto da invenção refere-se à utilização de uma sequência nucleotídica, de acordo com a invenção, para melhorar a produção em macrolídeos de um microorganismo. Assim, a invenção refere-se a um microorganismo mutante produtor de macrolídeo, caracterizado pelo fato de possuir uma modificação genética em pelo menos um gene que compreende uma sequência, tal como definida acima e/ou pelo fato de superexpressar pelo menos um gene que compreende uma sequência, tal como definida acima. A modificação genética pode consistir em uma supressão, uma substituição, uma deleção e/ou uma adição de uma ou várias bases no ou nos genes considerados com o objetivo de exprimir uma ou algumas proteínas tendo uma atividade superior ou de exprimir um nível mais elevado dessa ou dessas proteínas. A superexpressão do gene considerado pode ser obtida aumentando-se o número de cópias desse gene e/ou colocando-se um promotor mais ativo do que o promotor primitivo. Podem-se citar, a título de ilustração, o promotor p_{trc} ativo em *streptomyces ambofaciens* (Amann, E. et al., 1988) bem como o promotor ermE* (Bibb et al., 1985), (Bibb et al., 1994). Assim, a superexpressão do gene considerado pode ser obtida empregando-se, em um microorganismo produtor do macrolídeo considerado, uma cons-

5 trução de DNA recombinante, segundo a invenção, permitindo a superexpressão desse gene. Com efeito, certas etapas da biossíntese dos macrolídeos são limitadoras e casos e expresse uma ou várias proteínas mais ativas ou um nível de expressão mais importante da(s) proteína(s) primitiva(s) implicada(s) nessas etapas limitadoras, pode-se melhorar a produção do(s) macrolídeo(s) referido(s). Um aumento da taxa de produção da tilosina foi, por exemplo, obtida nas *streptomyces fradiae* por duplicação do gene codificando uma metil transferase limitadora que converte a macrocina em tilosina (Baltz R, 1997). A obtenção da expressão de uma proteína mais ativa pode
 10 ser obtida notadamente por mutagênese, o técnico poderá, por exemplo, se referir a esse respeito à obra de F.Ausubel et al. (2002). Preferencialmente, esses microorganismos mutantes melhorados para sua produção em macrolídeos são bactérias do gênero *streptomyces*. Mais preferencialmente, o macrolídeo em questão é a espiramicina e os microorganismos nos quais é/são
 15 feita(s) a(s) mutagênese(s) são cepas de *S.ambofaciens*. Mais preferencialmente, a modificação genética é feita em um ou vários genes que compreende(m) uma das seqüências correspondendo a uma ou várias das seqüências

SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147

e 149, ou uma de suas variantes ou uma das seqüências derivadas desta
 20 em razão da degenerescência do código genético. De forma preferida, o microorganismo superexpressa um ou vários genes que compreende(m) uma das seqüências correspondentes a uma ou várias das seqüências

SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 107, 109, 111, 113, 115, 118, 120, 141, 143, 145, 147

e 149, ou uma de suas variantes ou uma das seqüências derivadas desta
 25 em razão da degenerescência do código genético. De forma preferida, o microorganismo superexpressa o gene que compreende uma seqüência correspondente à SEQ ID N° 111 ou 141, ou uma de suas variantes ou uma das seqüências derivadas destas em razão da degenerescência do código gené-

tico. A sequência determinada em SEQ ID N° 111 é parcial, todavia, o técnico poderá facilmente completá-la, notadamente utilizando o ensinamento apresentado no exemplo 24. A sequência indeterminada em SEQ ID N° 111 foi determinada em uma segunda etapa e a sequência completa dessa orf (orf28c) é dada em SEQ ID N° 141. A tradução em proteína dessa sequência codificante é dada em SEQ ID N° 142. É assim apresentado o exemplo 24 um método para clonar o gene orf28c e para fabricar um vetor de expressão que permite a expressão de orf28c. É também mostrado nesse exemplo que a superexpressão do gene orf28c na cepa OSC2 conduz à melhoria da produção em espiramicinas dessa cepa. A superexpressão do gene orf28c pode ser obtida, aumentando-se o número de cópias desse gene e/ou colocando-se um promotor mais ativo que o promotor primitivo. De forma preferida, a superexpressão do gene orf28c é obtida, fornecendo-se no microorganismo uma construção de DNA recombinante, permitindo a superexpressão desse gene. De forma preferida, essa construção de DNA recombinante aumenta o número de cópia do gene orf28c e permite obter uma superexpressão do gene orf28c. Nessa construção de DNA recombinante, a sequência codificante de orf28c pode ser colocada sob o controle de um promotor mais ativo do que o promotor primitivo. Pode-se citar a título de ilustração o promotor ptrc ativo nos *streptomyces ambofaciens* (Amann, E. et al., 1988), assim como o promotor ermE*. Assim, de forma preferida, a(s) cópia(s) do gene orf28c fornecido é/são colocada(s) sob o controle do promotor ermE*, como é o caso na construção pSPM75 (conforme exemplo 24).

Um outro aspecto da invenção refere-se a um processo de produção de macrolídeos, utilizando os microorganismos descritos no parágrafo precedente. Esse processo consiste em cultivar em um meio de cultura apropriado um microorganismo definido no parágrafo precedente, recuperar o meio de cultura acondicionado ou um extrato celular, separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa precedente o(s) macrolídeo(s) produzido(s). As condições de cultura desse microorganismos poderão ser determinadas segundo técnicas bem conhecidas do técnico. O meio de cultura poderá, por exemplo, ser o meio

MP5 ou o meio SL11 para as *streptomyces* e notadamente para *streptomyces* ambofaciens (Pernodet et al., 1993). O técnico poderá notadamente se referir à obra de Kieser et al., 2000, no que refere-se à cultura das *streptomyces*. O(s) macrolídeo(s) produzido(s) pode(m) ser recuperado(s) por quaisquer técnicas conhecidas do técnico. O técnico poderá se referir, por exemplo, às técnicas ensinadas na patente americana US 3.000.785 e, mais particularmente, aos métodos de extração das espiramicinas descritas nessa patente. Preferencialmente, os microorganismos utilizados nesse processo são bactérias do gênero *streptomyces*. Mais preferencialmente, o macrolídeo em questão é a espiramicina e os microorganismos mutantes melhorados para sua produção em espiramicinas são cepas de *S. ambofaciens*.

Um outro aspecto da invenção refere-se à utilização de uma sequência e/ou de um vetor, de acordo com a invenção, para a preparação de antibióticos híbridos. Com efeito, os polinucleotídeos, segundo a invenção, podem servir para a obtenção de microorganismos que expressam uma ou várias proteínas mutantes que dão lugar a uma modificação na especificidade do substrato ou ainda ser expressos em múltiplos microorganismos produtores de antibióticos com a finalidade de gerar antibióticos híbridos. Assim, os polinucleotídeos, segundo a invenção, podem permitir, por transferência dos genes entre microorganismos produtores, fabricar antibióticos híbridos que têm propriedades farmacologicamente interessantes (Hopwood et al., 1985a, Hopwood et al., 1985b, Hutchinson et al., 1989). O princípio pelo qual a engenharia genética pode levar à produção de antibióticos híbridos foi inicialmente proposto por Hopwood (Hopwood 1981). Foi assim proposto que as enzimas que participam da biossíntese de antibióticos aceitam frequentemente substratos ligados estruturalmente, mas diferentes de seu substrato natural. Geralmente, é admitido (Hopwood 1981, Hutchinson 1988, Robinson 1988) que as enzimas codificadas pelos genes da via de biossíntese de antibióticos têm uma especificidade de substrato menos estrito que as enzimas do metabolismo primário. Pôde assim ser mostrado que um grande número de substratos não naturais são transformados por microorganismos produto-

res de antibiótico, seus mutantes ou enzimas purificadas da via de biossíntese desses antibióticos (Hutchinson 1988). Utilizando-se esse ensinamento, o técnico poderá construir microorganismos que expressam uma ou várias proteínas mutantes que dão lugar a uma modificação na especificidade do substrato com a finalidade de gerar antibióticos híbridos.

A invenção refere-se também à utilização de pelo menos um polinucleotídeo e/ou pelo menos um DNA recombinante e/ou pelo menos um vetor de expressão e/ou pelo menos um polipeptídeo e/ou pelo menos uma célula hospedeira, segundo a invenção, para realizar uma ou várias bioconversões. Assim, a invenção permite construir cepas bacterianas ou fúngicas nas quais se expressa, sob o controle de sinais de expressão apropriados, uma ou várias proteínas, de acordo com a invenção. Essas cepas podem, então, ser utilizadas para realizar uma ou bioconversões. Essas bioconversões podem ser feitas seja com o auxílio de células totais, seja como auxílio de extratos acelulares dessas células. Essas bioconversões podem permitir transformar uma molécula em uma forma derivada com uma enzima de uma via de biossíntese. Por exemplo, Carreras et al. Descreve a utilização de cepa de *Saccharopolyspora erythraea* e de *streptomyces coelicolor* para a produção de novos derivados de eritromicinas (Carreras et al., 2002). Walczar et al. Descreve a utilização de uma P450 monooxygenase de *streptomyces* para a bioconversão de desacetiladriamicina (um análogo da antraciclina) em novas antraciclina (Walczar et al, 2001). Olonao et al. Descreve a utilização de uma cepa de *streptomyces lividans* modificada pela bioconversão de rodomicinona-epsilom em rodomicina D (Olonao et al., 1999). O técnico poderá aplicar esse princípio a qualquer intermediário de biossíntese.

A invenção refere-se também a um DNA recombinante caracterizado pelo fato de compreender:

- um polinucleotídeo capaz de ser obtido por amplificação em cadeia por polimerase, utilizando-se o par de atrações da seguinte sequência:

5'AAGCTTGTGTGCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e
5'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces* ambofaciens, mais preferencialmente trata-se de um polinucleotídeo de seqüência SEQ ID N° 141;

- um fragmento de pelo menos 10, 12, 15, 18, 20 a 25, 30, 40,
5 50, 60, 70, 80, 90, 100,
150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950,
1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1460, 1470, 1480,
1490 ou 1500 nucleotídeos consecutivos desse polinucleotídeos.

Preferencialmente, esse DNA recombinante é um vetor. Ainda mais preferencialmente, o vetor é escolhido dentre os bacteriófagos, os plasmídeos, os fagemídeos, os vetores integrativos, os fosmídeos, os cos-
10 mídeos, os vetores navettes, os BAC (Bacterial Artificial Chromosome) ou os PAC (P1-derived Artificial Chromosome). A título ilustrativo, podem-se citar os plasmídeos que se replicam em E.coli, por exemplo pBR322 e seus derivados, pUC18 e seus derivados, pUC19 e seus derivados, pGB2 e seus derivados (Churchward G. et al., 1984), pACYC177 (número de acesso
15 GenBank: X06402) e seus derivados, pACYC184 (número de acesso GenBank: X06403) e seus derivados. Podem-se também citar os plasmídeos que se replicam em *streptomyces* como, por exemplo, pIJ101 e seus derivados, pSG5 e seus derivados, SLP1 e seus derivados, SCP2* e seus derivados (Kieser et al. 2000). Como fagemídeos podem-se citar a título ilustrativo
20 pBluescript II e seus derivados (comercializados notadamente pela sociedade Strategene (LaJolla, Califórnia, USA)), pGEM-T e seus derivados (comercializado pela sociedade Promega (Madison, Wisconsin, USA)), λZAPII e seus derivados (comercializados notadamente pela sociedade Stratagene (LaJolla, Califórnia, USA)). Como vetores integrativos, podem-se citar, a tí-
25 tulo ilustrativo, os vetores integrativos em *streptomyces* como, por exemplo, aqueles derivados de SLP1 (Kieser et al., 2000), aqueles derivados de pSAM2 (Kieser et al., 2000), os vetores utilizando os sistemas de integração

dos fágios PhiC31 (Kieser et al., 2000) (por exemplo pSET152 (Bierman et al., 1992)) ou de VWB (Van Mellaert L, et al. 1998), assim como os vetores que utilizam o sistema de integração de IS117 (Kieser et al., 2000). Como fasmídeos podem-se citar a título ilustrativo o fasmídeo pFOS1 (comerciali-

5 zado pela sociedade New England Biolabs Inc., Beverly, Massachussetts, USA) e seus derivados. Como cosmídeos, podem-se citar a título ilustrativo o cosmídeo SuperCos e seus derivados (comercializados notadamente pela sociedade Stratagene (LaJolla, Califórnia, USA)), o cosmídeo pWED15 (Wahl et al., 1987) e seus derivados. Como vetores navettes, podem-se citar

10 a título ilustrativo os plasmídeos navettes *E.coli* / *streptomyces* como, por exemplo, pIJ903 e seu derivados, a série dos plasmídeos pUWL, pCAO106, pWHM3, pOJ446 e seu derivados (Kieser et al. 2000), os BAC navettes *E.coli* / *streptomyces* como, por exemplo, aqueles descritos no pedido de patente WO 01/40497. Como BAC (Bacterial Artificial Chromosome), pode-

15 se citar, a título ilustrativo, o BAC pBeloBAC11 (número de acesso GenBank: U51113). Como PAC (P1-derived Artificial Chromosome), pode-se citar a título ilustrativo o vetor pCYPAC6 (número de acesso GenBank: AF133437). Mais preferencialmente, esse DNA recombinante é um vetor de expressão. Os vetores de expressão utilizáveis em diferentes sistemas de

20 expressão são bem conhecidos do técnico, no que refere-se às células procaríotas, podem-se citar a título ilustrativo os vetores de expressão em *E.coli* por exemplo, da família pET comercializados pela sociedade Stratagene (LaJolla, Califórnia, USA), os vetores da família GATEWAY comercializados pela sociedade Invitrogen (Carlsbad, Califórnia, USA), os vetores da família

25 pMAL comercializados pela sociedade New England Biolabs Inc. (Beverly, Massachussetts, USA), os vetores de expressão indutíveis pela ramnose citados na publicação Wilms B et al., 2001 e seu derivados, podem-se citar os vetores de expressão em *streptomyces* como, por exemplo, os vetores

pIJ4123,

pIJ6021, pPM927, pANT849, pANT 850, pANT 851, pANT1200, pANT1201, pANT1202

e seus derivados (Kieser et al., 2000). No que refere-se às células de lêve-

dos, pode-se citar a título ilustrativo o vetor pESC comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA). No que refere-se ao sistema de expressão baculovírus permitindo uma expressão em células de insetos, pode-se citar, a título ilustrativo, o vetor BacPAK 6 comercializado pela sociedade BD Biosciences Clontech, (Palo Alto, Califórnia, USA). No que refere-se às células de mamíferos, podem-se citar a título ilustrativo vetores que compreendem o promotor dos genes precoces do vírus CMV (Citomegalovírus) (Por exemplo o vetor pCMV e seu derivados comercializados pela sociedade Stratagene (LaJolla, Califórnia, USA) ou o promotor dos genes imediatos (SV40 early promoter) do vírus simiano vaculoisante SV40 (por exemplo, o vetor pSG5 comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA). Um outro aspecto da invenção refere-se às células hospedeiras na qual foi introduzido pelo menos um DNA recombinante descrito nesse parágrafo.

Um outro aspecto da invenção refere-se a um processo de produção de um polipeptídeo, caracterizado pelo fato de esse processo compreender as seguintes etapas:

- a) transformar uma célula hospedeira por pelo menos um vetor de expressão, tal como descrito no parágrafo acima;
- b) cultivar, em um meio de cultura apropriado, essa célula hospedeira;
- c) recuperar o meio de cultura acondicionado ou um extrato celular;
- d) separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa c), esse polipeptídeo;
- e) se for o caso, caracterizar o polipeptídeo recombinante produzido.

O polipeptídeo recombinante assim produzido pode ser purificado por passagem sobre uma série apropriada de colunas de cromatografia, segundo os métodos conhecidos do técnico e descritos, por exemplo, em F. Ausubel et al (2002). Pode-se citar a título indicativo a técnica do "Tag-Histidine" que consiste em acrescentar uma curta seqüência polihistidina ao

polipeptídeo a produzir, podendo este em seguida ser purificado sobre coluna de níquel. Esse polipeptídeo pode também ser preparado pelas técnicas de síntese in vitro. A título dessas técnicas, o polipeptídeo poderá ser preparado graças ao sistema "rapid translation system (RTS)", comercializado
 5 notadamente pela sociedade Roche Diagnostics France S.A., Meylan, France.

Um outro aspecto da invenção refere-se a um microorganismo produtor de pelo menos uma espiramicinas, caracterizado pelo fato de superexpressar:

- 10 - um gene capaz de ser obtido por amplificação em cadeia por polimerase (PCR), utilizando o par de atrações da seguinte seqüência:

5'AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e

5'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces* ambofaciens, mais preferencialmente trata-se do gene de seqüência codificante SEQ ID N° 141;

- 15 - ou um gene derivado destes em razão da degenerescência do código genético.

Um exemplo de seqüência desse gene é dado em SEQ ID N° 111 (DNA), todavia essa seqüência é parcial, já que ela não compreende a parte 3' da seqüência codificante correspondente. A tradução em proteína
 20 dessa parte de seqüência codificante é dada em SEQ ID N° 112. O técnico poderá facilmente completá-la utilizando o ensinamento apresentado no exemplo 24. Assim é apresentado nesse exemplo um método para clonar o gene orf28c e para fabricar um vetor de expressão que permita a expressão de orf28c. É também mostrado nesse exemplo que a superexpressão do
 25 gene orf28c na cepa OSC2 leva à melhoria da produção em espiramicinas dessa cepa. De forma preferida, o microorganismo superexpressa:

- um gene capaz de ser obtido por amplificação em cadeia por polimerase (PCR), utilizando o par de atrações da seguinte seqüência:

5'AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e
5'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces* ambofaciens, mais preferencialmente trata-se do gene de sequência codificante SEQ ID N° 141;

- ou um gene derivado deste em razão da degenerescência do código genético

e uma bactéria do gênero *streptomyces*, de maneira ainda mais preferida, trata-se de uma bactéria da espécie *streptomyces* ambofaciens. De forma
5 preferida, a superexpressão desse gene é obtida por transformação desse microorganismo por um vetor de expressão, de maneira inteiramente preferida, a cepa de microorganismo por um vetor de expressão, de maneira inteiramente preferida, a cepa de microorganismo é a cepa OSC2/pSPM75 (1) ou da cepa OSC2/pSPM75 (2) depositada junto à Collection Nationale de
10 Cultures de microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 6 de outubro de 2003, sob o número de registro I-3101.

Um outro aspecto da invenção refere-se a um processo de produção de espiramicinas, utilizando os microorganismos descritos no parágrafo
15 precedente. Esse processo consiste em cultivar em um meio de cultura apropriado um microorganismo definido no parágrafo precedente, recuperar o meio de cultura acondicionado ou um extrato celular, separar e purificar a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa precedente a(s) espiramicina(s) produzida(s). As condições de cultura
20 desses microorganismos poderão ser determinadas, segundo técnicas bem conhecidas do técnico. O meio de cultura poderá, por exemplo, ser o meio MP5 ou o meio SL11 para as *streptomyces* e notadamente para *streptomyces* ambofaciens (Pernodet et al., 1993). O técnico poderá notadamente se referir à obra de Kieser et al. 2000, no que refere-se à cultura
25 das *streptomyces*. A(s) espiramicina(s) produzida(s) pode(m) ser recuperada(s) por quaisquer técnicas conhecidas do técnico. O técnico poderá se

referir, por exemplo, às técnicas ensinadas na patente americana US 3.000.785 e mais particularmente aos métodos de extração das espiramicinas descritas nessa patente. Preferencialmente, os microorganismos utilizados nesse processo são bactérias do gênero *streptomyces*. Mais preferencialmente, os microorganismos são cepas de *S. ambofaciens*.

Um outro aspecto da invenção refere-se a um vetor de expressão caracterizado pelo fato de o polinucleotídeo de sequência SEQ ID N° 47 ou um polinucleotídeo derivado deste em razão da degenerescência do código genético é colocado sob o controle de um promotor que permite a expressão da proteína codificada por esse polinucleotídeo em *streptomyces ambofaciens*. Exemplos de vetores de expressão utilizáveis em *streptomyces* foram dados acima. Preferencialmente, esse vetor de expressão é caracterizado pelo fato de se tratar do plasmídeo pSPM524 ou pSPM525.

Um outro aspecto da invenção refere-se a uma cepa de *streptomyces ambofaciens* transformada por um vetor definido no parágrafo precedente.

Um outro aspecto da invenção refere-se a um polipeptídeo caracterizado pelo fato de sua sequência compreender a sequência SEQ ID N° 112. A invenção é também relativa a um polipeptídeo caracterizado pelo fato de sua sequência corresponder à sequência traduzida pela sequência codificante:

- de um gene capaz de ser obtido por amplificação em cadeia por polimerase (PCR), utilizando o par de atrações da seguinte sequência:

5'AAGCTTGTGTGCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e
5'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

e como matriz o cosmídeo pSPM36 ou o DNA total de *streptomyces ambofaciens*, mais preferencialmente se trata do gene de sequência codificante SEQ ID N° 141;

- ou de um gene derivado deste em razão da degenerescência do código genético.

Preferencialmente, esses polipeptídeos são expressos no estado natural por uma bactéria do gênero *streptomyces*, mais preferencialmente, esses polipeptídeos são implicados na biossíntese das espiramicinas.

Um outro aspecto da invenção refere-se a um vetor de expressão, permitindo a expressão de um polipeptídeo, tal como definido no parágrafo precedente em *streptomyces ambofaciens*. Exemplos de vetores de expressão utilizáveis em *streptomyces* foram dados acima. Preferencialmente, o vetor de expressão em questão é o plasmídeo pSPM75.

LISTA DAS FIGURAS

- 10 Figura 1: estrutura química das espiramicinas I, II e III.
 Figura 2: cosmídeos utilizados para a seqüenciação da região.
 Figura 3: organização de um grupo de genes implicados na via de biossíntese das espiramicinas.
- Figura 4: via de biossíntese proposta da micarose.
- 15 Figura 5: via de biossíntese proposta da micaminose.
 Figura 6: via de biossíntese proposta da forosamina.
 Figura 7: ordem preferencial de acréscimo dos açúcares na molécula de espiramicina e intermediários.
- Figura 8: via de biossíntese proposta do metoximalonila em
- 20 *S.ambofaciens*. Essa via é proposta por analogia com a biossíntese da metoximalonila em *streptomyces hygroscopicus* var. *ascomyceticus* (Wu,K. et al. 2000).
- Figura 9: etapas que levam à inativação de um gene:
- a) clonagem do gene alvo em um vetor que se replica em *E.coli*,
- 25 mas não nas *streptomyces*;
- B) inserção do cassette de resistência no gene alvo (Por clonagem ou recombinação entre curtas seqüências idênticas);
- C) introdução do plasmídeo em *streptomyces ambofaciens* (por transformação ou conjugação com *E.coli*) e seleção dos clones que integraram a cassette, depois alvo de crivação de clones que perderam a parte ve-
- 30 tor para ter uma substituição de gene;
- D) região do cromossoma da cepa mutante na qual o gene alvo

é inativado por substituição de gene.

Figura 10: etapas opcionais de acordo com a inativação de um gene, segundo o método descrito na Figura 9, podendo ser conduzidas caso um cassette excisável tenha sido empregada para interromper o gene alvo:

5 E) introdução na cepa mutante do plasmídeo pOSV508 portando os genes *xis* e *int* de pSAM2, cujos produtos vão permitir excisão eficaz por recombinação específica de sítios entre as seqüências attL e attR contornando a cassette;

10 F) obtenção de clones que perderam o cassette excisável e sensíveis ao antibiótico à qual o cassette conferiria a resistência;

G) após crescimento e esporulação sobre meio sólido sem antibiótico o plasmídeo pOSV508 é perdido à alta freqüência. Pode-se assim obter clones sensíveis ao tiosstrepton nos quais o gene alvo contém uma deleção em fase. A seqüência do gene alvo deletado pode ser controlada por
15 amplificação PCR e seqüenciação do produto PCR.

Figura 11: amplificação do cassette excisável com a finalidade de utilizá-la para uma experiência de recombinação homólogo. A técnica de recombinação homóloga, graças a curtas seqüências homólogas foi descrita por (Chaverroche et al., 2000). Os 39 ou 40 deóxi-nucleotídeos situados na
20 extremidade 5'desses oligonucleotídeos comportam uma seqüência correspondente a uma seqüência do gene a inativar e os 20 deóxi-nucleotídeos situados no máximo em 3' correspondem à seqüência de uma das extremidades do cassette excisável.

Figura 12: obtenção de uma construção para a inativação de um
25 gene alvo, graças à técnica descrita por (Chaverroche et al., 2000).

Figura 13: cartão do plasmídeo pWHM3. Strepto ori: origem de replicação *streptomyces*.

Figura 14: Cartão do plasmídeo pOSV508.

Figura 15: exemplo de estrutura de um cassette excisável;
30 Esta é constituída do cassette Ω hyg (Blondelet-Rouault et al., 1997) contornada dos lados attR e attL (Raynal et al., 1998), entre os quais um acontecimento de recombinação permitirá a excisão da cassette, graças à expres-

são dos genes xis e int.

Figura 16: cartão do plasmídeo pBXL1111.

Figura 17: Cartão do plasmídeo pBXL1112.

Figura 18: teste microbiológico de produção de espiramicina, baseado na sensibilidade de uma cepa de *Micrococcus luteus* à espiramicina. A cepa de *Micrococcus luteus* utilizada é uma cepa naturalmente sensível à espiramicina, mas resistente à congocidina. As diferentes cepas de *streptomyces* a testar foram cultivadas em erlenmeyers de 500 ml contendo 70 ml de meio MP5, inoculados a uma concentração inicial de 2.5×10^6 esporas/ml e colocadas para impulsionar a 27 °C sob agitação orbital de 250 rpm. Retiradas de mostos de fermentação foram feitas após 48, 72 e 96 horas de cultura e centrifugados. Uma diluição a um décimo do sobrenadante em meio de cultura estéril e utilizada para o teste. A cepa indicadora de *Micrococcus luteus* resistente à congocidina, mas sensível à espiramicina (Gourmelen et al., 1998) foi cultivada na caixa quadrada de 12 x 12 cm. Discos em papel Whatman AA foram embebidos com 70 µl da diluição com um décimo do sobrenadante e depositados sobre a superfície da caixa. Discos embebidos com uma solução de espiramicina de diferentes concentrações (2-4-8 µg/ml em meio de cultura MP5) são utilizados como faixa de aferição. As caixas são incubadas a 37°C, durante 24 a 48 horas. Se o disco contiver espiramicina, esta se espalha no agar e inibe o aumento da cepa indicadora de *Micrococcus luteus*. Essa inibição cria um "halo" em torno do disco, esse halo refletindo a zona na qual a cepa de *Micrococcus luteus* não cresceu. A presença desse halo é, portanto, uma indicação da presença de espiramicinas e permite determinar se a cepa de *S.ambofaciens* em questão é produtora ou não-produtora de espiramicina. Uma comparação com os diâmetro de inibição obtidos para a faixa de aferição permite obter uma indicação da quantidade de espiramicina produzida por essa cepa.

Figura 19: cromatograma CLHP do sobrenadante filtrado do meio de cultura da cepa OSC2.

Figura 20: cromatograma CLHP do sobrenadante filtrado do meio de cultura de cepa SPM501.

Figura 21: cromatograma CLHP do sobrenadante filtrado do meio de cultura de cepa SPM502.

Figura 22: cromatograma CLHP do sobrenadante filtrado do meio de cultura de cepa SPM507.

5 Figura 23: cromatograma CLHP do sobrenadante filtrado do meio de cultura de cepa SPM508.

Figura 24: cromatograma CLHP do sobrenadante filtrado do meio de cultura de cepa SPM509.

10 Figura 25: alinhamento da proteína Orf3 (SEQ ID N° 29) com a proteína TylB (SEQ ID N° 87) de *S. fradiae* realizado graças ao programa FASTA.

Figura 26: alinhamento da proteína MdmA de *S. mycarofaciens* (SEQ ID N° 88) com a proteína SMD (SEQ ID N° 16) realizado graças ao programa FASTA.

15 Figura 27. Exemplo de seqüências dos sítios residuais após excisão do cassette excisável. A grosso modo, é indicado o sítio att26 mínimo, tal como definido em Raynal et al., 1998. A seqüência da fase 1 (att1) de 33 nucleotídeos é apresentada em SEQ ID N° 104, a seqüência da fase 2 (att2) de 34 nucleotídeos é apresentada em SEQ ID N° 105, a seqüência da fase 3
20 (att3) de 35 nucleotídeos é apresentada em SEQ ID N° 95.

Figura 28: representação esquemática do gene orf10, localização das atrações PCR utilizadas e construção obtida com cada par de atrações.

Figura 29: construção do cassette pac-oriT;

25 Figura 30: Cartão do cosmídeo pWED2.

Figura 31: representação esquemática do grupo de genes implicados na via de biossíntese das espiramicinas e da localização das três sondas utilizadas para isolar os cosmídeos do banco de DNA genômico da cepa OSC2 de *streptomyces ambofaciens* em *E. coli* (conforme exemplo 19).

30 Figura 32: localização dos enxertos dos cosmídeos isolados do banco de DNA genômico da cepa OSC2 de *streptomyces ambofaciens* em *E. coli* (conforme exemplo 19). Novo banco de DNA cosmídico.

Figura 33: sob clonagem do fragmento PstI-PstI do cosmídeo pSPM36 (enxerto do plasmídeo (pSPM58), sob clonagem do fragmento StuI-StuI do cosmídeo pSPM36 (enxerto do plasmídeo pSPM72) e sob clonagem de um fragmento EcoRI-StuI (enxerto do plasmídeo pSPM73).

5 Figura 34: localização das fases abertas de leitura identificadas no enxerto PstI-PstI do plasmídeo pSPM58 e no enxerto EcoRI-StuI do plasmídeo pSPM73.

Figura 35: superposição dos cromatogramas CLHP que flutua filtrado do meio de cultura da cepa OS49.67 realizados a 238 e 280 nm (no
10 alto) e espectros UV das moléculas purificadas em 33,4 minutos e 44,8 minutos (em baixo).

Figura 36: estrutura molecular das moléculas platenolídeo A e platenolídeo B.

Figura 37: organização do grupo de genes implicados na via de
15 biossíntese das espiamicinas.

Figura 38: estrutura molecular de um intermediário de biossíntese produzido pela cepa SPM507.

Figura 39: estrutura de um intermediário de biossíntese produzido por uma cepa de *S. ambofaciens* de genótipo orf6*: att1Ωhyg+ obtida a
20 partir de uma cepa superproduzindo as espiamicinas. A inserção do cassette att1Ωhyg+ em orf6* exerce um efeito polar, impedindo a expressão de orf5*.

Figura 40: estrutura molecular das moléculas platenolídeo A + micarose e platenolídeo B + micarose produzidas pela cepa OS49.67.

Figura 41: sob clonagem de um fragmento PstI-PstI do cosmídeo pSPM36 (enxerto do plasmídeo (pSPM79)), localização das fases abertas de
25 leitura identificadas no enxerto PstI-PstI do plasmídeo pSPM79 e localização da sequência SEQ ID N° 140.

A presente invenção é ilustrada com o auxílio dos exemplos que se seguem, que devem ser considerados como ilustrativos e não limitativos.

30 Em resumo, os genes da via de biossíntese das espiamicinas foram isolados a partir de um banco de DNA genômico de *streptomyces ambofaciens*. Esse banco foi obtido por digestão parcial do DNA genômico de

streptomyces ambofaciens, pela enzima de restrição BamHI. Largos fragmentos de DNA, de 35 a 45 kb em média, foram clonados no cosmídeo pWED1 (Gourmelen et al., 1998) derivado do cosmídeo pWED15 (Wahl et al., 1987). Esses cosmídeos foram introduzidos em *E. Coli*, graças a partículas fágicas. O banco assim obtido foi hibridizado com uma sonda (de sequência SEQ ID N° 86) correspondente a uma parte do gene *tylB* de *S. fradiae* (merson-Davies & Cundliffe, 1994, número de acesso GenBank: U08223). Após hibridização, um cosmídeo em 4 que se hibrida com a sonda foi particularmente selecionado. Esse cosmídeo chamado pOS49.1 foi em seguida digerido por SacI e um fragmento de 3,3 kb contendo a região, o qual se hibrida com a sonda, foi clonado no vetor pUC19 e seqüenciado. Quatro fases abertas de leituras foram identificadas e uma dentre elas codifica uma proteína (SEQ ID N° 29) apresentando fortes similitudes de sequência com a proteína *TylB* de *S. fradiae* (SEQ ID N° 87) (conforme Figura 25). Esse gene foi chamado *orf3* (SEQ ID N° 28) e foi inativado em *S. ambofaciens*. Pôde ser demonstrado que os clones inativados no gene *orf3* não produzem mais espiramicinas. Isto mostra a implicação do gene *orf3* ou de genes situados a jusante na biossíntese das espiramicinas.

Uma vez conseguida essa confirmação, uma região maior do cosmídeo pOS49.1 foi seqüenciada, de ambos os lados do fragmento SacI anteriormente estudado. Assim, a partir do cosmídeo pOS49.1, a sequência de uma região que comporta sete fases abertas de leitura inteiras e duas outras incompletas, situadas de ambos os lados dessas sete fases abertas completas, pôde ser obtida. Graças a uma pesquisa nas bases de dados, pôde ser mostrado que uma das fases abertas de leitura incompleta correspondia ao locus *srmG* (região codificando uma enzima denominada "*polyketide synthase*" (PKS)). Os genes correspondentes foram clonados por Burgett S. et al. Em 1996 (Patente US 5.945.320). Por outro lado, as outras fases abertas de leitura, as sete ORFs inteiras denominadas: *orf1*, *orf2*, *orf3*, *orf4*, *orf5*, *orf6*, *orf7* (SEQ ID N° 23, 25, 28, 30, 34, 36, 40) e o começo da oitava ORF denominado *orf8* (seqüência ID N° 43) não foram encontradas nas bases de dados.

Na segunda etapa e com a finalidade de clonar outros genes implicados na biossíntese das espiramicinas, outros cosmídeos comportando fragmentos do genoma *S. ambofaciens* nessa mesma região foram isolados. Por isto uma nova série de hibridizações foi feita utilizando três sondas.

- 5 A primeira sonda corresponde a um fragmento de DNA de 3,7kb, contendo orf1, orf2 e o começo de orf3, subclonado a partir de pOS49.1. A segunda sonda corresponde a um fragmento de 2 kb de DNA, contendo uma parte de orf7 e uma parte de orf8 e subclonado a partir de pOS49.1. Uma terceira sonda foi também utilizada. Esta corresponde a um fragmento de DNA de
- 10 1.8 kb contendo o gene *srmD*. O gene *srmD* é um gene isolado de *S. ambofaciens* capaz de conferir a resistência a espiramicinas. Com efeito, trabalhos anteriores permitiram a clonagem de vários determinantes de resistência de *S. ambofaciens*, conferindo a resistência a espiramicinas a uma cepa de *S. griseofuscus* (cepa sensível à espiramicinas) (Pernodet et al.,
- 15 1993) (Pernodet et al., 1999). Para o isolamento de genes de resistências, um banco cosmídico do DNA genômico da cepa *S. ambofaciens* fora realizada no cosmídeo pKC505 (Richardson MA et al., 1987). Esse "pool" de cosmídeos fora introduzido em *S. griseofuscus*, naturalmente sensível à espiramicinas. Cinco cosmídeos capazes de conferir a resistência à apramicina e a
- 20 espiramicinas em *S. griseofuscus* foram assim obtidos. Dentre esses 5 cosmídeos, um cosmídeo denominado pOS44.1 possui em seu enxerto o gene *srmD* que codifica uma proteína que apresenta uma certa similitude com a proteína codificada pelo gene *mdmA* de *streptomyces mycarofaciens* e implicada na resistência à midecamicina nesse organismo produtor (Hara et al.,
- 25 1990, número de acesso genBank: A60725) (Figura 26). A terceira sonda utilizada para localizar os genes de biossíntese de espiramicina foi um enxerto de aproximadamente 1.8kb, compreendendo o gene *srmD*.

- Essas três sondas foram utilizadas para hibridar o banco de DNA genômico descrito acima e permitiram escolher dois cosmídeos
- 30 (pSPM7 e pSPM5) capazes de conter os enxertos os mais longos e não tendo faixas comuns. O cosmídeo pSMP5 se hibridizava com a primeira e a segunda sondas, mas não se hibridizava com a terceira sonda, enquanto

que o cosmídeo pSPM7 se hibridizava com a terceira sonda somente. Esses dois cosmídeos foram seqüenciados na totalidade pela técnica da seqüenciação às cegas ("shotgun sequencing"). As seqüências dos enxertos desses dois cosmídeos: pSPM7 e pSPM5 puderam ser ligados, pois embora não se sobrepondo, cada um dos enxertos compreendia uma seqüência conhecida em uma de suas extremidades. Com efeito, cada um desses enxertos comportava um fragmento da seqüência de um dos genes codificando uma enzima denominada "polyketide synthase" (PKS). Esses 5 genes foram clonados por Burgett S. et al. em 1996 (Patente americana US 5.945.320) (conforme Figura 2). Assim, pôde ser determinado uma seqüência única de DNA genômico de *S.ambofaciens*. Uma seqüência de 30943 nucleotídeos partindo, em 5', de um sítio EcoRI situado no primeiro gene PKS e que vai até um sítio BAMHI em 3' é apresentada em SEQ ID N° 1. Essa seqüência corresponde à região a montante dos genes da PKS (conforme Figura 2 e 3). Uma segunda região de 11171 nucleotídeos, partindo de um sítio PstI em 5' e indo até um sítio BstEII em 3' situado no quinto gene da PKS é apresentada em SEQ ID N° 2. Essa região é a região a jusante dos genes da PKS (a jusante e a montante definidos pela orientação dos 5 genes de PKS todos orientados no mesmo sentido) (conforme Figuras 2 e 3).

Em uma terceira etapa e com a finalidade de clonar outros genes implicados na biossíntese das espiramicinas, outros cosmídeos que comportam fragmentos do genoma de *S.ambofaciens* nessa mesma região foram isolados (conforme exemplo 18 e 19).

EXEMPLO 1: Construção de um banco de DNA genômico da cepa ATCC23877 de *streptomyces ambofaciens* em E.Coli

1.1. Extração do DNA genômico da cepa ATCC23877 de *streptomyces ambofaciens*.

A cepa ATCC23877 de *streptomyces ambofaciens* (acessível notadamente junto à American Type Culture Collection (ATCC) (Manassas, Virginie, USA), sob o número 23877, foi cultivada em meio YEME (Yeast Extract-Malt Extract) (Kieser, T., et al., 2000) e o DNA genômico dessa cepa foi extraído e purificado segundo as técnicas padrões de lise e precipitação

(Kieser T. et al., 2000).

1.2. Construção do banco de DNA genômico

O DNA genômico da cepa ATCC23877 de *streptomyces ambofaciens* tal como isolado acima foi digerido parcialmente pela enzima de restrição BamH de maneira a serem obtidos fragmentos de DNA de tamanho compreendido entre aproximadamente 35 e 45 kb. Esse fragmentos foram clonados no cosmídeo pWED1 (Gourmelen et al., 1998) digerido previamente por BamHI. O cosmídeo pWED1 é derivado do cosmídeo pWE15 (Wahl, et al., 1987) por deleção do fragmento HpaI-HpaI de 4.1 kb contendo o módulo de expressão ativo nos mamíferos (Gourmelen et al., 1998). A mistura de ligação foi em seguida encapsidada *in vitro* em partículas de fagos lambda graças ao sistema "Pacotene® Lambda DNA packaging system" comercializado pela sociedade Promega, segundo recomendações do fabricante. As partículas fágicas obtidas foram utilizadas para infectar a cepa SURE® de E.coli comercializada pela sociedade Stratagene (LaJolla, Califórnia, USA). A seleção dos clones foi feita sobre meio LB + ampicilina (50 µg/ml), já que o cosmídeo pWED1 confere a resistência à ampicilina.

EXEMPLO 2: isolamento e caracterização de genes implicados na biossíntese de espiramicinas em *streptomyces ambofaciens*

2.1 Hibridização sobre coluna dos clones de E.coli do banco genômico de *streptomyces ambofaciens* ATCC23877

Aproximadamente 2000 clones de E.coli do banco obtido acima foram transferidos sobre filtro para hibridização sobre colônias. A sonda utilizada para a hibridização é um fragmento Nael-Nael de DNA (SEQ ID N° 86), comportando uma parte do gene tylB de *streptomyces fradiae*. Esse fragmento corresponde aos nucleotídeos 2663 a 3702 do fragmento de DNA descrito por Merson-Davies, L.A. & Cundliffe E. (Merson-Davies, L.A. & Cundliffe E., 1994, número de acesso genBank: U08223), no qual a sequência codificante do gene tylB corresponde aos nucleotídeos 2677 a 3843.

O fragmento Nael-Nael de DNA portando uma parte do gene TylB de *streptomyces fradiae* (SEQ ID N° 86) foi marcada com ³²P pela técnica do "random piming" (Kit comercializado pela sociedade Roche) e utili-

zado como sonda para a hibridização dos 2000 clones do banco, transferidos sobre filtro. As membranas utilizadas são membranas nylon Hybond N comercializados pela sociedade Amersham (Amersham Biosciences, Orsay, França) e a hibridização foi feita a 55°C no tampão descrito por Church & Gilbert (Church & Gilbert, 1984). Uma lavagem foi feita em 2XSSC a 55°C, durante 15 minutos e duas lavagens sucessivas em 0.5X SSC foram em seguida efetuadas a 55°C, com uma duração de 15 minutos cada uma. Nessas condições de hibridização e de lavagem, 4 clones dentre os 2000 hibridados apresentavam um forte sinal de hibridização. Esses 4 clones foram cultivados em meio LB+ ampicilina (50 µg/ml) e os 4 cosmídeos correspondentes foram extraídos por lise alcalina padrão (Sambrook et al., 1989). Foi em seguida verificado que a hibridização era bem devido a um fragmento de DNA presente no enxerto desses quatro cosmídeos. Para isso, os cosmídeos foram digeridos independentemente por várias enzimas (BamHI, PstI e SacI). Os produtos de digestões foram separados sobre gel de agarose, transferidos sobre membrana Nylon e hibridados pelo fragmento Nael-Nael de DNA comportando uma parte do gene *tylB* de *streptomyces fradiae* (conforme acima) nas mesmas condições que acima. Os quatro cosmídeos puderam ser validados e um desses cosmídeos foi mais particularmente retido e foi nomeado pOS49.1.

2.2 Verificação da implicação da região identificada e seqüenciação do enxerto do cosmídeo pOS49.1

Vários fragmentos do enxerto do cosmídeo pOS49.1 foram subclonados e suas seqüências foram determinadas. O cosmídeo pOS49.1 foi digerido pela enzima SacI e foi mostrado por Southern blot, nas condições descritas anteriormente, que um fragmento de 3,3 kb contém a região que se hibrida com a sonda *tylB*. Esse fragmento de 3,3 kb foi isolado por eletroeleção a partir de um gel de agarose 0,8%, depois clonado no vetor pUC19 (número de acesso GenBank M77789) e seqüenciado. O plasmídeo assim obtido foi denominado pOS49.11. Quatro fase abertas de leitura que apresentam um uso de códons típico de *streptomyces* puderam ser identificadas nesse fragmento (duas completas e duas truncadas), graças ao programa

FramePlot (Ishikawa J & Hotta K., 1999). Comparações de seqüências graças ao programa FASTA (conforme (Pearson W.R & D.J. Lipman, 1988) e (Pearson W.R., 1990), acessível notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França) permitiram mostrar que a proteína deduzida de uma dessas 4 fases abertas de leitura apresenta fortes semelhança de seqüência com a proteína TylB (SEQ ID N° 87; Número de acesso GenBank: U08223) de *S. fradiae* (conforme Figura 25). Essa proteína foi denominada Orf3 (SEQ ID N° 29).

Com a finalidade de testar se o gene correspondente (o gene orf3 (SEQ ID N° 28)) era implicado na biossíntese das espiramicinas em *S. ambofaciens*, esse gene foi interrompido por um cassette Ω hyg (Blondelet-Rouault M-H. et al., 1997, número de acesso GenBank: X99315). Para isso, o plasmídeo pOS49.11 foi digerido pela enzima XhoI e o fragmento contendo as quatro fases abertas de leitura (duas completas e duas truncadas, das quais orf3 em inteiro) foi subclonado ao nível do sítio XhoI do vetor pBC SK + comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA). O plasmídeo assim obtido foi denominado pOS49.12. Com a finalidade de inativar orf3, um fragmento PmlI-BstEII interno em orf3 foi substituído pelo cassette Ω hyg por clonagem extremidade livre nesse último plasmídeo.

Para isso, o plasmídeo pOS49.12 foi digerido pelas enzimas PmlI e BstEII cujo sítio único se acha na seqüência codificante do gene orf3. As extremidades do fragmento correspondente ao vetor foram extremidades livres por um tratamento pela enzima de Klenow (grande fragmento da DNA polimerase I). O cassette Ω hyg foi obtido por digestão pela enzima BamHI do plasmídeo pHP45 Ω hyg (Blondelet-Rouault et al., 1997, número de acesso GenBank: X99315). O fragmento correspondente à cassette Ω hyg foi recuperado sobre gel de agarose e suas extremidades foram tornadas extremidades livres por um tratamento pela enzima de Klenow. Os dois fragmentos extremidades livres assim obtidos (o cassette Ω hyg e o plasmídeo pOS49.12) foram ligados e o produto de ligação foi utilizado para a transformação de bactérias *E. coli*. O plasmídeo assim obtido foi denominado pOS49.14 e contém o gene orf3 interrompido pelo cassette Ω hyg.

O enxerto do plasmídeo pOS49.14, sob a forma de um fragmento XhoI-XhoI, cujas extremidades foram tomadas não coesivas por um tratamento pela enzima de Klenow, foi clonado ao nível do sítio EcoRV do plasmídeo pOJ260 (o plasmídeo pOJ260 é um plasmídeo conjugativo capaz de se replicar em *E. coli*, mas incapaz de se replicar em *S. ambifaciens* (Bierman, M. et al., 1992). Esse plasmídeo confere a resistência à apramicina em *E. coli* e *streptomyces*). O plasmídeo obtido (enxerto do plasmídeo pOS49.14 clonado no plasmídeo pOJ260) foi denominado pOS49.16. Este foi transferido na cepa ATCC23877 de *S. ambifaciens* por conjugação como auxílio da cepa conjugativa *E. coli* S17-1, conforme descrito por Mazodier et al., (Mazodier et al., 1989). A cepa *E. coli* S17-1 é derivada da cepa *E. coli* 294 (Simon et al., 1983) (Simon, et al., 1986). Clones transconjugantes que possuem o marcador de resistência à higromicina portado pelo cassette Ω hyg e que perdeu o marcador de resistência à apramicina portado pelo vetor pOJ260 puderam ser obtidos. Para isso, após conjugação, os clones selecionados por sua resistência à hidromicina. Os clones resistentes à higromicina são em seguida repicados respectivamente sobre meio com higromicina (antibiótico B) e sobre meio com apramicina (antibiótico B) e sobre meio com apramicin (antibiótico A) (conforme Figura 9). Os clones resistentes à higromicina (HygR) e sensíveis à apramicina (ApraS) são em princípio aqueles em que um duplo acontecimento de recombinação se produziu e que possuem, portanto, o gene *orf3* interrompido pelo cassette Ω hyg. A substituição da cópia primitiva de *orf3* pela cópia interrompida foi verificada por duas hibridizações sucessivas. Assim, o DNA total dos clones obtidos foi digerido por diferentes enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette Ω hyg (conforme acima) para verificar a presença do cassette no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita, utilizando-se como sonda o enxerto XhoI-XhoI do plasmídeo pOS49.11 contendo as quatro fases abertas de leitura (duas completas e duas truncadas, das quais *orf3* por completo). A verificação do genótipo pode também ser realizado por qualquer método conhecido do técnico e notadamente por PCR, utilizando

oligonucleotídeos apropriados e seqüenciação do produto de PCR. Um dos clones orf3::Ωhyg assim obtido e cujo genótipo foi verificado, foi escolhido e denominado OS49.16.

5 A produção em espiramicina do clone OS49.16 assim obtido foi testada, graças ao teste de produção descrito mais abaixo (conforme exemplo 15). Pôde assim ser demonstrado que essa cepa não produz mais espiramicina, confirmando a implicação de orf3 e/ou genes situados a jusante como, por exemplo orf4, na biossíntese das espiramicinas.

10 Uma vez essa confirmação obtida, uma região maior do cosmídeo pOS49.1 foi seqüenciada, de ambos os lados do fragmento SacI anteriormente estudado. Assim, a partir do cosmídeo pOS49.1, a seqüência de uma região que comporta sete fases abertas de leitura completas e duas outras incompletas, situadas de ambos os lados dessas sete fases abertas completas, pôde ser obtida. Graças a uma pesquisa nas bases de dados, 15 pôde ser mostrado que uma das fases abertas de leitura incompleta correspondia ao locus srmG (região codificando uma enzima denominada "polyketide synthase"(PKS)). Os genes correspondentes foram clonados por Burgett S. et al., em 1996 (Patente americana US 5.945.320). Por outro lado, as outras fases abertas de leitura: as sete ORFs completas denominadas: orf1, 20 orf2, orf3, orf4, orf5, orf6, orf7 (SEQ ID N° 23, 25, 28, 30, 34, 36 e 40) e o começo da oitava ORF denominada orf8 (a seqüência completa dessa orf é dada em SEQ ID N° 43) não foram encontradas nas bases de dados.

EXEMPLO 3: isolamento e caracterização de outros genes implicados na biossíntese de espiramicinas nas *streptomyces ambofaciens*

25 Em uma segunda etapa e com a finalidade de clonar outros genes implicados na biossíntese das espiramicinas, outros cosmídeos que comportam fragmentos do genoma de *S.ambofaciens* nessa mesma região foram isolados. Para isso, procedeu-se a uma nova série de hibridizações sobre colônias, utilizando três sondas:

30 - a primeira sonda utilizada corresponde a um fragmento de DNA BamHI-PstI de 3,7kb, contendo um fragmento do gene da PKS (Os genes correspondentes à PKS foram clonados por Burgett Set al., em 1996 (Pa-

tente americana US 5.945.320), orf1, orf2 e o começo de orf3, subclonado a partir de pOS49.1 e que vai de um local BamHI situado 1300 pares de bases a montante do sítio EcoRI definindo a posição 1 de SEQ ID N° 1, até o local PstI situado em posições 2472 (SEQ ID N° 1). Esse fragmento BamHI-PstI
 5 foi subclonado a partir de pOS49.1 no plasmídeo pBC SK+, o que permitiu obter o plasmídeo pOS49.28;

- a segunda sonda utilizada corresponde a um fragmento de DNA PstI-BamHI de aproximadamente 2kb, contendo um fragmento de orf7 e de orf8, subclonado a partir de pOS49.1 e que vai de um sítio PstI situado
 10 em posição 6693 de SEQ ID N° 1 até o sítio BamHI situado em posição 8714 de SEQ ID N° 1. Esse fragmento PstI-BamHI foi subclonado a partir de pOS49.1 no plasmídeo pBC SK+, que permitiu obter o plasmídeo pOS49.76;

- uma terceira sonda foi também utilizada. Esta corresponde a um fragmento de DNA de 1.8kb EcoRI-HindIII contendo o gene *srnD*. O gene *srnD* é um gene isolado de *S.ambofaciens* capaz de conferir a resistência à espiramicina. Com efeito, trabalhos anteriores permitiram a clonagem de vários determinantes de resistência de *S.ambofaciens*, conferindo a resistência à espiramicina a uma cepa de *S. griseofuscus* (cepa sensível à espiramicina) (Pernodet et al., 1993) (Pernodet et al., 1999). Para isolar genes de resistência, um banco cosmídico do DNA genômico da cepa *S.ambofaciens* ATCC23877 foi realizada no cosmídeo pKC505 (Richardson MA et al., 1987). Para isso o DNA genômico da cepa *S.ambofaciens* ATCC23877 foi digerido parcialmente por Sau3AI, de maneira a serem obtidos fragmentos de tamanho compreendido entre aproximadamente 30 e 40 kb. O DNA genômico assim digerido (3 µg) foi ligado com 1 µg do pKC505 previamente digerido pela enzima BamHI (Pernodet et al., 1999). A mistura de ligação foi em seguida encapsulada *in vitro* em partículas fágicas. As partículas fágicas obtidas foram utilizadas para infectar a cepa de *E.coli* HB101 (acessível notadamente junto ao American Type Culture Collection (ATCC) (Manassas, Virgínia, USA), sob o número 33694).
 25 Aproximadamente 20000 clones de *E.coli* resistente à apramicina foram reunidos e os cosmídeos desses clones foram extraídos. Esse "pool" de cosmi-

deos foi introduzido por transformação de protoplastos na cepa DSM 10191 de *S. griseofuscus* (Cox KL & Baltz RH, 1984), naturalmente junto à German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikro-organismen und Zellkulturen GmbH, DSMZ), (Braunschweig, Alemanha), sob o número DSM 10191). Os transformadores foram selecionados sobre um meio contendo a apramicina. 1300 dos clones que crescem sobre o meio que contém a apramicina foram transferidos sobre o meio que contém 5 µg/ml de espiramicina. Vários clones resistentes à apramicina cresceram também sobre o meio que contém a espiramicina e os cosmídeos dessas colônias foram extraídos e utilizados para transformar *E.coli* e *S.griseofuscus* (Pernodet et al., 1999). Cinco cosmídeos capazes de conferir a resistência à apramicina em *E.coli* e a co-resistência à apramicina e a espiramicina nas *S. griseofuscus* foram assim obtidos. Dentre esses 5 cosmídeos, foi determinado que um cosmídeo denominado pOS44.1 possui em seu enxerto um gene (SEQ ID N° 15) que codifica uma proteína (SEQ ID N° 16) que apresenta uma certa similitude com uma proteína codificada pelo gene *mdmA* de *streptomyces mycarofaciens* (SEQ ID N° 88), esse gene foi denominado *srmD* (conforme alinhamento apresentado na Figura 26 realizado graças ao programa FASTA (conforme (Pearson W.R. & D.J. Lipman, 1988) e (Pearson W.R., 1990), acessíveis notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França).

Para isolar o determinante de resistência contido no plasmídeo pOS44.1, este foi digerido parcialmente pela enzima de restrição *Sau3AI*, de maneira a serem obtidos fragmentos do tamanho de aproximadamente 1,5 a 3 kb, esses fragmentos foram ligados no vetor *PIJ486* linearizado pela enzima *BamHI* (Ward et al., 1986). Um plasmídeo foi selecionado por sua capacidade de conferir a resistência a espiramicina na cepa DSM 10191 de *S.griseofuscus* (Rao RN et al., 1987), naturalmente sensível à espiramicina (conforme acima). Para isso, o "pool" de plasmídeos correspondentes aos fragmentos *Sau3AI* de pOS44.1 ligados no vetor *plJ486* (conforme acima), foi introduzido por transformação de protoplastos na cepa DSM 10191 e os transformadores foram selecionados por sua resistência ao tiostrepton (devi-

do ao gene *tsr* portado por pIJ486). Os clones que cresceram no meio que contém o tiostrepton foram transferidos sobre o meio que contém a espiramicina. Vários clones resistentes ao tiostrepton cresceram também sobre meio que contém a espiramicina e os plasmídeos dessas colônias foram extraídos. Um plasmídeo que confere a resistência e que contém um enxerto de próxima 1.8 kb foi selecionado e denominado pOS44.2. Esse enxerto de 1.8 kb pode ser tirado facilmente, graças a um sítio HindIII e um sítio EcoRI presentes no vetor de ambos os lados do enxerto. Esse enxerto de 1.8kb HindIII-EcoRI foi seqüenciado e o gene de resistência que ele contém foi denominado *srnD*. Esse fragmento que contém o gene *srnD* pôde assim ser facilmente subclonado no vetor pUC19 (número de acesso GenBank M77789) aberto por EcoRI-HindIII, o plasmídeo obtido foi denominado pOS44.4. O enxerto de 1.8kb HindIII-EcoRI, contendo o gene *srnD*, desse plasmídeo foi utilizado como sonda para localizar os genes de biossíntese de espiramicina (conforme abaixo).

Uma amostra de uma cepa *Escherichia Coli* DH5alfa contendo o plasmídeo pOS44.4 foi depositado junto à Collection Nationale de Cultures de microorganismos (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, 10 de julho de 2002 sob o número de registro I-2918.

Aproximadamente 2000 clones do banco, obtidos acima (conforme exemplo 1) foram transferidos sobre filtro para hibridização sobre colônias segundo as técnicas clássicas (Sambrook et al., 1989).

As três sondas descritas acima foram marcadas com ^{32}P pela técnica do "random priming" (Kit comercializado por Roche) e utilizadas para a hibridização foi feita a 65°C no tampão descrito por Church & Gilbert (Church & Gilbert, 1984). Uma lavagem foi feita em 2XSSC a 65°C durante 15 minutos e duas lavagens sucessivas foram em seguida feitas em 0,5 X SSC a 65°C, de uma duração de 15 minutos cada uma. Nessas condições de hibridização e de lavagem, 16 clones dentre os 2000 hibridados apresentavam um, forte sinal de hibridização com pelo menos uma das sondas. Todavia, nenhum cosmídeo não era hibridizado com as três sondas. Os 16

cosmídeos foram extraídos e digeridos pela enzima de restrição BamHI. A comparação dos perfis de restrição desses diferentes cosmídeos entre si levou a escolher dois cosmídeo capazes de conter os enxertos os mais longos da região e não tendo faixas comuns. Assim, dois cosmídeos um denominado pSPM5 e o outro pSPM7 foram escolhidos. O cosmídeo pSPM5 se hibridizava com as sondas orf1 a orf4 e a sonda orf8, mas não se hibridizava com a sonda srmD. pSPM7 se hibridizava somente com a sonda srmD e não com as duas outras sondas.

Esses dois cosmídeos foram seqüenciados na totalidade pela técnica de seqüenciação às cegas ("*shotgun sequencing*"). As seqüências dos enxertos desses dois cosmídeos: pSPM7 e pSPM5 puderam ser ligadas, pois embora não se sobreponham, cada um dos enxertos compreendia uma seqüência conhecida de um dos genes que codificam uma enzima denominada "polyketide synthase", (PKS). Esses 5 genes foram clonados por Burgett S. et al. em 1996 (Patente US 5.945.320) (Conforme Figura 2). Assim, pôde ser determinado uma seqüência única de DNA genômico de *S. ambofaciens*. Uma seqüência de 30943 nucleotídeos partindo em 5' de um sítio EcoRI situado no primeiro gene PK5 e indo até um sítio BamHI em 3' é apresentada em SEQ ID Nº 1. Essa seqüência corresponde à região a montante dos genes da PKS (conforme Figura 2 e 3). Uma segunda região de 11171 nucleotídeos, partindo de um sítio PstI em 5' e que vai até um sítio NcoI em 3' situado no quinto gene da PKS é apresentada em SEQ ID Nº 2. Essa segunda região é a região a jusante dos genes da PKS (a jusante e a montante sendo definido pela orientação dos 5 genes de pKS todos orientados no mesmo sentido) (Conforme Figuras 2 e 3).

EXEMPLO 4: análise das seqüências nucleotídicas, determinação das fases abertas de leitura e caracterização dos genes implicados na biossíntese das espiramicinas.

As seqüências obtidas foram analisadas graças ao programa FramePlot (Ishikawa J & Hotta K. 1999). Isto permitiu identificar, dentre as fases abertas de leitura as fases abertas de leitura que apresentam um uso dos códons típico de *streptomyces*. Essa análise permitir determinar que

- essa região comporta 35 ORFs localizadas de ambos os lados de cinco genes que codificam a enzima "*polyketide synthase*" (PKS). Respectivamente 10 e 25 ORFs foram identificadas a jusante e a montante desses genes (a jusante e a montante sendo definidos pela orientação dos 5 genes de PKS todos orientados no mesmo sentido) (conforme Figura 3). Assim, as 25 fases abertas de leitura desse tipo, ocupando uma região de aproximadamente 31 kb (SEQ ID N° 1 e Figura 3) foram identificadas a montante dos 5 genes codificando a PKS e 10 ocupando uma região de aproximadamente 11,1 kb (SEQ ID N° 2 e Figura 3) foram identificadas a jusante dos genes da PKS.
- 10 Os genes da região a montante foram denominados
orf1, orf2, orf3, orf4, orf5, orf6, orf7, orf8, orf9c, orf10, orf11c, orf12, orf13c, orf14, orf15c, orf16, orf17, orf18, orf19, orf20, orf21c, orf22c, orf23c, orf24c e orf25c (SEQ ID N° 23, 25, 28, 30, 34, 36, 40, 43, 45, 47, 49, 53, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 e 84).
- Os genes da região a jusante foram denominados
*orf1*c, orf2*c, orf3*c, orf4*c, orf5*, orf6*, orf7*c, orf8*, orf9*, orf10** (SEQ ID N° 3, 5, 7, 9, 11, 13, 15, 17, 19 e 21).
- O "c" acrescentado ao nome do gene significando para a ORF em questão que a sequência codificante está na orientação inversa (o filamento codificante é, portanto, o filamento complementar da sequência dada em SEQ ID
- 15 N° 1 ou SEQ ID N° 2 para esses genes) (conforme Figura 3).

- As seqüências protéicas deduzidas dessas fases abertas de leitura foram comparadas com aquelas presentes em diferentes bases de dados graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et al., 1997), CD-search, COGs (Cluster of Orthologous Groups) (esses três
- 20 programas são acessíveis notadamente junto ao National Center for Biotechnology Information (NCBI) (Bethesda, Maryland, USA), FASTA (Pearson W.R & D.J. Lipman, 1988) e (Pearson W.R., 1990), BEAUTY (Worley K.C. et al., 1995), (esses dois programas são acessíveis notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França). Essas comparações permiti-
- 25 ram formular hipóteses sobre a função dos produtos desses genes e identificar aqueles capazes de serem implicados na biossíntese de espiramicina.

EXEMPLO 5: inativação de gene: princípio da construção de uma cepa de

streptomyces ambofaciens interrompida

Os métodos aplicados consistem em efetuar uma substituição de gene. O gene alvo a interromper é substituído por uma cópia desse gene interrompida por uma cassette, conferindo a resistência a um antibiótico (por exemplo, a apramicina ou a higromicina), conforme ilustrado na Figura 9. As cassettes utilizadas são contornadas de ambos os lados por códons de terminação da tradução em todas as fases de leitura e por terminadores de transcrição ativos nas *streptomyces*.

A inserção do cassette no gene alvo pode ser acompanhada ou não de uma deleção nesse gene alvo. O tamanho das regiões que contornam o cassette pode ir de algumas centenas a vários milhares de pares de bases.

As construções necessárias à inativação do gene pelo cassette foram obtidas em *E. coli*, organismo de referência para a obtenção de construções de DNA recombinante. O gene interrompido foi obtido em um plasmídeo que se replica em *E. coli*, mas não podendo se replicar nas *streptomyces*.

As construções foram em seguida subclonados em vetores para permitir a transformação e a inativação do gene desejado nas *S. ambofaciens*. Para isso, dois plasmídeos foram utilizados:

. pOJ260 (Bierman M. et al., 1992) (conforme exemplo 2) que confere a resistência à apramicina nas *E. coli* e *streptomyces* e que foi empregado, quando o gene alvo foi interrompido por um cassette que confere a resistência à higromicina.

. pOSK 1205 (4726pb). Esse plasmídeo deriva do plasmídeo pBK-CMV (comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA), no qual um fragmento AvrII contendo a seqüência que codifica resistência à Neomicina/Canamicina foi substituída por uma seqüência que codifica a resistência à higromicina, conservando o promotor P SV40. Para isso, o plasmídeo pHP45-Ωhyg (Blondelet-Rouault et al., 1997) foi digerido pelas enzimas NotI e PflmI e o fragmento conferindo a resistência à higromicina foi subclonado ao nível do sítio AvrII do vetor pBK-CMV, depois que todas as

extremidades tenham sido tornadas extremidade livre por tratamento com enzima de Klenow. Em pOSK1205, o cassette que confere a resistência à higromicina é precedida do promotor pSV40. Esse plasmídeo confere a resistência à higromicina em *E. coli* e *streptomyces* e foi empregado quando gene alvo foi interrompido por um cassette que confere a resistência à apramicina.

As cassettes foram introduzidas no gene alvo seja por clonagem, utilizando sítios de restrição presentes no gene alvo, seja por recombinação entre curtas seqüências idênticas conforme descrito, por exemplo, por Chaveron et al (Chaveron MK et al., 2000).

O plasmídeo que porta o gene interrompido pelo cassette pode ser introduzido nas *streptomyces ambofaciens*, por exemplo por conjugação entre *E. coli* e *streptomyces* (Mazodier P. et al., 1989). Essa técnica foi utilizada no caso de o vetor de base ser o vetor pOJ260. Uma segunda técnica pode ser aplicada: a técnica da transformação de protoplastos após desnaturação por tratamento alcalino do DNA (Kieser, T. et al., 2000), para aumentar a frequência de recombinação conforme descrito, por exemplo, por Oh & Chater (Oh & Chater, 1997). Essa técnica foi utilizada no caso de o vetor de base ser pOJ260 ou pOSK1205 (conforme abaixo). Os transformadores são selecionados com o antibiótico correspondente à cassette presente no gene alvo (conforme Figura 9, antibiótico B). Seleciona-se assim uma mistura de clones dentre os quais houve integração por um só ou por dois acontecimentos de recombinação. Em uma segunda etapa, buscam-se os clones sensíveis ao antibiótico para o qual o gene de resistência está presente no vetor (fora do cassette de recombinação) (conforme Figura 9, antibiótico A). Podem-se assim selecionar os clones para os quais houve em princípio dois acontecimentos de recombinação que chegam à substituição do gene selvagem pela cópia interrompida pela cassette. Essas etapas foram esquematizadas na Figura 9.

Várias cassettes podem ser utilizadas para a interrupção dos genes alvos. Pode-se, por exemplo, utilizar o cassette Ω hyg que confere a resistência à higromicina (Blondelet-Rouault et al., 1997, número de acesso GenBank: X99315).

EXEMPLO 6: construção de uma cepa de *streptomyces ambofaciens* interrompida em fase no gene orf3

O gene orf3 fora interrompido pelo cassette Ω hyg (conforme exemplo 2.2) e pôde ser demonstrado que uma cepa orf3:: Ω hyg não produz mais espiramicina, confirmando a implicação de um ou vários genes da região clonada na biossíntese das espiramicinas (conforme exemplo 2.2). Considerando-se a sua orientação, uma cotranscrição das ORFs 1 a 7 é provável (conforme Figura 3) e o fenótipo observado (não-produtor de espiramicinas) pode ser devido à inativação de um ou vários genes co-transcritos com orf3. Para confirmar a implicação de orf3 na biossíntese de espiramicinas, uma nova inativação do gene orf3 foi feita, está sendo realizada em fase. Para isso, um fragmento Dralll interno à orf3 de 504 pares de bases foi deletado. Um fragmento de DNA obtido a partir de pOS49.1, indo do sítio EcoRI situado em posição 1 (SEQ ID Nº 1) até o sítio SaCI situado em posição 5274 (SEQ ID Nº 1) e que comporta uma deleção entre os dois sítios Dralll nas posições 2563 e 3067 (504 nucleotídeos retirados) foi clonado no plasmídeo pOJ260 (Bierman M. et al., 1992). O plasmídeo assim obtido foi denominado pOS49.67.

O enxerto de pOS49.67 é, portanto, constituído por um fragmento de DNA de *S. ambofaciens* contendo os genes orf1, orf2, orf3 com a deleção em fase, orf4 e uma parte de orf5. O vetor no qual esse enxerto foi subclonado é pOJ260, o plasmídeo pOS49.67 confere portanto a resistência à apramicina e foi introduzido por transformação de protoplastos na cepa OS49.16 (conforme exemplo 2). A cepa OS49.16 sendo resistente à higromicina, transformadores hygR e apraR foram obtidos. Após duas passagens desses clones sobre meio não seletivo, os clones sensíveis à apramicina e à higromicina (apraS e hygS) foram buscados. Em certos desses clones, se espera com efeito que um acontecimento de recombinação entre seqüências homólogas leva à substituição da cópia de orf3 interrompida pelo cassette Ω hyg (contido no genoma da cepa OS49.16) pela cópia de orf3 com a deleção em fase presente sobre o vetor. Os clones resultantes dessa recombinação são esperados como apraS e hygS, após a eliminação das seqüên-

cias do vetor. O genótipo das cepas assim obtidas pode ser verificado por hibridização ou por PCR e seqüenciação do produto de PCR (para verificar que só uma cópia de orf3 deletada em fase está presente no genoma dos clones obtidos). Clones que não têm a cópia de orf3 com uma deleção em fase foram assim obtidos e seu genótipo foi verificado. Um clone que apresenta as características buscadas foi mais particularmente selecionado e foi chamado OS49.67.

Uma amostra da cepa OS49.67 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, aos 10 de julho de 2002 sob o número de registro I-2916.

EXEMPLO 7: construção de uma cepa de *streptomyces ambofaciens* interrompida no gene orf8.

Para realizar a inativação do gene orf8, uma construção na qual o cassette Ω hyg foi introduzida na seqüência codificante de orf8 foi obtida. Para isto, o plasmídeo pOS49.88 foi inicialmente construído. O plasmídeo pOS49.88 é derivado do plasmídeo pUC19 (número de acesso GenBank M77789) por inserção de um fragmento de 3.7 kb (fragmento PstI-EcoRI obtido a partir do cosmídeo pSPM5), contendo o fim de orf7, orf8 e o começo de orf9, clonado nos sítios PstI-Eco RI de pUC19. O cassette Ω hyg (sob a forma de um fragmento BamHI, tornado extremidade livre por tratamento à enzima de Klenow) foi clonada ao nível do sítio único Sall de pOS49.88 situado em orf8, depois que todas as extremidades tenham sido tornadas extremidade livre por tratamento à enzima de Klenow.

A clonagem sendo extremidade livre, dois tipos de plasmídeos foram obtidos, segundo o sentido de inserção da cassette: pOS49.106 no qual os genes hyg e orf8 estão na mesma orientação e pOS49.120 no qual os genes hyg e orf8 estão nas orientações opostas. O enxerto do plasmídeo pOS49.106 foi em seguida subclonado no plasmídeo pOJ260 para dar pOS49.107. Para isso, o plasmídeo pOS49.106 foi digerido pela enzima Asp7181 e as extremidades foram tornadas extremidade livre por tratamento com enzima de Klenow, esse produto de digestão foi redigerido pela enzima

PstI e o fragmento contendo o gene *orf8* no qual o cassette Ω hyg foi inserida, foi clonado no vetor pOJ260 (conforme acima). Para isso, o vetor pOJ260 foi digerido pelas enzimas EcoRV e PstI e utilizado para a ligação. Essa manipulação permite, portanto, obter uma ligação orientada, já que
5 cada um dos dois fragmentos é extremidade livre de um lado e PstI do outro. O plasmídeo obtido foi chamado pOS49.107.

O plasmídeo pOS49.107 foi introduzido na cepa ATCC23877 de *S.ambofaciens* por transformação de protoplastos (Kieser, T. et al., 2000). Após transformação dos protoplastos, os clones são selecionados por sua
10 resistência à higromicina. Os clones resistentes à higromicina são em seguida repicados respectivamente sobre meio com higromicina (antibiótico B) e sobre meio com apramicina (antibiótico A) (conforme Figura 9). Os clones resistentes à higromicina (HygR) e sensíveis à apramicina (ApraS) são em princípio aqueles nos quais um duplo acontecimento de *cruzamento* se produziu e que possuem o gene *orf8* interrompido pelo cassette Ω hyg. A substituição da cópia primitiva de *orf8* pela cópia interrompida pelo cassette Ω hyg foi verificada por Southern Blot. Assim, o DNA total dos clones obtidos, foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette -
15 Ω hyg para verificar a presença do cassette no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita utilizando-se como sonda o enxerto PstI-EcoRI, contendo o fim de *orf7*, *orf8* e o começo de *orf9* de um tamanho de aproximadamente 3,7 kb do plasmídeo pOS49.88. A verificação do genótipo pode ser realizada por qualquer método conhecido do técnico e
20 notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone *orf8:: Ω hyg* foi escolhido e denominado OS49.107. Uma amostra da cepa OS49.107 foi depositado junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur
30 Roux 75724 Paris Cedex 15, France, aos 10 de julho de 2002, sob o número de registro I-2917.

EXEMPLO 8: Construção de uma cepa de *streptomyces ambofaciens* inter-

rompida no gene orf10

Um fragmento de DNA de 1,5 kb interno ao gene orf10 foi obtido por pCR, utilizando como matriz, o DNA genômico de *S.ambofaciens* e as seguintes atrações:

SRMR1 : 5' CTGCCAGTCCTCTCCCAGCAGTACG 3' (SEQ ID Nº 89)

SRMR2 : 5' TGAAGCTGGACGTCTCCTACGTCGG 3' (SEQ ID Nº 90)

- 5 Esse fragmento de DNA proveniente da PCR foi clonado no vetor pCR2.1 comercializado pela sociedade Invitrogen (Carlsbad, Califórnia, USA). O plasmídeo assim obtido foi denominado pOS49.32. O cassette Ω hyg (sob a forma de um fragmento BamHI, conforme acima) foi clonada ao nível do sítio único BstEII interno ao fragmento do gene orf10, depois que
- 10 todas as extremidades tenham sido tornadas extremidade livre por tratamento com a enzima de Klenow. A clonagem sendo extremidade livre, dois tipos de plasmídeos foram obtidos, segundo o sentido de inserção da cassette: pOS49.43 no qual os genes hyg e orf10 estão na mesma orientação e pOS49.44 no qual os genes hyg e orf10 estão em orientações opostas. O
- 15 enxerto do plasmídeo pOS49,43 for transferido (sob a forma de um fragmento Asp7181-XbaI, cujas extremidades foram tornadas extremidade livre por tratamento com a enzima de Klenow) no sítio EcoRV do plasmídeo pOJ260, o que permitiu obter o plasmídeo pOS49.50. O plasmídeo pOS49.50 contendo o fragmento do gene orf10 interrompido pelo cassette
- 20 Ω hyg foi introduzido na cepa ATCC23877 de *streptomyces ambofaciens*. Após transformação, os clones são selecionados por sua resistência à higromicina. Os clones resistentes à higromicina (antibiótico B) e sobre meio com apramicina (antibiótico A) (conforme Figura 9). Os clones resistentes à higromicina (HygR) e sensíveis à apramicina (Apras) são em princípio
- 25 aqueles nos quais um duplo acontecimento de cruzamento ocorreu e que possuem o gene orf10 interrompido pelo cassette Ω hyg. Clones que possuíam o gene orf10 interrompido pelo cassette Ω hyg. Clones que possuíam o marcador de resistência à higromicina portado pelo cassette e que perderam o marcador de resistência à apramicina portado pelo vetor pOJ260 foram

assim obtidos. O acontecimento de substituição da cópia primitiva de orf10 pela cópia interrompida orf10::-Ωhyg foi verificado por Southern Blot. Assim, o DNA genômico total dos clones obtidos, foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette Ωhyg para verificar a presença do cassette no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita, utilizando como sonda o produto PCR de 1,5 kb interna ao gene orf10 (conforme acima).

Um clone que apresenta as características esperadas (orf10::Ωhyg) foi mais particularmente selecionado e denominado OS49.50. Pôde, com efeito, ser verificado, graças às duas hibridizações, que o cassette Ωhyg estava bem presente no genoma desse clone e que se obtém o perfil de digestão esperado no caso de uma substituição, após um duplo acontecimento de recombinação, do gene primitivo pela cópia interrompida pelo cassette Ωhyg no genoma desse clone. A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR.

EXEMPLO 9: inativação de genes: princípio da construção de uma cepa de *streptomyces ambofaciens* interrompida segundo a técnica dos "cassettes excisáveis" (conforme Figura 9 e 10)

Um segundo tipo de cassettes pode ser utilizado para a inativação de genes: as cassettes ditas cassettes excisáveis. Essas cassettes apresentam a vantagem de poder ser excisadas nas *streptomyces* por um acontecimento de recombinação específica de sítio, após terem sido introduzidas no genoma de *S. ambofaciens*. A finalidade é de inativar certos genes em cepas de *streptomyces*, sem deixar na cepa final de marcadores de seleção ou de grandes seqüências de DNA que não pertencem à cepa. Após excisão subsiste unicamente uma curta seqüência de aproximadamente trinta pares de bases (denominada sítio "cicatrizial") no genoma da cepa (conforme Figura 10).

A utilização desse sistema consiste, em uma primeira etapa, na

substituição da cópia primitiva do gene alvo (graças a dois acontecimentos de recombinação homóloga, conforme Figura 9) por uma construção na qual um cassette excisável foi inserida nesse gene alvo. A inserção desse cassette é acompanhada de uma deleção no gene alvo (conforme Figura 9). Em
 5 uma segunda etapa, a excisão do cassette excisável do genoma da cepa é provocada. O cassette excisável funciona graças a um sistema de recombinação específica de sítio e tem por vantagem permitir a obtenção de mutantes de *streptomyces* não portando finalmente gene de resistência. Fica-se também livre de eventuais efeitos polares sobre a expressão dos genes situ-
 10 ados a jusante do(s) gene(s) inativado(s) (conforme Figura 10).

O emprego de cassettes excisáveis foi descrito e utilizado em numerosos organismos cujas células de mamífero, de lêvedos e em *E.coli* (Bayleu et al., 1992; Brunelli et Pall, 1993; Camilli et al., 1994; Dale et Ow, 1991; Russel et al., 1992; Lakso et al., 1992). Esses cassettes excisáveis
 15 utilizam todas a recombinase específica de sítio Cre, agindo sobre os sítios lox. Esse sistema de recombinação provém do bacteriófago P1.

Para a construção de um sistema de tipo "cassette excisável" em *streptomyces*, foi tirado partido do sistema de recombinação específica de sítio descrito para o elemento genético móvel pSAM2 de *streptomyces*
 20 *ambofaciens* (Boccard et al., 1989a e b). O sistema utilizado consiste em uma primeira etapa em construir um vetor recombinante que comporta o gene a interromper no qual é inserida um cassette excisável. A inserção no gene alvo do cassette excisável é acompanhada de uma deleção no gene alvo. Ela pode ser realizada por clonagem, utilizando sítios de restrição pre-
 25 sentes no gene alvo ou por recombinação entre curtas seqüências idênticas conforme descrito por exemplo por Chaverroche et al. (Chaverroche MK et al., 2000). O cassette excisável pode ser construída, utilizando-se por exemplo o cassette Ω hyg (Blondelet Rouault et al., 1997). Esse cassette foi contornado por seqüências attR e attL que contornam normalmente a cópia integrada de
 30 pSAM2 (conforme Figura 15). As seqüências attL e attR contêm todos os sítios necessários à recombinação sítio-específico, permitindo a excisão de pSAM2 ou de qualquer fragmento de DNA situado entre essas duas regiões

(Sezonov et al., 1997, Raynal et al., 1998). A construção dEsse cassette não está evidentemente limitada à utilização de um cassette Ω_{hyg} , mas outras cassettes de resistências podem servir de base à construção dEsse cassette (por exemplo, o cassette Ω_{aac} ou Ω_{vyh} (Blondelet Rouault et al., 1997).

5 Após a obtenção dessa construção, a cepa de *streptomyces* é transformada como plasmídeo recombinante. Os transformadores são em seguida selecionados com o antibiótico correspondente à cassette presente no gene alvo (conforme Figura 9, antibiótico B, trata-se, por exemplo, de uma seleção pela higromicina, caso o cassette excisável derive do cassette
10 Ω_{hyg}). Seleciona-se assim uma mistura de clones dentre os quais houve integração por um só ou por dois acontecimentos de recombinação. Em uma segunda etapa, buscam-se os clones sensíveis ao antibiótico para os quais o gene de resistência está presente no vetor (fora do cassette de recombinação) (conforme Figura 9, antibiótico A). Podem-se assim selecionar os
15 clones para o qual houve em princípio dois acontecimentos de recombinação que chega à substituição do gene primitivo pela cópia interrompida pela cassette. Essas etapas estão esquematizadas na Figura 9, o genótipo dos clones assim obtidos é verificado por Southern Blot e uma cepa que apresenta as características desejadas (a substituição do gene primitivo pela cópia interrompida pelo cassette excisável) é selecionada.
20

Em uma segunda etapa, a cepa selecionada acima, é transformada por um plasmídeo que permite a expressão dos genes *xis* e *int* que são ambos necessários à recombinação específica de sítio entre os sítios attR e attL. Essa recombinação acarreta a retirada do cassette excisável do
25 genoma da cepa, graças a um acontecimento de recombinação (conforme Figura 10) (Raynal et al., 1998). É interessante escolher o vetor que porta os genes *xis* e *int* dentre os vetores relativamente instáveis nas *streptomyces* (por exemplo derivado do vetor *streptomyces* pWHM3, (Vara et al., 1989), isto permite obter uma cepa que tendo perdido este vetor após alguns ciclos
30 de esporulação na ausência de pressão de seleção.

Para excisar a cassette, pode-se por exemplo utilizar o plasmídeo pOSV508 (conforme Figura 14) que é introduzido por transformação de

protoplastos na cepa de *S.ambofaciens* contendo um gene interrompido pelo cassette excisável. O plasmídeo pOSV508 é derivado do plasmídeo pWHM3 (Vara J. et al., 1989) (conforme Figura 13), no qual foram acrescentados os genes *xis* e *int* de pSAM2 (Boccard F. et al., 1989b) colocado sob o controle dopromotor *ptrc* (Amann, E. et al., 1988). Os genes *xis* e *int* colocados sob o controle do promotor *ptrc* foram subclonados no plasmídeo pWHM3 a partir do plasmídeo pOSint3 (Raynal et al., 1998) (conforme Figura 14). A introdução na cepa mutante doplas pOSV508 portando os genes *xis* e *int* de pSAM2 vai permitir a excisão eficaz por recombinação específica de sítio do cassette excisável entre os sítios *attL* e *attR* contornando Esse cassette (Raynal A. et al., 1998) (figura 10). Dentre os transformadores, selecionados por sua resistência ao tiostrepton devido ao gene *tsr* portado por pOSV508, escolhem-se aqueles tornados sensíveis ao antibiótico ao qual a presença do cassette confere a resistência (conforme Figura 10). A excisão é eficaz e foi observado que mais de 90% dos transformadores são desse tipo. Após um ou vários ciclos de aumento e esporulação sobre um meio sólido desprovido de tiostrepton. São obtidos clones que perderam o plasmídeo pOSV508. Esses clones são notáveis por sua sensibilidade ao tiostrepton. A sequência do gene alvo deletado pode ser verificada por PCR e seqüenciação do produto PCR.

Finalmente, a cepa resultante porta ao nível do gene inativado (deleção interna por exemplo) um sítio *att* "cicatrizial" correspondente ao sítio *attB* mínimo (Raynal et al., 1998), proveniente da recombinação entre os sítios *attR* e *attL*. Esse sítio *attB* mínimo que subsiste é semelhante àquele presente naturalmente nas cepas de *streptomyces ambofaciens*, *streptomyces pristinaespiralis* e *streptomyces lividans* (Sezonov et al., 1997).

O gene que se quer inativar pode ser cotranscrito com outros genes situados a jusante. Para evitar que a inativação de um dos genes tenha um efeito polar sobre a expressão dos genes a jusante no operon, é importante conseguir uma deleção em fase após excisão da cassette. O sistema das cassette excisáveis, tal como descrito acima permite responder a essa exigência. O técnico poderá, com efeito, facilmente construir três cas-

sete excisáveis distintas, estas deixando após excisão uma sequência de 33, 34 ou 35 nucleotídeos respectivamente, sem códon de parada/terminação, independentemente da fase de leitura. Conhecendo-se a sequência do gene alvo e o tamanho da deleção associada à inserção do cassette , é possível escolher entre essas três cassette excisáveis, de forma que a excisão leve a uma deleção em fase. Nos 33, 34 ou 35 nucleotídeos acrescentados, 26 correspondem à sequência attB mínima (conforme Figura 27).

No caso do presente pedido, dois cassettes excisáveis foram utilizadas. Esses dois cassettes são os seguintes: att1Ωhyg+ (SEQ ID N° 91) e att3Ωaac- (SEQ ID N° 92), essas cassettes deixam respectivamente 33 e 35 nucleotídeos após excisão. Elas comportam respectivamente o cassette Ωhyg ou o cassette Ωaac, os sinais + e - correspondendo à orientação do cassette de resistência. Essas duas cassettes foram construídas e clonadas ao nível do sítio EcoRV do vetor pBC SK+, cujo sítio HindIII foi suprimido previamente. Os plasmídeos obtidos foram chamados patt1Ωhyg+ e patt3Ωaac- respectivamente, As cassettes excisáveis podem ser facilmente tiradas por uma digestão do plasmídeo por EcoRV.

EXEMPLO 10: construção de uma cepa de *streptomyces ambofaciens* interrompida no gene orf2

A inativação do gene orf2 foi realizada graças à técnica das cassettes excisáveis (conforme acima). A cepa de partida utilizada é a cepa *streptomyces ambofaciens* OSC2 que deriva da cepa ATCC23877. Todavia, a cepa OSC2 difere da cepa ATCC23877 pelo fato de ter perdido o elemento genético móvel pSAM2 (Boccard et al., 1989a e b). Esse elemento móvel pode ser perdido de maneira espontânea durante a protoplastização (ação do lisozima para digerir a parede bacteriana e fragmentar o micélio (Kieser et al., 2000) e da regeneração dos protoplastos da cepa ATCC23877. Para selecionar os clones que perderam o elemento pSAM2, foi colocado um alvo, baseado na repressão do gene *pra* pelo repressor de transcrição KorSA (Sezonov et al., 1995) (Sezonov G. et al., 2000). Para isto, um fragmento de DNA contendo o promotor do gene *pra* colocado a montante do gene aph

(conferindo a resistência à canamicina e desprovido de seu próprio promotor) foi clonado no vetor instável pWHM3Hyg, este derivado do plasmídeo pWHM3 (Vara et al., 1989) no qual o gene *tsr* foi substituído pelo gene *hyg* (conferindo a resistência à higromicina). O plasmídeo assim obtido foi denominado pOSV510. A cepa ATCC23877 foi transformada após protoplastização pelo plasmídeo pOSV510. O promotor *pra* é um promotor reprimido pelo repressor *KorSa*, o gene que codifica este ficando situado no meio do elemento móvel pSAM2 (Sezonov G et al., 2000). Após transformação pelo plasmídeo pOSV510, as bactérias transformadas são selecionadas por sua resistência à canamicina (devido ao gene *aph* portado por pOSV510). Os clones que perderam o elemento integrativo pSAM2 perderam o repressor *KorSA* e o promotor *Pra* não é, portanto, mais reprimido e permite a expressão do gene de resistência à canamicina *aph*. Uma seleção pela canamicina, após transformação pelo plasmídeo pOSV510 permite, portanto, selecionar os clones que perderam o elemento integrativo pSAM2 (e portando *KorSA*) e possuindo o plasmídeo pOSV510. O plasmídeo pOSV510 sendo instável, após alguns ciclos de esporulação sem antibiótico, clones isolados são repicados sobre meio com canamicina e sobre meio sem antibiótico. Os clones sensíveis à canamicina e à higromicina perderam pOSV510. A perda do elemento pSAM2 foi verificada por hibridização e PCR. Um clone que apresenta as características desejadas foi selecionado e foi denominado OSC2.

Uma amostra da cepa OSC2 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2908.

A inativação do gene *orf2* foi realizada graças à técnica das cassetes excisáveis (conforme acima e Figura 10). Para isto, um enxerto de 4,5 kb, cuja sequência parte do local *EcoRI* situado em posição 1 até o sítio *BamHI* situado em posição 4521 (SEQ ID N° 1) foi subclonado ao nível dos sítios *EcoRI* e *BamHI* do plasmídeo pUC19 (número de acesso GenBank M77789), a partir do cosmídeo pSPM5. O plasmídeo assim obtido foi denominado pOS49.99.

Esse plasmídeo foi introduzido na cepa E.coli KS272 que contém o plasmídeo pKOBEG (Chaverroche et al., 2000) (cf Figura 12).

Em paralelo, o cassette excisável att3Ωaac- (SEQ ID N° 92, cf, acima) foi amplificada por PCR, utilizando como matriz o plasmídeo pOSK1102 (O plasmídeo pOSK1102 é um plasmídeo derivado do vetor pGP704Not (Chaverroche et al., 2000) (Miller VL & Mekalanos JJ, 1988) no qual o cassette att3Ωaac- foi clonada como um fragmento EcoRV no sítio único EcoRV de pGP704Not) e com o auxílio das seguintes atrações:

ORF2A

5' CCCGCGCGGCAGCCTCTCCGTCGATCGAGTCCGGCGTGACCATCGCGCGCGCTTCGTTTCGG-3' (SEQ ID N° 93)

ORF2B

5' GCTCCGTCGCGTCATGCAGGAAGGTGTCGTAGTCGCGGTAGATCTGCCTCTTCGTCGCCGAA-3' (SEQ ID N° 94)

Os 40 deóxi-nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma sequência correspondente a uma sequência no gene alvo (orf2 no caso presente) e os 20 deóxi-nucleotídeos situados além disso em 3' (representado em forte acima) correspondem à sequência de uma das extremidades do cassette excisável att3Ωaac- (conforme Figura 11).

O produto de PCR assim obtido foi utilizado para transformar a cepa E.coli contendo os plasmídeos pKOBEG e pOS49.99 conforme descrito (Chaverroche et al., 2000) (conforme Figura 12). Assim, as bactérias foram transformadas por eletroporação e selecionadas por sua resistência à apramicina. Os plasmídeos dos clones obtidos foram extraídos e digeridos por várias enzimas de restrição, com a finalidade de verificar que o perfil de digestão obtido corresponde ao perfil esperado se houver inserção do cassette (att3Ωaac-) no gene alvo (orf2), isto é, se tiver tido recombinação homóloga entre as extremidades do produto PCR e o gene alvo (Chaverroche et al., 2000). A verificação da construção pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüencição do produto de PCR. Um clone cujo plasmídeo possui o perfil esperado foi selecionado e o plasmídeo correspondente foi denominado pSPM17. Esse plasmídeo deriva de pOS49.99 no qual

orf2 é interrompida pelo cassette apramicina (conforme Figura 12).

A inserção do cassette é acompanhada de uma deleção em orf2, entre os nucleotídeos 211 e 492 da parte codificante de orf2.

O plasmídeo pSPM17 foi digerido pela enzima EcoRI, depois as
5 extremidades foram tornadas extremidade livre por tratamento com a enzima de Klenow, esse produto de digestão foi em seguida digerido pela enzima XbaI e o enxerto que contém o gene orf2 deletado foi clonado no vetor pOSK1205 (conforme acima). Para isso, o vetor pOSK1205 foi digerido pela
10 enzima XbaI e o enxerto contendo o gene orf2 deletado foi clonado no vetor pOSK1205 (conforme acima). Para isto, o vetor pOSK1205 foi digerido pela enzima BamHI, depois as extremidades foram tornadas extremidade livre por tratamento com a enzima de Klenow, esse produto foi em seguida digerido pela enzima XbaI, e utilizado para a ligação como enxerto obtido a partir de pSPM17 conforme acima. Essa manipulação permite, portanto, obter uma
15 ligação orientada, já que cada um dos dois fragmentos é extremidade livre de um lado e XbaI do outro. O plasmídeo assim obtido foi denominado pSPM21, ele porta um gene de resistência à higromicina (parte vetor) e um enxerto no qual o gene orf2 deletado é substituído pelo cassette att3Ωaac-.

O vetor pSPM21 foi introduzido na cepa *streptomyces ambofaciens* OSC2 (cf acima) por transformação de protoplastos (Kieser, T. et al.,
20 2000). Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones resistentes à apramicina são em seguida repicados respectivamente sobre o meio com apramicina (antibiótico B) e sobre meio com higromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à
25 apramicina (ApraR) e sensíveis à higromicina (HygS) são em princípio aqueles nos quais um duplo acontecimento decruzamento ocorreu e que possuem o gene orf2 interrompido pelo cassette att3Ωaac-. Esses clones foram selecionados e a substituição da cópia primitiva de orf2 pela cópia interrompida pelo cassette foi verificada. A presença do cassette att3Ωaac- foi
30 verificada por PCR sobre colônia. Uma hibridização foi também feita. Para isto, o DNA total dos clones obtidos foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizada com

uma sonda de 3 kb correspondente a um fragmento EcoRI-BamHI do enxerto do plasmídeo pOS49.99 (conforme acima). A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e sequênciação doproduto de PCR.

Um clone que apresenta as características esperadas foi selecionado e denominado SPM21. Pôde ser verificado graças à PCR e à hibridização que a cassette att3Ωaac- estava bem presente no genoma desse clone e que se obtém bem o perfil de digestão esperado no caso de uma substituição, após um duplo acontecimento de recombinação, do gene primitivo pela cópia interrompida pela cassette att3Ωaac- no genoma desse clone. Esse clone possui, portanto, o genótipo: orf2::att3Ωaac-.

Uma amostra da cepa SPM21 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2914.

A cepa SPM21 foi transformada pelo vetor pOSV508 por transformação de protoplastos para provocar a excisão do cassette (conforme Figura 14). O plasmídeo pOSV508 é derivado do plasmídeo pWHM3 (Vara J. et al., 1989) (conforme Figura 13) no qual foram acrescentados os genes xis e int de pSAM2 (Boccard F. et al., 1989b) colocado sob o controle do promotor ptrc (Amann, E. et al., 1988) (conforme Figura 14). A introdução na cepa SPM21 do plasmídeo pOSV508 portando os genes xis e int de pSAM2 permite a excisão eficaz por recombinação específica de sítio do cassette excisável entre os sítios attL e attR contornando Esse cassette (Raynal A. et al., 1998) (figura 10). Dentre os transformadores selecionados por sua resistência ao tiostrepon devido ao gene tsr portado por pOSV508, escolhem-se aqueles tornados sensíveis à apramicina cujo gene de resistência é portado pelo cassette att3Ωaac-, a excisão acarreta com efeito a perda desse gene de resistência (conforme Figura 10). O plasmídeo pOSV508 é instável e após duas passagens sucessivos sobre meio, sem antibiótico, clones isolados são repicados sobre meio com tiostrepton e sobre meio sem tiostrep-

ton. Os clones sensíveis ao tiostrepton perderam pOSV508. Foi verificado que a excisão do cassette chega bem a uma deleção em fase no gene orf2 por pCR e seqüenciação do produto PCR, a excisão do cassette deixa, com efeito, uma seqüência "cicatrizial" att3 característica (que é similar ao sítio attB de origem, após recombinação entre os sítios attL e attR):

5' ATCGCGCGCGCTTCGTTCGGGACGAAGAGGTAGAT 3' (SEQ ID N° 95).

A cepa assim obtida e possuindo o genótipo desejado (orf2::att3) foi denominado SPM22.

Uma amostra da cepa SPM22 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2915.

EXEMPLO 11: construção de uma cepa de *streptomyces ambofaciens* interrompida no gene orf12

Para a inativação de orf12, orf13c e orf14 um mesmo plasmídeo de partida (pSPM504) foi empregado para introduzir em diferentes posições um cassette de tipo cassette excisável. Esse plasmídeo possui um enxerto de 15,1 kb que corresponde à região orf7 à região orf17. Para construir esse plasmídeo um fragmento Bg/II e 15,a kb, proveniente da digestão do cosmídeo pSPM7 (conforme acima) foi clonado no plasmídeo pMBL18 (Nakano et al., 1995) digerido por BamHI. As extremidades BamHI e BG/II sendo compatíveis, obtém-se após ligação, o plasmídeo pSPM502. A totalidade do enxerto de pSPM502, obtém-se após ligação, o plasmídeo pSPM502. A totalidade do enxerto de pSPM502 foi em seguida subclonado (sob a foram de um fragmento HindIII/NheI) no plasmídeo pOSK1205 (digerido por HindIII/NheI), o que permitiu obter o plasmídeo pSPM504.

Esse plasmídeo foi introduzido na cepa E.Coli KS272 que continua já o plasmídeo pKOBEG (Chaverote et al., 2000) (conforme Figura 12).

Em paralelo, o cassette excisável att3Ωaac- foi amplificada por PCR, utilizando como matriz o plasmídeo pOSK1102 (o plasmídeo pOSK1102 é um plasmídeo derivado do vetor pGP704Not (Chaverote et al.,

2000) (Miller VL & Mekalanos JJ, 1988) no qual o cassette att3 Ω aac- foi clonada como fragmento EcoRV no sítio único EcoRV de pGP704Not), as atrações utilizadas são as seguintes:

EDR8:5' CGGGATGATCGCTTGTCGGCGGCCGGATGCCTAGCCTCATCGCGCGCGCTTCGTTCCG 3'
(SEQ ID N° 96)

EDR9:5' CCCGATCCAGAACGTCTGGTCGGTGATCAGGTCGCTGTTTCATCTGCCTCTTCGTCCCGAA 3'
(SEQ ID N° 97)

Os 40 (somente 39 para EDR8) deóxi-nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma seqüência correspondente a uma seqüência no gene alvo (orf12 no caso presente) e os 20 deóxi-nucleotídeos situados além disso em 3'(figurado em cheio acima) correspondem à seqüência de uma das extremidades do cassette excisável att3 Ω aac- (conforme Figura 11).

O produto de PCR assim obtido foi utilizado para transformar a cepa *E. coli* KS272 contendo os plasmídeos pKOBEG e pSPM504 (conforme acima), conforme descrito por Chaverroche et al. (Chaverroche et al., 2000) (conforme Figura 12 para o princípio, o plasmídeo pOS49.99 deve ser substituído pelo plasmídeo pSPM504 e o plasmídeo obtido não é mais pSPM17, mas pSPM507). Assim, as bactérias foram transformadas por eletroporação por esse produto de PCR e os clones foram selecionados por sua resistência à apramicina. Os plasmídeos dos clones obtidos foram extraídos e digeridos por várias enzimas de restrição, com a finalidade de verificar que o perfil de digestão obtido corresponde ao perfil esperado, caso tenha havido inserção do cassette (att3 Ω aac-) no gene alvo (orf12), isto é, se tiver tido recombinação homólogo entre as extremidades do produto PCR e o gene alvo (Chaverroche et al., 2000). A verificação da construção pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR. Um clone cujo plasmídeo possui um perfil esperado foi selecionado e o plasmídeo correspondente foi denominado pSPM507. Esse plasmídeo derivada de pSPM504 no qual orf12 é interrompida pelo cassette att3 Ω aac- (conforme Figura 12). A inserção do cassette é acompanhada de um dele-

ção no gene orf12 é interrompida pelo cassette att3 Ω aac- (conforme Figura 12). A inserção do cassette é acompanhada de uma deleção no gene orf12, a interrupção começa ao nível do trigésimo códon de orf12. Resta após o cassette os 46 últimos codons de orf12.

5 O vetor pSPM507 foi introduzido na cepa *streptomyces ambofaciens* OSC2 (conforme acima) por transformação de protoplastos (Kieser, T. et al., 2000). Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones resistentes à apramicina são em seguida replicados respectivamente sobre meio com apramicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (antibiótico B) e sobre
10 meio com higromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (ApraR) e sensíveis à higromicina (HygS) são em princípio aqueles nos quais um duplo acontecimento de *cruzamento* ocorreu e que possuem o gene orf12 interrompido pelo cassette att3 Ω aac-. Esses clones
15 foram mais particularmente selecionados e a substituição da cópia primitiva de orf12 pela cópia interrompida pelo cassette foi verificada por hibridização. Assim, o DNA total dos clones obtidos, foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette att3 Ω aac- para verificar a presença
20 do cassette no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita, utilizando-se como sonda um fragmento de DNA obtido por PCR e correspondendo a uma parte muito ampla da sequência codificante do gene orf12.

A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por pCR, utilizando oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone que apresenta as características esperadas (orf12::att3 Ω aac-) foi mais particularmente selecionado e denominado SPM507. Pôde com efeito ser verificado graças às duas hibridizações que o
30 cassette att3 Ω aac- estava bem presente no genoma desse clone e que se obtém bem o perfil de digestão esperado no caso de uma substituição, após um duplo acontecimento de recombinação, do gene primitivo pela cópia in-

terrompida pelo cassette att3Ωaac- no genoma desse clone. Esse clone possui, portanto, o genótipo: orf12::att3Ωaac- e foi denominado SPM507. É inútil proceder à excisão do cassette para o estudo do efeito da inativação de orf12, considerando-se a orientação dos genes (conforme Figura 3). Com efeito, o fato de o orf13c seja orientado em sentido oposto ao orf12 mostra que esses genes não são co-transcritos. A utilização de um cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer momento, notadamente por transformação pelo plasmídeo pOSV508. Uma amostra da cepa SPM507 foi depositado junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2911.

EXEMPLO 12: Construção de uma cepa de *streptomyces ambofaciens* interrompida no gene orf13c

A cassette excisável att3Ωaac- foi amplificada por PCR, utilizando como matriz o plasmídeo pOSK1102 (conforme acima) como auxílio das seguintes atrações:

EDR3 : 5' ACCGGGGCGGTCTCTCCCTCCGGGGCGTCACGGCCGCGGAATCTGCCTCTTCGTCCCGAA 3'
(SEQ ID N° 98)
EDR4 : 5' CACGCAGCGAGCCGACGCACTGATGGACGACACGATGGCCATCGCGCGCGCTTCGTTCGG 3'
(SEQ ID N° 99)

Os 40 deóxi-nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma sequência correspondente a uma sequência no gene alvo (orf13c no caso presente) e os 20 deóxi-nucleotídeos situados além disso em 3' (figurado a grosso modo acima) correspondem à sequência de uma das extremidades do cassette excisável att3Ωaac- (conforme Figura 11).

O produto de PCR assim obtido foi utilizado para transformar a cepa E.coli KS272 contendo os plasmídeos pKOBEG e pSPM504 (conforme acima), conforme descrito por Chaverroche et al. (Chaverroche et al., 2000) (conforme Figura 12 para o princípio, o plasmídeo pOS49.99) deve ser substituído pelo plasmídeo pSPM504 e o plasmídeo obtido não é mais pSPM17, mas pSPM508). Assim, as bactérias foram transformadas por ele-

troporação pelo produto de PCR e os clones foram selecionados por sua resistência à apramicina. Os plasmídeo dos clones obtidos foram extraídos e digeridos por várias enzimas de restrição, com a finalidade de verificar se o perfil de digestão obtido corresponde ao perfil esperado, caso tenha havido a

5 inserção do cassette (att3Ωaac-) no gene alvo (orf13c), isto é, se tiver tido recombinação homóloga entre as extremidades do produto PCR e o gene alvo (Chaveruche et al., 2000). A verificação da construção pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto

10 de PCR. Um clone cujo plasmídeo possui o perfil esperado foi selecionado e o plasmídeo correspondente foi denominado pSPM508. Esse plasmídeo deriva de pSPM504 no qual orf13c é interrompida pelo cassette apramicina (conforme Figura 12). A inserção do cassette é acompanhada de uma deleção no gene orf13c, a interrupção começa ao nível do sexto códon de

15 prf13c. Resta após o cassette os 3 últimos códons de orf13c.

O vetor pSPM508 foi introduzido na cepa *streptomyces ambofaciens* OSC2 (conforme acima) por transformação de protoplastos (Kieser, T. et al., 2000). Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones resistentes à apramicina são em seguida

20 repicados respectivamente sobre meio com apramicina (antibiótico B) e sobre meio com higromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (ApraR) e sensíveis à higromicina (HygS) são em princípios aqueles nos quais um duplo acontecimento de cruzamento ocorreu e que possuem o gene orf13c interrompido pelo cassette att3Ωaac-. Es-

25 ses clones foram mais particularmente selecionados e a substituição da cópia primitiva de orf13c pela cópia interrompida pelo cassette foi verificado por hibridização. Assim, o DNA total dos clones obtidos, foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette att3Ωaac- para verificar

30 a presença do cassette no DNA genômica dos clones obtidos. Uma segunda hibridização foi feita, utilizando como sonda um produto de PCR correspondente a uma seqüência que contorna uma centena de pares de base a

montante e a jusante da sequência codificante de orf13c. A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando-se oligonucleotídeos apropriados e seqüenciação do produto de PCR.

5 Um clone que apresenta as características esperadas (orf13c::att3Ωaac-) foi mais particularmente selecionado e denominado SPM508. Pôde, com efeito, ser verificado, graças às duas hibridizações que o cassette att3Ωaac- estava bem presente no genoma desse clone e que se obtém bem o perfil de digestão esperado no caso de uma substituição, após
10 um duplo acontecimento de recombinação, do gene primitivo pela cópia interrompida pelo cassette att3Ωaac- no genoma desse clone. Esse clone possui, portanto, o genótipo: orf13c::att3Ωaac- e foi nomeado SPM508. É inútil proceder a excisão do cassette para o estudo do efeito da inativação de orf13c, considerando-se a orientação dos genes (conforme Figura 3), o fato
15 de orf14 seja orientado em sentido oposto à orf13c mostra que esse genes não são co-transcritos. A utilização de um cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer momento. Uma amostra da cepa SPM508 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue
20 du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2912.

EXEMPLO 13: construção de uma cepa de *streptomyces ambofaciens* interrompida no gene orf14

O cassette excisável att3Ωaac- foi amplificada por pCR, utilizando
25 do como matriz o plasmídeo pOSK1102 (conforme acima) com o auxílio das seguintes atrações:

EDR5: 5' GGGCGTGAAGCGGGCGAGTGTGGATGTTCATGCGAGTACTCATCGCGCGCGCTTCGTTTCGG 3'
(SEQ ID N° 100)

EDR6: 5' CGGGAAACGGCGTCCGACTCCTCGGGGGCCGCGTCAGCCCATCTGCCTCTTCGTCCCGAA 3'
(SEQ ID N° 101)

Os 40 deóxi-nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma sequência correspondente a uma se-

qüência no gene alvo (orf14 no caso presente) e os 20 deóxi-nucleotídeos situados, além disso, em 3'(figurado a grosso modo acima) correspondem à seqüência de uma das extremidades do cassette excisável att3Ωaac- (conforme Figura 11).

5 O produto de PCR assim obtido foi utilizado para transformar a cepa E.coli KS272, contendo os plasmídeo pKOBEG e pSPM504 (conforme acima), conforme descrito por Chaverroche et al. (Chaverroche et al., 2000) (conforme Figura 12 para o princípio, o plasmídeo pOS49.99 deve ser substituído pelo plasmídeo pSPM504 e o plasmídeo obtido não é mais pSPM17,
10 mas pSPM509). Assim, as bactérias foram transformadas por eletroporação pelo produto de PCR e os clones foram selecionados por sua resistência à apramicina. Os plasmídeos dos clones obtidos foram extraídos e digeridos por várias enzimas de restrição, com a finalidade de verificar se o perfil de digestão obtido corresponde ao perfil esperado, caso tenha havido inserção
15 do cassette (att3Ωaac-) no gene alvo (orf14), isto é, se houve recombinação homóloga entre as extremidades do produto PCR e o gene alvo (Chaverroche et al., 2000). A verificação da construção pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR. Um
20 clone, cujo plasmídeo possui o perfil esperado, foi selecionado e o plasmídeo correspondente foi denominado pSPM509. Esse plasmídeo deriva de pSPM504 no qual orf14 é interrompido pelo cassette apramicina (conforme Figura 12). A inserção do cassette é acompanhada de uma deleção no gene orf14, a interrupção começa no nível do quarto códon de orf14. Resta após o
25 cassette o último códon de orf14.

O vetor pSPM509 foi introduzido na cepa *streptomyces ambofaciens* OSC2 (conforme acima) por transformação de protoplastos (Kieser, t. et al., 2000). Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones resistentes à apramicina são em seguida
30 repicados, respectivamente sobre meio com apramicina (antibiótico B) e sobre meio com higromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (ApraR) e sensíveis à higromicinas (HygS) são em

princípio aqueles nos quais um duplo acontecimento de cruzamento ocorreu e que possuem o gene orf14 interrompido pelo cassette att3Ωaac-. Esses clones foram mais particularmente selecionados e a substituição da cópia primitiva de orf14 pela cópia interrompida pelo cassette foi verificada por hibridização. Assim, o DNA total dos clones obtidos foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette att3Ωaac- para verificar a presença do cassette no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita, utilizando como sonda um produto de PCR, correspondendo a uma sequência que ultrapassa uma centena de pares de base a montante e a jusante da sequência codificante de orf14. A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone que apresenta as características esperadas (orf14::att3Ωaac-) foi mais particularmente selecionado e denominado SPM509. Pôde com efeito ser verificado, graças às duas hibridizações, que o cassette att3Ωaac- estava presente no genoma desse clone e que se obtém o perfil de digestão esperado no caso de uma substituição, em consequência de um duplo acontecimento de recombinação, do gene primitivo pela cópia interrompida pelo cassette att3Ωaac- no genoma desse clone. Esse clone possui, portanto, o genótipo: orf14::att4 e foi denominado SPM509. É inútil proceder à excisão do cassette para o estudo do efeito da inativação de orf14, considerando-se a orientação dos genes (conforme Figura 3), o fato de orf15c ser orientado em sentido oposto ao orf14 mostra que esses genes não são co-transcritos. A utilização de um cassette excisável; permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer momento. Uma amostra da cepa SPM509 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2913.

EXEMPLO 14: construção de uma cepa de *streptomyces ambofaciens* inter-

rompida no gene *orf6**

A inativação do gene *orf6** foi realizada graças à técnica das cassette excisáveis (conforme acima e Figura 10). Para isso, o cosmídeo pSPM7 foi utilizado como matriz para amplificar um fragmento do gene *orf6**,
5 graças aos seguintes oligonucleotídeos:

C9583 : 5' CTGCAGGTGCTCCAGCGCGTCGATCT 3' (oligo sens) (SEQ ID N° 102)

C9584 : 5' CTGCAGACGGAGGCGGACCTGCGGCT 3' (oligo antisens) (SEQ ID N° 103)

Os 20 deóxi-oligo situados na extremidade 3' desses oligonucleotídeos correspondem a uma sequência situada na parte codificante do gene *orf6** (SEQ ID N° 13) e os 6 deóxi-nucleotídeos situados além disso em 5' correspondem à sequência de um sítio PdtI, permitindo facilitar a clonagem
10 na sequência. O fragmento de DNA amplificado constitui um tamanho de aproximadamente 1,11 kb. Esse produto de PCR é clonado no vetor pGEM-T Easy (comercializado pela sociedade Promega (Madison, Wisconsin, USA), o que permitiu obter o plasmídeo denominado pBXL1111 (conforme Figura 16).

15 A cassette excisável att1Ωhyg⁺ foi em seguida introduzida na sequência codificante do gene *orf6**. Para isso, o plasmídeo pBXL1111 foi digerido pelas enzimas de restrição SmaI e Asp718I e o produto de digestão foi tratado pela enzima de Klenow. Essa manipulação permite realizar uma deleção interna de 120 pb na sequência codificante do gene *orf6** (conforme
20 Figura 15). Além disso, de ambos os lados dos sítios de restrições, restam respectivamente 511 pb e 485 pb da sequência de *orf6** que permitirão a recombinação homóloga para a inativação do gene *orf6**. O cassette excisável att1Ωhyg⁺ foi preparada a partir do plasmídeo patt1Ωhyg⁺ (conforme acima) por digestão desse plasmídeo por EcoRV. Esta última foi em seguida
25 subclonada no vetor pBXL1111 previamente preparado conforme descrito acima (digestão SmaI e Asp718I, depois tratamento com a enzima de Klenow). O plasmídeo obtido foi denominado pBXL1112 (conforme Figura 17). Nessa construção, o gene *orf6** comporta uma deleção de 120 pb e interrompido pelo cassette att1Ωhyg⁺.

30 O plasmídeo pBXL1112 foi em seguida digerido pela enzima PstI

(sítio que contorna a cassette, já que presente nos oligonucleotídeos PCR) e o enxerto PstI de 3,7 kb compreendendo uma parte de orf6* interrompido pelo cassette att1Ωhyg+ foi em seguida clonado ao nível do sítio PstI do plasmídeo pOJ260 (conforme acima). O plasmídeo assim obtido foi denominado pBXL1113.

O vetor pBXL1113 foi introduzido na cepa *Streptomyces ambifaciens* OSC2 (conforme acima) por transformação de protoplastos (Kieser, T. et al., 2000). Após transformação, os clones são selecionados por sua resistência à higromicina. Os clones resistentes à higromicina são em seguida repicados respectivamente sobre meio com higromicina (antibiótico B) e sobre meio com apramicina (antibiótico A) (conforme Figura 9). Os clones resistentes à higromicina (HygR) e sensíveis à apramicina (ApraS) são em princípio aqueles nos quais um duplo acontecimento de cruzamento ocorreu e que possuem o gene orf6* interrompido pelo cassette att1Ωhyg+. Esses clones foram mais particularmente selecionados e a substituição da cópia primitiva de orf6* pela cópia interrompida pelo cassette foi verificada pela técnica do Southern Blot. Assim, o DNA total dos clones obtidos foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente ao gene hyg (obtido por PCR) para verificar a presença do cassette no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita, utilizando-se como sonda o enxerto PstI-PstI, contendo o gene orf6*, de um tamanho de aproximadamente 1,1 kb e obtido a partir do plasmídeo pBXL1111 (conforme acima e Figura 6). A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por pCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone que apresenta as características esperadas (orf6*::att1Ωhyg+) foi mais particularmente selecionado e denominado SPM501. Pôde com efeito ser verificado, graças às duas hibridizações que o cassette att1Ωhyg+ estava presente no genoma desse clone e que se obtém o perfil de digestão esperado no caso de uma substituição, em consequência de um duplo acontecimento de recombinação, do gene primitivo pela cópia

interrompida pelo cassette att1Ωhyg+ no genoma desse clone. Esse clone possui , portanto, o genótipo: orf6*::att1Ωhyg+ e foi denominado SPM501. Uma amostra da cepa SPM501 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur
 5 Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2909.

A cepa SPM501 foi transformada pelo vetor pOSV508 por transformação de protoplastos para provocar a excisão do cassette (conforme Figura 14). O plasmídeo pOSV508 é derivado do plasmídeo pWHM3
 10 (Vara J. et al., 1989) (conforme Figura 13) no qual os genes *xis* e *int* de pSAM2 (Boccard F. et al., 1989b) colocados sob o controle do promotor ptrc (Amann, E. et al., 1988) foram acrescentados (conforme Figura 14). A introdução na cepa SPM501 do plasmídeo pOSV508 portando os genes *xis* e *int* de pSAM2 permite a excisão eficaz por recombinação sítio específico do
 15 cassette excisável entre os sítios attL e attR que contornam Esse cassette (Raynal A. et al., 1998) (figura 10). Dentre os transformadores, selecionados por sua resistência ao tiostrepton, devido ao gene *tsr* portado por pOSV508, escolhem-se aqueles tornados sensíveis à higromicina, cujo gene de resistência é portado pelo cassette att1Ωhyg+, a excisão acarreta com efeito a
 20 perda desse gene de resistência 9conforme Figura 10). O plasmídeo pOSV508 é instável e após duas passagens sucessivas sobre meio sem antibiótico dos clones isolados são repicados sobre meio com tiostrepton e sobre meio sem tiostrepton. Os clones sensíveis ao tiostrepton perderam pOSV508. Foi verificado que a excisão do cassette ocorreu em fase no gene
 25 orf6* por PCR e seqüenciação do produto PCR. A interrupção começa no 158º códon, 40 códons são deletados (120pb) e a excisão do cassette deixa uma seqüência "cicatrizial" att1 característica de 33pb:

5'ATCGCGCGCTTCGTTCGGGACGAAGAGGTAGAT 3' (SEQ ID Nº 104).

A cepa assim obtida e possuindo o genótipo desejado (orf6*::att1) foi denominada SPM502.

30 Uma amostra da cepa SPM502 foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue

du Docteur Roux 75724 Paris Cedex 15, France, em 10 de julho de 2002 sob o número de registro I-2910.

EXEMPLO 15: análise das cepas de *streptomyces ambofaciens* interrompida nos genes orf2, orf3, orf8, orf10, orf12, orf13c, orf14, orf6*

5 Para testar a produção em espiramicinas das diversas cepas obtidas, um teste microbiológico baseado na sensibilidade de uma cepa de *Micrococcus luteus* foi desenvolvido (conforme(Goumelen A. et al., 1998)). A cepa de *Micrococcus luteus* utilizada é uma cepa derivada da cepa DSM1790 naturalmente sensível à espiramicina (essa cepa está disponível
10 notadamente junto à German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikro-organismen und Zellkulturen GmbH, DSMZ), (Braunschweig, Alemanha), sob o número DSM1790), a cepa utilizada no presente teste difere da cepa DSM1790 pelo fato de ser resistente à congocidina. Essa cepa é um mutante espontâneo obtido por seleção sobre
15 meio que contém doses crescentes de congocidina. Essa cepa é um mutante espontâneo obtido por seleção sobre meio que contém doses crescentes de congocidina. Essa cepa foi selecionada devido ao fato de *streptomyces ambofaciens* produzir ao mesmo tempo a espiramicina e da congocidina. A finalidade sendo dosar a produção em espiramicina das di-
20 versas cepas obtidas graças a um teste microbiológico baseado na sensibilidade de uma cepa de *Micrococcus luteus*, é necessário dispor de uma cepa resistente à congocidina.

 As diferentes cepas de *streptomyces* a testar foram cultivadas em erlenmeyers com chicans (erlenmeyers chicanados) de 500 ml, contendo 70
25 ml de meio MP5 (Pernodet et al., 1993). Os erlenmeyers chicanados foram inoculados a uma concentração inicial de $2.5 \cdot 10^6$ esporas/ml das diferentes cepas de *S.ambofaciens* e colocadas para crescer a 27°C sob agitação orbital de 250 rpm. Retiradas de 2 ml de suspensões foram feitas após 48, 72 e 96 horas de cultura e centrifugadas. Os diferentes resíduos sobrenadantes foram em se-
30 guida congelados a -20°C. Uma diluição ao décimo desses resíduos em meio de cultura estéril é utilizada para o teste (conforme Figura 18).

 A cepa indicadora de *Micrococcus luteus* resistente à congocidi-

na, mas sensível à espiramicinas, foi cultivada em meio 2TY (Sambrook et al., 1989) contendo a congocidina a 5 µg/ml durante 48 horas a 37°C. A densidade óptica (DO) da cultura é medida e essa cultura é diluída de forma a ajustar a densidade óptica em 4. 0,4 ml dessa pré-cultura é diluída em 40 ml de meio DAM5 (Difco Antibiotic Medium 5, comercializado pela sociedade Difco), previamente levado a uma temperatura de aproximadamente 45°C. Esse meio é em seguida derramado em uma caixa quadrada de 12 x 12 cm e deixado em repouso à temperatura ambiente.

Uma vez o meio resfriado e solidificado, discos em papel Whatman AA (Conforme Gourmelen A. et al., 1998) de 12 mm de diâmetro foram embebidos com 70 µl da diluição ao décimo de cada resíduo sobrenadante e depositados sobre a superfície da caixa. Discos embebidos com uma solução de espira de diferentes concentrações (2-4-8 µg/ml em meio de cultura MP5) são utilizados como faixa de aferição. As caixas são deixadas a 4 °C durante 2 horas, de forma a permitir a difusão dos antibióticos no agar, depois são incubadas a 37°C, durante 24 a 48 horas.

Caso o disco contenha espiramicina, esta se propaga no agar e inibe o aumento da cepa indicadora de *Micrococcus luteus*. Essa inibição cria um "halo" em torno do disco, esse halo refletindo a zona na qual a cepa de *Micrococcus luteus* não cresceu. A presença desse halo é, portanto, uma indicação da presença de espiramicinas e permite determinar se a cepa de *S. ambofaciens* correspondente ao disco em questão é produtora ou não de espiramicina. Uma comparação com os diâmetros de inibição obtidos para a faixa de aferição permite obter uma indicação da quantidade de espiramicinas produzida por essa cepa.

As diferentes cepas descritas nos Exemplos precedentes foram utilizadas nesse teste para detectar sua produção em espiramicinas. Os resultados obtidos foram reagrupados na Redutase.

Tabela 38

Cepa	Gene inativado	Exemplo no qual a cepa é descrita	Fenótipo: Produtor (+) ou não-produtor (-) de espiramicina
ATCC23877	Nenhum	1	(+)
OS49.16	<i>orf3::Ωhyg</i>	2	(-)
OS49.67	<i>orf3</i> deleção em fase	6	(-)
OS49.107	<i>orf8::Ωhyg</i>	7	(-)
OS49.50	<i>orf10::Ωhyg</i>	8	(-)
OSC2	Nenhum	10	(+)
SPM21	<i>orf2::att3Ωaac-</i>	10	(-)
SPM22	<i>orf2::att3</i> deleção em fase	10	(-)
SPM501	<i>orf6*::att1Ωhyg+</i>	14	(-)
SPM502	<i>orf6*::att1</i> deleção em fase	14	(+)
SPM507	<i>orf12::att3Ωaac-</i>	11	(-)
SPM508	<i>orf13c::att3Ωaac-</i>	12	(+)
SPM509	<i>orf14::att3Ωaac-</i>	13	(-)

Esses resultados permitem tirar um certo número de conclusões no que refere-se à função dos diferentes genes implicados na biossíntese da espiramicinas. Assim, o gene *orf3* é essencial à biossíntese da espiramicinas. Com efeito, uma inativação em fase nesse gene leva a uma cepa (OS49.67, (conforme exemplo 6)) não produzindo espiramicina. A inativação em fase permite afastar a hipótese de uma eventual influência do cassette introduzida sobre a expressão dos genes situados a jusante de *orf3*.

Da mesma forma, os genes orf8 e orf10 codificam proteínas essenciais à biossíntese da espiramicinas, já que as cepas OS49.107 e OS49.50 têm um fenótipo não-produtor. Além disso, nestas duas últimas cepas, é a inativação do gene correspondente que é responsável por esse
5 fenótipo não-produtor, já que, considerando-se diferentes orfs (conforme Figura 3), a construção introduzida não pode ter efeito polar.

O estudo das cepas, possuindo um cassette excisável, permite tirar um certo número de conclusões no que refere-se à função dos genes interrompidos. A cepa SPM507 possui o genótipo: orf12::att3Ωaac-. É inútil
10 proceder à excisão do cassette para o estudo do efeito da inativação de orf12, considerando-se a orientação dos genes (conforme Figura 3). O fato de orf13c ser orientado em sentido oposto ao orf12 mostra que esses genes não são co-transcritos. A utilização de um cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer
15 momento. O fenótipo da cepa SPM507 é não-produtor, pode-se, portanto, concluir que o gene orf12 é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*.

A cepa SPM 508 possui o genótipo: orf13c::att3. É inútil proceder à excisão do cassette para o estudo do efeito da inativação de orf13c,
20 considerando-se a orientação dos genes (conforme Figura 3). O fato de orf14 ser orientado em sentido oposto ao orf13c mostra que esses genes não são co-transcritos. A utilização de um cassette excisável permite ao contrário ter a possibilidade de se livrar do marcador de seleção a qualquer momento. O fenótipo da cepa SPM508 é produtor, pode-se, portanto, con-
25 cluir que o gene orf13c não é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*.

A cepa SPM509 possui o genótipo: orf14::att3Ωaac-. É inútil proceder à excisão do cassette para o estudo do efeito da inativação de orf14, considerando-se a orientação dos genes (conforme Figura 3), o fato de
30 orf15c ser orientado em sentido oposto ao orf14 mostra que esses genes não são co-transcritos. A utilização de um cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer

momento. O fenótipo da cepa SPM509 é não-produtor, pode-se, portanto, concluir que o gene orf14 é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*.

A cepa SPM21 possui o genótipo: orf2::att3 Ω aac-. O fenótipo
5 dessa cepa é não-produtor de espiramicinas. Todavia, a orientação dos genes orf1 ao orf8 (conforme Figura 3) deixa pensar que esses genes são co-transcritos. Também, o fenótipo observado pode ser devido a um efeito polar do cassette introduzida em orf2 sobre a expressão de genes situados a jusante no operon. A cepa SPM22 possui o genótipo orf2::att3 e foi obtida
10 após excisão em fase do cassette introduzida. A excisão do cassette deixa somente uma seqüência "cicatricial" característica (conforme exemplo 10). A cepa SPM22 sendo também de fenótipo não-produtor, pode-se concluir que o gene orf2 é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*. Observa-se no caso unicamente o efeito devido à inativação
15 de orf2.

A cepa SPM501 possui o genótipo: orf6*::att1 Ω hyg+. O fenótipo
dessa cepa é não-produtor de espiramicinas. Todavia, os genes orf5* e orf6*
(conforme Figura 3) que tem a mesma orientação, o fenótipo observado
pode ser devido a um efeito polar do cassette introduzida em orf6* sobre a
20 expressão de orf5*. A disposição desses genes deixar pensar que podem ser co-transcritos. Para responder a essa questão, a cepa SPM502 foi obtida após excisão em fase do cassette introduzida. Nessa cepa, observa-se unicamente o efeito da inativação de orf6*. Essa cepa possui o genótipo orf6*::att1 (conforme exemplo 14). A excisão do cassette deixa somente uma
25 seqüência cicatricial em fase (conforme exemplo 14). A cepa SPM502 possui um fenótipo produtor (essa cepa todavia produz apenas a espiramicina I (cf*. exemplo 16)). Pode-se portanto concluir que o gene orf5* é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*, já que sua inativação indireta na cepa SPM501 leva a um fenótipo não-produtor. Ao contrário,
30 o gene orf6* não é um gene essencial à biossíntese da espiramicina I em *S.ambofaciens* (ao contrário, é essencial à produção de espiramicina II e III (conforme exemplo 16)).

EXEMPLO 16: dosagem da produção das espiramicinas I, II e III nas cepas mutantes obtidas.

As diferentes cepas a testar foram cultivadas em cada um dos 7 erlenmeyers chicanados de 500 mL contendo 70 ml de meio MP5 (Pernodet et al., 1993). Os erlens foram inoculados por $2.5 \cdot 10^6$ esporas/ml das diferentes cepas de *S. ambofaciens* e colocadas para crescer a 27°C sob agitação orbital de 250 rpm, durante 72 horas. As culturas correspondentes ao mesmo clone reunidas, eventualmente filtradas sobre filtro plissado, e centrifugadas durante 15 minutos a 7000 rpm. Os diferentes resíduos sobrenadantes foram em seguida armazenados a -30°C.

As dosagens foram feitas por Cromatografia Líquida de Elevado Desempenho (CLHD) por emparelhamento de íons. A análise CLHD do meio de cultura permite determinar precisamente a concentração das três formas de espiramicina. A coluna utilizada (Macherey-Nagel) é cheia com uma fase Nucleosil de sílica enxertada octila. A granulometria é de 5 µm e o tamanho dos poros 100 Å. O diâmetro interno da coluna é de 4,6 mm e seu comprimento é de 25 cm. A fase móvel é uma mistura de tampão H_3PO_4 (pH2,2) e de acetonitrila 70/30 (v/v) contendo 6,25 g/L de perclorato de $NaClO_4$. A análise é feita em regime isocrático com uma vazão fixada em 1 ml/min. A coluna é termo-regulada em 23°C. A detecção é assegurada por espectrofotometria UV a 238 nm. A amostra é refrigerada a +10°C e a quantificação é determinada a partir da área dos picos (por aferição externa). Nessas condições, os tempos de retenção da espiramicina I, II e III são respectivamente de aproximadamente 17, 21 e 30 minutos, conforme pôde ser verificado, graças a uma mostra comercial que compreende as três formas de espiramicina.

A cepa OSC2 possui um fenótipo produtor de espiramicina. É a cepa parenteral utilizada para a obtenção das cepas que possuem uma cassette excisável (conforme exemplo 15). Essa cepa foi, portanto, utilizada como prova positiva de produção das três formas de espiramicina. Essa cepa produz as três formas de espiramicinas conforme foi verificado por CLHP (conforme Figura 19).

O estudo das cepas que possuem uma cassette excisável permite tirar um certo número de conclusões no que refere-se à função dos genes interrompidos. A cepa SPM507 possui o genótipo: orf12::att3Ωaac-. O fenótipo da cepa SPM507 é não-produtor (conforme exemplo 15), pode-se, portanto, concluir que o gene orf12 é um gene essencial à biossíntese da
 5 espiramicina em *S.ambofaciens*. Essa cepa não produz mais espiramicinas conforme foi verificado por CLHP (conforme Figura 22). Esse resultado confirma o caráter essencial do gene orf12 na biossíntese da espiramicina.

A cepa SPM508 possui o genótipo: orf13c::att3Ωaac-. O fenótipo da cepa SPM508 é produtor de espiramicina (conforme exemplo 15), pode-se portanto concluir que o gene orf14c não é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*. Essa cepa produz espiramicinas I, II e III conforme foi verificado por CLHP (conforme Figura 23). Esse resultado confirma que o gene orf13c não é um gene essencial à biossíntese das
 10 espiramicinas I, II e III em *S.ambofaciens*.

A cepa SPM509 possui o genótipo: orf14::att3Ωaac-. O fenótipo da cepa SPM509 é não-produtor, pode-se portanto concluir que o gene orf14 é um gene essencial à biossíntese da espiramicina em *S.ambofaciens*. Essa cepa não produz mais espiramicinas conforme foi verificado por CLHP
 20 (conforme Figura 24). Esse resultado confirma a essencialidade do gene orf14 na biossíntese da espiramicinas.

A cepa SPM501 possui o genótipo: orf6*::att1Ωhyg+. O fenótipo dessa cepa é não-produtor de espiramicinas. Essa cepa não produz mais espiramicinas conforme foi verificado por CLHP (conforme Figura 20). Todavia, os genes orf5* e orf6* (conforme Figura 3) tendo a mesma orientação, o
 25 fenótipo observado pode ser devido a um efeito polar do cassette introduzida em orf6* sobre a expressão de orf5* no operon. Isto deixa pensar que esses genes são co-transcritos. Para responder a essa questão, a cepa SPM502 foi obtida por excisão do cassette introduzida, produzindo uma deleção em fase no gene orf6* e restaurando a expressão de orf5*. Essa cepa possui o
 30 genótipo orf6*::att1 (conforme exemplo 14 e 15). A excisão do cassette deixa somente uma seqüência att "cicatrizial" em fase (conforme exemplo 14). A

cepa SOM502 possui um fenótipo produtor de espiramicina. Todavia, conforme foi provado por CLHP, esta cepa produz apenas espiramicina I e não produz espiramicina II e III (conforme Figura 21). Pode-se portanto concluir desses resultados, que o gene orf5* é um gene essencial à biossíntese da espiramicina em *S. ambofaciens*, já que sua inativação indireta na cepa SPM501 leva a um fenótipo não-produtor de espiramicina (conforme Figura 20). Por outro lado, o gene orf6* não é um gene essencial à biossíntese da espiramicina I em *S. ambofaciens*, já que a inativação desse gene leva a um fenótipo produtor de espiramicina I (conforme Figura 21). Todavia, orf6* é essencial à produção de espiramicina II e III (conforme exemplo 16). O gene orf6* codifica bem, portanto, uma acil-transferase responsável pela modificação do platenólídeo na posição 3 (conforme Figura 1).

EXEMPLO 17: Determinação do ponto de acionamento da tradução de orf10 e melhoria da produção em espiramicinas

17.1 Construção dos plasmídeo pSPM523, pSPM524 e pSPM525

O gene orf10 foi identificado em *streptomyces ambofaciens* e foi designado srmR por Gleistlich et al. (Gleistlich, M. et al., 1992). A inativação do gene orf10 foi feita (conforme exemplo 8). Pôde assim ser mostrado que a cepa resultante não produz mais espiramicinas (conforme exemplo 15). Isto confirma que o gene orf10 é implicado na biossíntese das espiramicinas. A proteína codificada por esse gene é, portanto, bem essencial à biossíntese das espiramicinas. Todavia, a análise da sequência mostra que dois códons ATG situados na mesma fase de leitura podem ser utilizados para a tradução de orf10 (conforme Figura 28). Um dos dois códons possível (o códon o mais a montante) aciona a posição 10656 da sequência dada em SEQ ID N° 1, enquanto que o outro códon possível situado mais a jusante aciona a posição 10809 (conforme SEQ ID N° 1). Antes de testar um eventual efeito da superexpressão de srmR sobre a produção de espiramicina, é importante determinar em uma primeira etapa o ponto de acionamento da tradução.

Com a finalidade de determinar o sítio de acionamento da tradução, três construções que comportam três formas de orf10 foram realizadas.

Estas foram obtidas por PCR, utilizando oligonucleotídeos, comportando seja um sítio de restrição *Hind*III, seja um sítio de restrição *Bam*HI.

O primeiro par utilizado para a amplificação corresponde aos seguintes oligonucleotídeos:

EDR39 :

5'CCCAAGCTTGAGAAGGGAGCGGACATTCATGGCCCGCGCCGAACGC3'

(SEQ ID N° 122) (o sítio *Hind*III figura em negrito)

EDR42 :

5'CGGGATCCGGCTGACCATGGGAGACGGGCGCATCGCCGAGTTCAGC3'

(SEQ ID N° 123) (o sítio *Bam*HI figura em negrito)

- 5 O par de atrações EDR39-EDR42 permite a amplificação de um fragmento de orf10, comportando a ATG situada o mais em 3' (posição 10809 na sequência dada em SEQ ID N° 1) (Conforme Figura 28). O fragmento obtido forma um tamanho de aproximadamente 2 kb e será denominado na sequência orf10 curta, ele não contém o promotor de orf10. Esse
- 10 fragmento de 2kb foi clonado no vetor pGEM-T easy para dar origem ao plasmídeo pSPM520.

O segundo par utilizado para a amplificação corresponde aos seguintes oligonucleotídeos:

EDR40 :

5'CCCAAGCTTGAGAAGGGAGCGGACATTCAATGCTTTGGTAAAGCAC3'

(SEQ ID N° 124) (o sítio *Hind*III figura em negrito)

EDR42 :

5'CGGGATCCGGCTGACCATGGGAGACGGGCGCATCGCCGAGTTCAGC3'

(SEQ ID N° 123) (o sítio *Bam*HI figura em negrito)

- 15 O par de atrações EDR 40-EDR42 permite a amplificação de um fragmento de orf10, comportando o ATG situado além disso em 5' (posição 10656 na sequência dada em SEQ ID N° 1) (conforme Figura 28). Esse fragmento de 2,2 kb denominado na sequência orf10 longo foi clonado no vetor pGEM-T easy para dar origem ao plasmídeo pSPM521, esse plasmí-
- 20 deo não contém o promotor de orf10.

O terceiro par utilizado para a amplificação corresponde aos se-

guintes oligonucleotídeos:

EDR41 :

5'-CCCAAGCTTTCAAGGAACGACGGGGTGGTCAGTCAAGT-3'

(SEQ ID Nº 125) (o sítio *HindIII* figura em negrito)

EDR42 :

5'CGGGATCCGGCTGACCATGGGAGACGGGCGCATCGCCGAGTTCAGC3'

(SEQ ID Nº 123) (o sítio *BamHI* figura em negrito)

O par de atrações EDR41-EDR42 permite a amplificação do gene orf10 com as duas AGT, assim como seu próprio promotor (conforme Figura 28). Esse fragmento de 2,8 kb denominado na sequência "orf10 pro" foi clonado no vetor pGEM-T easy para dar origem ao plasmídeo pSPM522.

O fragmento orf10 pro foi obtido utilizando-se como matriz o DNA cromossômico da cepa OSC2. Os fragmentos orf10 curto e orf10 longo foram obtidos utilizando-se como matriz o DNA do fragmento orf10 pro previamente purificado.

Os enxertos *HindIII*-*BamHI* dos plasmídeos pSPM520, pSPM521 e pSPM522 foram em seguida subclonados no vetor pUWL201 (plasmídeo derivado do plasmídeo pUWL199 (Wehmeier UF, 1995), no qual o fragmento *KpnI*-*BamHI* da região do promotor *ermE* (conforme Bibb et al., 1985, notadamente a Figura 2) portando uma mutação que aumenta a força do promotor (promotor *ermE**) (Bibb et al., 1994) foi introduzido (conforme Doumith et al., 2000)) previamente digerido pelas enzimas *HindIII*-*BamHI*. Assim, foram obtidos três plasmídeos: pSPM523 (derivado do pUWL201 com a forma orf10 curto como enxerto), pSPM524 (derivado do pUWL201 com a forma orf10 longo, como enxerto) e pSPM525 ((derivado do pUWL201 com a forma orf10 pro) (figura 28).

17.2 Transformação da cepa OS49.50 pelas construções pSPM523, pSPM524 e pSPM525

A cepa OS49,50 (cepa interrompida no gene orf10, conforme exemplo 8) foi transformada independentemente por transformação de protoplastos (Kieser, T. et al., 2000) por cada um dos plasmídeos pSPM523, pSPM524 e pSPM525. Uma prova negativa foi igualmente construída,

transformando-se a cepa OS49.50 pelo plasmídeo pUWL201 sem enxerto. Após transformação dos protoplastos, os clones são selecionados por sua resistência ao tiostrepton. A transformação dos clones por cada um dos plasmídeos é verificada por extração desses plasmídeos. Assim, quatro novas cepas foram obtidas: a cepa OSC49.50 pUWL201, o proveniente da transformação pelo plasmídeo pUWL201 sem enxerto, a cepa OSC49.50 pSPM523, proveniente da transformação pelo plasmídeo pSPM523, a cepa OSC49.50 pSPM524, proveniente da transformação pelo plasmídeo pSPM524 e enfim a cepa OS49.50 pSPM525, proveniente da transformação pelo plasmídeo pSPM525.

A produção de espiramicinas de cada uma dessa quatro cepas foi testada por CLHP. Para isto, as diferentes cepas de *streptomyces* a testar foram cultivadas em erlenmeyers com chicanas (erlenmeyers chicanados) de 500 ml contendo 70 ml de meio MP5 (Pernodet et al., 1993). Quando a cepa contém o plasmídeo pUWL201 ou um de seus derivados, são acrescentados 5 µg/ml de tiostrepton. Os erlenmeyers chicanados foram inoculados a uma concentração inicial de $2.5 \cdot 10^6$ esporas/ml das diferentes cepas de *S.ambofaciens* e as culturas foram incubadas a 27 °C sob agitação orbital de 250 rpm durante 96 horas. As células foram em seguida separadas do meio por centrifugação e o sobrenadante foi analisado por CLHP (conforme exemplo 16) para determinar a quantidade de espiramicina produzida. Graças a uma amostra de aferição e à medida do ar dos picos, foi possível determinar a quantidade de cada uma das espiramicinas produzidas por essas cepas. Os resultados dessa análise são apresentados na tabela 39, os dados são expressos em mg por litro do sobrenadante. Os resultados correspondem à produção total em espiramicinas (obtida adicionando-se a produção em espiramicinas I, II e III).

Tabela 39. Produção em espiramicinas das cepas derivadas de OS49.50, (resultados expressos em mg/l)

Cepa	Produção em espiramicinas
OS49,50 pUWL201	0
OS49,50 pSPM523	0
OS49,50 pSPM524	93
OS49,50 pSPM525	149

Conforme mostram os resultados dados na tabela 39, a prova negativa (cepa OS49.50 transformada pelo plasmídeo pUWL201) não produz espiramicina. Quando o plasmídeo pSPM523 (que contém a forma orf10 curto) é introduzido na cepa OS49.50, nenhuma produção de espiramicina é observada. Ao contrário, a presença do plasmídeo pSPM524 (que contém a forma orf10 longo) e do plasmídeo pSPM525 (que contém a form orf10 pro) restaura a produção de espiramicina na cepa hospedeira OS49.50. Assim, só os fragmentos de orf10 contendo o ATG mais a montante permitem restaurar a síntese de espiramicina.

Com a finalidade de confirmar esses resultados, o plasmídeo pSPM521 (plasmídeo pGEM-T easy contendo a forma orf10 longo) foi digerido pela enzima de restrição XhoI (essa enzima possui um sítio único nesse plasmídeo, localizado entre as duas ATG (conforme Figura 28)). As extremidades XhoI foram em seguida tornados extremidade livre por tratamento na enzima de Klenow. O plasmídeo foi então contido sobre ele mesmo por ação da T4 DNA ligase para dar origem ao plasmídeo pSPM527. Caso a ATG a mais a montante (posição 10656 na sequência dada em SEQ ID Nº 1) seja bem utilizada como sítio de iniciação da tradução, esse tratamento acarretará uma defasagem do âmbito de leitura ao nível do sítio XhoI e terá por efeito a produção de uma proteína que não apresenta atividade ativadora. Ao contrário, se a iniciação da tradução ocorrer ao nível da ATG mais a jusante (posição 10809 na sequência dada em SEQ ID Nº 1), esse tratamento deverá ter pouco efeito ou quase nenhum sobre a expressão de Orf10 (considerando-se a localização do ponto de acionamento da transcrição) e nenhum efeito sobre a proteína produzida.

O enxerto BamHI-HindIII de pSPM527 foi em seguida subclonado no vetor pUWL201 para dar origem ao plasmídeo pSPM528. Esse plasmídeo foi introduzido na cepa OS49.50 e um clone que possui o plasmídeo desejado foi mais particularmente selecionado. A produção de espiramicina da cepa resultante foi em seguida testada por CLHP (conforme exemplo 16 e acima). Contrariamente ao que fora observado com o plasmídeo pSPM524 (conforme tabela 39), a presença do plasmídeo pSPM528 na cepa OS49.50 não restabelece a produção de espiramicina. Isto confirma que o acionamento da tradução do gene orf10 é a ATG situado mais a jusante (ATG 1, conforme Figura 28).

17.3 Melhoria da produção em espiramicinas da cepa OSC2 de *S.ambofaciens*

A fim de testar o efeito da superexpressão do gene orf10 sobre a produção de espiramicinas, os plasmídeos pSPM523, pSPM524, pSPM525 e pSPM528 foram introduzidos na cepa OSC2. Para isto, protoplastos da cepa OSC2 foram transformados (Kieser, T. et al., 2000) independentemente por cada um dos plasmídeos pSPM523, pSPM524, pSPM525 e pSPM528. Uma prova negativa foi também realizada, transformando a cepa OSC2 pelo plasmídeo pUWL201 sem enxerto. Após transformação de protoplastos, os clones são selecionados por sua resistência ao tiosstrepton. Assim, cinco novas cepas foram obtidas: a cepa OSC2 pUWL201, proveniente da transformação pelo plasmídeo pUWL201 sem enxerto, a cepa OSC2 pSPM523, proveniente da transformação pelo plasmídeo pSPM523, a cepa OSC2 pSPM524, proveniente da transformação pelo plasmídeo pSPM524, a cepa OSC2 pSPM525, proveniente da transformação pelo plasmídeo pSPM525, e enfim a cepa OSC2 pSPM528, proveniente da transformação pelo plasmídeo pSPM528. A produção em espiramicinas dessas cepas foi então analisada por CLHP (do mesmo modo que no exemplo 17.2). A análise da produção em espiramicinas da cepa OSC2 foi também feita em paralelo para comparação. Os resultados dessa análise são apresentadas na tabela 40, os dados são expressos em mg por litro do sobrenadante. Os resultados correspondem à produção total em espiramicinas (obtida adicionando-se a pro-

dução em espiramicina I, II e III).

Tabela 40. Produção em espiramicinas das cepas derivadas de OSC2 (resultados expressos em mg/l).

Cepa	Produção em espiramicinas
OSC2	69
OSC2 pUWL201	103
OSC2 pSPM523	19
OSC2 pSPM524	135
OSC2 pSPM525	278
OSC2 pSPM528	68

Assim, foi observado que a presença do plasmídeo pSPM524 ou do plasmídeo pSPM525 aumenta significativamente a produção em espiramicinas da cepa OSC2. Isto demonstra bem que a superexpressão de Orf10 tem um efeito positivo sobre a produção em espiramicinas. O plasmídeo pSPM528 não tem, ao contrário, efeito sobre a produção em espiramicinas.

Os plasmídeos pSPM525 e pUWL201 foram do mesmo modo introduzidos na cepa SPM502 (conforme exemplo 14). Assim duas novas cepas foram obtidas: a cepa SPM502 pUWL 201, proveniente da transformação pelo plasmídeo pUWL201 sem enxerto, e a cepa SPM502 pSPM525, proveniente da transformação pelo plasmídeo pSPM525.

Uma amostra da cepa SPM502 pSPM525 (essa cepa contém o plasmídeo pSPM525, conforme acima) foi depositada junto à Collection Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, em 26 de fevereiro de 2003 sob o número de registro I-2977.

A produção em espiramicinas das cepas SPM502 pUWL201 e SPM502 pSPM525 foi analisada por CLHP (da mesma maneira que no exemplo 17.2). A análise da produção em espiramicinas da cepa SPM502 foi também feita em paralelo para comparação. Os resultados dessa análise são apresentados na tabela 41, os dados são expressos em mg por litro do sobrenadante. Os resultados correspondem à produção em espiramicinas I. Com efeito, nenhuma dessas cepas produz espiramicina II e III.

Tabela 41. Produção em espiramicina I das cepas derivadas de SPM502, (resultados expressos em mg/l)

Cepa	Espiramicina I
SPM502	47
SPM502 pUWL201	72
SPM502 pSPM525	130

Assim, pôde ser observado que a superexpressão do gene orf10 na cepa SPM502 aumenta de forma importante a produção de espiramicina I.

EXEMPLO 18: construção de uma banco de DNA genômico da cepa OSC2 de *streptomyces ambofaciens* em E.coli no cosmídeo pWED2.

18.1 Construção do cosmídeo pWED2

Com a finalidade de facilitar a inativação de genes em *streptomyces*, um cosmídeo que porta a seqüência oriT do plasmídeo RK2 (o que permite sua introdução por conjugação em *streptomyces* a partir de uma cepa apropriada de E. coli) e portando também um gene de resistência a um antibiótico conferindo um fenótipo marcável em *streptomyces* foi construído. Esse cosmídeo contendo grandes enxertos de DNA genômica de *streptomyces ambofaciens* pode ser utilizado em experiências de inativação de genes.

Para construir esse vetor, um cassette pac-oriT (fragmento EcoRV) foi introduzido no cosmídeo pWED1 (Gourmelen et., 1998), derivado do cosmídeo pWED15 (Wahl et al., 1987), ao nível do sítio único HpaI. O cassette pac-oriT foi obtida por PCR. Para isto, o gene pac foi amplificado por PCR a partir do plasmídeo pVF 10.4 (Vara et al., 1985; Lacalle et al., 1989), utilizando como primeira atração, a atração A (de seqüência

5' CCAGTAGATATCCCGCCAACCCGGAGCTGCAC-3' (SEQ ID N° 126),

o sítio de restrição EcoRV foi sublinhado e a seqüência em negritode 20 nucleotídeos corresponde a uma região a montante do promotor do gene pac) e como segunda atração, a atração B (de seqüência

5' GAAAAGATCCGTCATGGGGTCGTGCGCTCCTT-3' (SEQ ID N° 127),

que comporta em sua extremidade 5' uma seqüência de 12 nucleotídeos

correspondentes ao começo da seqüência oriT (sublinhado duplo) e uma seqüência de 20 nucleotídeos (em traços cheios) correspondente ao fim do gene pac (conforme Figura 29, 1ª PCR).

O gene oriT foi amplificado por PCR a partir do plasmídeo pPM803 (Mazodier, P. et al., 1989), utilizando como primeira atração, a atração C (de seqüência

5'-CACGACCCCATGACGGATCTTTTCCGCTGCAT-3'

(SEQ ID Nº 128)), que comporta em sua extremidade 5' uma seqüência

De 12 nucleotídeos correspondendo a uma seqüência a jusante da seqüência codificante do gene pac (em traços cheios) e uma seqüência de 20 nu-

cleotídeos correspondendo ao começo da seqüência oriT) e como segunda atração, a atração D (de seqüência

5' GAGCCGGATATCATCGGTCTTGCTCGT-3' (SEQ ID Nº 129))

que comporta o sítio de restrição EcoRV (sublinhado duplo) (conforme Figura 29, 2ª PCR).

O produto de amplificação obtido com as atrações A e B e aquele obtido com as atrações C e D têm em uma de suas extremidades, uma seqüência comum de 24 nucleotídeos. Uma terceira PCR foi realizada, misturando-se os dois produtos de amplificação obtidos anteriormente e utilizando-se como atrações, as atrações A e D (conforme Figura 29, 3ª PCR). Isto permitiu obter um produto de amplificação correspondente ao conjunto pac+oriT. Esse fragmento pac-oriT foi então clonado no vetor pGEM-T Easy (comercializado pela sociedade Promega (Madison, Wisconsin, USA)), o que permitiu obter o plasmídeo pGEM-T-pac-oriT. O enxerto desse plasmídeo foi em seguida subclonado no cosmídeo pWEDI. Para isto, o plasmídeo pGEM-pac-oriT foi digerido pela enzima EcoRV e o enxerto EcoRV contendo o conjunto pac+oriT foi inserido no cosmídeo pWED1 aberto previamente pela enzima HpaI. O cosmídeo assim obtido foi batizado pWED2 (conforme Figura 30).

Esse cosmídeo permite facilitar a inativação de genes em *streptomyces*. Com efeito, ele porta a seqüência oriT (o que permite sua introdução por conjugação em *streptomyces*, a partir de uma cepa apropriada

de *E.coli*), mas também um gene de resistência a um antibiótico, conferindo um fenótipo marcável em *streptomyces*. Esse cosmídeo, contendo grandes enxertos de DNA genômico de *streptomyces ambofaciens* pode ser utilizado em experiências de inativação de genes.

5 Assim, um cosmídeo derivado de pWED2, contendo o gene alvo, poderá, por exemplo, ser introduzido na cepa de *E. coli* KS272, contendo o plasmídeo pKOBEG (conforme Chaverroche et al., 2000) e um cassette será introduzida no gene alvo, segundo a técnica descrita por Chaverroche et al., 2000. O cosmídeo obtido por essa técnica (cosmídeo no qual o gene
10 alvo é inativado), poderá em seguida ser introduzido em uma cepa de *E.coli*, tal como a cepa S17.1 ou qualquer outra cepa, permitindo transferir por conjugação dos plasmídeos contendo a sequência oriT em direção ao *streptomyces*.

Após conjugação entre *E.coli* e *streptomyces*, clones de
15 *streptomyces* nas quais a cópia primitiva do gene alvo terá sido substituída pela cópia interrompida poderão ser obtidas conforme descrito no exemplo 2.

O gene de resistência expresso em *streptomyces* presente sobre esse novo cosmídeo é o gene *pac* de *streptomyces alboniger* (Vara, J. et al.,
20 1985; Lacalle et al., 1989) que codifica a puromicina N-acetil transferase e que confere a resistência à puromicina. Quando de experiências de inativação de gene, buscar-se-ão clones nos quais um duplo acontecimento de recombinação terá ocorrido. Esses clones terão eliminado o cosmídeo pWED2 e serão, portanto, tornados de novo sensíveis à puromicina.

25 18.2 Construção de um banco de DNA genômico da cepa OSC2 de *streptomyces ambofaciens* em *E.coli* no cosmídeo pWED2

O DNA genômico da cepa OSC2 de *streptomyces ambofaciens* foi digerido parcialmente pela enzima de restrição BamHI, de maneira a serem obtidos fragmentos de DNA de tamanho compreendido entre aproxima-
30 damente 35 e 45 kb. Esses fragmentos foram em seguida clonados no cosmídeo pWED2, este tendo sido previamente digerido por BamHI, depois tratado pela fosfatase alcalina. A mistura de ligação foi em seguida encapsi-

dada *in vitro* em partículas de fagos lambda, graças ao sistema "*Packagene*[®] *Lambda DNA packaging system*" comercializado pela sociedade Promega (Madison, Wisconsin, USA) seguindo as recomendações do fabricante. As partículas fágicas obtidas foram utilizadas para infectar a cepa SURE[®] de *E. coli* comercializada pela sociedade Stratagene (LaJolla, Califórnia, USA). A seleção dos clones foi feita sobre meio LB + ampicilina (50 µg/ml), o cosmídeo pWED2, conferindo a resistência à ampicilina.

EXEMPLO 19: isolamento de cosmídeo do novo banco que cobre a região da via de biossíntese das espiromicinas. Sob clonagem e seqüenciação de fragmentos desses cosmídeos.

19.1 Hibridização sobre colônias do banco genômico de *streptomyces ambofaciens* OSC2

Cosmídeos do novo banco de *streptomyces ambofaciens* OSC2 (conforme exemplo 18), cobrindo os orf*1 ao orf10* ou uma parte ou a totalidade dos orf1 a orf25c, ou uma região mais a montante do orf25c foram isolados. Para isso, hibridizações sucessivas sobre as colunas de *E. coli* obtidas segundo o exemplo 18 foram realizadas, utilizando-se as 3 seguintes sondas (conforme Figura 31):

- a primeira sonda utilizada corresponde a um fragmento de DNA de aproximadamente 0,8kb amplificado por PCR, utilizando como matriz o cosmídeo pSPM5 e as seguintes atrações:

ORF23c: 5'-ACGTGCGCGGTGAGTTCGCCGTTGC-3' (SEQ ID N° 130) e
ORF25c: 5'-CTGAACGACGCCATCGCGGTGGTGC-3' (SEQ ID N° 131).

O produto de PCR assim obtido contém um fragmento do começo do orf23c, orf24c por inteiro e o fim do orf25c (conforme Figura 31, sonda I);

- a segunda sonda utilizada corresponde a um fragmento de DNA de aproximadamente 0,7kb amplificado por PCR, utilizando como matriz o DNA total da cepa *S. ambofaciens* ATCC23877 e as seguintes atrações:

ORF1*c: 5'-GACCACCTCGAACCGTCCGGCGTCA-3' (SEQ ID N° 132) e
ORF2*c: 5'-GGCCCGGTCCAGCGTGCCGAAGC-3' (SEQ ID N° 133).

O produto de PCR assim obtido contém um fragmento do fim do orf1*c e do começo da orf2*c (conforme Figura 31, sonda II);

- a terceira sonda utilizada corresponde a um fragmento EcoRI-BamHI de aproximadamente 3Kb contendo os orf1, orf2 e orf3 e obtido por digestão do plasmídeo pOS49.99 (cf Figura 31, sonda III).

Aproximadamente 2000 clones do banco obtido no exemplo 18.2 foram transferidos sobre filtro para hibridização sobre colônias, segundo as técnicas clássicas (Sambrook et al., 1989).

- A primeira sonda (conforme Figura 31, sonda I) foi marcada com ^{32}P pela técnica do "random priming" (Kit comercializado pela sociedade Roche) e utilizadas para a hibridização sobre 2000 clones do banco, após transferência sobre filtro. A hibridização foi feita a 65°C no tampão descrito por Church & Gilbert (Church & Gilbert, 1984). Duas lavagens foram feitas em 2x SSC, 0,1% SDS a 65°C, a primeira durante 10 minutos e a segunda durante 20 minutos e uma terceira lavagem de uma duração de 30 minutos foi em seguida feita em 0,2X SSC, 0,1% SDS a 65°C. Nessas condições de hibridização e de lavagem, 20 clones dentre os 2000 hibridados apresentavam um forte sinal de hibridização com a primeira sonda. Esses 20 clones foram cultivados em meio LB+ ampicilina (50 µg/ml) e os 20 cosmídeos correspondentes foram extraídos por lise alcalina (Sambrook et al., 1989), depois foram digeridos pela enzima de restrição BamHI. Os produtos de digestões foram em seguida separados sobre gel de agarose, transferidos sobre membrana de Nylon e hibridados pela primeira sonda (cfr. acima: o produto de PCR ORF23c-ORF25c, sonda I) nas mesmas condições que acima.
- Doze cosmídeos possuíam um fragmento pSPM34, pSPM35, pSPM36, pSPM37, pSPM38, pSPM39, pSPM40, pSPM41, pSPM42, pSPM43, pSPM44 e pSPM45. Os perfis, após digestão por BamHI, desses 12 cosmídeos foram comparados entre si e com aquele do cosmídeo pSPM5. Além disso, experiências de amplificação por PCR, utilizando diferentes atrações correspondentes a diferentes genes já identificados na região orf1-orf25c permitiram posicionar o enxerto de certos desses cosmídeos uns em relação aos outros e determinar também a localização desses enxertos na região

orf1-orf25c já conhecida (conforme Figura 32). O cosmídeo pSPM36 foi mais particularmente escolhido, pois era capaz de conter uma grande região a montante do orf25c (conforme Figura 31 e 32).

Em uma segunda etapa, e utilizando as mesmas condições que
5 aquelas descritas acima, os 200 clones do banco de *streptomyces ambofaciens* OSC2 foram hibridados com a segunda sonda correspondente ao produto de PCR: ORF1*c-ORF2*c (conforme Figura 31, sonda II). Essa hibridização permitiu isolar cosmídeos, cujo enxerto fica situado na região de orf1*c ao orf10*c. Nas condições de hibridização utilizadas, 16 clones dentre
10 os 2000 hibridados apresentavam um forte sinal de hibridização com essa segunda sonda. Esses 16 clones foram cultivados em meio LB+ ampicilina (50 µg/ml) e os 16 cosmídeo correspondentes foram extraídos por lise alcalina (Sambrook et al., 1989), depois foram digeridos pela enzima de restrição BamHI. Os perfis de digestão (após digestão por BamHI) desses 16 cosmí-
15 deos foram comparados entre si e com aquele do cosmídeo pSPM7. Experiências de amplificação por PCR com as atrações ORF1*c e ORF2*c permitiram escolher dois cosmídeos que continham bem os genes orf1*c e orf2*c e cujos perfis possuíam faixas comuns, mas também faixas diferentes. Além disso, outras experiências de amplificação por PCR, utilizando diferentes
20 atrações correspondentes a diferentes genes já identificados na região orf1*c à orf10*c permitiram posicionar o enxerto desses cosmídeos uns em relação aos outros e determinar também a localização desses enxertos na região orf1*c ao orf10*c já conhecida (conforme Figura 32). Os dois cosmídeos mais particularmente selecionados foram denominados pSPM47 e
25 pSPM48 (conforme Figura 32).

Utilizando-se as mesmas condições que aquelas descritas acima, os 2000 clones do banco de *streptomyces ambofaciens* OSC2 foram também hibridados com a terceira sonda correspondente ao fragmento de DNA EcoRI-BamHI do plasmídeo pOS49.99 (conforme Figura 31 sonda III).
30 Essa hibridização permitiu isolar os cosmídeo contendo a região do orf1 até orf3 e capazes de conter seja uma grande parte dos genes dos PKS, seja uma grande parte dos genes orf1 ao orf25c da via de biossíntese das espi-

ramicinas. Nessas condições de hibridização, 35 clones dentre os 2000 hibridados apresentavam um forte sinal de hibridização com a terceira sonda. Esses 35 clones foram cultivados em meio LB+ ampicilina (50 µg/ml) e os 35 cosmídeos correspondentes foram extraídos por lise alcalina (Sambrook et al., 1989), depois digeridos pela enzima de restrição BamHI. Os perfis, após digestão por BamHI, desses 35 cosmídeos foram comparados entre si e com aquele do cosmídeo pSPM5. Além disso, experiências de amplificação por PCR, utilizando diferentes atrações correspondentes a diferentes genes já identificados na região orf1 ao orf25c permitiram verificar que esses cosmídeos continham bem enxertos provenientes da região orf1 ao orf25c e posicionar os enxertos desses cosmídeos uns em relação aos outros e em relação à região orf1 ao orf25c já conhecida (conforme Figura 32). Cinco cosmídeos foram mais particularmente selecionados, foram denominados pSPM50, pSPM51, pSPM53, pSPM55 e pSPM56 (conforme Figura 32).

19.2 Subclonagem e seqüenciação de uma parte do enxerto do cosmídeo pSPM36

A sonda de aproximadamente 0,8 kb obtida por PCR com as atrações ORF23c e ORF25c (conforme acima e Figura 31, sonda I) foi também utilizada em experiências de Southern Blot sobre o DNA total de *S. ambofaciens* OSC2 digerido pela enzima PstI. Nas condições de hibridização descrita acima, essa sonda revela um fragmento único PstI de aproximadamente 6 kb, quando hibridizado sobre o DNA total de *S. ambo* OSC2 digerido pela enzima PstI. Existe um local PstI ao nível do orf23c (conforme SEQ ID Nº 80), mas nenhum outro local PstI até à extremidade (site BamHI) da seqüência conhecida (conforme SEQ ID Nº 1). Esse fragmento PstI-BamHI tem um tamanho de aproximadamente 1.4 kb. O fragmento PstI de 6 kb hibridizado sobre o DNA total de *S. ambofaciens* digerido pela enzima PstI contém, portanto, uma região de aproximadamente 4,6 kb situada a montante de orf25c. Essa região é capaz de conter outros genes, cujos produtos são implicados na via de biossíntese da espiramicina. Foi verificado por digestão que o cosmídeo pSPM36 continha bem esse fragmento pstI de 6 kb. Esse fragmento foi isolado a partir de pSPM36, com a finalidade de determi-

nar a seringa mais a montante de orf25c. Para isso, o cosmídeo pSPM36 foi digerido pela enzima de restrição PstI. O fragmento PstI-PstI de um tamanho de aproximadamente 6 kb foi isolado por eletropurificação, a partir de um gel de agarose 0,8%, depois clonado no vetor pBK-CMV (4512bp) (comercializado pela sociedade Stratagene (La Jolla, Califórnia, USA). O plasmídeo assim obtido foi denominado pSPM58 (conforme Figura 33) e a sequência de seu enxerto foi determinada. A seringa desse enxerto é apresentada em SEQ ID N° 134.

Todavia, toda a sequência não foi determinada e subsiste um orifício de aproximadamente 450 nucleotídeos, a parte da sequência não-determinada foi anotada por uma sucessão de "N" na sequência correspondente.

19.3 Análise das novas sequências nucleotídicas, determinação das fases abertas de leitura e caracterização dos genes implicados na biossíntese das espiramicinas

A sequência do enxerto do cosmídeo pSPM58 obtida foi analisada graças ao programa FramePlot (Ishikawa J. & Hotta K. 1999). Isto permitir identificar, dentre as fases abertas de leitura, as fases abertas de leitura apresentando um uso dos códons típico de *streptomyces*. Essa análise permitiu determinar que esse enxerto comporta 4 novos ORFs a montante da orf25c (conforme Figura 34). Esses genes foram denominados orf26 (SEQ ID N° 107), orf27 (SEQ ID N° 109), orf28c (SEQ ID N° 111, a sequência desse orf não foi determinada completamente já que um orifício de aproximadamente 450 nucleotídeos subsiste, quando da seqüenciação do enxerto de pSPM58, esses 450 nucleotídeos figuram sob a forma de uma série de "N" na sequência SEQ ID N° 111) e orf29 (a sequência deste último orf era incompleta nesse enxerto). O "c" acrescentado no nome do gene significativo para a ORF em questão que a sequência codificante está orientação inversa (conforme Figura 34).

As sequências protéicas deduzidas dessas fases abertas de leitura foram comparadas com aquelas presentes em diferentes bases de dados graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et

al., 1997), CD-serach, (esses dois programas são acessíveis notadamente junto ao National Center for Biotechnology Information (NCBI) (Bethesda, Maryland, USA)), FASTA ((Pearson W.R & D.J. Lipman, 1988) e (Pearson W.R., 1990) (esse programa é acessível notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França). Essas comparações permitiram formular hipóteses sobre a função dos produtos desses genes e identificar aqueles capazes de serem implicados na biossíntese de espiamicinas.

19.4 Subclonagem e seqüenciação de uma outra do enxerto do cosmídeo pSPM36

A sonda de aproximadamente 0,8 kb obtida por PCR com as atrações ORF23c e ORF25c (conforme acima e Figura 31, sonda I) foi também utilizada em experiências de Southern Blot sobre o DNA total da cepa OSC2 digerido pela enzima *StuI*. Nas condições de hibridização descrita acima para essa sonda, essa sonda revela um fragmento único *StuI* de aproximadamente 10 kb, quando hibridizado sobre o DNA total da cepa OSC2 digerido pela enzima *StuI*. Considerando-se a presença de um sítio *StuI* no orf23c (conforme SEQ ID N° 80) e da localização desse sítio em relação ao sítio *PstI*, esse fragmento de 10 kb inclui a totalidade do fragmento *PstI* anteriormente estudado (enxerto de pSPM58) e permite ter acesso a uma região não ainda estudada de aproximadamente 4 kb (conforme Figura 33). Foi verificado por digestão que o cosmídeo pSPM36 continha bem esse fragmento *StuI* do 10 kb. Esse fragmento foi isolado a partir do cosmídeo pSPM36, com a finalidade de determinar a seqüência do fim de orf29 e de outros genes mais a montante de orf29. Para isto, o cosmídeo pSPM36 foi digerido pela enzima de restrição *StuI*. O fragmento *StuI*-*StuI* de um tamanho de aproximadamente 10 kb foi isolado por eletropurificação a partir de um gel de agarose 0,8%, depois clonado no vetor pBK-CMV (4512bp) (comercializado pela sociedade Stratagene (LaJolla, Califórnia, USA)). O plasmídeo assim obtido foi denominado pSPM72 (conforme Figura 33). Este último foi em seguida digerido por *EcoRI* (sítio *EcoRI* no enxerto de pSPM58) e *HindIII* (sítio *HindIII* no sítio de clonagem múltipla do vetor, imediatamente após o sítio *StuI* da extremidade do enxerto) (conforme Figura 33). O frag-

mento de DNA EcoRI-HindIII assim obtido corresponde a um fragmento do enxerto do plasmídeo pSPM72 (conforme Figura 33) e foi subclonado no vetor pBC-SK+ (comercializado pela sociedade Stratagene (La Jolla, Califórnia, USA)) digerido previamente por EcoRI e HindIII. O plasmídeo assim obtido foi denominado pSPM73 e a seqüência de seu enxerto foi determinada. A seqüência desse enxerto é apresentada em SEQ ID N° 135.

Uma ligação das seqüências dos enxertos de pSPM58 e pSPM73 está presente em SEQ ID N° 106. Essa seqüência parte do sítio PstI ao nível da orf23c (conforme SEQ ID N° 80) e vai até o sítio StuI ao nível do orf32c (conforme Figura 34). A seqüência do orf28c (SEQ ID N° 111) não estando completa (conforme exemplo 19.3), uma região de aproximadamente 450 nucleotídeos não é seqüenciada, esses 450 nucleotídeos figuram sob a forma de uma série de "N" na seqüência SEQ ID N° 106)

19.5 Análise das novas seqüências de nucleotídeos, determinação das fases abertas de leitura e caracterização dos genes implicados na biossíntese de espiramicinas

A seqüência parcial do enxerto do cosmídeo pSPM73 obtida foi analisada graças ao programa FramePlot (Ishikawa J & Hotta K., 1999). Isto permitiu identificar, dentre as fases abertas de leitura, as fases abertas de leitura apresentando um uso de códons típica de *streptomyces*. Essa análise permitiu determinar que esse enxerto comporta 4 ORFs, uma incompleta e três completas (conforme Figura 34). A ORF incompleta corresponde à parte 3' da seqüência codificante de orf29, o que permitiu completar a seqüência desse gene, graças à seqüência parcial desse mesmo orf obtida quando da seqüenciação do enxerto do plasmídeo pSPM58 (conforme exemplo 19.2 e 19.3), o conjunto das duas seqüências permitiu assim obter a seqüência completa de orf29. Os 4 genes foram assim chamados orf29 (SEQ ID N° 113), orf30c (SEQ ID N° 115), orf31 (SEQ ID N° 118) e orf32c (SEQ ID N° 120). O "c" acrescentado no nome do gene significa para o ORF em questão indica que a seqüência codificante está na orientação inversa (conforme Figura 34).

As seqüência protéicas deduzidas dessas fases abertas de leitura

ra (SEQ ID N° 114 para orf29, SEQ ID N° 116 e 117 para orf30c, SEQ ID N° 119 para orf31 e SEQ ID N° 121 para orf32c) foram comparadas com aquelas presentes em diferentes bases de dados, graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et al, 1997), CD-search, (esses dois programas são acessíveis notadamente junto ao National Center for Biotechnology Information (NCBI) (Bethesda, Maryland, USA)), FASTA ((Pearson W.R & D.J. Lipman, 1988) e (Pearson W.R., 1990) (esse programa é acessível notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França). Essas comparações permitiram formular hipóteses sobre a função dos produtos desses genes e identificar aqueles capazes de ser implicados na biossíntese de espiramicinas.

19.6. Subclonagem e seqüenciação de uma terceira parte do enxerto do cosmídeo pSPM36.

Uma sonda (fragmento de DNA de 0,8 kb) correspondente a uma seqüência interna ao orf32c foi obtida por PCR, utilizando como matriz o DNA total da cepa de *streptomyces ambofaciens* e as seguintes atrações:

KF36: 5'- TTGCCGTAGCCGAGGACCAGCG-3' (SEQ ID N° 151) e

KF37: 5'- CACATGGCCCTGGAGGACCCTG-3' (SEQ ID N° 152).

O produto PCR assim obtido representa uma seqüência interna do orf32c. Essa sonda foi utilizada em experiências de Southern Blot sobre o DNA cromossômico da cepa OSC2 e sobre o DNA dos cosmídeos pSPM36 digeridos pela enzima PstI. Utilizando-se as mesmas condições de hibridização que aquelas descritas acima (conforme exemplo 19.1), essa sonda revela dois fragmentos PstI de aproximadamente 3,4 kb e 2,5 kb, quando hibridizada sobre o DNA total da cepa OSC2 e sobre o DNA do cosmídeo pSPM36 digeridos pela enzima PstI. Considerando-se a presença de um sítio PstI na sonda utilizada, podem ser explicados esses resultados. O primeiro fragmento de DNA que tem um tamanho de aproximadamente 3,4 kb é um fragmento cuja seqüência já é conhecida na totalidade. A seqüência do fragmento de 2,5 kb é conhecida apenas parcialmente, sobre uma região de 700 pb. Esse fragmento foi isolado a partir do cosmídeo pSPM36 com a finalidade de determinar a seqüência do fim da orf32c e de outros genes a

montante deste. Para isto, o cosmídeo pSPM36 foi digerido pela enzima de restrição PstI. O fragmento PstI-PstI de um tamanho de aproximadamente 2,5 kb foi isolado por purificação a partir de um gel de agarose 0,8%, depois clonado no vetor pBK-CMV (4518bp) (comercializado pela sociedade Strata-
 5 gene (La Jolla, Califórnia, USA)). O plasmídeo assim obtido foi denominado pSPM79 (conforme Figura 41) e a seqüência de seu enxerto foi determinada.

A seqüência do orf28c (SEQ ID N° 111) não era completa (conforme exemplo 19.3). Com efeito, uma região de aproximadamente 450 nucleotídeos não pudera ser determinada, esses 450 nucleotídeos figuram sob
 10 a forma de uma série de "N" na seqüência SEQ ID N° 106. A seqüência da totalidade da região que falta foi determinada por nova seqüenciação dessa região. A seqüência dos enxertos de pSPM58 e pSPM73 foi, portanto, determinada por inteiro. A seqüência completa da parte codificante do orf28c é
 15 apresentada em SEQ ID N° 141 e a proteína deduzida dessa seqüência em SEQ ID N° 142. A seqüência do enxerto do plasmídeo pSPM79 é apresentada em SEQ ID N° 161.

Uma ligação das seqüências dos enxertos de pSPM58, pSPM73 e pSPM79 é apresentada em SEQ ID N° 140 (conforme Figura 41). Essa
 20 seqüência parte do sítio PstI ao nível do orf23c (conforme SEQ ID N° 80) e vai até o sítio PstI ao nível do orf34c (conforme Figura 41).

19.7 Análise das novas seqüências nucleotídicas, determinação das fases abertas de leitura e caracterização dos genes que podem ser implicados na biossíntese das espiramicinas

25 A seqüência do enxerto do plasmídeo pSPM79 obtida foi analisada, graças ao programa Frame Plot (Ishikawa J & Hotta K., 1999). Isto permitiu identificar, dentre as fases abertas de leitura, as fases abertas de leitura que apresentam um uso dos códons típico de *streptomyces*. Essa análise permitiu determinar que esse enxerto comporta 3ORFs, dois incom-
 30 pletos (orf32c e orf34c) e um completo (orf33) (conforme Figura 41).

O primeiro ORF incompleto corresponde à parte 5' da seqüência codificante de orf32c. Isto permitiu completar a seqüência desse gene, gra-

ças à sequência parcial desse mesmo orf obtido quando da seqüenciação do enxerto do plasmídeo pSPM73, (conforme exemplo 19.4 e 19.5), o conjunto das duas seqüências permitiu assim obter a seqüência completa de orf32c. (SEQ ID N° 145). O "c" acrescentado no nome do gene significa para o ORF em questão que a seqüência codificante está na orientação inversa (conforme Figura 41). O orf completo foi denominado orf33 (SEQ ID N° 147). O terceiro ORF foi denominado orf34 (SEQ ID N° 149). O "c" acrescentado no nome do gene significa para o ORF em questão que a seqüência codificante está na orientação inversa (conforme Figura 37). As comparações feitas entre o produto desse orf e os bancos de dados sugerem que a parte C-terminal dessa proteína não está no produto deduzido da seqüência nucleotídica e, portanto, que esse orf seria mais longo e continuaria fora da região seqüenciada.

As seqüências protéicas deduzidas dessas fases abertas de leitura foram comparadas com aquelas presentes em diferentes bases de dados graças a diferentes programas: BLAST (Altschul et al., 1990) (Altschul et al., 1997), CD-search (esses dois programas são acessíveis notadamente junto ao National Center for Biotechnology Information (NCBI) (Bethesda, Maryland, USA)), FASTA ((Pearson W. R & D.J. Lipman, 1988) e (Pearson W. R., 1990) (esse programa é acessível notadamente junto ao centro de pesquisas INFOBIOGEN, Evry, França). Essas comparações permitiram formular hipóteses sobre a função dos produtos desses genes e identificar aqueles capazes de serem implicados na biossíntese de espiramicinas.

EXEMPLO 20: Análise da produção de intermediários de biossíntese de espiramicinas

20.1 Preparação das amostras

As diferentes cepas a testar foram cultivadas cada uma em 7 erlenmeyers chicanados de 500 ml, contendo 70 ml de meio MP5 (Pernodet et al., 1993). Os erlens foram inoculados por $2.5 \cdot 10^6$ esporas/ml das diferentes cepas de *S. ambofaciens* e colocadas para crescer a 27°C sob agitação orbital de 250 rpm durante 96 horas. As culturas correspondentes ao mesmo clone são reunidas, eventualmente filtradas sobre filtro plissado, e

centrifugadas 15 minutos a 7000 rpm.

O pH do mosto é em seguida ajustado a 9 com a soda e o sobrenadante é extraído pela metil iso-butil cetona (MIBK). A fase orgânica (MIBK) é então recuperada e evaporada. O extrato seco é em seguida retomado em 1 ml de acetonitrila, depois diluído em 1/10 (100 µl qsp 1 ml com água), antes de ser utilizado para as análises em cromatografia líquida acoplada à espectrometria de massa (CL/SM).

20.2 Análise das amostras em CL/SM

As amostras foram analisadas em CL/SM com a finalidade de determinar a massa dos diferentes produtos sintetizados pelas cepas a testar.

A coluna de cromatografia líquida de elevado desempenho utilizada é uma coluna Kromasil C8 150*4,6mm, 5mm, 100 Å.

A fase móvel é um gradiente constituído de uma mistura de acetonitrila e de uma solução aquosa de ácido trifluoroacético a 0,05%, a vazão é fixada em 1 ml/min. A temperatura do forno da coluna é mantida em 30°C.

A detecção UV na saída da coluna é feita com dois comprimentos de onda diferentes: 238 nm e 280 nm.

O espectrômetro de massa acoplado à coluna de cromatografia é um aparelho Simples Quadripolo comercializado pela sociedade Agilent, com tensões de cone a 30 e a 70V.

20.3 Análise dos intermediários de biossíntese produzidos pela cepa OS49.67

A sonda OS49.67 na qual o gene orf3 é inativado por uma deleção em fase não produz espiamicinas (conforme exemplo 6 e 15).

Uma amostra foi preparada segundo o método descrito acima (conforme parágrafo 20.1) e foi analisada por CL/SM conforme descrito acima (conforme o parágrafo 20.2).

Mais particularmente, a análise por cromatografia foi feita em gradiente de solvente com como fase móvel: 20% de acetonitrila do tempo T = 0 a 5 min, depois subida linear a 30% a T = 35 minutos, seguida de um

patamar até $T = 50$ minutos.

Nessas condições, dois produtos foram mais particularmente observados: um absorvente com 238 nm (tempo de retenção de 33,4 minutos) e um absorvente com 280 nm (tempo de retenção de 44,8 minutos) (conforme Figura 35). A Figura 35 mostra a superposição dos cromatogramas CLHP realizados em 238 e 280 nm (alto), assim como os espectros UV das moléculas purificadas em 33,4 minutos e 44,8 minutos (embaixo).

As condições de análises em espectrometria de massa acoplada eram as seguintes: a varredura foi feita em modo scan, abrangendo uma faixa de massa compreendida entre 100 e 1000 Da. O ganho do eletromultiplicador era de 1V. Com referência à fonte eletrospray, a pressão do gás de nebulização era de 241,3 KpA manométrico (35 psig), a vazão do gás secante de $12.0 \text{ l} \cdot \text{min}^{-1}$, a temperatura do gás secante de 350°C , a tensão do capilar era levada a $\pm 3000 \text{ V}$. Essas experiências permitiram determinar a massa dos dois produtos separados. Essas massas são respectivamente de 370 g/mol para o produto purificado primeiro ($[\text{M}-\text{H}_2\text{O}]^+ = 353$ produto majoritário) e 368 g/mol para o segundo produto ($[\text{M}-\text{H}_2\text{O}]^+ = 351$ produto majoritário).

A fim de se obter a estrutura, os produtos mencionados acima foram isolados e purificados nas seguintes condições: a fase móvel é uma mistura de uma solução aquosa de ácido trifluoroacético a 0,05% e de acetonitrila 70/30 (v/v). A cromatografia é feita em regime isocrático com uma vazão fixada em 1 ml/min . Nessas condições, os tempos de retenção dos 2 produtos são respectivamente de 8 e 13,3 minutos. Além disso, nesse caso, a amostra preparada (conforme parágrafo 20.1) não é diluída na água antes da injeção de $10 \mu\text{l}$.

Os 2 produtos são recuperados à saída da coluna cromatográfica e isolados nas seguintes condições: um cartucho Oasis HLB $1 \text{ cc } 30 \text{ mg}$ (Waters) é acondicionado seqüencialmente por 1 ml de acetonitrila, depois 1 ml de água/acetonitrila (20v/80v) e 1 ml de água/acetonitrila 80/20. A amostra é em seguida introduzida e o cartucho lavado sucessivamente por 1 ml de água/acetonitrila (95/5), 1 ml de água/acetonitrila deuterada (95/5), depois

purificada com 600 µl de água/acetonitrila deuterada 40/60. A solução recuperada é em seguida diretamente analisada em RMN.

Os espectros RMN obtidos para esses dois compostos são os seguintes:

5 - Primeiro produto purificado: platenolídeo A: (Espectro 9646V)

Espectro ¹H em CD₃CN (deslocamentos químicos em ppm):
 0,90 (3H, t, J=6Hz), 0,93 (3H, t, J=5Hz), 1,27 (3H, t, J=5Hz), entre 1,27 e
 1,40 (3H, m), 1,51 (1H, m), 1,95 (1H, m), 2,12 (1H, m), 2,30 (1H, d, J=12Hz),
 2,50 (1H, d, J=11Hz), 2,58 (1H, dd, J=9 e 12 Hz), 2,96 (1H, d, J=7Hz), 3,43
 10 (3H, s), 3,70 (1H, d, J=9Hz), 3,86 (1H, d, J=7Hz), 4,10 (1H, m), 5,08 (1H, m),
 5,58 (1H, dt, J=3 e 12Hz), 5,70 (1H, dd, J=8 e 12Hz), 6,05 (1H, dd, J=9 e
 12Hz), 6,24 (1H, dd, J=9 e 12Hz).

- Segundo produto purificado: platenolídeo B: (espectro 9647V)

Espectro ¹H em CD₃CN (deslocamentos químicos em ppm):
 15 0,81 ((3H, t, J=6Hz), 0,89 (1H, m), 1,17 (3H, d, J=5Hz), 1,30 (4H, m), 1,47
 (2H, m), 1,61 (1H, t, J=10Hz), 2,20 (1H, m), 2,38 (1H, d, J=13Hz), 2,52 (1H,
 m), 2,58 (1H, m), 2,68 (1H, dd, J=8 e 13Hz), 3,10 (1H, d, J=7Hz), 3,50 (3H,
 s), 3,61 (1H, d, J=8Hz), 3,82 (1H, d, J=7Hz), 5,09 (1H, m), 6,20 (1H, m), 6,25
 (1H, dd, J=9 e 12Hz), 6,58 (1H, d, J=12Hz), 7,19 (1H, dd, J=9 e 12Hz).

20 Essas experiências permitiram assim determinar a estrutura
 desses dois compostos. O primeiro produto purificado é o platenolídeo A e o
 segundo é o platenolídeo B, a estrutura deduzida dessas suas moléculas é
 apresentada na Figura 36.

Pôde também ser determinado, graças à técnica de CL/SM as-
 25 sociada à RMN, em condições ligeiramente diferentes daquelas descritas
 anteriormente, mas cujo desenvolvimento é bem conhecido do técnico, que
 a cepa OS49.67 produz além do platenolídeo A e B um derivado desses dois
 compostos. Trata-se do platenolídeo A+ micarose e do platenolídeo B+ mi-
 carose (a estrutura desses dois compostos é apresentada na Figura 40).
 30 Os resultados da análise do mosto da cepa OS49.67 são apresentadas na
 tabela 42.

TABELA 42. Os resultados da análise CL/SM do mosto da cepa OS49.67

Identidade	[] (mg/L)	Massa dos íons	Max de absorção
Platenolídeo A Massa exata=370 Fórmula molecular= $C_{20}H_{34}O_6$	16,1	$[M + Na]^+ 393,0$ $[M + K]^+ 408,9$ $[M - H_2O + H]^+ 353,0$ $[2M + Na]^+ 763,2$	231 nm
Platenolídeo B Massa exata=368 Fórmula molecular= $C_{20}H_{32}O_6$	1,4	$[[M - H_2O + H]^+ 351,0$ $[M + Na]^+ 391,0$ $[M + K]^+ 406,9$ $[2M + Na]^+ 759,1$	283 nm
Platenolídeo A + 'mi-carose' Massa exata=514 Fórmula molecular= $C_{27}H_{46}O_9$	4,27	$[M + K]^+ 537,0$ $[M + K]^+ 553,0$	230 nm
Platenolídeo B + 'mi-carose' Massa exata=512 Fórmula molecular= $C_{27}H_{44}O_9$	ND	$[M + Na]^+ 535,0$ $[M + K]^+ 550,9$ $[PlatB - H_2O + H]^+ 350,9$	284 nm

20.4 Análise dos intermediários de biossíntese produzidos pela cepa SPM509

- A cepa SPM509 na qual o gene orf14 é inativado (orf14::att3Ωaac-) não produz espiramicinas (conforme exemplo 13, 15 e 16). Uma amostra foi preparada segundo o método descrito acima (conforme parágrafo 20.1) e foi analisada por CL/SM conforme descrito acima (conforme o parágrafo 20.2 e 20.3). A análise dos intermediários de biossíntese presentes no sobrenadante de cultura da cepa SPM509 cultivada em meio MP5 mostrou que essa cepa produzia apenas a forma B do platenolídeo ("platenolídeo B", conforme Figura 36), mas não a forma A ("platenolídeo A", conforme Figura 36).

EXEMPLO 21: Interrupção do gene orf14 em uma cepa interrompida no gene orf3 (OS49.67)

O produto do gene orf14 é indispensável à biossíntese das espi-
ramicinas (conforme exemplo 13, 15 e 16: a cepa SPM509 na qual esse ge-
ne é interrompido não produz mais espi-
ramicinas). A análise dos intermediá-
rios de biossíntese presentes no sobrenadante de cultura da cepa SPM509
cultivada em meio MP5 mostrou que essa cepa produzia apenas a forma B
do platenolídeo, mas não a forma A (conforme exemplo 20). Uma das hipó-
teses que podem explicar essa observação é que o produto de orf14 seja
implicado na conversão da forma B do platenolídeo em forma A por uma
etapa enzimática de redução. Para testar essa hipótese, o gene orf14 foi
inativado em um mutante que não produz mais espi-
ramicina, mas produzindo as formas A e B do platenolídeo. É o caso da cepa OS49.67 (conforme
exemplo 6 e 20) na qual o gene orf3 é inativado por uma deleção em fase
(Δ orf3). Para inativar o gene orf14 nessa cepa, o plasmídeo pSPM509 foi
introduzido por transformação de protoplastos da cepa OS49.67 (Kieser, T.
et al., 2000). A inativação do gene orf14 foi descrita no caso da cepa OSC2
(conforme exemplo 13) e foi procedida do mesmo modo para a inativação do
gene orf14 na cepa OS49.67. Após transformação pelo plasmídeo
pSPM509, os clones são selecionados por sua resistência à apramicina. Os
clones resistentes à apramicina são em seguida repicados respectivamente
sobre meio com apramicina e sobre meio com higromicina. Os clones resis-
tentes à apramicina (ApraR) e sensíveis à higromicina (HygS) são em princí-
pio aqueles nos quais um duplo acontecimento de *cruzamento* ocorreu e nos
quais o gene orf14 foi substituído por uma cópia de orf14 interrompida pelo
cassete att3 Ω aac-. Esses clones foram mais particularmente selecionados e
a substituição da cópia selvagem de orf14 pela cópia interrompida pelo cas-
sette foi verificada por hibridização. Assim, o DNA total dos clones obtidos foi
digerido por várias enzimas, separado sobre gel de agarose, transferido sobre
membrana e hibridizado com uma sonda correspondente à cassete
aat3 Ω aac- para verificar que a substituição do gene ocorrera. A verificação
do genótipo pode também ser realizada por qualquer método conhecido do

técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto PCR.

Um clone que apresenta características esperadas (Δ orf3, orf14::att3 Ω aac-) foi mais particularmente selecionado e denominado SPM510. Pôde com efeito ser verificado graças às duas hibridizações que o cassette att3 Ω aac- estava bem presente no genoma desse clone e que se obtém o perfil de digestão esperado no caso de uma substituição, em consequência de um duplo acontecimento de recombinação, do gene selvagem orf14 pela cópia interrompida pelo cassette att3 Ω aac- no genoma desse clone.

EXEMPLO 22: Complementação funcional da interrupção do gene orf14

22.1 Construção do plasmídeo pSPM519

O gene orf14 foi amplificado por PCR, utilizando o seguinte par de oligonucleotídeos:

EDR31 : 5' CCCAAGCTTCTGCGCCCGCGGGCGTGAA 3' (SEQ ID N° 136) e EDR37 : 5' GCTCTAGAACCGTGTAGCCGCGCCCCGG 3' (SEQ ID N°

137) como matriz o DNA cromossômico da cepa OSC2. Os oligonucleotídeos EDR31 e EDR37 portam respectivamente o sítio de restrição HindIII e XbaI (seqüência em traços escuros). O fragmento de 1,2 kb assim obtido foi clonado no vetor pGEM-T easy (comercializado pela sociedade Promega (Madison, Wisconsin, USA)) para dar origem ao plasmídeo pSPM515. Esse plasmídeo foi em seguida digerido pelas enzimas de restrição HindIII e XbaI. O enxerto HindIII/XbaI de 1,2 kb obtido foi clonado no vetor pUWL201 (conforme exemplo 17.1) previamente digerido pelas mesmas enzimas. O plasmídeo assim obtido foi denominado pSPM519.

22.1 Transformação das cepas SPM509 e SPM510 pelo plasmídeo pSPM519.

O plasmídeo pSPM519 foi introduzido nas cepas SPM509 (conforme exemplo 13) e SPM510 (conforme exemplo 17) por transformação de protoplastos (Kieser, T. et al., 2000). Após transformação, os clones são selecionados por sua resistência ao tiostrepton. Os clones são em seguida repicados sobre um meio contendo tiostrepton.

A cepa SPM509 é uma cepa não-produtora de espiramicinas (conforme exemplo 13, 15 e 16 e Figura 24). A produção em espiramicinas da cepa SPM509 transformada pelo vetor pSPM519 (cepa denominada SPM509 pSPM519) foi analisada, cultivando-se essa cepa em meio MP5 em presença de tiostrepton. O sobrenadante de cultura foi em seguida analisado por CLHP (conforme exemplo 16 e 17). Os resultados dessa análise são apresentados na tabela 43, os dados são expressos em mg por litro do sobrenadante. Os resultados correspondem à produção total em espiramicinas (obtida adicionando-se a produção em espiramicina I, II e III). Foi observado que a presença do vetor pSPM519 na cepa SPM509 restaura a produção de espiramicinas (conforme tabela 43).

TABELA 43. Produção em espiramicinas da cepa SPM509 transformada pelo vetor pSPM519 (resultados expressos em mg/l do sobrenadante).

Cepa	Produção em espiramicinas
SPM509 pSPM519	58

A cepa SPM509 transformada pelo plasmídeo pSPM519 foi denominada SPM510 pSPM509.

EXEMPLO 23: complementação funcional da interrupção do gene orf3 pelo gene *tylB* de *S. fradiae*

23.1 Construção do plasmídeo pOSP49.52

O plasmídeo pOS49.52 corresponde a um plasmídeo, permitindo a expressão da proteína TylB em *S. ambofaciens*. Para construí-lo, a sequência codificante do gene TylB de *S. fradiae* (Merson-Davies & Cundliffe, 1994, número de acesso GenBank: U08223 (sequência da região), SFU08223 (sequência DNA) e AAA21342 (sequência protéica) foi introduzida no plasmídeo pKC1218 (Bierman et al., 1992, Kieser et al., 2000, uma cepa de *E. coli* contendo esse plasmídeo é acessível notadamente junto à ARS (NRRL) Agricultural Research Service Culture Collection) (Peoria, Illinois, USA), sob o número B-14790). Além disso, essa sequência codificante foi colocada sob controle do promotor *ermE** (conforme acima, notadamente exemplo 17.1, Bibb et al., 1985, Bibb et al., 1994).

23.1 Transformação da cepa OS49.67 pelo plasmídeo pOS49.52

A cepa OS49.67 na qual o gene orf3 é inativada por uma deleção em fase não produz espiramicinas (conforme exemplo 6 e 15). O plasmídeo pOS49.52 foi introduzido na cepa OS49.67 por transformação de protoplastos (Kieser, T et al., 2000). Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones são em seguida repicados sobre um meio contendo a apramicina. Um clone foi mais particularmente selecionado e foi denominado OS49.67 e pOS49.52.

Conforme foi demonstrado acima, a cepa OS49.67 não produz espiramicinas (conforme exemplo 6 e 15). A produção em espiramicinas da cepa OS49.67 transformada pelo vetor pOS49.52 foi analisada pela técnica descrita no exemplo 15. Pôde assim ser demonstrado que essa cepa possui um fenótipo produtor de espiramicinas. Assim, a proteína TylB permite a complementação funcional da interrupção do gene orf3.

EXEMPLO 24: Melhoria da produção em espiramicinas por superexpressão do gene orf28c

24.1 Construção do plasmídeo pSPM75

O gene orf28c foi amplificado por PCR, utilizando um par de oligonucleotídeos, comportando um sítio de restrição HindIII ou um sítio de restrição BamHI. Essas atrações têm a seguinte sentença:

KF30 : 5' AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138)

com um sítio de restrição HindIII (que Figura em traços cheios)

KF31 : 5' GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139)

com um sítio de restrição HindIII (que Figura em traços cheios)

As atrações KF30 e KF31 portam respectivamente o sítio de restrição HindIII e BamHI (seqüência em traços cheios). O par de atrações KF30 e KF31 permite amplificar um fragmento de DNA de um tamanho de aproximadamente 1,5 kb contendo o gene orf28c, utilizando como matriz o cosmídeo pSPM36 (conforme acima). O fragmento de 1,5 kb assim obtido foi clonado no vetor pGEM-T easy (comercializado pela sociedade Promega (Madison, Wisconsin, USA)) para dar origem ao plasmídeo pSPM74. O plasmídeo pSPM74 foi em seguida digerido pelas enzimas de restrição HindIII e BamHI e o enxerto HindIII/BamHI de aproximadamente 1,5 kb obti-

do foi subclonado no vetor pUWL201 (conforme exemplo 17.1) previamente digerido pelas mesmas enzimas. O plasmídeo assim obtido foi denominado pSPM75, ele contém o conjunto da sequência codificante de orf28c colocada sob o controle do promotor ermE*.

5 24.2 Transformação da cepa OSC2 pelo plasmídeo pSPM75

O plasmídeo pSPM75 foi introduzido nas cepas OSC2 por transformação de protoplastos (Kieser, T. et al., 2000). Após transformação dos protoplastos, os clones são selecionados por sua resistência ao tios-trepton. Os clones são em seguida repicados sobre um meio contendo o ti-
10 ostrepton e a transformação pelo plasmídeo é verificada por extração de plasmídeos. Dois clones foram mais particularmente selecionados e deno-minados OSC2/pSPM75(1) e OSC2/pSPM75(2).

Uma amostra da cepa OSC2/pSPM75(2) foi depositada junto à Collection Nationale de Cultures et Microorganismes (CNCM) Institut Pas-
15 teur, 25, rue du Docteur Roux 75724 Paris Cedex 15, France, 6 de outubro de 2003 sob o número de registro I-3101.

A fim de testar o efeito da superexpressão do gene orf28c sobre a produção de espiramicinas, a produção em espiramicinas dos clones OSC2/pSPM75(1) e OSC2/pSPM75(2) foi testada pela técnica descrita no
20 exemplo 15. A análise da produção em espiramicinas da cepa OSC2 foi também efetuada em paralelo para comparação. Foi possível assim ser ob-servada que a presença do plasmídeo pSPM75 aumenta significativamente a produção em espiramicinas da cepa OSC2. Isto demonstra que a superex-pressão de orf28c leva a um aumento da produção em espiramicinas e con-
25 firma seu papel como regulador.

A produção em espiramicinas dos clones OSC2/pSPM75(1) e OSC2/pSPM75(2) foi também analisada por CLHP (da mesma maneira que no exemplo 17.2). A análise da produção em espiramicinas da cepa OSC2 foi também efetuada em paralelo para comparação. Os resultados dessa
30 análise são apresentados na tabela 44. Os dados são expressos em mg por litro do sobrenadante. Os resultados correspondem à produção total em es-piramicinas (obtida adicionando-se a produção em espiramicina I, II e III).

TABELA 44. Produção em espiramicinas das cepas derivadas de OSC2 transformado por pSPM75, (resultados expressos em mg/l).

Cepa	Espiramicinas
OSC2	50
OSC2/pSPM75(1)	120
OSC2/pSPM75(2)	155

Assim, foi observado que a presença do plasmídeo pSPM75 aumenta significativamente a produção em espiramicinas da cepa OSC2.

- 5 Isto demonstra bem que a superexpressão de orf28c tem um efeito positivo em espiramicinas.

EXEMPLO 25: Análise da produção de intermediários de biossíntese de espiramicinas de uma cepa inativada no gene orf8

- A cepa OS49.107, na qual o gene orf8 é inativado por inserção
 10 do cassette Ω hyg não produz espiramicinas (conforme exemplo 7 e 15). O gene orf8 codifica uma proteína que apresenta uma relativa forte similitude com várias aminotransferases e sugere fortemente que o gene orf8 codifica uma 4 aminotransferase responsável pela reação de transaminação necessária à biossíntese da forosamina (conforme Figura 6). Espera-se, portanto,
 15 que a biossíntese das espiramicinas seja bloqueada no estágio da forocidina (conforme Figura 7). A cepa OS49.107 que é não-produtora de espiramicina deveria portanto produzir a forocidina.

- Uma amostra do sobrenadante da cepa OS49.107 foi preparada segundo o método descrito acima (conforme exemplo 16, sem extração do
 20 MIBK) e foi analisada por CL/SM conforme descrito acima (conforme o parágrafo 20.2 e 20.3). Em modo SIM, a massa 558 relativa ao íon molecular da forocidina foi selecionada e vários picos foram detectados. A presença de compostos de massa 588 é compatível com a hipótese do papel de orf8 em síntese de forosamina.

- 25 **EXEMPLO 26:** Análise da produção de intermediários de biossíntese de espiramicinas de uma cepa inativada no gene orf12

A cepa SPM507, na qual o gene orf12 é inativado, não produz espiramicinas (conforme exemplo 11 e 15). O gene orf12 codificaria uma 3,4

desidratase responsável pela reação de desidratação necessária à biossíntese da forosamina (conforme Figura 6). Espera-se, portanto que a biossíntese das espiramicinas seja bloqueada no estágio da forocidina (conforme Figura 7). A cepa SPM507 que é não-produtora de espiramicina deveria, portanto, produzir a forocidina.

Uma amostra do sobrenadante da cepa SPM507 foi preparada segundo método descrito acima (conforme exemplo 16, sem extração ao MIBK) e foi analisado por CL/SM conforme descrito acima (conforme o parágrafo 20.2 e 20.3). Nessas condições, o tempo de retenção da forocidina é de aproximadamente 12,9 minutos. Em modo SIM, a massa 558 relativa ao íon molecular $[M+H]^+$ da forocidina foi selecionado e um pico foi detectado. A presença de um composto com 558 permite validar a hipótese do papel do produto de orf12 na biossíntese das espiramicinas.

Todavia, a forocidina está presente em quantidade relativamente fraca e nessas condições, um produto absorvente com 238 nm foi mais particularmente observado (tempo de retenção de 17,1 min). A análise em LC/SM permitiu determinar a massa desse composto que é de 744,3 g/mol ($[M+H]^+ = 744,3$ produto majoritário).

A fim de se obter a estrutura, o produto mencionado acima foi isolado e purificado nas condições descritas anteriormente (conforme parágrafo 20.1)). A fase orgânica (MIBK) é então recuperada e evaporada. O extrato seco é retomado com água e extraído com heptano. A solução aquosa é em seguida extraída por fixação sobre cartucho Oasis HLB 1 g (Waters ASA, St-Quentin en-Yvelines, França). O composto é recuperado por purificação com uma mistura água/acetonitrila 30/70. Essa solução é em seguida injetada (100 μ L) sobre a coluna analítica e as frações recuperadas sobre cartucho Oasis HLB 1 cc 30 mg (Waters). Antes da utilização, os cartuchos Oasis HLB 1cc 30 mg (Waters) são acondicionados sequencialmente pela acetonitrila, depois uma mistura água/acetonitrila (20v/80v) e uma mistura de água/acetonitrila 80/20.

O cartucho Oasis HLB 1 cc 30 mg (Waters) é em seguida lavado sucessivamente por 1 ml de água/acetonitrila (95/5) , 1 ml de

água/acetonitrila deuterada (95/5), depois purificado com 600 µl de água/acetonitrila deuterada 40/60. A solução recuperada é em seguida diretamente analisada em RMN.

O espectro RMN obtido para esse composto é o seguinte
5 (Espectro RMN 19312V):

Espectro ¹H em CD₃CN/D₂O (deslocamentos químicos em ppm): 0,92 (3H, d, J=6Hz), 1,10 (1H, m), 1,14 (3H, s), 1,17 (3H, d, J=6Hz), 1,22 (3H, d, J=6Hz), 1,25 (3H, d, J=6Hz), 1,40 (1H, m), 1,75 (1H, dd, J=12 e 2Hz), 1,81 (1H, m), 1,90 (1H, d, J=12Hz), 2,05 (1H, m), 2,12 (3H, s), 2,15
10 (1H, m), 2,35 (2H, m), 2,45 (6H, s amplo), 2,53 (1H, m), 2,64 (1H, dd, J=12 e 9Hz), 2,80 (1H, dd, J=9 e 16Hz), 2,95 (1H, d, J=8Hz), 3,23 (2H, m), 3,34 (1H, d, J=7Hz), 3,45 (3H, s), 3,49 (1H, m), 3,93 (1H, dd, J=7 e 3Hz), 4,08 (1H, m), 4,37 (1H, d, J=6Hz), 4,88 (1H, m), 5,05 (2H, m), 5,65 (2H, m), 6,08 (1H, dd, J=8 e 12Hz), 6,40 (1H, dd, J=12 e 9Hz), 9,60 (1H, s).

15 Essas experiências permitiram assim determinar a estrutura desse composto, esta sendo apresentada na Figura 38.

EXEMPLO 28: Análise da produção de intermediários de biossíntese de espiramicinas de uma cepa inativada no gene *orf5**

A cepa SPM501 possui o genótipo *orf6*::att1Ωhyg+*. Graças ao
20 efeito polar da inserção do cassette *att1Ωhyg+* no gene *orf6**, pôde ser determinado que o gene *orf5** é indispensável à via de biossíntese das espiramicinas. Com efeito, a inserção do cassette excisável na parte codificante do gene *orf6** acarreta uma parada total da produção em espiramicinas por efeito polar sobre a expressão do gene *orf5**. Todavia, uma vez que o cas-
25 sette inserida foi excisada (e portanto quando só o gene *orf6** é inativado, cf Exemplos 14 e 15), restaura-se uma produção de espiramicina I. Isto demonstra que o gene *orf5** é indispensável à biossíntese das espiramicinas, já que sua inativação acarreta um parada total da produção em espiramicinas.

O gene *orf5** codifica uma proteína que apresenta uma relativa
30 forte similitude com várias O-metiltransferases. O gene *orf5** seria uma O-metil transferase implicada na biossíntese do platenolídeo. Para verificar essa hipótese, experiências de análise CL/SM e de RMN foram feitas sobre

uma cepa *S.ambofaciens* de genótipo *orf6*::att1Ωhyg+* obtida a partir de uma cepa que superproduz as espiramicinas.

Uma amostra do sobrenadante dessa cepa foi preparada, segundo o método descrito acima (conforme exemplo 16, sem extração ao MIBK) e foi analisada por CL/SM conforme descrito acima (conforme o parágrafo 20.2 e 20.3). Todavia, a coluna utilizada é uma coluna X-Terra (Waters AS, St-Quentin en Yvelines, França) e a tensão de cone do espectrômetro é regulada em 380 V para se obter a fragmentação do composto analisado. Nessas condições, um produto cujo tempo de retenção é de aproximadamente 13,1 minutos é observado. O espectro de massa desse composto tem um andamento similar àquele da espiramicina I, mas o íon molecular é de 829. A diferença de massa de 14 em relação à massa da espiramicina pode ser explicada pela ausência de grupo metila sobre oxigênio portada pelo carbono N° 4 do ciclolactona (a estrutura desse composto é apresentada na Figura 39). A presença de um composto em 829 permite validar a hipótese do papel de *orf5** na biossíntese das espiramicinas.

Por um teste microbiológico feito sobre uma cepa sensível de *M. luteus* (conforme exemplo 15 e Figura 18) foi demonstrado que a molécula intermediária (espiramicina menos grupo metila cuja estrutura é apresentada na Figura 39) produzida pela cepa *orf6*::Ωatthyg+* é muito menos ativa (de um fator 10) que a espiramicina de origem com o grupo metila em posição 4.

EXEMPLO 28: Construção de novas "cassettes excisáveis"

Novas cassettes excisáveis foram construídas. Essas cassettes são muito semelhantes às cassettes excisáveis descritas no exemplo 9. A diferença principal entre as antigas e as novas cassettes é a ausência nestas das seqüências correspondentes às extremidades da interposon Ω, seqüências que contêm um terminador de transcrição proveniente do fago T4.

Nas cassetes sem terminador, o gene que confere a resistência a um antibiótico é contornado pelas seqüências *attR* e *attL* que permitem a excisão. O gene de resistência é o gene *aac(3)IV* que codifica uma acetil-

transferase que confere a resistência à apramicina. Esse gene está presente no cassette Ω aac (número de acesso GenBank: X99313, Blondelet-Rouault. M.H. et al, 1997) e foi amplificado por pCR, utilizando como matriz o plasmídeo pOSK1102 (conforme acima) e como atrações os oligonucleotídeos

5 KF42 e KF43 contendo cada um o sítio de restrição HindIII (em traço cheios) (AAGCTT) em 5'.

KF42: 5'-AAGCTTGTACGGCCACAGAATGATGTCAC-3' (SEQ ID N° 153)

e

KF43: 5'-AAGCTTCGACTACCTTGGTGATCTCGCCTT-3' (SEQ ID N° 154).

O produto PCR obtido de aproximadamente 1kb foi clonado no vetor E.Coli pGEMT Easy dando origem ao plasmídeo pSPM83.

O vetor pSPM83 foi digerido pela enzima de restrição HindIII. O

10 fragmento HindIII-HindIII do enxerto foi isolado por purificação a partir de um gel de agarose 0,8%, depois clonado no sítio HindIII situado entre as seqüências attL e attR dos diferentes plasmídeos que portam as diferentes cassette excisáveis possíveis (conforme exemplo 9 e Figura 27) de forma a substituir o fragmento HindIII correspondente a Ω acc pelo fragmento HindIII

15 correspondente ao gene aac sozinho. Isto permitiu obter as cassettes att1aac, att2aac e att3aac (segundo a fase desejada, conforme exemplo 9). Segundo a orientação do gene aac em relação às seqüências attL e attR, distinguem-se att1aac+, att1aac-, att2aac+, att2aac-, att3aac+ e att3aac- (segundo as mesmas convenções que aquelas adotadas no exemplo 9).

20 EXEMPLO 29: Construção de uma cepa de *S.ambifaciens* interrompida no gene orf28c

A inativação do gene orf28c foi feita, graças à técnica das cassettes excisáveis. O cassette excisável att3aac+ (conforme exemplo 28) foi amplificada por PCR, utilizando como matriz o plasmídeo pSPM101 (O

25 plasmídeo pSPM101 é um plasmídeo derivado do vetor pGP704Not (Chaverroche et al., 2000) (Miller VL & Mekalanos JJ, 1988) no qual o cassette att3aac+ foi clonada como um fragmento EcoRV no sítio único EcoRV de pGP704Not) e com o auxílio das seguintes atrações:

KF32:

5' CAACCGCTTGAGCTGCTCCATCAACTGCTGGGCCGAGGTATCGCGCGCG
CTTCGTTCCGGGACGAA 3' (SEQ ID N° 155)

e KF33:

5' TGGGTCCCGCCGCGCGGCACGACTTCGACTCGCTCGTCTATCTGCCTCTT
CGTCCCGAAGCAACT 3' (SEQ ID N° 156)

Os 39 nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma sequência correspondente a uma sequência no gene orf28c e os 26 nucleotídeos situados mais em 3' (representado em **negrito e sublinhados acima**) correspondem à seringa de uma das extremidades do cassette excisável att3aac+.

O produto de PCR assim obtido foi utilizado para transformar a cepa E.coli hiper-recombinante DY330 (Yu et al., 2000) (essa cepa contém os genes exo, bet e gam do fago lambda integrados em seu cromossoma, esses genes são expressos em 42°C, ela foi utilizada no lugar da cepa E.coli KS272 (Chaverche et al., 2000) contendo o cosmídeo pSPM36. Assim, as bactérias foram transformadas por eletroporação por esse produto de PCR e os clones foram selecionados por seu reservatório à apramicina. Os cosmídeos dos clones obtidos foram extraídos e digeridos pela enzima de restrição BamHI, com a finalidade de verificar se o perfil de digestão obtido correspondia ao perfil esperado, caso houvesse inserção do cassette (att3aac+) no gene orf28c, isto é, se tivesse havido recombinação homóloga entre as extremidades do produto PCR e o gene alvo. A verificação da construção pode também, ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR. Um clone, cujo cosmídeo possui um perfil esperado foi selecionado e o cosmídeo correspondente foi denominado pSPM107. Esse cosmídeo é um derivado de pSPM36 no qual orf28c é interrompido pelo cassette att3aac-. A inserção do cassette é acompanhada de uma deleção no gene orf28c, a interrupção começa ao nível do 28º códon de orf28c. Restam após o cassette os 137 últimos códons de orf28c.

O cosmídeo pSPM107 foi em uma primeira etapa introduzido na

cepa *E.coli* DH5 α , depois na cepa *streptomyces ambofaciens* OSC2 por transformação de protoplastos. Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones resistentes à apramicina são em seguida repicados respectivamente sobre meio com apramicina (antibiótico B) e sobre meio com puromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (ApraR) e sensíveis à puromicina (PuroS) são em princípio aqueles nos quais um duplo acontecimento de cruzamento ocorreu e que possuem o gene *orf28c* interrompido pela cassette *att3aac-*. Esses clones foram mais particularmente selecionados e a substituição da cópia primitiva de *orf28c* pela cópia interrompida pelo cassette foi verificado por hibridização. Assim, o DNA total dos clones obtidos, foi digerido por várias enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette *att3aac+* para verificar a presença do cassette no local esperado no DNA genômico dos clones obtidos. A verificação do genótipo pode também ser realizada por qualquer método conhecido e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone que apresenta as características esperadas (*orf28c::att3aac+*) foi mais particularmente selecionado e denominado SPM107. Esse clone possui, portanto, o genótipo: *orf28c::att3aac+* e foi denominado SPM107. É inútil proceder à excisão do cassette para o estudo do efeito da inativação de *orf28c*, considerando-se a orientação dos genes (conforme Figura 3). Com efeito, o fato de *orf29* ser orientado em sentido oposto ao *orf28c* mostra que esses genes não são co-transcritos. A utilização de uma cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer momento, notadamente por transformação pelo plasmídeo pOSV508.

A fim de testar o efeito da extinção do gene *orf28c* sobre a produção de espiramicinas da cepa SPM107 foi testado pela técnica descrita no exemplo 15. Pôde assim ser demonstrado que essa cepa possui um fenótipo não-produtor de espiramicinas. Isto demonstra que o gene *orf28c* é um gene essencial à biossíntese da espiramicina em *S. ambofaciens*.

EXEMPLO 30: Construção de uma cepa de *S.ambofaciens* interrompida no gene orf31

A inativação do gene orf31 foi feita graças à técnica das cassetes excisáveis. O cassette excisável att3aac+ foi amplificada por PCR, utilizando como matriz o plasmídeo pSPM101 e os oligonucleotídeos EDR71 e EDR72.

EDR71:

**5' CGTCATCGACGTGCGGGGAAGACAGAGGTGATACCGATGATCGCGCGC
GCTTCGTTTCGGGACGAA 3' (SEQ ID Nº 157)**

EDR72:

**5' GCCAGCACCTCGTCCAGCTGCTCGACGGAACTCACCCCCATCTGCCTCT
TCGTCCCGAAGCAACT 3' (SEQ ID Nº 158)**

Os 39 nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma sequência correspondente a uma sequência no gene orf31 e os 26 nucleotídeos situados mais em 3' (figurados em negrito e sublinhados acima) correspondem à sequência de uma das extremidades do cassette excisável att3aac+.

O produto de PCR assim obtido foi utilizado para transformar a cepa *E.coli* KS272 contendo o plasmídeo pKOBEG e o cosmídeo pSPM36, conforme descrito por Chaverroche et al. (Chaverroche et al., 2000) (conforme Figura 12 para o princípio, o plasmídeo pOS49.99 deve ser substituído pelo cosmídeo pSPM36 e o plasmídeo obtido não é mais pSPM17, mas pSPM543). Assim, as bactérias foram transformadas por eletroporação por esse produto de PCR e os clones foram selecionados por sua resistência à apramicina. Os cosmídeos dos clones obtidos foram extraídos e digeridos por várias enzimas de restrição, com a finalidade de verificar se o perfil de digestão obtido corresponde ao perfil esperado se tiver tido inserção do cassette (att3aac+) no gene orf31, isto é, se tiver tido recombinação homóloga entre as extremidades do produto PCR e o gene alvo. A verificação da construção pode também ser feita por qualquer método conhecido do técnico e notadamente por pCR, utilizando os oligonucleotídeos apropriados e

seqüenciação do produto de PCR. Um clone cujo cosmídeo possui o perfil esperado foi selecionado e o cosmídeo correspondente foi denominado pSPM543. Esse cosmídeo é um derivado de pSPM36 no qual orf31 é interrompido pelo cassette att3aac⁺ (conforme Figura 12). A inserção do cassette
5 é acompanhada de uma deleção no gene orf31, a interrupção começa ao nível do trigésimo sexto códon de orf31. Restam após o cassette os 33 últimos códons de orf31.

O cosmídeo pSPM543 foi introduzido na cepa *streptomyces ambofaciens* OSC2 (conforme acima) por transformação de protoplastos (Kieser, T. al., 2000). Após transformação, os clones são selecionados por sua
10 resistência à apramicina. Os clones resistentes à apramicina são em seguida repicados respectivamente sobre meio com apramicina (antibiótico B) e sobre meio com puromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (ApraR) e sensíveis à puromicina (PuroS) são em
15 princípio aqueles nos quais um duplo acontecimento de cruzamento ocorreu e que possuem o gene orf31 interrompido pelo cassette att3aac⁺. Esses clones foram mais particularmente selecionados e a substituição da cópia primitiva de orf31 pela cópia interrompida pelo cassette foi verificada por hibridização. Assim, o DNA total dos clones obtidos foi digerido por várias en-
20 zimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette att3aac⁺ para verificar a presença do cassette no local esperado no DNA genômico dos clones obtidos. Uma segunda hibridização foi feita, utilizando como sonda um fragmento de DNA obtido por PCR e correspondente a uma parte muito larga da
25 seqüência codificante do gene orf31.

A verificação do genótipo pode também ser realizada por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone que apresenta as características esperadas
30 (orf31::att3aac⁺) foi mais particularmente selecionado e denominado SPM543. Pôde, com efeito, ser verificado, graças às duas hibridizações que o cassette att3aac⁺ estava bem presente no genoma desse clone que se

obtem bem o perfil de digestão esperado no caso de uma substituição, após um duplo acontecimento de recombinação, do gene selvagem pela cópia interrompida pelo cassette att3aac+ no genoma desse clone. Esse clone possui, portanto, o genótipo: orf31::att3aac+ e foi denominado SPM543. É

5 inútil proceder à excisão do cassette para o estudo do efeito da inativação de orf31, considerando-se a orientação dos genes (conforme Figura 3). Com efeito, o fato de orf32c ser orientado em sentido oposto ao orf31 mostra que esses genes não são co-transcritos. A utilização de uma cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a

10 qualquer momento, notadamente por transformação pelo plasmídeo pOSV508.

A fim de testar o efeito da extinção do gene orf31 sobre a produção de espiramicinas, a produção em espiramicinas da cepa SPM543 foi testada pela técnica descrita no exemplo 15. Pôde assim ser demonstrado

15 que essa cepa possui um fenótipo não-produtor de espiramicinas. Isto demonstra que o gene orf31 é um gene essencial à biossíntese da espiramicina em *S. ambofaciens*.

EXEMPLO 31: Construção de uma cepa de *S. ambofaciens* interrompida no gene orf32c

20 A inativação do gene orf32c foi realizada graças à técnica das cassettes excisáveis. O cassette excisável att3aac+ foi amplificada por PCR utilizando-se como matriz o plasmídeo pSPM101 e com o auxílio das seguintes atrações:

KF52:

5' GATCCGCCAGCCTCACGTCACGCCGCGCCGCCTCCCTGAC**ATCGCGCGC**
GCTTCGTTCGGGACGAA 3' (SEQ ID N° 159).

e KF53:

5' GAGGCGGACGTCGGTACGCGGTGGGAGCCGGAGTTCGACA**ATCTGCCTC**
TTCGTCCCGAAGCAACT 3' (SEQ ID N° 160).

Os 40 nucleotídeos situados na extremidade 5' desses oligonucleotídeos comportam uma sequência que corresponde a uma sequência no

25 gene orf32c e os 26 nucleotídeos situados mais em 3' (figurado em negrito e

sublinhado acima) correspondem à seqüência de uma das extremidades do cassette excisável att3aac+.

O produto de PCR assim obtido foi utilizado para transformar a cepa *E. coli* hiper-recombinante DY330 (Yu et al., 2000) contendo o cosmídeo pSPM36. Assim, as bactérias foram transformadas por eletroporação por esse produto de PCR e os clones foram selecionados por sua resistência à apramicina. Os cosmídeos dos clones obtidos foram extraídos e digeridos pelas enzimas de restrição BamHI, com o objetivo de verificar se o perfil de digestão obtido corresponde ao perfil esperado, se houve inserção do cassette (att3aac+) no gene orf32c, isto é, se houve recombinação homóloga entre as extremidades do produto PCR e o gene alvo. A verificação da construção pode também ser realizada por qualquer método conhecido do técnico, e notadamente por PCR, utilizando-se os oligonucleotídeos apropriados e seqüenciação do produto de PCR. Um clone, cujo cosmídeo possui o perfil esperado, foi selecionado e o cosmídeo correspondente foi denominado pSPM106. Esse cosmídeo é um derivado de pSPM36 no qual orf32c é interrompido pelo cassette att3aac+. A inserção do cassette é acompanhada de uma deleção no gene orf32c, a interrupção começa ao nível do 112º códon de orf32c. Permanecem, após a cassette, os 91 últimos códons de orf32c.

O cosmídeo pSPM106 foi, num primeiro tempo, introduzido na cepa *E. coli* DH5α depois, na cepa *streptomyces ambofaciens* OSC2 por transformação. Após transformação, os clones são selecionados por sua resistência à apramicina. Os clones resistentes à apramicina são, em seguida, repicados respectivamente sobre meio com apramicina (antibiótico B) e sobre meio com puromicina (antibiótico A) (conforme Figura 9). Os clones resistentes à apramicina (ApraR) e sensíveis à puromicina (PuroS) são em princípio aqueles nos quais um duplo acontecimento de cruzamento se produziu e que possuem o gene orf32c interrompido pela cassette att3aac+ . Esses clones foram mais particularmente selecionados e a substituição da cópia primitiva de orf32c pela cópia interrompida pelo cassette foi verificado por hibridização. Assim, o DNA total dos clones obtidos foi digerido por vári-

as enzimas, separado sobre gel de agarose, transferido sobre membrana e hibridizado com uma sonda correspondente à cassette att3aac+ para verificar a presença do cassette no DNA genômico dos clones obtidos. A verificação do genótipo pode também ser realizado por qualquer método conhecido do técnico e notadamente por PCR, utilizando os oligonucleotídeos apropriados e seqüenciação do produto de PCR.

Um clone que apresenta as características esperadas (orf32c::att3aac+) foi mais particularmente selecionado. Esse clone possui, portanto, o genótipo: orf32c::att3aac+ e foi denominado SPM106. É inútil proceder à excisão do cassette para o estudo do efeito da inativação de orf32c, considerando-se a orientação dos genes (conforme Figura 3). Com efeito, o fato de orf33 ser orientado em sentido oposto a orf32c mostra que esses genes não são co-transcritos. A utilização de uma cassette excisável permite, ao contrário, ter a possibilidade de se livrar do marcador de seleção a qualquer momento, notadamente por transformação pelo plasmídeo pOSV508.

A fim de testar o efeito da extinção do gene orf32c sobre a produção de espiramicinas, a produção em espiramicinas da cepa SPM106 foi testada pela técnica descrita no exemplo 15. Pôde assim ser demonstrado que essa cepa possui um fenótipo produtor de espiramicinas. Isto demonstra que o gene orf32c não é um gene essencial à biossíntese da espiramicina em *S. ambofaciens*.

Lista das construções descritas no presente pedido

Lista das abreviações: Am: ampicilina; Hyg: higromicina; Sp: espiramicina; Ts: tiostrepton; Cm: cloranfenicol. Kn: canamicina, Apra: apramicina.

Nome da construção	Marcador de seleção	Principais características	Referência
pWE15	Am		(Wahl, et., 1987)
pWED1	Am	pWE15 no qual um fragmento <i>HpaI-HpaI</i> de 4,1 kb foi deletado	(Gourmelen et al., 1998)
pOJ260	Apra	Conjugativo, não replicativo em <i>Streptomyces</i>	(Bierman et al., 1992)
pHP45 Ω hyg	Hyg	Cassette Ω hyg em pHP45.	(Blondelet-Rouault et al., 1997)
pKC505	Apra	Cosmídeo	(Richardson MA et al., 1987)
pIJ486	Ts	Plasmídeo replicativo multicópias <i>Streptomyces</i>	(Ward et al., 1986)
pOSint3	Am	<i>P_{trc}-xis-int</i> em pTrc99A	(Raynal et al., 1998)
pWHM3	Am, Ts	Vetor <i>E. coli</i> / <i>Streptomyces</i>	(Vara et al., 1989)
pKOBEG	Cm		(Chaveroche et al., 2000)
pGP704Not	Am		(Chaveroche et al., 2000)
pMBL18	Am		(Nakano et al., 1995)
pGEM-T Easy	Am	Vetor <i>E. coli</i> para clonagem de produtos PCR	Mezei et al., 1994
pOS49,1	Am	pWED1 com enxerto ao nível do sítio <i>BamHI</i>	Exemplo 2
pOS49,11	Am	Fragmento <i>SacI</i> de pOS49,1 em pUC19.	Exemplo 2
pOS49,12	Ch	Fragmento <i>XhoI</i> de pOS49,11 em pBC SK+	Exemplo 2

Nome da construção	Marcador de seleção	Principais características	Referência
pOS49,14	Cm, Hyg	pOS49,12 com o gene <i>orf3</i> interrompido pela cassette Ω_{hyg}	Exemplo 2
pOS49,16	Apra, Hyg	Enxerto de pOS49,14 em pOJ260	Exemplo 2
pOS49,28	Cm	Fragmento <i>Bam</i> HI- <i>Pst</i> II de 3,7 kb de pOS49,1 em pBC SK+	Exemplo 3
pOS44,1	Apra, Sp	PKC505 contendo um enxerto que confere a resistência à espiramicina em <i>S. griseofuscus</i>	(Pernodet et al., 1999)
pOS44,2	Ts, Sp	Fragmento <i>Sau</i> 3AI de 1,8 kg pOS44,1 em pIJ486	Exemplo 3
pOS44,4	Am	Enxerto de pOS44,2 em pUC19	Exemplo 3
pSPM5	Am	pWED1 enxerto com d'DNA de <i>S. ambofaciens</i> no nível do sítio <i>BAM</i> HI	Exemplo 3
pSPM7	Am	pWED1 com enxerto com d'DNA de <i>S. ambofaciens</i> no nível do sítio <i>BAM</i> HI	Exemplo 3
pOSK1205	Hyg	pBK-CMB em que <i>hyg</i> replace <i>neo</i>	Exemplo 5
pOS49,67	Apra	Fragmento <i>Eco</i> RI- <i>Sac</i> I de pOS49,1 comportando uma deleção interna de 504 nucleotídeos, em pOJ260	Exemplo 6
pOS49,88	Am	Fragmento de 3,7 kb <i>Pst</i> II- <i>Eco</i> RI de pOS49,1 em	Exemplo 7

Nome da construção	Marcador de seleção	Principais características	Referência
		pUC19	
pOS49,106	Am	pOS49,88 com <i>hyg</i> em <i>orf8</i> (<i>hyg</i> e <i>orf8</i> na mesma orientação)	Exemplo 7
pOS49,120	Am	pOS49,88 com <i>hyg</i> em <i>orf8</i> (<i>hyg</i> e <i>orf8</i> na mesma orientação)	Exemplo 7
pOS49,107	Apra, Hyg	Enxerto de pOS49,106 em pOJ260	Exemplo 7
pOS49,32	Am, Kn	Fragmento de 1,5 kb interno <i>orf10</i> em pCR2,1-TOPO	Exemplo 8
pOS49,43	Am, Kn	pOS49,32 com <i>hyg</i> em <i>orf10</i> (<i>hyg</i> e <i>orf10</i> na mesma orientação)	Exemplo 8
pOS49,44	Am, Kn	pOS49,32 com <i>hyg</i> em <i>orf10</i> (<i>hyg</i> e <i>orf10</i> na orientação oposta)	Exemplo 8
pOS49,50	Apra, Hyg	Enxerto de pOS49,43 em pOJ260	Exemplo 8
pWHM3Hyg	Am, Hyg	PWHM3 em que <i>tsr</i> é substituído por <i>hyg</i>	Exemplo 10

pOSV508	Am, Ts	<i>P_{trc-xis-int}</i> no qual pWHM3	Exemplo 9
patt1 Ω hyg+	Cm, Hyg	Cassette <i>patt1Ωhyg+</i> no qual pBC SK+ no sítio <i>HindIII</i> foi suprimido	Exemplo 9
patt3 Ω aac-	Cm, Gn	Cassette <i>att3Ωaac-</i> no qual pBC SK+ no sítio <i>HindIII</i> foi suprimido	Exemplo 9
pOSV510	Am, Hyg	<i>pra-Amh</i> em pWHM3Hyg	Exemplo 10
pOS49,99	Am	Fragmento <i>EcoRI-</i> <i>BamHI</i> de 4,5 kb de pSPM5 no qual pUC19	Exemplo 10
pOSK1102	Am, Apra	pGP704Not contendo uma cassette <i>att3Ωaac-</i>	Exemplo 10
pSPM17	Am, Apra	pOS49,99 no qual <i>orf2</i> é interrompido pela cas- sette <i>att3Ωaac-</i>	Exemplo 10
pSPM21	Hyg, Apra	OSK1205 contendo o enxerto <i>EcoRI-XbaI</i> de pSPM17 (no qual <i>orf2</i> é interrompido pela cas- sette <i>att3Ωaac-</i>)	Exemplo 10
pSPM502	Am	Fragmento <i>BglIII</i> de 15,1 kb de pSPM7 em pMBL18	Exemplo 11
pSPM504	Hyg	Enxerto de pSPM502 em pOSK1205	Exemplo 11
pSPM507	Hyg, Para	pSPM504 no qual <i>orf12</i> é interrompido pela cas- sette <i>att3Ωaac</i>	Exemplo 11

pSPM508	Hyg, Para	pSPM504 no qual <i>orf13</i> é interrompido pela cas-sette <i>att3Ωaac</i>	Exemplo 12
pSPM509	Hyg, Para	pSPM504 no qual <i>orf14</i> é interrompido pela cas-sette <i>att3Ωaac</i> -	Exemplo 13
pBXL1111	Am	Fragmento de 1,11 kb contendo <i>orf6</i> amplifica-do para PCR a partir de pSPM7, no vetor pGEM-T Easy	Exemplo 14
pBXL1112	Am, Hyg	pBXL1111 no qual a cassette <i>att1Ωhyg+</i> foi introduzido após dele-ção de 120 pb na se-quência codante do ge-ne <i>orf6</i> *	Exemplo 14
pBXL1113	Apra, Hyg	Enxerto PstI de 3,7 kb de pBXL1112 em pOJ260	Exemplo 14
pSPM520	Am	Fragmento de PCR amplifi-cado para oligonucleótidos EDR40-EDR42 em pGEM-T easy	Exemplo 17
pSPM521	Am	Fragmento de PCR amplifi-cado para oligonucleotideos EDR40- EDR42 em pGEM-T easy	Exemplo 17
pSPM522	Am	Fragmento de PCR am-plificado para oligonu-cleotídeos EDR41-EDR42 em pGEM-T Easy	Exemplo 17

pUWL201	Am, Ts		(Doumith et al, 2000)
pSPM523	Am, Ts	Fragmento HindIII-BamHI de enxerto do plasmídeo pSPM520 o vetor pUWL201	Exemplo 17
pSPM524	Am, Ts	Fragmento HindIII-BamHI de enxerto do plasmídeo pSPM521 o vetor pUWL201	Exemplo 17
pSPM525	Am, Ts	Fragmento HindIII-BamHI de enxerto do plasmídeo pSPM522 o vetor pUWL201	Exemplo 17
pSPM527	Am	pSPM521 com recolocação do quadro de leitura no nível do sítio XhoI	Exemplo 17
pSPM528	Am, Ts	Fragmento HindIII-BamHI de enxerto do plasmídeo pSPM522 o vetor pUWL201	Exemplo 17
pVF 10,4			(Vara et al., 1985; Lacalle et al., 1989)
pPM803	Ts		(Mazodier, P et al., 1989)
pGEM-T-pac-oriT	Am	Cassette pac-oriT (amplificado para PCR a partir de pVF 10,4 e pPM803) no pGEM-T Easy	Exemplo 18
pWED2	Am	Cassette pac-oriT obtido a partir de pGEM-T-pac-oriT inserido no pWED1	Exemplo 18
pSPM34	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM35	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19

pSPM36	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM37	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM38	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM39	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM40	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM41	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM42	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM43	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM44	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM45	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM47	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM48	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM50	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM51	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM52	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM53	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19

pSPM55	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM56	Am	pWED2 com enxerto no nível do sítio BamHI	Exemplo 19
pSPM58	Kn	Fragmento PstI-PstI de aproximadamente 6kb de enxerto de pSPM36 em pBK-CMV	Exemplo 19
pSPM72	Kn	Fragmento PstI-PstI de aproximadamente 10kb de enxerto de pSPM36 clonado em pBK-CMV	Exemplo 19
pSPM73	Cm	Fragmento EcoRI-HindIII de enxerto de pSPM72 em pBC-SK+	Exemplo 19
pSPM515	Am	Fragmento de PCR amplificado para EDR31-EDR37 em pGEM-T easy	Exemplo 22
pSPM519	Am, Ts	Enxerto HindIII/XbaI de pSPM515 em pUWL201	Exemplo 22
pOS49,52	Apra	Seqüência codificante de tyIB sob controle do promotor ermE* no plasmídeo pKC1218	Exemplo 23
pSPM74	Am	Fragmento de PCR amplificado para KF30-KF31 em pGEM-T easy	Exemplo 24
pSPM75	Am, Ts	Enxerto HindIII/BamHI de pSPM74 em pUWL201	Exemplo 24
pSPM79	Kn	Fragmento PstI-PstI de aproximadamente 2,5kb de enxerto de pSPM36 em pBK-CMV	Exemplo 19

pSPM83	Am	Fragmento de PCR amplificado por KF42-KF43 em pGEM-T easy	Exemplo 28
pSPM107	Am, Apra	pSPM36 no qual orf28c é interrompido pela cassette att3acc+	Exemplo 29
pSPM543	Am, Apra	pSPM36 no qual orf31 é interrompido pela cassette att3acc+	Exemplo 30
pSPM106	Am, Apra	pSPM36 no qual orf 32c é interrompido pela cassette att3aac+	Exemplo 31

Depósito de material biológico

Os seguintes organismos foram depositados em 10 de julho de 2002 junto à Collection Nationale de Cultures de Microorganismes (CNCM), 25 rue du Docteur Roux, 75724 Paris Cedex 15, France, segundo as disposições do Tratado de Budapeste.

- 5
- Cepa OSC2 sob o número de registro I-2908
 - Cepa SPM501 sob o número de registro I-2909
 - Cepa SPM502 sob o número de registro I-2910
 - Cepa SPM507 sob o número de registro I-2911
 - 10 - Cepa SPM508 sob o número de registro I-2912
 - Cepa SPM509 sob o número de registro I-2913
 - Cepa SPM21 sob o número de registro I-2914
 - Cepa SPM22 sob o número de registro I-2915
 - Cepa OS49.67 sob o número de registro I-2916
 - 15 - Cepa OS49107 sob o número de registro I-2917
 - Cepa *Escherichia coli* DH5 α contendo o plasmídeo pOS44.4 sob o número de registro I-2918

Os seguintes organismos foram depositados em 26 de fevereiro de 2003 junto à Collection Nationale de Cultures de Microorganismes

(CNCM), 25 rue du Docteur Roux, 75724 Paris Cedex 15, France, segundo as disposições do Tratado de Budapeste.

- Cepa SPM502 pSPM525 sob o número de registro I-2977.

Os seguintes organismos foram depositados junto à Collection

- 5 Nationale de Cultures de Microorganismes (CNCM) Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, França, segundo as disposições do Tratado de Budapeste.

- Cepa OSC2/pSPM75(2) sob o número de registro I-3101.

- 10 Todas as publicações e patentes citadas incorporadas ao presente pedido por referência.

BIBLIOGRAFIA

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, Basic local alignment search tool. *J mol Biol.* (1990). 215 (3): 403-410.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Milier W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* (1997). 25 (17):3389-402.
- Amann, E., Ocbs, B. and Abel, K. J. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Eseberichia coli* Gene (1988) 69 (2), 301-315.
- 10 - Arisawa, A., Kawamura,N., Tsunekawa, H., Okamura, K., Tone,H. and Okamoto,R. Cloning and nucleotide sequences of two genes involved in the 4"-O-acylation of macrolide antibiotics from *Streptomyces thermotolerans*. *Biosci. Biotechnol. Biochem.* (1993). 57 (12) 2020-2025.
- Arisawa A, Kawamura N, Takeda K, Tsunekawa H, Okamura K, Okamoto R. Cloning of the macrolide antibiotic biosynthesis gene *acyA*, which encodes 3-O-acyltransferase, from *Streptomyces thermotolerans* and its use for direct fermentative production of a hybrid macrolide antibiotic. *Appl Environ Microbiol.* (1994). 60 (7):2657-2660.
- 15 - August, P.R., Tang, L., Yoon, Y. J., Ning, Streptomyces, Mueller, R., Yu, T. W., Taylor, M., Hoffmann, D., Kim, C.G., Zhang, X., Hutchinson, C.R. et Floss, H.G. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* (1998) 5 (2), 69-79.
- Ausubel Fred M., Brent Roger, Kingston Robert E., Moore David D., Seidman J. G., Smith John A., Struhl Kevin (Editeurs). *Current Protocols in Molecular Biology*, publié par John Wiley & Sons Inc. Current Protocols Customer Service, 605 Third Avenue, 9th Floor New York, NY 10158 USA (édition mise à jour de Mars 2002).
- 25 - Baltz RH, McHenney MA, Cantwell CA, Queener SW, Solenberg PJ. Applications of transposition mutagenesis in antibiotic producing streptomycetes. *Antonie Van Leeuwenhoek.* (1997). 71 (1-2): 179-187.
- Bate N, Butler AR, Gandeche AR, Cundliffe E. Multiple regu-
- 30

latory genes in the tylosin biosynthetic cluster of *Streptomyces fradiae*. *Chem Biol.* (1999). 6 (9): 617-624.

- Bate, N., Butler, A.R., Smith, I. P. et Cundliffe, E. The mycarose-biosynthetic genes of *Streptomyces fradiae*, producer of tylosin. *Microbiology* 5 (2000). 146 (Pt 1), 139-146.

- Bayley C., Morgan, Dale E. C., Ow D. W. Exchange of gene activity in transgenic plants catalysed by the Cre-lox site specific system. *Plant Mol. Biol* (1992). 18 :353-361.

- Bentley, S. D., Chater, K.F., C rdeno-Tarraga, A. M., Challis, G.
10 L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H, Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbino-
15 witsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J. et Hopwood, D. A. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* (2002). 417 (6885), 141-147.

- Bibb MJ, Findlay PR, Johnson MW. The relationship between
20 base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene.* (1984) 30 (1-3):157-166.

- Bibb MJ, Janssen GR, Ward JM. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene.* (1985). 38 (1-3): 215-26.
25

- Bibb MJ, White J, Ward JM, Janssen GR. The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Mol Microbiol.* (1994). 14 (3):533-45.

- Bierman M, Logan R, O'Brien K, Seno ET, Rao RN, Schoner BE. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene.* 1992;116(1):43-9.
30

- Blondelet-Rouault M-H., Weiser J., Lebrihi A., Branny P. and Pernodet J-L. Antibiotic resistance gene cassettes derived from the Ω interposon for use in *E. coli* and *Streptomyces*. *Gene* (1997) 190 : 315-317.
- Boccard F., Smokvina T., J.-L., Pernodet Fridmann A., and Guérineau M., Structural analysis of loci involved in pSAM2 site-specific integration in *Streptomyces*. *Plasmid*, (1989a). 21 : 59-70.
- Boccard F., Smokvina T., J.-L., Pernodet Fridmann A., and Guérineau M.. The integrated conjugative plasmid pSAM2 of *Streptomyces ambofaciens* is related to temperate bacteriophages. *EMBO J.* (1989b) 8 : 973-980.
- Brunelli J.P., Pall M.L., A series of Yeast/ *E. coli* lambda expression vectors designed for directional cloning of cDNA and cre/lox mediated plasmid excision. *Yeast.* (1993) 9:1309-1318.
- Camilli A., Beattie D. T. et Mekalanos J. J. Use of genetic recombination as a reporter of gene expression. *Proc. Natl. Acad. Sci. USA* (1994) 91(7), 2634-2638.
- Campelo,A.B. and Gil,J.A. The candicidin gene cluster from *Streptomyces griseus* IMRU 3570. *Microbiology* (2002) 148 (Pt 1), 51-59.
- Carreras C, Frykman S, Ou S, Cadapan L, Zavala S, Woo E, Leaf T, Carney J, Burlingame M, Patel S, Ashley G, Licari P. *Saccharopolyspora erythraea*-catalyzed bioconversion of 6-deoxyerythronolide B analogs for production of novel erythromycins. *J Biotechnol.* (2002). 92(3): 217-28.
- Chao K.-M., Pearson W.R., Milier W. Aligning two sequences within a specified diagonal band. *Comput. Appl. Biosci.* (1992) 8 : 481-487.
- Chater K. F. The improving prospects for yield increase by genetic engineering in antibiotic-producing *Streptomyces*. *Biotechnology.* (1990). 8 (2): 115-121.
- Chaverroche MK, Ghigo JM, d'Enfert C. A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res.* 2000; 28(22) :E97.
- Church CM, Gilbert W. Genomic sequencing. *Proc Natl Acad*

- Sci USA. (1984) 81 (7):1991-5.
- Churchward G, Belin D, Nagamine Y. A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene*. (1984) 31 (1-3): 165-71.
 - 5 - Comstock, L.E., Coyne, M.J., Tzianabos, A.O. and Kasper, D.L. Interstrain variation of the polysaccharide B biosynthesis locus of *Bacteroides fragilis*: characterization of the region from strain 638R. *J. Bacteriol.* (1999). 181 (19): 6192-6196.
 - Cox KL, Baltz RH. Restriction of bacteriophage plaque formation in *Streptomyces* spp. *J Bacteriol.* (1984) 159(2): 499-504.
 - 10 - Dale E.C., Ow D.W., Gene transfer with subsequent removal of the selection gene from the host genome. *Proc. Natl. Acad. Sci. USA* (1991). 88:10558-10562.
 - Doumith M, Weingarten P, Wehmeier UF, Salah-Bey K, Benhamou B, Capdevila C, Michel JM, Piepersberg W, Raynal MC. Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in *Saccharopolyspora erythraea*. *Mol Gen Genet.* (2000) 264(4): 477-485.
 - 15 - Draeger, G., Park, Streptomyces-H.H. et Floss, H.G. Mechanism of the 2-deoxygenation step in the biosynthesis of the deoxyhexose moieties of the antibiotics granaticin and oleandomycin. *J. Am. Chem. Soc.* (1999) 121: 2611- 2612.
 - 20 - Gandecha, A.R., Large, S.L. et Cundliffe, E. .Analysis of four tylosin biosynthetic genes from the tyllM region of the *Streptomyces fradiae* genome. *Gene*. (1997). 184 (2) :197-203.
 - 25 - Geistlich, M., Losick, R., Turner, J.R. et Rao, R.N. Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in *Streptomyces ambofaciens*. *Mol. Microbiol.* (1992). 6 (14) 2019-2029.
 - Gourmelen A, Blondelet-Rouault MH, Permodet JL. Characterization of a glycosyl transferase inactivating macrolides, encoded by *gimA* from *Streptomyces ambofaciens*. *Antimicrob Agents Chemother.* (1998). 42(10): 2612-9.
 - 30

- Huang X. et Miller W. A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.* (1991). 12 337-357.
- Hara O et Hutchinson CR. Cloning of midecamycin(MLS)-resistance genes from *Streptomyces mycarofaciens*, *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Antibiot (Tokyo)*. (1990). 43 (8):977-991.
- Hara O et Hutchinson CR. A macrolide 3-O-acyltransferase gene from the midecamycin-producing species *Streptomyces mycarofaciens*. *J Bacteriol.* (1992). 174(15) :5141-5144.
- 10 - Hoffmeister D, Ichinose K, Domann S, Faust B, Trefzer A, Drager G, Kirschning A, Fischer C, Kunzel E, Bearden D, Rohr J et Bechthold A. The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. *Chem Biol.* (2000). 7 (11):821-831.
- 15 - Hopwood, D.A. Future possibilities for the discovery of new antibiotics by genetic engineering. *Iri Beta-lactam antibiotics*. Edited by MR.J. Salton and G.D. Shockman. New York: Academic Press. (1981). 585-598.
- Hopwood D. A., Malpartida F., Kieser H. M., Ikeda H., Duncan J., Fujii I., Rudd A. M., Floss H. G. et Omura S. Production of 'hybrid' antibiotics by genetic engineering. *Nature*. (1985a). 314 (6012): 642-644.
- 20 - Hopwood D. A., Malpartida F., Kieser H. M., Ikeda H. et Omura S. (1985b) In *Microbiology* (ed S. Silver). American Society for Microbiology, Washington D. C., 409-413.
- 25 - Houben Weyl, 1974, in *Meuthode der Organischen Chemie*, E. Wunsch Ed., Volume 15-I et 15-II,
- Hutchinson, C.R. Prospects for the discovery of new (hybrid) antibiotics by genetic engineering of antibiotic-producing bacteria. *Med Res Rev.* (1988). 8 (4): 557-567.
- 30 - Hutchinson C. R., Borell C. W., Otten S. L., Stutzman-Engwall K. J. et Wang Y. Drug discovery and development through the genetic engineering of antibiotic-producing microorganisms *J. Med. Chem.* (1989)32 (5) 929-937.

- Ishikawa J, Hotta K. FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol Lett.* (1999) 174(2):25 1-3.
- Kieser, T, Bibb, MJ, Buttner MJ, Chater KF, Hopwood DA. Practical *Streptomyces* Genetics. 2000. The John Innes Foundation, Norwich UK.
- Kuhstoss et al. Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene.* (1996). 183(1-2): 23 1-236.
- Lacalle RA, Pulido D, Vara J, Zalacain M, Jimenez A. Molecular analysis of the *pac* gene encoding a puromycin N-acetyl transferase from *Streptomyces alboniger*. *Gene.* (1989). 79(2):375-380.
- Lakso M, Sauer B, Mosinger B Jr, Lee EJ, Manning RW, Yu SH, Mulder KL, Westphal H. Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc Natl Acad Sci USA.* (1992) 89 (14) :6232-6.
- Li,T.B., Shang,G.D., Xia,H.Z. and Wang,Y.G. Cloning of the sugar related biosynthesis gene cluster from *Streptomyces tenebrarius* H6. *Sheng Wu Gong Cheng Xue Bao* (2001). 17 (3) : 329-33 1.
- Ligon,J., Hill,S., Beck,J., Zirkle,R., Molnar,I., Zawodny,J., Money,S. and Schupp,T. Characterization of the biosynthetic gene cluster for the antifungal polyketide soraphen A from *Sorangium cellulosum* So ce26. *Gene* (2002). 285 (1-2), 257-267.
- Liu L, Saevels J, Louis P, Nelis H, Rico S, Dierick K, Guyomard S, Roets E, Hoogmartens J. Interlaboratory study comparing the microbiological potency of spiramycins I, II and III. *J Pharm Biomed Anal.* (1999). 20 (1-2):217-24.
- Mazodier P, Petter R, Thompson C. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J Bacteriol.* (1989). 171 (6):3583-5.
- Merrifield RB, 1965a, *Nature*, 207(996): 522-523.
- Merrifield RB., 1965b, *Science*, 150(693): 178-185.

- Merson-Davies, L. A. et Cundliffe, E. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome, *Molecular microbiology*. (1994). 13 (2) : 349-355.
- Miller VL, Mekalanos JJ. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol*. (1988) 170(6):2575-83.
- Nakano, Y., Yoshida, Y., Yamashita, Y. & Koga, T. Construction of a series of pACYC-derived plasmid vectors. *Gene* (1995). 162:157-8.
- Nielsen H., Engelbrecht J., Brunak S. et von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering*. (1997). 10:1-6.
- Oh SH, Chater KF. Denaturation of circular or linear DNA facilitates targeted integrative transformation of *Streptomyces coelicolor* A3(2): possible relevance to other organisms. *J Bacteriol*. 1997; 179 (1):122-7.
- Olano C, Lomovskaya N, Fonstein L, Roll JT, Hutchinson CR. A two-plasmid system for the glycosylation of polyketide antibiotics: bioconversion of epsilon-rhodomyacinone to rhodomyacin D. *Chem Biol*. (1999). 6 (12):845-55.
- Omura S, Kitao C, Hamada H, Ikeda H. Bioconversion and biosynthesis of 16-membered macrolide antibiotics. X. Final steps in the biosynthesis of spiramycin, using enzyme inhibitor: cerulenin. *Chem Pharm Bull (Tokyo)*. (1979a). 27(1): 176-82.
- Omura S, Ikeda H, Kitao C. Isolation and properties of spiramycin 1 3-hydroxyl acylase from *Streptomyces* and *ambofaciens*. *J Biochem (Tokyo)*. (1979b). 86(6): 1753-8.
- Omura,S., Ikeda,H., Ishikawa,J., Hanamoto,A., Takahashi,C., Shinose,M., Takahashi,Y., Horikawa,H., Nakazawa,H., Osonoe,T., Kikuchi,H., Shiba,T., Sakaki,Y. and Hattori,M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U.S.A.* (2001). 98 (21), 12215-12220.

- Pearson, W.R. and D.J. Lipman. Improved Tools for Biological Sequence Analysis. Proc. Natl. Acad. Sci. USA 1988, 85: 2444-2448.
- Pearson, W.R. Rapid and Sensitive Sequence Comparison with FASTP and FASTA. Methods in Enzymology. 1990, 183: 63-98.
- 5 - Pernodet JL, Simonet JM, Guerineau M. Plasmids in different strains of *Streptomyces ambofaciens*; free and integrated form of plasmid pSAM2. Mol Gen Genet. (1984); 198 (1): 35-41.
- Pernodet JL, Alegre MT, Blondelet-Rouault MH, Guerineau M. Resistance to spiramycin in *Streptomyces ambofaciens*, the producer organ-
10 ism, involves at least two different mechanisms. J Gen Microbiol. (1993); 139 (Pt 5): 1003-11.
- Pernodet JL, Fish S, Blondelet-Rounault MH, Cundliffe E. The macrolide-lincosamide-streptogramin B resistance phenotypes characterized by using a specifically deleted, antibiotic-sensitive strain of *Streptomyces*
15 lividans. Antimicrob Agents Chemother. (1996), 40 (3): 581-5.
- Pernodet JL, Gourmelen A, Blondelet-Rouault MH, Cundliffe E. Dispensable ribosomal resistance to spiramycin conferred by *srmA* in the spiramycin producer *Streptomyces ambofaciens*. Microbiology. (1999), 145 (Pt 9): 2355-64.
- 20 - Pfoestl, A., Hofinger, A., Kosma, P. and Messner, P. Biosynthesis of dTDP-3-acetamido-3,6-dideoxy- α -D-galactose in *Aneurinibacillus thermoaerophilus* L420-91T. J. Biol. Chem. (2003), 278 (29): 26410-26417.
- Rao RN, Richardson MA, Kuhstoss S. Cosmid shuttle vectors for cloning and analysis of *Streptomyces* DNA. Methods Enzymol. (1987);
25 153: 166-98.
- Raynal A, Tuphile K, Gerbaud C, Luther T, Guerineau M, and Pernodet JL.. Structure of the chromosomal insertion site for pSAM2: functional analysis in *Escherichia coli*. Molecular Microbiology (1998) 28 (2) 333-342.
- 30 - Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cul- lum, J., Kinashi, H. et Hopwood, D.A. A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2)

- chromosome *Mol. Microbiol.* (1996). 21 (1): 77-96.
- Richardson MA, Kuhstoss S, Solenberg P, Schaus NA, Rao RN. A new shuttle cosmid vector, pKC505, for streptomycetes: its use in the cloning of three different spiramycin-resistance genes from a *Streptomyces ambofaciens* library. *Gene*. (1987); 61 (3): 231-41.
 - Robinson, J. A. Enzymes of Secondary Metabolism in Microorganisms. *Chem Soc Rev.* (1988). 17:383-452.
 - Russell SH, Hoopes JL, Odell JT. Directed excision of a transgene from the plant genome. *Mol Gen Genet.* (1992). 234 (1): 49-59.
 - Sambrook J, Frisch EF, Maniatis T. *Molecular Cloning: a laboratory manual* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York USA (1989).
 - Schoner B, Geistlich M, Rosteck P Jr, Rao RN, Seno E, Reynolds P, Cox K, Burgett S et Hersherberger C. Sequence similarity between macrolide-resistance determinants and ATP-binding transport proteins. *Gene*. (1992). 115 (1-2), 93-96.
 - Sezonov G, Hagege J, Pernodet JL, Friedmann A and Guerinneau M. Characterization of *pra*, a gene for replication control in pSAM2, the integrating element of *Streptomyces ambofaciens*. *Mol Microbiol.* (1995). 17(3): 533-44.
 - Sezonov G., Blanc V., Bamas-Jacques N., Friedmann A., Pernodet J.L. and M. Guerinneau, Complete conversion of antibiotic precursor to pristinamycin IIA by over expression of *Streptomyces pristinaespiralis* biosynthetic genes. *Nature Biotechnology.* (1997) 15: 349-353.
 - Sezonov G, Possoz C, Friedmann A, Pernodet JL, Guerinneau M. KorSA from the *Streptomyces* integrative element pSAM2 is a central transcriptional repressor: target genes and binding sites. *J Bacteriol.* (2000). 182 (5): 1243-50.
 - Simon R., Priefer U and Pühler. A broad host range mobilisation system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/Technology* (1983). 1:784-791.
 - Simon, et al., Plasmid vectors for the genetic analysis and ma-

nipulation of rhizobia and other gram-negative bacteria, p. 640-659. In A. Weissbach, and H. Weissbach (eds.), *Methods in enzymology*, vol 118, Academic Press, Inc., Orlando, 1986.

- Summers, R.G., Donadio, Streptomyces, Staver, M.J., Wendt-
- 5 Pienkowski, E., Hutchinson, C.R et Katz, L. Sequencing and mutagenesis of genes from the erythromycin biosynthetic gene cluster of *Saccharopolyspora erythraea* that are involved in L-mycarose and D-desosamine production. *Microbiology* (1997). 143 (Pt 10): 3251-3262.
- Van Mellaert L, Mei L, Lammertyn E, Schacht S, Anne J. Site-
- 10 specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. *Microbiology*. (1998). 144 (Pt 12): 3351-8.
- Vara J, Lewandowska-Skarbek M, Wang YG, Donadio S, Hutchinson CR. Clonning of genes governing the deoxysugar portion of the
- 15 erythromycin biosynthesis pathway in *Saccharopolyspora erythraea* (*Streptomyces erythreus*). *J. Bacteriol.* (1989). 171 (11): 5872-81.
- Vara J, Malpartida F, Hopwood DA, Jimenez A. Cloning and expression of a puromycin N-acetyl transferase gene from *Streptomyces alboniger* in *Streptomyces lividans* and *Escherichia coli*. *Gene*. (1985). 33(2):
- 20 197-206.
- Wahl, G. M., K. A. Lewis, J. C. Ruiz, B. Rothenberg, J. Zhao, and G. A. Evans. Cosmid vectors for rapid genomic walking restriction mapping, and gene transfer. *Proc. Natl. Acad. Sci. USA* (1987). 84: 2160-2164.
- Waldron, C., Matsushima, P., Rosteck, P.R., Broughton, M.C.,
- 25 Turner, J., Madduri, K., Crawford, K. P., Merlo, D.J. et Baltz, R. H. Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa* (1) *Chem. Biol.* (2001) 8 (5): 487-499.
- Walczak RJ, Hines JV, Strohl WR, Priestley ND. Bioconversion of the anthracycline analogue desacetyladiamycin by recombinant DoxA. a
- 30 P450-monooxygenase from *Streptomyces* sp. strain C5. *Org Lett.* (2001). 3 (15): 2277-9.
- Wang, Z. X., Li, S.M. et Heide, L. Identification of the coumer-

mycin A(1) biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489. *Antimicrob. Agents Chemother.* (2000) 44 (11), 3040-3048.

- Ward, J.M., Janssen, G.R., Kieser, T, Bibb, M.J., Buttner, M.J. & Bibb, M.J. Construction and characterization of a series of multi-copy promoter-probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase from Tn5 as indicator. *Mol Gen Genet* (1986). 203: 468-478.
- Wehmeier UF. New multifunctional *Escherichia coli*-*Streptomyces* shuttle vectors allowing blue-white screening on XGa1 plates. *Genes*. (1995). 165 (1): 149-150.
- Wilms B, Hauck A, Reuss M, Sylatk C, Mattes R, Siemann M, Altenbuchner J. High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli* rhaBAD promoter. *Biotechnol Bioeng.* (2001). 73(2): 95-103.
- Worley K. C., Wiese B. A., and Smith R. F. BEAUTY: An enhanced BLAST-based search tool that integrates multiple biological information resources into sequence similarity search results. *Genome Research* (1995). 5: 173-184.
- Wu K, Chung L, Revill WP, Katz L et Reeves CD. The FK520 gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene*. (2000). 251 (1): 81-90.
- Xue Y, Zhao L, Liu HW et Sherman DH. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc Natl Acad Sci USA.* (1998). 95 (21): 12111-12116.
- Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG et Court DL. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* (2000). 97(11): 5978-83.

REIVINDICAÇÕES

1. Célula hospedeira, caracterizada pelo fato de que é transformada ou transfectada com um vetor de expressão, permitindo a expressão de um ou vários polipeptídeos implicados na biossíntese das espiramicinas, em que os ditos polipeptídeos resultam da expressão de polinucleotídeos inseridos no referido vetor de expressão, a sequência dos ditos polinucleotídeos sendo:

(a) uma das sequências de SEQ ID N° 111 ou 141; ou

(b) uma das sequências derivadas da sequência (a) em razão da degenerescência do código genético;

em que a referida célula hospedeira é uma célula hospedeira procariótica ou uma célula hospedeira de levedura.

2. Célula hospedeira de acordo com a reivindicação 1, caracterizada pelo fato de que o polinucleotídeo é isolado a partir de uma bactéria do gênero *Streptomyces*.

3. Célula hospedeira de acordo com a reivindicação 1, caracterizada pelo fato de que o polipeptídeo apresenta uma sequência compreendendo a SEQ ID N° 142.

4. Célula hospedeira de acordo com a reivindicação 1, caracterizada pelo fato de que o polipeptídeo apresenta uma sequência compreendendo a SEQ ID N° 112.

5. Célula hospedeira de acordo com a reivindicação 1, caracterizada pelo fato de que a célula hospedeira procariótica é *Escherichia coli*.

6. Processo de produção de um polipeptídeo implicado na biossíntese das espiramicinas, em que o dito polipeptídeo possui uma sequência compreendendo SEQ ID NO: 112 ou 142 e resulta da expressão de um polinucleotídeo, a sequência do dito polinucleotídeo sendo:

(a) uma das sequências de SEQ ID N° 111 ou 141; ou

(b) uma das sequências derivadas da sequência (a) em razão da degenerescência do código genético

o processo sendo caracterizado pelo fato de compreender as seguintes etapas:

a) inserir pelo menos um dos referidos polinucleotídeos codificando o referido polipeptídeo em um vetor de expressão que permite a expressão do referido polipeptídeo;

5 b) cultivar, em um meio de cultura apropriado, uma célula hospedeira procariótica ou de levedura previamente transformada ou transfectada com o vetor da etapa a);

c) recuperar o meio de cultura acondicionado ou um extrato celular;

10 d) separar e purificar o dito polipeptídeo a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa c); e

e) se for o caso, caracterizar o polipeptídeo recombinante produzido.

7. Processo de acordo com a reivindicação 6, caracterizado pelo fato de que o polinucleotídeo é isolado a partir de uma bactéria do gênero
15 *Streptomyces*.

8. Processo de acordo com a reivindicação 6, caracterizado pelo fato de que a célula hospedeira procariótica é *Escherichia coli*.

9. Micro-organismo transgênico mutante produtor de macrolídeo, caracterizado pelo fato de que superexpressa pelo menos um gene que
20 compreende um polinucleotídeo, a sequência do dito polinucleotídeo sendo:

(a) uma das sequências de SEQ ID N° 111 ou 141; ou

(b) uma das sequências derivadas da sequência (a) em razão da degenerescência do código genético;

em que a superexpressão do referido gene é obtida pelo aumento do número de cópias para esse gene e pela introdução de um promotor
25 que é mais ativo que o promotor selvagem.

10. Micro-organismo transgênico mutante produtor de macrolídeo de acordo com a reivindicação 9, caracterizado pelo fato de que superexpressa um ou vários genes que compreende(m) uma sequência correspondendo a uma ou mais das sequências SEQ ID N° 111 ou 141, ou uma
30 das sequências derivadas destas em razão da degenerescência do código genético.

11. Micro-organismo transgênico mutante produtor de macrolídeo de acordo com a reivindicação 9 ou 10, caracterizado pelo fato de que o polinucleotídeo é isolado a partir de uma bactéria do gênero *Streptomyces*.

5 12. Micro-organismo transgênico mutante produtor de macrolídeo de acordo com qualquer uma das reivindicações 9 a 11, caracterizado pelo fato de que é uma cepa de *Streptomyces ambofaciens*.

13. Micro-organismo de acordo com a reivindicação 12, caracterizado pelo fato de que é uma cepa de *Streptomyces ambofaciens* OSC2/pSPM75(1) ou uma cepa OSC2/pSPM75(2) depositada junto à
10 Collection Nationale de Cultures de Microorganismes (CNCM) sob o número de registro I-3101.

14. Processo de produção de macrolídeos, caracterizado pelo fato de compreender as seguintes etapas:

(a) cultivar, em um meio de cultura apropriado, um micro-
15 organismo como definido em qualquer uma das reivindicações 9 a 13,

(b) recuperar o meio de cultura acondicionado ou um extrato celular, e

(c) separar e purificar esse(s) macrolídeo(s) produzido(s) a partir do meio de cultura ou ainda a partir do extrato celular obtido na etapa (b).

20 15. Utilização de uma sequência e/ou de um DNA recombinante correspondendo a uma ou mais das sequências SEQ ID N° 111 ou 141, ou uma das sequências derivadas destas em razão da degenerescência do código genético e/ou de uma célula hospedeira coma definida em qualquer uma das reivindicações 1 a 5, caracterizado por ser para a preparação de
25 antibióticos híbridos.

16. DNA recombinante, caracterizado pelo fato de compreender:

- um polinucleotídeo obtido por amplificação em cadeia por polimerase, utilizando o seguinte par de sequências iniciadoras:

5 'AAGCTTGTGTGCCCCGGTGTACCTGGGGAGC 3' (SEQ ID N° 138) e

30 5 'GGATCCCGCGACGGACACGACCGCCGCGCA 3' (SEQ ID N° 139) e, como matriz, o cosmídeo pSPM36,

- ou um fragmento de pelo menos 1200, 1400, 1450 ou 1500

nucleotídeos consecutivos desse polinucleotídeo,

o dito fragmento sendo opcionalmente usado para amplificar as sequências SEQ ID N° 111 ou 141, ou uma das sequências derivadas destas em razão da degenerescência do código genético.

5 17. Célula hospedeira, caracterizada pelo fato de que foi introduzido na mesma pelo menos um DNA recombinante como definido na reivindicação 16, a dita célula hospedeira sendo uma célula procariótica ou uma célula de levedura.

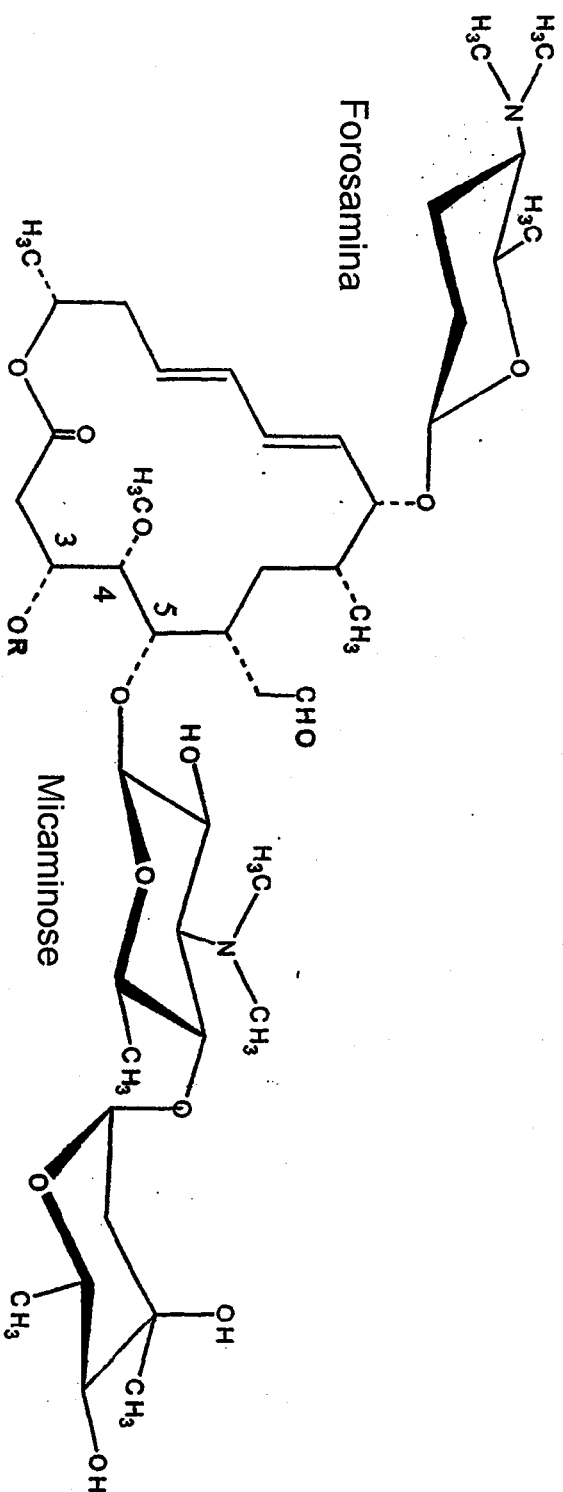
10 18. Célula hospedeira de acordo com a reivindicação 17, caracterizada pelo fato de que a célula hospedeira procariótica é *Escherichia coli*.

19. Processo de produção de espiramicina(s), caracterizado pelo fato de compreender as seguintes etapas:

(a) cultivar, em um meio de cultura apropriado, uma célula hospedeira como definida na reivindicação 17 ou 18,

15 (b) recuperar o meio de cultura acondicionado ou um extrato celular, e

(c) separar e purificar a(s) espiramicina(s) a partir desse meio de cultura ou ainda a partir do extrato celular obtido na etapa (b).



1/41

R=H espiramicina I

R=COCH₃ espiramicina II

R=COCH₂CH₃ espiramicina III

FIG. 1

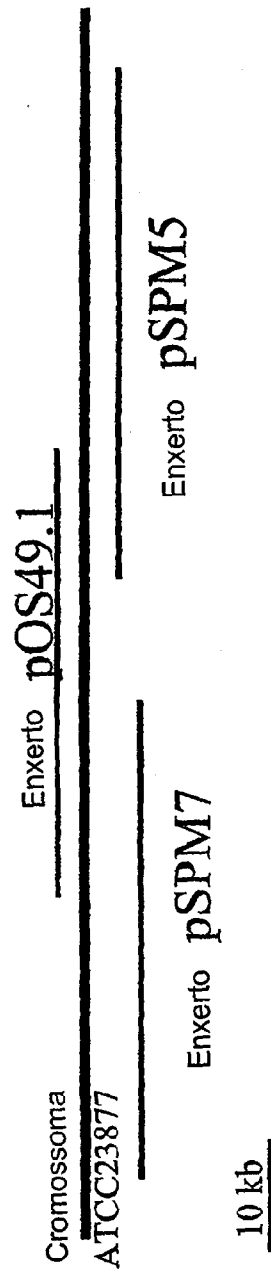


FIG. 2

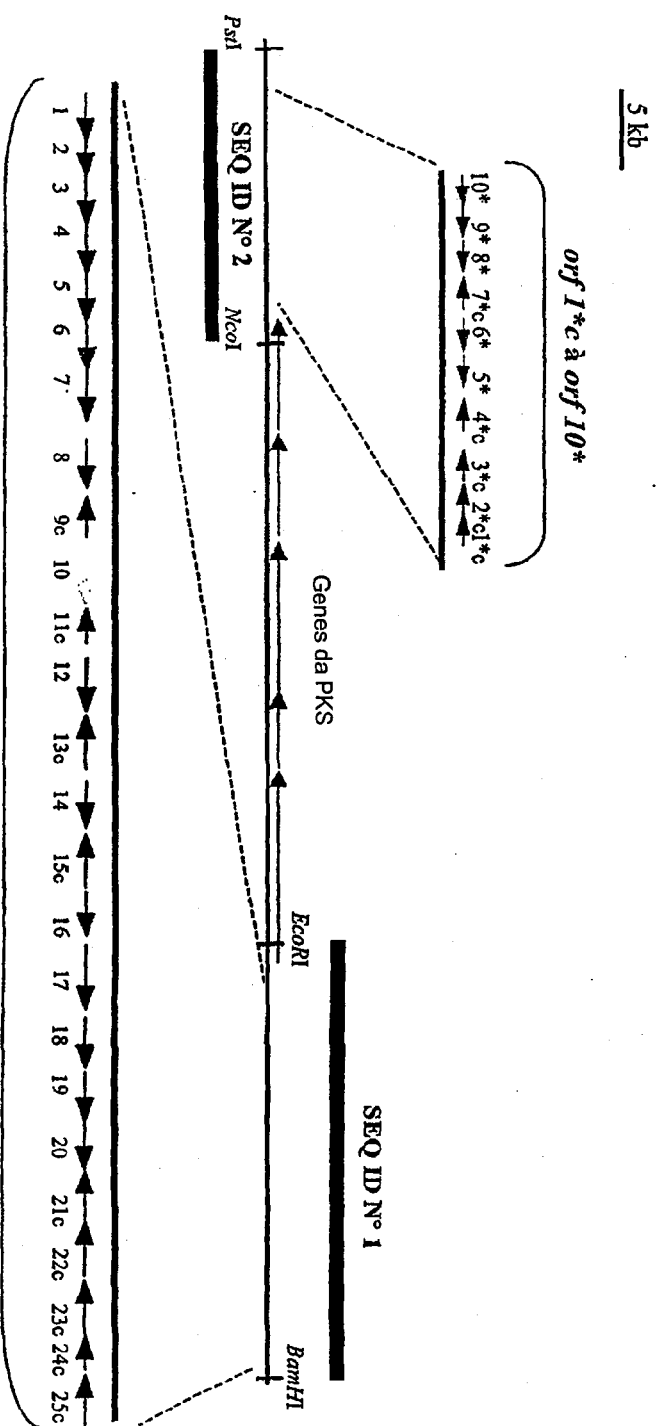


FIG. 3

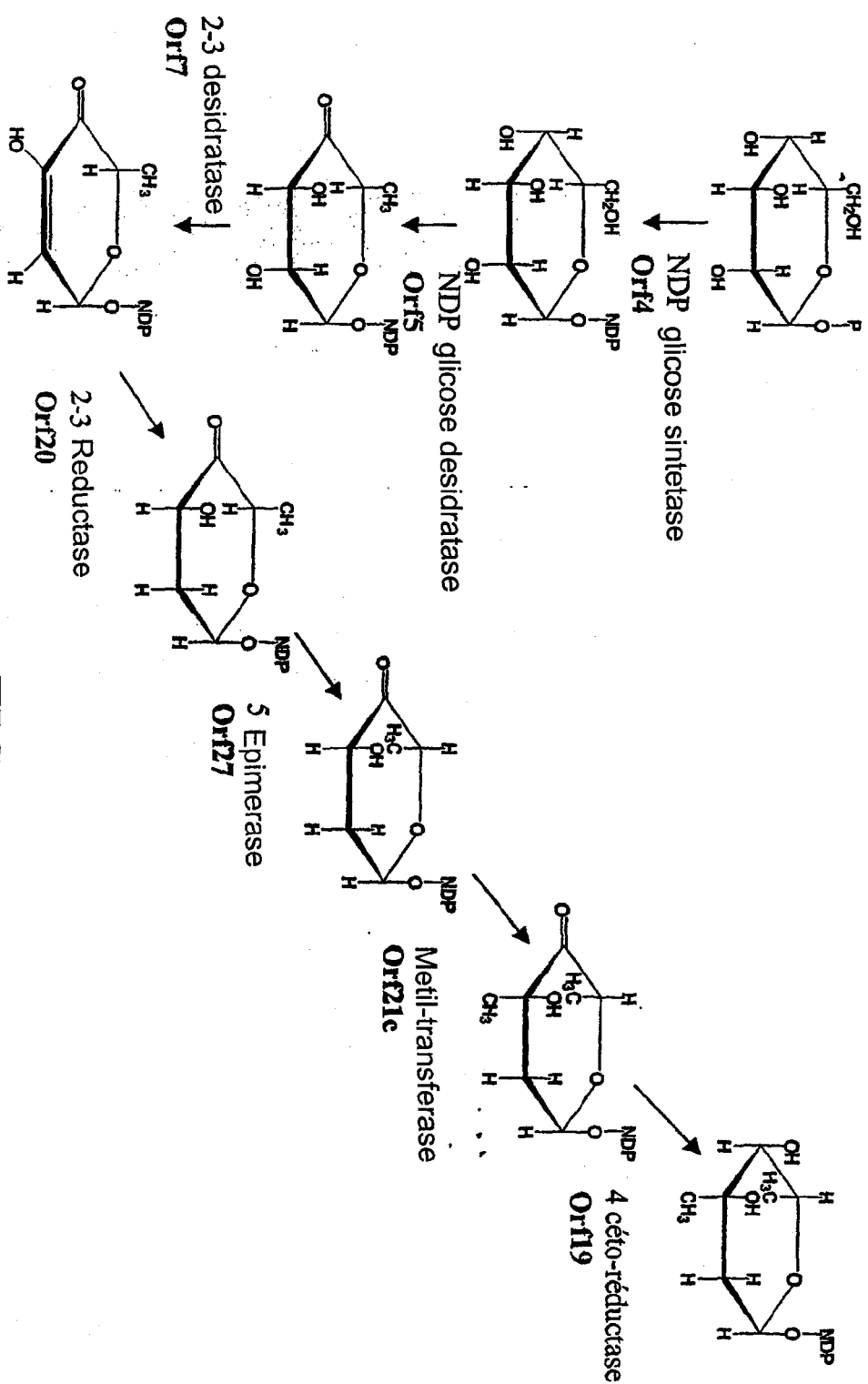


FIG. 4

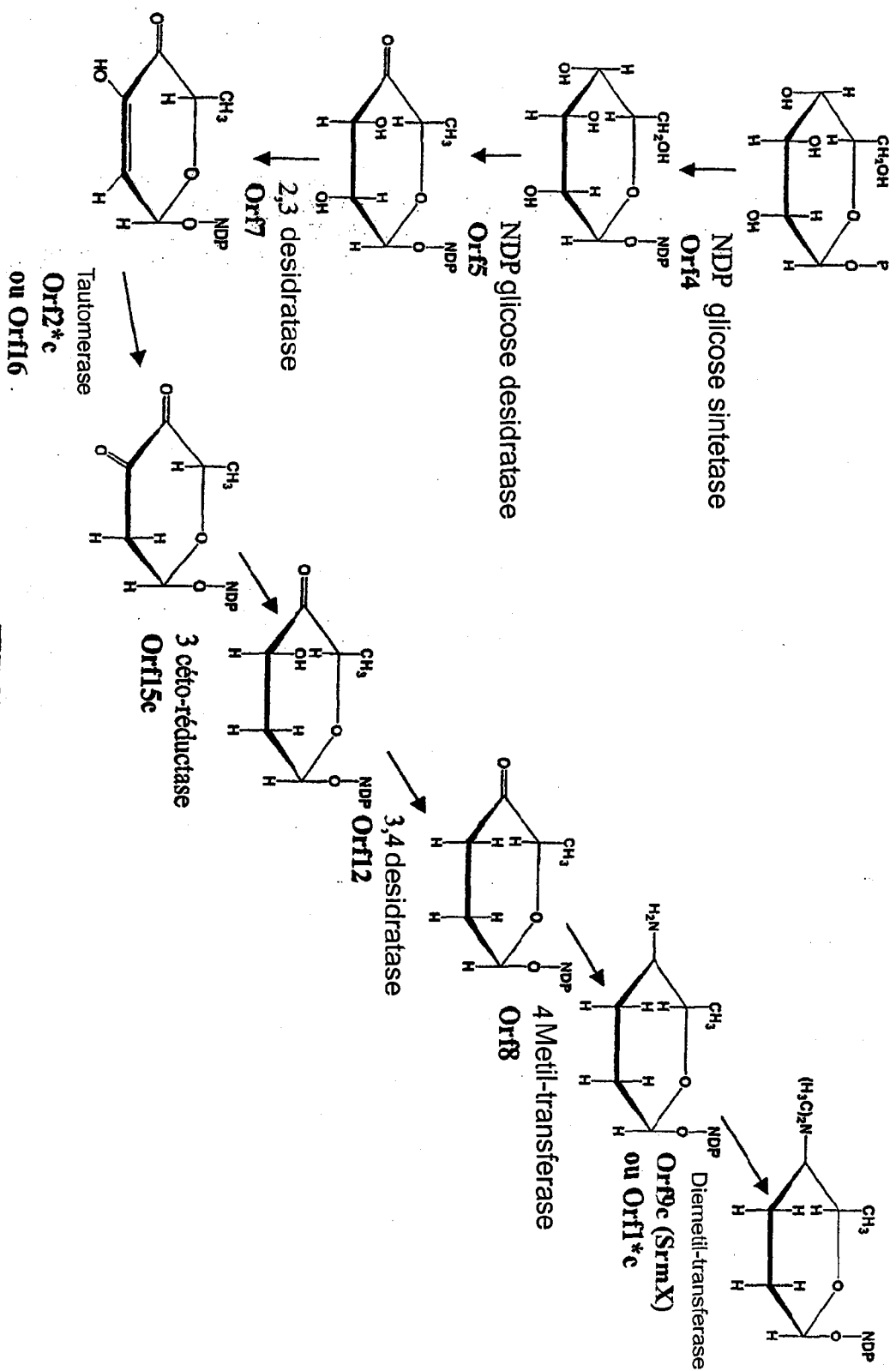


FIG. 6

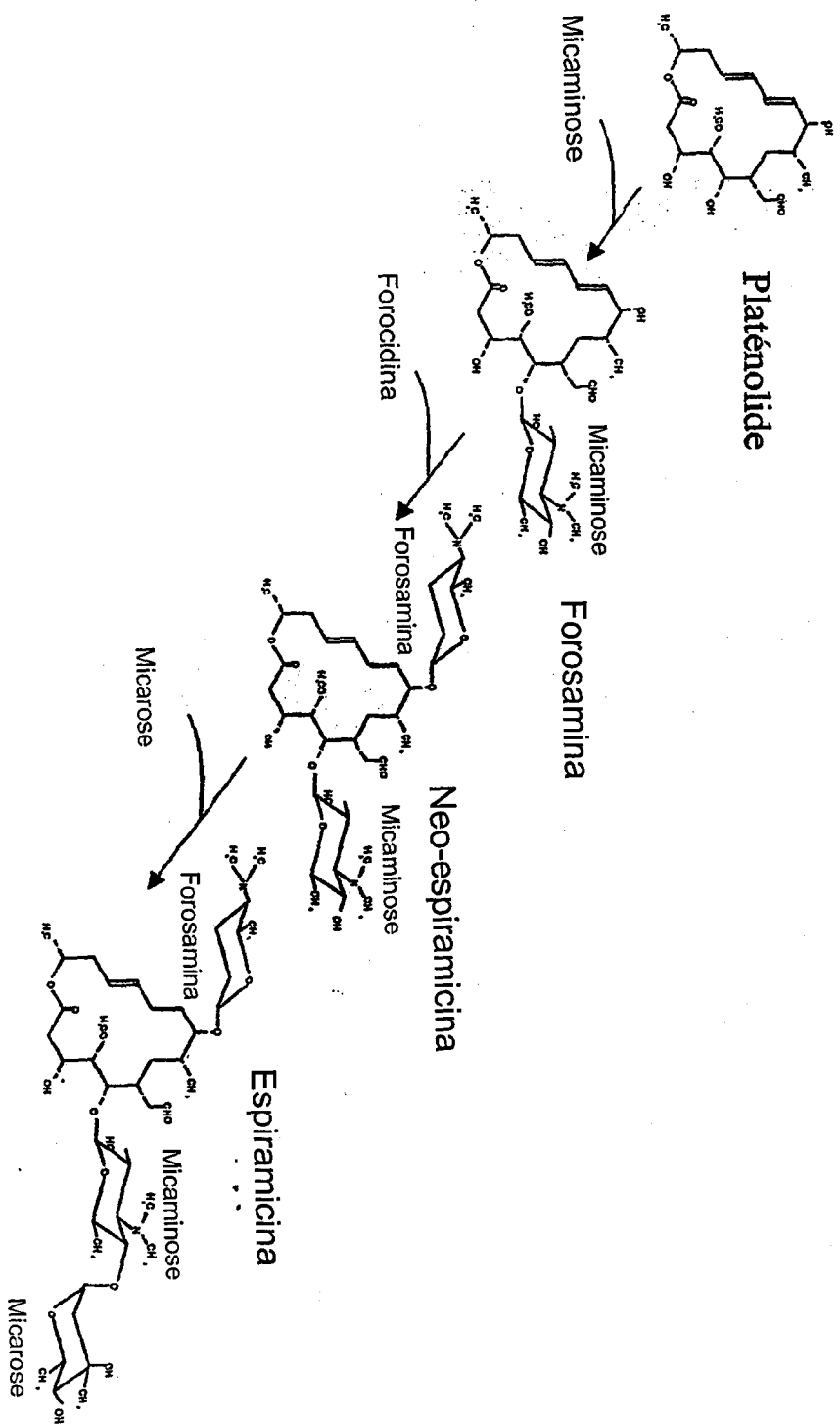


FIG. 7

Intermediário
de uma via
de glicólise

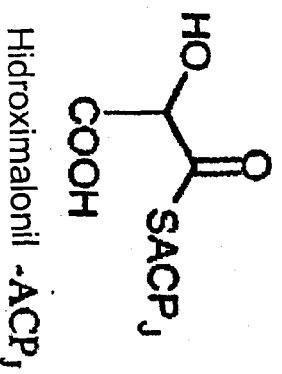
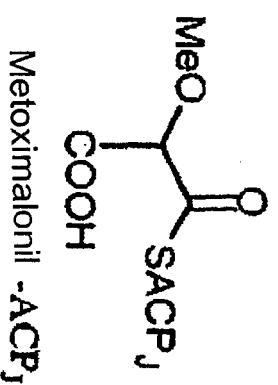
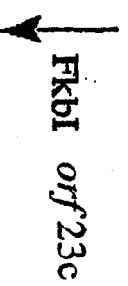
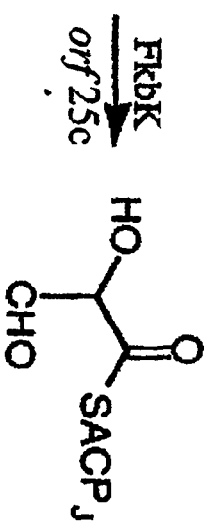
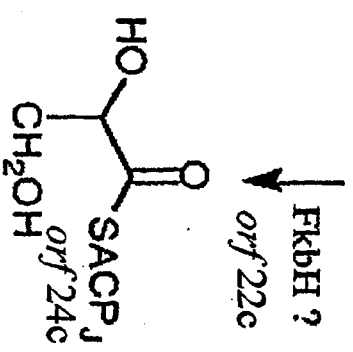


FIG. 8

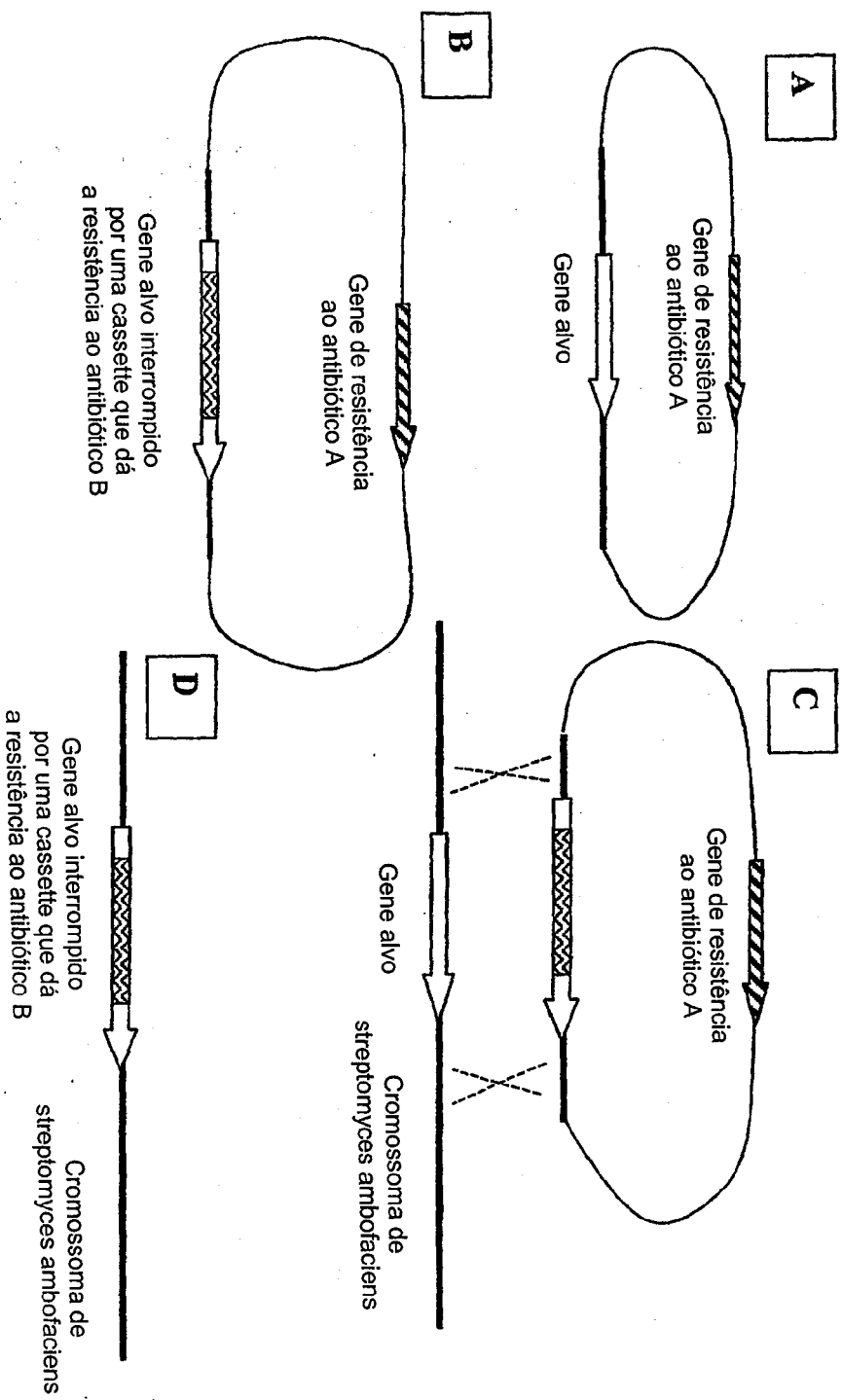


FIG. 9

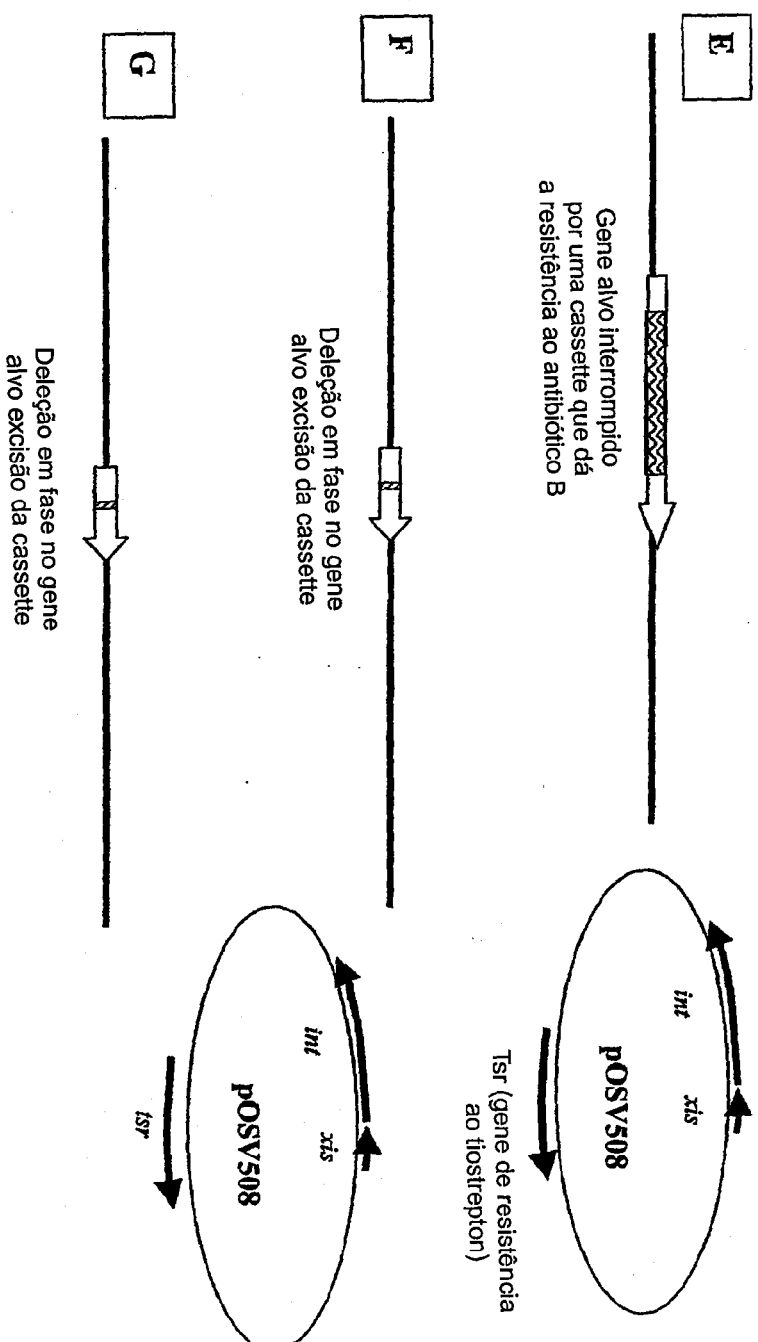


FIG. 10

Amplificação da cassette excisável

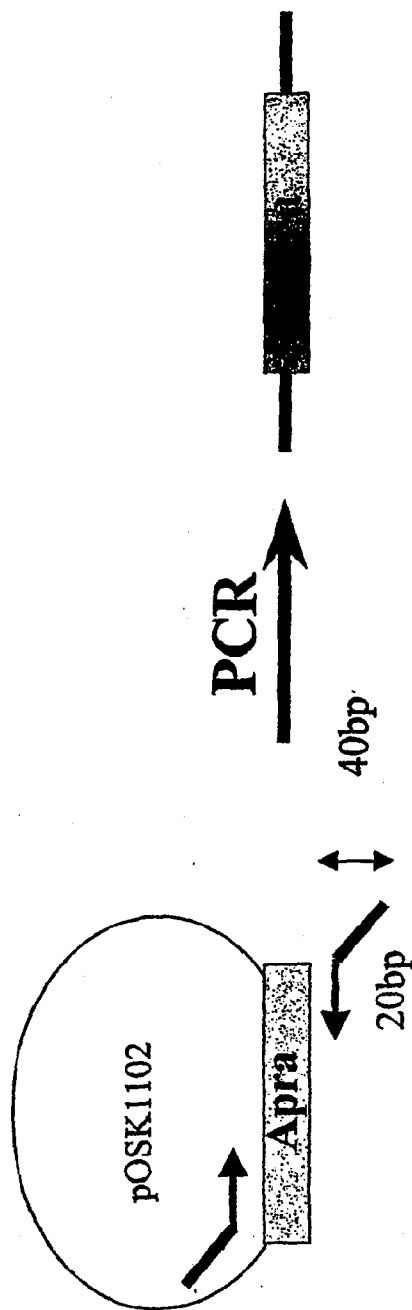


FIG. 11

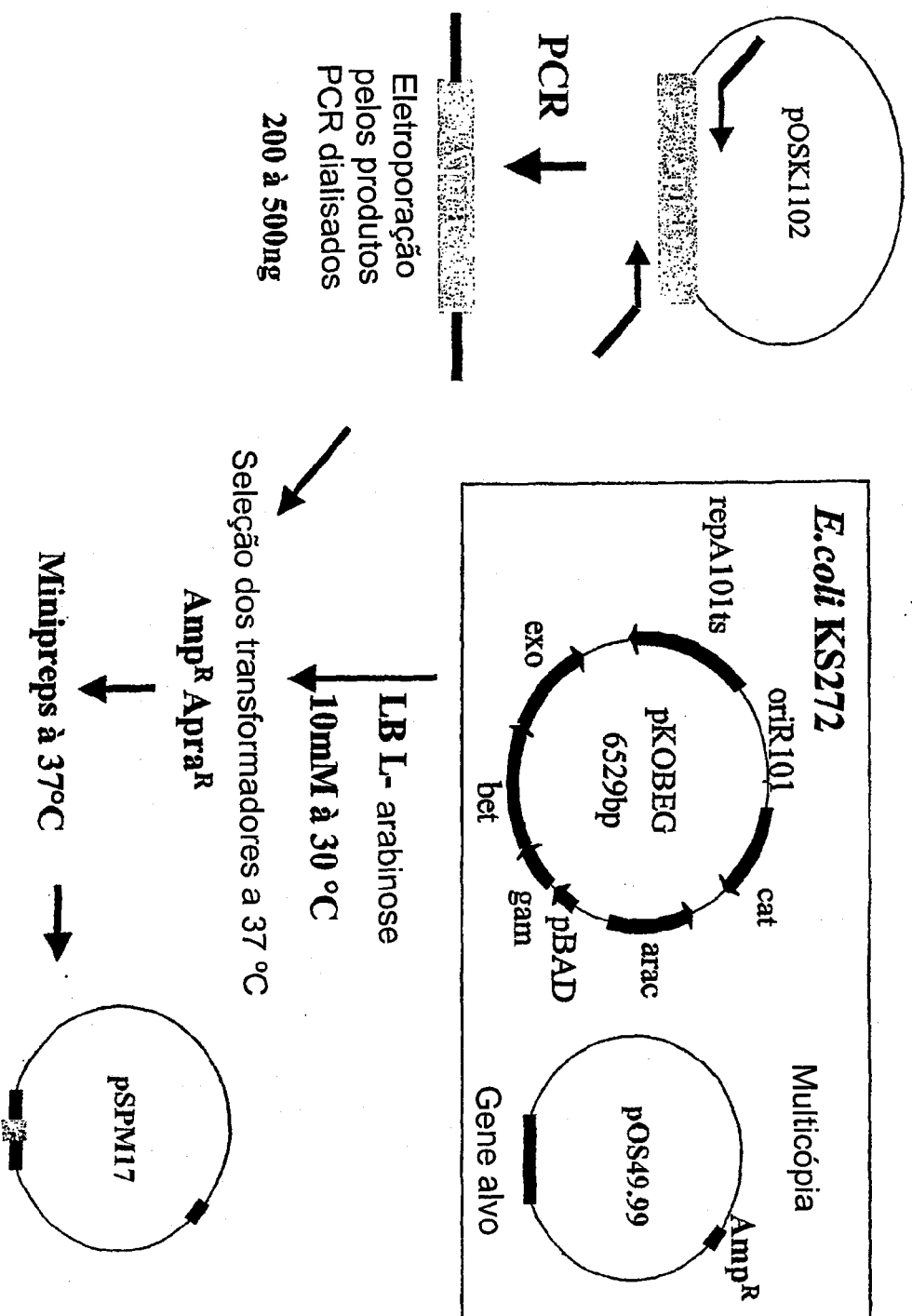


FIG. 12

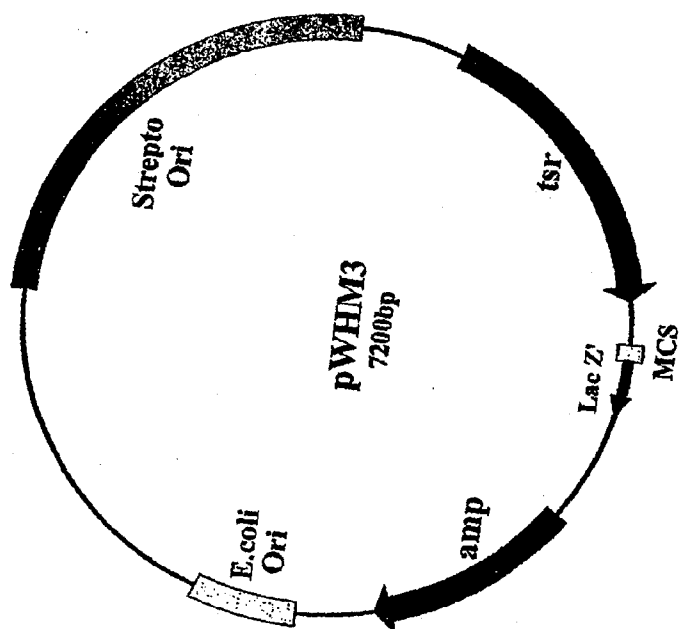


FIG. 13

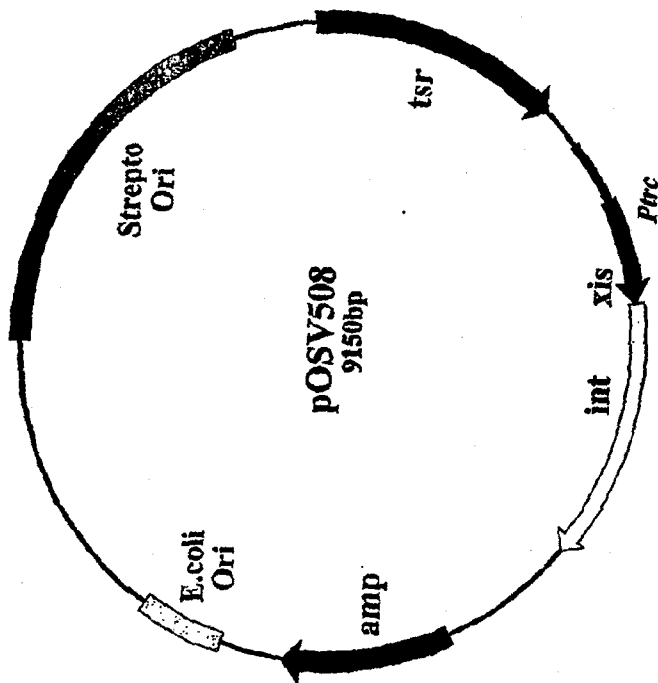


FIG. 14

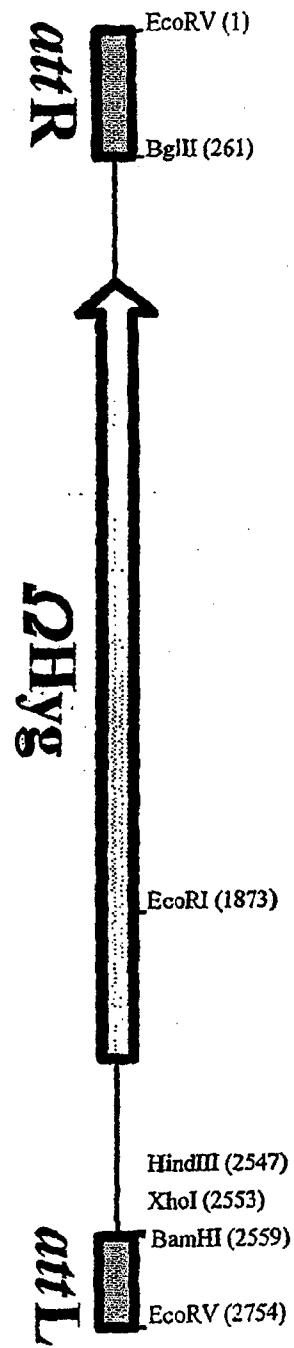


FIG. 15

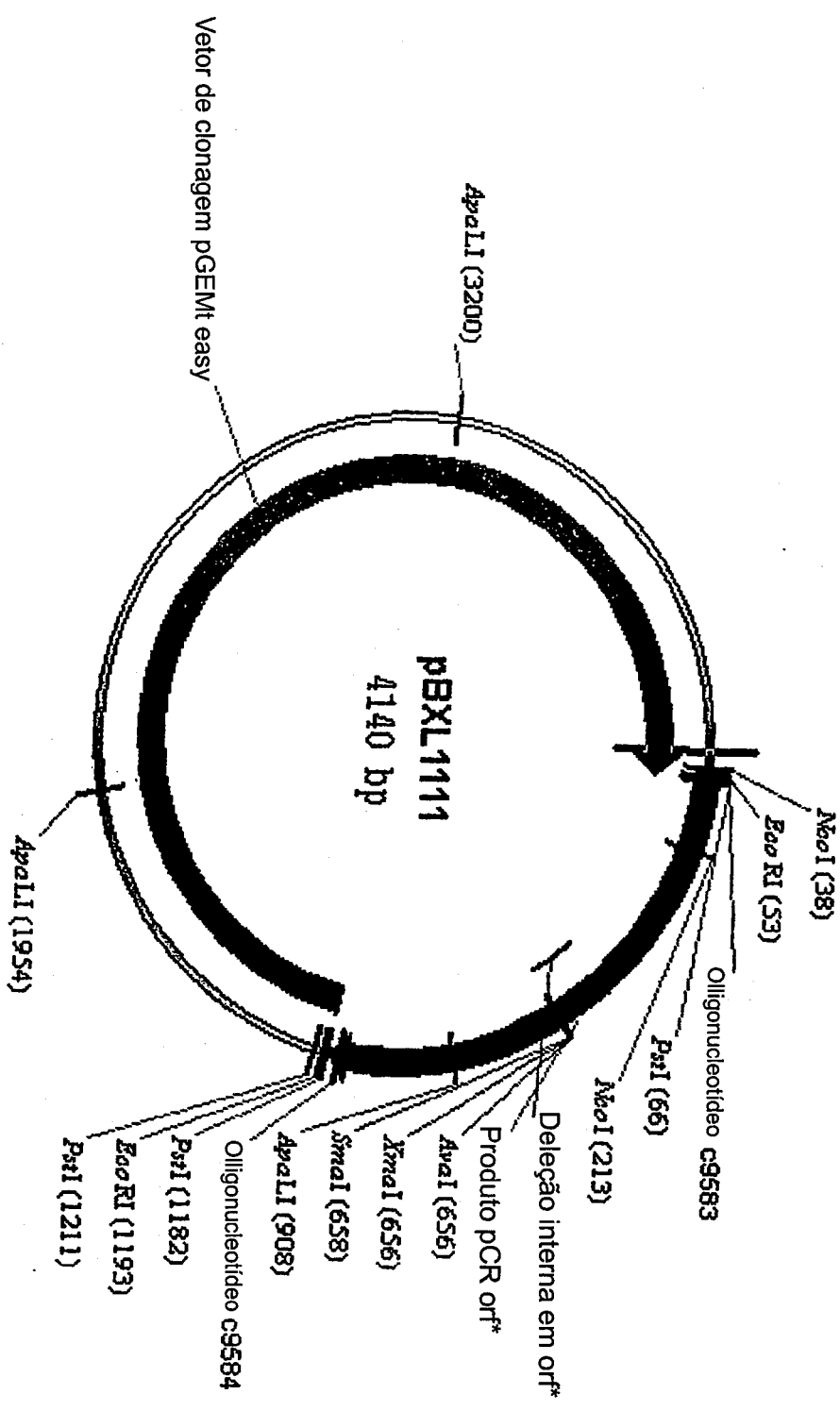


FIG. 16

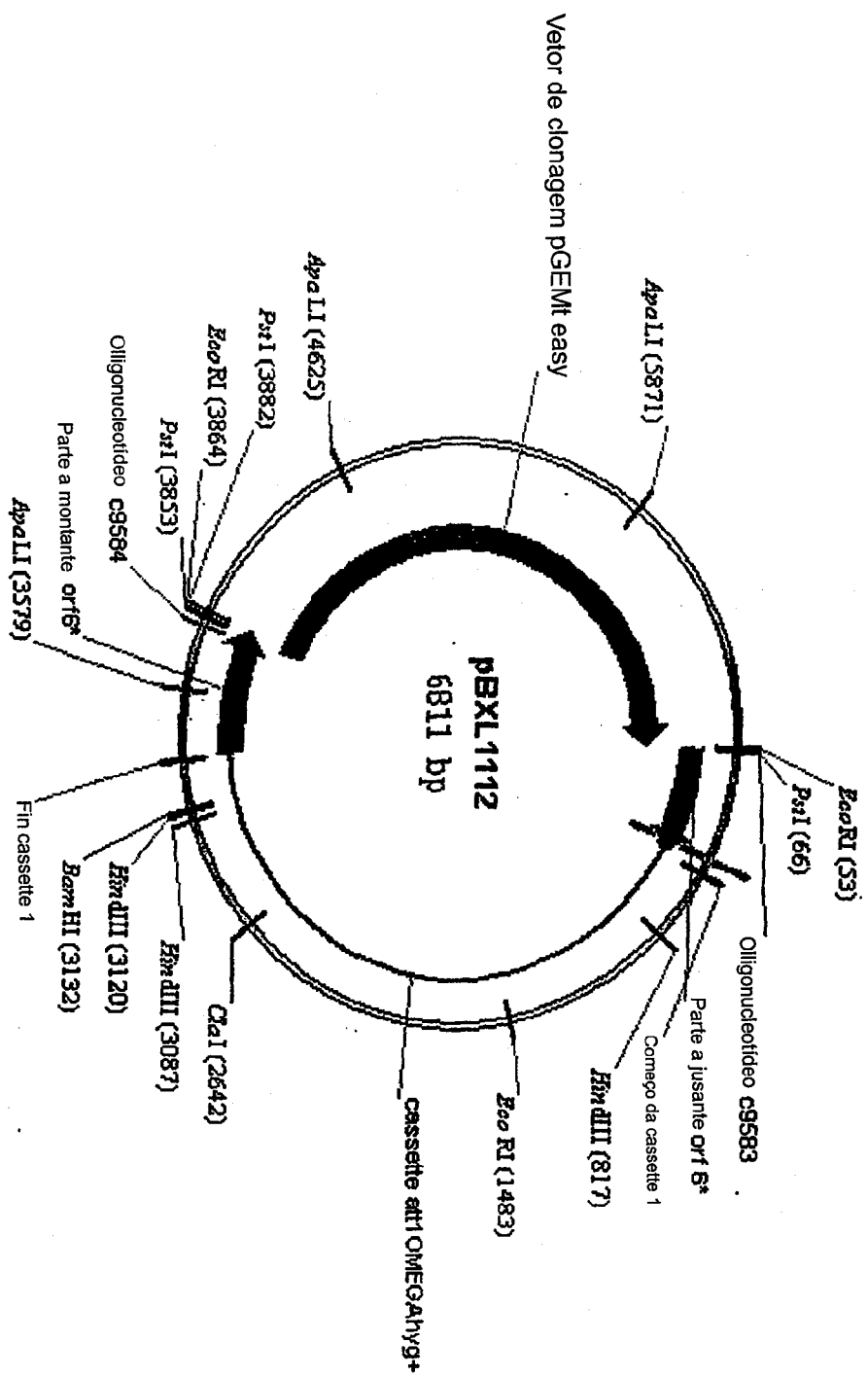
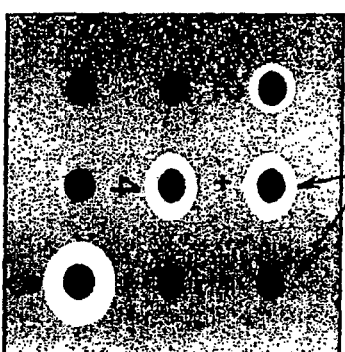


FIG. 17

70 ml MP5+2.10⁶ spores/ml
27°C; 250 rpm

t = 48h; 72h; 96h



Caixa de micrococcus luteus
37°C, 24-48h

FIG. 18

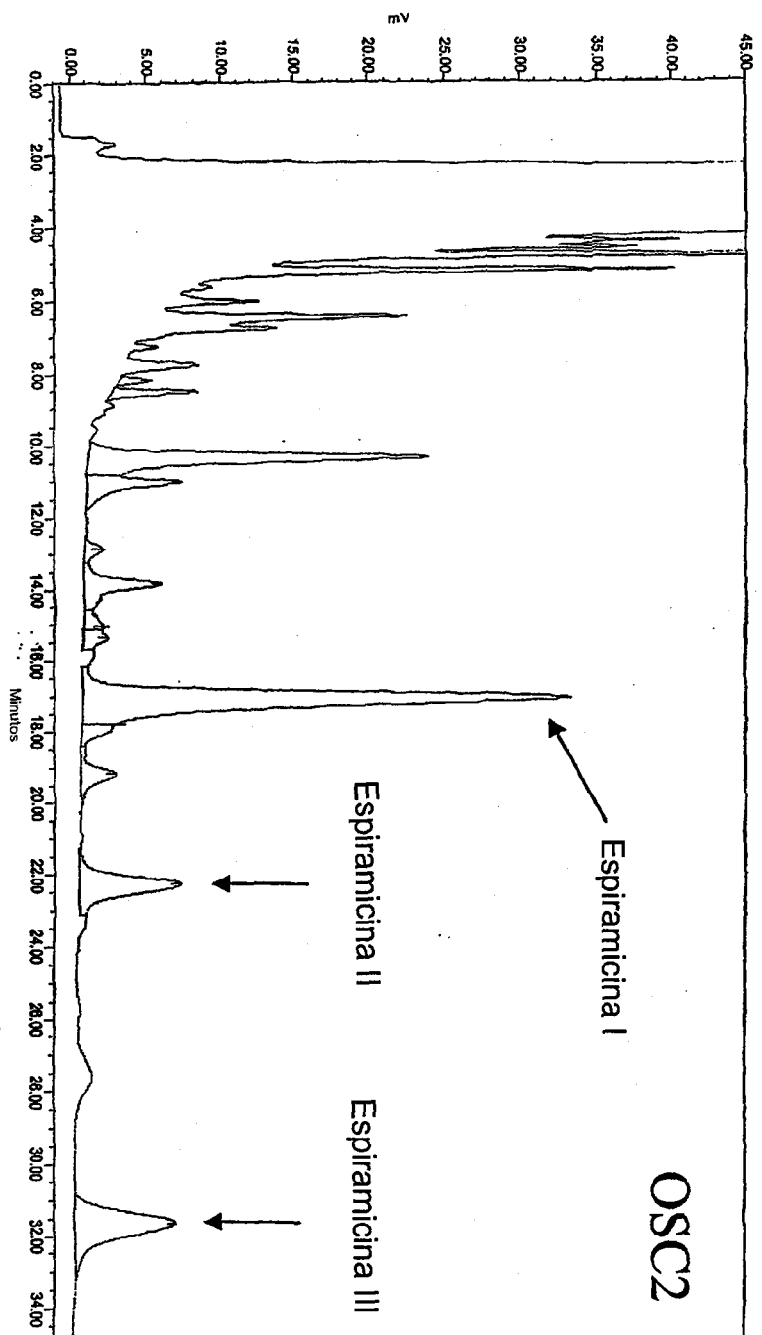


FIG. 19

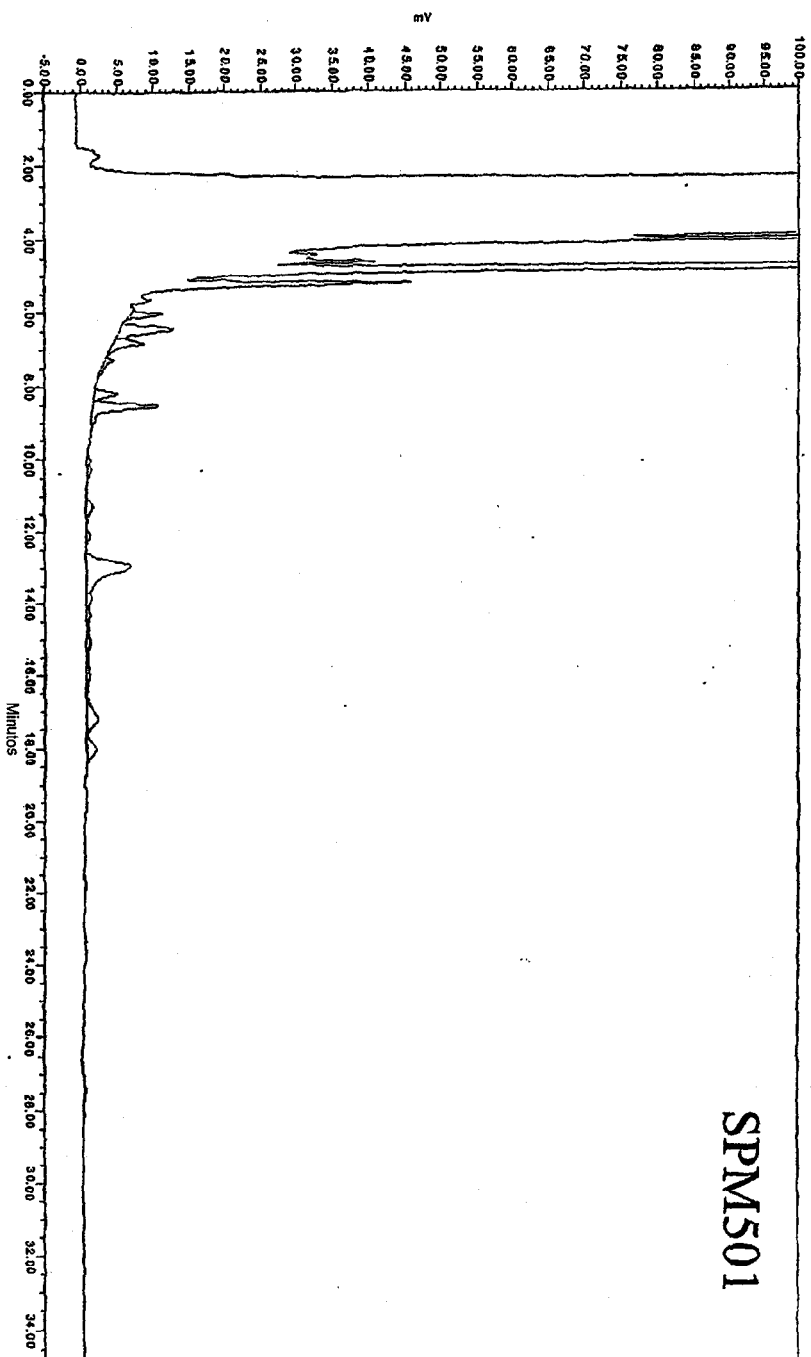


FIG. 20

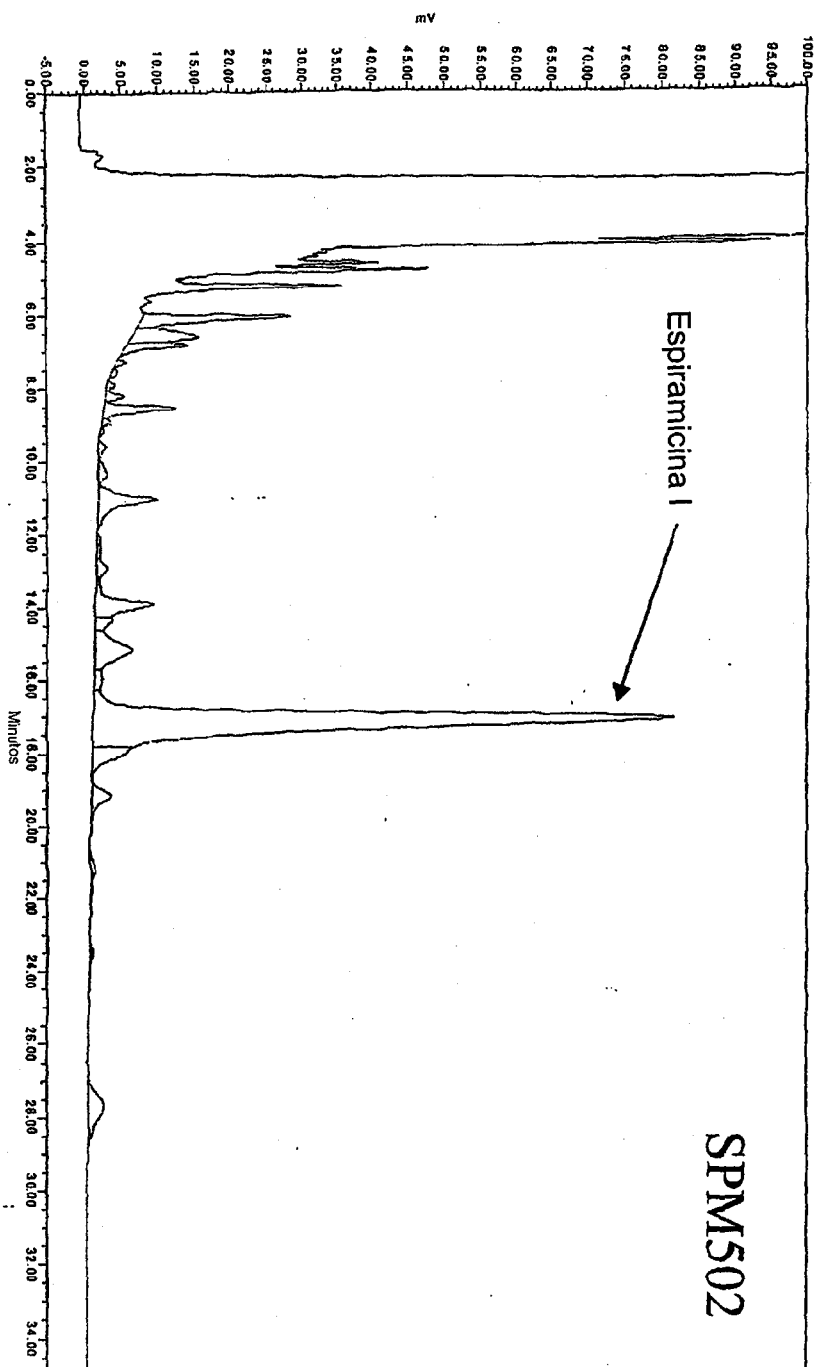


FIG. 21

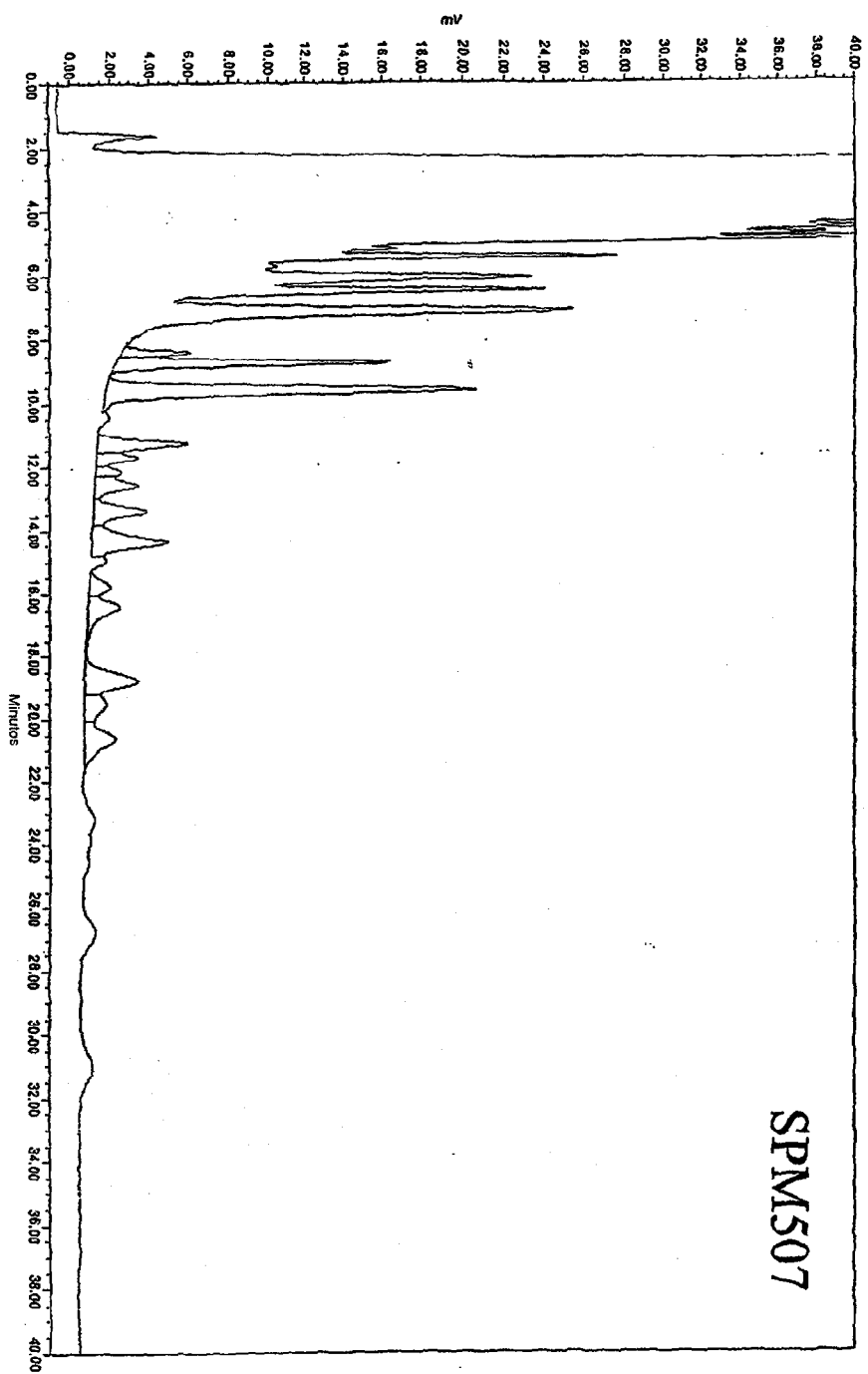


FIG. 22

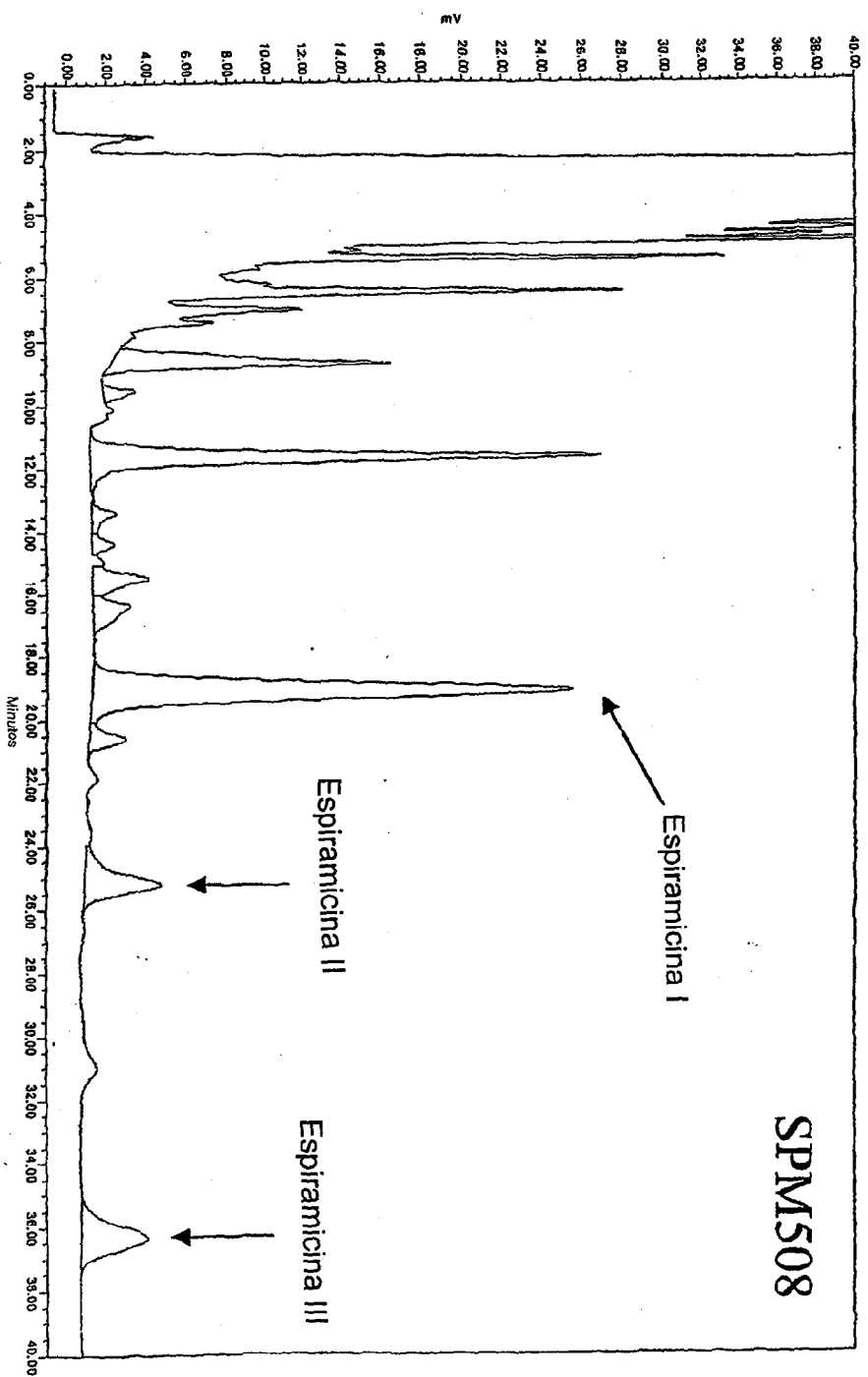


FIG. 23

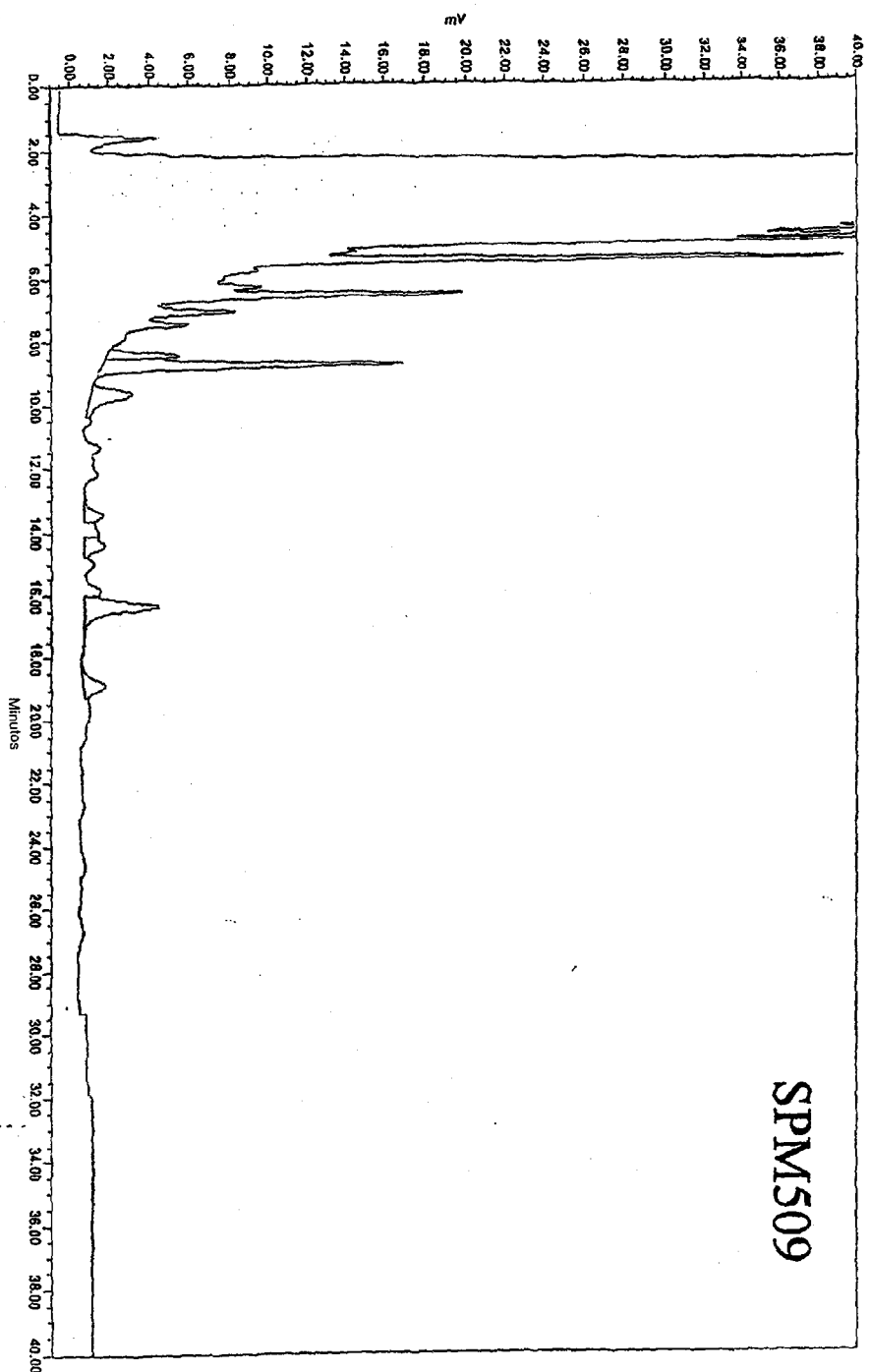


FIG. 24

FASTA (3.42 Sept 2001) function [optimized, BL50 matrix (15:-5)]
 ktup: 2 join: 37, opt: 25, open/ext: -10/-2, width: 16

Smith-Waterman score: 1467; 63.130% identity (66.480% ungapped) in
 377 aa overlap (3-367:11-380)

```

                                10      20      30      40
orf3      MTIPFLD-----AGAGYRELRAEIDAALQVVSASGRYLLDAELAAFEFEFAA
           .:. :      .:.:      .:.:      .:.:      .:.:      .:.:
tylB      MTGLPRPAVRVPFHDLRDVHAATG---VESEIGGALLRVAARGRYLLGAELAAFEERFAE
           10      20      30      40      50

           50      60      70      80      90      100
orf3      YCDNDHCVAVGSGCDALELSLRALDIGPGDEVVVEAHTFIGTWLAVSATGARPVAV----
           .:. :      .:.:      .:.:      .:.:      .:.:      .:.:
tylB      YCGNAHCVAVGSLDDARLALWALGVGEGDEVIVPSHTFIASWLAVSATGATPVPVPEPGD
           60      70      80      90      100      110

           110     120     130     140     150     160
orf3      --DPTPDGLSLDPALVEAALTPTPTRALMPVHLHGH PADLDPILLAIAGRHLAVVEDAAQA
           .:. :      .:.:      .:.:      .:.:      .:.:      .:.:
tylB      PGEPCGCAFLLDPRLEAALTPTPTRAVMPVHLYGHFVDLDPVGAFAEPHGLAVVEDAAQA
           120     130     140     150     160     170

           170     180     190     200     210     220
orf3      HGARYRGRRIGSGHVVAFSFYPGKNLGAMGDGGAVVTGDSGVAERIRLLRNCGSREKYRH
           .:.:      .:.:      .:.:      .:.:      .:.:      .:.:
tylB      T-ARYRGRRIGSGHRTAFSFYPGKNLGALGDGGAVVTSDEPELADRLRLRNYGAREKYRH
           180     190     200     210     220     230

           230     240     250     260     270     280
orf3      EVRSTHSRLDEFQAAVLRAKLPRLDAWNARRAGTAERYGRALGPVPQIAVPVTAPWADPV
           .:.:      .:.:      .:.:      .:.:      .:.:      .:.:
tylB      EERGTSRLDELQAAVLSVKLPYLDAWNTRRREIAARYGEALAGLPGVTVPEGRVAEPV
           240     250     260     270     280     290

           290     300     310     320     330     340
orf3      WHLYVIRCAERDELRRRLERAGVQTLIHYPVPPHRS PAYADDPAGA-PAGTHPLSERLAA
           .:.:      .:.:      .:.:      .:.:      .:.:      .:.:
tylB      WHQYVLRSPYRDRLRRRLAEAGVETLVHYPVAVHASGAYAG--AGPCPAGGLPRAERLAG
           300     310     320     330     340     350

           350     360     370     380
orf3      QSLSLPLGPHLGEDEARAVVAAVRAASAGLAAYPTPDGQRFPLVTEKR
           .:.:      .:.:      .:.:      .:.:
tylB      EVLSLPIGPHLPDEAVEVIAAVQSAALDSWEEGP
           360     370     380

```

Comparison of:
(A) SRMD=Query

(B) MDMA=Subject

>gi|320536|pir||A60725 midecamycin-resistance protein mdma - Streptomyces mycarofaciens
length = 271

Score = 380 bits (975), Expect = e-104
Identities = 211/269 (78%), Positives = 232/269 (85%)

(A) Query: 7 LPPQSVSAPADSRSTARREWGQNFRTAAACRFSAQIDGSDTTPPDSPNDLMTVEIGAG 66
+ P S SAPA SRSTARRE GQNFRTAAACRFSAQIDGSDTTPPDSPNDLMTVEIGAG
(B) Subject: 1 MSPISASAPASRSTARRELGQNFRTAAACRFSDQDAFCADLPGLADVLTVETIGAG 60

(A) Query: 67 SGRVTKVLASPGTPLLAVEIDRRMARLAEISLPDVTVNEDFLTQLPGQPVRLIGNLP 126
SGRVTK IAS G LLAVEID MARL AEISLPDVTVNEDFL LQLP QP+RLIGNLP
(B) Subject: 61 SGRVTKALASAGSLLAVEIDAYMARLAEISLPDVTVNEDFLNLQLPQPIKILIGNLP 120

(A) Query: 127 FVTGTRMLRCLDMGPAPMRQGVFLQREYVKGRTGAWGNLFNNAQWEPWYSFDRGLAFS 186
FV+GT++LRCL++GP RM Q VFLQREYVKGRTGAWGNLFNNAQWEPWY+F+ GLAFS
(B) Subject: 121 FVSGTKILRCLDELGNRMCAVFLQREYVKGRTGAWGNLFNNAQWEPWYTFEGGLAFS 180

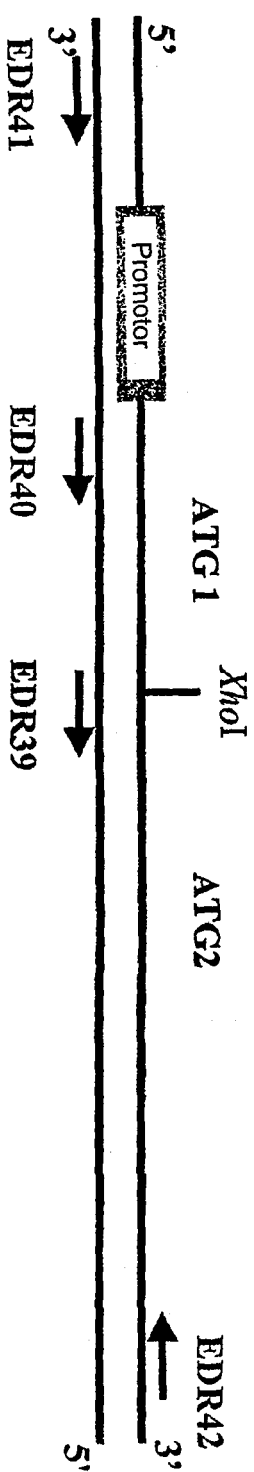
(A) Query: 187 RQDFTPPVPRA DTQITLMA PHRPSPVPMREKRAYQRFVQKVPD TGQMTVGDAARKYLRRGH 246
R +F+PVPRADTQITL+V P RRPSPVPMRE+ YQRF Q++FDTGQMT+G+AARKYLRRGH
(B) Subject: 181 RNEFSVPVRADTQITLVMPRRRPSVPWMRETDYQRFVQKLPD TGQMTIGEARKYLRRGH 240

(A) Query: 247 AQFVRGAGVRPADRVKDLTVPEWTALFRA 275
AQFVR AGVRPADRVKDLTV +W ALFRA
(B) Subject: 241 AQFVRSAGVRPADRVKDLTVDMRALFRA 269

FIG. 26

Fase 1 (at11) :	atcgcgcgccttcggttcgggacgaaagagtagat
Fase 2 (at22) :	atcgcgcgcgccttcggttcgggacgaaagagtagat
Fase 3 (at33) :	atcgcgcgcgccttcggttcgggacgaaagagtagat

FIG. 27



ATG 1: ATG situado mais em 5'
 ATG 2: ATG mais em 5'

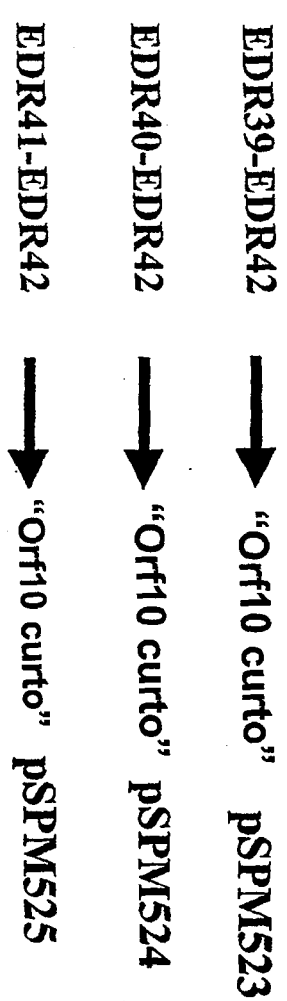


FIG. 28

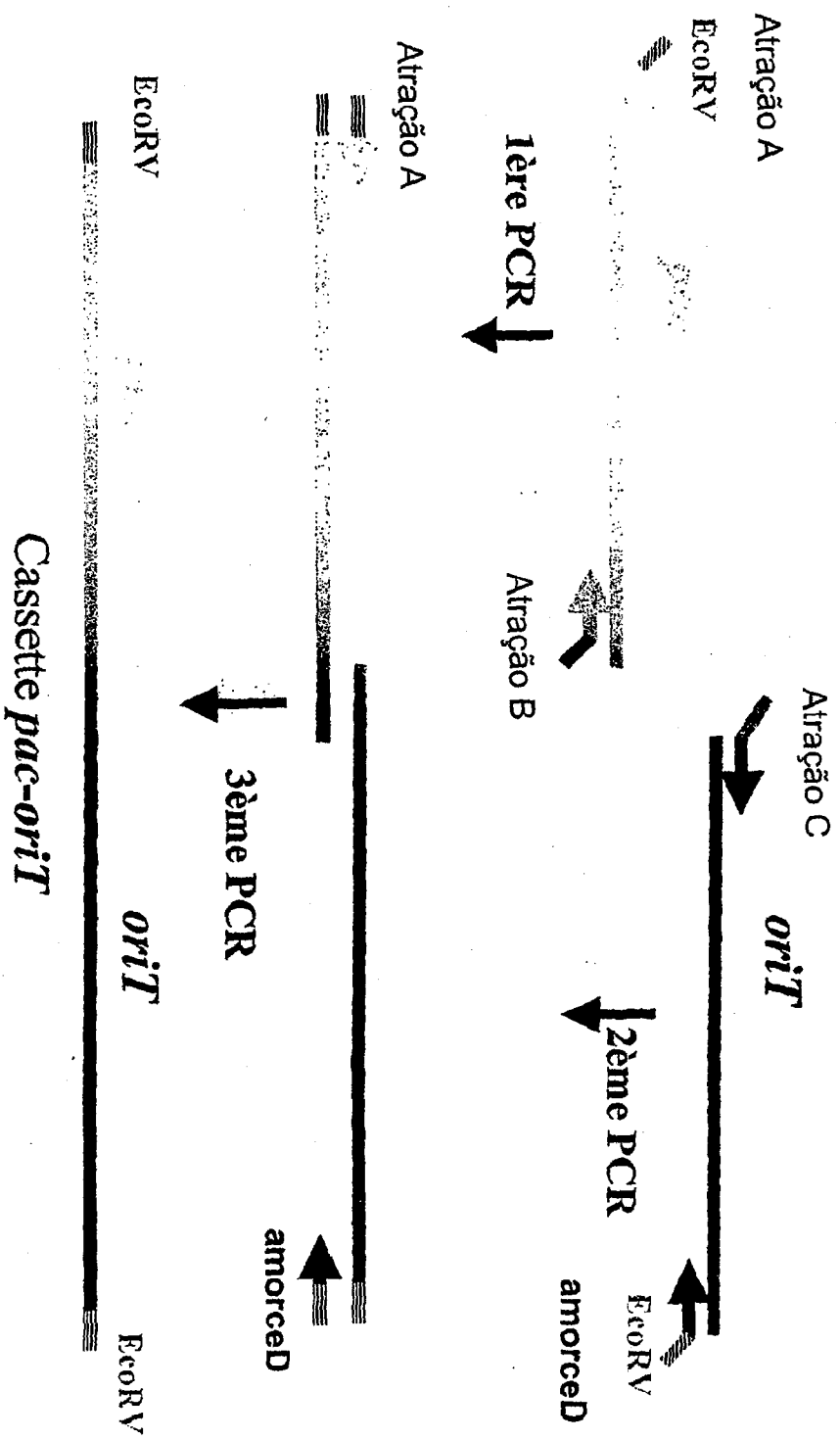


FIG. 29

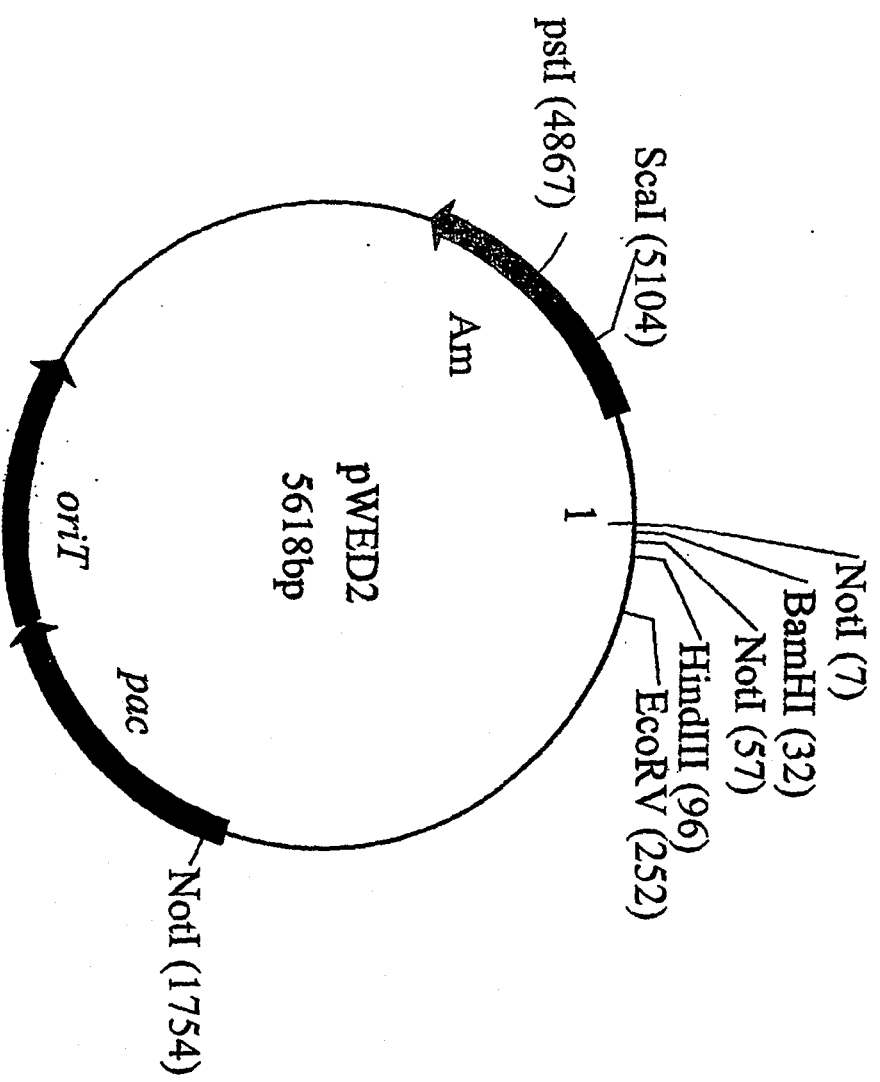


FIG. 30

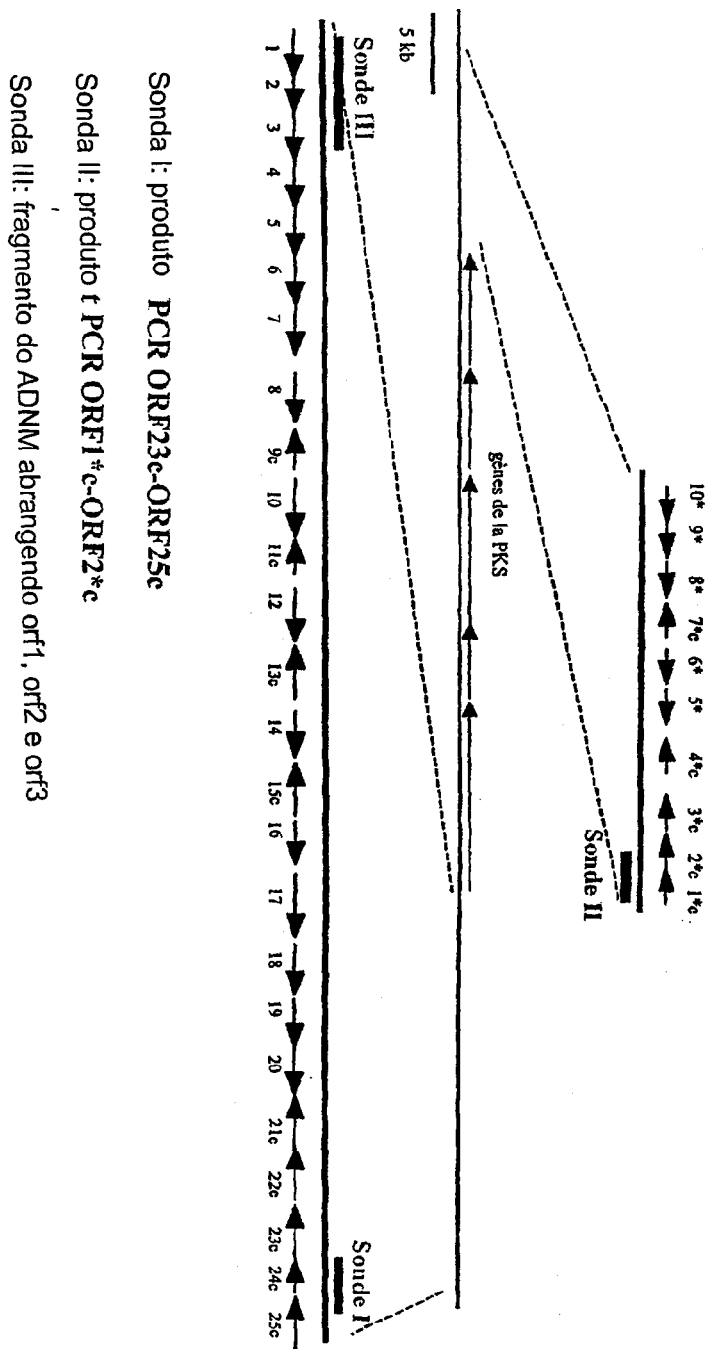


FIG. 31

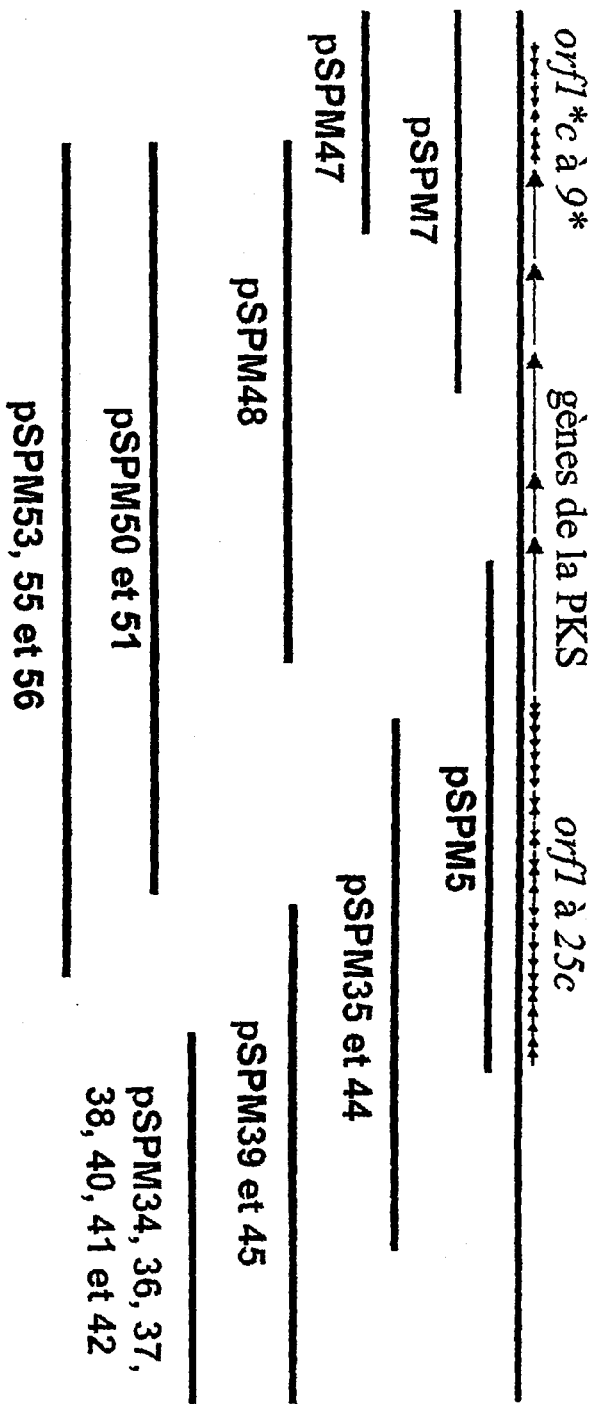
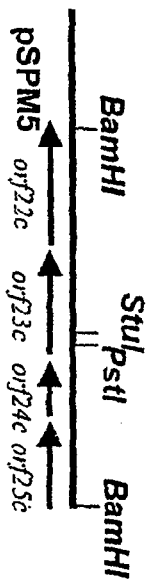


FIG. 32

Cromossoma

Sonde I



500pb

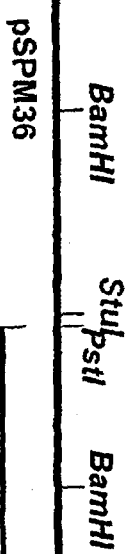


FIG. 33

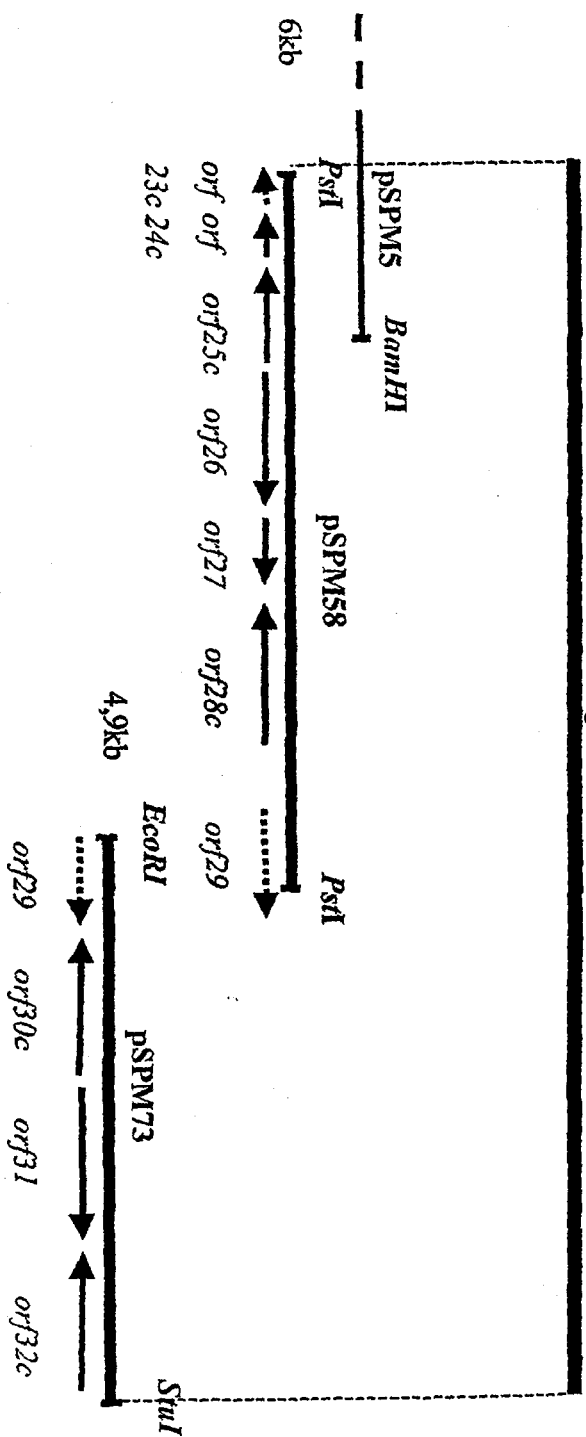


FIG. 34

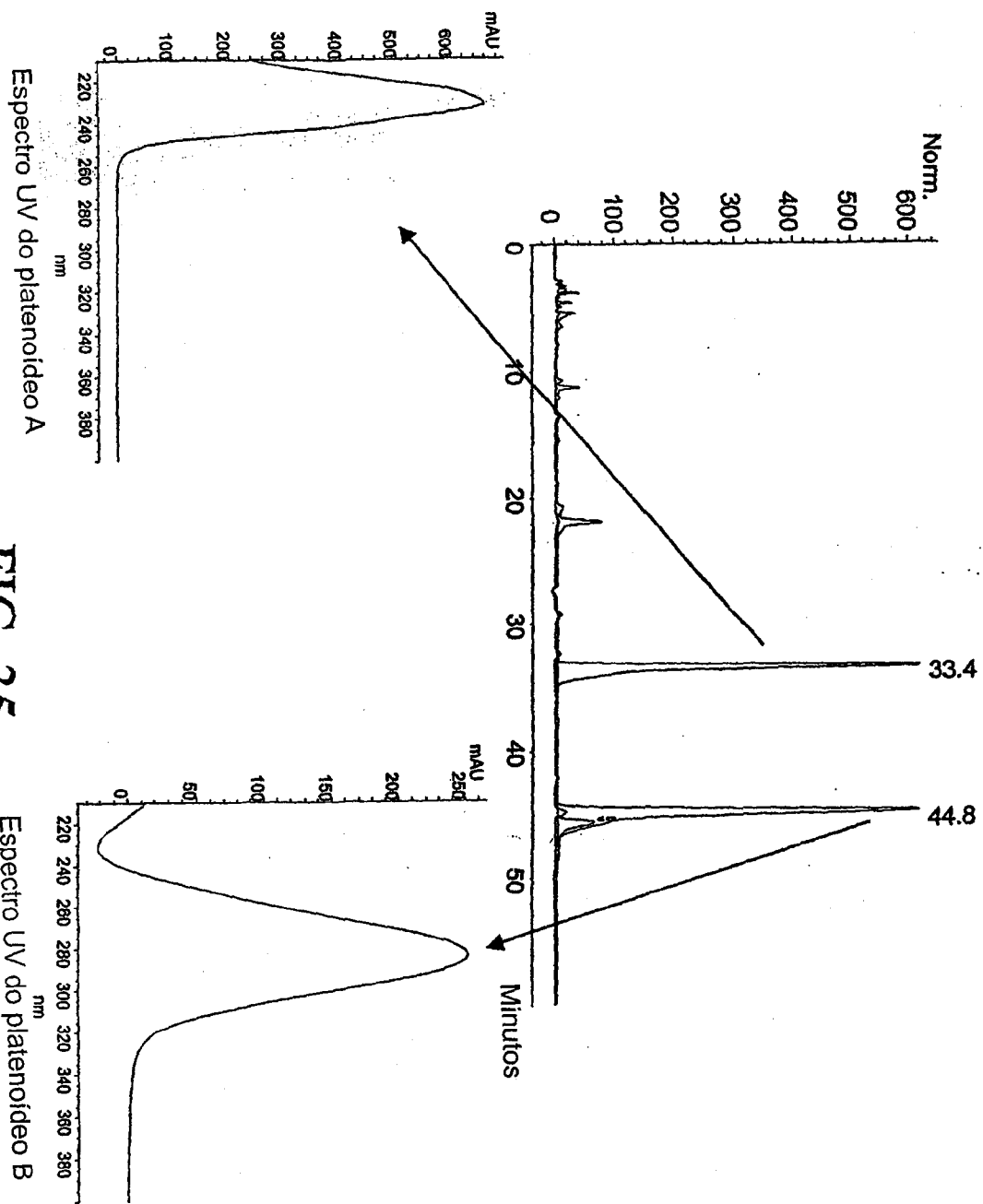
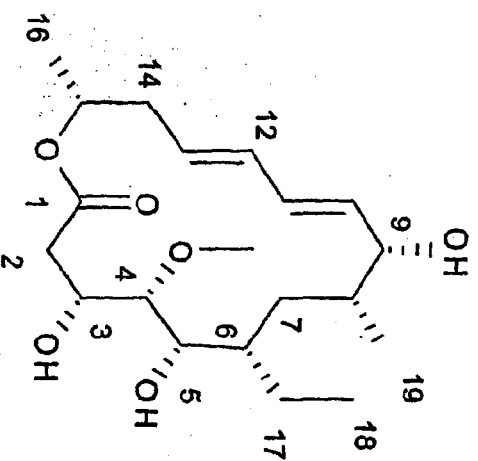


FIG. 35

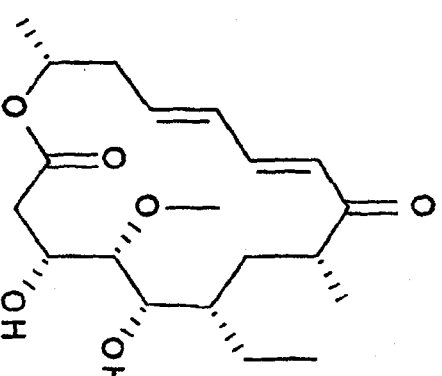
Platenolideo A MM=370



Massa molecular =370

Fórmula molecular=C₂₀H₃₄O₆

Platenolideo B MM=368



Massa molecular =368

Fórmula molecular=C₂₀H₃₂O₆

FIG. 36

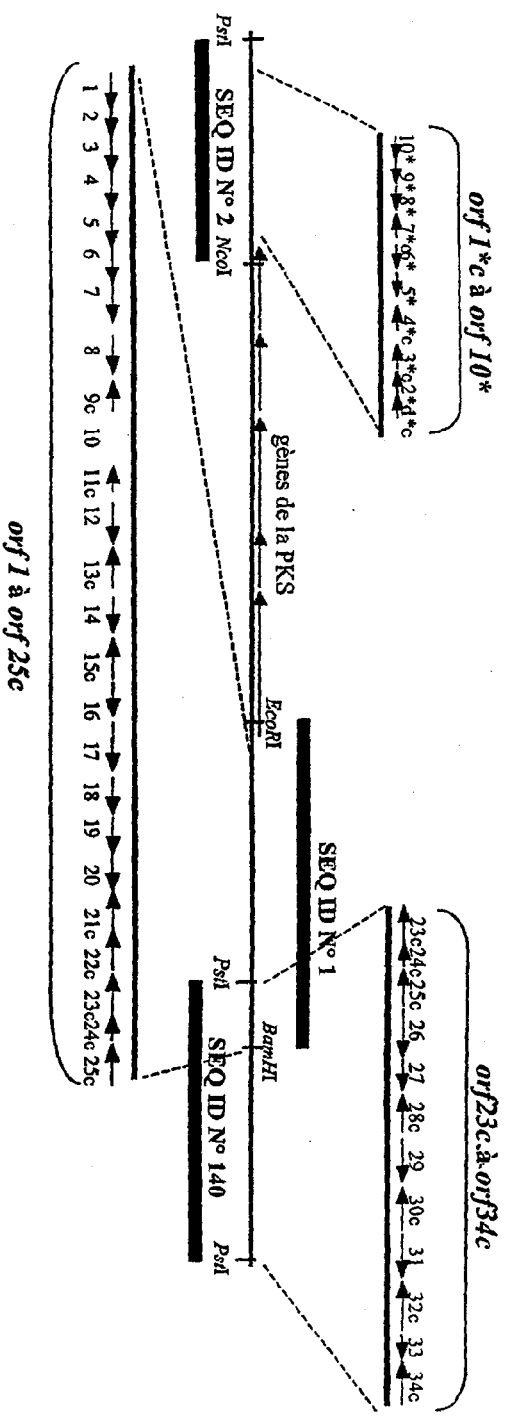


FIG. 37

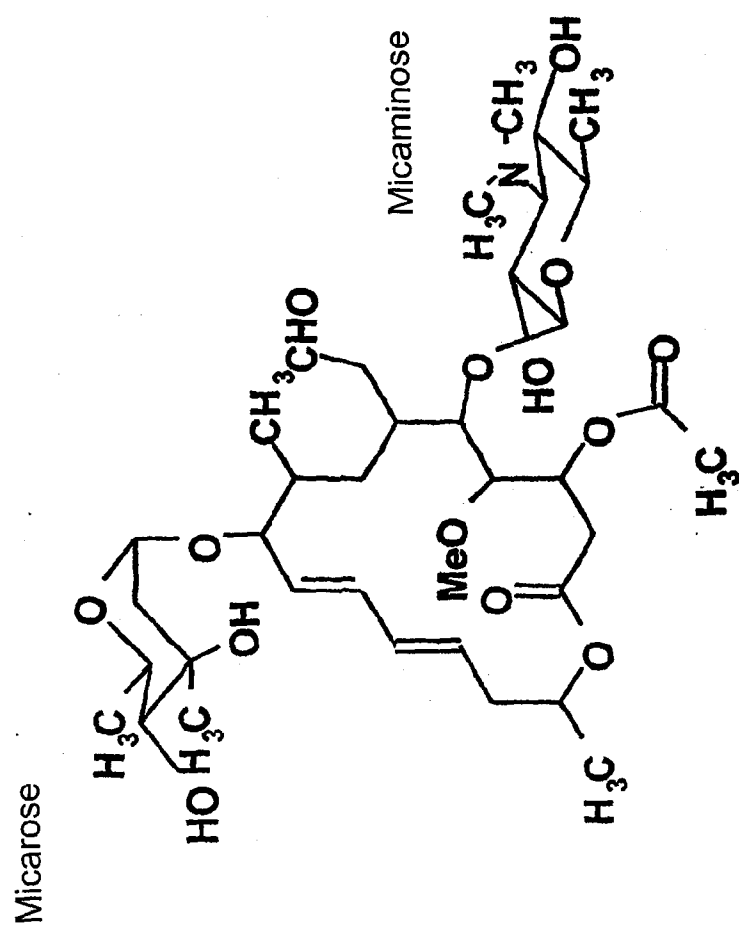


FIG. 38

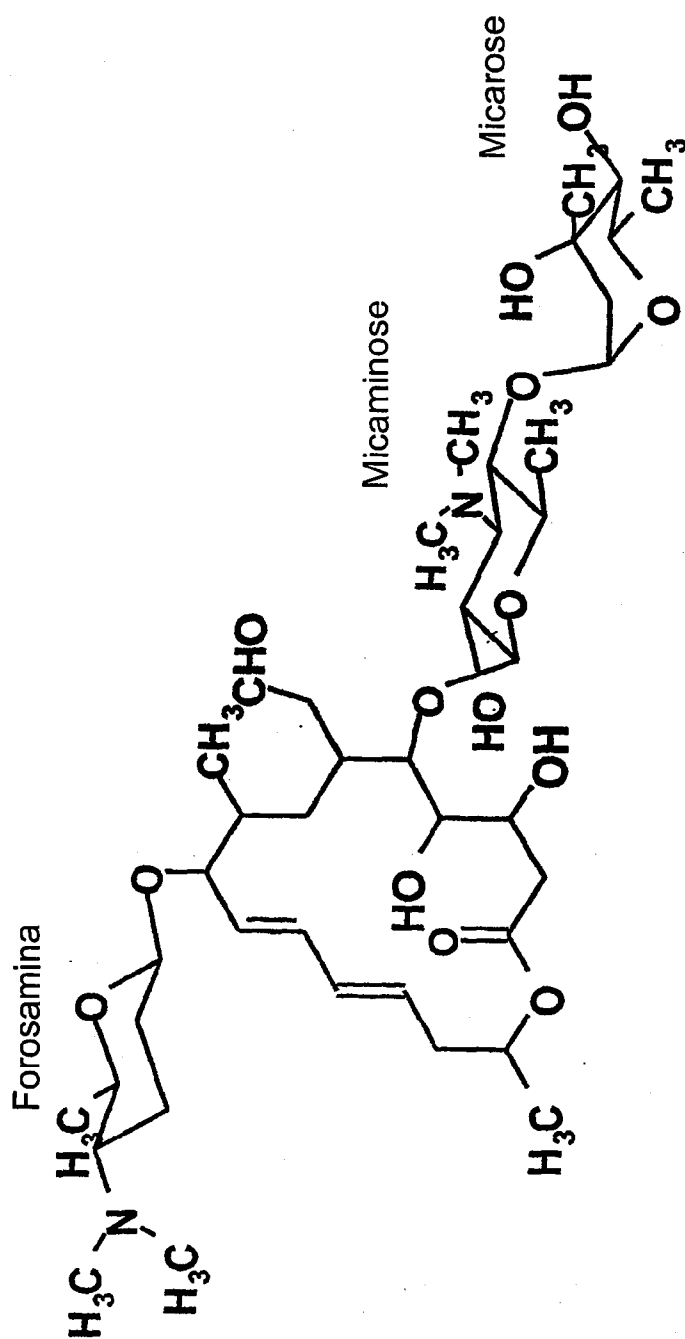
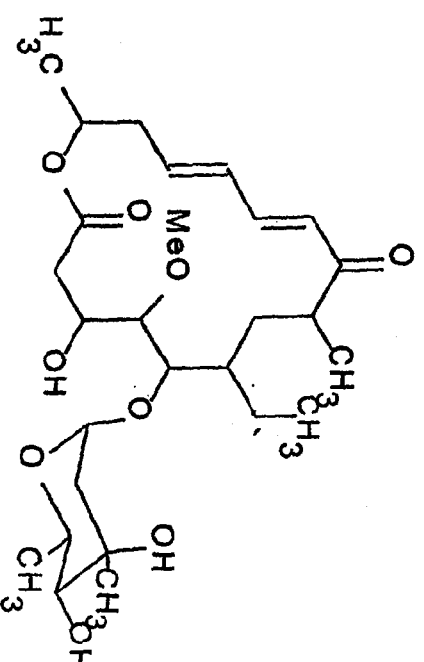
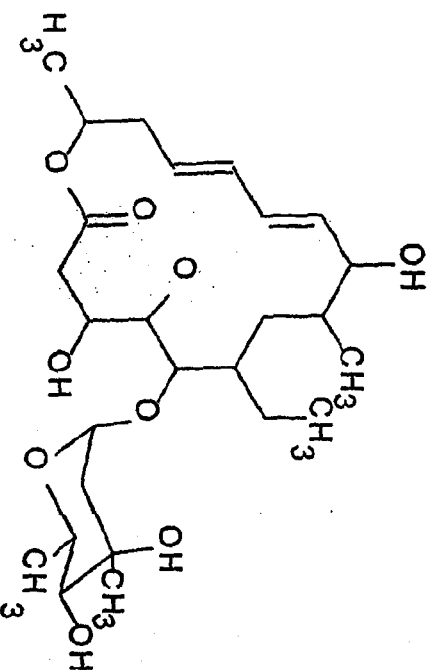


FIG. 39



Platenolide A + micarose

Platenolide B + micarose

FIG. 40

Cromossoma

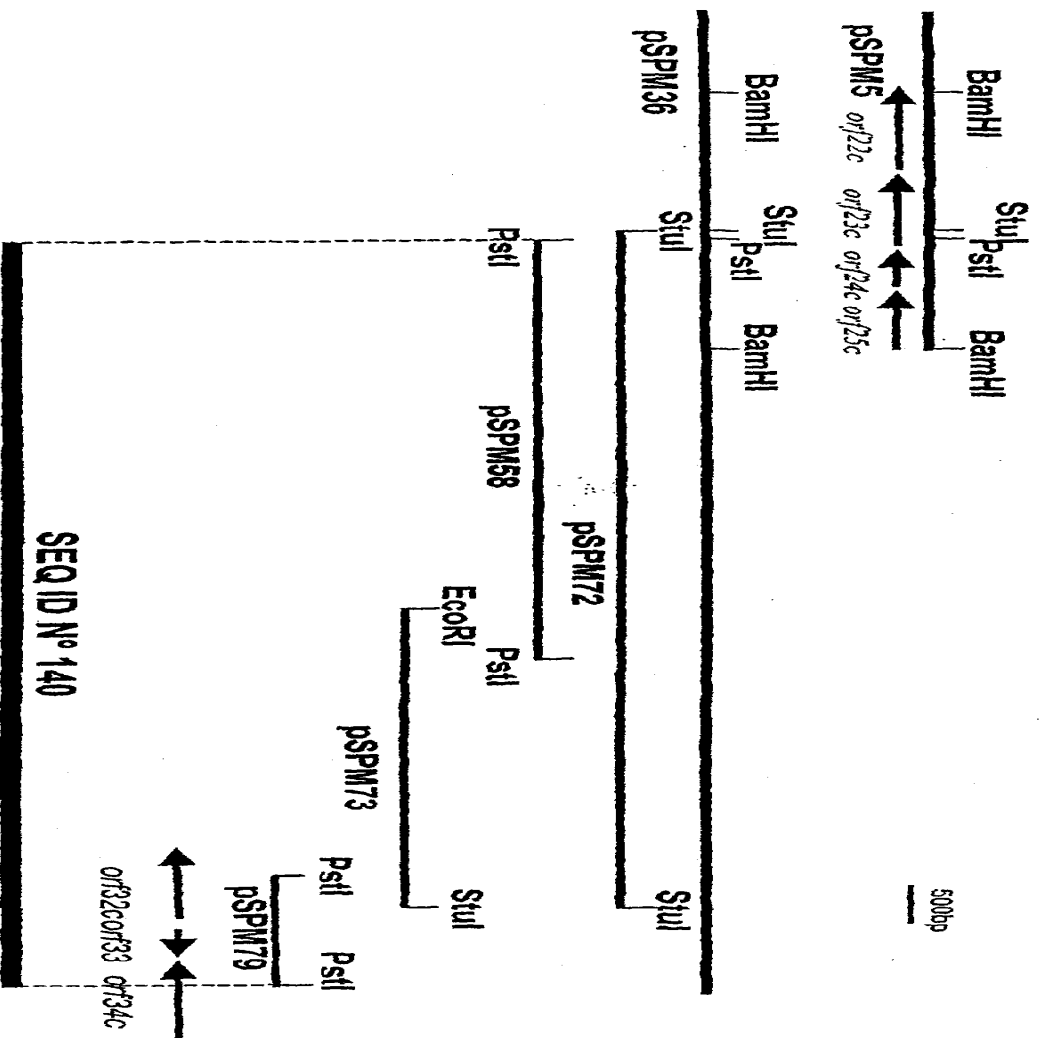


FIG. 41

RESUMO

Patente de Invenção: **"CÉLULAS HOSPEDEIRAS, DNA RECOMBINANTE, UTILIZAÇÃO DOS MESMOS, PROCESSOS DE PRODUÇÃO DE UM POLIPEPTÍDEO IMPLICADO NA BIOSÍNTESE DAS ESPIRAMICINAS, DE PRODUÇÃO DE MACROLÍDEOS E DE PRODUÇÃO DE ESPIRAMICINA(S), E MICRO-ORGANISMO TRANSGÊNICO MUTANTE PRODUTOR DE MACROLÍDEO".**

A presente invenção refere-se ao isolamento e à identificação de novos genes da via de biossíntese das espiramicinas e de novos polipeptídeos implicados nessa biossíntese. A invenção é também relativa a um processo de produção desses polipeptídeos. Ela é também relativa à utilização desses genes com a finalidade de aumentar as taxas de produção e a pureza da espiramicina produzida. A invenção refere-se notadamente a um micro-organismo que produz a espiramicina I, mas que não produz a espiramicina II e III e a utilização desse micro-organismo. A invenção refere-se também à utilização dos genes da via de biossíntese das espiramicinas para a construção de mutantes que podem levar à síntese de novos antibióticos ou a formas derivadas de espiramicinas. A invenção refere-se ainda às moléculas produzidas graças à expressão desses genes e das composições farmacologicamente ativas de uma molécula produzida graças à expressão desses genes.

LISTAGEM DE SEQUENCIAS

<110> Aventis Pharma S.A.
Centre National de la Recherche Scientifique (C.N.R.S.)

<120> CULAS HOSPEDEIRAS, DNA RECOMBINANTE, UTILIZAO DOS MESMOS,
PROCESSOS DE PRODUO DE UM POLIPEPTO IMPLI-CADO NA BIOSSETESE
DAS ESPIRAMICINAS, DE PRODUO DE MACROLEOS, DE PRODUO DE
ESPIRAMICINA(S) E MICRO-ORGANISMO TRANSGICO MUTANTE PRODUTOR DE
MACROLEO

<130> FRAV2002/0028

<140> PI0314543-3
<141> 08-10-2003

<150> FR 0212489
<151> 08-10-2002

<150> FR 0302439
<151> 27-02-2003

<160> 161

<170> PatentIn vers 3.2

<210> 1
<211> 30943
<212> DNA
<213> Streptomyces ambofaciens

<400> 1

gaattcggca attcccgggg cgccgggaaa ccggcacgcc attccgacca cggcaacggc	60
gtcggaccgg tcgtcactgc gggaaatcgc gagttctcca gacacctgca tccctcatat	120
gttcgccgta ccacgccgtg gcttttctgc cttttcttga tcttcccgag cgtacagcgg	180
gcgaactgcc gggcgacagg aacggccgggt gaacgacaat caattgtgtc aagttcggag	240
attgtccacc gattctcgtc accggagacg gcgcggggatc atgcggggacc acacggcatc	300
gcaccaggtc cgcagggtcg gcgctccgac gtcccggccg gtcgcgcact ccggtgacct	360
gcacgcggag tcctgggcga gcggagaaga tttcagtacc tcgcgggcgcg ggacaaccac	420
ttcgcgacga atatgtcagg tctccccgag cggtcctgtc cgccccgcc tgaccttc	480

gcgcccgctcc	cgctccgccc	cggtcgatcat	ctcgcccgga	atttcagtcg	gcagctcatt	540
gtgtcagggtt	cgccttgccg	acattctccg	gagattccta	agctctgccg	gtaaccggga	600
ccggaaccac	cgtgccgcgc	gttcggtcca	cacaccgctt	ttcgaggagt	ccgactgatg	660
ggtgaggccg	tgacgggacc	gatggagctg	agcaaggacg	cggacgcccg	ggggctgctt	720
gagtggttcg	cgtacaacag	gacgcgtcat	ccggtgttct	gggacgagac	ccgacaggcg	780
tggcaggctt	tcggctacga	cgactacgtg	acggtgtoga	acaaccgcga	gttcttctcc	840
tcggacttca	acatggtgat	gccgacgccg	cccgaactgg	agatgatcat	cggtccgggc	900
acgatcggcg	cgctggaccc	gcccgcgcac	ggaccgatgc	gcaagctggg	gagccaggcg	960
ttcaccccc	gacggatcgc	ccggctggag	cccaggggtgc	gcgcgatcac	cgaggagctc	1020
ctggacaagg	tggggcagca	ggacgtcgtc	gacgccgtgg	gtgacctgtc	ctacgcgctg	1080
ccggtcatcg	tgatcgccga	actgctgggc	ataccgcgcg	gcgaccgtga	cctgttccgg	1140
gagtgggtcg	acaccctgct	gacgaacgag	ggcctggagt	acccgaacct	cccggacaac	1200
ttcaccgaga	cgatcgcgcc	cgcgctcaag	gagatgaccg	actacctcct	gaagcagatc	1260
cacgccaagc	gggacgcgcc	cgccgacgac	ctggtcagcg	ggctgggtcca	ggcggagcag	1320
gacggccgcc	ggctgaccga	cgtcgagatc	gtcaacatcg	tcgcgctgct	cctgacggcg	1380
gggcacgtct	cctccagcac	cctgctcagc	aacctgttcc	tggctcctgga	ggagaacccg	1440
caggcgctgg	aggacctgcy	ggccgatcgc	tccttggtgc	ccggcgcgat	cgaggagacg	1500
ctgcgctacc	gcagcccctt	caacaacatc	ttccggttcg	tcaaggagga	caccaccgtc	1560
ctcggtcgcg	tcattggagaa	gggccagatg	gtgatcgcc	ggagccagtc	cgccaaccgg	1620
gacccccggc	acttcccga	cccggacacc	ttcgacatcc	gccgctcgga	cggcaccggg	1680
cacatggcct	tcgggcacgg	catccaccac	tgccctgggtg	ccgccctcgc	ccgcctggag	1740
ggcaaggcca	tgctcgaact	cctcctggac	cggttccaag	gcttccgcat	cgaccacgag	1800
cacaccgtgt	tctacgaggc	cgaccagctc	actccgaagt	acctgcccg	ccgggtcgac	1860

tggaactgaa cccgaggggtc tcgtcccggga gtccaggggcc gtcccagagcc ggccctggac	1920
ctcacgaccg cccgataagg agcgccgcca tcgccagagaa cacagccgag ctccctgccc	1980
ggcggggtcgg caggatcaag ccgtgccggc tgatcaggct cgagcagcac atcgacccgc	2040
gcggcagcct ctccgtgata gaggccggcg tgaccgtgga cttccccgtc cgacgcgtct	2100
actacatgca tggccagacc cagtcctctc ccccgcgagg cctgcacgcg caccgcaccc	2160
tggaacaact cgtcatcgcc gtccacggcg cttctccat caccctcgac gacggcttcc	2220
agcacgccac ctaccgtctg gacgaaccgg gagccggact ctgcatcggc cccatggtct	2280
ggcgcgtcct gaaggacttc gaccccgaca ccgtggccct ggtcctcgcc tcgcagcact	2340
acgaggagtc cgactactac cgcgactacg acaccttctt gcatgacgca cggagcctca	2400
catgaccatc cccttcctcg acgcggggcg cggctaccgg gaggttgcgag ccgagatcga	2460
cgcggccctg cagcgggtgt ccgcctccgg ccgctatctg ctcgacgcgg aactcgcggc	2520
cttcgaggag gaggtcggcg cgtactgcga caacgaccac tgtgtggcgg tgggcagtgg	2580
ctgcgacgcg ctggagctgt ccctgcgggc gctggacatc ggtcccgggg acgagggtgt	2640
ggtgcccgcg cacaccttca tcgggacctg gctggccgtg tccgctaccg gggcacggcc	2700
ggtggccgtc gacccgacgc cggacgggct ctccctcgac cgggcgctgg tggaggcggc	2760
gctcaccctt cggaccagag ccctgatgcc ggtgcacctg cacgggcacc cggccgacct	2820
cgacccgcta ctggcgatcg ccggacggca cggcctggcc gtggtcgagg acgccgcgca	2880
ggcccacggc gcccggtacc ggggcccag gatcggtctg ggccacgtgg tcgcgttcag	2940
cttctacccc ggcaagaacc tcggcgccat gggggacggc ggcgcggtgg tcacgggtga	3000
ctccggtgtg gccgagcggg tccggttctt gcgcaactgc ggctcgcggg agaagtaccg	3060
gcacgagggt cgctcgacct actcccggct cgacgagttc caggcgggcg tgctgcgggc	3120
caaactgccg cggctcgacg cgtggaacgc ccgcccgggc ggcacggccg aacggtacgg	3180
gcggggccctg ggtccggtac cgcagatcgc cgtcccggtc accgctccct gggccgacct	3240
ggtgtggcac ctgtacgtga tccgctgcgc ggagcgcgac gagctgcgcc gccggctgga	3300

acgagccggg	gtccagaccc	tgatccacta	ccccgtgcc	cgcacccggt	ccccggccta	3360
cgccgacgac	ccggccggcg	caccggcggg	gacccacccg	ctcagtgagc	gcctggcggc	3420
gcagagcctc	agccttcccc	tgggaccgca	cctcggggag	gacgaggccc	gcgccgtcgt	3480
ggcggcggtc	cgggcggcgt	ccgcagggct	ggcggcgtag	ccgacgccgg	acggccagcg	3540
ttttcctcta	gtgacggaga	aacgatgacc	gagggtcatgt	cagggcgctc	cggaatgaaa	3600
gggatcatcc	tcgcaggcgg	cggagggacc	cgcctaagcc	ccttgaccgg	cacgctgtcc	3660
aagcaactgc	tgccccgtcta	cgacaagccg	atgatctact	acccgctgtc	cgtcctgatg	3720
ctgggcggca	tccgcgagat	cctcgtcgtc	tcctccaccc	agcacatcga	gctgttccag	3780
cggctgctgg	gcgacggctc	ccgcctcggc	ctcgacatca	cctacgccga	acaggccgag	3840
cccgagggca	tagcgcaggc	catcaccatc	ggcaccgacc	acatcggcga	ctcaccggctc	3900
gcgctcatcc	tgggcgacaa	catcttccac	ggccccggct	tctcggccgt	gctccagggc	3960
agcatccgcc	acctcgacgg	ctgtgtgctg	ttcggctacc	cggtcagcga	cccgaagcgc	4020
tacggcgtcg	gcgagatcga	cgaccagggc	gtactgctgt	ccctggagga	gaaaccggcc	4080
cggccccgct	ccaacctcgc	cgtcaccggc	ctctacctct	acgacaacga	cgtggtcgac	4140
atcgccaaga	acatccggcc	ctcggcgcgc	ggcgaactcg	agatcacgga	cgtcaacagg	4200
acctacctgg	agcagaaacg	cgccccggctc	atcgaactgg	gccacggctt	cgcctggctc	4260
gacatgggca	cccacgactc	cctcctccag	ggcggccagt	acgtccagct	catcgagcag	4320
cgccagggag	tgcggatcgc	ctgcatcgag	gagatcgccc	tgcgcatggg	cttcatcgac	4380
gccgacaccc	tccaccggct	cggccgcgaa	ctgggcacct	cgggatacgg	cgcgtacctg	4440
atggaggtgg	ccacccgtgc	aggcaccgaa	tgagacgcgc	cgcgggcccg	cccgtccgc	4500
cggccgacgg	ccgccggccc	ggatcctcgt	caccgggggc	gccggcttca	tcggctcgcg	4560
cttcgtgaac	gcgctgctgg	acggctccct	gccggagttc	ggcaaaccgc	aggtgagggc	4620
gctcgacgcg	ctcacctacg	cgggcaacct	ggccaatctg	gccccgggtg	gcgactgtcc	4680

ccggctgcgg atcttcccgg gggacatccg cgaccgcggc gcggtcacc	aggcgatggc	4740
gggggtcgac ctggtggtgc acttcgcggc cgagtcgcac gtggaccgct	cgatcgacga	4800
cgccgacgcc ttcgtgcgca ccaacgtgct gggcaccag gtcctcctcc	aggaggcact	4860
ggccgtacgc cccgggctgt tcgtgcacgt ctgcacggac gaggtgtacg	gctccatcga	4920
ggaggggtcc tggcccagag agcaccgcgt gaaccccaac tcgccctacg	ccgcctcgaa	4980
ggcgtcctcc gacctgctgg cgctggccca ccaccgcacg cacggactgc	cggtgtgcgt	5040
caccgcgtgc tccaacaact acgggcccta ccagtaccgc gagaagatca	tcccgctgtt	5100
caccagcagc ctctcgacg gcgggaccgt cccgctctac ggggacggcg	gcaaccggcg	5160
cgactggctg cacgtggacg accactgccg gggcatcgcc ctggtggccc	ggggcgggcg	5220
gcccggcgag gtctacaaca tcggcgggcg caccgagctg agcaacgtcg	agctcacgga	5280
gcgtctgctg aaactgtgcg gagccgactg gtcggcggtg cggcggggtg	ccgaccgcaa	5340
gggccacgac cggcgctact ccgtcgacta caccaagatc gcggacgagc	tgggttacgc	5400
gccgcggatc accatcgacg aagggtgga gcggaccgtg cactggtacc	gggagaaccg	5460
cgcggtggtg gcgcccgcga agagggggcg atgacggtga cgaccgcac	cgtggaccgc	5520
ctcgacctgt ggctccgccg gtaccagccg tccgcgtcac ccgcgctccg	gctggtgtgc	5580
ttcccgcacg cgggcggctc ggcgagtctg ttctgccgt tcaccggga	gctgccggac	5640
cggatcgagg tcgtggccgt ccagtacccc gggcgccagg accgcaggag	cgaaccgctg	5700
gtcgacacca tcgagggact ggccgagccc ctggccggcc tgctggaggc	gcaggccggc	5760
cccccggtgg tgctgttcgg gcacagcatg ggcgcgctgg tggcctacga	ggtcgcccgc	5820
gcgctccagc ggcggggagc ggctccggtg cgcctggtgg tctccgggcg	ccgggcccc	5880
gccgtcgacc ggccgatgac cgtgcacctc tacgacgacg accggctggt	cgaggaaactc	5940
cgcaagctcg acggcaccga cagccaggtg ttgcgcgac cggagctgct	ccggctggtg	6000
ctgcccgtga tccgcaacga ctaccgggcc gtggcggcct acgccaccg	cccgggggcg	6060
ccgctggact gccccctcac cgtgttcacc ggcgcgcgacg accccaccgt	gaccgcggcc	6120

gagggcggcgg cctggcacga ggcggcggcg tccgacgtcg agacgcgcac cttccccggt	6180
ggccacttct tcccgtagca gcggaccgcg gaggtgtgcg gggccctggt cgacacgctc	6240
gagccgctgc tgtcggcccg gacgcgcggt gtccggcggg tccgcccggg gtgacgtggg	6300
cacggtcgag tacgccgtcc accggcgtag cgcggaacgg gtgaggggtct ccgccgacac	6360
cctggacagc ccggtcaccg cgctggcggg ggtgccccgc tggctggagg aataccaccg	6420
ggcgcaccgc ttccacgtcg agccgatccc ctccgacggg ctccggcgggt ggtccttcga	6480
gccgggcacc ggcgacctgc ggcacgagac gggccgcttc ttctccgtgg aggggctgcg	6540
caccagctcg gacgccgatc cggtcgcccg tgtccagccg atcatcgtgc agcccaggt	6600
ggggctgctc ggcatcctgg cccgggagtt cgacgggggtg ctgcacttcc tgatgcaggc	6660
caaaccgag cccggcaacg tcaacgggct gcagatctcc cccacgggtgc aggccacgcg	6720
cagcaacttc gacgaggtgc accacggccg gtccaccccg ttctctgacc acttcatcca	6780
ccgccccggc cgccgggtcc tgatcgacag catccagtcc gaacagggcg actggttctt	6840
gcacaagcgc aaccgcaaca tggtcgtcga gatcgacacc gacatcgagg ccgacgccgc	6900
gttccgctgg ctgaccctcg ggcagatccg ccggctgatg ctccaggacg acctcgtcaa	6960
catggacacc cgcagtgtgc tggcctgtct gcccaccgcg cacggcacgc ccgacgacgg	7020
tgacgactcc ttcccggcg cgctgcgccg ctccctctac ggggagaccg cgccgttgca	7080
cgatctgcac gccatcacca gctgcctcac cgacgtccgg gcgctgcggg tgctgcgcca	7140
gcagagcgtg ccgctcgacg acgcccggcg ggacggctgg gagcggaccg ggagcgcgat	7200
ccggcatcgc agcggcaggc atttcgagat catggcgggtg gaggtgaccg cggagcgccg	7260
tgaagtggcc tcgtggaccc agccgttgct gcgcccgtgc tcgcaggga tggcggccct	7320
gatcaccg cgcatcaacg ggggtctgca cgccctgggtg gggcgcggt cggaggtcgg	7380
cacgtcaac gtcgccgagt tcggaccgac cgtccagtgc cggcccgcg aggcggacgg	7440
ccagtcgccc ccgtacctgg accgggtgct gacggccgga gccgaccgcg tccgctacga	7500

cgtggtgcag	tcggaggagg	gcgggcgctt	ctaccacgcg	cgcaaccgct	atctggtggt	7560
cgaggcgggg	ccggagctcg	acacgggctg	cccgcccggc	ttctgctggg	ctaccttcgg	7620
ccagctcacc	gaactgctcg	cgcacggcaa	ctatctcaac	gtcgaactcc	gcacctcat	7680
ggcgtgcgca	cacgcctcct	actgaatggt	cacgaaagct	gcaccgcgcg	ggagaatcgg	7740
cagcgcgcca	ccggccggcc	ggcacccgga	aggtaaagcg	ccgttctccc	gcacgcgcgc	7800
cctgcgggaa	acggcggaac	ggccggcccc	gaccgcgcgc	aattcccggc	gggacacggt	7860
gggagcccg	acgaggaacc	gctttccccg	ccttcggtgc	gcccggccgc	gggaccaccc	7920
ccgcctccc	gccgggccc	ggaatacgac	gggggcggcc	gaggacattc	ctttcccgcc	7980
tccgaaaaag	cgccccga	ggcccccg	atgccggcg	ggacggacgg	cgactgcgcg	8040
cggacggcg	cccggcgctc	aacgcacctg	cccgagtcgc	gacgagacag	cgcgacgcga	8100
gaggcgaaaa	tgatcaatct	cttcacgccc	cagatggggg	ccgaggaact	ggcgcggtg	8160
tccgaggtct	tcgacgacca	atggctcggt	cacggacccc	ggaccgcggc	gttcgagtc	8220
gcgttcgccc	agcacctcgg	ggtcggcccc	gagcacgtcg	tcttcctcaa	ctcgggcacc	8280
gccggcctct	tcctggccct	ggagtcgctc	ggcctgcggc	ccggcgacga	ggtcgtgctc	8340
ccctcgcccc	gcttcctcgc	cgcggcggaac	gccgtacagc	tctcgggagc	gcgcccgggtg	8400
ttctgcgaca	ccgaccgcgc	gacgctgaac	cccgccctgg	agcacatcga	ggcgggccgtc	8460
accccgcgca	ccagggccgt	catcgcgctc	cactacggcg	gccaccccg	cgacatcggtg	8520
cgcacgcgcg	agcgtgccc	ggagcggggc	atcacctga	tcgaggacgc	cgcgtgctcc	8580
gtggcctccc	gcgtcgacgg	ccgaccggtc	ggcaccttcg	gcgacctcgc	catgtggagc	8640
ttcgacgcca	tgaaggtcct	ggtcaccggc	gacggagggg	tgatctacgt	caaggacccc	8700
ggggcgggcc	cccggatccg	gcgcctcgcc	taccacggcc	tcacgcggtc	cagcggcctg	8760
ggatacgcca	gggtctcggc	gcgctgggtg	gagatggacg	tccccgaacc	gggcccgcgc	8820
gtcatcgggg	acgacctcac	cgcggccatc	ggcgcggtcc	agttgcgcgc	gcttcccggc	8880
ttcgtggccc	gccgcaggga	gatcgctcgc	ctgtacgaca	gcgaactgag	ctcgtggag	8940

ggcgtgctga caccgcccgc gccacccgcg gggcacgagt ccacgcacta cttctactgg	9000
atccagctgg cccccggcgt ccgggaccgc gtggcacgcg acctgctcac cgacggcatc	9060
tacaccacct tccgctacgc acctctgcac aagggtgccc cctacggcca caccggaggc	9120
gaactgcccg gcgtaggagc ggcgtccgaa cggaccctgt gcctgcccct gcaccccggc	9180
ctgtcggacg ccgacgtccg caccgtcgtg tcctccctgc gcagagccct gagcgccgcg	9240
gatccggccc ccgcctgacg cgcgcacgcg ccacggcccc tgtcccggcg gtgaccgccg	9300
ggacagggggc cgtggcgcgt cctgcggacg gtccggggcca ccccgccgtc ccgtcgggtg	9360
gtcagcgacg cgtgccgagg aagaagcccg gtgagccctc gcccgtcacg acgtactcga	9420
cgtccaggcc ggccttgccg tacgcggcct cgtactccgc acgcgagaac agcgtgaggt	9480
agtccacctc gctgcgggtg cgcatgccct ccgcggtgtg cgcgatcagg tagtggatct	9540
ccatccgggt ccgcccggcc tcccgggtcg agtgggagac acggggcgac ccctggtcgc	9600
ccggtgccgt ccgcagggcg tgggtggaga cgtggccgtc caggaagggt tcggggaagt	9660
accacggttc gacggccagg acaccgtccg cggtcagggt ccgcgccatg gcggcgacgg	9720
cgtcctccag gtcggccgtg gtctccagggt acccgatcga gctgaacatg cagaccacgg	9780
cgtcgaacgt ctgcgcgagg tcgaacgccc gcatgtcacc gcggtggagg gtgacgccgg	9840
ggagccgctc ctgcggcgcg gccgccatcc actcggacag ctccaggccg ctacgcgggt	9900
cgtagagctt ggcgaaggcc tccagggtggg taccggtgcc gcacgcgatg tcgagcaggc	9960
tggccgcgtc cggcgtacgt tcccggatca ggtcgggtgac ccgggaggcc tcacccgcgt	10020
agtccttgcg gtccctggtag agcaggtcgt agacctcggc ggcaactgtc ttctcgtaca	10080
tgggaatcct ccggggccgtg aagggggaac cgtgggtctg tcgaagagga gtgcgccgtg	10140
cgctcggggc gggggcgggc tgccgcgggc ccctcgcgat gatctccgta cggacaccca	10200
acagttcacg ccgagccggg gtcaaggaaac gacgggggtg tcagtcaagt cgggcgctcc	10260
gccccggggg cggggcgcgc cggcgacgcg cacggattcg gccaacgggt tgctcgttcc	10320

gccggaatc acggtgtggc ccccgggcca ccgggtagct tatgcctcgt tcaccgcagc 10380
 ggttgaagag gcagccttca accccggccc ggcctttatg gaattcattt ccaccgtgcc 10440
 gcaacaccca ctgaaggacg gccggatata ggccatgaag ccccggcctt tcagccaggc 10500
 accctctctt gtcgaataga gtatgtcctc cgctgaagcc gccgaagacg gacgaagggg 10560
 acgaacggtc acctcggtcg atctagacgg aatccttgaa agcgtaatag cctgtcaatg 10620
 ctttggtaaa gcacagggat gggggtgcct gcgggatgag tgacctgggt tctggtgaag 10680
 aaggggccga agaagacgag tcggacgacg cactcgcctt cctcgagttc atcgcccggc 10740
 cggcaccacg gagcgaatac gaccggctca tggcccgccg cgaacgctcg ggcgccgacg 10800
 aggaccggat gcgccgactg gagcgcttca accggctcgc cctcaccgcg cagtcgatga 10860
 tcgagtaccg ccgcgaccgg gaggcggagc tcgcggccct ggtcgacgcc gcgcacgagt 10920
 tcgtcgccgc ccggcggggc aaggacctgc tggagtccat cggccgcaga gcacggctgc 10980
 tgctgaagct ggacgtctcc tacgtcggcc tgacagagga ggaccggccc ggcacggtgg 11040
 tgctgagcgc cgacggcaac gcggtcaagg tcgccgagag ctaccggctg ccggccgacg 11100
 gcggactggg cgccatggtg cgcacctgcc gcgctccctt ctggaccccg gactacctcg 11160
 gggacaacag cttcacgcac gtcgaggccg tcgacgacat cgtccgcgcc gaaggcctgc 11220
 gcgcggctct ggccgtcccg ctgtgcgccg ggggcgaacc gatgggggtc ctctacgtcg 11280
 ccgaccgtca ggtgcggcat ctgaccccca acgaggtcac cctgctgtgc tcgctcgccg 11340
 atctggccgc ggtggcgatc gagcgcaacc ggctggctga ggagctccac gacaccatcg 11400
 ggcaactgcg ccaggacatc ggcgaggccc gcaccgccct cgcgcgcacc cgcaggtccg 11460
 ccgacctcca gtcgcacctg gtcacgcagg tgatggacag gcgcggcgcc gactcgttac 11520
 tcgcgacggc cgccgaggcg ctgcggcgcg gagccggcct gtgcagcccg ctgcggcgcc 11580
 cgctcgccga gtacgggacc ctgcgccccg tcgccccac ggaactgcgc gcggcggtgc 11640
 gccgggccgc cgagaccggc cggcccacct ccgtggcccc gggggctctg acggtgcccc 11700
 tgcttcccgg gggcaacgcc ggcttcctgc tgaccgacct cggtcggac gcggaccaca 11760

ccgccgtccc cctgctcccc atggtcgccc gcaccctcgc gctgcacctg cgcgtccagc 11820
 acgacgactc ccccaaggcg cagagccacc aggagttctt cgacgacctg atcggggagc 11880
 cccgctcacc cacgctcctc agggaacgcg ccctgatggt ctccctcagc ttccgccgcc 11940
 cgcacgtggt gctggtggcg gacggacccc gcgggacctc gccgcggctg gaggcctccg 12000
 gcgccgacta cgcgaaggag ctcggcgggc tgtgcagcgt gcgggacggc gccgtcgtcc 12060
 tgctgctgcc cggcgacgac cccgtcgccg tggcgagac cgccgccccg gagctgaccg 12120
 accgcgccgg gcaccccgtc accgtggggg tcgcggggcc cgccctcgacc gtcgacggca 12180
 tcgccgacgc gcaccgtgag gccgcgaagt gtctggagac cctccgcgcg ctcggcggcg 12240
 acggcggcac cgcgtgcgcc tccgacctgg gtttcctcgg catgctctc gccgaggaga 12300
 acgacgtccc cggttacatc aggacgacga tcggccccgt ggtcgactac gacaccacc 12360
 gcttcacgga tctggttccc actctgaggg tgtacctgga gtcgggcagg agccccacgc 12420
 gtgccgcaga gacactgcgc gtgcacccga acaccgtctc acggcggctg gagcgcatcg 12480
 gcgtactgct gggagaggac tggcagtcac cggagcgggt gctggacata caactggccc 12540
 tcgggtctta tcaggtgccg tcggcgctct cctcgcaacc ggcgtccgag acccgggccg 12600
 tgctcgatc gctgcgcgag tgacctctcc ggcacgggag gcgccccgcc cgtcggcgcg 12660
 gatcgcgccg acaggcgggc gttgaccacc ggccaccggt ccccgctacg acgaggtccc 12720
 gggcccgct ggttcagccg gcgctgaact cggcgatgcg cccgtctccc atggtcagcc 12780
 gggcgccggt gaaccgggac cgcattccgac ggtcgtgggt gacgaccag acggcgccgc 12840
 ggtagtccgc gagtgcctgc tccaactcct ccaccagcac cggggtgagg tggttggtgg 12900
 gctcgtccag cagcagcagg tccatcggt cgctcaccag cggggcgatc tcgatccggc 12960
 ggcgctgccc gtaggacaga tccttcacgc gtcgccgcag gtcggacggg ctgaacaggc 13020
 cgagcgacag cagtttctcc gcgtggtcct ccaggtagcc ctcccggccc tgggcgaagg 13080
 cccgcagcac ggtcagtccg ggcgccagg gcgtctcgtc ctgccgcaga tgaccgacc 13140

ggcagccgac ggcgaccgag ccgccgtccg gctccagttc cccggacagc acccgcaaca 13200
 aggtggactt gcccgcaccg ttgggacctg tgacgagcag ccgttcgccg ggccggatcg 13260
 tcagggagtc cacggcgagc cgacccgcga cgcgcacgtc ggtgagttcg gccaccgcct 13320
 cctccgcctc cgggcccgcg gtgtcgatgc gggcggcgaa ggacaacggg tcggcgggag 13380
 cgtggaccgg gttctcggtc agctgcgcca cgcgttgctt cgcgttgccg atccgcacca 13440
 tcgcgccgtg gtcgcgccct cgcctgcggt aggcgccgtg gccgaacacg gagagggaca 13500
 tcttgccggg gatgccgtcc atccgcgcca cgttgagggt gatcagcccg cggttgcggt 13560
 cgagttcggc acgccactcc tcgtactccc gcagccgccg ctgcgcttcc acggccttgg 13620
 ccgtcaggta gccctcgtag ccgttgccgt agcgggtgac gctgccggag tcgacctcca 13680
 ggatcgtggt ggtgagccgg tcgaggaaga cccggtcgtg ggtgaccgcg atcaccgtgc 13740
 cgcggtggcc ggccagggtg tcctccagcc attccatcgc ccggtcgtcg aggtcgttgg 13800
 tcggttcgtc caggagcagc agctccggcg acgaggcgag ggtcgcggcg agcgcgaggc 13860
 gggagcgttc gccaccggag agggttccga gcttgccggtc gcggtccagg ctccggcagtc 13920
 cgaggccgtg cagcgcgacc tccacgcgca cgtccggcctc gtagccgcca cgcgcctggt 13980
 actgctcgac cagagcggcg tagcgtgga ggccggcgga cagctcgcgc tcggagccgt 14040
 tctcgtcgtc ctgcgccagc tccgcctcgg cctcgcgcac cgcgccttcg agctcgcgca 14100
 ggtcggacag ggccagggtc acggcgtcct ggacgggtggc gtcgaggggc agttccagtg 14160
 tctgcgccag gtagccgacg ccgccgggag cgaccacggt gagcgcgccg ttgtcgggct 14220
 ccacgcggcc ggcgaggatc ttgagcagcg tggacttgcc ggaaccggtg tcgccgatca 14280
 cgccgacctt ctgcgccggc ttgatgctga aaccgaccgc gtcgagcacg acacagtcgt 14340
 ggtagcgctt cgtgatgtcg tgtagggcgt attgcgcaat cgacacgcgt aagtctcctg 14400
 tttccacgat gaggatgagt ggatgcgtga gcgcgctcgc agaaagaacg gaaagcagaa 14460
 gggacgccac cactgcggac atggccggtc aggggggtgtc acgagcacgc tgctggatgc 14520
 ggcaggcgga gtcagctcaa cgccgggcat cctcatctat cacagagatc ccatgcgaag 14580

aactatagcc gtgctacccg gtgccgcgca acggtatgcg tgtcgcaccg gccgacgtga 14640
tgcagtcgga ccgggatgat cgcttgtccg gcggccggat gcctagcctc gggagcaacc 14700
acagcggctc ttcacgagag gggtcgacca tgggcgatct caggaaccgc atcaccgagc 14760
tggtcgcgc gtaccaccg gaacaggcgc ccgggggctt cgttcccggg acgacgcacg 14820
taccggcttc cggcgcggtg ctgagcgagg aggaccggct ggcgctggtg gagacggcgc 14880
tggagatgcg gatcgcggcc ggcccggcct ccgggggctt cgagcggcag ttcggccggt 14940
acctcgggct ccggaaggcg cacctgacca actcgggttc ctccgccaac ctccctgccc 15000
tcggcgcgct cacctcgccg cagctggagg agagacggct gcgtccgggg gacgaggtcg 15060
tcacggctgc cgccgggttc cccacgacgg tcaaccgat cttccacaac gggctggtgc 15120
ccgtcttcgt ggacgtcgag ctcggcacgt acaacacgac gcccgagcgc atcgagcggg 15180
ccatcggccc ccggaccagg gcgatcatga tcgcgcacgc cctgggcaac cccttcgagg 15240
ccgaagaggt ggcccgctc gcggacgagc gggagctggt cctcatcgag gacaactgcg 15300
acgcggtggg gtcccgtac cggggcaggc tcaccggctc cttcggcgac ctgtcgaccg 15360
tcagcttcta tcccgcgcac cacatcgca tgggtgaggg gggctgcgtg ctcaccgaca 15420
acctggccct ggcgcggatc gtggaatcac tgccgcactg ggggcgcgac tgctggtgcg 15480
agccgggtga ggacaaccgc tgcctcaagc ggttcgacca gaagatgggt gacctgccgc 15540
ccgggtacga ccacaagtac atcttctcgc acgtcgggta caacctgaag tcgaccgacc 15600
tgcaggcggc cctcgggctg tcccagctga ccgggatoga ggagttcacc gaggccaggc 15660
gcgccaactg gcggcatctg cgcgccgcgt tggacgggct gcccggtctg ctgctgcctc 15720
atgccacacc gggcagcgat ccgagctggt tcgggttcct catcaccgtg gaccgggacg 15780
ccgcgtacag cagggcgggc ctggtcgacc acctggaatc gcgcgggatc agcaccgcgc 15840
gcctgttcgg gggcaacctc gtgcggcacc ccgcctacac cgaccgtcgg taccgggtgt 15900
ccggctccct ggagaacagc gacctgatca ccgaccagac gttctggatc ggggtcttcc 15960

ccggcatcac cccggagatg atcgccctacg tcggcgacac gatccgggag tgcgtgctca 16020
 agcactcctg aggggcggcg gtacgggggt gcgcgtggag gggaggaccg gggcggtcct 16080
 cctccggggc gtcacggccg cggaagcgtc cgaagctccg aggcgaggtc ggggtggacc 16140
 cgcgcggtga gcagggtgcc ctccggcgtg tgctcgggtg tgagcacctc gccctcgtcg 16200
 tgcaccccg ccaccaggt cccctcgtcg taggggatca cgacctccac ctccacctcg 16260
 ggggtgcggca gcaggcggtc gatcagttcc cgcagctcgt cgatgccccg acccgaacgg 16320
 gcggacacga cgatcgctc cggtcctgc tccagcagac gggcgaggac gtccgggtcc 16380
 gcgacatcgg ccttgttgac gaccacgacc tcgggtggact cgccggcgcc cacgtcccg 16440
 agcacctcgc gcaccgaggc cagctgcgcg ccgggggtccg ggtgcgaacc gtcgaccacg 16500
 tgcagcacca gatgcgcgtc cgcgacctct tcgacgtgg aacggaacgc ctccaccagg 16560
 tgggtgcggga ggtggcgtac gaagcccacg gtgtcggcga tgggtgtaggg gcgcccgtc 16620
 ggcgtcgtcg cccgccgcac ggtcgtgtcc aggggtggcga acagggcggt ctccaccagc 16680
 acgccggctc cggtgaggcg gttgagcagc gatgacttgc cggcgttggt gtagccggcg 16740
 agggcgaccg acagcacctt gttgcgccgt cgctcctccc ggttcacgtc ccggccggtc 16800
 ttcagctgct ccagctcccc gcggagcctg gccatcttgt cgttgatccg ccgccggtcc 16860
 gtctcgatct tcgtctcacc gggaccgcgc gtggccatgc caccgccgc accgccgcc 16920
 atctgccggg acagcgactg gcccagccg cgcagccgc gcagcatgta ctgcatctgc 16980
 gccagcgcca cctgcgcctt gccctcccg gactgggcgt gctgcgcgaa gatgtccagg 17040
 atcagggccg tcgggtccac gaccttgacc ccgacgacct cctccagggt catcagctga 17100
 ctgggggtca gttccccgtc gcacaccacg gtgtcggcgc cggctctctc gacgatgtcg 17160
 cgcagctgcg acgccttgcc cgagccgatg tacgtgcgc ggtccggctt ctgccggcgc 17220
 tgcacgacgc cgtccagcac gagggcgccc gcggctctcg cgagcgccgc cagctcggcg 17280
 agcgaactgt cggcctccgc ggccgttccc gaggtccaga tgccgacgag caccacccgc 17340
 tcgagtcgca gctggcggtg ctcgacctcg gagacgtcgg tcagttcggg ggagagcccc 17400

gcgacgcggc ggagggaggc ccggtcctcg cgggtctaact gatcgccgtc ccattcctcg 17460
 gcgtcggcgc cggccatcgt gtcgtccatc agtgcgctcg ctcgctgcgt gccggcgaag 17520
 tcctcgggat gcgtcaaagt acttccaatc tgggggggttc gagagcgttc cggggcgggg 17580
 cgcgtccggc ctgcggggcc gggcgcgggg ccgggggggtc gggggccgacg tgtccgcacc 17640
 tcgccgtcgg gggccggtgc cgtggagggtc gctgccggaa ctcacatgcc gtccacccta 17700
 gccatgtcgc cggggtgtgc aaggacgtat cggcgatcac ctcccagct gacagagtgt 17760
 gcaccgggta ctgcgcccgc gggcgtgaag cgggcgagtg tggatgtcat gcgagtactc 17820
 attatcgggg gttcacagtt cgtggggccgg gccttcgccg ccgaggcact ggccgcgggg 17880
 caccgggtca ccacgttcaa ccggggtgtc agcggcacccg acctgcccgg cgtcgaggcg 17940
 gtcagggggc accgcgaggt ggccggcgac ctggagcggc tgggtgtccgg aaggcactgg 18000
 gacgcggtcg tggacacctg cggttacgtg ccccgcacgg tgggggcctc ggccgcggcg 18060
 ctgtccgggc acgcggacac ctacctctac gtctccagca tcgcctgcct gcccgactgg 18120
 gcgcaggcgg tccgtccggt ggacgacgac tcacctgcc acctgccc gccggacgcc 18180
 ggaccggacc acgcgacgg tgactacggc gtcctgaagg ccggctgcga gcgcgccgtg 18240
 gaccggcact tcgcggggcg gaccctgcac ctgcggggcg gtgtcatcct cgggccgcac 18300
 gacacatgc gcatgctcga cgcctggctg tggcgcatgc gcgtcgccga gggggagcac 18360
 cgccgggtgc tcgccccgg caaccccgag gtgggcatgc gcctgatcga cgtacgcgat 18420
 gtcgccgtct tcggcctcga ctgcctcgcg gacggccgta ccggcgccct catcgtcaac 18480
 ccgccggaga agaacaccac cttcggggag ttgctcacgg agtgcgtaa ggccaccggt 18540
 tcggccgcgg agccggtgtg ggtcgacgag gggttcctcg ccgagcacgg cgtgagtccg 18600
 tggacggacc tgccgatgtg ggtgcccgac accgcgcggg acaccctcgt gtgggcgggc 18660
 ggagcaccgc gcgccccggc cgcgggtctg gcctgccggc ccttctccga caccgtgcgg 18720
 gacgcctggg aggtcgtccg ggaccggccc gtcccgaac tgccgctcgc ggccggctgc 18780

ggcctgtccc tgagccggga gaaggagctg ctgcgcgcct gggacgctcg cggcgggtgcg 18840
 gcggcgggct gacgcggccc ccgaggagtg cgacgcggtt tcccgcctcg ggggagcggc 18900
 gctcgcgtgt ccgccccgcg gcccgggcggc gcgtcctgtg tcccggcgcc cggggcgcgg 18960
 ctacacgggtg aaggctctcg cgcgggccac gatctcgtcg accagccccg cctggcgag 19020
 cgcggcgtac gactcggcgg agacgtcctc ggcccggatc accgcacggc ggaagaccga 19080
 caggatgttc acgaagtgcc ggtccacggg caggacgcgc tcctcccggg ggtcctggcg 19140
 ggacagccgc agtacggggt ggtggtgtc cggggtcgtg aagacgtggc tcagggcgag 19200
 cgagccgggtg ctgccgtgca gctcgtacgc cgagcggtag ccgtgttcca tgccgaaggc 19260
 gaggtgggcg gccacgccct ggggggcggc gagcaggacg ctgccggaga ccacgacgcc 19320
 ccggcggcgg tcacggcgca gcacggcgcc ggtcaggcgc agttcgggtc cgaggaagtg 19380
 cagcgcggcc cgacgcgggt agacgccgtt gtcgagcagt gcgccgccgc cgatgtcggg 19440
 tcggtagcgc atgtcgtcgt cggcgcgcgg cgggatggtg aaggcggcgg agaaggtgcg 19500
 cagctctccg atggcaccgg cctccaacag ggctttgacg gtggcgtgtt gagagtgggtg 19560
 gaggaacatg aagttctcca tcagtaccag tccgcgtcc cgggccatgg cgaacaggcg 19620
 ggccgcgtcg gcgtggttg cggcggcggg cttctccacg aggacgtgtt tgcccgcgcg 19680
 cagcgcgcc gcggccatt cggcgtgcag catgctgggc accgcgatgt agacggcgtc 19740
 cacgtcgggg cgttccagca gggcctcgta gggcgcgacc gcctcgcagt cgaagtgcct 19800
 gcccagggcc ttggcccgtt ccgcgtcgcg gctgccgacg caggtcagca cggtgccggg 19860
 ggtggagagc agggcgggca ggtgcgacg gcccgcgatg tcgccgcagc cgatcgtcc 19920
 gaagcgcagc accggggacg cttccggccg cggcgggccg ggtgcggcga acggttcagc 19980
 agaagtaccc agggcgggtca tggatggttc cgtcaatcgg tccgatggt cacggcacag 20040
 cgtggtgaac ggccggcgca caccacagag tcgcttcgg caatgggcgg ccctgaagct 20100
 acccggcggc ccgacgcgac ggcaaggccg atctgacaca gtgccgcgcc gagggaccgt 20160
 cggtgccccg cgtaccgcgt gacctgacac tacgcgcgacc aatctgtgga cggggcgggtg 20220

gccgcccggg ttaggtgagc gcggcgcccg cggtgaccgc gacgccggtc cgaagcggcc 20280
 ctcgggggcac gccgttgag agaggagacg cgggtgacgg acgcgatcac gaccgagctg 20340
 gccgaccgcg aactggggcg cagactgcac cggatacgcg gcgtccactg gtatttcggc 20400
 aaccacggtg acccgtagc cctcatcctg cgcggtcaga ccgacgaccc gtcggtgtac 20460
 gaggagcggg tccgcgaggg cgggccgctg ttccgcagcc gtaccgggac ctgggtgacc 20520
 gcggacccgg aggtggccgc ggccgtgctg ggcgactcgc gcttcggtgc gctggaccgc 20580
 gccggacggc gcccgaggga gtacctccag ccgtcgccc cccacgtacct ggggctggac 20640
 cgcgccgcgt acgcgcgtct gcggcggtg gccgagccc tgctgggccc ggacgccccc 20700
 gccgcgtggc gccggctcgg cgaggacgtc gggcgccggc tgctcgccgg ccgcggttcc 20760
 ggccctcgacc tgacggcgga cttcgcccgc cggctgcggc cattgggtcct ggccgcgtgg 20820
 ctcggggtgc cgggcgaacg gtgcgacgag tgggaggagt cgctgcgggc ggcggggccg 20880
 ctgctggacg gtctgctgtg tccgcagacg ctggcgccca cccgtgcggc ggactcgccc 20940
 gccgaggggc tgcgcgcgct gttggacgag gtggctgcgc cgcgtcccgg cgggtccggc 21000
 gaggggtgcg tggcccgcac ggtcggcgcc ggagccgccc ccgacgacgc ggtggccgcc 21060
 gccgtgtgcc tggcgctctc ggccgtcgaa ccgacgacga ccctggtgtg cgaagcggtc 21120
 cggctgctgc tcgaccgacc cgagtgggtg cggcggttgt gcgactcccc cgctctggcg 21180
 ccggccgcgg tccggcacac cctgcggcac gcgcccccg tgcggtgga gagccgggtg 21240
 gccacgagg acgtgacggt ggcggatcgt ccgctgccc cggggagcca cgtggtggtg 21300
 ctctggggcg cggcacggcg cgcgggcgcc ccggccgcgg agccggcgga cctggcgggc 21360
 gcaccggcgg cggagctgcc ggacgacctg tggttcgcgc tgtccgggga gttcgtcggc 21420
 cgtgccgccg agaccgcgct gggcgtgctg gccgaggccg ccccgggact gcggcgggac 21480
 ggcgacatcg tccggcgcg cgttccccg gtccctcgga ggtacgcgc gttccccgtc 21540
 gcgtactcct gacgggccc cggccggcgt cccctcagtc cccacgacg tttcatgaaa 21600

ggagtgccgt gcgcgtcctg gtgacctcca tcccgcacca cacgcactac taccacctgg 21660
 taccgctgat ctgggctctg cgtgcctcgg ggcacgaggt ggtggcgggc ggccagccgt 21720
 cgctggtcga cgccatcacc gccagcggca tcccggcgtt cgccctggcc gaggaggagt 21780
 cgctggcgca gatcttcgag gaggtcgagg gcgatctcca gccgtatcag cacggcatcg 21840
 acgagttcga cttcttcggc accctgaagg acgagctgga ctgggagaag ctgctcgccc 21900
 agcaggtgat cctgtccggc ctgtggctgg aaccgctcaa cggcgccacg accctcgaca 21960
 gcatcgtcga cttcgcccgg gcctggaagc ccgacctggg gctgtgggag ccgttcacct 22020
 atgcggggcc ggtggcgggc cgggcgtgcg gggccgcgca cggccgcgtc ctgtgggggc 22080
 cggacacgat cgggctgctg cggacgaagt tccttcaggc ccaggcgcgt cagcccagag 22140
 agcaccggga cgaccgggtc gcggagtgga tgacctgggc cctggcgcgc tacgggtgcg 22200
 acttccggga ggaggacgtg ctcggtcagt ggagcgtgga cccgatggcg gagggcgtca 22260
 gtctgggcct cgacctgccg accgtcccga tgcgtacac cccgtacaac gggtcggcgg 22320
 tgatccccga ctggctgacc gaggaaccga aacggcctcg ggtctgcctg accctggggg 22380
 tgtcctcgcg ggagcacagt gaggacgagg tcccggtgca gaggtttatc gaggcgtgg 22440
 ccgatctcga catcgagctg gtggcgaccc tggacgacgc ccagcgggac ctgctgccga 22500
 ggatccccga caacacgcgc atcgtcgact tcgtgcccat ggacgcgttg ctgccgacgt 22560
 gctcggcgat catcaaccac agcggttcgg gcacgtgcaa caccgccgcg ctgcacgggg 22620
 tgccgcagat catcctcggc ggcatcctgg acgccgccgt acggcagcac atgttcgcgc 22680
 agaactccgc cgccctcacc ttcgctccgg aggaggtgac cggcgcgctc ctgaggagcg 22740
 cgctgggtgcg cctgctcgag gagccgcggg tccgcgacgg cgcgcggcgg ctgaaggagc 22800
 ggatgcgggc catgcccagc ccggccggga tcgtgccgac cctggagcgc ctcacggccc 22860
 agcaccgccg ggcgtgttga accggcgcgc gggcccgtgc cggcggtgac cgcccaccc 22920
 gactctcgcg tgtgatcgat ctcgtcgact cagccgtgga ccggttcgcc tgtccgcgcc 22980
 cgacactgga gtgctgatgc gggccctctt cagcaccgcg ccgctcgcgg gccacctgct 23040

tccgctggtg cccatcgcggt gggccctgcg ggcggccggc cacgaggtac tggtagcgac 23100
 ccgggaggac ttcgtgccgg tcgccctgcg gtcggggctg ccgtccgcct cgtgcggggc 23160
 gcccgccgcg gacctggcgg gcgcggccga ggcgggggcg ctgcgcggc cccgcggagc 23220
 ggcggaggct cgggggggtcc tgagcggggc gctggcgcg ctcgcccggg gcagtctggc 23280
 gggggtgcgg cggctggcgg acgcctggcg gccggatctg atcgtcagcg aacgggccga 23340
 gttcgccggg ccgctggtcg cggcggccct cgggggtccg tgggtccgct accactggtc 23400
 ggtctcgtcc ctggaggagt accggcgagc ggccgaggcc gagttcgcg ccgagctggc 23460
 ggcgctcggc ctgcaccggt tcccggaggc ggcgcgcgtg ctcgatccgt ggccggtgtc 23520
 gctgcgccgg ccggacgcgg tcgcccacga cggggtccg cactaccgg ccacgggga 23580
 cggccccgtc cccgactggg cgttcacgcg cggtcgcggg ccgcggtatct gcgtgacgct 23640
 cggcaccatg ctgccccggt acggcgccgc cgggatggcc gacttcctga cggagctggt 23700
 ggcggagacc cgcggagggg actgcgaact gctcgtggcg gtcgacgacg acgtcgtcgc 23760
 gcggtggccg tcgctgccct ccgcggtgcg gtacgccggc cggctgccgc tggcggaggc 23820
 gctgcccgcg tgcgacgcgg tggtagacca cggcgggcag ggcacgtccc tgaccgcgt 23880
 ggcgcggggt cggccgcagg tcgtcatggc gcggctcgac gaccagttcg acaacgcgcg 23940
 ggcactggcg gcggcggggg cggccctgct cgtaccgcgc tcccgggcca ctcccgcggc 24000
 cgtggcccgcg gggtagcgcg aagtgtgga gaacgcctg tatgccaagg cggcagccgg 24060
 gctcgccgag gagatggcg tgctgccgtc gccgtcggcg gcggtcggac tcctggaaca 24120
 cccggggccc gggccggaca tgccgcggag ttaccogaac gaggatgcgg tgtgacgtga 24180
 atctggaagt actcaaccgt tcgaacgatc cgcgcggggc ggtgatcacg gtggtcggcg 24240
 cgtccggctt catcgggtcc gccctggtag ccgagctggc gcgcattgccc gtgcggctgc 24300
 gggcgggtggc ccggcgcgag acccccgttc ccgcgggggc acgggccgcc gtcgaggtcc 24360
 gccgggcgga cctcgcccgg ccggacgagg tcggggccgc cgtcgagggg gcggacgccg 24420

tcgtgcacct	cgccgcccac	atcggcgggc	cgcggtcgtg	gcgcgcggcc	gacgagcgg	24480
cgctgcgggt	gaacgtcgg	ctgctgcgc	acgtggccga	cgcgttccg	gaccgctcg	24540
ggcccgc	ggccgtggtc	ctggccagta	ccctccaggc	cggcgtcgag	ctgtcccggc	24600
agggcccgt	cgcccggcag	aagtcggcgg	ccgaggagg	cctgctgcg	gccgcctcc	24660
aggaggtgg	ccgcggcgtc	gtgctgcgg	tgcgaccgt	ctacgggcgc	agcccgtga	24720
ccgggtggac	gggcccgcgg	gtggtcgcgt	cggtggcacg	gcaggccgtc	tcgggcgagc	24780
cggtcacgat	gtggcacgac	ggcacggctg	ggcgcgatct	gctccacgtg	gaggacgcgg	24840
cccgcgcct	cgcggcggcg	ctcggtcacg	tggagcggct	ggacggcggc	acgtggtccg	24900
tcggtacggg	ccggctggag	cccttgggag	aggtgttctc	ggccctcgcc	gggctggtgg	24960
ccgagcggac	ggggaggccc	cccgtaccgg	tgggtctccac	ggagccgccc	gaccatgccg	25020
aggcgggcga	cttcgagagc	gcggtctgtg	accctccgc	gttccgcgcg	gtgaccgggt	25080
ggtctccct	cgttccgttg	cgggcggggc	tcggcgccgt	ggtggagacg	atggtggccg	25140
acggagcgag	gggtgggatc	cgaacgtgag	cacggaccgg	gagcaggccg	cgcacacgcg	25200
gctcggtcgc	agcgcgaccc	tggtgagccg	gctctggctg	ggcacctga	acttcagcgg	25260
ccgggtcgag	gacggtgacg	cgatgcagct	gatggaggcg	gcggtcgacc	gcggcatcaa	25320
ctgcatcgac	accgcggaca	tctacggctg	gcggatccac	aagggccaca	ccgaggaact	25380
ggtgggccgg	tggctggcca	agagcgccgc	gcggcgggag	gacgtcctgc	tggccaccaa	25440
ggtcggcggg	gacatgagcg	aacggctcaa	cgacggcggc	ctgtcggcgc	ggcacatcgt	25500
cacggcctgc	gagcagtcgc	tgcggcgcc	gggcgtggac	cacatcgacc	tgtaccagat	25560
gcaccgcgtc	gaccacgccg	cgccgtggga	cgagatctgg	caggcgatgg	accgtctggt	25620
ggcgagcggc	aaggtgacct	acgtggggtc	gtcgaacttc	gccggctgga	acgtcgccgc	25680
cgcgaggac	gcggcccggc	ggcgccagtc	cctcgggtctg	gtgtccgagc	agtgcctgta	25740
caacctggcg	gtgcgccacg	ccgagctgga	actgctgccg	gccgcccagg	cgtacggact	25800
gggcgtgttc	gcctggtcgc	cgctgcacgg	cgggctgctc	agcgggggtgc	tgcgcaagct	25860

cgcgggcgggc gtcgcggtga agtcggcaca ggggcggggcc cagctgctgc tgcccagact 25920
 gcgcgcgacg atcgaggcgt acgaggggtt ctgcggccgg atcggcgcg atccggccga 25980
 ggtcgggtctg gcctgggtgc tgtcccggcc ggggatcagc ggcgcggtga tcgggtccgcg 26040
 cacggtggac cagctggact cggcgtgcg gtccctggac ctggtcctcg gggaggccga 26100
 actggccgag ctggacgcca tcttcccgcc cctgggcaag ggcggccggg cgccggacgc 26160
 gtggatcagc tgaaggggt gcatcgccg acgtcacgcc ggccgatgcc gccggtcaca 26220
 cgacgtcgag cgcgggcagc gggaagacca gtcggccgcc gccgtcgagg aactcccgt 26280
 cccgttcgac gaacccgtcc cggtagatcc agggcaggac cagcaactgg tccggcttct 26340
 gcgccttcgc gtcctcctcg gacacgatgg ggatgccgt cccgggggtg aaacgccccg 26400
 ccttctctc gtcacctcg ccgatgcacg gcaggccccg ttcggtgatc ccgcagtact 26460
 ggaggatgac gttgcccttg gtggaggcgc cgtacccgag ggtcagcagg ccctcttggc 26520
 gggagcggtc caggaagccg cgcaggcgt cccgctggtc ggcgacgcg cgggcgaagg 26580
 cctcgaacgg tgccatgccg tccagtcccg cggcggcctc ggcggcccgg atgcgggcca 26640
 ggcccgcctc gtccctcggg tgccgggaac cggtcctggc gagcgtgacg cacaggctgc 26700
 cgccgtacac ctcggtgagc tcggcccggg tgacgggtgag gccgacgcgt tccgccatcc 26760
 actcgatctg gcgcagcgc tagtactcca ggtgctcgtg gcagacgatg tcgtacgcgt 26820
 cggcttcgag catggcgggc aggtagctct gctocatcat ccagacgccg tcctcggcga 26880
 ggacgtcgcg gacgtcgtc atgaagcgca gcgggtccgg caggtcgtag aacatcgca 26940
 tggaggtgac ggccttcgcc cgccgcgcc cgaagcggtt ctggaaggc gcgcgggtga 27000
 agtagtccac gaccaactcg gcgttcggcg ggtacaggtc gcggaacttg ccgccggtcg 27060
 ggtcgatgcc gaccagtcgc gggccgtcgg cggggtagcc cttgaggagc gtggcgctcgt 27120
 tgctgccgat gtcgaggacc aggtcgtccg ggccgaggtc caccagccgg cggacggcgg 27180
 cgaccttgcc atgcagggtg tcgaccatga agggccggat gcccgagcgg tagccgtagc 27240

cctcgccgta catgagggtcc gggtcggggg tgtggcgag ttgcacgagg ccgcatccgg 27300
 ccggggaaca cgcgacgagt tccagcggga ccgacggcac gacctcgtcg cggtcggccg 27360
 ggaacacccc ggtgagcgcc tgttcgccc ggtcgagtag ggagagcagc tccttggtgc 27420
 cgcagacgcg gcacgcggtg gcaatcatgg ggtcctttcg ggatcctggc cggggcgccg 27480
 ccggcccgtg gccaggtcgg gagcttccag acggaggggtg tcggggggcg gccgggggga 27540
 cggcaccagg tgcaggcggc cgatcgcgcc ggtggcgggc gaggcgtcct ggcacgggca 27600
 gggatcgtcg gtgaatccgg cgaagcggtg ggccacttcc atcatccggt tgcgttcggt 27660
 ggcccgaag tcggctccga ggtgcacccc ggcccgggtc gcctgggtccg tcagccagcg 27720
 caggatcacc gtgccggccc ccagcgacac caccggcac gaggtggcca gcagcttgat 27780
 ccgccaggcc tcggggccgc gccgcagcag cacgacggcg accgcgccgt agggaccgaa 27840
 gcgatcgggt acggtggtga ccagcacctc gtggtcgggg tcgtcgatga gggcgcgag 27900
 ctcgctctcg gagtagtgca ctccggtggc gttcatctgg ctggtgcgca gggtcagttc 27960
 ctcgaccgc gacagctcca ggggcgtggc gcggctgata cgcattccga tgtccagcga 28020
 gcgcaggaag tccgcgtcgg gtccggtgaa gtcggaccgc tcggcggtccc ggcggaagga 28080
 cgctgtgtac atggagcggc ggcgggtcga gtcgacgggt acggtgccc ggctgaactc 28140
 cgggaggccg gtcagccggg tcgcctgctc ggcggcgtag gtgcggacct ccggcagttc 28200
 gtgggtgacc tcggcccgtt cgaagggtcg gtcgtcgatg aaggcgaggg tgctcggcgc 28260
 gaagttcagc cggtcggcga tctcgcggac cgacttcgac ttcgggcccc agccgatcct 28320
 cgggagcacg aagtactcgg cgacgccgag ctgttcgagc ttcgcccagg cgtgatcgtg 28380
 gtcgttcttg ctggccaccg cctgcaggat gccgcgcgcg tccagctcgg tgatggtccg 28440
 caggacgtcc ggggcgagcc ggacctcgtc ctctcgagg aggggtgccct gccagaggg 28500
 gttgtccagg tcccagacca ggcatttgac caacggctcg gcggcgttgt cccggtgcac 28560
 gtcgttccct tctctcgata cggcggttgc gaccgcgact cctcgacagg gcgtacgcgg 28620
 cccgcggccg gcggggacac gggccggttc acccggcgcg ggacatcacg tgccgggcca 28680

ggaccagttc gcagatctcg ttgctgccct cgatgatctc catgagcttg gcgtcgcggt 28740
 gcgcccgtgc gacgacgtgt ccctcccgcg ccccggccga cgccagcacc tgcacggccc 28800
 gttcggcccc gcgtgcggcg ccggtggccg cgacgtgctt ggccaggacg gcggcgacca 28860
 ccatgtcggg gctgccctcg tcccactggg cgctggcggtg ctgcgacgcc cgggcgcggt 28920
 gctgttcggc gacgaacagt tcggcgaggt gccgggcgac gagctggtgg tccgagagcc 28980
 gcgtgccgaa ctgctcccgt ccgccggcgt gacgtgcggc ggcgggcagg caggcgcgca 29040
 ggatgcccag ggagccccag gccaccgaca tccggccgta gctcagcgcg gtggtgacca 29100
 gcagggcggg ggtgcggtcg tggccctgga gcagggcgtc tgccggcagc cggaccccgt 29160
 ccaggtggat gtcggcggtg ccggccgcgc ggcagccgtg cgcgtcggtg atgcgctcga 29220
 cgcgtagccc gggggccgag gcgggcacca cgacggcgcc cgcgccctgc tccgtgcggc 29280
 cgaagaccac cagcaggtcc gcgtacgcgg cgttggtcgc ccacaccttg acgccgtcga 29340
 cgacgacttc gtcgccgtcg gaggtgatgc gggcgcgag ggcggacagg tcgctgccgg 29400
 cgccggcctc ggtgaacgcc acggcggcc a gttcgccgcc ggtcaaccgg ggcaccaggg 29460
 aggctgctg gtccgctccg gcgagccggc gcagcgccca ggcggccatg ccctgcgagg 29520
 tcatgacgct ccgcaacgag ctgcagaggg cgccgacgtg cgcggtgagt tcgccgttgc 29580
 gccggctggt ccagccgagg ccgccgtgga ccgcgggcgc ctgtgcgcac agcaggcccc 29640
 gggagccgag gtcgtgcagc aggccgaggg gcagctcgcc cgtccggtcc cactcggccg 29700
 cccggtcgcc gaccagcgcg gtgaacagct cctcggcctc ggtgccgtcg gcgtgggagg 29760
 tgtcagccac ctgctgccc ggtcgcctcg ccgggcgcgg ccagccgtcc gaccagccgg 29820
 accatcgct cgacggtgcg gaagttgtcg agcatcaggt cggcgccgct gatgacgatg 29880
 ccgtaggtct tctccaggtg cacgacgagc tgcatggcga acagcgagga catcccgccg 29940
 acggcgaaca ggtcctggtc gcgctcccag gtggtcttgg tcgggtcctc gaggaacccg 30000
 agcagttccc cggcgacctc gtcggcggtg ggggtcgtgc cgggtgggtc gggccgaccg 30060

gacgtcgtgg tcacgcgtg gcctcctggt agtcgtagaa tccccggccg ctcttgccgc 30120
 cgagcaggcc ctggcggacc ttgtccagca gcagctcgct cgggcggagc gccggatcgc 30180
 cggtccttgc gtgcatcacc cgcagcgagt cggccagggt gtccagtccg atcaggctcg 30240
 ccgtggccag gggtcgggtg cggtgggcca tgcagtcgag catcagcgcg tccacggtct 30300
 ccgggggtggc ccggccctcg tgcaccaccg cgatggcgtc gttcagcatc cggtgacgca 30360
 ggccggctggt cacgaagccg gcgccgtcgc cgacgacgat gccccggcgg cccaggccgg 30420
 acagcaggtc ccgggtggcc cgggcggccg cctctccgct gcgcgggtccg aggaccacct 30480
 cgaccgtggg gatcacgtac gcgggggttca tgaagtgcac gccgacgagg tcctcggggc 30540
 gggggacggc gtcggccagc tcgtcgatgg ggacgcccga ggtgttgctg acgagcagcg 30600
 tccccgggcg cgccacggac gccagggtccg ccagcacctc ggccctccgc tcgggggtcct 30660
 cggtgacggc ctcgatcacg gcggtcgcgg tggcgacggc ggccggcgcc tcctcgacgg 30720
 tcagctcccc gggcgggagg tcgtggggca gcgcgcccac cagccggggc gtccgcagat 30780
 gcagcgcgac cgcgtcgggg gcggccgcgc gcgccccggc ggagggtgtc accagtgtga 30840
 ccgggtgccc gtgtccgacg gcgagtgccg cgatggccgt gcccatgacg cccgcgccga 30900
 gcacgacgag cggagaattt tccttggaat cgggcacgga tcc 30943

<210> 2

<211> 11171

<212> DNA

<213> *Streptomyces ambofaciens*

<400> 2

ctgcaggcgc gcgggctgct cgggcaggag gcgcccgcg agagcctcga cgcgatgac 60
 gacgactact gcgcgcagat ccgcgaggtc cagcccaggg gcccgtagc gctgctcggc 120
 tgggtccatgg gcgggctgct cgcgcaccgg gtgcgccacc gcctccagcg ggccggggcag 180
 cgcgtcgacc tgctcgcgat cgtcgacgcc tatccgccc cctcgggtgc ggccgactgg 240
 gacgcggcgg agatggtggc gcggatcggg gaggaactgg gcttcgacgt ggaccgggtc 300

ggccccggcc aggaggaggc gctgctcgcg gatctgcggg cgaaggggca tcccctgggg	360
catctgccgg gcggtgacat cggcgcggcg gtccgcgtgt acgtcaacag ctccgcgcctg	420
acgaggaacc tgcggcccga ggcgttcgac ggggacgtcc tcttcttcgc ctccgggacc	480
tcgttcgccg aggacgacca ggttcacgac gtccgcgcct ggcgggcgta cgtcaccggg	540
cagatcaccg agcacgtgat cgaccacctc cacgaggacc tcttgatcga accggccgcg	600
gtcacggcca tctccgacgt gctcgtcggg catctgcccg cggacgacac accccgcgcg	660
tccggcgcac cccgcgcacc tcgcgcatcc cgatgaggaa ggaaacatca tgagcaaccc	720
gttcgaggac acggaagcca cctacgtcgt gctggtcaac gacgaggggc agcactcgt	780
gtggccgtcg ttcgcggagg tcccggcggg ctggtccgtc gtggtgccgg agacggaccg	840
gcagtcgtgc ctggactaca tcaacgagaa ctggaccgac atgcgcccc aagacctcgt	900
cgaggcgatg gcgacggccg ggcaggacgc cccttgagcc aggacgccat gatccgggcg	960
gtcggcgtcc gcaagagctt gcggcgacgt ccacgcctc gacggcgtgg acatcgaggt	1020
cgaacggggc cgggtgctcg ggctgctggg tcacaacggg gccggcaaga cgaccttgg	1080
gaacatcctt gccacggtct ccccggcgtc cgcgggcacg gtgaccgtcg ccggtttcga	1140
cgtcgcgacg cagggcgccg agatccgcgc gcgcacggg gtgaccggcc agttcgcgtc	1200
ggtggacgag tacctgagcg gattccgcaa cctcgtcctg atcgccgcgc tctcggggc	1260
gggacggcgt gaggcggcgg cccgggccac cgagctcctg gagctgttcg agctgaccgg	1320
ggcggcccac cagccctccc gcacctactc gggcgggatg cgccgacggc tcgacctcgc	1380
cgccagcctg gtcggccggc cggacgtgct gttcctcgac gagccgacga ccgggctgga	1440
cccggcgacc cggatcgccc tgtgggagac ggtggagaag ctggtggcgg gcggcacgac	1500
cgtcctgctg accaccagt acctggacga ggcggaccgg ctggccgact ggatcacctg	1560
cctgtcgaag ggccgggtgg tggcctcgga caccaccgac cggctcaagg ccgacctggg	1620
ccaccggtcg gtgcgggtgg tccttcgcc cgcgcgcgac ctgacggccg ccgccgccgc	1680
gctcaccgcc ggcgggttcc gtccgcggtc cgacgcgggg gagcacgcgc tgaccacgcc	1740

cgtggacacc	tcggccggtg	tcgcgggctg	catccgcgcg	ctggacaccg	tcggaacgca	1800
ggccgtcgag	ctgaccgtca	aggagccgtc	cctggacgac	gtctacctgg	cgctcaccca	1860
tccctcaccc	gccgccgacg	cggcctgac	cccaggagtt	ccgttggcca	tccaggagcg	1920
cgccgtgaag	gaggtggcgg	tcgacaccgg	taccccgggc	ggccccgtgt	ggcgtggtgc	1980
cggcctcggc	accagctgt	gggtactgac	cgcgcggcag	atccgttcca	tgtacggcga	2040
ccgccgcctg	gcgctgttca	gcctgatgca	accggtgac	atgctgttgc	tgctgagcga	2100
gatcttcggc	agcatggccg	acccggacga	cttcccgag	ggcgtgcgct	acatcgacta	2160
cgtggtgccc	gcgctgctgg	tcaccaccgg	catcggtcgc	gccagggcg	cgggggtggg	2220
cctggtcagg	gacatggaca	acgggatggt	ggcgcgcttc	cgcgtcctgc	cggcccggct	2280
gttctctggtg	ctggtcgccc	ggtcgtggc	cgatctggtc	cgtgtgttca	ccgagttggt	2340
cgtcctcgtg	gccgtcgggtg	tgatcctgct	gggcttccgt	ccggccgggg	gtttgtgggg	2400
cacgtccgcc	gccctgttgc	tcaccctggt	cgtcatctgg	tcgctgatct	gggggttcat	2460
cgcctcgcg	gcgtggctgc	gcagcgtgga	ggtgatgtcc	agcctcgcgg	ttctggtgat	2520
gttcccgcctc	atgttcgcct	ccagtgcgtt	cgtcccgcgtg	gacgccctcc	cggagtggct	2580
gcgctcgggtg	gcgcacctca	accccgtagc	gtacgcggtc	gacagcgccc	gccgcctggc	2640
gctggactgg	gacccggggt	ggagcgtgcc	cggcgcgctg	ctgaccagca	ccgcgctcat	2700
ggcgggtgggg	atgtacgtcg	ccgggcgttc	cttcaagagg	ccccgaacg	aatgattccg	2760
gcgaacgggt	ccgtcgctg	cccctccggg	ctcgcccggg	ggaggcgggc	gcggtctcag	2820
cggtcggccg	tccgcccgta	ggcgcggaag	agtgcggtcc	actccgggac	cgtgaggtcc	2880
ttgaccgggt	cggccggcct	cacgcccgcc	ccgcgcacga	actgggcgtg	tccgcggcgc	2940
agcaccttcc	gcgcggcgctc	gcccacgctc	atctggccgg	tgtcgaagac	ccgttggacg	3000
aaccgctggt	aggcggcctt	ctcacgccag	ggcacggacg	gcctgcgggtg	cggggcgacc	3060
atcaggtct	gggtgtccgc	gcgcggtacg	ggggtgaagt	cctggcgtga	gaaggccagg	3120

ccccggtcga acgagtacca cggctcccccac tgggcgttga agaggtttcc gccccaggct	3180
ccggtccgct ttcccacgta ctcccgtgc aacaggaaca cgcctgccg catgcgcgc	3240
ggacccatgt cgaggcagcg cctcagcatt ctggtgccg tgacgaagg aagattccc	3300
atgagtctga ccggctgccc gggcagttgc agggtcagga aatcctcgtt caccaccgtg	3360
acgtccggca gcgattcggc ggccagccgc cgagcccage ggggatctat ttccaccgcg	3420
agtaaaggcg tcccgggtga ggcgagcact ttggtcacc gccctgatcc cgcgcctatt	3480
tcgacggtca tcaggtcatt cggagaatcg ggcggaatgg tgtccgaacc gtccagctga	3540
gcggagaaac gacaggccgc ggcggtgta cgaaagaagt tctgaccca ttctctccgc	3600
gcggtgctgc gtgaatcagc gggtgccgag acggattgcg gtggcagtgg ggatttcaat	3660
gtcacctcgg cgacattacc aagtcttgac ccaacggtec atcaaccacc ggtatacccc	3720
atgcaattca gccaccgtcc ggcccgatcc cgacgttcgg ccgcgggcct ttcgggccc	3780
accggtgacc agcgcgcgaa ccggagttcc cggcgatatt tccggctcct ggcgactcgc	3840
cacagtctcg tcgggggscg ccggcgccat gactccgacg ggggcgattc ccggccgact	3900
tcgtcaggtc cggaggggtcc cccgggcggc gcgggcgat ggccgctgac tgtgtcatct	3960
gcgggttgct gcccgaaac accgcgcccc aggatgactc agccgttcct gcggtcgtcc	4020
cgaccccggc cccacccccg aaaagaggag cgagaatcca gtgctccagc gcgtcgatct	4080
gtcgtcactc accggcctcc gctggtatgc ggcgtgacg gtattcgct gtcacatcgc	4140
ccagcagggc ttcttcgccg accagcagg gggcagcgca ctgctgcaca tcccccgct	4200
cggttccatg gcggtctcga tcttcttcat actgagtga ttcgtcctcg cctggtcggc	4260
ccgcgacgag gactccgtgc cgactttctg gcggcgccgc atcgcgaga tctatccgct	4320
gcatctcgcg acgttcggca tcgcggctct catcatcttc tccctgtcgg agccggtact	4380
tcccggcggc tccgtatggg acgggctggg gcccaatgtt ctgctcgtgc agtcctggct	4440
tcccgacgcg accctcacgg ccagtttcaa cacgcccage tggtcgctct cctgtgagat	4500
cgccttctat ctgtcgttcc cgctgtggta ccggctggcg cgcaggattc ccgcacggcg	4560

gctgtggtgg	tgcgccgcgg	ggatcgccgt	ggccgtgacg	tgtgtgcccc	tgctggcggg	4620
cctgctcccc	gcgagcgagg	aggtggcccc	cgggatgtcg	ctcaacgagg	tctggttcgc	4680
gtactggctt	ccgccggtgc	gcatgctgga	gttcgtcctc	ggcatcgtga	tggcgctgat	4740
cctgcgcgcg	gggatctgga	agggcccccg	tccggcggtc	tgcacggcgc	tcctcgccgc	4800
gagttacggc	ctcaccaga	tggtgcccc	gatcttcacc	ctcgtcgcct	gctccgtcgt	4860
accggccgcg	ctgctgatca	cggcgctggc	cgacgcgcac	gtgcacggcc	ggcgacggg	4920
gctgcgttcg	gcgacgtgg	tgcggctggg	ccagtggctc	ttcgccttct	acctgggtcca	4980
cttcctgata	atccgctacg	gacaccggct	gatgggcggc	gatctgggct	acgagcggca	5040
gtggagcacc	ccggccgcga	tcgcgctgtc	cctgggggatg	ctgggggtgg	cggtcctggc	5100
cggcggtctg	ctgcacaccg	tcgtcgaaca	gccctgcatg	cgctgttcg	gcagccgcag	5160
gtccgcctcc	cgtccgaagc	ccggcgccac	cgcggtccc	cggaactcac	ccgcggccga	5220
cgcgcccggc	gtgcccctgc	tcccgggcgt	accggggccc	gcgcacaccc	ccgcagcgac	5280
gaacgaaccc	accccgagag	gatgatgagc	gtggcagacc	agacggctct	cagccccgcg	5340
ctgctggagt	acgcccgag	cgtcgcgctg	cgggacgacg	gcctgctgcg	cgaactgcac	5400
gaggtgaccg	ccgggctccc	cggcggcccg	gccatgcaga	tcatgcccga	ggaggcgag	5460
ttcctcgccc	tgctgatccg	gctcgtcggg	gcccggcggg	tgctggagat	cggcaccttc	5520
acgggggtaca	gcacgctgtg	catggcgcg	gactgcccg	ccgacggcac	cgtcgtcacc	5580
tgcgacatca	gcgacagggtg	gcccggcgtc	ggcgcacccg	actggcgccg	ggccgggggtg	5640
gagtcccga	tcgacctgcg	cgtcggcgac	gccgtccgga	ccctcgccga	gctccgcgag	5700
cacgaggggg	acggctcggt	cgacctggtc	ttcgtcgacg	ccgacaagac	cgggtacccg	5760
cactactacg	agcaggcgct	ggccctggta	cgccccggcg	gactgggtggc	ggtcgacaac	5820
accctgttct	tcggccgggt	ggccgacccg	gccgtcgagg	acgccgacac	cgtcgccgtg	5880
cgcgcgctca	acgagctgct	gcgcgacgac	gaacgcgtgg	acatcgccct	gctgacggtc	5940

gccgacggga	tcactctggc	ccgccggcgg	gagtgagtc	gcacggggtg	cggagcacct	6000
ggtgccgtca	ccggcccga	cacgtccgcc	gccgacctgc	ccgggcaggt	cgccgagctg	6060
ctgccctggc	gtggccggcc	ctcgtcgagg	gccgcgcggt	cagcggaacc	ggttgatcgc	6120
gtcgatgtgc	cgcgcccgt	tctcctcgtc	gcgcacgccc	agccctccg	tgggcgccag	6180
gcagagcacg	ccgaccttgc	cctggtgctg	gttgccgtgc	acgtcgtgca	cggcaaccgc	6240
cgtgtcggcc	agcgggtagg	tgcgcgagag	ggtggggtgg	atcctcccct	tggcgacgag	6300
ccggttcgcc	tcccacgcct	cgcggtagtt	ggcgaagtgg	gtgccgacga	tacgcttgag	6360
gtgcatccag	aggtagcgg	tgtcgaactc	gtggcggaag	cccgaagtgg	aggcgcaggt	6420
gacgatcgtg	ccgcccctgc	gggtgacgta	gacggacgcc	ccgaaggtct	cccgccggg	6480
gtgctcgaag	acgatgtcga	cgtcctcgcc	cccgtgacc	tcgcggatgc	gcttgccgaa	6540
gcgcttcac	tcccgcgggt	cctgggtgtc	ctcgtcgctc	cagaagcgg	agtcctcggc	6600
ggagcggtcg	atgatggcct	ccgcgcccac	ggcccggcac	acctcggcct	tgcgcgcgct	6660
ggagaccacg	cagacggggt	ggccccgcc	ggcgagggcc	agctgggtgg	cgtacgagcc	6720
caggccgccg	ctggcgcccc	agatcagcac	gttgctgccc	tgcttcatgc	cggcgccgtt	6780
gcgggagacc	agctggcgg	aggcgttga	gttgaccaga	ccgggcgcgg	cggcctcctc	6840
ccaggtcaga	tgggccgcct	tgggcatcag	ctggttgga	ttgacgagg	ccacctcggc	6900
caggcccccg	aagttggtct	cgaagcccca	gatgcgctgg	gccgggtcga	gcatggtgtc	6960
gtcgtgcccg	tcggggctct	ccagctccac	cgacagacag	tgcgcgacga	cctcgtcacc	7020
gggtttccag	acgttcacgc	cgggtcccgt	gcgcagcacc	acgccggaca	ggtcgggaacc	7080
gaggatgtgg	tacggcaggt	cgtggcgcgc	ggccagcggc	gaggtgcgcc	cgtagcgctc	7140
caggaagccg	aacgtcggga	ggggctcgaa	gatggaggac	cagaccgtgt	tgtagttgac	7200
ggagctggcc	atcaccgcga	ccagcgcctc	cccgggcccc	acttcgggca	gcggcacctc	7260
gtcgacgtgc	agggacttgc	gcgggtcctt	gtccgcgctc	ggcatcccgc	ggaacatgtc	7320
ggtgtcctct	ttgcggacgg	tcacggcgcg	gaaggactcg	ggcaggggca	gcgcccgat	7380

gtcctcgggg	gtgcggtccg	cggcggtgat	cgcggcgagc	agcgcgctct	gcgcatggct	7440
ttcgggcatg	gaacggtctc	cgatcgctcg	tgtcgtcagg	tggcccgggtg	cagggcggtg	7500
aggcgctcca	gggtggggat	gatcccggcg	ggcgtgggggt	cggcgagcat	ctcctcgctc	7560
agccgcccgcg	cgcccgccct	gatcgcgggg	tcgtgcacgg	ccgtgtgcac	cgcgtcggcc	7620
agaccgcgcg	ccgtgagcgt	cgccgcgggc	aggtcgaacc	cggcggacag	gcgctggagt	7680
tgctgcgctt	tgagcggcgc	gtcccacagg	gagggcagca	ggatctgcgg	cactccgtgg	7740
cgcagggcgg	tggaccaggt	gcccgtcccg	ccgtggtgga	tgatcgcggc	gcagctcggc	7800
agcagcgcggt	cgagcgggtac	gaagtccacg	ggcacgacgt	tgtcggggag	ggtgcccagg	7860
cggtcgagct	gggaggtgtc	caggggtggc	accacctoga	tgtccagccg	gccgagcgcc	7920
tccagcagtt	cggagtagga	gaccgcgtcg	cggccgtagg	tctcgcgggc	ggacacgccc	7980
aggggtgaggc	agacgcgggg	gcggtcgggc	ttcttgccga	gccagggttc	gatcacggag	8040
cgcccgttgt	agggcacgta	gcacatgggc	accgtggtct	ggcccaggtc	gaggcgggtg	8100
ctgcgcgggc	ccgggtcgat	caccagtggt	ccgagcacgt	cgcgttcgtc	gaaggcgagc	8160
ccgtaccggt	ccagcgtcca	gccgagccat	tcggcgatgg	ggtcctcgcg	cagttcctcc	8220
ggcatcccggt	cgagggcggc	gaggaagcgc	aggcgggagc	gtccgatggc	gtcggggccc	8280
cacaggatgc	gcgcgtgggc	ggcaccgcag	gcgcggggccg	cgaccggccc	cgcgtaggtc	8340
cagggttccc	acacgaccag	gtcggggccgc	cagccacggg	cgaagtcgac	gatctcgtcc	8400
atcgtggcgg	cgccgttgaa	cggggcgaaag	cacagggcgg	ccatcatgct	ctgctggccg	8460
agcaggtact	cccaggtggt	ctcctcctcg	ctctccacct	caccgaactc	gaagccccgc	8520
tggtagggga	cgaggtcgct	gcccattctcg	gcgagcagct	ccgcggccgg	tcggtcgctcg	8580
cccaccggta	cggcggtcag	gccggtggag	gtgatgacgt	cggtagggga	gggcgggctc	8640
gccacccgta	cctcgtgccc	ggcggcgcg	agcggcccagg	cgaggggcac	cagactgtag	8700
tagtgggtgt	tgtgggcgag	ggatgtcagc	aggacgcgca	cagcggctcc	ggtctgggag	8760

ggggcgctacg	gacggcgggg	tcattcccggg	tggaccocgga	ggcgggcgca	ggcgcccagg	8820
accggtgacc	gcggccgggt	caggggcgggt	ccctcggccc	gcaggccggg	caggcgggccg	8880
gccgtgacct	ggaccgcgggt	ggtggcgggcc	aggcggatca	gcgggagggc	gagcgccagg	8940
tgcggtccgt	cggcgccccc	gagcggtctc	ggcccgggga	tctcccgtcc	ggcggcggcc	9000
agcaccacga	cgtgctcgtc	ggcgggtgatc	cggcgggcgc	cgagtccgag	gccggtgtgc	9060
gcgacccggt	tctccagccg	ggccgggggg	cgtcgcgcga	gcacctggtc	gacggcgttc	9120
gcccgggccg	gcgtccggcg	ggccctctcc	cactcccccg	gccggccgag	cagccggtgg	9180
accgtgtgcg	cgacggcggt	gacgacgggt	tcgggtgcgc	cgaccgcgag	gagcagcgcg	9240
atgcgctcga	cgtcgccggg	ggcgacgccg	tcgtgcagga	ggagggcgag	cacgtcgtcg	9300
ggccgggggt	cggcgttccc	gaggcccccg	gacttctccg	cgacgagctc	ggggaccagg	9360
tcggccagtg	tgcggacggc	gtcggcgggac	tcccggggcca	cggtcagcag	ttgcggggcc	9420
atccgggggt	cgagctgggg	gccgcaggcg	gcgagtgcc	gcgcggcggt	ggcgcggtcg	9480
cggcccggca	ctccgaggag	cctgagcatc	agctcgacgg	cgtagggccg	ggcgacctcg	9540
ccgacaaggt	cgaaccgggc	gtcaccgggtg	ggcagcaggc	ggccgagcac	tctgcggggc	9600
gtggtgcgcg	ccgcgcaggg	cgcctggggc	ggggcgtaac	ggctgagcac	cggggcgggc	9660
agcgcccgca	ggcgtacgag	ctccgcgcgt	tcgtggtggg	ggaacgcctc	ggcgaggggc	9720
agcagttcct	cgtccggggc	gcgtccggcc	cggctccagcg	tgccgaagcg	cgggtcggcc	9780
aggacggccg	ccgccacctc	ggggtccgcg	gtcaccagg	cgtccagctg	ttcgtgcgg	9840
aaccacggtc	cgcgggcccc	gatctcccgt	tcgaacggct	cggggtcggc	gacggcgcg	9900
aggatcagcg	cgtacgggtc	gccctgggtg	ccggcgccacc	agtgtgcggc	ccgggtcagc	9960
tggagccgac	ggccgagcgc	acggacgccc	gtgctcgtgt	ccgcgctcgt	gctcgtgctc	10020
gcggggggccg	tttccgtggc	aagggtgggc	atggctgccc	tgcctttctt	ctcgggtcga	10080
tgggggcgct	cgcacgatcg	ggggaccccg	gtgtgcgggg	tgggtcgggg	caaccgcct	10140
ccggggccgcg	gcgggtccgcg	tcggcaccgg	gtcagcggcg	ggtgccgatg	aacagcccgc	10200

gaccggacgg gccgccttcc tggtagacca cgtcgcagcc ggcgcgctcg aaggccgcct 10260
 cgtagtcggc gcgcgggaac agggatgatg tgtggtcctc ggtcaggtga cggacgccgc 10320
 cgggtccggc gaggaggtag tgcacctcga tgcgggtggc gttcccctcc cgtacggagt 10380
 gcgagacccg gcagacggtg cgctcgcccg cctccgtgat gctggcgccg acgtaaccgg 10440
 gggatgaagga ctgcgggaac caccaggggt cgacgacgat gacgccggac ggttcgaggt 10500
 ggtcggatga ggcgcgcagc gtgctgtcga gttcgtcggg ggtccgcagg tggcctatgg 10560
 agctgaacat gcaggtcacc gcgtcgaacc ggcgtcccag ggagaacgag cgcattgtccc 10620
 cttggtggaa ggtgacaccg gggttccggc cggtcgcgag ggccagcatg tcggcggaga 10680
 gttccaggcc ctgcacgtgg tcgaagaggc cgtccaggtg gtgcaggtgc tggccggtgc 10740
 cgcaggccac gtcgagcagg gtccggggcg ccggccgggtg gacgcgcacg agtgccggcg 10800
 tctcctcggc ctctgcccg tagtccttcc ccttcccctc gtggaccagg tcgtagacgg 10860
 ccgcgatgtc gtcggcgtag atcagtggtt ccctccgggt agcggggcg gaccgggctg 10920
 gtgcgggagg ctgtccagcc attcgtgcac cagggaacgc gtgtgccggg cgtgttcggt 10980
 gagcatggtg aagtgggtgc cgggcacgtc ggcgacggtc cgtgcgaacg ggacctggga 11040
 ccgccagtcg ccccgcgccc cgcccgcggg cggccacgcg cacaggggtt cggacgcccg 11100
 ggccagcagg acgggggctt cgagggcggt gagccgggtt ccgagcacca gccgctgata 11160
 gccggccatg g 11171

<210> 3
 <211> 711
 <212> DNA
 <213> *Streptomyces ambofaciens*

<220>
 <221> CDS
 <222> (1) .. (711)

<400> 3

atg tac gcc gac gac atc gcg gcc gtc tac gac ctg gtc cac gag ggg	48
Met Tyr Ala Asp Asp Ile Ala Ala Val Tyr Asp Leu Val His Glu Gly	
1 5 10 15	
aag ggg aag gac tac cgg cag gag gcc gag gag atc gcc gca ctc gtg	96
Lys Gly Lys Asp Tyr Arg Gln Glu Ala Glu Glu Ile Ala Ala Leu Val	
20 25 30	
cgc gtc cac cgg ccg ggc gcc cgg acc ctg ctc gac gtg gcc tgc ggc	144
Arg Val His Arg Pro Gly Ala Arg Thr Leu Leu Asp Val Ala Cys Gly	
35 40 45	
acc ggc cag cac ctg cac cac ctg gac ggc ctc ttc gac cac gtc gag	192
Thr Gly Gln His Leu His His Leu Asp Gly Leu Phe Asp His Val Glu	
50 55 60	
ggc ctg gaa ctc tcc gcc gac atg ctg gcc ctc gcg acc ggc cgg aac	240
Gly Leu Glu Leu Ser Ala Asp Met Leu Ala Leu Ala Thr Gly Arg Asn	
65 70 75 80	
ccc ggt gtc acc ttc cac caa ggg gac atg cgc tcg ttc tcc ctg gga	288
Pro Gly Val Thr Phe His Gln Gly Asp Met Arg Ser Phe Ser Leu Gly	
85 90 95	
cgc cgg ttc gac gcg gtg acc tgc atg ttc agc tcc ata ggc cac ctg	336
Arg Arg Phe Asp Ala Val Thr Cys Met Phe Ser Ser Ile Gly His Leu	
100 105 110	
cgg acc acc gac gaa ctc gac agc acg ctg cgg gcc ttc acc gac cac	384
Arg Thr Thr Asp Glu Leu Asp Ser Thr Leu Arg Ala Phe Thr Asp His	
115 120 125	
ctc gaa ccg tcc ggc gtc atc gtc gtc gaa ccc tgg tgg ttc ccc gag	432
Leu Glu Pro Ser Gly Val Ile Val Val Glu Pro Trp Trp Phe Pro Glu	
130 135 140	
tcc ttc acc ccc ggt tac gtc ggc gcc agc atc acg gag gcg ggc gag	480
Ser Phe Thr Pro Gly Tyr Val Gly Ala Ser Ile Thr Glu Ala Gly Glu	
145 150 155 160	
cgc acc gtc tgc cgg gtc tcg cac tcc gta cgg gag ggg aac gcc acc	528
Arg Thr Val Cys Arg Val Ser His Ser Val Arg Glu Gly Asn Ala Thr	
165 170 175	
cgc atc gag gtg cac tac ctc ctc gcc gga ccc ggc ggc gtc cgt cac	576
Arg Ile Glu Val His Tyr Leu Leu Ala Gly Pro Gly Gly Val Arg His	
180 185 190	

ctg acc gag gac cac acc atc acc ctg ttc ccg cgc gcc gac tac gag 624
Leu Thr Glu Asp His Thr Ile Thr Leu Phe Pro Arg Ala Asp Tyr Glu
195 200 205

gcg gcc ttc gag cgc gcc ggc tgc gac gtg gtc tac cag gaa ggc ggc 672
Ala Ala Phe Glu Arg Ala Gly Cys Asp Val Val Tyr Gln Glu Gly Gly
210 215 220

ccg tcc ggt cgc ggg ctg ttc atc ggc acc cgc cgc tga 711
 Pro Ser Gly Arg Gly Leu Phe Ile Gly Thr Arg Arg
 225 230 235

```
<210> 4
<211> 236
<212> PRT
<213> Streptomyces ambofaciens
```

<400> 4

Met Tyr Ala Asp Asp Ile Ala Ala Val Tyr Asp Leu Val His Glu Gly
1 5 10 15

Lys Gly Lys Asp Tyr Arg Gln Glu Ala Glu Glu Ile Ala Ala Leu Val
20 25 30

Arg Val His Arg Pro Gly Ala Arg Thr Leu Leu Asp Val Ala Cys Gly
35 40 45

Thr Gly Gln His Leu His His Leu Asp Gly Leu Phe Asp His Val Glu
50 55 60

Gly Leu Glu Leu Ser Ala Asp Met Leu Ala Leu Ala Thr Gly Arg Asn
65 70 75 80

Pro Gly Val Thr Phe His Gln Gly Asp Met Arg Ser Phe Ser Leu Gly
85 90 95

Arg Arg Phe Asp Ala Val Thr Cys Met Phe Ser Ser Ile Gly His Leu
100 105 110

Arg Thr Thr Asp Glu Leu Asp Ser Thr Leu Arg Ala Phe Thr Asp His
 115 120 125

Leu Glu Pro Ser Gly Val Ile Val Val Glu Pro Trp Trp Phe Pro Glu
 130 135 140

Ser Phe Thr Pro Gly Tyr Val Gly Ala Ser Ile Thr Glu Ala Gly Glu
 145 150 155 160

Arg Thr Val Cys Arg Val Ser His Ser Val Arg Glu Gly Asn Ala Thr
 165 170 175

Arg Ile Glu Val His Tyr Leu Leu Ala Gly Pro Gly Gly Val Arg His
 180 185 190

Leu Thr Glu Asp His Thr Ile Thr Leu Phe Pro Arg Ala Asp Tyr Glu
 195 200 205

Ala Ala Phe Glu Arg Ala Gly Cys Asp Val Val Tyr Gln Glu Gly Gly
 210 215 220

Pro Ser Gly Arg Gly Leu Phe Ile Gly Thr Arg Arg
 225 230 235

<210> 5
 <211> 1272
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1272)

<400> 5
 atg ccc acc ctt gcc acg gaa acg gcc ccc gcg agc acg agc acg agc
 Met Pro Thr Leu Ala Thr Glu Thr Ala Pro Ala Ser Thr Ser Thr Ser

1	5	10	15	
gcg gac acg agc acg ggc gtc cgt gcg ctc ggc cgt cgg ctc cag ctg				96
Ala Asp Thr Ser Thr Gly Val Arg Ala Leu Gly Arg Arg Leu Gln Leu				
	20	25	30	
acc cgg gcc gca cac tgg tgc gcc ggc aac cag ggc gac ccg tac gcg				144
Thr Arg Ala Ala His Trp Cys Ala Gly Asn Gln Gly Asp Pro Tyr Ala				
	35	40	45	
ctg atc ctg cgc gcc gtc gcc gac ccc gag ccg ttc gaa cgg gag atc				192
Leu Ile Leu Arg Ala Val Ala Asp Pro Glu Pro Phe Glu Arg Glu Ile				
	50	55	60	
cgg gcc cgc gga ccg tgg ttc cgc agc gaa cag ctg gac gcc tgg gtg				240
Arg Ala Arg Gly Pro Trp Phe Arg Ser Glu Gln Leu Asp Ala Trp Val				
65	70	75	80	
acc gcg gac ccc gag gtg gcg gcg gcc gtc ctg gcc gac ccg cgc ttc				288
Thr Ala Asp Pro Glu Val Ala Ala Ala Val Leu Ala Asp Pro Arg Phe				
	85	90	95	
ggc acg ctg gac cgg gcc gga cgc cgc ccg gac gag gaa ctg ctg ccc				336
Gly Thr Leu Asp Arg Ala Gly Arg Arg Pro Asp Glu Glu Leu Leu Pro				
	100	105	110	
ctc gcc gag gcg ttc ccc cac cac gaa cgc gcg gag ctc gta cgc ctg				384
Leu Ala Glu Ala Phe Pro His His Glu Arg Ala Glu Leu Val Arg Leu				
	115	120	125	
cgg gcg ctg gcc gcc ccg gtg ctc agc cgg tac gcc ccg gcc cag gcg				432
Arg Ala Leu Ala Ala Pro Val Leu Ser Arg Tyr Ala Pro Ala Gln Ala				
130	135	140		
ccc tgc gcg gcg cgc acc acc gcc cgc aga gtg ctc ggc cgc ctg ctg				480
Pro Cys Ala Ala Arg Thr Thr Ala Arg Arg Val Leu Gly Arg Leu Leu				
145	150	155	160	
ccc acc ggt gac gcc ggg ttc gac ctt gtc ggc gag gtc gcc cgg ccc				528
Pro Thr Gly Asp Ala Gly Phe Asp Leu Val Gly Glu Val Ala Arg Pro				
	165	170	175	
tac gcc gtc gag ctg atg ctc agg ctc ctc gga gtg ccg ggc cgc gac				576
Tyr Ala Val Glu Leu Met Leu Arg Leu Leu Gly Val Pro Gly Arg Asp				
	180	185	190	
cgc gcc acc gcc gcg cgg gca ctc gcc gcc tgc ggc ccc cag ctc gac				624

Arg	Ala	Thr	Ala	Ala	Arg	Ala	Leu	Ala	Ala	Cys	Gly	Pro	Gln	Leu	Asp	
	195						200					205				
gcc	cgg	atg	gcc	ccg	caa	ctg	ctg	acc	gtg	gcc	cgg	gag	tcc	gcc	gac	672
Ala	Arg	Met	Ala	Pro	Gln	Leu	Leu	Thr	Val	Ala	Arg	Glu	Ser	Ala	Asp	
	210					215					220					
gcc	gtc	cgc	aca	ctg	gcc	gac	ctg	gtc	ccc	gag	ctc	gtc	gcg	gag	aag	720
Ala	Val	Arg	Thr	Leu	Ala	Asp	Leu	Val	Pro	Glu	Leu	Val	Ala	Glu	Lys	
	225				230				235					240		
tcc	cgg	ggc	ctc	ggg	aac	gcc	gag	ccc	cgg	ccc	gac	gac	gtg	ctc	gcc	768
Ser	Arg	Gly	Leu	Gly	Asn	Ala	Glu	Pro	Arg	Pro	Asp	Asp	Val	Leu	Ala	
			245				250						255			
ctc	ctc	ctg	cac	gac	ggc	gtc	gcc	ccc	ggc	gac	gtc	gag	cgc	atc	gcg	816
Leu	Leu	Leu	His	Asp	Gly	Val	Ala	Pro	Gly	Asp	Val	Glu	Arg	Ile	Ala	
		260					265					270				
ctg	ctc	ctc	gcg	gtc	ggc	gca	ccc	gaa	ccc	gtc	gtc	acc	gcc	gtc	gcg	864
Leu	Leu	Leu	Ala	Val	Gly	Ala	Pro	Glu	Pro	Val	Val	Thr	Ala	Val	Ala	
		275					280					285				
cac	acg	gtc	cac	cgg	ctg	ctc	ggc	cgg	ccg	ggg	gag	tgg	gag	agg	gcc	912
His	Thr	Val	His	Arg	Leu	Leu	Gly	Arg	Pro	Gly	Glu	Trp	Glu	Arg	Ala	
	290					295				300						
cgc	cgg	acg	ccg	gcc	gcg	gcg	aac	gcc	gtc	gac	cag	gtg	ctg	cgc	gag	960
Arg	Arg	Thr	Pro	Ala	Ala	Ala	Asn	Ala	Val	Asp	Gln	Val	Leu	Arg	Glu	
	305				310					315				320		
cgc	ccc	ccg	gcc	cgg	ctg	gag	aac	cgg	gtc	gcg	cac	acc	ggc	ctc	gaa	1008
Arg	Pro	Pro	Ala	Arg	Leu	Glu	Asn	Arg	Val	Ala	His	Thr	Gly	Leu	Glu	
			325						330				335			
ctc	ggc	ggc	cgc	cgg	atc	acc	gcc	gac	gag	cac	gtc	gtg	gtg	ctg	gcc	1056
Leu	Gly	Gly	Arg	Arg	Ile	Thr	Ala	Asp	Glu	His	Val	Val	Val	Val	Ala	
		340						345					350			
gcc	gcc	gga	cgg	gag	atc	ccc	ggg	ccg	gag	ccg	ctc	ggg	ggc	gcc	gac	1104
Ala	Ala	Gly	Arg	Glu	Ile	Pro	Gly	Pro	Glu	Pro	Leu	Gly	Gly	Ala	Asp	
		355					360				365					
gga	ccg	cac	ctg	gcg	ctc	gcc	ctc	ccg	ctg	atc	cgc	ctg	gcc	gcc	acc	1152
Gly	Pro	His	Leu	Ala	Leu	Ala	Leu	Pro	Leu	Ile	Arg	Leu	Ala	Ala	Thr	
	370					375					380					

acc gcg gtc cag gtc acg gcc ggc cgc ctg ccc ggc ctg cgg gcc gag 1200
 Thr Ala Val Gln Val Thr Ala Gly Arg Leu Pro Gly Leu Arg Ala Glu
 385 390 395 400

gga ccg ccc ctg acc cgg ccg cgg tca ccg gtc ctg ggc gcc tgc gcc 1248
 Gly Pro Pro Leu Thr Arg Pro Arg Ser Pro Val Leu Gly Ala Cys Ala
 405 410 415

cgc ctc cgg gtc cac ccg gga tga 1272
 Arg Leu Arg Val His Pro Gly
 420

<210> 6
 <211> 423
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 6

Met Pro Thr Leu Ala Thr Glu Thr Ala Pro Ala Ser Thr Ser Thr Ser
 1 5 10 15

Ala Asp Thr Ser Thr Gly Val Arg Ala Leu Gly Arg Arg Leu Gln Leu
 20 25 30

Thr Arg Ala Ala His Trp Cys Ala Gly Asn Gln Gly Asp Pro Tyr Ala
 35 40 45

Leu Ile Leu Arg Ala Val Ala Asp Pro Glu Pro Phe Glu Arg Glu Ile
 50 55 60

Arg Ala Arg Gly Pro Trp Phe Arg Ser Glu Gln Leu Asp Ala Trp Val
 65 70 75 80

Thr Ala Asp Pro Glu Val Ala Ala Ala Val Leu Ala Asp Pro Arg Phe
 85 90 95

Gly Thr Leu Asp Arg Ala Gly Arg Arg Pro Asp Glu Glu Leu Leu Pro
 100 105 110

Leu Ala Glu Ala Phe Pro His His Glu Arg Ala Glu Leu Val Arg Leu
 115 120 125

Arg Ala Leu Ala Ala Pro Val Leu Ser Arg Tyr Ala Pro Ala Gln Ala
 130 135 140

Pro Cys Ala Ala Arg Thr Thr Ala Arg Arg Val Leu Gly Arg Leu Leu
 145 150 155 160

Pro Thr Gly Asp Ala Gly Phe Asp Leu Val Gly Glu Val Ala Arg Pro
 165 170 175

Tyr Ala Val Glu Leu Met Leu Arg Leu Leu Gly Val Pro Gly Arg Asp
 180 185 190

Arg Ala Thr Ala Ala Arg Ala Leu Ala Ala Cys Gly Pro Gln Leu Asp
 195 200 205

Ala Arg Met Ala Pro Gln Leu Leu Thr Val Ala Arg Glu Ser Ala Asp
 210 215 220

Ala Val Arg Thr Leu Ala Asp Leu Val Pro Glu Leu Val Ala Glu Lys
 225 230 235 240

Ser Arg Gly Leu Gly Asn Ala Glu Pro Arg Pro Asp Asp Val Leu Ala
 245 250 255

Leu Leu Leu His Asp Gly Val Ala Pro Gly Asp Val Glu Arg Ile Ala
 260 265 270

Leu Leu Leu Ala Val Gly Ala Pro Glu Pro Val Val Thr Ala Val Ala
 275 280 285

His Thr Val His Arg Leu Leu Gly Arg Pro Gly Glu Trp Glu Arg Ala
 290 295 300

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

Arg Pro Pro Ala Arg Leu Glu Asn Arg Val Ala His Thr Gly Leu Glu
 325 330 335

Leu Gly Gly Arg Arg Ile Thr Ala Asp Glu His Val Val Val Leu Ala
 340 345 350

Ala Ala Gly Arg Glu Ile Pro Gly Pro Glu Pro Leu Gly Gly Ala Asp
 355 360 365

Gly Pro His Leu Ala Leu Ala Leu Pro Leu Ile Arg Leu Ala Ala Thr
 370 375 380

Thr Ala Val Gln Val Thr Ala Gly Arg Leu Pro Gly Leu Arg Ala Glu
 385 390 395 400

Gly Pro Pro Leu Thr Arg Pro Arg Ser Pro Val Leu Gly Ala Cys Ala
 405 410 415

Arg Leu Arg Val His Pro Gly
 420

<210> 7
 <211> 1266
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1266)

<400> 7
 gtg cgc gtc ctg ctg aca tcc ctc gcc cac aac acc cac tac tac agt
 Val Arg Val Leu Leu Thr Ser Leu Ala His Asn Thr His Tyr Tyr Ser

1	5	10	15	
ctg gtg ccc ctc gcc tgg gcg ctg cgc gcc gcc ggg cac gag gta cgg				96
Leu Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg				
	20	25	30	
gtg gcg agc ccg ccc tcc ctc acc gac gtc atc acc tcc acc ggc ctg				144
Val Ala Ser Pro Pro Ser Leu Thr Asp Val Ile Thr Ser Thr Gly Leu				
	35	40	45	
acc gcc gta ccg gtg ggc gac gac cga ccg gcc gcg gag ctg ctc gcc				192
Thr Ala Val Pro Val Gly Asp Asp Arg Pro Ala Ala Glu Leu Leu Ala				
	50	55	60	
gag atg ggc agc gac ctc gtc ccc tac cag cgg ggc ttc gag ttc ggt				240
Glu Met Gly Ser Asp Leu Val Pro Tyr Gln Arg Gly Phe Glu Phe Gly				
65	70	75	80	
gag gtg gag agc gag gag gag acc acc tgg gag tac ctg ctc ggc cag				288
Glu Val Glu Ser Glu Glu Glu Thr Thr Trp Glu Tyr Leu Leu Gly Gln				
	85	90	95	
cag agc atg atg gcc gcc ctg tgc ttc gcc ccg ttc aac ggc gcc gcc				336
Gln Ser Met Met Ala Ala Leu Cys Phe Ala Pro Phe Asn Gly Ala Ala				
	100	105	110	
acg atg gac gag atc gtc gac ttc gcc cgt ggc tgg cgg ccc gac ctg				384
Thr Met Asp Glu Ile Val Asp Phe Ala Arg Gly Trp Arg Pro Asp Leu				
	115	120	125	
gtc gtg tgg gaa ccc tgg acc tac gcg ggg ccg gtc gcg gcc cgc gcc				432
Val Val Trp Glu Pro Trp Thr Tyr Ala Gly Pro Val Ala Ala Arg Ala				
	130	135	140	
tgc ggt gcc gcc cac gcg cgc atc ctg tgg ggc ccc gac gcc atc gga				480
Cys Gly Ala Ala His Ala Arg Ile Leu Trp Gly Pro Asp Ala Ile Gly				
145	150	155	160	
cgc tcc cgc ctg cgc ttc ctc gcc gcc ctc gac ggg atg ccg gag gaa				528
Arg Ser Arg Leu Arg Phe Leu Ala Ala Leu Asp Gly Met Pro Glu Glu				
	165	170	175	
ctg cgc gag gac ccc atc gcc gaa tgg ctc ggc tgg acg ctg gac cgg				576
Leu Arg Glu Asp Pro Ile Ala Glu Trp Leu Gly Trp Thr Leu Asp Arg				
	180	185	190	
tac ggg ctc gcc ttc gac gaa cgc gac gtg ctc ggc cac tgg gtg atc				624

Tyr Gly Leu Ala Phe Asp Glu Arg Asp Val Leu Gly His Trp Val Ile	
195 200 205	
gac ccg ggc ccg cgc agc acc cgc ctc gac ctg ggc cag acc acg gtg	672
Asp Pro Gly Pro Arg Ser Thr Arg Leu Asp Leu Gly Gln Thr Thr Val	
210 215 220	
ccc atg tgc tac gtg ccc tac aac ggg cgc tcc gtg atc gaa ccc tgg	720
Pro Met Cys Tyr Val Pro Tyr Asn Gly Arg Ser Val Ile Glu Pro Trp	
225 230 235 240	
ctc ggc aag aag ccc gac cgc ccc cgc gtc tgc ctc acc ctc ggc gtg	768
Leu Gly Lys Lys Pro Asp Arg Pro Arg Val Cys Leu Thr Leu Gly Val	
245 250 255	
tcc gcc cgc gag acc tac ggc cgc gac gcg gtc tcc tac tcc gaa ctg	816
Ser Ala Arg Glu Thr Tyr Gly Arg Asp Ala Val Ser Tyr Ser Glu Leu	
260 265 270	
ctg gag gcg ctc ggc cgg ctg gac atc gag gtg gtg gcc acc ctg gac	864
Leu Glu Ala Leu Gly Arg Leu Asp Ile Glu Val Val Ala Thr Leu Asp	
275 280 285	
acc tcc cag ctc gac cgc ctg ggc acc ctc ccc gac aac gtc gtg ccc	912
Thr Ser Gln Leu Asp Arg Leu Gly Thr Leu Pro Asp Asn Val Val Pro	
290 295 300	
gtg gac ttc gta ccg ctc gac gcg ctg ctg ccg agc tgc gcc gcg atc	960
Val Asp Phe Val Pro Leu Asp Ala Leu Leu Pro Ser Cys Ala Ala Ile	
305 310 315 320	
atc cac cac ggc gga gcg ggc acc tgg tcc acc gcc ctg cgc cac gga	1008
Ile His His Gly Gly Ala Gly Thr Trp Ser Thr Ala Leu Arg His Gly	
325 330 335	
gtg ccg cag atc ctg ctg ccc tcc ctg tgg gac gcg ccg ctc aag gcg	1056
Val Pro Gln Ile Leu Leu Pro Ser Leu Trp Asp Ala Pro Leu Lys Ala	
340 345 350	
cag caa ctc cag cgc ctg tcc gcc ggg ttc gac ctg ccc gcg gcg acg	1104
Gln Gln Leu Gln Arg Leu Ser Ala Gly Phe Asp Leu Pro Ala Ala Thr	
355 360 365	
ctc acg gcg cgc ggt ctg gcc gac gcg gtg cac acg gcc gtg cac gac	1152
Leu Thr Ala Arg Gly Leu Ala Asp Ala Val His Thr Ala Val His Asp	
370 375 380	

ccc gcg atc agg gcg ggc gcg cgg cgg ctg agc gag gag atg ctc gcc 1200
 Pro Ala Ile Arg Ala Gly Ala Arg Arg Leu Ser Glu Glu Met Leu Ala
 385 390 395 400

gac ccc acg ccc gcc ggg atc atc ccc acc ctg gag cgc ctc acc gcc 1248
 Asp Pro Thr Pro Ala Gly Ile Ile Pro Thr Leu Glu Arg Leu Thr Ala
 405 410 415

ctg cac cgg gcc acc tga 1266
 Leu His Arg Ala Thr
 420

<210> 8
 <211> 421
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 8

Val Arg Val Leu Leu Thr Ser Leu Ala His Asn Thr His Tyr Tyr Ser
 1 5 10 15

Leu Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg
 20 25 30

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

Thr Ala Val Pro Val Gly Asp Asp Arg Pro Ala Ala Glu Leu Leu Ala
 50 55 60

Glu Met Gly Ser Asp Leu Val Pro Tyr Gln Arg Gly Phe Glu Phe Gly
 65 70 75 80

Glu Val Glu Ser Glu Glu Glu Thr Thr Trp Glu Tyr Leu Leu Gly Gln
 85 90 95

Gln Ser Met Met Ala Ala Leu Cys Phe Ala Pro Phe Asn Gly Ala Ala
 100 105 110

Thr Met Asp Glu Ile Val Asp Phe Ala Arg Gly Trp Arg Pro Asp Leu
 115 120 125

Val Val Trp Glu Pro Trp Thr Tyr Ala Gly Pro Val Ala Ala Arg Ala
 130 135 140

Cys Gly Ala Ala His Ala Arg Ile Leu Trp Gly Pro Asp Ala Ile Gly
 145 150 155 160

Arg Ser Arg Leu Arg Phe Leu Ala Ala Leu Asp Gly Met Pro Glu Glu
 165 170 175

Leu Arg Glu Asp Pro Ile Ala Glu Trp Leu Gly Trp Thr Leu Asp Arg
 180 185 190

Tyr Gly Leu Ala Phe Asp Glu Arg Asp Val Leu Gly His Trp Val Ile
 195 200 205

Asp Pro Gly Pro Arg Ser Thr Arg Leu Asp Leu Gly Gln Thr Thr Val
 210 215 220

Pro Met Cys Tyr Val Pro Tyr Asn Gly Arg Ser Val Ile Glu Pro Trp
 225 230 235 240

Leu Gly Lys Lys Pro Asp Arg Pro Arg Val Cys Leu Thr Leu Gly Val
 245 250 255

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

Leu Glu Ala Leu Gly Arg Leu Asp Ile Glu Val Val Ala Thr Leu Asp
 275 280 285

Thr Ser Gln Leu Asp Arg Leu Gly Thr Leu Pro Asp Asn Val Val Pro
 290 295 300

Val Asp Phe Val Pro Leu Asp Ala Leu Leu Pro Ser Cys Ala Ala Ile
 305 310 315 320

Ile His His Gly Gly Ala Gly Thr Trp Ser Thr Ala Leu Arg His Gly
 325 330 335

Val Pro Gln Ile Leu Leu Pro Ser Leu Trp Asp Ala Pro Leu Lys Ala
 340 345 350

Gln Gln Leu Gln Arg Leu Ser Ala Gly Phe Asp Leu Pro Ala Ala Thr
 355 360 365

Leu Thr Ala Arg Gly Leu Ala Asp Ala Val His Thr Ala Val His Asp
 370 375 380

Pro Ala Ile Arg Ala Gly Ala Arg Arg Leu Ser Glu Glu Met Leu Ala
 385 390 395 400

Asp Pro Thr Pro Ala Gly Ile Ile Pro Thr Leu Glu Arg Leu Thr Ala
 405 410 415

Leu His Arg Ala Thr
 420

<210> 9
 <211> 1350
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1350)

<400> 9
 atg ccc gaa agc cat gcg cag agc gcg ctg ctc gcc gcg atc acc gcc
 Met Pro Glu Ser His Ala Gln Ser Ala Leu Leu Ala Ala Ile Thr Ala

1	5	10	15	
gcg gac cgc acc ccc gag gac atc gcg gcg ctg ccc ctg ccc gag tcc				96
Ala Asp Arg Thr Pro Glu Asp Ile Ala Ala Leu Pro Leu Pro Glu Ser				
	20	25	30	
ttc cgc gcc gtg acc gtc cgc aaa gag gac acc gac atg ttc cgc ggg				144
Phe Arg Ala Val Thr Val Arg Lys Glu Asp Thr Asp Met Phe Arg Gly				
	35	40	45	
atg ccg agc gcg gac aag gac ccg cgc aag tcc ctg cac gtc gac gag				192
Met Pro Ser Ala Asp Lys Asp Pro Arg Lys Ser Leu His Val Asp Glu				
	50	55	60	
gtg ccg ctg ccc gaa gtg ggg ccc ggg gag gcg ctg gtc gcg gtg atg				240
Val Pro Leu Pro Glu Val Gly Pro Gly Glu Ala Leu Val Ala Val Met				
65	70	75	80	
gcc agc tcc gtc aac tac aac acg gtc tgg tcc tcc atc ttc gag ccc				288
Ala Ser Ser Val Asn Tyr Asn Thr Val Trp Ser Ser Ile Phe Glu Pro				
	85	90	95	
ctc ccg acg ttc ggc ttc ctg gag cgc tac ggg cgc acc tcg ccg ctg				336
Leu Pro Thr Phe Gly Phe Leu Glu Arg Tyr Gly Arg Thr Ser Pro Leu				
	100	105	110	
gcc gcg cgc cac gac ctg ccg tac cac atc ctc ggt tcc gac ctg tcc				384
Ala Ala Arg His Asp Leu Pro Tyr His Ile Leu Gly Ser Asp Leu Ser				
	115	120	125	
ggc gtg gtg ctg cgc acg gga ccc ggc gtg aac gtc tgg aaa ccc ggt				432
Gly Val Val Leu Arg Thr Gly Pro Gly Val Asn Val Trp Lys Pro Gly				
	130	135	140	
gac gag gtc gtc gcg cac tgt ctg tcg gtg gag ctg gag agc ccc gac				480
Asp Glu Val Val Ala His Cys Leu Ser Val Glu Leu Glu Ser Pro Asp				
145	150	155	160	
ggg cac gac gac acc atg ctc gac ccg gcc cag cgc atc tgg ggc ttc				528
Gly His Asp Asp Thr Met Leu Asp Pro Ala Gln Arg Ile Trp Gly Phe				
	165	170	175	
gag acc aac ttc ggg ggc ctg gcc gag gtg gcc ctc gtc aag tcc aac				576
Glu Thr Asn Phe Gly Gly Leu Ala Glu Val Ala Leu Val Lys Ser Asn				
	180	185	190	
cag ctg atg ccc aag gcg gcc cat ctg acc tgg gag gag gcc gcc gcg				624

Gln Leu Met Pro Lys Ala Ala His Leu Thr Trp Glu Glu Ala Ala Ala	
195	200 205
ccc ggt ctg gtc aac tcc acc gcc tac cgc cag ctg gtc tcc cgc aac	672
Pro Gly Leu Val Asn Ser Thr Ala Tyr Arg Gln Leu Val Ser Arg Asn	
210	215 220
ggc gcc ggc atg aag cag ggc gac aac gtg ctg atc tgg ggc gcc agc	720
Gly Ala Gly Met Lys Gln Gly Asp Asn Val Leu Ile Trp Gly Ala Ser	
225	230 235 240
ggc ggc ctg ggc tcg tac gcc acc cag ctg gcc ctc gcc ggc ggg gcc	768
Gly Gly Leu Gly Ser Tyr Ala Thr Gln Leu Ala Leu Ala Gly Gly Ala	
	245 250 255
cac ccc gtc tgc gtg gtc tcc agc gcg cgc aag gcc gag gtg tgc cgg	816
His Pro Val Cys Val Val Ser Ser Ala Arg Lys Ala Glu Val Cys Arg	
	260 265 270
gcc atg ggc gcg gag gcc atc atc gac cgc tcc gcc gag gac tac cgc	864
Ala Met Gly Ala Glu Ala Ile Ile Asp Arg Ser Ala Glu Asp Tyr Arg	
	275 280 285
ttc tgg agc gac gag gac acc cag gac ccg cgg gag tgg aag cgc ttc	912
Phe Trp Ser Asp Glu Asp Thr Gln Asp Pro Arg Glu Trp Lys Arg Phe	
	290 295 300
ggc aag cgc atc cgc gag gtc acc ggg ggc gag gac gtc gac atc gtc	960
Gly Lys Arg Ile Arg Glu Val Thr Gly Gly Glu Asp Val Asp Ile Val	
305	310 315 320
ttc gag cac ccc ggc cgg gag acc ttc ggg gcg tcc gtc tac gtc acc	1008
Phe Glu His Pro Gly Arg Glu Thr Phe Gly Ala Ser Val Tyr Val Thr	
	325 330 335
cgc agg ggc ggc acg atc gtc acc tgc gcc tcc acc tcg ggc ttc cgc	1056
Arg Arg Gly Gly Thr Ile Val Thr Cys Ala Ser Thr Ser Gly Phe Arg	
	340 345 350
cac gag ttc gac aac cgc tac ctc tgg atg cac ctc aag cgt atc gtc	1104
His Glu Phe Asp Asn Arg Tyr Leu Trp Met His Leu Lys Arg Ile Val	
	355 360 365
ggc acc cac ttc gcc aac tac cgc gag gcg tgg gag gcg aac cgg ctc	1152
Gly Thr His Phe Ala Asn Tyr Arg Glu Ala Trp Glu Ala Asn Arg Leu	
	370 375 380

gtc gcc aag ggg agg atc cac ccc acc ctc tcg cgc acc tac ccg ctg 1200
 Val Ala Lys Gly Arg Ile His Pro Thr Leu Ser Arg Thr Tyr Pro Leu
 385 390 395 400

gcc gac acg gcg gtt gcc gtg cac gac gtg cac ggc aac cag cac cag 1248
 Ala Asp Thr Ala Val Ala Val His Asp Val His Gly Asn Gln His Gln
 405 410 415

ggc aag gtc ggc gtg ctc tgc ctg gcg ccc acg gag ggg ctg ggc gtg 1296
 Gly Lys Val Gly Val Leu Cys Leu Ala Pro Thr Glu Gly Leu Gly Val
 420 425 430

cgc gac gag gag aag cgg gcg cgg cac atc gac gcg atc aac cgg ttc 1344
 Arg Asp Glu Glu Lys Arg Ala Arg His Ile Asp Ala Ile Asn Arg Phe
 435 440 445

cgc tga 1350
 Arg

<210> 10
 <211> 449
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 10

Met Pro Glu Ser His Ala Gln Ser Ala Leu Leu Ala Ala Ile Thr Ala
 1 5 10 15

Ala Asp Arg Thr Pro Glu Asp Ile Ala Ala Leu Pro Leu Pro Glu Ser
 20 25 30

Phe Arg Ala Val Thr Val Arg Lys Glu Asp Thr Asp Met Phe Arg Gly
 35 40 45

Met Pro Ser Ala Asp Lys Asp Pro Arg Lys Ser Leu His Val Asp Glu
 50 55 60

Val Pro Leu Pro Glu Val Gly Pro Gly Glu Ala Leu Val Ala Val Met
 65 70 75 80

Ala Ser Ser Val Asn Tyr Asn Thr Val Trp Ser Ser Ile Phe Glu Pro
 85 90 95

Leu Pro Thr Phe Gly Phe Leu Glu Arg Tyr Gly Arg Thr Ser Pro Leu
 100 105 110

Ala Ala Arg His Asp Leu Pro Tyr His Ile Leu Gly Ser Asp Leu Ser
 115 120 125

Gly Val Val Leu Arg Thr Gly Pro Gly Val Asn Val Trp Lys Pro Gly
 130 135 140

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

Gly His Asp Asp Thr Met Leu Asp Pro Ala Gln Arg Ile Trp Gly Phe
 165 170 175

Glu Thr Asn Phe Gly Gly Leu Ala Glu Val Ala Leu Val Lys Ser Asn
 180 185 190

Gln Leu Met Pro Lys Ala Ala His Leu Thr Trp Glu Glu Ala Ala Ala
 195 200 205

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

Gly Ala Gly Met Lys Gln Gly Asp Asn Val Leu Ile Trp Gly Ala Ser
 225 230 235 240

Gly Gly Leu Gly Ser Tyr Ala Thr Gln Leu Ala Leu Ala Gly Gly Ala
 245 250 255

His Pro Val Cys Val Val Ser Ser Ala Arg Lys Ala Glu Val Cys Arg
 260 265 270

Ala Met Gly Ala Glu Ala Ile Ile Asp Arg Ser Ala Glu Asp Tyr Arg
 275 280 285

Phe Trp Ser Asp Glu Asp Thr Gln Asp Pro Arg Glu Trp Lys Arg Phe
 290 295 300

Gly Lys Arg Ile Arg Glu Val Thr Gly Gly Glu Asp Val Asp Ile Val
 305 310 315 320

Phe Glu His Pro Gly Arg Glu Thr Phe Gly Ala Ser Val Tyr Val Thr
 325 330 335

Arg Arg Gly Gly Thr Ile Val Thr Cys Ala Ser Thr Ser Gly Phe Arg
 340 345 350

His Glu Phe Asp Asn Arg Tyr Leu Trp Met His Leu Lys Arg Ile Val
 355 360 365

Gly Thr His Phe Ala Asn Tyr Arg Glu Ala Trp Glu Ala Asn Arg Leu
 370 375 380

Val Ala Lys Gly Arg Ile His Pro Thr Leu Ser Arg Thr Tyr Pro Leu
 385 390 395 400

Ala Asp Thr Ala Val Ala Val His Asp Val His Gly Asn Gln His Gln
 405 410 415

Gly Lys Val Gly Val Leu Cys Leu Ala Pro Thr Glu Gly Leu Gly Val
 420 425 430

Arg Asp Glu Glu Lys Arg Ala Arg His Ile Asp Ala Ile Asn Arg Phe
 435 440 445

Arg

<210> 11
 <211> 675
 <212> DNA
 <213> *Streptomyces ambofaciens*

<220>
 <221> CDS
 <222> (1)..(675)

<400> 11
 atg atg agc gtg gca gac cag acg gct ctc agc ccc gcg ctg ctg gag 48
 Met Met Ser Val Ala Asp Gln Thr Ala Leu Ser Pro Ala Leu Leu Glu
 1 5 10 15
 tac gcc cgg agc gtc gcg ctg cgg gac gac ggc ctg ctg cgc gaa ctg 96
 Tyr Ala Arg Ser Val Ala Leu Arg Asp Asp Gly Leu Leu Arg Glu Leu
 20 25 30
 cac gag gtg acc gcc ggg ctc ccc ggc ggc cgg gcc atg cag atc atg 144
 His Glu Val Thr Ala Gly Leu Pro Gly Gly Arg Ala Met Gln Ile Met
 35 40 45
 ccc gag gag gcg cag ttc ctc gcc ctg ctg atc cgg ctc gtc ggt gcc 192
 Pro Glu Glu Ala Gln Phe Leu Ala Leu Leu Ile Arg Leu Val Gly Ala
 50 55 60
 cgg cgg gtg ctg gag atc ggc acc ttc acg ggg tac agc acg ctg tgc 240
 Arg Arg Val Leu Glu Ile Gly Thr Phe Thr Gly Tyr Ser Thr Leu Cys
 65 70 75 80
 atg gcg cgg gca ctg ccc gcc gac ggc acc gtc gtc acc tgc gac atc 288
 Met Ala Arg Ala Leu Pro Ala Asp Gly Thr Val Val Thr Cys Asp Ile
 85 90 95
 agc gac agg tgg ccc ggc gtc ggc gca ccg tac tgg cgc cgg gcc ggg 336
 Ser Asp Arg Trp Pro Gly Val Gly Ala Pro Tyr Trp Arg Arg Ala Gly
 100 105 110
 gtg gag tcc cgg atc gac ctg cgc gtc ggc gac gcc gtc cgg acc ctc 384
 Val Glu Ser Arg Ile Asp Leu Arg Val Gly Asp Ala Val Arg Thr Leu
 115 120 125
 gcc gag ctc cgc gag cac gag ggg gac ggc tcg ttc gac ctg gtc ttc 432

Ala Glu Leu Arg Glu His Glu Gly Asp Gly Ser Phe Asp Leu Val Phe
130 135 140

gtc gac gcc gac aag acc ggg tac ccg cac tac tac gag cag gcg ctg 480
Val Asp Ala Asp Lys Thr Gly Tyr Pro His Tyr Tyr Glu Gln Ala Leu
145 150 155 160

gcc ctg gta cgc ccc ggc gga ctg gtg gcg gtc gac aac acc ctg ttc 528
Ala Leu Val Arg Pro Gly Gly Leu Val Ala Val Asp Asn Thr Leu Phe
165 170 175

ttc ggc cgg gtg gcc gac ccg gcc gtc gag gac gcc gac acc gtc gcc 576
Phe Gly Arg Val Ala Asp Pro Ala Val Glu Asp Ala Asp Thr Val Ala
180 185 190

gtg cgc gcg ctc aac gag ctg ctg cgc gac gac gaa cgc gtg gac atc 624
Val Arg Ala Leu Asn Glu Leu Leu Arg Asp Asp Glu Arg Val Asp Ile
195 200 205

gcc ctg ctg acg gtc gcc gac ggg atc act ctg gcc cgc cgg cgg gag 672
Ala Leu Leu Thr Val Ala Asp Gly Ile Thr Leu Ala Arg Arg Arg Glu
210 215 220

tga 675

<210> 12
<211> 224
<212> PRT
<213> Streptomyces ambofaciens

<400> 12

Met Met Ser Val Ala Asp Gln Thr Ala Leu Ser Pro Ala Leu Leu Glu
1 5 10 15

Tyr Ala Arg Ser Val Ala Leu Arg Asp Asp Gly Leu Leu Arg Glu Leu
20 25 30

His Glu Val Thr Ala Gly Leu Pro Gly Gly Arg Ala Met Gln Ile Met
35 40 45

Pro Glu Glu Ala Gln Phe Leu Ala Leu Leu Ile Arg Leu Val Gly Ala
50 55 60

Arg Arg Val Leu Glu Ile Gly Thr Phe Thr Gly Tyr Ser Thr Leu Cys
65 70 75 80

Met Ala Arg Ala Leu Pro Ala Asp Gly Thr Val Val Thr Cys Asp Ile
85 90 95

Ser Asp Arg Trp Pro Gly Val Gly Ala Pro Tyr Trp Arg Arg Ala Gly
100 105 110

Val Glu Ser Arg Ile Asp Leu Arg Val Gly Asp Ala Val Arg Thr Leu
115 120 125

Ala Glu Leu Arg Glu His Glu Gly Asp Gly Ser Phe Asp Leu Val Phe
130 135 140

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

Ala Leu Val Arg Pro Gly Gly Leu Val Ala Val Asp Asn Thr Leu Phe
165 170 175

Phe Gly Arg Val Ala Asp Pro Ala Val Glu Asp Ala Asp Thr Val Ala
180 185 190

Val Arg Ala Leu Asn Glu Leu Leu Arg Asp Asp Glu Arg Val Asp Ile
195 200 205

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

<210> 13

<211> 1245

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1)..(1245)

<400> 13

gtg	ctc	cag	cgc	gtc	gat	ctg	tgc	tca	ctc	acc	ggc	ctc	cgc	tgg	tat	48
Val	Leu	Gln	Arg	Val	Asp	Leu	Ser	Ser	Leu	Thr	Gly	Leu	Arg	Trp	Tyr	
1				5					10					15		

gcg	gcg	ctg	acg	gta	ttc	gcc	tgt	cac	atc	gcc	cag	cag	ggc	ttc	ttc	96
Ala	Ala	Leu	Thr	Val	Phe	Ala	Cys	His	Ile	Ala	Gln	Gln	Gly	Phe	Phe	
			20					25					30			

gcc	gac	cag	cag	gtg	ggc	agc	gca	ctg	ctg	cac	atc	acc	ccg	ctc	ggt	144
Ala	Asp	Gln	Gln	Val	Gly	Ser	Ala	Leu	Leu	His	Ile	Thr	Pro	Leu	Gly	
		35					40					45				

tcc	atg	gcg	gtc	tgc	atc	ttc	ttc	ata	ctg	agt	gga	ttc	gtc	ctc	gcc	192
Ser	Met	Ala	Val	Ser	Ile	Phe	Phe	Ile	Leu	Ser	Gly	Phe	Val	Leu	Ala	
	50					55					60					

tgg	tgc	gcc	cgc	gac	gag	gac	tcc	gtg	ccg	act	ttc	tgg	cgg	cgc	cgc	240
Trp	Ser	Ala	Arg	Asp	Glu	Asp	Ser	Val	Pro	Thr	Phe	Trp	Arg	Arg	Arg	
65					70					75				80		

atc	gcg	aag	atc	tat	ccg	ctg	cat	ctc	gcg	acg	ttc	ggc	atc	gcg	gct	288
Ile	Ala	Lys	Ile	Tyr	Pro	Leu	His	Leu	Ala	Thr	Phe	Gly	Ile	Ala	Ala	
			85						90				95			

ctc	atc	att	ttc	tcc	ctg	tgc	gag	ccg	gta	ctt	ccc	ggc	ggt	tcc	gta	336
Leu	Ile	Ile	Phe	Ser	Leu	Ser	Glu	Pro	Val	Leu	Pro	Gly	Gly	Ser	Val	
			100					105					110			

tgg	gac	ggg	ctg	gtg	ccc	aat	gtt	ctg	ctc	gtg	cag	tcc	tgg	ctt	ccc	384
Trp	Asp	Gly	Leu	Val	Pro	Asn	Val	Leu	Leu	Val	Gln	Ser	Trp	Leu	Pro	
		115					120					125				

gac	gcg	acc	ctc	acg	gcc	agt	ttc	aac	acg	ccc	agc	tgg	tgc	ctc	tcc	432
Asp	Ala	Thr	Leu	Thr	Ala	Ser	Phe	Asn	Thr	Pro	Ser	Trp	Ser	Leu	Ser	
	130					135					140					

tgt	gag	atc	gcc	ttc	tat	ctg	tgc	ttc	ccg	ctg	tgg	tac	cgg	ctg	gtg	480
Cys	Glu	Ile	Ala	Phe	Tyr	Leu	Ser	Phe	Pro	Leu	Trp	Tyr	Arg	Leu	Val	
145					150					155				160		

cgc	agg	att	ccc	gca	cgg	cgg	ctg	tgg	tgg	tgc	gcc	gcg	ggg	atc	gcc	528

Arg Arg Ile Pro Ala Arg Arg Leu Trp Trp Cys Ala Ala Gly Ile Ala	
165 170 175	
gtg gcc gtg acg tgt gtg ccc ctg ctg gcg ggc ctg ctc ccg gcg agc	576
Val Ala Val Thr Cys Val Pro Leu Leu Ala Gly Leu Leu Pro Ala Ser	
180 185 190	
gag gag gtg gcc ccc ggg atg tcg ctc aac gag gtc tgg ttc gcg tac	624
Glu Glu Val Ala Pro Gly Met Ser Leu Asn Glu Val Trp Phe Ala Tyr	
195 200 205	
tgg ctt ccg ccg gtg cgc atg ctg gag ttc gtc ctc ggc atc gtg atg	672
Trp Leu Pro Pro Val Arg Met Leu Glu Phe Val Leu Gly Ile Val Met	
210 215 220	
gcg ctg atc ctg cgc gcg ggg atc tgg aag ggc ccc ggt ccg gcg gtc	720
Ala Leu Ile Leu Arg Ala Gly Ile Trp Lys Gly Pro Gly Pro Ala Val	
225 230 235 240	
tgc acg gcg ctc ctc gcc gcg agt tac ggc ctc acc cag atg gtg ccc	768
Cys Thr Ala Leu Leu Ala Ala Ser Tyr Gly Leu Thr Gln Met Val Pro	
245 250 255	
ccg atc ttc acc ctc gtc gcc tgc tcc gtc gta ccg gcc gcg ctg ctg	816
Pro Ile Phe Thr Leu Val Ala Cys Ser Val Val Pro Ala Ala Leu Leu	
260 265 270	
atc acg gcg ctg gcc gac gcc gac gtg cac ggc cgg cgc acg ggg ctg	864
Ile Thr Ala Leu Ala Asp Ala Asp Val His Gly Arg Arg Thr Gly Leu	
275 280 285	
cgt tcg gcg acg ctg gtg cgg ctg ggc cag tgg tcc ttc gcc ttc tac	912
Arg Ser Ala Thr Leu Val Arg Leu Gly Gln Trp Ser Phe Ala Phe Tyr	
290 295 300	
ctg gtc cac ttc ctg atc atc cgc tac gga cac cgg ctg atg ggc ggc	960
Leu Val His Phe Leu Ile Ile Arg Tyr Gly His Arg Leu Met Gly Gly	
305 310 315 320	
gat ctg ggc tac gag cgg cag tgg agc acc ccg gcc gcg atc gcg ctg	1008
Asp Leu Gly Tyr Glu Arg Gln Trp Ser Thr Pro Ala Ala Ile Ala Leu	
325 330 335	
tcc ctg ggg atg ctg ggg gtg gcg gtc ctg gcc ggc ggt ctg ctg cac	1056
Ser Leu Gly Met Leu Gly Val Ala Val Leu Ala Gly Gly Leu Leu His	
340 345 350	

acc gtc gtc gaa cag ccc tgc atg cgc ctg ttc ggc agc cgc agg tcc 1104
Thr Val Val Glu Gln Pro Cys Met Arg Leu Phe Gly Ser Arg Arg Ser
355 360 365

gcc tcc cgt ccg aag ccc ggc gcc acc gcg gct ccc cgg aac tca ccc 1152
Ala Ser Arg Pro Lys Pro Gly Ala Thr Ala Ala Pro Arg Asn Ser Pro
370 375 380

gcg gcc gac gcg gcc ggc gtg ccc ctg ctc ccg ggc gta ccc ggg ccc 1200
Ala Ala Asp Ala Ala Gly Val Pro Leu Leu Pro Gly Val Pro Gly Pro
385 390 395 400

gcg cac acc ccc gca gcg acg aac gaa ccc acc ccg aga gga tga 1245
Ala His Thr Pro Ala Ala Thr Asn Glu Pro Thr Pro Arg Gly
405 410

```
<210> 14
<211> 414
<212> PRT
<213> Streptomyces ambofaciens
```

<400> 14

Val	Leu	Gln	Arg	Val	Asp	Leu	Ser	Ser	Leu	Thr	Gly	Leu	Arg	Trp	Tyr
1			5						10					15	

Ala Ala Leu Thr Val Phe Ala Cys His Ile Ala Gln Gln Gly Phe Phe
20 25 30

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

Ser Met Ala Val Ser Ile Phe Phe Ile Leu Ser Gly Phe Val Leu Ala
50 55 60

Trp Ser Ala Arg Asp Glu Asp Ser Val Pro Thr Phe Trp Arg Arg Arg
65 70 75 80

Ile Ala Lys Ile Tyr Pro Leu His Leu Ala Thr Phe Gly Ile Ala Ala
85 90 95

Leu Ile Ile Phe Ser Leu Ser Glu Pro Val Leu Pro Gly Gly Ser Val
 100 105 110

Trp Asp Gly Leu Val Pro Asn Val Leu Leu Val Gln Ser Trp Leu Pro
 115 120 125

Asp Ala Thr Leu Thr Ala Ser Phe Asn Thr Pro Ser Trp Ser Leu Ser
 130 135 140

Cys Glu Ile Ala Phe Tyr Leu Ser Phe Pro Leu Trp Tyr Arg Leu Val
 145 150 155 160

Arg Arg Ile Pro Ala Arg Arg Leu Trp Trp Cys Ala Ala Gly Ile Ala
 165 170 175

Val Ala Val Thr Cys Val Pro Leu Leu Ala Gly Leu Leu Pro Ala Ser
 180 185 190

Glu Glu Val Ala Pro Gly Met Ser Leu Asn Glu Val Trp Phe Ala Tyr
 195 200 205

Trp Leu Pro Pro Val Arg Met Leu Glu Phe Val Leu Gly Ile Val Met
 210 215 220

Ala Leu Ile Leu Arg Ala Gly Ile Trp Lys Gly Pro Gly Pro Ala Val
 225 230 235 240

Cys Thr Ala Leu Leu Ala Ala Ser Tyr Gly Leu Thr Gln Met Val Pro
 245 250 255

Pro Ile Phe Thr Leu Val Ala Cys Ser Val Val Pro Ala Ala Leu Leu
 260 265 270

Ile Thr Ala Leu Ala Asp Ala Asp Val His Gly Arg Arg Thr Gly Leu
 275 280 285

Arg Ser Ala Thr Leu Val Arg Leu Gly Gln Trp Ser Phe Ala Phe Tyr
 290 295 300

Leu Val His Phe Leu Ile Ile Arg Tyr Gly His Arg Leu Met Gly Gly
 305 310 315 320

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

Ser Leu Gly Met Leu Gly Val Ala Val Leu Ala Gly Gly Leu Leu His
 340 345 350

Thr Val Val Glu Gln Pro Cys Met Arg Leu Phe Gly Ser Arg Arg Ser
 355 360 365

Ala Ser Arg Pro Lys Pro Gly Ala Thr Ala Ala Pro Arg Asn Ser Pro
 370 375 380

Ala Ala Asp Ala Ala Gly Val Pro Leu Leu Pro Gly Val Pro Gly Pro
 385 390 395 400

Ala His Thr Pro Ala Ala Thr Asn Glu Pro Thr Pro Arg Gly
 405 410

<210> 15
 <211> 849
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(849)

<400> 15
 gtg aca ttg aaa tcc cca ctg cca ccg caa tcc gtc tcc gca ccc gct
 Val Thr Leu Lys Ser Pro Leu Pro Pro Gln Ser Val Ser Ala Pro Ala

1	5	10	15	
gat tca cgc agc acc gcg cgg aga gaa tgg ggt cag aac ttc ttt cgt				96
Asp Ser Arg Ser Thr Ala Arg Arg Glu Trp Gly Gln Asn Phe Phe Arg				
	20	25	30	
aca gcc gcc gcg gcc tgt cgt ttc tcc gct cag ctg gac ggt tcg gac				144
Thr Ala Ala Ala Ala Cys Arg Phe Ser Ala Gln Leu Asp Gly Ser Asp				
	35	40	45	
acc att ccg ccc gat tct ccg aat gac ctg atg acc gtc gaa ata ggc				192
Thr Ile Pro Pro Asp Ser Pro Asn Asp Leu Met Thr Val Glu Ile Gly				
	50	55	60	
gcg gga tca ggg cgg gtg acc aaa gtg ctc gcc tca ccc ggg acg cct				240
Ala Gly Ser Gly Arg Val Thr Lys Val Leu Ala Ser Pro Gly Thr Pro				
65	70	75	80	
tta ctc gcg gtg gaa ata gat ccc cgc tgg gct cgg cgg ctg gcc gcc				288
Leu Leu Ala Val Glu Ile Asp Pro Arg Trp Ala Arg Arg Leu Ala Ala				
	85	90	95	
gaa tcg ctg ccg gac gtc acg gtg gtg aac gag gat ttc ctg acc ctg				336
Glu Ser Leu Pro Asp Val Thr Val Val Asn Glu Asp Phe Leu Thr Leu				
	100	105	110	
caa ctg ccc ggg cag ccg gtc aga ctc atc ggg aat ctt ccc ttc gtc				384
Gln Leu Pro Gly Gln Pro Val Arg Leu Ile Gly Asn Leu Pro Phe Val				
	115	120	125	
acc ggc acc aga atg ctg agg cgc tgc ctc gac atg ggt ccg gcg cgc				432
Thr Gly Thr Arg Met Leu Arg Arg Cys Leu Asp Met Gly Pro Ala Arg				
	130	135	140	
atg cgg cag ggc gtg ttc ctg ttg cag cgg gag tac gtg gga aag cgg				480
Met Arg Gln Gly Val Phe Leu Leu Gln Arg Glu Tyr Val Gly Lys Arg				
145	150	155	160	
acc gga gcc tgg ggc gga aac ctc ttc aac gcc cag tgg gag ccg tgg				528
Thr Gly Ala Trp Gly Gly Asn Leu Phe Asn Ala Gln Trp Glu Pro Trp				
	165	170	175	
tac tcg ttc gac cgg ggc ctg gcc ttc tca cgc cag gac ttc acc ccc				576
Tyr Ser Phe Asp Arg Gly Leu Ala Phe Ser Arg Gln Asp Phe Thr Pro				
	180	185	190	
gta ccg cgc gcg gac acc cag acc ctg atg gtc gcc ccg cac cgc agg				624

Val Pro Arg Ala Asp Thr Gln Thr Leu Met Val Ala Pro His Arg Arg
 195 200 205

ccg tcc gtg ccc tgg cgt gag aag gcc gcc tac cag cgg ttc gtc caa 672
 Pro Ser Val Pro Trp Arg Glu Lys Ala Ala Tyr Gln Arg Phe Val Gln
 210 215 220

cgg gtc ttc gac acc ggc cag atg acg gtg ggc gac gcc gcg cgg aag 720
 Arg Val Phe Asp Thr Gly Gln Met Thr Val Gly Asp Ala Ala Arg Lys
 225 230 235 240

gtg ctg cgc cgc gga cac gcc cag ttc gtg cgc ggg gcg ggc gtg agg 768
 Val Leu Arg Arg Gly His Ala Gln Phe Val Arg Gly Ala Gly Val Arg
 245 250 255

ccg gcc gac cgg gtc aag gac ctc acg gtc ccg gag tgg acc gca ctc 816
 Pro Ala Asp Arg Val Lys Asp Leu Thr Val Pro Glu Trp Thr Ala Leu
 260 265 270

ttc cgc gcc tac ggg cgg acg gcc gac cgc tga 849
 Phe Arg Ala Tyr Gly Arg Thr Ala Asp Arg
 275 280

<210> 16

<211> 282

<212> PRT

<213> Streptomyces ambofaciens

<400> 16

Val Thr Leu Lys Ser Pro Leu Pro Pro Gln Ser Val Ser Ala Pro Ala
 1 5 10 15

Asp Ser Arg Ser Thr Ala Arg Arg Glu Trp Gly Gln Asn Phe Phe Arg
 20 25 30

Thr Ala Ala Ala Ala Cys Arg Phe Ser Ala Gln Leu Asp Gly Ser Asp
 35 40 45

Thr Ile Pro Pro Asp Ser Pro Asn Asp Leu Met Thr Val Glu Ile Gly
 50 55 60

Ala Gly Ser Gly Arg Val Thr Lys Val Leu Ala Ser Pro Gly Thr Pro
 65 70 75 80

Leu Leu Ala Val Glu Ile Asp Pro Arg Trp Ala Arg Arg Leu Ala Ala
 85 90 95

Glu Ser Leu Pro Asp Val Thr Val Val Asn Glu Asp Phe Leu Thr Leu
 100 105 110

Gln Leu Pro Gly Gln Pro Val Arg Leu Ile Gly Asn Leu Pro Phe Val
 115 120 125

Thr Gly Thr Arg Met Leu Arg Arg Cys Leu Asp Met Gly Pro Ala Arg
 130 135 140

Met Arg Gln Gly Val Phe Leu Leu Gln Arg Glu Tyr Val Gly Lys Arg
 145 150 155 160

Thr Gly Ala Trp Gly Gly Asn Leu Phe Asn Ala Gln Trp Glu Pro Trp
 165 170 175

Tyr Ser Phe Asp Arg Gly Leu Ala Phe Ser Arg Gln Asp Phe Thr Pro
 180 185 190

Val Pro Arg Ala Asp Thr Gln Thr Leu Met Val Ala Pro His Arg Arg
 195 200 205

Pro Ser Val Pro Trp Arg Glu Lys Ala Ala Tyr Gln Arg Phe Val Gln
 210 215 220

Arg Val Phe Asp Thr Gly Gln Met Thr Val Gly Asp Ala Ala Arg Lys
 225 230 235 240

Val Leu Arg Arg Gly His Ala Gln Phe Val Arg Gly Ala Gly Val Arg
 245 250 255

Pro Ala Asp Arg Val Lys Asp Leu Thr Val Pro Glu Trp Thr Ala Leu
 260 265 270

Phe Arg Ala Tyr Gly Arg Thr Ala Asp Arg
 275 280

<210> 17
 <211> 831
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(831)

<400> 17
 gtg aag gag gtg gcg gtc gac acc ggt acc ccg gcc ggc ccc gtg tgg 48
 Val Lys Glu Val Ala Val Asp Thr Gly Thr Pro Ala Gly Pro Val Trp
 1 5 10 15

 cgt ggt gcc ggc ctc ggc acc cag ctg tgg gta ctg acc gcg cgg cag 96
 Arg Gly Ala Gly Leu Gly Thr Gln Leu Trp Val Leu Thr Ala Arg Gln
 20 25 30

 atc cgt tcc atg tac ggc gac cgc cgc ctg gcg ctg ttc agc ctg atg 144
 Ile Arg Ser Met Tyr Gly Asp Arg Arg Leu Ala Leu Phe Ser Leu Met
 35 40 45

 caa ccg gtg atc atg ctg ttg ctg ctg agc gag atc ttc ggc agc atg 192
 Gln Pro Val Ile Met Leu Leu Leu Leu Ser Glu Ile Phe Gly Ser Met
 50 55 60

 gcc gac ccg gac gac ttc ccg cag ggc gtg cgc tac atc gac tac gtg 240
 Ala Asp Pro Asp Asp Phe Pro Gln Gly Val Arg Tyr Ile Asp Tyr Val
 65 70 75 80

 gtg ccc gcg ctg ctg gtc acc acc ggc atc ggc tcg gcc cag ggc gcg 288
 Val Pro Ala Leu Leu Val Thr Thr Gly Ile Gly Ser Ala Gln Gly Ala
 85 90 95

 ggg gtg ggc ctg gtc agg gac atg gac aac ggg atg gtg gcg cgc ttc 336
 Gly Val Gly Leu Val Arg Asp Met Asp Asn Gly Met Val Ala Arg Phe
 100 105 110

cgc gtc ctg ccg gcc cgg ctg ttc ctg gtg ctg gtc gcc cgg tcg ctg	384
Arg Val Leu Pro Ala Arg Leu Phe Leu Val Leu Val Ala Arg Ser Leu	
115 120 125	
gcc gat ctg gtc cgt gtg ttc acc gag ttg gtc gtc ctc gtg gcc gtc	432
Ala Asp Leu Val Arg Val Phe Thr Glu Leu Val Val Leu Val Ala Val	
130 135 140	
ggg gtg atc ctg ctg ggc ttc cgt ccg gcc ggg ggt ttg tgg ggc acg	480
Gly Val Ile Leu Leu Gly Phe Arg Pro Ala Gly Gly Leu Trp Gly Thr	
145 150 155 160	
tcc gcc gcc ctg ttg ctc acc ctg ttc gtc atc tgg tcg ctg atc tgg	528
Ser Ala Ala Leu Leu Leu Thr Leu Phe Val Ile Trp Ser Leu Ile Trp	
165 170 175	
ggg ttc atc gcc ctc gcg gcg tgg ctg cgc agc gtg gag gtg atg tcc	576
Gly Phe Ile Ala Leu Ala Ala Trp Leu Arg Ser Val Glu Val Met Ser	
180 185 190	
agc ctc gcg gtt ctg gtg atg ttc ccg ctc atg ttc gcc tcc agt gcg	624
Ser Leu Ala Val Leu Val Met Phe Pro Leu Met Phe Ala Ser Ser Ala	
195 200 205	
ttc gtc ccg ctg gac gcc ctc ccg gag tgg ctg cgc tcg gtg gcg cac	672
Phe Val Pro Leu Asp Ala Leu Pro Glu Trp Leu Arg Ser Val Ala His	
210 215 220	
ctc aac ccc gtg acg tac gcg gtc gac agc gcc cgc cgc ctg gcg ctg	720
Leu Asn Pro Val Thr Tyr Ala Val Asp Ser Ala Arg Arg Leu Ala Leu	
225 230 235 240	
gac tgg gac ccg ggg tgg agc gtg ccc ggc gcg ctg ctg acc agc acc	768
Asp Trp Asp Pro Gly Trp Ser Val Pro Gly Ala Leu Leu Thr Ser Thr	
245 250 255	
gcg ctc atg gcg gtg ggg atg tac gtc gcc ggg cgt tcc ttc aag agg	816
Ala Leu Met Ala Val Gly Met Tyr Val Ala Gly Arg Ser Phe Lys Arg	
260 265 270	
ccc ccg aac gaa tga	831
Pro Pro Asn Glu	
275	

<211> 276

<212> PRT

<213> Streptomyces ambofaciens

<400> 18

Val Lys Glu Val Ala Val Asp Thr Gly Thr Pro Ala Gly Pro Val Trp
 1 5 10 15

Arg Gly Ala Gly Leu Gly Thr Gln Leu Trp Val Leu Thr Ala Arg Gln
 20 25 30

Ile Arg Ser Met Tyr Gly Asp Arg Arg Leu Ala Leu Phe Ser Leu Met
 35 40 45

Gln Pro Val Ile Met Leu Leu Leu Leu Ser Glu Ile Phe Gly Ser Met
 50 55 60

Ala Asp Pro Asp Asp Phe Pro Gln Gly Val Arg Tyr Ile Asp Tyr Val
 65 70 75 80

Val Pro Ala Leu Leu Val Thr Thr Gly Ile Gly Ser Ala Gln Gly Ala
 85 90 95

Gly Val Gly Leu Val Arg Asp Met Asp Asn Gly Met Val Ala Arg Phe
 100 105 110

Arg Val Leu Pro Ala Arg Leu Phe Leu Val Leu Val Ala Arg Ser Leu
 115 120 125

Ala Asp Leu Val Arg Val Phe Thr Glu Leu Val Val Leu Val Ala Val
 130 135 140

Gly Val Ile Leu Leu Gly Phe Arg Pro Ala Gly Gly Leu Trp Gly Thr
 145 150 155 160

Ser Ala Ala Leu Leu Leu Thr Leu Phe Val Ile Trp Ser Leu Ile Trp

165

170

175

Gly Phe Ile Ala Leu Ala Ala Trp Leu Arg Ser Val Glu Val Met Ser
 180 185 190

Ser Leu Ala Val Leu Val Met Phe Pro Leu Met Phe Ala Ser Ser Ala
 195 200 205

Phe Val Pro Leu Asp Ala Leu Pro Glu Trp Leu Arg Ser Val Ala His
 210 215 220

Leu Asn Pro Val Thr Tyr Ala Val Asp Ser Ala Arg Arg Leu Ala Leu
 225 230 235 240

Asp Trp Asp Pro Gly Trp Ser Val Pro Gly Ala Leu Leu Thr Ser Thr
 245 250 255

Ala Leu Met Ala Val Gly Met Tyr Val Ala Gly Arg Ser Phe Lys Arg
 260 265 270

Pro Pro Asn Glu
 275

<210> 19
 <211> 882
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(882)

<400> 19
 gtg gac atc gag gtc gaa cgg ggc cgg gtg ctc ggg ctg ctg ggt cac 48
 Val Asp Ile Glu Val Glu Arg Gly Arg Val Leu Gly Leu Leu Gly His
 1 5 10 15

aac ggg gcc ggc aag acg acc ttg gtg aac atc ctt gcc acg gtc tcc 96

Asn Gly Ala Gly Lys Thr Thr Leu Val Asn Ile Leu Ala Thr Val Ser	
20 25 30	
ccg gcg tcc gcg ggc acg gtg acc gtc gcc ggt ttc gac gtc gcg acg	144
Pro Ala Ser Ala Gly Thr Val Thr Val Ala Gly Phe Asp Val Ala Thr	
35 40 45	
cag ggc gcc gag atc cgc gcg cgc atc ggg gtg acc ggc cag ttc gcg	192
Gln Gly Ala Glu Ile Arg Ala Arg Ile Gly Val Thr Gly Gln Phe Ala	
50 55 60	
tcg gtg gac gag tac ctg agc gga ttc cgc aac ctc gtc ctg atc ggc	240
Ser Val Asp Glu Tyr Leu Ser Gly Phe Arg Asn Leu Val Leu Ile Gly	
65 70 75 80	
cgc ctc ctc ggg gcg gga cgg cgt gag gcg gcg gcc cgg gcc acc gag	288
Arg Leu Leu Gly Ala Gly Arg Arg Glu Ala Ala Ala Arg Ala Thr Glu	
85 90 95	
ctc ctg gag ctg ttc gag ctg acc ggg gcg gcc cac cag ccc tcc cgc	336
Leu Leu Glu Leu Phe Glu Leu Thr Gly Ala Ala His Gln Pro Ser Arg	
100 105 110	
acc tac tcg ggc ggg atg cgc cga cgg ctc gac ctc gcc gcc agc ctg	384
Thr Tyr Ser Gly Gly Met Arg Arg Arg Leu Asp Leu Ala Ala Ser Leu	
115 120 125	
gtc ggc cgg ccg gac gtg ctg ttc ctc gac gag ccg acg acc ggg ctg	432
Val Gly Arg Pro Asp Val Leu Phe Leu Asp Glu Pro Thr Thr Gly Leu	
130 135 140	
gac ccg gcg acc cgg atc gcc ctg tgg gag acg gtg gag aag ctg gtg	480
Asp Pro Ala Thr Arg Ile Ala Leu Trp Glu Thr Val Glu Lys Leu Val	
145 150 155 160	
gcg ggc ggc acg acc gtc ctg ctg acc acc cag tac ctg gac gag gcg	528
Ala Gly Gly Thr Thr Val Leu Leu Thr Thr Gln Tyr Leu Asp Glu Ala	
165 170 175	
gac cgg ctg gcc gac tgg atc acc gtc ctg tcg aag ggc cgg gtg gtg	576
Asp Arg Leu Ala Asp Trp Ile Thr Val Leu Ser Lys Gly Arg Val Val	
180 185 190	
gcc tcg gac acc acc gac cgg ctc aag gcc gac ctg ggc cac cgg tcg	624
Ala Ser Asp Thr Thr Asp Arg Leu Lys Ala Asp Leu Gly His Arg Ser	
195 200 205	

gtg cgg gtg gtc ctt ccg ccc gcc gcc gac ctg acg gcc gcc gcc gcc 672
 Val Arg Val Val Leu Pro Pro Ala Ala Asp Leu Thr Ala Ala Ala Ala
 210 215 220

gcg ctc acc gcc ggc ggg ttc cgt ccg cgg tcc gac gcc ggg gag cac 720
 Ala Leu Thr Ala Gly Gly Phe Arg Pro Arg Ser Asp Ala Gly Glu His
 225 230 235 240

gcg ctg acc acg ccc gtg gac acc tcg gcc ggt atc gcg ggc gtc atc 768
 Ala Leu Thr Thr Pro Val Asp Thr Ser Ala Gly Ile Ala Gly Val Ile
 245 250 255

cgc gcg ctg gac acc gtc gga acg cag gcc gtc gag ctg acc gtc aag 816
 Arg Ala Leu Asp Thr Val Gly Thr Gln Ala Val Glu Leu Thr Val Lys
 260 265 270

gag ccg tcc ctg gac gac gtc tac ctg gcg ctc acc cat ccc tca ccc 864
 Glu Pro Ser Leu Asp Asp Val Tyr Leu Ala Leu Thr His Pro Ser Pro
 275 280 285

gcc gcc gac gcg gcc tga 882
 Ala Ala Asp Ala Ala
 290

<210> 20
 <211> 293
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 20

Val Asp Ile Glu Val Glu Arg Gly Arg Val Leu Gly Leu Leu Gly His
 1 5 10 15

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

Pro Ala Ser Ala Gly Thr Val Thr Val Ala Gly Phe Asp Val Ala Thr
 35 40 45

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

Ser Val Asp Glu Tyr Leu Ser Gly Phe Arg Asn Leu Val Leu Ile Gly
 65 70 75 80

Arg Leu Leu Gly Ala Gly Arg Arg Glu Ala Ala Ala Arg Ala Thr Glu
 85 90 95

Leu Leu Glu Leu Phe Glu Leu Thr Gly Ala Ala His Gln Pro Ser Arg
 100 105 110

Thr Tyr Ser Gly Gly Met Arg Arg Arg Leu Asp Leu Ala Ala Ser Leu
 115 120 125

Val Gly Arg Pro Asp Val Leu Phe Leu Asp Glu Pro Thr Thr Gly Leu
 130 135 140

Asp Pro Ala Thr Arg Ile Ala Leu Trp Glu Thr Val Glu Lys Leu Val
 145 150 155 160

Ala Gly Gly Thr Thr Val Leu Leu Thr Thr Gln Tyr Leu Asp Glu Ala
 165 170 175

Asp Arg Leu Ala Asp Trp Ile Thr Val Leu Ser Lys Gly Arg Val Val
 180 185 190

Ala Ser Asp Thr Thr Asp Arg Leu Lys Ala Asp Leu Gly His Arg Ser
 195 200 205

Val Arg Val Val Leu Pro Pro Ala Ala Asp Leu Thr Ala Ala Ala Ala
 210 215 220

Ala Leu Thr Ala Gly Gly Phe Arg Pro Arg Ser Asp Ala Gly Glu His
 225 230 235 240

Ala Leu Thr Thr Pro Val Asp Thr Ser Ala Gly Ile Ala Gly Val Ile
 245 250 255

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

Glu Pro Ser Leu Asp Asp Val Tyr Leu Ala Leu Thr His Pro Ser Pro
 275 280 285

Ala Ala Asp Ala Ala
 290

<210> 21
 <211> 228
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(228)

<400> 21
 atg agc aac ccg ttc gag gac acg gaa gcc acc tac gtc gtg ctg gtc 48
 Met Ser Asn Pro Phe Glu Asp Thr Glu Ala Thr Tyr Val Val Leu Val
 1 5 10 15

 aac gac gag ggg cag cac tcg ctg tgg ccg tcg ttc gcg gag gtc ccg 96
 Asn Asp Glu Gly Gln His Ser Leu Trp Pro Ser Phe Ala Glu Val Pro
 20 25 30

 gcg ggc tgg tcc gtc gtg gtg ccg gag acg gac cgg cag tcg tgc ctg 144
 Ala Gly Trp Ser Val Val Val Pro Glu Thr Asp Arg Gln Ser Cys Leu
 35 40 45

 gac tac atc aac gag aac tgg acc gac atg cgc ccc aag agc ctc gtc 192
 Asp Tyr Ile Asn Glu Asn Trp Thr Asp Met Arg Pro Lys Ser Leu Val
 50 55 60

 gag gcg atg gcg acg gcc ggg cag gac gcc cct tga 228
 Glu Ala Met Ala Thr Ala Gly Gln Asp Ala Pro
 65 70 75

<210> 22

<211> 75
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 22

Met Ser Asn Pro Phe Glu Asp Thr Glu Ala Thr Tyr Val Val Leu Val
 1 5 10 15

Asn Asp Glu Gly Gln His Ser Leu Trp Pro Ser Phe Ala Glu Val Pro
 20 25 30

Ala Gly Trp Ser Val Val Val Pro Glu Thr Asp Arg Gln Ser Cys Leu
 35 40 45

Asp Tyr Ile Asn Glu Asn Trp Thr Asp Met Arg Pro Lys Ser Leu Val
 50 55 60

Glu Ala Met Ala Thr Ala Gly Gln Asp Ala Pro
 65 70 75

<210> 23
 <211> 1212
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1212)

<400> 23

atg ggt gag gcc gtg acg gga ccg atg gag ctg agc aag gac gcg gac 48
 Met Gly Glu Ala Val Thr Gly Pro Met Glu Leu Ser Lys Asp Ala Asp
 1 5 10 15

gcc cgg ggg ctg ctt gag tgg ttc gcg tac aac agg acg cgt cat ccg 96
 Ala Arg Gly Leu Leu Glu Trp Phe Ala Tyr Asn Arg Thr Arg His Pro
 20 25 30

gtg ttc tgg gac gag acc cga cag gcg tgg cag gtc ttc ggc tac gac 144
 Val Phe Trp Asp Glu Thr Arg Gln Ala Trp Gln Val Phe Gly Tyr Asp

35	40	45	
gac tac gtg acg gtg tcg aac aac ccg cag ttc ttc tcc tcg gac ttc			192
Asp Tyr Val Thr Val Ser Asn Asn Pro Gln Phe Phe Ser Ser Asp Phe			
50	55	60	
aac atg gtg atg ccg acg ccg ccc gaa ctg gag atg atc atc ggt ccg			240
Asn Met Val Met Pro Thr Pro Pro Glu Leu Glu Met Ile Ile Gly Pro			
65	70	75	80
ggc acg atc ggc gcg ctg gac ccg ccc gcg cac gga ccg atg cgc aag			288
Gly Thr Ile Gly Ala Leu Asp Pro Pro Ala His Gly Pro Met Arg Lys			
	85	90	95
ctg gtg agc cag gcg ttc acc ccc cga cgg atc gcc cgg ctg gag ccc			336
Leu Val Ser Gln Ala Phe Thr Pro Arg Arg Ile Ala Arg Leu Glu Pro			
	100	105	110
agg gtg cgc gcg atc acc gag gag ctc ctg gac aag gtg ggg cag cag			384
Arg Val Arg Ala Ile Thr Glu Glu Leu Leu Asp Lys Val Gly Gln Gln			
	115	120	125
gac gtc gtc gac gcc gtg ggt gac ctg tcc tac gcg ctg ccg gtc atc			432
Asp Val Val Asp Ala Val Gly Asp Leu Ser Tyr Ala Leu Pro Val Ile			
130	135	140	
gtg atc gcc gaa ctg ctg ggc ata ccc gcc ggc gac cgt gac ctg ttc			480
Val Ile Ala Glu Leu Leu Gly Ile Pro Ala Gly Asp Arg Asp Leu Phe			
145	150	155	160
cgg gag tgg gtc gac acc ctg ctg acg aac gag ggc ctg gag tac ccg			528
Arg Glu Trp Val Asp Thr Leu Leu Thr Asn Glu Gly Leu Glu Tyr Pro			
	165	170	175
aac ctc ccg gac aac ttc acc gag acg atc gcg ccc gcg ctc aag gag			576
Asn Leu Pro Asp Asn Phe Thr Glu Thr Ile Ala Pro Ala Leu Lys Glu			
	180	185	190
atg acc gac tac ctc ctg aag cag atc cac gcc aag cgg gac gcg ccc			624
Met Thr Asp Tyr Leu Leu Lys Gln Ile His Ala Lys Arg Asp Ala Pro			
	195	200	205
gcc gac gac ctg gtc agc ggg ctg gtc cag gcg gag cag gac ggc cgc			672
Ala Asp Asp Leu Val Ser Gly Leu Val Gln Ala Glu Gln Asp Gly Arg			
210	215	220	
cgg ctg acc gac gtc gag atc gtc aac atc gtc gcg ctg ctc ctg acg			720

[illegible]

<210> 24
 <211> 403
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 24

Met Gly Glu Ala Val Thr Gly Pro Met Glu Leu Ser Lys Asp Ala Asp
 1 5 10 15

Ala Arg Gly Leu Leu Glu Trp Phe Ala Tyr Asn Arg Thr Arg His Pro
 20 25 30

Val Phe Trp Asp Glu Thr Arg Gln Ala Trp Gln Val Phe Gly Tyr Asp
 35 40 45

Asp Tyr Val Thr Val Ser Asn Asn Pro Gln Phe Phe Ser Ser Asp Phe
 50 55 60

Asn Met Val Met Pro Thr Pro Pro Glu Leu Glu Met Ile Ile Gly Pro
 65 70 75 80

Gly Thr Ile Gly Ala Leu Asp Pro Pro Ala His Gly Pro Met Arg Lys
 85 90 95

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

Arg Val Arg Ala Ile Thr Glu Glu Leu Leu Asp Lys Val Gly Gln Gln
 115 120 125

Asp Val Val Asp Ala Val Gly Asp Leu Ser Tyr Ala Leu Pro Val Ile
 130 135 140

Val Ile Ala Glu Leu Leu Gly Ile Pro Ala Gly Asp Arg Asp Leu Phe
 145 150 155 160

Arg Glu Trp Val Asp Thr Leu Leu Thr Asn Glu Gly Leu Glu Tyr Pro
 165 170 175

Asn Leu Pro Asp Asn Phe Thr Glu Thr Ile Ala Pro Ala Leu Lys Glu
 180 185 190

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

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

Arg Leu Thr Asp Val Glu Ile Val Asn Ile Val Ala Leu Leu Leu Thr
 225 230 235 240

Ala Gly His Val Ser Ser Ser Thr Leu Leu Ser Asn Leu Phe Leu Val
 245 250 255

Leu Glu Glu Asn Pro Gln Ala Leu Glu Asp Leu Arg Ala Asp Arg Ser
 260 265 270

Leu Val Pro Gly Ala Ile Glu Glu Thr Leu Arg Tyr Arg Ser Pro Phe
 275 280 285

Asn Asn Ile Phe Arg Phe Val Lys Glu Asp Thr Thr Val Leu Gly Pro
 290 295 300

Leu Met Glu Lys Gly Gln Met Val Ile Ala Trp Ser Gln Ser Ala Asn
 305 310 315 320

Arg Asp Pro Arg His Phe Pro Asp Pro Asp Thr Phe Asp Ile Arg Arg
 325 330 335

Ser Asp Gly Thr Arg His Met Ala Phe Gly His Gly Ile His His Cys
 340 345 350

Leu Gly Ala Ala Leu Ala Arg Leu Glu Gly Lys Val Met Leu Glu Leu
 355 360 365

Leu Leu Asp Arg Val Gln Gly Phe Arg Ile Asp His Glu His Thr Val
 370 375 380

Phe Tyr Glu Ala Asp Gln Leu Thr Pro Lys Tyr Leu Pro Val Arg Val
 385 390 395 400

Asp Trp Asn

<210> 25
 <211> 540
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(540)

<220>
 <221> CDS2
 <222> (190)..(540)

<400> 25
 ctg aac ccg agg gtc tcg tcc cgg agt cca ggg ccg tcc cga gcc ggc 48
 Leu Asn Pro Arg Val Ser Ser Arg Ser Pro Gly Pro Ser Arg Ala Gly
 1 5 10 15
 cct gga cct cac gac cgc ccg ata agg agc gcc gcc atc gcc gag aac 96
 Pro Gly Pro His Asp Arg Pro Ile Arg Ser Ala Ala Ile Ala Glu Asn
 20 25 30
 aca gcc gag ctc cct gcc cgg cgg gtc ggc agg atc aag ccg tgc cgg 144
 Thr Ala Glu Leu Pro Ala Arg Arg Val Gly Arg Ile Lys Pro Cys Arg
 35 40 45
 ctg atc agg ctc gag cag cac atc gac ccg cgc ggc agc ctc tcc gtg 192
 Leu Ile Arg Leu Glu Gln His Ile Asp Pro Arg Gly Ser Leu Ser Val

50				55				60								
atc	gag	tcc	ggc	gtg	acc	gtg	gac	ttc	ccc	gtc	cga	cgc	gtc	tac	tac	240
Ile	Glu	Ser	Gly	Val	Thr	Val	Asp	Phe	Pro	Val	Arg	Arg	Val	Tyr	Tyr	
65				70				75				80				
atg	cat	ggc	cag	acc	cag	tcc	tct	ccc	ccg	cgc	ggc	ctg	cac	gcg	cac	288
Met	His	Gly	Gln	Thr	Gln	Ser	Ser	Pro	Pro	Arg	Gly	Leu	His	Ala	His	
				85				90				95				
cgc	acc	ctg	gaa	caa	ctc	gtc	atc	gcc	gtc	cac	ggc	gcc	ttc	tcc	atc	336
Arg	Thr	Leu	Glu	Gln	Leu	Val	Ile	Ala	Val	His	Gly	Ala	Phe	Ser	Ile	
				100				105				110				
acc	ctc	gac	gac	ggc	ttc	cag	cac	gcc	acc	tac	cgt	ctg	gac	gaa	ccc	384
Thr	Leu	Asp	Asp	Gly	Phe	Gln	His	Ala	Thr	Tyr	Arg	Leu	Asp	Glu	Pro	
				115				120				125				
gga	gcc	gga	ctc	tgc	atc	ggc	ccc	atg	gtc	tgg	cgc	gtc	ctg	aag	gac	432
Gly	Ala	Gly	Leu	Cys	Ile	Gly	Pro	Met	Val	Trp	Arg	Val	Leu	Lys	Asp	
				130				135				140				
ttc	gac	ccc	gac	acc	gtg	gcc	ctg	gtc	ctc	gcc	tgc	cag	cac	tac	gag	480
Phe	Asp	Pro	Asp	Thr	Val	Ala	Leu	Val	Leu	Ala	Ser	Gln	His	Tyr	Glu	
145				150				155				160				
gag	tcc	gac	tac	tac	cgc	gac	tac	gac	acc	ttc	ctg	cat	gac	gca	cgg	528
Glu	Ser	Asp	Tyr	Tyr	Arg	Asp	Tyr	Asp	Thr	Phe	Leu	His	Asp	Ala	Arg	
				165				170				175				
agc	ctc	aca	tga													540
Ser	Leu	Thr														

```
<210> 26
<211> 179
<212> PRT
<213> Streptomyces ambofaciens
```

<400> 26

Leu Asn Pro Arg Val Ser Ser Arg Ser Pro Gly Pro Ser Arg Ala Gly
1 5 10 15

Pro Gly Pro His Asp Arg Pro Ile Arg Ser Ala Ala Ile Ala Glu Asn

20

25

30

Thr Ala Glu Leu Pro Ala Arg Arg Val Gly Arg Ile Lys Pro Cys Arg
 35 40 45

Leu Ile Arg Leu Glu Gln His Ile Asp Pro Arg Gly Ser Leu Ser Val
 50 55 60

Ile Glu Ser Gly Val Thr Val Asp Phe Pro Val Arg Arg Val Tyr Tyr
 65 70 75 80

Met His Gly Gln Thr Gln Ser Ser Pro Pro Arg Gly Leu His Ala His
 85 90 95

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

Thr Leu Asp Asp Gly Phe Gln His Ala Thr Tyr Arg Leu Asp Glu Pro
 115 120 125

Gly Ala Gly Leu Cys Ile Gly Pro Met Val Trp Arg Val Leu Lys Asp
 130 135 140

Phe Asp Pro Asp Thr Val Ala Leu Val Leu Ala Ser Gln His Tyr Glu
 145 150 155 160

Glu Ser Asp Tyr Tyr Arg Asp Tyr Asp Thr Phe Leu His Asp Ala Arg
 165 170 175

Ser Leu Thr

<210> 27

<211> 116

<212> PRT

<213> Streptomyces ambofaciens

<400> 27

Val Ile Glu Ser Gly Val Thr Val Asp Phe Pro Val Arg Arg Val Tyr
 1 5 10 15

Tyr Met His Gly Gln Thr Gln Ser Ser Pro Pro Arg Gly Leu His Ala
 20 25 30

His Arg Thr Leu Glu Gln Leu Val Ile Ala Val His Gly Ala Phe Ser
 35 40 45

Ile Thr Leu Asp Asp Gly Phe Gln His Ala Thr Tyr Arg Leu Asp Glu
 50 55 60

Pro Gly Ala Gly Leu Cys Ile Gly Pro Met Val Trp Arg Val Leu Lys
 65 70 75 80

Asp Phe Asp Pro Asp Thr Val Ala Leu Val Leu Ala Ser Gln His Tyr
 85 90 95

Glu Glu Ser Asp Tyr Tyr Arg Asp Tyr Asp Thr Phe Leu His Asp Ala
 100 105 110

Arg Ser Leu Thr
 115

<210> 28

<211> 1167

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1)..(1167)

<400> 28

atg acc atc ccc ttc ctc gac ggc ggc gcc ggc tac cgg gag ttg cga

Met	Thr	Ile	Pro	Phe	Leu	Asp	Ala	Gly	Ala	Gly	Tyr	Arg	Glu	Leu	Arg	
1				5					10					15		
gcc	gag	atc	gac	gcg	gcc	ctg	cag	cgg	gtg	tcc	gcc	tcc	ggc	cgc	tat	96
Ala	Glu	Ile	Asp	Ala	Ala	Leu	Gln	Arg	Val	Ser	Ala	Ser	Gly	Arg	Tyr	
			20					25					30			
ctg	ctc	gac	gcg	gaa	ctc	gcg	gcc	ttc	gag	gag	gag	ttc	gcc	gcg	tac	144
Leu	Leu	Asp	Ala	Glu	Leu	Ala	Ala	Phe	Glu	Glu	Glu	Phe	Ala	Ala	Tyr	
		35					40					45				
tgc	gac	aac	gac	cac	tgt	gtg	gcg	gtg	ggc	agt	ggc	tgc	gac	gcg	ctg	192
Cys	Asp	Asn	Asp	His	Cys	Val	Ala	Val	Gly	Ser	Gly	Cys	Asp	Ala	Leu	
	50					55				60						
gag	ctg	tcc	ctg	cgg	gcg	ctg	gac	atc	ggt	ccc	ggg	gac	gag	gtg	gtg	240
Glu	Leu	Ser	Leu	Arg	Ala	Leu	Asp	Ile	Gly	Pro	Gly	Asp	Glu	Val	Val	
65					70					75				80		
gtg	ccc	gcg	cac	acc	ttc	atc	ggg	acc	tgg	ctg	gcc	gtg	tcc	gct	acc	288
Val	Pro	Ala	His	Thr	Phe	Ile	Gly	Thr	Trp	Leu	Ala	Val	Ser	Ala	Thr	
				85				90					95			
ggg	gca	cgg	ccg	gtg	gcc	gtc	gac	ccg	acg	ccg	gac	ggg	ctc	tcc	ctc	336
Gly	Ala	Arg	Pro	Val	Ala	Val	Asp	Pro	Thr	Pro	Asp	Gly	Leu	Ser	Leu	
			100					105					110			
gac	ccg	gcg	ctg	gtg	gag	gcg	gcg	ctc	acc	cct	cgg	acc	aga	gcc	ctg	384
Asp	Pro	Ala	Leu	Val	Glu	Ala	Ala	Leu	Thr	Pro	Arg	Thr	Arg	Ala	Leu	
		115					120					125				
atg	ccg	gtg	cac	ctg	cac	ggg	cac	ccg	gcc	gac	ctc	gac	ccg	cta	ctg	432
Met	Pro	Val	His	Leu	His	Gly	His	Pro	Ala	Asp	Leu	Asp	Pro	Leu	Leu	
	130					135					140					
gcg	atc	gcc	gga	cgg	cac	ggc	ctg	gcc	gtg	gtc	gag	gac	gcc	gcg	cag	480
Ala	Ile	Ala	Gly	Arg	His	Gly	Leu	Ala	Val	Val	Glu	Asp	Ala	Ala	Gln	
145					150				155					160		
gcc	cac	ggc	gcc	cgt	tac	cgg	ggc	cgc	agg	atc	ggc	tcg	ggc	cac	gtg	528
Ala	His	Gly	Ala	Arg	Tyr	Arg	Gly	Arg	Arg	Ile	Gly	Ser	Gly	His	Val	
			165					170					175			
gtc	gcg	ttc	agc	ttc	tac	ccc	ggc	aag	aac	ctc	ggc	gcc	atg	ggg	gac	576
Val	Ala	Phe	Ser	Phe	Tyr	Pro	Gly	Lys	Asn	Leu	Gly	Ala	Met	Gly	Asp	
			180					185					190			

ggc ggc gcg gtg gtc acg ggt gac tcc ggt gtg gcc gag cgg atc cgg	624
Gly Gly Ala Val Val Thr Gly Asp Ser Gly Val Ala Glu Arg Ile Arg	
195 200 205	
ttg ctg cgc aac tgc ggc tgc cgg gag aag tac cgg cac gag gtg cgc	672
Leu Leu Arg Asn Cys Gly Ser Arg Glu Lys Tyr Arg His Glu Val Arg	
210 215 220	
tcg acc cac tcc cgg ctc gac gag ttc cag gcg gcc gtg ctg cgg gcc	720
Ser Thr His Ser Arg Leu Asp Glu Phe Gln Ala Ala Val Leu Arg Ala	
225 230 235 240	
aaa ctg ccg cgg ctc gac gcg tgg aac gcc cgc cgg gcc ggc acg gcc	768
Lys Leu Pro Arg Leu Asp Ala Trp Asn Ala Arg Arg Ala Gly Thr Ala	
245 250 255	
gaa cgg tac ggg cgg gcc ctg ggt ccg gta ccg cag atc gcc gtc ccg	816
Glu Arg Tyr Gly Arg Ala Leu Gly Pro Val Pro Gln Ile Ala Val Pro	
260 265 270	
gtc acc gct ccc tgg gcc gac ccg gtg tgg cac ctg tac gtg atc cgc	864
Val Thr Ala Pro Trp Ala Asp Pro Val Trp His Leu Tyr Val Ile Arg	
275 280 285	
tgc gcg gag cgc gac gag ctg cgc cgc cgg ctg gaa cga gcc ggg gtc	912
Cys Ala Glu Arg Asp Glu Leu Arg Arg Arg Leu Glu Arg Ala Gly Val	
290 295 300	
cag acc ctg atc cac tac ccc gtg ccc ccg cac cgg tcc ccg gcc tac	960
Gln Thr Leu Ile His Tyr Pro Val Pro Pro His Arg Ser Pro Ala Tyr	
305 310 315 320	
gcc gac gac ccg gcc ggc gca ccg gcg ggg acc cac ccg ctc agt gag	1008
Ala Asp Asp Pro Ala Gly Ala Pro Ala Gly Thr His Pro Leu Ser Glu	
325 330 335	
cgc ctg gcg gcg cag agc ctc agc ctt ccc ctg gga ccg cac ctc ggg	1056
Arg Leu Ala Ala Gln Ser Leu Ser Leu Pro Leu Gly Pro His Leu Gly	
340 345 350	
gag gac gag gcc cgc gcc gtc gtg gcg gcg gtc cgg gcg gcg tcc gca	1104
Glu Asp Glu Ala Arg Ala Val Val Ala Ala Val Arg Ala Ala Ser Ala	
355 360 365	
ggg ctg gcg gcg tac ccg acg ccg gac ggc cag cgt ttt cct cta gtg	1152
Gly Leu Ala Ala Tyr Pro Thr Pro Asp Gly Gln Arg Phe Pro Leu Val	
370 375 380	

acg gag aaa cga tga
 Thr Glu Lys Arg
 385

1167

<210> 29
 <211> 388
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 29

Met Thr Ile Pro Phe Leu Asp Ala Gly Ala Gly Tyr Arg Glu Leu Arg
 1 5 10 15

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

Leu Leu Asp Ala Glu Leu Ala Ala Phe Glu Glu Glu Phe Ala Ala Tyr
 35 40 45

Cys Asp Asn Asp His Cys Val Ala Val Gly Ser Gly Cys Asp Ala Leu
 50 55 60

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

Val Pro Ala His Thr Phe Ile Gly Thr Trp Leu Ala Val Ser Ala Thr
 85 90 95

Gly Ala Arg Pro Val Ala Val Asp Pro Thr Pro Asp Gly Leu Ser Leu
 100 105 110

Asp Pro Ala Leu Val Glu Ala Ala Leu Thr Pro Arg Thr Arg Ala Leu
 115 120 125

Met Pro Val His Leu His Gly His Pro Ala Asp Leu Asp Pro Leu Leu
 130 135 140

Ala Ile Ala Gly Arg His Gly Leu Ala Val Val Glu Asp Ala Ala Gln
 145 150 155 160

Ala His Gly Ala Arg Tyr Arg Gly Arg Arg Ile Gly Ser Gly His Val
 165 170 175

Val Ala Phe Ser Phe Tyr Pro Gly Lys Asn Leu Gly Ala Met Gly Asp
 180 185 190

Gly Gly Ala Val Val Thr Gly Asp Ser Gly Val Ala Glu Arg Ile Arg
 195 200 205

Leu Leu Arg Asn Cys Gly Ser Arg Glu Lys Tyr Arg His Glu Val Arg
 210 215 220

Ser Thr His Ser Arg Leu Asp Glu Phe Gln Ala Ala Val Leu Arg Ala
 225 230 235 240

Lys Leu Pro Arg Leu Asp Ala Trp Asn Ala Arg Arg Ala Gly Thr Ala
 245 250 255

Glu Arg Tyr Gly Arg Ala Leu Gly Pro Val Pro Gln Ile Ala Val Pro
 260 265 270

Val Thr Ala Pro Trp Ala Asp Pro Val Trp His Leu Tyr Val Ile Arg
 275 280 285

Cys Ala Glu Arg Asp Glu Leu Arg Arg Arg Leu Glu Arg Ala Gly Val
 290 295 300

Gln Thr Leu Ile His Tyr Pro Val Pro Pro His Arg Ser Pro Ala Tyr
 305 310 315 320

Ala Asp Asp Pro Ala Gly Ala Pro Ala Gly Thr His Pro Leu Ser Glu

325

330

335

Arg Leu Ala Ala Gln Ser Leu Ser Leu Pro Leu Gly Pro His Leu Gly
 340 345 350

Glu Asp Glu Ala Arg Ala Val Val Ala Ala Val Arg Ala Ala Ser Ala
 355 360 365

Gly Leu Ala Ala Tyr Pro Thr Pro Asp Gly Gln Arg Phe Pro Leu Val
 370 375 380

Thr Glu Lys Arg
 385

<210> 30
 <211> 909
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(909)

<220>
 <221> CDS2
 <222> (10)..(909)

<220>
 <221> CDS3
 <222> (28)..(909)

<400> 30
 atg acc gag gtc atg tca ggg cgt ccc gga atg aaa ggg atc atc ctc 48
 Met Thr Glu Val Met Ser Gly Arg Pro Gly Met Lys Gly Ile Ile Leu
 1 5 10 15

gca ggc ggc gga ggg acc cgc cta cgc ccc ttg acc ggc acg ctg tcc 96
 Ala Gly Gly Gly Gly Thr Arg Leu Arg Pro Leu Thr Gly Thr Leu Ser
 20 25 30

aag caa ctg ctg ccc gtc tac gac aag ccg atg atc tac tac ccg ctg 144

Lys	Gln	Leu	Leu	Pro	Val	Tyr	Asp	Lys	Pro	Met	Ile	Tyr	Tyr	Pro	Leu	
	35						40					45				
tcc	gtc	ctg	atg	ctg	ggc	ggc	atc	cgc	gag	atc	ctc	gtc	gtc	tcc	tcc	192
Ser	Val	Leu	Met	Leu	Gly	Gly	Ile	Arg	Glu	Ile	Leu	Val	Val	Ser	Ser	
	50					55					60					
acc	cag	cac	atc	gag	ctg	ttc	cag	cgg	ctg	ctg	ggc	gac	ggc	tcc	cgc	240
Thr	Gln	His	Ile	Glu	Leu	Phe	Gln	Arg	Leu	Leu	Gly	Asp	Gly	Ser	Arg	
65					70				75					80		
ctc	ggc	ctc	gac	atc	acc	tac	gcc	gaa	cag	gcc	gag	ccc	gag	ggc	ata	288
Leu	Gly	Leu	Asp	Ile	Thr	Tyr	Ala	Glu	Gln	Ala	Glu	Pro	Glu	Gly	Ile	
				85					90					95		
gcg	cag	gcc	atc	acc	atc	ggc	acc	gac	cac	atc	ggc	gac	tca	ccg	gtc	336
Ala	Gln	Ala	Ile	Thr	Ile	Gly	Thr	Asp	His	Ile	Gly	Asp	Ser	Pro	Val	
			100					105					110			
gcg	ctc	atc	ctg	ggc	gac	aac	atc	ttc	cac	ggc	ccc	ggc	ttc	tcg	gcc	384
Ala	Leu	Ile	Leu	Gly	Asp	Asn	Ile	Phe	His	Gly	Pro	Gly	Phe	Ser	Ala	
		115					120					125				
gtg	ctc	cag	ggc	agc	atc	cgc	cac	ctc	gac	ggc	tgt	gtg	ctg	ttc	ggc	432
Val	Leu	Gln	Gly	Ser	Ile	Arg	His	Leu	Asp	Gly	Cys	Val	Leu	Phe	Gly	
	130					135					140					
tac	ccg	gtc	agc	gac	ccg	aag	cgc	tac	ggc	gtc	ggc	gag	atc	gac	gac	480
Tyr	Pro	Val	Ser	Asp	Pro	Lys	Arg	Tyr	Gly	Val	Gly	Glu	Ile	Asp	Asp	
145					150					155				160		
cag	ggc	gta	ctg	ctg	tcc	ctg	gag	gag	aaa	ccg	gcc	cgg	ccc	cgc	tcc	528
Gln	Gly	Val	Leu	Leu	Ser	Leu	Glu	Glu	Lys	Pro	Ala	Arg	Pro	Arg	Ser	
				165					170					175		
aac	ctc	gcc	gtc	acc	ggc	ctc	tac	ctc	tac	gac	aac	gac	gtg	gtc	gac	576
Asn	Leu	Ala	Val	Thr	Gly	Leu	Tyr	Leu	Tyr	Asp	Asn	Asp	Val	Val	Asp	
			180					185					190			
atc	gcc	aag	aac	atc	cgg	ccc	tcg	gcg	cgc	ggc	gaa	ctc	gag	atc	acg	624
Ile	Ala	Lys	Asn	Ile	Arg	Pro	Ser	Ala	Arg	Gly	Glu	Leu	Glu	Ile	Thr	
		195					200					205				
gac	gtc	aac	agg	acc	tac	ctg	gag	cag	aaa	cgc	gcc	cgg	ctc	atc	gaa	672
Asp	Val	Asn	Arg	Thr	Tyr	Leu	Glu	Gln	Lys	Arg	Ala	Arg	Leu	Ile	Glu	
	210						215					220				

ctg ggc cac ggc ttc gcc tgg ctc gac atg ggc acc cac gac tcc ctc 720
 Leu Gly His Gly Phe Ala Trp Leu Asp Met Gly Thr His Asp Ser Leu
 225 230 235 240

ctc cag ggc ggc cag tac gtc cag ctc atc gag cag cgc cag gga gtg 768
 Leu Gln Gly Gly Gln Tyr Val Gln Leu Ile Glu Gln Arg Gln Gly Val
 245 250 255

cgg atc gcc tgc atc gag gag atc gcc ctg cgc atg ggc ttc atc gac 816
 Arg Ile Ala Cys Ile Glu Glu Ile Ala Leu Arg Met Gly Phe Ile Asp
 260 265 270

gcc gac acc ctc cac cgg ctc ggc cgc gaa ctg ggc acc tcc gga tac 864
 Ala Asp Thr Leu His Arg Leu Gly Arg Glu Leu Gly Thr Ser Gly Tyr
 275 280 285

ggc gcg tac ctg atg gag gtg gcc acc cgt gca ggc acc gaa tga 909
 Gly Ala Tyr Leu Met Glu Val Ala Thr Arg Ala Gly Thr Glu
 290 295 300

<210> 31
 <211> 302
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 31

Met Thr Glu Val Met Ser Gly Arg Pro Gly Met Lys Gly Ile Ile Leu
 1 5 10 15

Ala Gly Gly Gly Gly Thr Arg Leu Arg Pro Leu Thr Gly Thr Leu Ser
 20 25 30

Lys Gln Leu Leu Pro Val Tyr Asp Lys Pro Met Ile Tyr Tyr Pro Leu
 35 40 45

Ser Val Leu Met Leu Gly Gly Ile Arg Glu Ile Leu Val Val Ser Ser
 50 55 60

Thr Gln His Ile Glu Leu Phe Gln Arg Leu Leu Gly Asp Gly Ser Arg
 65 70 75 80

Leu Gly Leu Asp Ile Thr Tyr Ala Glu Gln Ala Glu Pro Glu Gly Ile
 85 90 95

Ala Gln Ala Ile Thr Ile Gly Thr Asp His Ile Gly Asp Ser Pro Val
 100 105 110

Ala Leu Ile Leu Gly Asp Asn Ile Phe His Gly Pro Gly Phe Ser Ala
 115 120 125

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

Tyr Pro Val Ser Asp Pro Lys Arg Tyr Gly Val Gly Glu Ile Asp Asp
 145 150 155 160

Gln Gly Val Leu Leu Ser Leu Glu Glu Lys Pro Ala Arg Pro Arg Ser
 165 170 175

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

Ile Ala Lys Asn Ile Arg Pro Ser Ala Arg Gly Glu Leu Glu Ile Thr
 195 200 205

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

Leu Gly His Gly Phe Ala Trp Leu Asp Met Gly Thr His Asp Ser Leu
 225 230 235 240

Leu Gln Gly Gly Gln Tyr Val Gln Leu Ile Glu Gln Arg Gln Gly Val
 245 250 255

Arg Ile Ala Cys Ile Glu Glu Ile Ala Leu Arg Met Gly Phe Ile Asp
 260 265 270

Ala Asp Thr Leu His Arg Leu Gly Arg Glu Leu Gly Thr Ser Gly Tyr
 275 280 285

Gly Ala Tyr Leu Met Glu Val Ala Thr Arg Ala Gly Thr Glu
 290 295 300

<210> 32

<211> 299

<212> PRT

<213> Streptomyces ambofaciens

<400> 32

Val Met Ser Gly Arg Pro Gly Met Lys Gly Ile Ile Leu Ala Gly Gly
 1 5 10 15

Gly Gly Thr Arg Leu Arg Pro Leu Thr Gly Thr Leu Ser Lys Gln Leu
 20 25 30

Leu Pro Val Tyr Asp Lys Pro Met Ile Tyr Tyr Pro Leu Ser Val Leu
 35 40 45

Met Leu Gly Gly Ile Arg Glu Ile Leu Val Val Ser Ser Thr Gln His
 50 55 60

Ile Glu Leu Phe Gln Arg Leu Leu Gly Asp Gly Ser Arg Leu Gly Leu
 65 70 75 80

Asp Ile Thr Tyr Ala Glu Gln Ala Glu Pro Glu Gly Ile Ala Gln Ala
 85 90 95

Ile Thr Ile Gly Thr Asp His Ile Gly Asp Ser Pro Val Ala Leu Ile
 100 105 110

Leu Gly Asp Asn Ile Phe His Gly Pro Gly Phe Ser Ala Val Leu Gln
 115 120 125

Gly Ser Ile Arg His Leu Asp Gly Cys Val Leu Phe Gly Tyr Pro Val
 130 135 140

Ser Asp Pro Lys Arg Tyr Gly Val Gly Glu Ile Asp Asp Gln Gly Val
 145 150 155 160

Leu Leu Ser Leu Glu Glu Lys Pro Ala Arg Pro Arg Ser Asn Leu Ala
 165 170 175

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

Asn Ile Arg Pro Ser Ala Arg Gly Glu Leu Glu Ile Thr Asp Val Asn
 195 200 205

Arg Thr Tyr Leu Glu Gln Lys Arg Ala Arg Leu Ile Glu Leu Gly His
 210 215 220

Gly Phe Ala Trp Leu Asp Met Gly Thr His Asp Ser Leu Leu Gln Gly
 225 230 235 240

Gly Gln Tyr Val Gln Leu Ile Glu Gln Arg Gln Gly Val Arg Ile Ala
 245 250 255

Cys Ile Glu Glu Ile Ala Leu Arg Met Gly Phe Ile Asp Ala Asp Thr
 260 265 270

Leu His Arg Leu Gly Arg Glu Leu Gly Thr Ser Gly Tyr Gly Ala Tyr
 275 280 285

Leu Met Glu Val Ala Thr Arg Ala Gly Thr Glu
 290 295

<210> 33

<211> 293
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 33

Gly Met Lys Gly Ile Ile Leu Ala Gly Gly Gly Gly Thr Arg Leu Arg
 1 5 10 15

Pro Leu Thr Gly Thr Leu Ser Lys Gln Leu Leu Pro Val Tyr Asp Lys
 20 25 30

Pro Met Ile Tyr Tyr Pro Leu Ser Val Leu Met Leu Gly Gly Ile Arg
 35 40 45

Glu Ile Leu Val Val Ser Ser Thr Gln His Ile Glu Leu Phe Gln Arg
 50 55 60

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

Gln Ala Glu Pro Glu Gly Ile Ala Gln Ala Ile Thr Ile Gly Thr Asp
 85 90 95

His Ile Gly Asp Ser Pro Val Ala Leu Ile Leu Gly Asp Asn Ile Phe
 100 105 110

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

Asp Gly Cys Val Leu Phe Gly Tyr Pro Val Ser Asp Pro Lys Arg Tyr
 130 135 140

Gly Val Gly Glu Ile Asp Asp Gln Gly Val Leu Leu Ser Leu Glu Glu
 145 150 155 160

Lys Pro Ala Arg Pro Arg Ser Asn Leu Ala Val Thr Gly Leu Tyr Leu

	165		170		175
Tyr Asp Asn Asp Val Val Asp Ile Ala Lys Asn Ile Arg Pro Ser Ala					
	180		185		190
Arg Gly Glu Leu Glu Ile Thr Asp Val Asn Arg Thr Tyr Leu Glu Gln					
	195		200		205
Lys Arg Ala Arg Leu Ile Glu Leu Gly His Gly Phe Ala Trp Leu Asp					
	210		215		220
Met Gly Thr His Asp Ser Leu Leu Gln Gly Gly Gln Tyr Val Gln Leu					
	225		230		235
Ile Glu Gln Arg Gln Gly Val Arg Ile Ala Cys Ile Glu Glu Ile Ala					
	245		250		255
Leu Arg Met Gly Phe Ile Asp Ala Asp Thr Leu His Arg Leu Gly Arg					
	260		265		270
Glu Leu Gly Thr Ser Gly Tyr Gly Ala Tyr Leu Met Glu Val Ala Thr					
	275		280		285
Arg Ala Gly Thr Glu					
	290				

<210> 34
 <211> 1038
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1038)

<400> 34

gtg cag gca ccg aat gag acg ccg cgc cgg ccc gcc cgc tcc gcc ggc

Val	Gln	Ala	Pro	Asn	Glu	Thr	Pro	Arg	Arg	Pro	Ala	Arg	Ser	Ala	Gly	
1				5					10					15		
cga	cgg	ccg	ccg	gcc	cgg	atc	ctc	gtc	acc	ggg	ggc	gcc	ggc	ttc	atc	96
Arg	Arg	Pro	Pro	Ala	Arg	Ile	Leu	Val	Thr	Gly	Gly	Ala	Gly	Phe	Ile	
			20					25					30			
ggc	tcg	cgc	ttc	gtg	aac	gcg	ctg	ctg	gac	ggc	tcc	ctg	ccg	gag	ttc	144
Gly	Ser	Arg	Phe	Val	Asn	Ala	Leu	Leu	Asp	Gly	Ser	Leu	Pro	Glu	Phe	
		35					40					45				
ggc	aaa	ccc	gag	gtg	agg	gtg	ctc	gac	gcg	ctc	acc	tac	gcg	ggc	aac	192
Gly	Lys	Pro	Glu	Val	Arg	Val	Leu	Asp	Ala	Leu	Thr	Tyr	Ala	Gly	Asn	
	50					55					60					
ctg	gcc	aat	ctg	gcc	ccg	gtg	ggc	gac	tgt	ccc	cgg	ctg	cgg	atc	ttc	240
Leu	Ala	Asn	Leu	Ala	Pro	Val	Gly	Asp	Cys	Pro	Arg	Leu	Arg	Ile	Phe	
65					70				75					80		
ccg	ggg	gac	atc	cgc	gac	cgc	ggc	gcg	gtc	acc	cag	gcg	atg	gcg	ggg	288
Pro	Gly	Asp	Ile	Arg	Asp	Arg	Gly	Ala	Val	Thr	Gln	Ala	Met	Ala	Gly	
				85					90					95		
gtc	gac	ctg	gtg	gtg	cac	ttc	gcg	gcc	gag	tcg	cac	gtg	gac	cgc	tcg	336
Val	Asp	Leu	Val	Val	His	Phe	Ala	Ala	Glu	Ser	His	Val	Asp	Arg	Ser	
			100						105					110		
atc	gac	gac	gcc	gac	gcc	ttc	gtg	cgc	acc	aac	gtg	ctg	ggc	acc	cag	384
Ile	Asp	Asp	Ala	Asp	Ala	Phe	Val	Arg	Thr	Asn	Val	Leu	Gly	Thr	Gln	
			115					120					125			
gtc	ctc	ctc	cag	gag	gca	ctg	gcc	gta	cgc	ccc	ggg	ctg	ttc	gtg	cac	432
Val	Leu	Leu	Gln	Glu	Ala	Leu	Ala	Val	Arg	Pro	Gly	Leu	Phe	Val	His	
	130					135					140					
gtc	tcg	acg	gac	gag	gtg	tac	ggc	tcc	atc	gag	gag	ggg	tcc	tgg	ccc	480
Val	Ser	Thr	Asp	Glu	Val	Tyr	Gly	Ser	Ile	Glu	Glu	Gly	Ser	Trp	Pro	
145					150				155					160		
gag	gag	cac	ccg	ctg	aac	ccc	aac	tcg	ccc	tac	gcc	gcc	tcg	aag	gcg	528
Glu	Glu	His	Pro	Leu	Asn	Pro	Asn	Ser	Pro	Tyr	Ala	Ala	Ser	Lys	Ala	
				165					170					175		
tcc	tcc	gac	ctg	ctg	gcg	ctg	gcc	cac	cac	cgc	acg	cac	gga	ctg	ccg	576
Ser	Ser	Asp	Leu	Leu	Ala	Leu	Ala	His	His	Arg	Thr	His	Gly	Leu	Pro	
			180					185						190		

gtg tgc gtc acc cgc tgc tcc aac aac tac ggg ccc tac cag tac ccg 624
 Val Cys Val Thr Arg Cys Ser Asn Asn Tyr Gly Pro Tyr Gln Tyr Pro
 195 200 205

gag aag atc atc ccg ctg ttc acc agc agc ctc ctc gac ggc ggg acc 672
 Glu Lys Ile Ile Pro Leu Phe Thr Ser Ser Leu Leu Asp Gly Gly Thr
 210 215 220

gtc ccg ctc tac ggg gac ggc ggc aac cgg cgc gac tgg ctg cac gtg 720
 Val Pro Leu Tyr Gly Asp Gly Gly Asn Arg Arg Asp Trp Leu His Val
 225 230 235 240

gac gac cac tgc cgg ggc atc gcc ctg gtg gcc cgg ggc ggc cgg ccc 768
 Asp Asp His Cys Arg Gly Ile Ala Leu Val Ala Arg Gly Gly Arg Pro
 245 250 255

ggc gag gtc tac aac atc ggc ggc ggc acc gag ctg agc aac gtc gag 816
 Gly Glu Val Tyr Asn Ile Gly Gly Gly Thr Glu Leu Ser Asn Val Glu
 260 265 270

ctc acg gag cgt ctg ctg aaa ctg tgc gga gcc gac tgg tcg gcg gtg 864
 Leu Thr Glu Arg Leu Leu Lys Leu Cys Gly Ala Asp Trp Ser Ala Val
 275 280 285

cgg cgg gtg ccc gac cgc aag ggc cac gac cgg cgc tac tcc gtc gac 912
 Arg Arg Val Pro Asp Arg Lys Gly His Asp Arg Arg Tyr Ser Val Asp
 290 295 300

tac acc aag atc gcg gac gag ctg ggt tac gcg ccg cgg atc acc atc 960
 Tyr Thr Lys Ile Ala Asp Glu Leu Gly Tyr Ala Pro Arg Ile Thr Ile
 305 310 315 320

gac gaa ggg ctg gag cgg acc gtg cac tgg tac cgg gag aac cgc gcg 1008
 Asp Glu Gly Leu Glu Arg Thr Val His Trp Tyr Arg Glu Asn Arg Ala
 325 330 335

tgg tgg gcg ccc gcg aag agg ggg cga tga 1038
 Trp Trp Ala Pro Ala Lys Arg Gly Arg
 340 345

<210> 35

<211> 345

<212> PRT

<213> Streptomyces ambofaciens

<400> 35

Val Gln Ala Pro Asn Glu Thr Pro Arg Arg Pro Ala Arg Ser Ala Gly
 1 5 10 15

Arg Arg Pro Pro Ala Arg Ile Leu Val Thr Gly Gly Ala Gly Phe Ile
 20 25 30

Gly Ser Arg Phe Val Asn Ala Leu Leu Asp Gly Ser Leu Pro Glu Phe
 35 40 45

Gly Lys Pro Glu Val Arg Val Leu Asp Ala Leu Thr Tyr Ala Gly Asn
 50 55 60

Leu Ala Asn Leu Ala Pro Val Gly Asp Cys Pro Arg Leu Arg Ile Phe
 65 70 75 80

Pro Gly Asp Ile Arg Asp Arg Gly Ala Val Thr Gln Ala Met Ala Gly
 85 90 95

Val Asp Leu Val Val His Phe Ala Ala Glu Ser His Val Asp Arg Ser
 100 105 110

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

Val Leu Leu Gln Glu Ala Leu Ala Val Arg Pro Gly Leu Phe Val His
 130 135 140

Val Ser Thr Asp Glu Val Tyr Gly Ser Ile Glu Glu Gly Ser Trp Pro
 145 150 155 160

Glu Glu His Pro Leu Asn Pro Asn Ser Pro Tyr Ala Ala Ser Lys Ala
 165 170 175

Ser Ser Asp Leu Leu Ala Leu Ala His His Arg Thr His Gly Leu Pro
 180 185 190

Val Cys Val Thr Arg Cys Ser Asn Asn Tyr Gly Pro Tyr Gln Tyr Pro
 195 200 205

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

Val Pro Leu Tyr Gly Asp Gly Gly Asn Arg Arg Asp Trp Leu His Val
 225 230 235 240

Asp Asp His Cys Arg Gly Ile Ala Leu Val Ala Arg Gly Gly Arg Pro
 245 250 255

Gly Glu Val Tyr Asn Ile Gly Gly Gly Thr Glu Leu Ser Asn Val Glu
 260 265 270

Leu Thr Glu Arg Leu Leu Lys Leu Cys Gly Ala Asp Trp Ser Ala Val
 275 280 285

Arg Arg Val Pro Asp Arg Lys Gly His Asp Arg Arg Tyr Ser Val Asp
 290 295 300

Tyr Thr Lys Ile Ala Asp Glu Leu Gly Tyr Ala Pro Arg Ile Thr Ile
 305 310 315 320

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

Trp Trp Ala Pro Ala Lys Arg Gly Arg
 340 345

<210> 36

<211> 804

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1) .. (804)

<220>

<221> CDS2

<222> (7) .. (804)

<220>

<221> CDS3

<222> (22) .. (804)

<400> 36

atg	acg	gtg	acg	acc	gca	tcc	gtg	gac	ccg	ctc	gac	ctg	tgg	ctc	cgc	48
Met	Thr	Val	Thr	Thr	Ala	Ser	Val	Asp	Pro	Leu	Asp	Leu	Trp	Leu	Arg	
1				5					10					15		

cgg	tac	cag	ccg	tcc	gcg	tca	ccc	gcc	gtc	cgg	ctg	gtg	tgc	ttc	ccg	96
Arg	Tyr	Gln	Pro	Ser	Ala	Ser	Pro	Ala	Val	Arg	Leu	Val	Cys	Phe	Pro	
			20					25					30			

cac	gcg	ggc	ggc	tcg	gcg	agt	tcg	ttc	ctg	ccg	ttc	acc	cgg	cag	ctg	144
His	Ala	Gly	Gly	Ser	Ala	Ser	Ser	Phe	Leu	Pro	Phe	Thr	Arg	Gln	Leu	
		35					40					45				

ccg	gac	cgg	atc	gag	gtc	gtg	gcc	gtc	cag	tac	ccc	ggg	cgc	cag	gac	192
Pro	Asp	Arg	Ile	Glu	Val	Val	Ala	Val	Gln	Tyr	Pro	Gly	Arg	Gln	Asp	
	50					55				60						

cgc	agg	agc	gaa	ccg	ctg	gtc	gac	acc	atc	gag	gga	ctg	gcc	gag	ccc	240
Arg	Arg	Ser	Glu	Pro	Leu	Val	Asp	Thr	Ile	Glu	Gly	Leu	Ala	Glu	Pro	
65					70				75					80		

ctg	gcc	ggc	ctg	ctg	gag	gcg	cag	gcc	ggc	ccc	ccg	gtg	gtg	ctg	ttc	288
Leu	Ala	Gly	Leu	Leu	Glu	Ala	Gln	Ala	Gly	Pro	Pro	Val	Val	Leu	Phe	
			85						90					95		

ggg	cac	agc	atg	ggc	gcg	ctg	gtg	gcc	tac	gag	gtc	gcc	cgc	gcg	ctc	336
Gly	His	Ser	Met	Gly	Ala	Leu	Val	Ala	Tyr	Glu	Val	Ala	Arg	Ala	Leu	
			100					105					110			

cag	cgg	cgg	gga	gcg	gct	ccg	gtg	cgc	ctg	gtg	gtc	tcc	ggg	cgc	cgg	384
Gln	Arg	Arg	Gly	Ala	Ala	Pro	Val	Arg	Leu	Val	Val	Ser	Gly	Arg	Arg	
			115				120					125				

gcc	ccc	gcc	gtc	gac	cgg	ccg	atg	acc	gtg	cac	ctc	tac	gac	gac	gac	432

Ala	Pro	Ala	Val	Asp	Arg	Pro	Met	Thr	Val	His	Leu	Tyr	Asp	Asp	Asp				
130						135					140								
cg	g	g	g	g	a	c	c	a	a	c	c	a	a	c	a	c	g	t	g
Arg	Leu	Val	Glu	Glu	Leu	Arg	Lys	Leu	Asp	Gly	Thr	Asp	Ser	Gln	Val				
145					150					155					160				
ttc	gcc	gat	ccg	gag	ctg	ctc	cgg	ctg	gtg	ctg	ccc	gtg	atc	cgc	aac				
Phe	Ala	Asp	Pro	Glu	Leu	Leu	Arg	Leu	Val	Leu	Pro	Val	Ile	Arg	Asn				
				165					170					175					
gac	tac	cgg	gcc	gtg	gcg	gcc	tac	gcc	cac	cgc	ccg	ggg	gcg	ccg	ctg				
Asp	Tyr	Arg	Ala	Val	Ala	Ala	Tyr	Ala	His	Arg	Pro	Gly	Ala	Pro	Leu				
				180					185					190					
gac	tgc	ccc	ctc	acc	gtg	ttc	acc	ggc	gcc	gac	gac	ccc	acc	gtg	acc				
Asp	Cys	Pro	Leu	Thr	Val	Phe	Thr	Gly	Ala	Asp	Asp	Pro	Thr	Val	Thr				
		195					200					205							
gcg	gcc	gag	gcg	gcg	gcc	tgg	cac	gag	gcg	gcg	gcg	tcc	gac	gtc	gag				
Ala	Ala	Glu	Ala	Ala	Ala	Trp	His	Glu	Ala	Ala	Ala	Ser	Asp	Val	Glu				
		210					215					220							
acg	cgc	acc	ttc	ccc	ggt	ggc	cac	ttc	ttc	ccg	tac	cag	cgg	acc	gcg				
Thr	Arg	Thr	Phe	Pro	Gly	Gly	His	Phe	Phe	Pro	Tyr	Gln	Arg	Thr	Ala				
225					230					235					240				
gag	gtg	tgc	ggg	gcc	ctg	gtc	gac	acg	ctc	gag	ccg	ctg	ctg	tcg	gcc				
Glu	Val	Cys	Gly	Ala	Leu	Val	Asp	Thr	Leu	Glu	Pro	Leu	Leu	Ser	Ala				
				245					250					255					
ggg	acg	cgc	ggt	gtc	cgg	cgg	gtc	cgc	ccg	ggg	tga								
Gly	Thr	Arg	Gly	Val	Arg	Arg	Val	Arg	Pro	Gly									
				260				265											

<210> 37

<211> 267

<212> PRT

<213> Streptomyces ambofaciens

<400> 37

Met	Thr	Val	Thr	Thr	Ala	Ser	Val	Asp	Pro	Leu	Asp	Leu	Trp	Leu	Arg
1				5					10					15	

Arg Tyr Gln Pro Ser Ala Ser Pro Ala Val Arg Leu Val Cys Phe Pro
 20 25 30

His Ala Gly Gly Ser Ala Ser Ser Phe Leu Pro Phe Thr Arg Gln Leu
 35 40 45

Pro Asp Arg Ile Glu Val Val Ala Val Gln Tyr Pro Gly Arg Gln Asp
 50 55 60

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

Leu Ala Gly Leu Leu Glu Ala Gln Ala Gly Pro Pro Val Val Leu Phe
 85 90 95

Gly His Ser Met Gly Ala Leu Val Ala Tyr Glu Val Ala Arg Ala Leu
 100 105 110

Gln Arg Arg Gly Ala Ala Pro Val Arg Leu Val Val Ser Gly Arg Arg
 115 120 125

Ala Pro Ala Val Asp Arg Pro Met Thr Val His Leu Tyr Asp Asp Asp
 130 135 140

Arg Leu Val Glu Glu Leu Arg Lys Leu Asp Gly Thr Asp Ser Gln Val
 145 150 155 160

Phe Ala Asp Pro Glu Leu Leu Arg Leu Val Leu Pro Val Ile Arg Asn
 165 170 175

Asp Tyr Arg Ala Val Ala Ala Tyr Ala His Arg Pro Gly Ala Pro Leu
 180 185 190

Asp Cys Pro Leu Thr Val Phe Thr Gly Ala Asp Asp Pro Thr Val Thr
 195 200 205

Ala Ala Glu Ala Ala Ala Trp His Glu Ala Ala Ala Ser Asp Val Glu
 210 215 220

Thr Arg Thr Phe Pro Gly Gly His Phe Phe Pro Tyr Gln Arg Thr Ala
 225 230 235 240

Glu Val Cys Gly Ala Leu Val Asp Thr Leu Glu Pro Leu Leu Ser Ala
 245 250 255

Gly Thr Arg Gly Val Arg Arg Val Arg Pro Gly
 260 265

<210> 38
 <211> 265
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 38

Val Thr Thr Ala Ser Val Asp Pro Leu Asp Leu Trp Leu Arg Arg Tyr
 1 5 10 15

Gln Pro Ser Ala Ser Pro Ala Val Arg Leu Val Cys Phe Pro His Ala
 20 25 30

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

Arg Ile Glu Val Val Ala Val Gln Tyr Pro Gly Arg Gln Asp Arg Arg
 50 55 60

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

Gly Leu Leu Glu Ala Gln Ala Gly Pro Pro Val Val Leu Phe Gly His
 85 90 95

Ser Met Gly Ala Leu Val Ala Tyr Glu Val Ala Arg Ala Leu Gln Arg
 100 105 110

Arg Gly Ala Ala Pro Val Arg Leu Val Val Ser Gly Arg Arg Ala Pro
 115 120 125

Ala Val Asp Arg Pro Met Thr Val His Leu Tyr Asp Asp Asp Arg Leu
 130 135 140

Val Glu Glu Leu Arg Lys Leu Asp Gly Thr Asp Ser Gln Val Phe Ala
 145 150 155 160

Asp Pro Glu Leu Leu Arg Leu Val Leu Pro Val Ile Arg Asn Asp Tyr
 165 170 175

Arg Ala Val Ala Ala Tyr Ala His Arg Pro Gly Ala Pro Leu Asp Cys
 180 185 190

Pro Leu Thr Val Phe Thr Gly Ala Asp Asp Pro Thr Val Thr Ala Ala
 195 200 205

Glu Ala Ala Ala Trp His Glu Ala Ala Ala Ser Asp Val Glu Thr Arg
 210 215 220

Thr Phe Pro Gly Gly His Phe Phe Pro Tyr Gln Arg Thr Ala Glu Val
 225 230 235 240

Cys Gly Ala Leu Val Asp Thr Leu Glu Pro Leu Leu Ser Ala Gly Thr
 245 250 255

Arg Gly Val Arg Arg Val Arg Pro Gly
 260 265

<210> 39

<211> 260

<212> PRT

<213> Streptomyces ambofaciens

<400> 39

Val	Asp	Pro	Leu	Asp	Leu	Trp	Leu	Arg	Arg	Tyr	Gln	Pro	Ser	Ala	Ser
1				5					10					15	

Pro	Ala	Val	Arg	Leu	Val	Cys	Phe	Pro	His	Ala	Gly	Gly	Ser	Ala	Ser
			20					25					30		

Ser	Phe	Leu	Pro	Phe	Thr	Arg	Gln	Leu	Pro	Asp	Arg	Ile	Glu	Val	Val
	35						40					45			

Ala	Val	Gln	Tyr	Pro	Gly	Arg	Gln	Asp	Arg	Arg	Ser	Glu	Pro	Leu	Val
	50					55					60				

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

Gln	Ala	Gly	Pro	Pro	Val	Val	Leu	Phe	Gly	His	Ser	Met	Gly	Ala	Leu
			85						90					95	

Val	Ala	Tyr	Glu	Val	Ala	Arg	Ala	Leu	Gln	Arg	Arg	Gly	Ala	Ala	Pro
		100						105					110		

Val	Arg	Leu	Val	Val	Ser	Gly	Arg	Arg	Ala	Pro	Ala	Val	Asp	Arg	Pro
		115					120						125		

Met	Thr	Val	His	Leu	Tyr	Asp	Asp	Asp	Arg	Leu	Val	Glu	Glu	Leu	Arg
	130						135					140			

Lys	Leu	Asp	Gly	Thr	Asp	Ser	Gln	Val	Phe	Ala	Asp	Pro	Glu	Leu	Leu
145					150					155				160	

Arg	Leu	Val	Leu	Pro	Val	Ile	Arg	Asn	Asp	Tyr	Arg	Ala	Val	Ala	Ala
			165						170					175	

Tyr Ala His Arg Pro Gly Ala Pro Leu Asp Cys Pro Leu Thr Val Phe
 180 185 190

Thr Gly Ala Asp Asp Pro Thr Val Thr Ala Ala Glu Ala Ala Ala Trp
 195 200 205

His Glu Ala Ala Ala Ser Asp Val Glu Thr Arg Thr Phe Pro Gly Gly
 210 215 220

His Phe Phe Pro Tyr Gln Arg Thr Ala Glu Val Cys Gly Ala Leu Val
 225 230 235 240

Asp Thr Leu Glu Pro Leu Leu Ser Ala Gly Thr Arg Gly Val Arg Arg
 245 250 255

Val Arg Pro Gly
 260

<210> 40
 <211> 1410
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1410)

<220>
 <221> CDS2
 <222> (46)..(1410)

<400> 40
 gtg ggc acg gtc gag tac gcc gtc cac cgg cgt acc gcg gaa cgg gtg 48
 Val Gly Thr Val Glu Tyr Ala Val His Arg Arg Thr Ala Glu Arg Val
 1 5 10 15

agg gtc tcc gcc gac acc ctg gac agc ccg gtc acc gcg ctg gcg gag 96
 Arg Val Ser Ala Asp Thr Leu Asp Ser Pro Val Thr Ala Leu Ala Glu

20	25	30	
gtg ccc cgc tgg ctg gag gaa tac cac cgg gcg cac cgc ttc cac gtc			144
Val Pro Arg Trp Leu Glu Glu Tyr His Arg Ala His Arg Phe His Val			
35	40	45	
gag ccg atc ccc ttc gac cgg ctc cgg cgg tgg tcc ttc gag ccg ggc			192
Glu Pro Ile Pro Phe Asp Arg Leu Arg Arg Trp Ser Phe Glu Pro Gly			
50	55	60	
acc ggc gac ctg cgg cac gag acg ggc cgc ttc ttc tcc gtg gag ggg			240
Thr Gly Asp Leu Arg His Glu Thr Gly Arg Phe Phe Ser Val Glu Gly			
65	70	75	80
ctg cgc acc agc tcg gac gcc gat ccg gtc gcc cgt gtc cag ccg atc			288
Leu Arg Thr Ser Ser Asp Ala Asp Pro Val Ala Arg Val Gln Pro Ile			
85	90	95	
atc gtg cag ccc gag gtg ggg ctg ctc ggc atc ctg gcc cgg gag ttc			336
Ile Val Gln Pro Glu Val Gly Leu Leu Gly Ile Leu Ala Arg Glu Phe			
100	105	110	
gac ggg gtg ctg cac ttc ctg atg cag gcc aaa ccc gag ccc ggc aac			384
Asp Gly Val Leu His Phe Leu Met Gln Ala Lys Pro Glu Pro Gly Asn			
115	120	125	
gtc aac ggg ctg cag atc tcc ccc acg gtg cag gcc acg cgc agc aac			432
Val Asn Gly Leu Gln Ile Ser Pro Thr Val Gln Ala Thr Arg Ser Asn			
130	135	140	
ttc gac gag gtg cac cac ggc cgg tcc acc ccg ttc ctc gac cac ttc			480
Phe Asp Glu Val His His Gly Arg Ser Thr Pro Phe Leu Asp His Phe			
145	150	155	160
atc cac cgc ccc ggc cgc cgg gtc ctg atc gac agc atc cag tcc gaa			528
Ile His Arg Pro Gly Arg Arg Val Leu Ile Asp Ser Ile Gln Ser Glu			
165	170	175	
cag ggc gac tgg ttc ctg cac aag cgc aac cgc aac atg gtc gtc gag			576
Gln Gly Asp Trp Phe Leu His Lys Arg Asn Arg Asn Met Val Val Glu			
180	185	190	
atc gac acc gac atc gag gcc gac gcc gcg ttc cgc tgg ctg acc ctc			624
Ile Asp Thr Asp Ile Glu Ala Asp Ala Ala Phe Arg Trp Leu Thr Leu			
195	200	205	
ggg cag atc cgc cgg ctg atg ctc cag gac gac ctc gtc aac atg gac			672

Gly	Gln	Ile	Arg	Arg	Leu	Met	Leu	Gln	Asp	Asp	Leu	Val	Asn	Met	Asp	
210						215					220					
acc	cgc	agt	gtg	ctg	gcc	tgt	ctg	ccc	acc	gcg	cac	ggc	acg	ccc	gac	720
Thr	Arg	Ser	Val	Leu	Ala	Cys	Leu	Pro	Thr	Ala	His	Gly	Thr	Pro	Asp	
225					230				235					240		
gac	ggc	gac	gac	tcc	ttc	ccg	gcg	gcg	ctg	cgc	cgc	tcc	ctc	tac	ggg	768
Asp	Gly	Asp	Asp	Ser	Phe	Pro	Ala	Ala	Leu	Arg	Arg	Ser	Leu	Tyr	Gly	
				245					250					255		
gag	acc	gcg	ccg	ttg	cac	gat	ctg	cac	gcc	atc	acc	agc	tgc	ctc	acc	816
Glu	Thr	Ala	Pro	Leu	His	Asp	Leu	His	Ala	Ile	Thr	Ser	Cys	Leu	Thr	
			260						265					270		
gac	gtc	cgg	gcg	ctg	cgg	gtg	ctg	cgc	cag	cag	agc	gtg	ccg	ctc	gac	864
Asp	Val	Arg	Ala	Leu	Arg	Val	Leu	Arg	Gln	Gln	Ser	Val	Pro	Leu	Asp	
		275						280					285			
gac	gcc	cgg	cgg	gac	ggc	tgg	gag	cgg	acc	ggg	agc	gcg	atc	cgg	cat	912
Asp	Ala	Arg	Arg	Asp	Gly	Trp	Glu	Arg	Thr	Gly	Ser	Ala	Ile	Arg	His	
	290					295					300					
cgc	agc	ggc	agg	cat	ttc	gag	atc	atg	gcg	gtg	gag	gtg	acc	gcg	gag	960
Arg	Ser	Gly	Arg	His	Phe	Glu	Ile	Met	Ala	Val	Glu	Val	Thr	Ala	Glu	
305					310						315				320	
cgc	cgt	gaa	gtg	gcc	tcg	tgg	acc	cag	ccg	ttg	ctg	cgc	ccg	tgc	tcg	1008
Arg	Arg	Glu	Val	Ala	Ser	Trp	Thr	Gln	Pro	Leu	Leu	Arg	Pro	Cys	Ser	
				325						330				335		
cag	gga	ctg	gcg	gcc	ctg	atc	acc	cgg	cgg	atc	aac	ggg	gtg	ctg	cac	1056
Gln	Gly	Leu	Ala	Ala	Leu	Ile	Thr	Arg	Arg	Ile	Asn	Gly	Val	Leu	His	
			340					345					350			
gcc	ctg	gtg	gcg	gcg	cgg	tcg	gag	gtc	ggc	acg	ctc	aac	gtc	gcc	gag	1104
Ala	Leu	Val	Ala	Ala	Arg	Ser	Glu	Val	Gly	Thr	Leu	Asn	Val	Ala	Glu	
		355						360					365			
ttc	gga	ccg	acc	gtc	cag	tgc	cgg	ccc	gac	gag	gcg	gac	ggc	cag	tcg	1152
Phe	Gly	Pro	Thr	Val	Gln	Cys	Arg	Pro	Asp	Glu	Ala	Asp	Gly	Gln	Ser	
	370					375						380				
ccc	ccg	tac	ctg	gac	cgg	gtg	ctg	acg	gcc	gga	gcc	gac	cgc	gtc	cgc	1200
Pro	Pro	Tyr	Leu	Asp	Arg	Val	Leu	Thr	Ala	Gly	Ala	Asp	Arg	Val	Arg	
385					390					395					400	

tac gac gtg gtg cag tcg gag gag ggc ggg cgc ttc tac cac gcg cgc 1248
 Tyr Asp Val Val Gln Ser Glu Glu Gly Gly Arg Phe Tyr His Ala Arg
 405 410 415

aac cgc tat ctg gtg gtc gag gcg ggg ccg gag ctc gac acg ggc tgc 1296
 Asn Arg Tyr Leu Val Val Glu Ala Gly Pro Glu Leu Asp Thr Gly Cys
 420 425 430

ccg ccc ggc ttc tgc tgg gct acc ttc ggc cag ctc acc gaa ctg ctc 1344
 Pro Pro Gly Phe Cys Trp Ala Thr Phe Gly Gln Leu Thr Glu Leu Leu
 435 440 445

gcg cac ggc aac tat ctc aac gtc gaa ctc cgc acc ctc atg gcg tgc 1392
 Ala His Gly Asn Tyr Leu Asn Val Glu Leu Arg Thr Leu Met Ala Cys
 450 455 460

gca cac gcc tcc tac tga 1410
 Ala His Ala Ser Tyr
 465

<210> 41
 <211> 469
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 41

Val Gly Thr Val Glu Tyr Ala Val His Arg Arg Thr Ala Glu Arg Val
 1 5 10 15

Arg Val Ser Ala Asp Thr Leu Asp Ser Pro Val Thr Ala Leu Ala Glu
 20 25 30

Val Pro Arg Trp Leu Glu Glu Tyr His Arg Ala His Arg Phe His Val
 35 40 45

Glu Pro Ile Pro Phe Asp Arg Leu Arg Arg Trp Ser Phe Glu Pro Gly
 50 55 60

Thr Gly Asp Leu Arg His Glu Thr Gly Arg Phe Phe Ser Val Glu Gly
 65 70 75 80

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

Ile Val Gln Pro Glu Val Gly Leu Leu Gly Ile Leu Ala Arg Glu Phe
 100 105 110

Asp Gly Val Leu His Phe Leu Met Gln Ala Lys Pro Glu Pro Gly Asn
 115 120 125

Val Asn Gly Leu Gln Ile Ser Pro Thr Val Gln Ala Thr Arg Ser Asn
 130 135 140

Phe Asp Glu Val His His Gly Arg Ser Thr Pro Phe Leu Asp His Phe
 145 150 155 160

Ile His Arg Pro Gly Arg Arg Val Leu Ile Asp Ser Ile Gln Ser Glu
 165 170 175

Gln Gly Asp Trp Phe Leu His Lys Arg Asn Arg Asn Met Val Val Glu
 180 185 190

Ile Asp Thr Asp Ile Glu Ala Asp Ala Ala Phe Arg Trp Leu Thr Leu
 195 200 205

Gly Gln Ile Arg Arg Leu Met Leu Gln Asp Asp Leu Val Asn Met Asp
 210 215 220

Thr Arg Ser Val Leu Ala Cys Leu Pro Thr Ala His Gly Thr Pro Asp
 225 230 235 240

Asp Gly Asp Asp Ser Phe Pro Ala Ala Leu Arg Arg Ser Leu Tyr Gly
 245 250 255

Glu Thr Ala Pro Leu His Asp Leu His Ala Ile Thr Ser Cys Leu Thr
 260 265 270

Asp Val Arg Ala Leu Arg Val Leu Arg Gln Gln Ser Val Pro Leu Asp
 275 280 285

Asp Ala Arg Arg Asp Gly Trp Glu Arg Thr Gly Ser Ala Ile Arg His
 290 295 300

Arg Ser Gly Arg His Phe Glu Ile Met Ala Val Glu Val Thr Ala Glu
 305 310 315 320

Arg Arg Glu Val Ala Ser Trp Thr Gln Pro Leu Leu Arg Pro Cys Ser
 325 330 335

Gln Gly Leu Ala Ala Leu Ile Thr Arg Arg Ile Asn Gly Val Leu His
 340 345 350

Ala Leu Val Ala Ala Arg Ser Glu Val Gly Thr Leu Asn Val Ala Glu
 355 360 365

Phe Gly Pro Thr Val Gln Cys Arg Pro Asp Glu Ala Asp Gly Gln Ser
 370 375 380

Pro Pro Tyr Leu Asp Arg Val Leu Thr Ala Gly Ala Asp Arg Val Arg
 385 390 395 400

Tyr Asp Val Val Gln Ser Glu Glu Gly Gly Arg Phe Tyr His Ala Arg
 405 410 415

Asn Arg Tyr Leu Val Val Glu Ala Gly Pro Glu Leu Asp Thr Gly Cys
 420 425 430

Pro Pro Gly Phe Cys Trp Ala Thr Phe Gly Gln Leu Thr Glu Leu Leu
 435 440 445

Ala His Gly Asn Tyr Leu Asn Val Glu Leu Arg Thr Leu Met Ala Cys

450

455

460

Ala His Ala Ser Tyr

465

<210> 42

<211> 454

<212> PRT

<213> Streptomyces ambofaciens

<400> 42

Val	Arg	Val	Ser	Ala	Asp	Thr	Leu	Asp	Ser	Pro	Val	Thr	Ala	Leu	Ala
1				5					10					15	

Glu	Val	Pro	Arg	Trp	Leu	Glu	Glu	Tyr	His	Arg	Ala	His	Arg	Phe	His
			20					25					30		

Val	Glu	Pro	Ile	Pro	Phe	Asp	Arg	Leu	Arg	Arg	Trp	Ser	Phe	Glu	Pro
		35						40				45			

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

Gly	Leu	Arg	Thr	Ser	Ser	Asp	Ala	Asp	Pro	Val	Ala	Arg	Val	Gln	Pro
65					70					75				80	

Ile	Ile	Val	Gln	Pro	Glu	Val	Gly	Leu	Leu	Gly	Ile	Leu	Ala	Arg	Glu
			85						90					95	

Phe	Asp	Gly	Val	Leu	His	Phe	Leu	Met	Gln	Ala	Lys	Pro	Glu	Pro	Gly
			100					105					110		

Asn	Val	Asn	Gly	Leu	Gln	Ile	Ser	Pro	Thr	Val	Gln	Ala	Thr	Arg	Ser
		115					120					125			

Asn	Phe	Asp	Glu	Val	His	His	Gly	Arg	Ser	Thr	Pro	Phe	Leu	Asp	His

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

Ser Gln Gly Leu Ala Ala Leu Ile Thr Arg Arg Ile Asn Gly Val Leu
 325 330 335

His Ala Leu Val Ala Ala Arg Ser Glu Val Gly Thr Leu Asn Val Ala
 340 345 350

Glu Phe Gly Pro Thr Val Gln Cys Arg Pro Asp Glu Ala Asp Gly Gln
 355 360 365

Ser Pro Pro Tyr Leu Asp Arg Val Leu Thr Ala Gly Ala Asp Arg Val
 370 375 380

Arg Tyr Asp Val Val Gln Ser Glu Glu Gly Gly Arg Phe Tyr His Ala
 385 390 395 400

Arg Asn Arg Tyr Leu Val Val Glu Ala Gly Pro Glu Leu Asp Thr Gly
 405 410 415

Cys Pro Pro Gly Phe Cys Trp Ala Thr Phe Gly Gln Leu Thr Glu Leu
 420 425 430

Leu Ala His Gly Asn Tyr Leu Asn Val Glu Leu Arg Thr Leu Met Ala
 435 440 445

Cys Ala His Ala Ser Tyr
 450

<210> 43
 <211> 1248
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1) .. (1248)

<400> 43

atg ccg ggc ggg acg gac ggc gac tgc gcg cgg acg gcg gcc cgg cgt	48
Met Pro Gly Gly Thr Asp Gly Asp Cys Ala Arg Thr Ala Ala Arg Arg	
1 5 10 15	
cga acg cac ctg ccc gag tcc gga cga gac agc gcg acg cga gag gcg	96
Arg Thr His Leu Pro Glu Ser Gly Arg Asp Ser Ala Thr Arg Glu Ala	
20 25 30	
aaa atg atc aat ctc ttc cag ccc cag atg ggg gcc gag gaa ctg gcg	144
Lys Met Ile Asn Leu Phe Gln Pro Gln Met Gly Ala Glu Glu Leu Ala	
35 40 45	
gcg gtg tcc gag gtc ttc gac gac caa tgg ctc ggt cac gga ccc cgg	192
Ala Val Ser Glu Val Phe Asp Asp Gln Trp Leu Gly His Gly Pro Arg	
50 55 60	
acc gcg gcg ttc gag tcc gcg ttc gcc gag cac ctc ggg gtc ggc ccc	240
Thr Ala Ala Phe Glu Ser Ala Phe Ala Glu His Leu Gly Val Gly Pro	
65 70 75 80	
gag cac gtc gtc ttc ctc aac tcg ggc acc gcc ggc ctc ttc ctg gcc	288
Glu His Val Val Phe Leu Asn Ser Gly Thr Ala Gly Leu Phe Leu Ala	
85 90 95	
ctg gag tcg ctc ggc ctg cgg ccc ggc gac gag gtc gtg ctc ccc tcg	336
Leu Glu Ser Leu Gly Leu Arg Pro Gly Asp Glu Val Val Leu Pro Ser	
100 105 110	
ccc agc ttc ctc gcc gcg gcg aac gcc gta cag ctc tcg gga gcg cgc	384
Pro Ser Phe Leu Ala Ala Ala Asn Ala Val Gln Leu Ser Gly Ala Arg	
115 120 125	
ccg gtg ttc tgc gac acc gac ccg cgg acg ctg aac ccc gcc ctg gag	432
Pro Val Phe Cys Asp Thr Asp Pro Arg Thr Leu Asn Pro Ala Leu Glu	
130 135 140	
cac atc gag gcg gcc gtc acc ccg cgc acc agg gcc gtc atc gcg ctc	480
His Ile Glu Ala Ala Val Thr Pro Arg Thr Arg Ala Val Ile Ala Leu	
145 150 155 160	
cac tac ggc ggc cac ccc ggc gac atc gtg cgc atc gcc gag cgc tgc	528
His Tyr Gly Gly His Pro Gly Asp Ile Val Arg Ile Ala Glu Arg Cys	
165 170 175	
cgg gag cgg ggc atc acc ctg atc gag gac gcc gcg tgc tcc gtg gcc	576
Arg Glu Arg Gly Ile Thr Leu Ile Glu Asp Ala Ala Cys Ser Val Ala	
180 185 190	

tcc cgc gtc gac ggc cga ccg gtc ggc acc ttc ggc gac ctc gcc atg	624
Ser Arg Val Asp Gly Arg Pro Val Gly Thr Phe Gly Asp Leu Ala Met	
195 200 205	
tgg agc ttc gac gcc atg aag gtc ctg gtc acc ggc gac gga ggg atg	672
Trp Ser Phe Asp Ala Met Lys Val Leu Val Thr Gly Asp Gly Gly Met	
210 215 220	
atc tac gtc aag gac ccc ggg gcg gcc gcc cgg atc cgg cgc ctc gcc	720
Ile Tyr Val Lys Asp Pro Gly Ala Ala Ala Arg Ile Arg Arg Leu Ala	
225 230 235 240	
tac cac ggc ctc acg cgg tcc agc ggc ctg gga tac gcc agg gtc tcg	768
Tyr His Gly Leu Thr Arg Ser Ser Gly Leu Gly Tyr Ala Arg Val Ser	
245 250 255	
gcg cgc tgg tgg gag atg gac gtc ccc gaa ccg ggc cgc cgc gtc atc	816
Ala Arg Trp Trp Glu Met Asp Val Pro Glu Pro Gly Arg Arg Val Ile	
260 265 270	
ggg aac gac ctc acc gcg gcc atc ggc gcg gtc cag ttg cgc cgg ctt	864
Gly Asn Asp Leu Thr Ala Ala Ile Gly Ala Val Gln Leu Arg Arg Leu	
275 280 285	
ccc ggc ttc gtg gcc cgc cgc agg gag atc gtc gcc ctg tac gac agc	912
Pro Gly Phe Val Ala Arg Arg Arg Glu Ile Val Ala Leu Tyr Asp Ser	
290 295 300	
gaa ctg agc tcg ctg gag ggc gtg ctg aca ccg ccc gcg cca ccc gcg	960
Glu Leu Ser Ser Leu Glu Gly Val Leu Thr Pro Pro Ala Pro Pro Ala	
305 310 315 320	
ggg cac gag tcc acg cac tac ttc tac tgg atc cag ctg gcc ccc ggc	1008
Gly His Glu Ser Thr His Tyr Phe Tyr Trp Ile Gln Leu Ala Pro Gly	
325 330 335	
gtc cgg gac cgg gtg gca cgc gac ctg ctc acc gac ggc atc tac acc	1056
Val Arg Asp Arg Val Ala Arg Asp Leu Leu Thr Asp Gly Ile Tyr Thr	
340 345 350	
acc ttc cgc tac gca cct ctg cac aag gtg ccc gcc tac ggc cac acc	1104
Thr Phe Arg Tyr Ala Pro Leu His Lys Val Pro Ala Tyr Gly His Thr	
355 360 365	
gga ggc gaa ctg ccc ggc gtg gag cgg gcg tcc gaa cgg acc ctg tgc	1152
Gly Gly Glu Leu Pro Gly Val Glu Arg Ala Ser Glu Arg Thr Leu Cys	

370 375 380
 ctg ccc ctg cac ccc ggc ctg tcg gac gcc gac gtc cgc acc gtc gtg 1200
 Leu Pro Leu His Pro Gly Leu Ser Asp Ala Asp Val Arg Thr Val Val
 385 390 395 400

 tcc tcc ctg cgc aga gcc ctg agc gcc gcg gat ccg gcc ccc gcc tga 1248
 Ser Ser Leu Arg Arg Ala Leu Ser Ala Ala Asp Pro Ala Pro Ala
 405 410 415

 <210> 44
 <211> 415
 <212> PRT
 <213> Streptomyces ambofaciens

 <400> 44

 Met Pro Gly Gly Thr Asp Gly Asp Cys Ala Arg Thr Ala Ala Arg Arg
 1 5 10 15

 Arg Thr His Leu Pro Glu Ser Gly Arg Asp Ser Ala Thr Arg Glu Ala
 20 25 30

 Lys Met Ile Asn Leu Phe Gln Pro Gln Met Gly Ala Glu Glu Leu Ala
 35 40 45

 Ala Val Ser Glu Val Phe Asp Asp Gln Trp Leu Gly His Gly Pro Arg
 50 55 60

 Thr Ala Ala Phe Glu Ser Ala Phe Ala Glu His Leu Gly Val Gly Pro
 65 70 75 80

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

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

 Pro Ser Phe Leu Ala Ala Ala Asn Ala Val Gln Leu Ser Gly Ala Arg

115

120

125

Pro Val Phe Cys Asp Thr Asp Pro Arg Thr Leu Asn Pro Ala Leu Glu
 130 135 140

His Ile Glu Ala Ala Val Thr Pro Arg Thr Arg Ala Val Ile Ala Leu
 145 150 155 160

His Tyr Gly Gly His Pro Gly Asp Ile Val Arg Ile Ala Glu Arg Cys
 165 170 175

Arg Glu Arg Gly Ile Thr Leu Ile Glu Asp Ala Ala Cys Ser Val Ala
 180 185 190

Ser Arg Val Asp Gly Arg Pro Val Gly Thr Phe Gly Asp Leu Ala Met
 195 200 205

Trp Ser Phe Asp Ala Met Lys Val Leu Val Thr Gly Asp Gly Gly Met
 210 215 220

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

Tyr His Gly Leu Thr Arg Ser Ser Gly Leu Gly Tyr Ala Arg Val Ser
 245 250 255

Ala Arg Trp Trp Glu Met Asp Val Pro Glu Pro Gly Arg Arg Val Ile
 260 265 270

Gly Asn Asp Leu Thr Ala Ala Ile Gly Ala Val Gln Leu Arg Arg Leu
 275 280 285

Pro Gly Phe Val Ala Arg Arg Arg Glu Ile Val Ala Leu Tyr Asp Ser
 290 295 300

Glu Leu Ser Ser Leu Glu Gly Val Leu Thr Pro Pro Ala Pro Pro Ala
 305 310 315 320

Gly His Glu Ser Thr His Tyr Phe Tyr Trp Ile Gln Leu Ala Pro Gly
 325 330 335

Val Arg Asp Arg Val Ala Arg Asp Leu Leu Thr Asp Gly Ile Tyr Thr
 340 345 350

Thr Phe Arg Tyr Ala Pro Leu His Lys Val Pro Ala Tyr Gly His Thr
 355 360 365

Gly Gly Glu Leu Pro Gly Val Glu Arg Ala Ser Glu Arg Thr Leu Cys
 370 375 380

Leu Pro Leu His Pro Gly Leu Ser Asp Ala Asp Val Arg Thr Val Val
 385 390 395 400

Ser Ser Leu Arg Arg Ala Leu Ser Ala Ala Asp Pro Ala Pro Ala
 405 410 415

<210> 45
 <211> 720
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(720)

<400> 45
 atg tac gag aac gac agt gcc gcc gag gtc tac gac ctg ctc tac cag 48
 Met Tyr Glu Asn Asp Ser Ala Ala Glu Val Tyr Asp Leu Leu Tyr Gln
 1 5 10 15
 gac cgc aag gac tac gcg ggt gag gcc gcc cgg gtc acc gac ctg atc 96
 Asp Arg Lys Asp Tyr Ala Gly Glu Ala Ala Arg Val Thr Asp Leu Ile
 20 25 30

cgg gaa cgt acg ccg gac gcg gcc agc ctg ctc gac atc gcg tgc ggc	144
Arg Glu Arg Thr Pro Asp Ala Ala Ser Leu Leu Asp Ile Ala Cys Gly	
35 40 45	
acc ggt acc cac ctg gag gcc ttc gcc aag ctc tac gac cgc gtg agc	192
Thr Gly Thr His Leu Glu Ala Phe Ala Lys Leu Tyr Asp Arg Val Ser	
50 55 60	
ggc ctg gag ctg tcc gag tgg atg gcg gcc cgc gcc gag gag cgg ctc	240
Gly Leu Glu Leu Ser Glu Trp Met Ala Ala Arg Ala Glu Glu Arg Leu	
65 70 75 80	
ccc ggc gtc acc ctc cac cgc ggt gac atg cgg gcg ttc gac ctc ggc	288
Pro Gly Val Thr Leu His Arg Gly Asp Met Arg Ala Phe Asp Leu Gly	
85 90 95	
gag acg ttc gac gcc gtg gtc tgc atg ttc agc tcg atc ggg tac ctg	336
Glu Thr Phe Asp Ala Val Val Cys Met Phe Ser Ser Ile Gly Tyr Leu	
100 105 110	
gag acc acg gcc gac ctg gag gac gcc gtc gcc gcc atg gcg cgg cac	384
Glu Thr Thr Ala Asp Leu Glu Asp Ala Val Ala Ala Met Ala Arg His	
115 120 125	
ctg acc gcg gac ggt gtc ctg gcc gtc gaa ccg tgg tac ttc ccc gac	432
Leu Thr Ala Asp Gly Val Leu Ala Val Glu Pro Trp Tyr Phe Pro Asp	
130 135 140	
acc ttc ctg gac ggc cac gtc tcc acc cac gcc ctg cgg acg gca ccg	480
Thr Phe Leu Asp Gly His Val Ser Thr His Ala Leu Arg Thr Ala Pro	
145 150 155 160	
ggc gac cag ggc gtc gcc cgt gtc tcc cac tcg acc cgg gag ggc ggg	528
Gly Asp Gln Gly Val Ala Arg Val Ser His Ser Thr Arg Glu Gly Gly	
165 170 175	
cgg acc cgg atg gag atc cac tac ctg atc gcg cac acc gcg gag ggc	576
Arg Thr Arg Met Glu Ile His Tyr Leu Ile Ala His Thr Ala Glu Gly	
180 185 190	
atc cgg cac cgc agc gag gtg gac tac ctc acg ctg ttc tcg cgt gcg	624
Ile Arg His Arg Ser Glu Val Asp Tyr Leu Thr Leu Phe Ser Arg Ala	
195 200 205	
gag tac gag gcc gcg tac cgc aag gcc ggc ctg gac gtc gag tac gtc	672
Glu Tyr Glu Ala Ala Tyr Arg Lys Ala Gly Leu Asp Val Glu Tyr Val	
210 215 220	

gtg acg ggc gag ggc tca ccg ggc ttc ttc ctc ggc acg cgt cgc tga 720
 Val Thr Gly Glu Gly Ser Pro Gly Phe Phe Leu Gly Thr Arg Arg
 225 230 235

<210> 46
 <211> 239
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 46

Met Tyr Glu Asn Asp Ser Ala Ala Glu Val Tyr Asp Leu Leu Tyr Gln
 1 5 10 15

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

Arg Glu Arg Thr Pro Asp Ala Ala Ser Leu Leu Asp Ile Ala Cys Gly
 35 40 45

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

Gly Leu Glu Leu Ser Glu Trp Met Ala Ala Arg Ala Glu Glu Arg Leu
 65 70 75 80

Pro Gly Val Thr Leu His Arg Gly Asp Met Arg Ala Phe Asp Leu Gly
 85 90 95

Glu Thr Phe Asp Ala Val Val Cys Met Phe Ser Ser Ile Gly Tyr Leu
 100 105 110

Glu Thr Thr Ala Asp Leu Glu Asp Ala Val Ala Ala Met Ala Arg His
 115 120 125

Leu Thr Ala Asp Gly Val Leu Ala Val Glu Pro Trp Tyr Phe Pro Asp
 130 135 140

Thr Phe Leu Asp Gly His Val Ser Thr His Ala Leu Arg Thr Ala Pro
 145 150 155 160

Gly Asp Gln Gly Val Ala Arg Val Ser His Ser Thr Arg Glu Gly Gly
 165 170 175

Arg Thr Arg Met Glu Ile His Tyr Leu Ile Ala His Thr Ala Glu Gly
 180 185 190

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

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

Val Thr Gly Glu Gly Ser Pro Gly Phe Phe Leu Gly Thr Arg Arg
 225 230 235

<210> 47
 <211> 1968
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1968)

<400> 47
 atg agt gac ctg ggt tct ggt gaa gaa ggg tcc gaa gaa gac gag tcg 48
 Met Ser Asp Leu Gly Ser Gly Glu Glu Gly Ser Glu Glu Asp Glu Ser
 1 5 10 15
 gac gac gca ctc gcc ttc ctc gag ttc atc gcc cgg tcg gca cca cgg 96
 Asp Asp Ala Leu Ala Phe Leu Glu Phe Ile Ala Arg Ser Ala Pro Arg
 20 25 30
 agc gaa tac gac cgg ctc atg gcc cgc gcc gaa cgc tcg ggc gcc gac 144
 Ser Glu Tyr Asp Arg Leu Met Ala Arg Ala Glu Arg Ser Gly Ala Asp

35	40	45	
gag gac cgg atg cgc cga ctg gag cgc ttc aac cgg ctc gcc ctc acc			192
Glu Asp Arg Met Arg Arg Leu Glu Arg Phe Asn Arg Leu Ala Leu Thr			
50	55	60	
gcg cag tcg atg atc gag tac cgc cgc gac cgg gag gcg gag ctc gcg			240
Ala Gln Ser Met Ile Glu Tyr Arg Arg Asp Arg Glu Ala Glu Leu Ala			
65	70	75	80
gcc ctg gtc gac gcc gcg cac gag ttc gtc gcc gcc cgg cgg ggc aag			288
Ala Leu Val Asp Ala Ala His Glu Phe Val Ala Ala Arg Arg Gly Lys			
	85	90	95
gac ctg ctg gag tcc atc gcc cgc aga gca cgg ctg ctg ctg aag ctg			336
Asp Leu Leu Glu Ser Ile Ala Arg Arg Ala Arg Leu Leu Leu Lys Leu			
	100	105	110
gac gtc tcc tac gtc ggc ctg cac gag gag gac cgg ccc ggc acg gtg			384
Asp Val Ser Tyr Val Gly Leu His Glu Glu Asp Arg Pro Gly Thr Val			
	115	120	125
gtg ctg agc gcc gac ggc aac gcg gtc aag gtc gcc gag agc tac cgg			432
Val Leu Ser Ala Asp Gly Asn Ala Val Lys Val Ala Glu Ser Tyr Arg			
130	135	140	
ctg ccg gcc gac ggc gga ctg ggc gcc atg gtg cgc acc tgc cgc gct			480
Leu Pro Ala Asp Gly Gly Leu Gly Ala Met Val Arg Thr Cys Arg Ala			
145	150	155	160
ccc ttc tgg acc ccg gac tac ctc ggg gac aac agc ttc acg cac gtc			528
Pro Phe Trp Thr Pro Asp Tyr Leu Gly Asp Asn Ser Phe Thr His Val			
	165	170	175
gag gcc gtc gac gac atc gtc cgc gcc gaa ggc ctg cgc gcg gtc ctg			576
Glu Ala Val Asp Asp Ile Val Arg Ala Glu Gly Leu Arg Ala Val Leu			
	180	185	190
gcc gtc ccg ctg tgc gcc ggg ggc gaa ccg atg ggg gtc ctc tac gtc			624
Ala Val Pro Leu Cys Ala Gly Gly Glu Pro Met Gly Val Leu Tyr Val			
	195	200	205
gcc gac cgt cag gtg cgg cat ctg acc ccc aac gag gtc acc ctg ctg			672
Ala Asp Arg Gln Val Arg His Leu Thr Pro Asn Glu Val Thr Leu Leu			
210	215	220	
tgc tcg ctc gcc gat ctg gcc gcg gtg gcg atc gag cgc aac cgg ctg			720

Cys Ser Leu Ala Asp Leu Ala Ala Val Ala Ile Glu Arg Asn Arg Leu	
225	230 235 240
gtc gag gag ctc cac gac acc atc ggg caa ctg cgc cag gac atc ggc	768
Val Glu Glu Leu His Asp Thr Ile Gly Gln Leu Arg Gln Asp Ile Gly	
245	250 255
gag gcc cgc acc gcc ctc gcg cgc acc cgc agg tcc gcc gac ctc cag	816
Glu Ala Arg Thr Ala Leu Ala Arg Thr Arg Arg Ser Ala Asp Leu Gln	
260	265 270
tcg cac ctg gtc acg cag gtg atg gac agg cgc ggc gcc gac tcg tta	864
Ser His Leu Val Thr Gln Val Met Asp Arg Arg Gly Ala Asp Ser Leu	
275	280 285
ctc gcg acg gcc gcc gag gcg ctc ggc ggc gga gcc ggc ctg tgc agc	912
Leu Ala Thr Ala Ala Glu Ala Leu Gly Gly Gly Ala Gly Leu Cys Ser	
290	295 300
ccg ctc ggg cgc ccg ctc gcc gag tac ggg acc ctg cgc ccc gtc gcc	960
Pro Leu Gly Arg Pro Leu Ala Glu Tyr Gly Thr Leu Arg Pro Val Ala	
305	310 315 320
ccc acg gaa ctg cgc gcg gcg tgc cgc cgg gcc gcc gag acc ggc cgg	1008
Pro Thr Glu Leu Arg Ala Ala Cys Arg Arg Ala Ala Glu Thr Gly Arg	
325	330 335
ccc acc tcc gtg gcc ccg ggg gtc tgg acg gtg ccc ctg ctt ccc ggg	1056
Pro Thr Ser Val Ala Pro Gly Val Trp Thr Val Pro Leu Leu Pro Gly	
340	345 350
ggc aac gcc ggc ttc ctg ctg acc gac ctc ggt ccg gac gcg gac cac	1104
Gly Asn Ala Gly Phe Leu Leu Thr Asp Leu Gly Pro Asp Ala Asp His	
355	360 365
acc gcc gtc ccc ctg ctc ccg atg gtc gcc cgc acc ctc gcg ctg cac	1152
Thr Ala Val Pro Leu Leu Pro Met Val Ala Arg Thr Leu Ala Leu His	
370	375 380
ctg cgc gtc cag cac gac gac tcc ccc aag gcg cag agc cac cag gag	1200
Leu Arg Val Gln His Asp Asp Ser Pro Lys Ala Gln Ser His Gln Glu	
385	390 395 400
ttc ttc gac gac ctg atc ggg gcg ccc cgc tca ccc acg ctc ctc agg	1248
Phe Phe Asp Asp Leu Ile Gly Ala Pro Arg Ser Pro Thr Leu Leu Arg	
405	410 415

gaa cgc gcc ctg atg ttc tcc ctc agc ttc cgc cgc ccg cac gtg gtg	1296
Glu Arg Ala Leu Met Phe Ser Leu Ser Phe Arg Arg Pro His Val Val	
420 425 430	
ctg gtg gcg gac gga ccc cgc ggg acc tcg ccg cgg ctg gag gcc tcc	1344
Leu Val Ala Asp Gly Pro Arg Gly Thr Ser Pro Arg Leu Glu Ala Ser	
435 440 445	
ggc gcc gac tac gcg aag gag ctc ggc ggg ctg tgc agc gtg cgg gac	1392
Gly Ala Asp Tyr Ala Lys Glu Leu Gly Gly Leu Cys Ser Val Arg Asp	
450 455 460	
ggc gcc gtc gtc ctg ctg ctg ccc ggc gac gac ccc gtc gcc gtg gcg	1440
Gly Ala Val Val Leu Leu Leu Pro Gly Asp Asp Pro Val Ala Val Ala	
465 470 475 480	
cag acc gcc gcc ccg gag ctg acc gac cgc gcc ggg cac ccc gtc acc	1488
Gln Thr Ala Ala Pro Glu Leu Thr Asp Arg Ala Gly His Pro Val Thr	
485 490 495	
gtg ggg gtc gcg ggc ccc gcc tcg acc gtc gac ggc atc gcc gac gcg	1536
Val Gly Val Ala Gly Pro Ala Ser Thr Val Asp Gly Ile Ala Asp Ala	
500 505 510	
cac cgt gag gcc gcg aag tgt ctg gag acc ctc cgc gcg ctc ggc ggc	1584
His Arg Glu Ala Ala Lys Cys Leu Glu Thr Leu Arg Ala Leu Gly Gly	
515 520 525	
gac ggc ggc acc gcg tgc gcc tcc gac ctg ggt ttc ctc ggc atg ctc	1632
Asp Gly Gly Thr Ala Cys Ala Ser Asp Leu Gly Phe Leu Gly Met Leu	
530 535 540	
ctc gcc gag gag aac gac gtc ccc ggt tac atc agg acg acg atc ggc	1680
Leu Ala Glu Glu Asn Asp Val Pro Gly Tyr Ile Arg Thr Thr Ile Gly	
545 550 555 560	
ccc gtg gtc gac tac gac acc cac cgc ttc acg gat ctg gtt ccc act	1728
Pro Val Val Asp Tyr Asp Thr His Arg Phe Thr Asp Leu Val Pro Thr	
565 570 575	
ctg agg gtg tac ctg gag tcg ggc agg agc ccc acg cgt gcc gca gag	1776
Leu Arg Val Tyr Leu Glu Ser Gly Arg Ser Pro Thr Arg Ala Ala Glu	
580 585 590	
aca ctg cgc gtg cac ccg aac acc gtc tca cgg cgg ctg gag cgc atc	1824
Thr Leu Arg Val His Pro Asn Thr Val Ser Arg Arg Leu Glu Arg Ile	
595 600 605	

ggc gta ctg ctg gga gag gac tgg cag tca ccg gag cgg gtg ctg gac 1872
 Gly Val Leu Leu Gly Glu Asp Trp Gln Ser Pro Glu Arg Val Leu Asp
 610 615 620

ata caa ctg gcc ctg cgg ctc tat cag gtg cgc tcg gcg ctc tcc tcg 1920
 Ile Gln Leu Ala Leu Arg Leu Tyr Gln Val Arg Ser Ala Leu Ser Ser
 625 630 635 640

caa ccg gcg tcc gag acc cgg gcc gtg ctc gga tcg ctg cgc gag tga 1968
 Gln Pro Ala Ser Glu Thr Arg Ala Val Leu Gly Ser Leu Arg Glu
 645 650 655

<210> 48
 <211> 655
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 48

Met Ser Asp Leu Gly Ser Gly Glu Glu Gly Ser Glu Glu Asp Glu Ser
 1 5 10 15

Asp Asp Ala Leu Ala Phe Leu Glu Phe Ile Ala Arg Ser Ala Pro Arg
 20 25 30

Ser Glu Tyr Asp Arg Leu Met Ala Arg Ala Glu Arg Ser Gly Ala Asp
 35 40 45

Glu Asp Arg Met Arg Arg Leu Glu Arg Phe Asn Arg Leu Ala Leu Thr
 50 55 60

Ala Gln Ser Met Ile Glu Tyr Arg Arg Asp Arg Glu Ala Glu Leu Ala
 65 70 75 80

Ala Leu Val Asp Ala Ala His Glu Phe Val Ala Ala Arg Arg Gly Lys
 85 90 95

Asp Leu Leu Glu Ser Ile Ala Arg Arg Ala Arg Leu Leu Leu Lys Leu
 100 105 110

Asp Val Ser Tyr Val Gly Leu His Glu Glu Asp Arg Pro Gly Thr Val
 115 120 125

Val Leu Ser Ala Asp Gly Asn Ala Val Lys Val Ala Glu Ser Tyr Arg
 130 135 140

Leu Pro Ala Asp Gly Gly Leu Gly Ala Met Val Arg Thr Cys Arg Ala
 145 150 155 160

Pro Phe Trp Thr Pro Asp Tyr Leu Gly Asp Asn Ser Phe Thr His Val
 165 170 175

Glu Ala Val Asp Asp Ile Val Arg Ala Glu Gly Leu Arg Ala Val Leu
 180 185 190

Ala Val Pro Leu Cys Ala Gly Gly Glu Pro Met Gly Val Leu Tyr Val
 195 200 205

Ala Asp Arg Gln Val Arg His Leu Thr Pro Asn Glu Val Thr Leu Leu
 210 215 220

Cys Ser Leu Ala Asp Leu Ala Ala Val Ala Ile Glu Arg Asn Arg Leu
 225 230 235 240

Val Glu Glu Leu His Asp Thr Ile Gly Gln Leu Arg Gln Asp Ile Gly
 245 250 255

Glu Ala Arg Thr Ala Leu Ala Arg Thr Arg Arg Ser Ala Asp Leu Gln
 260 265 270

Ser His Leu Val Thr Gln Val Met Asp Arg Arg Gly Ala Asp Ser Leu
 275 280 285

Leu Ala Thr Ala Ala Glu Ala Leu Gly Gly Gly Ala Gly Leu Cys Ser

290		295		300
Pro Leu Gly Arg Pro Leu Ala Glu Tyr Gly Thr Leu Arg Pro Val Ala				
305		310		315 320
Pro Thr Glu Leu Arg Ala Ala Cys Arg Arg Ala Ala Glu Thr Gly Arg				
	325		330	335
Pro Thr Ser Val Ala Pro Gly Val Trp Thr Val Pro Leu Leu Pro Gly				
	340		345	350
Gly Asn Ala Gly Phe Leu Leu Thr Asp Leu Gly Pro Asp Ala Asp His				
	355		360	365
Thr Ala Val Pro Leu Leu Pro Met Val Ala Arg Thr Leu Ala Leu His				
	370		375	380
Leu Arg Val Gln His Asp Asp Ser Pro Lys Ala Gln Ser His Gln Glu				
385		390		395 400
Phe Phe Asp Asp Leu Ile Gly Ala Pro Arg Ser Pro Thr Leu Leu Arg				
	405		410	415
Glu Arg Ala Leu Met Phe Ser Leu Ser Phe Arg Arg Pro His Val Val				
	420		425	430
Leu Val Ala Asp Gly Pro Arg Gly Thr Ser Pro Arg Leu Glu Ala Ser				
	435		440	445
Gly Ala Asp Tyr Ala Lys Glu Leu Gly Gly Leu Cys Ser Val Arg Asp				
	450		455	460
Gly Ala Val Val Leu Leu Leu Pro Gly Asp Asp Pro Val Ala Val Ala				
465		470		475 480

Gln Thr Ala Ala Pro Glu Leu Thr Asp Arg Ala Gly His Pro Val Thr
 485 490 495

Val Gly Val Ala Gly Pro Ala Ser Thr Val Asp Gly Ile Ala Asp Ala
 500 505 510

His Arg Glu Ala Ala Lys Cys Leu Glu Thr Leu Arg Ala Leu Gly Gly
 515 520 525

Asp Gly Gly Thr Ala Cys Ala Ser Asp Leu Gly Phe Leu Gly Met Leu
 530 535 540

Leu Ala Glu Glu Asn Asp Val Pro Gly Tyr Ile Arg Thr Thr Ile Gly
 545 550 555 560

Pro Val Val Asp Tyr Asp Thr His Arg Phe Thr Asp Leu Val Pro Thr
 565 570 575

Leu Arg Val Tyr Leu Glu Ser Gly Arg Ser Pro Thr Arg Ala Ala Glu
 580 585 590

Thr Leu Arg Val His Pro Asn Thr Val Ser Arg Arg Leu Glu Arg Ile
 595 600 605

Gly Val Leu Leu Gly Glu Asp Trp Gln Ser Pro Glu Arg Val Leu Asp
 610 615 620

Ile Gln Leu Ala Leu Arg Leu Tyr Gln Val Arg Ser Ala Leu Ser Ser
 625 630 635 640

Gln Pro Ala Ser Glu Thr Arg Ala Val Leu Gly Ser Leu Arg Glu
 645 650 655

<210> 49

<211> 1749

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1) .. (1749)

<220>

<221> CDS2

<222> (76) .. (1749)

<220>

<221> CDS3

<222> (97) .. (1749)

<400> 49

atg tcc gca gtg gtg gcg tcc ctt ctg ctt tcc gtt ctt tct gcg agc 48

Met Ser Ala Val Val Ala Ser Leu Leu Leu Ser Val Leu Ser Ala Ser

1 5 10 15

gcg ctc acg cat cca ctc atc ctc atc gtg gaa aca gga gac tta cgc 96

Ala Leu Thr His Pro Leu Ile Leu Ile Val Glu Thr Gly Asp Leu Arg

20 25 30

gtg tcg att gcg caa tac gcc cta cac gac atc acg aag cgc tac cac 144

Val Ser Ile Ala Gln Tyr Ala Leu His Asp Ile Thr Lys Arg Tyr His

35 40 45

gac tgt gtc gtg ctc gac cgg gtc ggt ttc agc atc aag ccg ggc gag 192

Asp Cys Val Val Leu Asp Arg Val Gly Phe Ser Ile Lys Pro Gly Glu

50 55 60

aag gtc ggc gtg atc ggc gac aac ggt tcc ggc aag tcc acg ctg ctc 240

Lys Val Gly Val Ile Gly Asp Asn Gly Ser Gly Lys Ser Thr Leu Leu

65 70 75 80

aag atc ctc gcc ggc cgc gtg gag ccc gac aac ggc gcg ctc acc gtg 288

Lys Ile Leu Ala Gly Arg Val Glu Pro Asp Asn Gly Ala Leu Thr Val

85 90 95

gtc gct ccc ggc ggc gtc ggc tac ctg gcg cag aca ctg gaa ctg ccc 336

Val Ala Pro Gly Gly Val Gly Tyr Leu Ala Gln Thr Leu Glu Leu Pro

100 105 110

ctc gac gcc acc gtc cag gac gcc gtc gac ctg gcc ctg tcc gac ctg 384

Leu Asp Ala Thr Val Gln Asp Ala Val Asp Leu Ala Leu Ser Asp Leu

115 120 125

cgc gag ctc gaa gcg gcg atg cgc gag gcc gag gcg gag ctg ggc gag	432
Arg Glu Leu Glu Ala Ala Met Arg Glu Ala Glu Ala Glu Leu Gly Glu	
130 135 140	
agc gac gag aac ggc tcc gag cgc gag ctg tcc gcc ggc ctc cag cgc	480
Ser Asp Glu Asn Gly Ser Glu Arg Glu Leu Ser Ala Gly Leu Gln Arg	
145 150 155 160	
tac gcc gct ctg gtc gag cag tac cag gcg cgt ggc ggc tac gag gcc	528
Tyr Ala Ala Leu Val Glu Gln Tyr Gln Ala Arg Gly Gly Tyr Glu Ala	
165 170 175	
gac gtg cgc gtg gag gtc gcg ctg cac ggc ctc gga ctg ccg agc ctg	576
Asp Val Arg Val Glu Val Ala Leu His Gly Leu Gly Leu Pro Ser Leu	
180 185 190	
gac cgc gac cgc aag ctc gga acc ctc tcc ggt ggc gaa cgc tcc cgc	624
Asp Arg Asp Arg Lys Leu Gly Thr Leu Ser Gly Gly Glu Arg Ser Arg	
195 200 205	
ctc gcg ctc gcc gcg acc ctc gcc tcg tcg ccg gag ctg ctg ctc ctg	672
Leu Ala Leu Ala Ala Thr Leu Ala Ser Ser Pro Glu Leu Leu Leu Leu	
210 215 220	
gac gaa ccg acc aac gac ctc gac gac cgg gcg atg gaa tgg ctg gag	720
Asp Glu Pro Thr Asn Asp Leu Asp Asp Arg Ala Met Glu Trp Leu Glu	
225 230 235 240	
gac cac ctg gcc ggc cac cgc ggc acg gtg atc gcg gtc acc cac gac	768
Asp His Leu Ala Gly His Arg Gly Thr Val Ile Ala Val Thr His Asp	
245 250 255	
cgg gtc ttc ctc gac cgg ctc acc acc acg atc ctg gag gtc gac tcc	816
Arg Val Phe Leu Asp Arg Leu Thr Thr Thr Ile Leu Glu Val Asp Ser	
260 265 270	
ggc agc gtc acc cgc tac ggc aac ggc tac gag ggc tac ctg acg gcc	864
Gly Ser Val Thr Arg Tyr Gly Asn Gly Tyr Glu Gly Tyr Leu Thr Ala	
275 280 285	
aag gcc gtg gaa cgc gag cgg cgg ctg cgg gag tac gag gag tgg cgt	912
Lys Ala Val Glu Arg Glu Arg Arg Leu Arg Glu Tyr Glu Glu Trp Arg	
290 295 300	
gcc gaa ctc gac cgc aac cgc ggg ctg atc acc tcc aac gtg gcg cgg	960
Ala Glu Leu Asp Arg Asn Arg Gly Leu Ile Thr Ser Asn Val Ala Arg	

305	310	315	320	
atg gac ggc atc ccg cgc aag atg tcc ctc tcc gtg ttc ggc cac ggc				1008
Met Asp Gly Ile Pro Arg Lys Met Ser Leu Ser Val Phe Gly His Gly				
	325	330	335	
gcc tac cgc agg cga ggg cgc gac cac ggc gcg atg gtg cgg atc cgc				1056
Ala Tyr Arg Arg Arg Gly Arg Asp His Gly Ala Met Val Arg Ile Arg				
	340	345	350	
aac gcg aag caa cgc gtg gcg cag ctg acc gag aac ccg gtc cac gct				1104
Asn Ala Lys Gln Arg Val Ala Gln Leu Thr Glu Asn Pro Val His Ala				
	355	360	365	
ccc gcc gac ccg ttg tcc ttc gcc gcc cgc atc gac acc gcg ggc ccg				1152
Pro Ala Asp Pro Leu Ser Phe Ala Ala Arg Ile Asp Thr Ala Gly Pro				
	370	375	380	
gag gcg gag gag gcg gtg gcc gaa ctc acc gac gtg cgc gtc gcg ggt				1200
Glu Ala Glu Glu Ala Val Ala Glu Leu Thr Asp Val Arg Val Ala Gly				
385	390	395	400	
cgg ctc gcc gtg gac tcc ctg acg atc cgg ccc ggc gaa cgg ctg ctc				1248
Arg Leu Ala Val Asp Ser Leu Thr Ile Arg Pro Gly Glu Arg Leu Leu				
	405	410	415	
gtc aca ggt ccc aac ggt gcg ggc aag tcc acc ttg ttg cgg gtg ctg				1296
Val Thr Gly Pro Asn Gly Ala Gly Lys Ser Thr Leu Leu Arg Val Leu				
	420	425	430	
tcc ggg gaa ctg gag ccg gac ggc ggc tcg gtg cgc gtc ggc tgc cgg				1344
Ser Gly Glu Leu Glu Pro Asp Gly Gly Ser Val Arg Val Gly Cys Arg				
	435	440	445	
gtc ggt cat ctg cgg cag gac gag acg ccc tgg gcg ccc gga ctg acc				1392
Val Gly His Leu Arg Gln Asp Glu Thr Pro Trp Ala Pro Gly Leu Thr				
	450	455	460	
gtg ctg cgg gcc ttc gcc cag ggc cgg gag ggc tac ctg gag gac cac				1440
Val Leu Arg Ala Phe Ala Gln Gly Arg Glu Gly Tyr Leu Glu Asp His				
465	470	475	480	
gcg gag aaa ctg ctg tcg ctc ggc ctg ttc agc ccg tcc gac ctg cgg				1488
Ala Glu Lys Leu Leu Ser Leu Gly Leu Phe Ser Pro Ser Asp Leu Arg				
	485	490	495	
cga cgc gtg aag gat ctg tcc tac ggg cag cgc cgc cgg atc gag atc				1536

Arg Arg Val Lys Asp Leu Ser Tyr Gly Gln Arg Arg Arg Ile Glu Ile
 500 505 510

gcc cgg ctg gtg agc gac ccg atg gac ctg ctg ctg ctg gac gag ccc 1584
 Ala Arg Leu Val Ser Asp Pro Met Asp Leu Leu Leu Leu Asp Glu Pro
 515 520 525

acc aac cac ctc acc ccg gtg ctg gtg gag gag ttg gag cag gca ctc 1632
 Thr Asn His Leu Thr Pro Val Leu Val Glu Glu Leu Glu Gln Ala Leu
 530 535 540

gcg gac tac cgc ggc gcc gtc gtg gtc gtc acc cac gac cgt cgg atg 1680
 Ala Asp Tyr Arg Gly Ala Val Val Val Val Thr His Asp Arg Arg Met
 545 550 555 560

cgg tcc cgg ttc acc ggc gcc cgg ctg acc atg gga gac ggg cgc atc 1728
 Arg Ser Arg Phe Thr Gly Ala Arg Leu Thr Met Gly Asp Gly Arg Ile
 565 570 575

gcc gag ttc agc gcc ggc tga 1749
 Ala Glu Phe Ser Ala Gly
 580

<210> 50

<211> 582

<212> PRT

<213> Streptomyces ambofaciens

<400> 50

Met Ser Ala Val Val Ala Ser Leu Leu Leu Ser Val Leu Ser Ala Ser
 1 5 10 15

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

Val Ser Ile Ala Gln Tyr Ala Leu His Asp Ile Thr Lys Arg Tyr His
 35 40 45

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

Lys Val Gly Val Ile Gly Asp Asn Gly Ser Gly Lys Ser Thr Leu Leu
65 70 75 80

Lys Ile Leu Ala Gly Arg Val Glu Pro Asp Asn Gly Ala Leu Thr Val
85 90 95

Val Ala Pro Gly Gly Val Gly Tyr Leu Ala Gln Thr Leu Glu Leu Pro
100 105 110

Leu Asp Ala Thr Val Gln Asp Ala Val Asp Leu Ala Leu Ser Asp Leu
115 120 125

Arg Glu Leu Glu Ala Ala Met Arg Glu Ala Glu Ala Glu Leu Gly Glu
130 135 140

Ser Asp Glu Asn Gly Ser Glu Arg Glu Leu Ser Ala Gly Leu Gln Arg
145 150 155 160

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

Asp Val Arg Val Glu Val Ala Leu His Gly Leu Gly Leu Pro Ser Leu
180 185 190

Asp Arg Asp Arg Lys Leu Gly Thr Leu Ser Gly Gly Glu Arg Ser Arg
195 200 205

Leu Ala Leu Ala Ala Thr Leu Ala Ser Ser Pro Glu Leu Leu Leu Leu
210 215 220

Asp Glu Pro Thr Asn Asp Leu Asp Asp Arg Ala Met Glu Trp Leu Glu
225 230 235 240

Asp His Leu Ala Gly His Arg Gly Thr Val Ile Ala Val Thr His Asp
245 250 255

Arg Val Phe Leu Asp Arg Leu Thr Thr Thr Ile Leu Glu Val Asp Ser
 260 265 270

Gly Ser Val Thr Arg Tyr Gly Asn Gly Tyr Glu Gly Tyr Leu Thr Ala
 275 280 285

Lys Ala Val Glu Arg Glu Arg Arg Leu Arg Glu Tyr Glu Glu Trp Arg
 290 295 300

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

Met Asp Gly Ile Pro Arg Lys Met Ser Leu Ser Val Phe Gly His Gly
 325 330 335

Ala Tyr Arg Arg Arg Gly Arg Asp His Gly Ala Met Val Arg Ile Arg
 340 345 350

Asn Ala Lys Gln Arg Val Ala Gln Leu Thr Glu Asn Pro Val His Ala
 355 360 365

Pro Ala Asp Pro Leu Ser Phe Ala Ala Arg Ile Asp Thr Ala Gly Pro
 370 375 380

Glu Ala Glu Glu Ala Val Ala Glu Leu Thr Asp Val Arg Val Ala Gly
 385 390 395 400

Arg Leu Ala Val Asp Ser Leu Thr Ile Arg Pro Gly Glu Arg Leu Leu
 405 410 415

Val Thr Gly Pro Asn Gly Ala Gly Lys Ser Thr Leu Leu Arg Val Leu
 420 425 430

Ser Gly Glu Leu Glu Pro Asp Gly Gly Ser Val Arg Val Gly Cys Arg
 435 440 445

Val Gly His Leu Arg Gln Asp Glu Thr Pro Trp Ala Pro Gly Leu Thr
 450 455 460

Val Leu Arg Ala Phe Ala Gln Gly Arg Glu Gly Tyr Leu Glu Asp His
 465 470 475 480

Ala Glu Lys Leu Leu Ser Leu Gly Leu Phe Ser Pro Ser Asp Leu Arg
 485 490 495

Arg Arg Val Lys Asp Leu Ser Tyr Gly Gln Arg Arg Arg Ile Glu Ile
 500 505 510

Ala Arg Leu Val Ser Asp Pro Met Asp Leu Leu Leu Leu Asp Glu Pro
 515 520 525

Thr Asn His Leu Thr Pro Val Leu Val Glu Glu Leu Glu Gln Ala Leu
 530 535 540

Ala Asp Tyr Arg Gly Ala Val Val Val Val Thr His Asp Arg Arg Met
 545 550 555 560

Arg Ser Arg Phe Thr Gly Ala Arg Leu Thr Met Gly Asp Gly Arg Ile
 565 570 575

Ala Glu Phe Ser Ala Gly
 580

<210> 51
 <211> 557
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 51

Val Glu Thr Gly Asp Leu Arg Val Ser Ile Ala Gln Tyr Ala Leu His
 1 5 10 15

Asp Ile Thr Lys Arg Tyr His Asp Cys Val Val Leu Asp Arg Val Gly
 20 25 30

Phe Ser Ile Lys Pro Gly Glu Lys Val Gly Val Ile Gly Asp Asn Gly
 35 40 45

Ser Gly Lys Ser Thr Leu Leu Lys Ile Leu Ala Gly Arg Val Glu Pro
 50 55 60

Asp Asn Gly Ala Leu Thr Val Val Ala Pro Gly Gly Val Gly Tyr Leu
 65 70 75 80

Ala Gln Thr Leu Glu Leu Pro Leu Asp Ala Thr Val Gln Asp Ala Val
 85 90 95

Asp Leu Ala Leu Ser Asp Leu Arg Glu Leu Glu Ala Ala Met Arg Glu
 100 105 110

Ala Glu Ala Glu Leu Gly Glu Ser Asp Glu Asn Gly Ser Glu Arg Glu
 115 120 125

Leu Ser Ala Gly Leu Gln Arg Tyr Ala Ala Leu Val Glu Gln Tyr Gln
 130 135 140

Ala Arg Gly Gly Tyr Glu Ala Asp Val Arg Val Glu Val Ala Leu His
 145 150 155 160

Gly Leu Gly Leu Pro Ser Leu Asp Arg Asp Arg Lys Leu Gly Thr Leu
 165 170 175

Ser Gly Gly Glu Arg Ser Arg Leu Ala Leu Ala Ala Thr Leu Ala Ser
 180 185 190

Ser Pro Glu Leu Leu Leu Leu Asp Glu Pro Thr Asn Asp Leu Asp Asp

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

Arg Pro Gly Glu Arg Leu Leu Val Thr Gly Pro Asn Gly Ala Gly Lys
 385 390 395 400

Ser Thr Leu Leu Arg Val Leu Ser Gly Glu Leu Glu Pro Asp Gly Gly
 405 410 415

Ser Val Arg Val Gly Cys Arg Val Gly His Leu Arg Gln Asp Glu Thr
 420 425 430

Pro Trp Ala Pro Gly Leu Thr Val Leu Arg Ala Phe Ala Gln Gly Arg
 435 440 445

Glu Gly Tyr Leu Glu Asp His Ala Glu Lys Leu Leu Ser Leu Gly Leu
 450 455 460

Phe Ser Pro Ser Asp Leu Arg Arg Arg Val Lys Asp Leu Ser Tyr Gly
 465 470 475 480

Gln Arg Arg Arg Ile Glu Ile Ala Arg Leu Val Ser Asp Pro Met Asp
 485 490 495

Leu Leu Leu Leu Asp Glu Pro Thr Asn His Leu Thr Pro Val Leu Val
 500 505 510

Glu Glu Leu Glu Gln Ala Leu Ala Asp Tyr Arg Gly Ala Val Val Val
 515 520 525

Val Thr His Asp Arg Arg Met Arg Ser Arg Phe Thr Gly Ala Arg Leu
 530 535 540

Thr Met Gly Asp Gly Arg Ile Ala Glu Phe Ser Ala Gly
 545 550 555

<210> 52

<211> 550

<212> PRT

<213> Streptomyces ambofaciens

<400> 52

Val Ser Ile Ala Gln Tyr Ala Leu His Asp Ile Thr Lys Arg Tyr His
1 5 10 15

Asp Cys Val Val Leu Asp Arg Val Gly Phe Ser Ile Lys Pro Gly Glu
20 25 30

Lys Val Gly Val Ile Gly Asp Asn Gly Ser Gly Lys Ser Thr Leu Leu
35 40 45

Lys Ile Leu Ala Gly Arg Val Glu Pro Asp Asn Gly Ala Leu Thr Val
50 55 60

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

Leu Asp Ala Thr Val Gln Asp Ala Val Asp Leu Ala Leu Ser Asp Leu
85 90 95

Arg Glu Leu Glu Ala Ala Met Arg Glu Ala Glu Ala Glu Leu Gly Glu
100 105 110

Ser Asp Glu Asn Gly Ser Glu Arg Glu Leu Ser Ala Gly Leu Gln Arg
115 120 125

Tyr Ala Ala Leu Val Glu Gln Tyr Gln Ala Arg Gly Gly Tyr Glu Ala
130 135 140

Asp Val Arg Val Glu Val Ala Leu His Gly Leu Gly Leu Pro Ser Leu
145 150 155 160

Asp Arg Asp Arg Lys Leu Gly Thr Leu Ser Gly Gly Glu Arg Ser Arg
165 170 175

Leu Ala Leu Ala Ala Thr Leu Ala Ser Ser Pro Glu Leu Leu Leu Leu
 180 185 190

Asp Glu Pro Thr Asn Asp Leu Asp Asp Arg Ala Met Glu Trp Leu Glu
 195 200 205

Asp His Leu Ala Gly His Arg Gly Thr Val Ile Ala Val Thr His Asp
 210 215 220

Arg Val Phe Leu Asp Arg Leu Thr Thr Thr Ile Leu Glu Val Asp Ser
 225 230 235 240

Gly Ser Val Thr Arg Tyr Gly Asn Gly Tyr Glu Gly Tyr Leu Thr Ala
 245 250 255

Lys Ala Val Glu Arg Glu Arg Arg Leu Arg Glu Tyr Glu Glu Trp Arg
 260 265 270

Ala Glu Leu Asp Arg Asn Arg Gly Leu Ile Thr Ser Asn Val Ala Arg
 275 280 285

Met Asp Gly Ile Pro Arg Lys Met Ser Leu Ser Val Phe Gly His Gly
 290 295 300

Ala Tyr Arg Arg Arg Gly Arg Asp His Gly Ala Met Val Arg Ile Arg
 305 310 315 320

Asn Ala Lys Gln Arg Val Ala Gln Leu Thr Glu Asn Pro Val His Ala
 325 330 335

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

Glu Ala Glu Glu Ala Val Ala Glu Leu Thr Asp Val Arg Val Ala Gly
 355 360 365

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

Val Thr Gly Pro Asn Gly Ala Gly Lys Ser Thr Leu Leu Arg Val Leu
 385 390 395 400

Ser Gly Glu Leu Glu Pro Asp Gly Gly Ser Val Arg Val Gly Cys Arg
 405 410 415

Val Gly His Leu Arg Gln Asp Glu Thr Pro Trp Ala Pro Gly Leu Thr
 420 425 430

Val Leu Arg Ala Phe Ala Gln Gly Arg Glu Gly Tyr Leu Glu Asp His
 435 440 445

Ala Glu Lys Leu Leu Ser Leu Gly Leu Phe Ser Pro Ser Asp Leu Arg
 450 455 460

Arg Arg Val Lys Asp Leu Ser Tyr Gly Gln Arg Arg Arg Ile Glu Ile
 465 470 475 480

Ala Arg Leu Val Ser Asp Pro Met Asp Leu Leu Leu Leu Asp Glu Pro
 485 490 495

Thr Asn His Leu Thr Pro Val Leu Val Glu Glu Leu Glu Gln Ala Leu
 500 505 510

Ala Asp Tyr Arg Gly Ala Val Val Val Val Thr His Asp Arg Arg Met
 515 520 525

Arg Ser Arg Phe Thr Gly Ala Arg Leu Thr Met Gly Asp Gly Arg Ile
 530 535 540

Ala Glu Phe Ser Ala Gly

545

550

<210> 53
 <211> 1431
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1) .. (1431)

<220>
 <221> CDS2
 <222> (16) .. (1431)

<220>
 <221> CDS3
 <222> (37) .. (1431)

<220>
 <221> CDS4
 <222> (40) .. (1431)

<220>
 <221> CDS5
 <222> (79) .. (1431)

<220>
 <221> CDS6
 <222> (130) .. (1431)

<400> 53
 gtg ccg cgc aac ggt atg cgt gtc gca ccg gcc gac gtg atg cag tcg 48
 Val Pro Arg Asn Gly Met Arg Val Ala Pro Ala Asp Val Met Gln Ser
 1 5 10 15

gac cgg gat gat cgc ttg tcc ggc ggc cgg atg cct agc ctc ggg agc 96
 Asp Arg Asp Asp Arg Leu Ser Gly Gly Arg Met Pro Ser Leu Gly Ser
 20 25 30

aac cac agc ggt ctt tca cga gag ggg tcg acc atg ggc gat ctc agg 144
 Asn His Ser Gly Leu Ser Arg Glu Gly Ser Thr Met Gly Asp Leu Arg
 35 40 45

aac cgc atc acc gag ctg gtc cgc gcg tac cac cgg gaa cag gcg ccc 192

Asn	Arg	Ile	Thr	Glu	Leu	Val	Arg	Ala	Tyr	His	Arg	Glu	Gln	Ala	Pro	
50						55					60					
ggg	ggc	ttc	gtt	ccc	ggg	acg	acg	cac	gta	ccg	gtc	tcc	ggc	gcg	gtg	240
Gly	Gly	Phe	Val	Pro	Gly	Thr	Thr	His	Val	Pro	Val	Ser	Gly	Ala	Val	
65					70				75						80	
ctg	agc	gag	gag	gac	cgg	ctg	gcg	ctg	gtg	gag	acg	gcg	ctg	gag	atg	288
Leu	Ser	Glu	Glu	Asp	Arg	Leu	Ala	Leu	Val	Glu	Thr	Ala	Leu	Glu	Met	
				85					90					95		
cgg	atc	gcg	gcc	ggc	ccg	gcc	tcc	cgg	ggc	ttc	gag	cgg	cag	ttc	gcc	336
Arg	Ile	Ala	Ala	Gly	Pro	Ala	Ser	Arg	Gly	Phe	Glu	Arg	Gln	Phe	Ala	
			100					105						110		
cgg	tac	ctc	ggg	ctc	cgg	aag	gcg	cac	ctg	acc	aac	tcc	ggg	tcc	tcc	384
Arg	Tyr	Leu	Gly	Leu	Arg	Lys	Ala	His	Leu	Thr	Asn	Ser	Gly	Ser	Ser	
		115					120					125				
gcc	aac	ctc	ctc	gcc	ctc	ggc	gcg	ctc	acc	tgc	ccg	cag	ctg	gag	gag	432
Ala	Asn	Leu	Leu	Ala	Leu	Gly	Ala	Leu	Thr	Ser	Pro	Gln	Leu	Glu	Glu	
		130				135					140					
aga	cgg	ctg	cgt	ccg	ggg	gac	gag	gtc	gtc	acg	gtc	gcc	gcc	ggg	ttc	480
Arg	Arg	Leu	Arg	Pro	Gly	Asp	Glu	Val	Val	Thr	Val	Ala	Ala	Gly	Phe	
145					150					155					160	
ccc	acg	acg	gtc	aac	ccg	atc	ttc	cac	aac	ggg	ctg	gtg	ccc	gtc	ttc	528
Pro	Thr	Thr	Val	Asn	Pro	Ile	Phe	His	Asn	Gly	Leu	Val	Pro	Val	Phe	
				165					170					175		
gtg	gac	gtc	gag	ctc	ggc	acg	tac	aac	acg	acg	ccc	gag	cgc	atc	gag	576
Val	Asp	Val	Glu	Leu	Gly	Thr	Tyr	Asn	Thr	Thr	Pro	Glu	Arg	Ile	Glu	
			180					185					190			
cgg	gcc	atc	ggc	ccc	cgg	acc	agg	gcg	atc	atg	atc	gcg	cac	gcc	ctg	624
Arg	Ala	Ile	Gly	Pro	Arg	Thr	Arg	Ala	Ile	Met	Ile	Ala	His	Ala	Leu	
		195					200					205				
ggc	aac	ccc	ttc	gag	gcc	gaa	gag	gtg	gcc	cgc	ctc	gcg	gac	gag	cgg	672
Gly	Asn	Pro	Phe	Glu	Ala	Glu	Glu	Val	Ala	Arg	Leu	Ala	Asp	Glu	Arg	
		210				215				220						
gag	ctg	ttc	ctc	atc	gag	gac	aac	tgc	gac	gcg	gtg	ggg	tcc	cgc	tac	720
Glu	Leu	Phe	Leu	Ile	Glu	Asp	Asn	Cys	Asp	Ala	Val	Gly	Ser	Arg	Tyr	
225					230				235						240	

cgg ggc agg ctc acc ggc tcc ttc ggc gac ctg tcg acc gtc agc ttc	768
Arg Gly Arg Leu Thr Gly Ser Phe Gly Asp Leu Ser Thr Val Ser Phe	
245 250 255	
tat ccc gcg cac cac atc gcg atg ggt gag ggg ggc tgc gtg ctc acc	816
Tyr Pro Ala His His Ile Ala Met Gly Glu Gly Gly Cys Val Leu Thr	
260 265 270	
gac aac ctg gcc ctg gcg cgg atc gtg gaa tca ctg cgc gac tgg ggg	864
Asp Asn Leu Ala Leu Ala Arg Ile Val Glu Ser Leu Arg Asp Trp Gly	
275 280 285	
cgc gac tgc tgg tgc gag ccg ggt gag gac aac cgc tgc ctc aag cgg	912
Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp Asn Arg Cys Leu Lys Arg	
290 295 300	
ttc gac cag aag atg ggt gac ctg ccg ccc ggg tac gac cac aag tac	960
Phe Asp Gln Lys Met Gly Asp Leu Pro Pro Gly Tyr Asp His Lys Tyr	
305 310 315 320	
atc ttc tcg cac gtc ggt tac aac ctg aag tcg acc gac ctg cag gcg	1008
Ile Phe Ser His Val Gly Tyr Asn Leu Lys Ser Thr Asp Leu Gln Ala	
325 330 335	
gcc ctc ggg ctg tcc cag ctg acc cgg atc gag gag ttc acc gag gcc	1056
Ala Leu Gly Leu Ser Gln Leu Thr Arg Ile Glu Glu Phe Thr Glu Ala	
340 345 350	
agg cgc gcc aac tgg cgg cat ctg cgc gcc gcg ttg gac ggg ctg ccc	1104
Arg Arg Ala Asn Trp Arg His Leu Arg Ala Ala Leu Asp Gly Leu Pro	
355 360 365	
ggt ctg ctg ctg cct cat gcc aca ccg ggc agc gat ccg agc tgg ttc	1152
Gly Leu Leu Leu Pro His Ala Thr Pro Gly Ser Asp Pro Ser Trp Phe	
370 375 380	
ggg ttc ctc atc acc gtg gac ccg gac gcc gcg tac agc agg gcg gcc	1200
Gly Phe Leu Ile Thr Val Asp Pro Asp Ala Ala Tyr Ser Arg Ala Ala	
385 390 395 400	
ctg gtc gac cac ctg gaa tcg cgc cgg atc agc acc cgc cgc ctg ttc	1248
Leu Val Asp His Leu Glu Ser Arg Arg Ile Ser Thr Arg Arg Leu Phe	
405 410 415	
ggg ggc aac ctc gtg cgg cac ccc gcc tac acc gac cgt cgg tac cgg	1296
Gly Gly Asn Leu Val Arg His Pro Ala Tyr Thr Asp Arg Arg Tyr Arg	
420 425 430	

gtg tcc ggc tcc ctg gag aac agc gac ctg atc acc gac cag acg ttc 1344
Val Ser Gly Ser Leu Glu Asn Ser Asp Leu Ile Thr Asp Gln Thr Phe
435 440 445

tgg atc ggg gtc ttc ccc ggc atc acc ccg gag atg atc gcc tac gtc 1392
 Trp Ile Gly Val Phe Pro Gly Ile Thr Pro Glu Met Ile Ala Tyr Val
 450 455 460

ggc gac acg atc cgg gag ttc gtg ctc aag cac tcc tga 1431
Gly Asp Thr Ile Arg Glu Phe Val Leu Lys His Ser
465 470 475

<210>	54
<211>	476
<212>	PRT
<213>	Streptomyces ambofaciens

<400> 54

Val	Pro	Arg	Asn	Gly	Met	Arg	Val	Ala	Pro	Ala	Asp	Val	Met	Gln	Ser
1				5					10					15	

Asp Arg Asp Asp Arg Leu Ser Gly Gly Arg Met Pro Ser Leu Gly Ser
20 25 30

Asn His Ser Gly Leu Ser Arg Glu Gly Ser Thr Met Gly Asp Leu Arg
35 40 45

Asn Arg Ile Thr Glu Leu Val Arg Ala Tyr His Arg Glu Gln Ala Pro
50 55 60

Gly Gly Phe Val Pro Gly Thr Thr His Val Pro Val Ser Gly Ala Val
65 70 75 80

Leu Ser Glu Glu Asp Arg Leu Ala Leu Val Glu Thr Ala Leu Glu Met
85 90 95

Arg Ile Ala Ala Gly Pro Ala Ser Arg Gly Phe Glu Arg Gln Phe Ala
100 105 110

Arg Tyr Leu Gly Leu Arg Lys Ala His Leu Thr Asn Ser Gly Ser Ser
 115 120 125

Ala Asn Leu Leu Ala Leu Gly Ala Leu Thr Ser Pro Gln Leu Glu Glu
 130 135 140

Arg Arg Leu Arg Pro Gly Asp Glu Val Val Thr Val Ala Ala Gly Phe
 145 150 155 160

Pro Thr Thr Val Asn Pro Ile Phe His Asn Gly Leu Val Pro Val Phe
 165 170 175

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

Arg Ala Ile Gly Pro Arg Thr Arg Ala Ile Met Ile Ala His Ala Leu
 195 200 205

Gly Asn Pro Phe Glu Ala Glu Glu Val Ala Arg Leu Ala Asp Glu Arg
 210 215 220

Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp Ala Val Gly Ser Arg Tyr
 225 230 235 240

Arg Gly Arg Leu Thr Gly Ser Phe Gly Asp Leu Ser Thr Val Ser Phe
 245 250 255

Tyr Pro Ala His His Ile Ala Met Gly Glu Gly Gly Cys Val Leu Thr
 260 265 270

Asp Asn Leu Ala Leu Ala Arg Ile Val Glu Ser Leu Arg Asp Trp Gly
 275 280 285

Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp Asn Arg Cys Leu Lys Arg

290		295		300
Phe Asp Gln Lys Met Gly Asp Leu Pro Pro Gly Tyr Asp His Lys Tyr				
305		310		315 320
Ile Phe Ser His Val Gly Tyr Asn Leu Lys Ser Thr Asp Leu Gln Ala				
	325		330	335
Ala Leu Gly Leu Ser Gln Leu Thr Arg Ile Glu Glu Phe Thr Glu Ala				
	340		345	350
Arg Arg Ala Asn Trp Arg His Leu Arg Ala Ala Leu Asp Gly Leu Pro				
	355		360	365
Gly Leu Leu Leu Pro His Ala Thr Pro Gly Ser Asp Pro Ser Trp Phe				
	370		375	380
Gly Phe Leu Ile Thr Val Asp Pro Asp Ala Ala Tyr Ser Arg Ala Ala				
385		390		395 400
Leu Val Asp His Leu Glu Ser Arg Arg Ile Ser Thr Arg Arg Leu Phe				
	405		410	415
Gly Gly Asn Leu Val Arg His Pro Ala Tyr Thr Asp Arg Arg Tyr Arg				
	420		425	430
Val Ser Gly Ser Leu Glu Asn Ser Asp Leu Ile Thr Asp Gln Thr Phe				
	435		440	445
Trp Ile Gly Val Phe Pro Gly Ile Thr Pro Glu Met Ile Ala Tyr Val				
	450		455	460
Gly Asp Thr Ile Arg Glu Phe Val Leu Lys His Ser				
465		470		475

<210> 55
 <211> 471
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 55

Met Arg Val Ala Pro Ala Asp Val Met Gln Ser Asp Arg Asp Asp Arg
 1 5 10 15

Leu Ser Gly Gly Arg Met Pro Ser Leu Gly Ser Asn His Ser Gly Leu
 20 25 30

Ser Arg Glu Gly Ser Thr Met Gly Asp Leu Arg Asn Arg Ile Thr Glu
 35 40 45

Leu Val Arg Ala Tyr His Arg Glu Gln Ala Pro Gly Gly Phe Val Pro
 50 55 60

Gly Thr Thr His Val Pro Val Ser Gly Ala Val Leu Ser Glu Glu Asp
 65 70 75 80

Arg Leu Ala Leu Val Glu Thr Ala Leu Glu Met Arg Ile Ala Ala Gly
 85 90 95

Pro Ala Ser Arg Gly Phe Glu Arg Gln Phe Ala Arg Tyr Leu Gly Leu
 100 105 110

Arg Lys Ala His Leu Thr Asn Ser Gly Ser Ser Ala Asn Leu Leu Ala
 115 120 125

Leu Gly Ala Leu Thr Ser Pro Gln Leu Glu Glu Arg Arg Leu Arg Pro
 130 135 140

Gly Asp Glu Val Val Thr Val Ala Ala Gly Phe Pro Thr Thr Val Asn
 145 150 155 160

Pro Ile Phe His Asn Gly Leu Val Pro Val Phe Val Asp Val Glu Leu
 165 170 175

Gly Thr Tyr Asn Thr Thr Pro Glu Arg Ile Glu Arg Ala Ile Gly Pro
 180 185 190

Arg Thr Arg Ala Ile Met Ile Ala His Ala Leu Gly Asn Pro Phe Glu
 195 200 205

Ala Glu Glu Val Ala Arg Leu Ala Asp Glu Arg Glu Leu Phe Leu Ile
 210 215 220

Glu Asp Asn Cys Asp Ala Val Gly Ser Arg Tyr Arg Gly Arg Leu Thr
 225 230 235 240

Gly Ser Phe Gly Asp Leu Ser Thr Val Ser Phe Tyr Pro Ala His His
 245 250 255

Ile Ala Met Gly Glu Gly Gly Cys Val Leu Thr Asp Asn Leu Ala Leu
 260 265 270

Ala Arg Ile Val Glu Ser Leu Arg Asp Trp Gly Arg Asp Cys Trp Cys
 275 280 285

Glu Pro Gly Glu Asp Asn Arg Cys Leu Lys Arg Phe Asp Gln Lys Met
 290 295 300

Gly Asp Leu Pro Pro Gly Tyr Asp His Lys Tyr Ile Phe Ser His Val
 305 310 315 320

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

Gln Leu Thr Arg Ile Glu Glu Phe Thr Glu Ala Arg Arg Ala Asn Trp
 340 345 350

Arg His Leu Arg Ala Ala Leu Asp Gly Leu Pro Gly Leu Leu Leu Pro
 355 360 365

His Ala Thr Pro Gly Ser Asp Pro Ser Trp Phe Gly Phe Leu Ile Thr
 370 375 380

Val Asp Pro Asp Ala Ala Tyr Ser Arg Ala Ala Leu Val Asp His Leu
 385 390 395 400

Glu Ser Arg Arg Ile Ser Thr Arg Arg Leu Phe Gly Gly Asn Leu Val
 405 410 415

Arg His Pro Ala Tyr Thr Asp Arg Arg Tyr Arg Val Ser Gly Ser Leu
 420 425 430

Glu Asn Ser Asp Leu Ile Thr Asp Gln Thr Phe Trp Ile Gly Val Phe
 435 440 445

Pro Gly Ile Thr Pro Glu Met Ile Ala Tyr Val Gly Asp Thr Ile Arg
 450 455 460

Glu Phe Val Leu Lys His Ser
 465 470

<210> 56

<211> 464

<212> PRT

<213> Streptomyces ambofaciens

<400> 56

Val Met Gln Ser Asp Arg Asp Asp Arg Leu Ser Gly Gly Arg Met Pro
 1 5 10 15

Ser Leu Gly Ser Asn His Ser Gly Leu Ser Arg Glu Gly Ser Thr Met
 20 25 30

Gly Asp Leu Arg Asn Arg Ile Thr Glu Leu Val Arg Ala Tyr His Arg
 35 40 45

Glu Gln Ala Pro Gly Gly Phe Val Pro Gly Thr Thr His Val Pro Val
 50 55 60

Ser Gly Ala Val Leu Ser Glu Glu Asp Arg Leu Ala Leu Val Glu Thr
 65 70 75 80

Ala Leu Glu Met Arg Ile Ala Ala Gly Pro Ala Ser Arg Gly Phe Glu
 85 90 95

Arg Gln Phe Ala Arg Tyr Leu Gly Leu Arg Lys Ala His Leu Thr Asn
 100 105 110

Ser Gly Ser Ser Ala Asn Leu Leu Ala Leu Gly Ala Leu Thr Ser Pro
 115 120 125

Gln Leu Glu Glu Arg Arg Leu Arg Pro Gly Asp Glu Val Val Thr Val
 130 135 140

Ala Ala Gly Phe Pro Thr Thr Val Asn Pro Ile Phe His Asn Gly Leu
 145 150 155 160

Val Pro Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Thr Thr Pro
 165 170 175

Glu Arg Ile Glu Arg Ala Ile Gly Pro Arg Thr Arg Ala Ile Met Ile
 180 185 190

Ala His Ala Leu Gly Asn Pro Phe Glu Ala Glu Glu Val Ala Arg Leu
 195 200 205

Ala Asp Glu Arg Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp Ala Val
 210 215 220

Gly Ser Arg Tyr Arg Gly Arg Leu Thr Gly Ser Phe Gly Asp Leu Ser
 225 230 235 240

Thr Val Ser Phe Tyr Pro Ala His His Ile Ala Met Gly Glu Gly Gly
 245 250 255

Cys Val Leu Thr Asp Asn Leu Ala Leu Ala Arg Ile Val Glu Ser Leu
 260 265 270

Arg Asp Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp Asn Arg
 275 280 285

Cys Leu Lys Arg Phe Asp Gln Lys Met Gly Asp Leu Pro Pro Gly Tyr
 290 295 300

Asp His Lys Tyr Ile Phe Ser His Val Gly Tyr Asn Leu Lys Ser Thr
 305 310 315 320

Asp Leu Gln Ala Ala Leu Gly Leu Ser Gln Leu Thr Arg Ile Glu Glu
 325 330 335

Phe Thr Glu Ala Arg Arg Ala Asn Trp Arg His Leu Arg Ala Ala Leu
 340 345 350

Asp Gly Leu Pro Gly Leu Leu Leu Pro His Ala Thr Pro Gly Ser Asp
 355 360 365

Pro Ser Trp Phe Gly Phe Leu Ile Thr Val Asp Pro Asp Ala Ala Tyr
 370 375 380

Ser Arg Ala Ala Leu Val Asp His Leu Glu Ser Arg Arg Ile Ser Thr
 385 390 395 400

Arg Arg Leu Phe Gly Gly Asn Leu Val Arg His Pro Ala Tyr Thr Asp

405

410

415

Arg Arg Tyr Arg Val Ser Gly Ser Leu Glu Asn Ser Asp Leu Ile Thr
 420 425 430

Asp Gln Thr Phe Trp Ile Gly Val Phe Pro Gly Ile Thr Pro Glu Met
 435 440 445

Ile Ala Tyr Val Gly Asp Thr Ile Arg Glu Phe Val Leu Lys His Ser
 450 455 460

<210> 57

<211> 463

<212> PRT

<213> Streptomyces ambofaciens

<400> 57

Met Gln Ser Asp Arg Asp Asp Arg Leu Ser Gly Gly Arg Met Pro Ser
 1 5 10 15

Leu Gly Ser Asn His Ser Gly Leu Ser Arg Glu Gly Ser Thr Met Gly
 20 25 30

Asp Leu Arg Asn Arg Ile Thr Glu Leu Val Arg Ala Tyr His Arg Glu
 35 40 45

Gln Ala Pro Gly Gly Phe Val Pro Gly Thr Thr His Val Pro Val Ser
 50 55 60

Gly Ala Val Leu Ser Glu Glu Asp Arg Leu Ala Leu Val Glu Thr Ala
 65 70 75 80

Leu Glu Met Arg Ile Ala Ala Gly Pro Ala Ser Arg Gly Phe Glu Arg
 85 90 95

Gln Phe Ala Arg Tyr Leu Gly Leu Arg Lys Ala His Leu Thr Asn Ser

100	105	110
Gly Ser Ser Ala Asn Leu Leu Ala Leu Gly Ala Leu Thr Ser Pro Gln		
115	120	125
Leu Glu Glu Arg Arg Leu Arg Pro Gly Asp Glu Val Val Thr Val Ala		
130	135	140
Ala Gly Phe Pro Thr Thr Val Asn Pro Ile Phe His Asn Gly Leu Val		
145	150	155
Pro Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Thr Thr Pro Glu		
	165	170
Arg Ile Glu Arg Ala Ile Gly Pro Arg Thr Arg Ala Ile Met Ile Ala		
	180	185
His Ala Leu Gly Asn Pro Phe Glu Ala Glu Glu Val Ala Arg Leu Ala		
	195	200
Asp Glu Arg Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp Ala Val Gly		
	210	215
Ser Arg Tyr Arg Gly Arg Leu Thr Gly Ser Phe Gly Asp Leu Ser Thr		
225	230	235
Val Ser Phe Tyr Pro Ala His His Ile Ala Met Gly Glu Gly Gly Cys		
	245	250
Val Leu Thr Asp Asn Leu Ala Leu Ala Arg Ile Val Glu Ser Leu Arg		
	260	265
Asp Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp Asn Arg Cys		
	275	280
		285

150

Leu Lys Arg Phe Asp Gln Lys Met Gly Asp Leu Pro Pro Gly Tyr Asp
 290 295 300

His Lys Tyr Ile Phe Ser His Val Gly Tyr Asn Leu Lys Ser Thr Asp
 305 310 315 320

Leu Gln Ala Ala Leu Gly Leu Ser Gln Leu Thr Arg Ile Glu Glu Phe
 325 330 335

Thr Glu Ala Arg Arg Ala Asn Trp Arg His Leu Arg Ala Ala Leu Asp
 340 345 350

Gly Leu Pro Gly Leu Leu Leu Pro His Ala Thr Pro Gly Ser Asp Pro
 355 360 365

Ser Trp Phe Gly Phe Leu Ile Thr Val Asp Pro Asp Ala Ala Tyr Ser
 370 375 380

Arg Ala Ala Leu Val Asp His Leu Glu Ser Arg Arg Ile Ser Thr Arg
 385 390 395 400

Arg Leu Phe Gly Gly Asn Leu Val Arg His Pro Ala Tyr Thr Asp Arg
 405 410 415

Arg Tyr Arg Val Ser Gly Ser Leu Glu Asn Ser Asp Leu Ile Thr Asp
 420 425 430

Gln Thr Phe Trp Ile Gly Val Phe Pro Gly Ile Thr Pro Glu Met Ile
 435 440 445

Ala Tyr Val Gly Asp Thr Ile Arg Glu Phe Val Leu Lys His Ser
 450 455 460

<210> 58

<211> 450

<212> PRT

<213> Streptomyces ambofaciens

<400> 58

Met Pro Ser Leu Gly Ser Asn His Ser Gly Leu Ser Arg Glu Gly Ser
1 5 10 15

Thr Met Gly Asp Leu Arg Asn Arg Ile Thr Glu Leu Val Arg Ala Tyr
20 25 30

His Arg Glu Gln Ala Pro Gly Gly Phe Val Pro Gly Thr Thr His Val
35 40 45

Pro Val Ser Gly Ala Val Leu Ser Glu Glu Asp Arg Leu Ala Leu Val
50 55 60

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

Phe Glu Arg Gln Phe Ala Arg Tyr Leu Gly Leu Arg Lys Ala His Leu
85 90 95

Thr Asn Ser Gly Ser Ser Ala Asn Leu Leu Ala Leu Gly Ala Leu Thr
100 105 110

Ser Pro Gln Leu Glu Glu Arg Arg Leu Arg Pro Gly Asp Glu Val Val
115 120 125

Thr Val Ala Ala Gly Phe Pro Thr Thr Val Asn Pro Ile Phe His Asn
130 135 140

Gly Leu Val Pro Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Thr
145 150 155 160

Thr Pro Glu Arg Ile Glu Arg Ala Ile Gly Pro Arg Thr Arg Ala Ile
165 170 175

Met Ile Ala His Ala Leu Gly Asn Pro Phe Glu Ala Glu Glu Val Ala
 180 185 190

Arg Leu Ala Asp Glu Arg Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp
 195 200 205

Ala Val Gly Ser Arg Tyr Arg Gly Arg Leu Thr Gly Ser Phe Gly Asp
 210 215 220

Leu Ser Thr Val Ser Phe Tyr Pro Ala His His Ile Ala Met Gly Glu
 225 230 235 240

Gly Gly Cys Val Leu Thr Asp Asn Leu Ala Leu Ala Arg Ile Val Glu
 245 250 255

Ser Leu Arg Asp Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp
 260 265 270

Asn Arg Cys Leu Lys Arg Phe Asp Gln Lys Met Gly Asp Leu Pro Pro
 275 280 285

Gly Tyr Asp His Lys Tyr Ile Phe Ser His Val Gly Tyr Asn Leu Lys
 290 295 300

Ser Thr Asp Leu Gln Ala Ala Leu Gly Leu Ser Gln Leu Thr Arg Ile
 305 310 315 320

Glu Glu Phe Thr Glu Ala Arg Arg Ala Asn Trp Arg His Leu Arg Ala
 325 330 335

Ala Leu Asp Gly Leu Pro Gly Leu Leu Leu Pro His Ala Thr Pro Gly
 340 345 350

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

Ala Tyr Ser Arg Ala Ala Leu Val Asp His Leu Glu Ser Arg Arg Ile
 370 375 380

Ser Thr Arg Arg Leu Phe Gly Gly Asn Leu Val Arg His Pro Ala Tyr
 385 390 395 400

Thr Asp Arg Arg Tyr Arg Val Ser Gly Ser Leu Glu Asn Ser Asp Leu
 405 410 415

Ile Thr Asp Gln Thr Phe Trp Ile Gly Val Phe Pro Gly Ile Thr Pro
 420 425 430

Glu Met Ile Ala Tyr Val Gly Asp Thr Ile Arg Glu Phe Val Leu Lys
 435 440 445

His Ser
 450

<210> 59
 <211> 433
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 59

Met Gly Asp Leu Arg Asn Arg Ile Thr Glu Leu Val Arg Ala Tyr His
 1 5 10 15

Arg Glu Gln Ala Pro Gly Gly Phe Val Pro Gly Thr Thr His Val Pro
 20 25 30

Val Ser Gly Ala Val Leu Ser Glu Glu Asp Arg Leu Ala Leu Val Glu
 35 40 45

Thr Ala Leu Glu Met Arg Ile Ala Ala Gly Pro Ala Ser Arg Gly Phe
 50 55 60

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

Asn Ser Gly Ser Ser Ala Asn Leu Leu Ala Leu Gly Ala Leu Thr Ser
 85 90 95

Pro Gln Leu Glu Glu Arg Arg Leu Arg Pro Gly Asp Glu Val Val Thr
 100 105 110

Val Ala Ala Gly Phe Pro Thr Thr Val Asn Pro Ile Phe His Asn Gly
 115 120 125

Leu Val Pro Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Thr Thr
 130 135 140

Pro Glu Arg Ile Glu Arg Ala Ile Gly Pro Arg Thr Arg Ala Ile Met
 145 150 155 160

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

Leu Ala Asp Glu Arg Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp Ala
 180 185 190

Val Gly Ser Arg Tyr Arg Gly Arg Leu Thr Gly Ser Phe Gly Asp Leu
 195 200 205

Ser Thr Val Ser Phe Tyr Pro Ala His His Ile Ala Met Gly Glu Gly
 210 215 220

Gly Cys Val Leu Thr Asp Asn Leu Ala Leu Ala Arg Ile Val Glu Ser
 225 230 235 240

Leu Arg Asp Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp Asn

Ser

<210> 60
 <211> 1398
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1398)

<400> 60
 atg gac gac acg atg gcc ggc gcc gac gcc gag gaa tgg gac ggc gat 48
 Met Asp Asp Thr Met Ala Gly Ala Asp Ala Glu Glu Trp Asp Gly Asp
 1 5 10 15
 cag tta gac cgc gag gac cgg gcc tcc ctc cgc cgc gtc gcg ggg ctc 96
 Gln Leu Asp Arg Glu Asp Arg Ala Ser Leu Arg Arg Val Ala Gly Leu
 20 25 30
 tcc acc gaa ctg acc gac gtc tcc gag gtc gag tac cgc cag ctg cga 144
 Ser Thr Glu Leu Thr Asp Val Ser Glu Val Glu Tyr Arg Gln Leu Arg
 35 40 45
 ctc gag cgg gtg gtg ctc gtc ggc atc tgg acc tcg gga acg gcc gcg 192
 Leu Glu Arg Val Val Leu Val Gly Ile Trp Thr Ser Gly Thr Ala Ala
 50 55 60
 gag gcc gac agt tcg ctc gcc gag ctg gcg gcg ctc gcc gag acc gcg 240
 Glu Ala Asp Ser Ser Leu Ala Glu Leu Ala Ala Leu Ala Glu Thr Ala
 65 70 75 80
 ggc gcc ctc gtg ctg gac ggc gtc gtg cag cgc cgg cag aag ccg gac 288
 Gly Ala Leu Val Leu Asp Gly Val Val Gln Arg Arg Gln Lys Pro Asp
 85 90 95
 ccg gcg acg tac atc ggc tcg ggc aag gcg tcg cag ctg cgc gac atc 336
 Pro Ala Thr Tyr Ile Gly Ser Gly Lys Ala Ser Gln Leu Arg Asp Ile
 100 105 110
 gtc gag gag acc ggc gcc gac acc gtg gtg tgc gac ggg gaa ctg agc 384
 Val Glu Glu Thr Gly Ala Asp Thr Val Val Cys Asp Gly Glu Leu Ser
 115 120 125

ccc agt cag ctg atg cac ctg gag gag gtc gtc ggg gtc aag gtc gtg	432
Pro Ser Gln Leu Met His Leu Glu Glu Val Val Gly Val Lys Val Val	
130 135 140	
gac cgc acg gcc ctg atc ctg gac atc ttc gcg cag cac gcc cag tcc	480
Asp Arg Thr Ala Leu Ile Leu Asp Ile Phe Ala Gln His Ala Gln Ser	
145 150 155 160	
cgg gag ggc aag gcg cag gtg gcg ctg gcg cag atg cag tac atg ctg	528
Arg Glu Gly Lys Ala Gln Val Ala Leu Ala Gln Met Gln Tyr Met Leu	
165 170 175	
ccg cgg ctg cgc ggc tgg ggc cag tcg ctg tcc cgg cag atg ggc ggc	576
Pro Arg Leu Arg Gly Trp Gly Gln Ser Leu Ser Arg Gln Met Gly Gly	
180 185 190	
ggt ggc ggc ggt ggc atg gcc acg cgc ggt ccc ggt gag acg aag atc	624
Gly Gly Gly Gly Gly Met Ala Thr Arg Gly Pro Gly Glu Thr Lys Ile	
195 200 205	
gag acg gac cgg cgg cgg atc aac gac aag atg gcc agg ctc cgc cgg	672
Glu Thr Asp Arg Arg Arg Ile Asn Asp Lys Met Ala Arg Leu Arg Arg	
210 215 220	
gag ctg gag cag ctg aag acc ggc cgg gac gtg aac cgg gag gag cga	720
Glu Leu Glu Gln Leu Lys Thr Gly Arg Asp Val Asn Arg Glu Glu Arg	
225 230 235 240	
cgg cgc aac aag gtg ctg tcg gtc gcc ctc gcc ggc tac acc aac gcc	768
Arg Arg Asn Lys Val Leu Ser Val Ala Leu Ala Gly Tyr Thr Asn Ala	
245 250 255	
ggc aag tca tcg ctg ctc aac cgc ctc acc gga gcc ggc gtg ctg gtg	816
Gly Lys Ser Ser Leu Leu Asn Arg Leu Thr Gly Ala Gly Val Leu Val	
260 265 270	
gag aac gcc ctg ttc gcc acc ctg gac acg acc gtg cgg cgg gcg acg	864
Glu Asn Ala Leu Phe Ala Thr Leu Asp Thr Thr Val Arg Arg Ala Thr	
275 280 285	
acg ccg agc ggg cgc ccc tac acc atc gcc gac acc gtg ggc ttc gta	912
Thr Pro Ser Gly Arg Pro Tyr Thr Ile Ala Asp Thr Val Gly Phe Val	
290 295 300	
cgc cac ctc ccg cac cac ctg gtg gag gcg ttc cgt tcc acg atc gaa	960
Arg His Leu Pro His His Leu Val Glu Ala Phe Arg Ser Thr Ile Glu	
305 310 315 320	

gag gtc gcg gac gcg cat ctg gtg ctg cac gtg gtc gac ggt tgc cac	1008
Glu Val Ala Asp Ala His Leu Val Leu His Val Val Asp Gly Ser His	
325 330 335	
ccg gac ccc ggc gcg cag ctg gcc tgc gtg cgc gag gtg ctg cgg gac	1056
Pro Asp Pro Gly Ala Gln Leu Ala Ser Val Arg Glu Val Leu Arg Asp	
340 345 350	
gtg ggc gcc ggc gag tcc acc gag gtc gtg gtc gtc aac aag gcc gat	1104
Val Gly Ala Gly Glu Ser Thr Glu Val Val Val Val Asn Lys Ala Asp	
355 360 365	
gtc gcg gac ccg gac gtc ctc gcc cgt ctg ctg gag cag gag ccg gac	1152
Val Ala Asp Pro Asp Val Leu Ala Arg Leu Leu Glu Gln Glu Pro Asp	
370 375 380	
gcg atc gtc gtg tcc gcc cgt tgc ggt cgg ggc atc gac gag ctg cgg	1200
Ala Ile Val Val Ser Ala Arg Ser Gly Arg Gly Ile Asp Glu Leu Arg	
385 390 395 400	
gaa ctg atc gac cgc ctg ctg ccg cac ccc gag gtg gag gtg gag gtc	1248
Glu Leu Ile Asp Arg Leu Leu Pro His Pro Glu Val Glu Val Glu Val	
405 410 415	
gtg atc ccc tac gac gag ggg agc ctg gtg gcg cgg gtg cac gac gag	1296
Val Ile Pro Tyr Asp Glu Gly Ser Leu Val Ala Arg Val His Asp Glu	
420 425 430	
ggc gag gtg ctc agc acc gag cac acg ccg gag ggc acc ctg ctc acc	1344
Gly Glu Val Leu Ser Thr Glu His Thr Pro Glu Gly Thr Leu Leu Thr	
435 440 445	
gcg cgg gtc cac ccc gac ctc gcc tgc gag ctt cgg acg ctt ccg cgg	1392
Ala Arg Val His Pro Asp Leu Ala Ser Glu Leu Arg Thr Leu Pro Arg	
450 455 460	
ccg tga	1398
Pro	
465	

<210> 61

<211> 465

<212> PRT

<213> Streptomyces ambofaciens

<400> 61

Met Asp Asp Thr Met Ala Gly Ala Asp Ala Glu Glu Trp Asp Gly Asp
 1 5 10 15

Gln Leu Asp Arg Glu Asp Arg Ala Ser Leu Arg Arg Val Ala Gly Leu
 20 25 30

Ser Thr Glu Leu Thr Asp Val Ser Glu Val Glu Tyr Arg Gln Leu Arg
 35 40 45

Leu Glu Arg Val Val Leu Val Gly Ile Trp Thr Ser Gly Thr Ala Ala
 50 55 60

Glu Ala Asp Ser Ser Leu Ala Glu Leu Ala Ala Leu Ala Glu Thr Ala
 65 70 75 80

Gly Ala Leu Val Leu Asp Gly Val Val Gln Arg Arg Gln Lys Pro Asp
 85 90 95

Pro Ala Thr Tyr Ile Gly Ser Gly Lys Ala Ser Gln Leu Arg Asp Ile
 100 105 110

Val Glu Glu Thr Gly Ala Asp Thr Val Val Cys Asp Gly Glu Leu Ser
 115 120 125

Pro Ser Gln Leu Met His Leu Glu Glu Val Val Gly Val Lys Val Val
 130 135 140

Asp Arg Thr Ala Leu Ile Leu Asp Ile Phe Ala Gln His Ala Gln Ser
 145 150 155 160

Arg Glu Gly Lys Ala Gln Val Ala Leu Ala Gln Met Gln Tyr Met Leu
 165 170 175

Pro Arg Leu Arg Gly Trp Gly Gln Ser Leu Ser Arg Gln Met Gly Gly

180

185

190

Gly Gly Gly Gly Gly Met Ala Thr Arg Gly Pro Gly Glu Thr Lys Ile
 195 200 205

Glu Thr Asp Arg Arg Arg Ile Asn Asp Lys Met Ala Arg Leu Arg Arg
 210 215 220

Glu Leu Glu Gln Leu Lys Thr Gly Arg Asp Val Asn Arg Glu Glu Arg
 225 230 235 240

Arg Arg Asn Lys Val Leu Ser Val Ala Leu Ala Gly Tyr Thr Asn Ala
 245 250 255

Gly Lys Ser Ser Leu Leu Asn Arg Leu Thr Gly Ala Gly Val Leu Val
 260 265 270

Glu Asn Ala Leu Phe Ala Thr Leu Asp Thr Thr Val Arg Arg Ala Thr
 275 280 285

Thr Pro Ser Gly Arg Pro Tyr Thr Ile Ala Asp Thr Val Gly Phe Val
 290 295 300

Arg His Leu Pro His His Leu Val Glu Ala Phe Arg Ser Thr Ile Glu
 305 310 315 320

Glu Val Ala Asp Ala His Leu Val Leu His Val Val Asp Gly Ser His
 325 330 335

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

Val Gly Ala Gly Glu Ser Thr Glu Val Val Val Val Asn Lys Ala Asp
 355 360 365

Val Ala Asp Pro Asp Val Leu Ala Arg Leu Leu Glu Gln Glu Pro Asp
 370 375 380

Ala Ile Val Val Ser Ala Arg Ser Gly Arg Gly Ile Asp Glu Leu Arg
 385 390 395 400

Glu Leu Ile Asp Arg Leu Leu Pro His Pro Glu Val Glu Val Glu Val
 405 410 415

Val Ile Pro Tyr Asp Glu Gly Ser Leu Val Ala Arg Val His Asp Glu
 420 425 430

Gly Glu Val Leu Ser Thr Glu His Thr Pro Glu Gly Thr Leu Leu Thr
 435 440 445

Ala Arg Val His Pro Asp Leu Ala Ser Glu Leu Arg Thr Leu Pro Arg
 450 455 460

Pro
 465

<210> 62
 <211> 1044
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1044)

<400> 62
 atg cga gta ctc att atc ggg ggt tca cag ttc gtg ggc cgg gcc ttc 48
 Met Arg Val Leu Ile Ile Gly Gly Ser Gln Phe Val Gly Arg Ala Phe
 1 5 10 15
 gcc gcc gag gca ctg gcc gcg ggg cac cgg gtc acc acg ttc aac cgg 96
 Ala Ala Glu Ala Leu Ala Ala Gly His Arg Val Thr Thr Phe Asn Arg
 20 25 30

ggt gtc agc ggc acc gac ctg ccc ggc gtc gag gcg gtc agg ggc gac	144
Gly Val Ser Gly Thr Asp Leu Pro Gly Val Glu Ala Val Arg Gly Asp	
35 40 45	
cgc gag gtg gcc ggc gac ctg gag cgg ctg gtg tcc gga agg cac tgg	192
Arg Glu Val Ala Gly Asp Leu Glu Arg Leu Val Ser Gly Arg His Trp	
50 55 60	
gac gcg gtc gtg gac acc tgc ggt tac gtg ccc cgc acg gtg ggg gcc	240
Asp Ala Val Val Asp Thr Cys Gly Tyr Val Pro Arg Thr Val Gly Ala	
65 70 75 80	
tcg gcc gcg gcg ctg tcc ggg cac gcg gac acc tac ctc tac gtc tcc	288
Ser Ala Ala Ala Leu Ser Gly His Ala Asp Thr Tyr Leu Tyr Val Ser	
85 90 95	
agc atc gcc tgc ctg ccc gac tgg gcg cag gcg gtc cgt ccg gtg gac	336
Ser Ile Ala Cys Leu Pro Asp Trp Ala Gln Ala Val Arg Pro Val Asp	
100 105 110	
gac gac tca cct gcc cac gac tgc ccg ccg gac gcc gga ccg gac cac	384
Asp Asp Ser Pro Ala His Asp Cys Pro Pro Asp Ala Gly Pro Asp His	
115 120 125	
gcc gac ggt gac tac ggc gtc ctg aag gcc ggc tgc gag cgc gcc gtg	432
Ala Asp Gly Asp Tyr Gly Val Leu Lys Ala Gly Cys Glu Arg Ala Val	
130 135 140	
gac cgg cac ttc gcg ggc cgg acc ctg cac ctg cgg gcc ggt gtc atc	480
Asp Arg His Phe Ala Gly Arg Thr Leu His Leu Arg Ala Gly Val Ile	
145 150 155 160	
ctc ggg ccg cac gac acc atg cgc atg ctc gac gcc tgg ctg tgg cgc	528
Leu Gly Pro His Asp Thr Met Arg Met Leu Asp Ala Trp Leu Trp Arg	
165 170 175	
atg cgc gtc gcc gag ggg gag cac cgc cgg gtg ctc gcc ccg ggc aac	576
Met Arg Val Ala Glu Gly Glu His Arg Arg Val Leu Ala Pro Gly Asn	
180 185 190	
ccc gag gtg ggc atg cgc ctg atc gac gta cgc gat gtc gcc gtc ttc	624
Pro Glu Val Gly Met Arg Leu Ile Asp Val Arg Asp Val Ala Val Phe	
195 200 205	
ggc ctc gac tgc ctc gcg gac ggc cgt acc ggc gcc ttc atc gtc aac	672
Gly Leu Asp Cys Leu Ala Asp Gly Arg Thr Gly Ala Phe Ile Val Asn	
210 215 220	

ccg ccg gag aag aac acc acc ttc ggg gag ttg ctc acg gag tgc gtc 720
 Pro Pro Glu Lys Asn Thr Thr Phe Gly Glu Leu Leu Thr Glu Cys Val
 225 230 235 240

aag gcc acc ggt tcg gcc gcg gag ccg gtg tgg gtc gac gag ggg ttc 768
 Lys Ala Thr Gly Ser Ala Ala Glu Pro Val Trp Val Asp Glu Gly Phe
 245 250 255

ctc gcc gag cac ggc gtg agt ccg tgg acg gac ctg ccg atg tgg gtg 816
 Leu Ala Glu His Gly Val Ser Pro Trp Thr Asp Leu Pro Met Trp Val
 260 265 270

ccc gac acc gcg cgg gac acc ctc gtg tgg gcg gcc gga gca ccg cgc 864
 Pro Asp Thr Ala Arg Asp Thr Leu Val Trp Ala Ala Gly Ala Pro Arg
 275 280 285

gcc cgg gcc gcg ggt ctg gcc tgc cgg ccc ttc tcc gac acc gtg cgg 912
 Ala Arg Ala Ala Gly Leu Ala Cys Arg Pro Phe Ser Asp Thr Val Arg
 290 295 300

gac gcc tgg gag gtc gtc cgg gac cgg ccc gtc ccg gaa ctg ccg ctc 960
 Asp Ala Trp Glu Val Val Arg Asp Arg Pro Val Pro Glu Leu Pro Leu
 305 310 315 320

gcg gcc ggc tgc ggc ctg tcc ctg agc cgg gag aag gag ctg ctc gcc 1008
 Ala Ala Gly Cys Gly Leu Ser Leu Ser Arg Glu Lys Glu Leu Leu Ala
 325 330 335

gcc tgg gac gct cgc ggc ggt gcg gcg gcg ggc tga 1044
 Ala Trp Asp Ala Arg Gly Gly Ala Ala Gly
 340 345

<210> 63

<211> 347

<212> PRT

<213> Streptomyces ambofaciens

<400> 63

Met Arg Val Leu Ile Ile Gly Gly Ser Gln Phe Val Gly Arg Ala Phe
 1 5 10 15

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

Gly Val Ser Gly Thr Asp Leu Pro Gly Val Glu Ala Val Arg Gly Asp
 35 40 45

Arg Glu Val Ala Gly Asp Leu Glu Arg Leu Val Ser Gly Arg His Trp
 50 55 60

Asp Ala Val Val Asp Thr Cys Gly Tyr Val Pro Arg Thr Val Gly Ala
 65 70 75 80

Ser Ala Ala Ala Leu Ser Gly His Ala Asp Thr Tyr Leu Tyr Val Ser
 85 90 95

Ser Ile Ala Cys Leu Pro Asp Trp Ala Gln Ala Val Arg Pro Val Asp
 100 105 110

Asp Asp Ser Pro Ala His Asp Cys Pro Pro Asp Ala Gly Pro Asp His
 115 120 125

Ala Asp Gly Asp Tyr Gly Val Leu Lys Ala Gly Cys Glu Arg Ala Val
 130 135 140

Asp Arg His Phe Ala Gly Arg Thr Leu His Leu Arg Ala Gly Val Ile
 145 150 155 160

Leu Gly Pro His Asp Thr Met Arg Met Leu Asp Ala Trp Leu Trp Arg
 165 170 175

Met Arg Val Ala Glu Gly Glu His Arg Arg Val Leu Ala Pro Gly Asn
 180 185 190

Pro Glu Val Gly Met Arg Leu Ile Asp Val Arg Asp Val Ala Val Phe
 195 200 205

Gly Leu Asp Cys Leu Ala Asp Gly Arg Thr Gly Ala Phe Ile Val Asn

210		215		220
Pro Pro Glu Lys Asn Thr Thr Phe Gly Glu Leu Leu Thr Glu Cys Val				
225		230		240
Lys Ala Thr Gly Ser Ala Ala Glu Pro Val Trp Val Asp Glu Gly Phe				
	245		250	255
Leu Ala Glu His Gly Val Ser Pro Trp Thr Asp Leu Pro Met Trp Val				
	260		265	270
Pro Asp Thr Ala Arg Asp Thr Leu Val Trp Ala Ala Gly Ala Pro Arg				
	275		280	285
Ala Arg Ala Ala Gly Leu Ala Cys Arg Pro Phe Ser Asp Thr Val Arg				
	290		295	300
Asp Ala Trp Glu Val Val Arg Asp Arg Pro Val Pro Glu Leu Pro Leu				
305		310		320
Ala Ala Gly Cys Gly Leu Ser Leu Ser Arg Glu Lys Glu Leu Leu Ala				
	325		330	335
Ala Trp Asp Ala Arg Gly Gly Ala Ala Ala Gly				
	340		345	

<210> 64
 <211> 1041
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1041)

<400> 64
 atg acc gcc ctg ggt act tct gct gaa ccg ttc gcc gca ccc ggc ccg

Met	Thr	Ala	Leu	Gly	Thr	Ser	Ala	Glu	Pro	Phe	Ala	Ala	Pro	Gly	Pro	
1				5					10					15		
ccg	cgg	ccg	gaa	gcg	tcc	ccg	gtg	ctg	cgc	ttc	gga	gcg	atc	ggc	tgc	96
Pro	Arg	Pro	Glu	Ala	Ser	Pro	Val	Leu	Arg	Phe	Gly	Ala	Ile	Gly	Cys	
			20					25					30			
ggc	gac	atc	gcg	ggc	cgt	cgc	acc	ctg	ccc	gcc	ctg	ctc	tcc	acc	ccc	144
Gly	Asp	Ile	Ala	Gly	Arg	Arg	Thr	Leu	Pro	Ala	Leu	Leu	Ser	Thr	Pro	
		35					40					45				
ggc	acc	gtg	ctg	acc	tgc	gtc	ggc	agc	cgc	gac	gcg	gac	cgg	gcc	aag	192
Gly	Thr	Val	Leu	Thr	Cys	Val	Gly	Ser	Arg	Asp	Ala	Asp	Arg	Ala	Lys	
	50					55				60						
gcc	ctg	ggc	agg	cac	ttc	gac	tgc	gag	gcg	gtc	gcg	ccc	tac	gag	gcc	240
Ala	Leu	Gly	Arg	His	Phe	Asp	Cys	Glu	Ala	Val	Ala	Pro	Tyr	Glu	Ala	
65					70					75				80		
ctg	ctg	gaa	cgc	ccc	gac	gtg	gac	gcc	gtc	tac	atc	gcg	gtg	ccc	agc	288
Leu	Leu	Glu	Arg	Pro	Asp	Val	Asp	Ala	Val	Tyr	Ile	Ala	Val	Pro	Ser	
				85					90					95		
atg	ctg	cac	gcc	gaa	tgg	gcc	gcg	gcg	gcg	ctg	cgg	gcg	ggc	aaa	cac	336
Met	Leu	His	Ala	Glu	Trp	Ala	Ala	Ala	Ala	Leu	Arg	Ala	Gly	Lys	His	
			100					105					110			
gtc	ctc	gtg	gag	aag	ccc	gcc	gcc	gcc	aac	cac	gcc	gac	gcg	gcc	cgc	384
Val	Leu	Val	Glu	Lys	Pro	Ala	Ala	Ala	Asn	His	Ala	Asp	Ala	Ala	Arg	
			115					120					125			
ctg	ttc	gcc	atg	gcc	cgg	gag	cgc	gga	ctg	gta	ctg	atg	gag	aac	ttc	432
Leu	Phe	Ala	Met	Ala	Arg	Glu	Arg	Gly	Leu	Val	Leu	Met	Glu	Asn	Phe	
		130						135				140				
atg	ttc	ctc	cac	cac	tct	caa	cac	gcc	acc	gtc	aaa	gcc	ctg	ttg	gag	480
Met	Phe	Leu	His	His	Ser	Gln	His	Ala	Thr	Val	Lys	Ala	Leu	Leu	Glu	
145					150					155				160		
gcc	ggt	gcc	atc	gga	gag	ctg	cgc	acc	ttc	tcc	gcc	gcc	ttc	acc	atc	528
Ala	Gly	Ala	Ile	Gly	Glu	Leu	Arg	Thr	Phe	Ser	Ala	Ala	Phe	Thr	Ile	
				165					170				175			
ccg	ccg	cgc	gcc	gac	gac	gac	atg	cgc	tac	cga	ccc	gac	atc	ggc	ggc	576
Pro	Pro	Arg	Ala	Asp	Asp	Asp	Met	Arg	Tyr	Arg	Pro	Asp	Ile	Gly	Gly	
			180					185					190			

ggc gca ctg ctc gac aac ggc gtc tac ccg ctg cgg gcc gcg ctg cac	624
Gly Ala Leu Leu Asp Asn Gly Val Tyr Pro Leu Arg Ala Ala Leu His	
195 200 205	
ttc ctc gga ccc gaa ctg cgc ctg acc ggc gcc gtg ctg cgc cgt gac	672
Phe Leu Gly Pro Glu Leu Arg Leu Thr Gly Ala Val Leu Arg Arg Asp	
210 215 220	
cgc cgc cgg ggc gtc gtg gtc tcc ggc agc gtc ctg ctc gcc gcc ccc	720
Arg Arg Arg Gly Val Val Val Ser Gly Ser Val Leu Leu Ala Ala Pro	
225 230 235 240	
cag ggc gtg gcc gcc cac ctc gcc ttc ggc atg gaa cac ggc tac cgc	768
Gln Gly Val Ala Ala His Leu Ala Phe Gly Met Glu His Gly Tyr Arg	
245 250 255	
tcg gcg tac gag ctg cac ggc agc acc ggc tcg ctc gcc ctg agc cac	816
Ser Ala Tyr Glu Leu His Gly Ser Thr Gly Ser Leu Ala Leu Ser His	
260 265 270	
gtc ttc acg acc ccg gac agc cac cac ccc gta ctg cgg ctg tcc cgc	864
Val Phe Thr Thr Pro Asp Ser His His Pro Val Leu Arg Leu Ser Arg	
275 280 285	
cag gac cac cgg gag gag cgc gtc ctg ccc gtg gac cgg cac ttc gtg	912
Gln Asp His Arg Glu Glu Arg Val Leu Pro Val Asp Arg His Phe Val	
290 295 300	
aac atc ctg tcg gtc ttc cgc cgt gcg gtg atc cgg gcc gag gac gtc	960
Asn Ile Leu Ser Val Phe Arg Arg Ala Val Ile Arg Ala Glu Asp Val	
305 310 315 320	
tcc gcc gag tcg tac gcc gcg ctg cgc cag gcg ggg ctg gtc gac gag	1008
Ser Ala Glu Ser Tyr Ala Ala Leu Arg Gln Ala Gly Leu Val Asp Glu	
325 330 335	
atc gtg gcc cgc gcc gag acc ttc acc gtg tag	1041
Ile Val Ala Arg Ala Glu Thr Phe Thr Val	
340 345	

<210> 65

<211> 346

<212> PRT

<213> Streptomyces ambofaciens

<400> 65

Met Thr Ala Leu Gly Thr Ser Ala Glu Pro Phe Ala Ala Pro Gly Pro
 1 5 10 15

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

Gly Asp Ile Ala Gly Arg Arg Thr Leu Pro Ala Leu Leu Ser Thr Pro
 35 40 45

Gly Thr Val Leu Thr Cys Val Gly Ser Arg Asp Ala Asp Arg Ala Lys
 50 55 60

Ala Leu Gly Arg His Phe Asp Cys Glu Ala Val Ala Pro Tyr Glu Ala
 65 70 75 80

Leu Leu Glu Arg Pro Asp Val Asp Ala Val Tyr Ile Ala Val Pro Ser
 85 90 95

Met Leu His Ala Glu Trp Ala Ala Ala Ala Leu Arg Ala Gly Lys His
 100 105 110

Val Leu Val Glu Lys Pro Ala Ala Ala Asn His Ala Asp Ala Ala Arg
 115 120 125

Leu Phe Ala Met Ala Arg Glu Arg Gly Leu Val Leu Met Glu Asn Phe
 130 135 140

Met Phe Leu His His Ser Gln His Ala Thr Val Lys Ala Leu Leu Glu
 145 150 155 160

Ala Gly Ala Ile Gly Glu Leu Arg Thr Phe Ser Ala Ala Phe Thr Ile
 165 170 175

Pro Pro Arg Ala Asp Asp Asp Met Arg Tyr Arg Pro Asp Ile Gly Gly
 180 185 190

Gly Ala Leu Leu Asp Asn Gly Val Tyr Pro Leu Arg Ala Ala Leu His
 195 200 205

Phe Leu Gly Pro Glu Leu Arg Leu Thr Gly Ala Val Leu Arg Arg Asp
 210 215 220

Arg Arg Arg Gly Val Val Val Ser Gly Ser Val Leu Leu Ala Ala Pro
 225 230 235 240

Gln Gly Val Ala Ala His Leu Ala Phe Gly Met Glu His Gly Tyr Arg
 245 250 255

Ser Ala Tyr Glu Leu His Gly Ser Thr Gly Ser Leu Ala Leu Ser His
 260 265 270

Val Phe Thr Thr Pro Asp Ser His His Pro Val Leu Arg Leu Ser Arg
 275 280 285

Gln Asp His Arg Glu Glu Arg Val Leu Pro Val Asp Arg His Phe Val
 290 295 300

Asn Ile Leu Ser Val Phe Arg Arg Ala Val Ile Arg Ala Glu Asp Val
 305 310 315 320

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

Ile Val Ala Arg Ala Glu Thr Phe Thr Val
 340 345

<210> 66

<211> 1239

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1) .. (1239)

<400> 66

gtg	acg	gac	gcg	atc	acg	acc	gag	ctg	gcc	gac	cgc	gaa	ctg	ggg	cgc	48
Val	Thr	Asp	Ala	Ile	Thr	Thr	Glu	Leu	Ala	Asp	Arg	Glu	Leu	Gly	Arg	
1				5					10					15		

aga	ctg	cac	cgg	ata	cgc	ggc	gtc	cac	tgg	tat	ttc	ggc	aac	cac	ggc	96
Arg	Leu	His	Arg	Ile	Arg	Gly	Val	His	Trp	Tyr	Phe	Gly	Asn	His	Gly	
			20				25						30			

gac	ccg	tac	gcc	ctc	atc	ctg	cgc	ggc	cag	acc	gac	gac	ccg	tcg	gtg	144
Asp	Pro	Tyr	Ala	Leu	Ile	Leu	Arg	Gly	Gln	Thr	Asp	Asp	Pro	Ser	Val	
			35				40						45			

tac	gag	gag	cgg	gtc	cgc	gag	ggc	ggg	ccg	ctg	ttc	cgc	agc	cgt	acc	192
Tyr	Glu	Glu	Arg	Val	Arg	Glu	Gly	Gly	Pro	Leu	Phe	Arg	Ser	Arg	Thr	
	50					55					60					

ggg	acc	tgg	gtg	acc	gcg	gac	ccg	gag	gtg	gcc	gcg	gcc	gtg	ctg	ggc	240
Gly	Thr	Trp	Val	Thr	Ala	Asp	Pro	Glu	Val	Ala	Ala	Ala	Val	Leu	Gly	
65					70					75				80		

gac	tcg	cgc	ttc	ggc	gcg	ctg	gac	cgc	gcc	gga	cgg	cgc	ccg	gag	gag	288
Asp	Ser	Arg	Phe	Gly	Ala	Leu	Asp	Arg	Ala	Gly	Arg	Arg	Pro	Glu	Glu	
			85						90					95		

tac	ctc	cag	ccg	tcg	ccc	gcc	acg	tac	ctg	ggg	ctg	gac	cgc	gcc	gcg	336
Tyr	Leu	Gln	Pro	Ser	Pro	Ala	Thr	Tyr	Leu	Gly	Leu	Asp	Arg	Ala	Ala	
			100						105					110		

tac	gcg	cgt	ctg	cgg	cgg	gtg	gcc	gag	ccc	gtg	ctg	ggc	gcg	gac	gcc	384
Tyr	Ala	Arg	Leu	Arg	Arg	Val	Ala	Glu	Pro	Val	Leu	Gly	Ala	Asp	Ala	
			115				120					125				

gcc	gcc	gcg	tgg	cgc	cgg	ctc	ggc	gag	gac	gtc	ggg	cgc	cgg	ctg	ctc	432
Ala	Ala	Ala	Trp	Arg	Arg	Leu	Gly	Glu	Asp	Val	Gly	Arg	Arg	Leu	Leu	
		130				135					140					

gcc	ggc	cgc	ggc	tcc	ggc	ctc	gac	ctg	acg	gcg	gac	ttc	gcc	cgc	cgg	480
Ala	Gly	Arg	Gly	Ser	Gly	Leu	Asp	Leu	Thr	Ala	Asp	Phe	Ala	Arg	Arg	
145					150					155				160		

ctg	ccg	gca	ttg	gtc	ctg	gcc	gcg	tgg	ctc	ggg	gtg	ccg	ggc	gaa	cgg	528

Leu	Pro	Ala	Leu	Val	Leu	Ala	Ala	Trp	Leu	Gly	Val	Pro	Gly	Glu	Arg	
				165					170					175		
tgc	gac	gag	tgg	gag	gag	tcg	ctg	cgg	gcg	gcg	ggg	ccg	ctg	ctg	gac	576
Cys	Asp	Glu	Trp	Glu	Glu	Ser	Leu	Arg	Ala	Ala	Gly	Pro	Leu	Leu	Asp	
			180					185					190			
ggt	ctg	ctg	tgt	ccg	cag	acg	ctg	gcg	gcc	acc	cgt	gcg	gcg	gac	tcg	624
Gly	Leu	Leu	Cys	Pro	Gln	Thr	Leu	Ala	Ala	Thr	Arg	Ala	Ala	Asp	Ser	
			195					200					205			
gcc	gcc	gag	ggg	ctg	cgc	gcg	ctg	ttg	gac	gag	gtg	gtc	gcc	gcg	cgt	672
Ala	Ala	Glu	Gly	Leu	Arg	Ala	Leu	Leu	Asp	Glu	Val	Val	Ala	Ala	Arg	
		210					215				220					
ccc	ggc	ggg	tcc	ggc	gag	ggt	gcg	gtg	gcc	cgc	atg	gtc	ggc	gcc	gga	720
Pro	Gly	Gly	Ser	Gly	Glu	Gly	Ala	Val	Ala	Arg	Met	Val	Gly	Ala	Gly	
225					230					235				240		
gcc	gcc	ccc	gac	gac	gcg	gtg	gcc	gcc	gcc	gtg	tgc	ctg	gcg	ctc	tcg	768
Ala	Ala	Pro	Asp	Asp	Ala	Val	Ala	Ala	Ala	Val	Cys	Leu	Ala	Leu	Ser	
				245					250					255		
gcc	gtc	gaa	ccg	acg	acg	acc	ctg	gtg	tgc	gaa	gcg	gtc	cgg	ctg	ctg	816
Ala	Val	Glu	Pro	Thr	Thr	Thr	Leu	Val	Cys	Glu	Ala	Val	Arg	Leu	Leu	
			260					265					270			
ctc	gac	cga	ccc	gag	tgg	tgg	cgg	cgg	ttg	tgc	gac	tcc	ccc	gct	ctg	864
Leu	Asp	Arg	Pro	Glu	Trp	Trp	Arg	Arg	Leu	Cys	Asp	Ser	Pro	Ala	Leu	
			275					280					285			
gcg	ccg	gcc	gcg	gtc	cgg	cac	acc	ctg	cgg	cac	gcg	ccc	ccg	gtg	cgg	912
Ala	Pro	Ala	Ala	Val	Arg	His	Thr	Leu	Arg	His	Ala	Pro	Pro	Val	Arg	
		290					295				300					
ctg	gag	agc	cgg	gtg	gcc	cac	gag	gac	gtg	acg	gtg	gcg	gat	cgt	ccg	960
Leu	Glu	Ser	Arg	Val	Ala	His	Glu	Asp	Val	Thr	Val	Ala	Asp	Arg	Pro	
305					310				315					320		
ctg	ccc	gcc	ggg	agc	cac	gtg	gtg	gtg	ctc	gtg	ggc	gcg	gca	cgg	cgc	1008
Leu	Pro	Ala	Gly	Ser	His	Val	Val	Val	Leu	Val	Gly	Ala	Ala	Arg	Arg	
				325					330				335			
gcg	ggc	gcc	ccg	gcc	gcg	gag	ccg	gcg	gac	ctg	gcg	ggc	gca	ccg	gcg	1056
Ala	Gly	Ala	Pro	Ala	Ala	Glu	Pro	Ala	Asp	Leu	Ala	Gly	Ala	Pro	Ala	
			340					345					350			

gcg gag ctg ccg gac gac ctg tgg ttc gcg ctg tcc ggg gag ttc gtc 1104
 Ala Glu Leu Pro Asp Asp Leu Trp Phe Ala Leu Ser Gly Glu Phe Val
 355 360 365

ggc cgt gcc gcc gag acc gcg ctg ggc gtg ctg gcc gag gcc gcc ccg 1152
 Gly Arg Ala Ala Glu Thr Ala Leu Gly Val Leu Ala Glu Ala Ala Pro
 370 375 380

gga ctg cgg cgg gac ggc gac atc gtc cgg cgg cgc cgt tcc ccg gtc 1200
 Gly Leu Arg Arg Asp Gly Asp Ile Val Arg Arg Arg Arg Ser Pro Val
 385 390 395 400

ctc ggc agg tac gcg cgg ttc ccc gtc gcg tac tcc tga 1239
 Leu Gly Arg Tyr Ala Arg Phe Pro Val Ala Tyr Ser
 405 410

<210> 67
 <211> 412
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 67

Val Thr Asp Ala Ile Thr Thr Glu Leu Ala Asp Arg Glu Leu Gly Arg
 1 5 10 15

Arg Leu His Arg Ile Arg Gly Val His Trp Tyr Phe Gly Asn His Gly
 20 25 30

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

Tyr Glu Glu Arg Val Arg Glu Gly Gly Pro Leu Phe Arg Ser Arg Thr
 50 55 60

Gly Thr Trp Val Thr Ala Asp Pro Glu Val Ala Ala Ala Val Leu Gly
 65 70 75 80

Asp Ser Arg Phe Gly Ala Leu Asp Arg Ala Gly Arg Arg Pro Glu Glu
 85 90 95

Tyr Leu Gln Pro Ser Pro Ala Thr Tyr Leu Gly Leu Asp Arg Ala Ala
 100 105 110

Tyr Ala Arg Leu Arg Arg Val Ala Glu Pro Val Leu Gly Ala Asp Ala
 115 120 125

Ala Ala Ala Trp Arg Arg Leu Gly Glu Asp Val Gly Arg Arg Leu Leu
 130 135 140

Ala Gly Arg Gly Ser Gly Leu Asp Leu Thr Ala Asp Phe Ala Arg Arg
 145 150 155 160

Leu Pro Ala Leu Val Leu Ala Ala Trp Leu Gly Val Pro Gly Glu Arg
 165 170 175

Cys Asp Glu Trp Glu Glu Ser Leu Arg Ala Ala Gly Pro Leu Leu Asp
 180 185 190

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

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

Pro Gly Gly Ser Gly Glu Gly Ala Val Ala Arg Met Val Gly Ala Gly
 225 230 235 240

Ala Ala Pro Asp Asp Ala Val Ala Ala Ala Val Cys Leu Ala Leu Ser
 245 250 255

Ala Val Glu Pro Thr Thr Thr Leu Val Cys Glu Ala Val Arg Leu Leu
 260 265 270

Leu Asp Arg Pro Glu Trp Trp Arg Arg Leu Cys Asp Ser Pro Ala Leu
 275 280 285

Ala Pro Ala Ala Val Arg His Thr Leu Arg His Ala Pro Pro Val Arg
 290 295 300

Leu Glu Ser Arg Val Ala His Glu Asp Val Thr Val Ala Asp Arg Pro
 305 310 315 320

Leu Pro Ala Gly Ser His Val Val Val Leu Val Gly Ala Ala Arg Arg
 325 330 335

Ala Gly Ala Pro Ala Ala Glu Pro Ala Asp Leu Ala Gly Ala Pro Ala
 340 345 350

Ala Glu Leu Pro Asp Asp Leu Trp Phe Ala Leu Ser Gly Glu Phe Val
 355 360 365

Gly Arg Ala Ala Glu Thr Ala Leu Gly Val Leu Ala Glu Ala Ala Pro
 370 375 380

Gly Leu Arg Arg Asp Gly Asp Ile Val Arg Arg Arg Arg Ser Pro Val
 385 390 395 400

Leu Gly Arg Tyr Ala Arg Phe Pro Val Ala Tyr Ser
 405 410

<210> 68
 <211> 1272
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1272)

<400> 68
 gtg cgc gtc ctg gtg acc tcc atc ccg cac cac acg cac tac tac cac
 Val Arg Val Leu Val Thr Ser Ile Pro His His Thr His Tyr Tyr His

1	5	10	15	
ctg gta ccg ctg atc tgg gct ctg cgt gcc tcg ggg cac gag gtg gtg				96
Leu Val Pro Leu Ile Trp Ala Leu Arg Ala Ser Gly His Glu Val Val				
	20	25	30	
gcg gcc ggc cag ccg tcg ctg gtc gac gcc atc acc gcc agc ggc atc				144
Ala Ala Gly Gln Pro Ser Leu Val Asp Ala Ile Thr Ala Ser Gly Ile				
	35	40	45	
ccg gcg ttc gcc ctg gcc gag gag gag tcg ctg gcg cag atc ttc gag				192
Pro Ala Phe Ala Leu Ala Glu Glu Glu Ser Leu Ala Gln Ile Phe Glu				
	50	55	60	
gag gtc gag ggc gat ctc cag ccg tat cag cac ggc atc gac gag ttc				240
Glu Val Glu Gly Asp Leu Gln Pro Tyr Gln His Gly Ile Asp Glu Phe				
65	70	75	80	
gac ttc ttc ggc acc ctg aag gac gag ctg gac tgg gag aag ctg ctc				288
Asp Phe Phe Gly Thr Leu Lys Asp Glu Leu Asp Trp Glu Lys Leu Leu				
	85	90	95	
gcc cag cag gtg atc ctg tcc ggc ctg tgg ctg gaa ccg ctc aac ggc				336
Ala Gln Gln Val Ile Leu Ser Gly Leu Trp Leu Glu Pro Leu Asn Gly				
	100	105	110	
gcc acg acc ctc gac agc atc gtc gac ttc gcc cgg gcc tgg aag ccc				384
Ala Thr Thr Leu Asp Ser Ile Val Asp Phe Ala Arg Ala Trp Lys Pro				
	115	120	125	
gac ctg gtg ctg tgg gag ccg ttc acc tat gcg ggg ccg gtg gcg gcc				432
Asp Leu Val Leu Trp Glu Pro Phe Thr Tyr Ala Gly Pro Val Ala Ala				
130	135	140		
cgg gcg tgc ggg gcc gcg cac gcc cgc gtc ctg tgg ggg ccg gac acg				480
Arg Ala Cys Gly Ala Ala His Ala Arg Val Leu Trp Gly Pro Asp Thr				
145	150	155	160	
atc ggg ctg ctg cgg acg aag ttc ctt cag gcc cag gcg cgt cag ccc				528
Ile Gly Leu Leu Arg Thr Lys Phe Leu Gln Ala Gln Ala Arg Gln Pro				
	165	170	175	
gag gag cac cgg gac gac ccg gtc gcg gag tgg atg acc tgg gcc ctg				576
Glu Glu His Arg Asp Asp Pro Val Ala Glu Trp Met Thr Trp Ala Leu				
	180	185	190	
gcg cgc tac ggg tgc gac ttc cgg gag gag gac gtg ctc ggt cag tgg				624

Ala	Arg	Tyr	Gly	Cys	Asp	Phe	Arg	Glu	Glu	Asp	Val	Leu	Gly	Gln	Trp	
	195						200					205				
agc	gtg	gac	ccg	atg	gcg	gag	ggc	gtc	agt	ctg	ggc	ctc	gac	ctg	ccg	672
Ser	Val	Asp	Pro	Met	Ala	Glu	Gly	Val	Ser	Leu	Gly	Leu	Asp	Leu	Pro	
	210					215				220						
acc	gtc	ccg	atg	cgc	tac	acc	ccg	tac	aac	ggg	tcg	gcg	gtg	atc	ccc	720
Thr	Val	Pro	Met	Arg	Tyr	Thr	Pro	Tyr	Asn	Gly	Ser	Ala	Val	Ile	Pro	
	225				230				235					240		
gac	tgg	ctg	acc	gag	gaa	ccg	aaa	cgg	cct	cgg	gtc	tgc	ctg	acc	ctg	768
Asp	Trp	Leu	Thr	Glu	Glu	Pro	Lys	Arg	Pro	Arg	Val	Cys	Leu	Thr	Leu	
				245				250					255			
ggg	gtg	tcc	tcg	cgg	gag	cac	agt	gag	gac	gag	gtc	ccg	gtg	cag	agg	816
Gly	Val	Ser	Ser	Arg	Glu	His	Ser	Glu	Asp	Glu	Val	Pro	Val	Gln	Arg	
			260					265				270				
ttt	atc	gag	gcg	ctg	gcc	gat	ctc	gac	atc	gag	ctg	gtg	gcg	acc	ctg	864
Phe	Ile	Glu	Ala	Leu	Ala	Asp	Leu	Asp	Ile	Glu	Leu	Val	Ala	Thr	Leu	
		275					280					285				
gac	gac	gcc	cag	cgg	gac	ctg	ctg	ccg	agg	atc	ccg	gac	aac	acg	cgc	912
Asp	Asp	Ala	Gln	Arg	Asp	Leu	Leu	Pro	Arg	Ile	Pro	Asp	Asn	Thr	Arg	
	290					295					300					
atc	gtc	gac	ttc	gtg	ccc	atg	gac	gcg	ttg	ctg	ccg	acg	tgc	tcg	gcg	960
Ile	Val	Asp	Phe	Val	Pro	Met	Asp	Ala	Leu	Leu	Pro	Thr	Cys	Ser	Ala	
	305					310				315					320	
atc	atc	aac	cac	agc	ggt	tcg	ggc	acg	tgc	aac	acc	gcc	gcg	ctg	cac	1008
Ile	Ile	Asn	His	Ser	Gly	Ser	Gly	Thr	Cys	Asn	Thr	Ala	Ala	Leu	His	
			325					330						335		
ggg	gtg	ccg	cag	atc	atc	ctc	ggc	ggc	atc	ctg	gac	gcc	gcc	gta	cgg	1056
Gly	Val	Pro	Gln	Ile	Ile	Leu	Gly	Gly	Ile	Leu	Asp	Ala	Ala	Val	Arg	
			340					345					350			
cag	cac	atg	ttc	gcg	cag	aac	tcc	gcc	gcc	ctc	acc	ttc	gct	ccg	gag	1104
Gln	His	Met	Phe	Ala	Gln	Asn	Ser	Ala	Ala	Leu	Thr	Phe	Ala	Pro	Glu	
		355						360					365			
gag	gtg	acc	ggc	gcg	tcg	ctg	agg	agc	gcg	ctg	gtg	cgc	ctg	ctc	gag	1152
Glu	Val	Thr	Gly	Ala	Ser	Leu	Arg	Ser	Ala	Leu	Val	Arg	Leu	Leu	Glu	
	370					375					380					

gag ccg cgg ttc cgc gac ggc gcg cgg cgg ctg aag gag cgg atg cgg 1200
 Glu Pro Arg Phe Arg Asp Gly Ala Arg Arg Leu Lys Glu Arg Met Arg
 385 390 395 400

gcc atg ccc agc ccg gcc ggg atc gtg ccg acc ctg gag cgc ctc acg 1248
 Ala Met Pro Ser Pro Ala Gly Ile Val Pro Thr Leu Glu Arg Leu Thr
 405 410 415

gcc cag cac cgc cgg gcg tgt tga 1272
 Ala Gln His Arg Arg Ala Cys
 420

<210> 69
 <211> 423
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 69

Val Arg Val Leu Val Thr Ser Ile Pro His His Thr His Tyr Tyr His
 1 5 10 15

Leu Val Pro Leu Ile Trp Ala Leu Arg Ala Ser Gly His Glu Val Val
 20 25 30

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

Pro Ala Phe Ala Leu Ala Glu Glu Glu Ser Leu Ala Gln Ile Phe Glu
 50 55 60

Glu Val Glu Gly Asp Leu Gln Pro Tyr Gln His Gly Ile Asp Glu Phe
 65 70 75 80

Asp Phe Phe Gly Thr Leu Lys Asp Glu Leu Asp Trp Glu Lys Leu Leu
 85 90 95

Ala Gln Gln Val Ile Leu Ser Gly Leu Trp Leu Glu Pro Leu Asn Gly
 100 105 110

Ala Thr Thr Leu Asp Ser Ile Val Asp Phe Ala Arg Ala Trp Lys Pro
 115 120 125

Asp Leu Val Leu Trp Glu Pro Phe Thr Tyr Ala Gly Pro Val Ala Ala
 130 135 140

Arg Ala Cys Gly Ala Ala His Ala Arg Val Leu Trp Gly Pro Asp Thr
 145 150 155 160

Ile Gly Leu Leu Arg Thr Lys Phe Leu Gln Ala Gln Ala Arg Gln Pro
 165 170 175

Glu Glu His Arg Asp Asp Pro Val Ala Glu Trp Met Thr Trp Ala Leu
 180 185 190

Ala Arg Tyr Gly Cys Asp Phe Arg Glu Glu Asp Val Leu Gly Gln Trp
 195 200 205

Ser Val Asp Pro Met Ala Glu Gly Val Ser Leu Gly Leu Asp Leu Pro
 210 215 220

Thr Val Pro Met Arg Tyr Thr Pro Tyr Asn Gly Ser Ala Val Ile Pro
 225 230 235 240

Asp Trp Leu Thr Glu Glu Pro Lys Arg Pro Arg Val Cys Leu Thr Leu
 245 250 255

Gly Val Ser Ser Arg Glu His Ser Glu Asp Glu Val Pro Val Gln Arg
 260 265 270

Phe Ile Glu Ala Leu Ala Asp Leu Asp Ile Glu Leu Val Ala Thr Leu
 275 280 285

Asp Asp Ala Gln Arg Asp Leu Leu Pro Arg Ile Pro Asp Asn Thr Arg
 290 295 300

Ile Val Asp Phe Val Pro Met Asp Ala Leu Leu Pro Thr Cys Ser Ala
 305 310 315 320

Ile Ile Asn His Ser Gly Ser Gly Thr Cys Asn Thr Ala Ala Leu His
 325 330 335

Gly Val Pro Gln Ile Ile Leu Gly Gly Ile Leu Asp Ala Ala Val Arg
 340 345 350

Gln His Met Phe Ala Gln Asn Ser Ala Ala Leu Thr Phe Ala Pro Glu
 355 360 365

Glu Val Thr Gly Ala Ser Leu Arg Ser Ala Leu Val Arg Leu Leu Glu
 370 375 380

Glu Pro Arg Phe Arg Asp Gly Ala Arg Arg Leu Lys Glu Arg Met Arg
 385 390 395 400

Ala Met Pro Ser Pro Ala Gly Ile Val Pro Thr Leu Glu Arg Leu Thr
 405 410 415

Ala Gln His Arg Arg Ala Cys
 420

<210> 70
 <211> 1179
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1179)

<400> 70
 atg cgg gcc ctc ttc acg acc gcg ccg ctc gcg ggc cac ctg ctt ccg
 Met Arg Ala Leu Phe Thr Thr Ala Pro Leu Ala Gly His Leu Leu Pro

1	5	10	15	
ctg gtg ccc atc gcg tgg gcc ctg cgg gcg gcc ggc cac gag gta ctg				96
Leu Val Pro Ile Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Leu				
20	25	30		
gtg gcg acc cgg gag gac ttc gtg ccg gtc gcc ctg cgg tcg ggg ctg				144
Val Ala Thr Arg Glu Asp Phe Val Pro Val Ala Leu Arg Ser Gly Leu				
35	40	45		
ccg tcc gcc tcg tgc ggg ccg ccc gcc gcg gac ctg gcg ggc gcg gcc				192
Pro Ser Ala Ser Cys Gly Pro Pro Ala Ala Asp Leu Ala Gly Ala Ala				
50	55	60		
gag gcg ggg gcg ctc gcg cgg ccc cgc gga gcg gcg gag gct cgg ggg				240
Glu Ala Gly Ala Leu Ala Arg Pro Arg Gly Ala Ala Glu Ala Arg Gly				
65	70	75	80	
gtc ctg agc ggg gcg ctg gcg cgc gtc gcc cgg ggc agt ctg gcg ggg				288
Val Leu Ser Gly Ala Leu Ala Arg Val Ala Arg Gly Ser Leu Ala Gly				
85	90	95		
gtg cgg cgg ctg gcg gac gcc tgg cgg ccg gat ctg atc gtc agc gaa				336
Val Arg Arg Leu Ala Asp Ala Trp Arg Pro Asp Leu Ile Val Ser Glu				
100	105	110		
cgg gcc gag ttc gcc ggg ccg ctg gtc gcg gcg gcc ctc ggg gtc ccg				384
Arg Ala Glu Phe Ala Gly Pro Leu Val Ala Ala Ala Leu Gly Val Pro				
115	120	125		
tgg gtc cgc tac cac tgg tcg gtc tcg tcc ctg gag gag tac cgg cga				432
Trp Val Arg Tyr His Trp Ser Val Ser Ser Leu Glu Glu Tyr Arg Arg				
130	135	140		
gcg gcc gag gcc gag ttc gcg ccc gag ctg gcg gcg ctc ggc ctc gac				480
Ala Ala Glu Ala Glu Phe Ala Pro Glu Leu Ala Ala Leu Gly Leu Asp				
145	150	155	160	
cgg ttc ccg gag gcg gcg cgc gtg ctc gat ccg tgg ccg gtg tcg ctg				528
Arg Phe Pro Glu Ala Ala Arg Val Leu Asp Pro Trp Pro Val Ser Leu				
165	170	175		
cgc cgg ccg gac gcg gtc gcc cac gac ggg gtc cgg cac gta ccg gcc				576
Arg Arg Pro Asp Ala Val Ala His Asp Gly Val Arg His Val Pro Ala				
180	185	190		
cac ggg gac gcc ccc gtc ccc gac tgg gcg ttc acg cgc ggt cgc ggg				624

His Gly Asp Ala Pro Val Pro Asp Trp Ala Phe Thr Arg Gly Arg Gly	
195 200 205	
ccg cgg atc tgc gtg acg ctc ggc acc atg ctg ccc cgg tac ggc gcc	672
Pro Arg Ile Cys Val Thr Leu Gly Thr Met Leu Pro Arg Tyr Gly Ala	
210 215 220	
gcc ggg atg gcc gac ttc ctg acg gag ctg gtg gcg gag acc cgc gga	720
Ala Gly Met Ala Asp Phe Leu Thr Glu Leu Val Ala Glu Thr Arg Gly	
225 230 235 240	
ggg gac tgc gaa ctg ctc gtg gcg gtc gac gac gac gtc gtc gcg cgg	768
Gly Asp Cys Glu Leu Leu Val Ala Val Asp Asp Asp Val Val Ala Arg	
245 250 255	
tgg ccg tcg ctg ccc tcc gcg gtg cgg tac gcc ggc cgg ctg ccg ctg	816
Trp Pro Ser Leu Pro Ser Ala Val Arg Tyr Ala Gly Arg Leu Pro Leu	
260 265 270	
gcg gag gtg ctg ccc gcg tgc gac gcg gtg gtg cac cac ggc ggg cag	864
Ala Glu Val Leu Pro Ala Cys Asp Ala Val Val His His Gly Gly Gln	
275 280 285	
ggc acg tcc ctg acc gcg ctg gcc gcg ggt cgg ccg cag gtc gtc atg	912
Gly Thr Ser Leu Thr Ala Leu Ala Ala Gly Arg Pro Gln Val Val Met	
290 295 300	
gcg cgg ctc gac gac cag ttc gac aac gcg cgg gca ctg gcg gcg gcg	960
Ala Arg Leu Asp Asp Gln Phe Asp Asn Ala Arg Ala Leu Ala Ala Ala	
305 310 315 320	
ggg gcg gcc ctg ctc gta ccg ccg tcc cgg gcc act ccc gcg gcc gtg	1008
Gly Ala Ala Leu Leu Val Pro Pro Ser Arg Ala Thr Pro Ala Ala Val	
325 330 335	
gcc gcg ggg tgc gcc gaa gtg ctg gag aac gcc ctg tat gcc aag gcg	1056
Ala Ala Gly Cys Ala Glu Val Leu Glu Asn Ala Leu Tyr Ala Lys Ala	
340 345 350	
gca gcc ggg ctc gcc gag gag atg gcg ctg ctg ccg tcg ccg tcg gcg	1104
Ala Ala Gly Leu Ala Glu Glu Met Ala Leu Leu Pro Ser Pro Ser Ala	
355 360 365	
gcg gtc gga ctc ctg gaa cac ccg ggg ccc ggg ccg gac atg ccg cgg	1152
Ala Val Gly Leu Leu Glu His Pro Gly Pro Gly Pro Asp Met Pro Arg	
370 375 380	

agt tac ccg aac gag gat gcg gtg tga
 Ser Tyr Pro Asn Glu Asp Ala Val
 385 390

1179

<210> 71
 <211> 392
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 71

Met Arg Ala Leu Phe Thr Thr Ala Pro Leu Ala Gly His Leu Leu Pro
 1 5 10 15

Leu Val Pro Ile Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Leu
 20 25 30

Val Ala Thr Arg Glu Asp Phe Val Pro Val Ala Leu Arg Ser Gly Leu
 35 40 45

Pro Ser Ala Ser Cys Gly Pro Pro Ala Ala Asp Leu Ala Gly Ala Ala
 50 55 60

Glu Ala Gly Ala Leu Ala Arg Pro Arg Gly Ala Ala Glu Ala Arg Gly
 65 70 75 80

Val Leu Ser Gly Ala Leu Ala Arg Val Ala Arg Gly Ser Leu Ala Gly
 85 90 95

Val Arg Arg Leu Ala Asp Ala Trp Arg Pro Asp Leu Ile Val Ser Glu
 100 105 110

Arg Ala Glu Phe Ala Gly Pro Leu Val Ala Ala Ala Leu Gly Val Pro
 115 120 125

Trp Val Arg Tyr His Trp Ser Val Ser Ser Leu Glu Glu Tyr Arg Arg
 130 135 140

Ala Ala Glu Ala Glu Phe Ala Pro Glu Leu Ala Ala Leu Gly Leu Asp
 145 150 155 160

Arg Phe Pro Glu Ala Ala Arg Val Leu Asp Pro Trp Pro Val Ser Leu
 165 170 175

Arg Arg Pro Asp Ala Val Ala His Asp Gly Val Arg His Val Pro Ala
 180 185 190

His Gly Asp Ala Pro Val Pro Asp Trp Ala Phe Thr Arg Gly Arg Gly
 195 200 205

Pro Arg Ile Cys Val Thr Leu Gly Thr Met Leu Pro Arg Tyr Gly Ala
 210 215 220

Ala Gly Met Ala Asp Phe Leu Thr Glu Leu Val Ala Glu Thr Arg Gly
 225 230 235 240

Gly Asp Cys Glu Leu Leu Val Ala Val Asp Asp Asp Val Val Ala Arg
 245 250 255

Trp Pro Ser Leu Pro Ser Ala Val Arg Tyr Ala Gly Arg Leu Pro Leu
 260 265 270

Ala Glu Val Leu Pro Ala Cys Asp Ala Val Val His His Gly Gly Gln
 275 280 285

Gly Thr Ser Leu Thr Ala Leu Ala Ala Gly Arg Pro Gln Val Val Met
 290 295 300

Ala Arg Leu Asp Asp Gln Phe Asp Asn Ala Arg Ala Leu Ala Ala Ala
 305 310 315 320

Gly Ala Ala Leu Leu Val Pro Pro Ser Arg Ala Thr Pro Ala Ala Val
 325 330 335

Ala Ala Gly Cys Ala Glu Val Leu Glu Asn Ala Leu Tyr Ala Lys Ala
 340 345 350

Ala Ala Gly Leu Ala Glu Glu Met Ala Leu Leu Pro Ser Pro Ser Ala
 355 360 365

Ala Val Gly Leu Leu Glu His Pro Gly Pro Gly Pro Asp Met Pro Arg
 370 375 380

Ser Tyr Pro Asn Glu Asp Ala Val
 385 390

<210> 72
 <211> 993
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(993)

<400> 72
 gtg aat ctg gaa gta ctc aac cgt tcg aac gat ccg cgc ggg ccg gtg 48
 Val Asn Leu Glu Val Leu Asn Arg Ser Asn Asp Pro Arg Gly Pro Val
 1 5 10 15

atc acg gtg gtc ggc gcg tcc ggc ttc atc ggg tcc gcc ctg gtc gcc 96
 Ile Thr Val Val Gly Ala Ser Gly Phe Ile Gly Ser Ala Leu Val Ala
 20 25 30

gag ctg gcg cgc atg ccg gtg cgg ctg cgg gcg gtg gcc cgg cgc gag 144
 Glu Leu Ala Arg Met Pro Val Arg Leu Arg Ala Val Ala Arg Arg Glu
 35 40 45

acc ccc gtt ccc gcg ggg gca cgg gcc gcc gtc gag gtc cgc cgg gcg 192
 Thr Pro Val Pro Ala Gly Ala Arg Ala Ala Val Glu Val Arg Arg Ala
 50 55 60

gac ctc gcc cgg ccg gac gag gtc ggg gcc gcc gtc gag ggg gcg gac 240
 Asp Leu Ala Arg Pro Asp Glu Val Gly Ala Ala Val Glu Gly Ala Asp

65	70	75	80	
gcc gtc gtg cac ctc gcc gcc cac atc ggc ggc gcg cgg tcg tgg cgc				288
Ala Val Val His Leu Ala Ala His Ile Gly Gly Ala Arg Ser Trp Arg				
	85	90	95	
gcg gcc gac gag cgg tcg ctg cgg gtg aac gtc ggt ctg ctg cgc gac				336
Ala Ala Asp Glu Arg Ser Leu Arg Val Asn Val Gly Leu Leu Arg Asp				
	100	105	110	
gtg gcc gac gcg ttc cgg gac cgc tcg ggg ccc gcc ccg gcc gtg gtc				384
Val Ala Asp Ala Phe Arg Asp Arg Ser Gly Pro Ala Pro Ala Val Val				
	115	120	125	
ctg gcc agt acc ctc cag gcc ggc gtc gag ctg tcc cgg cag ggc ccg				432
Leu Ala Ser Thr Leu Gln Ala Gly Val Glu Leu Ser Arg Gln Gly Pro				
	130	135	140	
tac gcc cgg cag aag tcg gcg gcc gag gag gtc ctg ctg cgg gcc gcc				480
Tyr Ala Arg Gln Lys Ser Ala Ala Glu Glu Val Leu Leu Arg Ala Ala				
145	150	155	160	
tcc gag gag gtg gtc cgc ggc gtc gtg ctg cgg ctg ccg acc gtc tac				528
Ser Glu Glu Val Val Arg Gly Val Val Leu Arg Leu Pro Thr Val Tyr				
	165	170	175	
ggg cgc agc ccg ctg acc ggg tgg acg ggc cgc ggg gtg gtc gcg tcg				576
Gly Arg Ser Pro Leu Thr Gly Trp Thr Gly Arg Gly Val Val Ala Ser				
	180	185	190	
gtg gca cgg cag gcc gtc tcg ggc gag ccg gtc acg atg tgg cac gac				624
Val Ala Arg Gln Ala Val Ser Gly Glu Pro Val Thr Met Trp His Asp				
	195	200	205	
ggc acg gtc ggg cgc gat ctg ctc cac gtg gag gac gcg gcc cgc gcc				672
Gly Thr Val Gly Arg Asp Leu Leu His Val Glu Asp Ala Ala Arg Ala				
	210	215	220	
ttc gcg gcg gcg ctc ggt cac gtg gag cgg ctg gac ggc ggc acg tgg				720
Phe Ala Ala Ala Leu Gly His Val Glu Arg Leu Asp Gly Gly Thr Trp				
225	230	235	240	
tcc gtc ggt acg ggc cgg ctg gag ccc ttg gga gag gtg ttc tcg gcc				768
Ser Val Gly Thr Gly Arg Leu Glu Pro Leu Gly Glu Val Phe Ser Ala				
	245	250	255	
ctc gcc ggg ctg gtg gcc gag cgg acg ggg agg ccc ccc gta ccg gtg				816

Leu Ala Gly Leu Val Ala Glu Arg Thr Gly Arg Pro Pro Val Pro Val
 260 265 270

gtc tcc acg gag ccg ccc gac cat gcc gag gcg ggc gac ttc gag agc 864
 Val Ser Thr Glu Pro Pro Asp His Ala Glu Ala Gly Asp Phe Glu Ser
 275 280 285

gcg gtc tgt gac ccc tcc gcg ttc cgc gcg gtg acc ggg tgg tct ccc 912
 Ala Val Cys Asp Pro Ser Ala Phe Arg Ala Val Thr Gly Trp Ser Pro
 290 295 300

ctc gtt ccg ttg cgg gcg ggg ctc ggc gcc gtg gtg gag acg atg gtg 960
 Leu Val Pro Leu Arg Ala Gly Leu Gly Ala Val Val Glu Thr Met Val
 305 310 315 320

gcc gac gga gcg agg ggt ggg atc cga acg tga 993
 Ala Asp Gly Ala Arg Gly Gly Ile Arg Thr
 325 330

<210> 73
 <211> 330
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 73

Val Asn Leu Glu Val Leu Asn Arg Ser Asn Asp Pro Arg Gly Pro Val
 1 5 10 15

Ile Thr Val Val Gly Ala Ser Gly Phe Ile Gly Ser Ala Leu Val Ala
 20 25 30

Glu Leu Ala Arg Met Pro Val Arg Leu Arg Ala Val Ala Arg Arg Glu
 35 40 45

Thr Pro Val Pro Ala Gly Ala Arg Ala Ala Val Glu Val Arg Arg Ala
 50 55 60

Asp Leu Ala Arg Pro Asp Glu Val Gly Ala Ala Val Glu Gly Ala Asp
 65 70 75 80

187

Ala Val Val His Leu Ala Ala His Ile Gly Gly Ala Arg Ser Trp Arg
85 90 95

Ala Ala Asp Glu Arg Ser Leu Arg Val Asn Val Gly Leu Leu Arg Asp
100 105 110

Val	Ala	Asp	Ala	Phe	Arg	Asp	Arg	Ser	Gly	Pro	Ala	Pro	Ala	Val	Val
		115					120					125			

Leu Ala Ser Thr Leu Gln Ala Gly Val Glu Leu Ser Arg Gln Gly Pro
130 135 140

Tyr Ala Arg Gln Lys Ser Ala Ala Glu Glu Val Leu Leu Arg Ala Ala
145 150 155 160

Ser Glu Glu Val Val Arg Gly Val Val Leu Arg Leu Pro Thr Val Tyr
165 170 175

Gly Arg Ser Pro Leu Thr Gly Trp Thr Gly Arg Gly Val Val Ala Ser
180 185 190

Val	Ala	Arg	Gln	Ala	Val	Ser	Gly	Glu	Pro	Val	Thr	Met	Trp	His	Asp
		195					200					205			

Gly Thr Val Gly Arg Asp Leu Leu His Val Glu Asp Ala Ala Arg Ala
210 215 220

Phe Ala Ala Ala Leu Gly His Val Glu Arg Leu Asp Gly Gly Thr Trp
225 230 235 240

Ser Val Gly Thr Gly Arg Leu Glu Pro Leu Gly Glu Val Phe Ser Ala
245 250 255

Leu Ala Gly Leu Val Ala Glu Arg Thr Gly Arg Pro Pro Val Pro Val
260 265 270

Val Ser Thr Glu Pro Pro Asp His Ala Glu Ala Gly Asp Phe Glu Ser
 275 280 285

Ala Val Cys Asp Pro Ser Ala Phe Arg Ala Val Thr Gly Trp Ser Pro
 290 295 300

Leu Val Pro Leu Arg Ala Gly Leu Gly Ala Val Val Glu Thr Met Val
 305 310 315 320

Ala Asp Gly Ala Arg Gly Gly Ile Arg Thr
 325 330

<210> 74
 <211> 1008
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1008)

<400> 74
 gtg agc acg gac cgg gag cag gcc gcg cac acg cgg ctc ggt cgc agc 48
 Val Ser Thr Asp Arg Glu Gln Ala Ala His Thr Arg Leu Gly Arg Ser
 1 5 10 15
 gcg acc ctg gtg agc cgg ctc tgg ctg ggc acc gtg aac ttc agc ggc 96
 Ala Thr Leu Val Ser Arg Leu Trp Leu Gly Thr Val Asn Phe Ser Gly
 20 25 30
 cgg gtc gag gac ggt gac gcg atg cag ctg atg gag gcg gcg gtc gac 144
 Arg Val Glu Asp Gly Asp Ala Met Gln Leu Met Glu Ala Ala Val Asp
 35 40 45
 cgc ggc atc aac tgc atc gac acc gcg gac atc tac ggc tgg cgg atc 192
 Arg Gly Ile Asn Cys Ile Asp Thr Ala Asp Ile Tyr Gly Trp Arg Ile
 50 55 60
 cac aag ggc cac acc gag gaa ctg gtg ggc cgg tgg ctg gcc aag agc 240
 His Lys Gly His Thr Glu Glu Leu Val Gly Arg Trp Leu Ala Lys Ser
 65 70 75 80

gcc gcg cgg cgg gag gac gtc ctg ctg gcc acc aag gtc ggc ggg gac	288
Ala Ala Arg Arg Glu Asp Val Leu Leu Ala Thr Lys Val Gly Gly Asp	
85 90 95	
atg agc gaa cgg ctc aac gac ggc ggc ctg tcg gcg cgg cac atc gtc	336
Met Ser Glu Arg Leu Asn Asp Gly Gly Leu Ser Ala Arg His Ile Val	
100 105 110	
acg gcc tgc gag cag tcg ctg cgg cgc ctg ggc gtg gac cac atc gac	384
Thr Ala Cys Glu Gln Ser Leu Arg Arg Leu Gly Val Asp His Ile Asp	
115 120 125	
ctg tac cag atg cac cgc gtc gac cac gcc gcg ccg tgg gac gag atc	432
Leu Tyr Gln Met His Arg Val Asp His Ala Ala Pro Trp Asp Glu Ile	
130 135 140	
tgg cag gcg atg gac cgt ctg gtg gcg agc ggc aag gtg acc tac gtg	480
Trp Gln Ala Met Asp Arg Leu Val Ala Ser Gly Lys Val Thr Tyr Val	
145 150 155 160	
ggg tcg tcg aac ttc gcc ggc tgg aac gtc gcc gcc gcg cag gac gcg	528
Gly Ser Ser Asn Phe Ala Gly Trp Asn Val Ala Ala Ala Gln Asp Ala	
165 170 175	
gcc cgg cgg cgc cag tcc ctc ggt ctg gtg tcc gag cag tgc ctg tac	576
Ala Arg Arg Arg Gln Ser Leu Gly Leu Val Ser Glu Gln Cys Leu Tyr	
180 185 190	
aac ctg gcg gtg cgc cac gcc gag ctg gaa ctg ctg ccg gcc gcc cag	624
Asn Leu Ala Val Arg His Ala Glu Leu Glu Leu Leu Pro Ala Ala Gln	
195 200 205	
gcg tac gga ctg ggc gtg ttc gcc tgg tcg ccg ctg cac ggc ggg ctg	672
Ala Tyr Gly Leu Gly Val Phe Ala Trp Ser Pro Leu His Gly Gly Leu	
210 215 220	
ctc agc ggg gtg ctg cgc aag ctc gcg gcg ggc gtc gcg gtg aag tcg	720
Leu Ser Gly Val Leu Arg Lys Leu Ala Ala Gly Val Ala Val Lys Ser	
225 230 235 240	
gca cag ggg cgg gcc cag ctg ctg ctg ccc gag ctg cgc gcg acg atc	768
Ala Gln Gly Arg Ala Gln Leu Leu Leu Pro Glu Leu Arg Ala Thr Ile	
245 250 255	
gag gcg tac gag ggg ttc tgc ggc cgg atc ggc gcg gat ccg gcc gag	816
Glu Ala Tyr Glu Gly Phe Cys Gly Arg Ile Gly Ala Asp Pro Ala Glu	

260	265	270	
gtc ggt ctg gcc tgg gtg ctg tcc cgg ccg ggg atc agc ggc gcg gtg			864
Val Gly Leu Ala Trp Val Leu Ser Arg Pro Gly Ile Ser Gly Ala Val			
275	280	285	
atc ggt ccg cgc acg gtg gac cag ctg gac tcg gcg ctg cgg tcc ctg			912
Ile Gly Pro Arg Thr Val Asp Gln Leu Asp Ser Ala Leu Arg Ser Leu			
290	295	300	
gac ctg gtc ctc ggg gag gcc gaa ctg gcc gag ctg gac gcc atc ttc			960
Asp Leu Val Leu Gly Glu Ala Glu Leu Ala Glu Leu Asp Ala Ile Phe			
305	310	315	320
ccg ccc ctg ggc aag ggc ggc cgg gcg ccg gac gcg tgg atc agc tga			1008
Pro Pro Leu Gly Lys Gly Gly Arg Ala Pro Asp Ala Trp Ile Ser			
325	330	335	
<210>	75		
<211>	335		
<212>	PRT		
<213>	Streptomyces ambofaciens		
<400>	75		
Val Ser Thr Asp Arg Glu Gln Ala Ala His Thr Arg Leu Gly Arg Ser			
1	5	10	15
Ala Thr Leu Val Ser Arg Leu Trp Leu Gly Thr Val Asn Phe Ser Gly			
20	25	30	
Arg Val Glu Asp Gly Asp Ala Met Gln Leu Met Glu Ala Ala Val Asp			
35	40	45	
Arg Gly Ile Asn Cys Ile Asp Thr Ala Asp Ile Tyr Gly Trp Arg Ile			
50	55	60	
His Lys Gly His Thr Glu Glu Leu Val Gly Arg Trp Leu Ala Lys Ser			
65	70	75	80
Ala Ala Arg Arg Glu Asp Val Leu Leu Ala Thr Lys Val Gly Gly Asp			

85

90

95

Met Ser Glu Arg Leu Asn Asp Gly Gly Leu Ser Ala Arg His Ile Val
 100 105 110

Thr Ala Cys Glu Gln Ser Leu Arg Arg Leu Gly Val Asp His Ile Asp
 115 120 125

Leu Tyr Gln Met His Arg Val Asp His Ala Ala Pro Trp Asp Glu Ile
 130 135 140

Trp Gln Ala Met Asp Arg Leu Val Ala Ser Gly Lys Val Thr Tyr Val
 145 150 155 160

Gly Ser Ser Asn Phe Ala Gly Trp Asn Val Ala Ala Ala Gln Asp Ala
 165 170 175

Ala Arg Arg Arg Gln Ser Leu Gly Leu Val Ser Glu Gln Cys Leu Tyr
 180 185 190

Asn Leu Ala Val Arg His Ala Glu Leu Glu Leu Leu Pro Ala Ala Gln
 195 200 205

Ala Tyr Gly Leu Gly Val Phe Ala Trp Ser Pro Leu His Gly Gly Leu
 210 215 220

Leu Ser Gly Val Leu Arg Lys Leu Ala Ala Gly Val Ala Val Lys Ser
 225 230 235 240

Ala Gln Gly Arg Ala Gln Leu Leu Leu Pro Glu Leu Arg Ala Thr Ile
 245 250 255

Glu Ala Tyr Glu Gly Phe Cys Gly Arg Ile Gly Ala Asp Pro Ala Glu
 260 265 270

Val Gly Leu Ala Trp Val Leu Ser Arg Pro Gly Ile Ser Gly Ala Val
 275 280 285

Ile Gly Pro Arg Thr Val Asp Gln Leu Asp Ser Ala Leu Arg Ser Leu
 290 295 300

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

Pro Pro Leu Gly Lys Gly Gly Arg Ala Pro Asp Ala Trp Ile Ser
 325 330 335

<210> 76
 <211> 1233
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1233)

<400> 76
 atg att gcc acc gcg tgc cgc gtc tgc ggc aac aag gag ctg ctc tcc 48
 Met Ile Ala Thr Ala Cys Arg Val Cys Gly Asn Lys Glu Leu Leu Ser
 1 5 10 15
 gta ctc gac ctg ggc gaa cag gcg ctc acc ggg gtg ttc ccg gcc gac 96
 Val Leu Asp Leu Gly Glu Gln Ala Leu Thr Gly Val Phe Pro Ala Asp
 20 25 30
 cgc gac gag gtc gtg ccg tcg gtc ccg ctg gaa ctc gtc gcg tgt tcc 144
 Arg Asp Glu Val Val Pro Ser Val Pro Leu Glu Leu Val Ala Cys Ser
 35 40 45
 ccg gcc gga tgc ggc ctc gtg caa ctg cgc cac acc ccc gac ccg gac 192
 Pro Ala Gly Cys Gly Leu Val Gln Leu Arg His Thr Pro Asp Pro Asp
 50 55 60
 ctc atg tac ggc gag ggc tac ggc tac cgc tcg ggc atc cgg ccc ttc 240
 Leu Met Tyr Gly Glu Gly Tyr Gly Tyr Arg Ser Gly Ile Arg Pro Phe
 65 70 75 80

atg gtc gac cac ctg cat ggc aag gtc gcc gcc gtc cgc cgg ctg gtg	288
Met Val Asp His Leu His Gly Lys Val Ala Ala Val Arg Arg Leu Val	
85 90 95	
gac ctc ggc ccg gac gac ctg gtc ctc gac atc ggc agc aac gac gcc	336
Asp Leu Gly Pro Asp Asp Leu Val Leu Asp Ile Gly Ser Asn Asp Ala	
100 105 110	
acg ctc ctc aag ggc tac ccc gcc gac ggc ccg cga ctg gtc ggc atc	384
Thr Leu Leu Lys Gly Tyr Pro Ala Asp Gly Pro Arg Leu Val Gly Ile	
115 120 125	
gac ccg acc ggc ggc aag ttc cgc gac ctg tac ccg ccg aac gcc gag	432
Asp Pro Thr Gly Gly Lys Phe Arg Asp Leu Tyr Pro Pro Asn Ala Glu	
130 135 140	
ttg gtc gtg gac tac ttc acc cgc gcg acc ttc gag aac cgc ttc ggg	480
Leu Val Val Asp Tyr Phe Thr Arg Ala Thr Phe Glu Asn Arg Phe Gly	
145 150 155 160	
gcg cgg cgg gcg aag gcc gtc acc tcc atc gcg atg ttc tac gac ctg	528
Ala Arg Arg Ala Lys Ala Val Thr Ser Ile Ala Met Phe Tyr Asp Leu	
165 170 175	
ccg gac ccg ctg cgc ttc atg agc gac gtc cgc gac gtc ctc gcc gag	576
Pro Asp Pro Leu Arg Phe Met Ser Asp Val Arg Asp Val Leu Ala Glu	
180 185 190	
gac ggc gtc tgg atg atg gag cag agc tac ctg ccc gcc atg ctc gaa	624
Asp Gly Val Trp Met Met Glu Gln Ser Tyr Leu Pro Ala Met Leu Glu	
195 200 205	
gcc gac gcg tac gac atc gtc tgc cac gag cac ctg gag tac tac gcg	672
Ala Asp Ala Tyr Asp Ile Val Cys His Glu His Leu Glu Tyr Tyr Ala	
210 215 220	
ctg cgc cag atc gag tgg atg gcg gaa cgc gtc ggc ctc acc gtc atc	720
Leu Arg Gln Ile Glu Trp Met Ala Glu Arg Val Gly Leu Thr Val Ile	
225 230 235 240	
cgg gcc gag ctc acc gag gtg tac ggc ggc agc ctg tgc gtc acg ctc	768
Arg Ala Glu Leu Thr Glu Val Tyr Gly Gly Ser Leu Cys Val Thr Leu	
245 250 255	
gcc agg acc ggt tcc cgg cac ccg agg gac gag gcg ggc ctg gcc cgc	816
Ala Arg Thr Gly Ser Arg His Pro Arg Asp Glu Ala Gly Leu Ala Arg	
260 265 270	

atc cgg gcc gcc gag gcc gcc gcg gga ctg gac ggc atg gca ccg ttc 864
 Ile Arg Ala Ala Glu Ala Ala Ala Gly Leu Asp Gly Met Ala Pro Phe
 275 280 285

gag gcc ttc gcc cgc cgc gtc gcc gac cag cgg gac gcc ctg cgc ggc 912
 Glu Ala Phe Ala Arg Arg Val Ala Asp Gln Arg Asp Ala Leu Arg Gly
 290 295 300

ttc ctg gac cgc tcc cgc caa gag ggc ctg ctg acc ctc ggg tac ggc 960
 Phe Leu Asp Arg Ser Arg Gln Glu Gly Leu Leu Thr Leu Gly Tyr Gly
 305 310 315 320

gcc tcc acc aag ggc aac gtc atc ctc cag tac tgc ggg atc acc gaa 1008
 Ala Ser Thr Lys Gly Asn Val Ile Leu Gln Tyr Cys Gly Ile Thr Glu
 325 330 335

cgg gac ctg ccg tgc atc ggc gag gtg agc gag gag aag gcg ggg cgt 1056
 Arg Asp Leu Pro Cys Ile Gly Glu Val Ser Glu Glu Lys Ala Gly Arg
 340 345 350

ttc acc ccc ggg acg ggc atc ccc atc gtg tcc gag gag gac gcg aag 1104
 Phe Thr Pro Gly Thr Gly Ile Pro Ile Val Ser Glu Glu Asp Ala Lys
 355 360 365

gcg cag aag ccg gac cag ttg ctg gtc ctg ccc tgg atc tac ccg gac 1152
 Ala Gln Lys Pro Asp Gln Leu Leu Val Leu Pro Trp Ile Tyr Arg Asp
 370 375 380

ggg ttc gtc gaa ccg gag cgg gag ttc ctc gac ggc ggc ggc cga ctg 1200
 Gly Phe Val Glu Arg Glu Arg Glu Phe Leu Asp Gly Gly Gly Arg Leu
 385 390 395 400

gtc ttc ccg ctg ccc gcg ctc gac gtc gtg tga 1233
 Val Phe Pro Leu Pro Ala Leu Asp Val Val
 405 410

<210> 77

<211> 410

<212> PRT

<213> Streptomyces ambofaciens

<400> 77

Met Ile Ala Thr Ala Cys Arg Val Cys Gly Asn Lys Glu Leu Leu Ser
 1 5 10 15

Val Leu Asp Leu Gly Glu Gln Ala Leu Thr Gly Val Phe Pro Ala Asp
 20 25 30

Arg Asp Glu Val Val Pro Ser Val Pro Leu Glu Leu Val Ala Cys Ser
 35 40 45

Pro Ala Gly Cys Gly Leu Val Gln Leu Arg His Thr Pro Asp Pro Asp
 50 55 60

Leu Met Tyr Gly Glu Gly Tyr Gly Tyr Arg Ser Gly Ile Arg Pro Phe
 65 70 75 80

Met Val Asp His Leu His Gly Lys Val Ala Ala Val Arg Arg Leu Val
 85 90 95

Asp Leu Gly Pro Asp Asp Leu Val Leu Asp Ile Gly Ser Asn Asp Ala
 100 105 110

Thr Leu Leu Lys Gly Tyr Pro Ala Asp Gly Pro Arg Leu Val Gly Ile
 115 120 125

Asp Pro Thr Gly Gly Lys Phe Arg Asp Leu Tyr Pro Pro Asn Ala Glu
 130 135 140

Leu Val Val Asp Tyr Phe Thr Arg Ala Thr Phe Glu Asn Arg Phe Gly
 145 150 155 160

Ala Arg Arg Ala Lys Ala Val Thr Ser Ile Ala Met Phe Tyr Asp Leu
 165 170 175

Pro Asp Pro Leu Arg Phe Met Ser Asp Val Arg Asp Val Leu Ala Glu
 180 185 190

Asp Gly Val Trp Met Met Glu Gln Ser Tyr Leu Pro Ala Met Leu Glu

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

Gly Phe Val Glu Arg Glu Arg Glu Phe Leu Asp Gly Gly Gly Arg Leu
 385 390 395 400

Val Phe Pro Leu Pro Ala Leu Asp Val Val
 405 410

<210> 78
 <211> 1116
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1116)

<400> 78
 gtg cac cgg gac aac gcc gcc gag ccg ttg gtc aaa tgc ctg gtc tgg 48
 Val His Arg Asp Asn Ala Ala Glu Pro Leu Val Lys Cys Leu Val Trp
 1 5 10 15
 gac ctg gac aac acc ctc tgg cag ggc acc ctc ctc gag gag gac gag 96
 Asp Leu Asp Asn Thr Leu Trp Gln Gly Thr Leu Leu Glu Glu Asp Glu
 20 25 30
 gtc cgg ctc gcc ccg gac gtc ctg cgg acc atc acc gag ctg gac gcg 144
 Val Arg Leu Ala Pro Asp Val Leu Arg Thr Ile Thr Glu Leu Asp Ala
 35 40 45
 cgc ggc atc ctg cag gcg gtg gcc agc aag aac gac cac gat cac gcc 192
 Arg Gly Ile Leu Gln Ala Val Ala Ser Lys Asn Asp His Asp His Ala
 50 55 60
 tgg gcg aag ctc gaa cag ctc ggc gtc gcc gag tac ttc gtg ctc ccg 240
 Trp Ala Lys Leu Glu Gln Leu Gly Val Ala Glu Tyr Phe Val Leu Pro
 65 70 75 80
 agg atc ggc tgg ggc ccg aag tcg aag tcg gtc cgc gag atc gcc gac 288
 Arg Ile Gly Trp Gly Pro Lys Ser Lys Ser Val Arg Glu Ile Ala Asp
 85 90 95
 cgg ctg aac ttc gcg ccg agc acc ctc gcc ttc atc gac gac cag ccc 336
 Arg Leu Asn Phe Ala Pro Ser Thr Leu Ala Phe Ile Asp Asp Gln Pro
 100 105 110

ttc gaa cgg gcc gag gtc acc cac gaa ctg ccg gag gtc cgc acc tac	384
Phe Glu Arg Ala Glu Val Thr His Glu Leu Pro Glu Val Arg Thr Tyr	
115 120 125	
gcc gcc gag cag gcg acc cgg ctg acc ggc ctc ccg gag ttc agc ccg	432
Ala Ala Glu Gln Ala Thr Arg Leu Thr Gly Leu Pro Glu Phe Ser Pro	
130 135 140	
ggc acc gtc acc gtc gac tcg acc cgc cgc cgc tcc atg tac cag gcg	480
Gly Thr Val Thr Val Asp Ser Thr Arg Arg Arg Ser Met Tyr Gln Ala	
145 150 155 160	
tcc ttc cgc cgg gac gcc gag cgg tcc gac ttc acc gga ccc gac gcg	528
Ser Phe Arg Arg Asp Ala Glu Arg Ser Asp Phe Thr Gly Pro Asp Ala	
165 170 175	
gac ttc ctg cgc tcg ctg gac atc cgg atg cgg atc agc cgc gcc acg	576
Asp Phe Leu Arg Ser Leu Asp Ile Arg Met Arg Ile Ser Arg Ala Thr	
180 185 190	
ccc ctg gag ctg tcg cgg gtc gag gaa ctg acc ctg cgc acc agc cag	624
Pro Leu Glu Leu Ser Arg Val Glu Glu Leu Thr Leu Arg Thr Ser Gln	
195 200 205	
atg aac gcc acc gga gtg cac tac tcc gag gac gag ctg cgc gcc ctc	672
Met Asn Ala Thr Gly Val His Tyr Ser Glu Asp Glu Leu Arg Ala Leu	
210 215 220	
atc gac gac ccc gac cac gag gtg ctg gtc acc acc gtc acc gat cgc	720
Ile Asp Asp Pro Asp His Glu Val Leu Val Thr Thr Val Thr Asp Arg	
225 230 235 240	
ttc ggt ccc tac ggc gcg gtc ggc gtc gtg ctg ctg cgg cgc ggc ccc	768
Phe Gly Pro Tyr Gly Ala Val Gly Val Val Leu Leu Arg Arg Gly Pro	
245 250 255	
gag gcc tgg cgg atc aag ctg ctg gcc acc tcg tgc cgg gtg gtg tcg	816
Glu Ala Trp Arg Ile Lys Leu Leu Ala Thr Ser Cys Arg Val Val Ser	
260 265 270	
ctg ggg gcc ggc acg gtg atc ctg cgc tgg ctg acg gac cag gcg cac	864
Leu Gly Ala Gly Thr Val Ile Leu Arg Trp Leu Thr Asp Gln Ala His	
275 280 285	
cgg gcc ggg gtg cac ctc gga gcc gac ttc cgg gcc acc gaa cgc aac	912
Arg Ala Gly Val His Leu Gly Ala Asp Phe Arg Ala Thr Glu Arg Asn	
290 295 300	

cgg atg atg gaa gtg gcc tac cgc ttc gcc gga ttc acc gac gat ccc 960
 Arg Met Met Glu Val Ala Tyr Arg Phe Ala Gly Phe Thr Asp Asp Pro
 305 310 315 320

tgc ccg tgc cag gac gcc tcc gcc gcc acc ggc gcg atc ggc cgc ctg 1008
 Cys Pro Cys Gln Asp Ala Ser Ala Ala Thr Gly Ala Ile Gly Arg Leu
 325 330 335

cac ctg gtg ccg tcc ccc cgg ccc gcc ccc gac acc ctc cgt ctg gaa 1056
 His Leu Val Pro Ser Pro Arg Pro Ala Pro Asp Thr Leu Arg Leu Glu
 340 345 350

gct ccc gac ctg gcc acg ggc cgg cgg cgc ccc ggc cag gat ccc gaa 1104
 Ala Pro Asp Leu Ala Thr Gly Arg Arg Arg Pro Gly Gln Asp Pro Glu
 355 360 365

agg acc cca tga 1116
 Arg Thr Pro
 370

<210> 79
 <211> 371
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 79

Val His Arg Asp Asn Ala Ala Glu Pro Leu Val Lys Cys Leu Val Trp
 1 5 10 15

Asp Leu Asp Asn Thr Leu Trp Gln Gly Thr Leu Leu Glu Glu Asp Glu
 20 25 30

Val Arg Leu Ala Pro Asp Val Leu Arg Thr Ile Thr Glu Leu Asp Ala
 35 40 45

Arg Gly Ile Leu Gln Ala Val Ala Ser Lys Asn Asp His Asp His Ala
 50 55 60

Trp Ala Lys Leu Glu Gln Leu Gly Val Ala Glu Tyr Phe Val Leu Pro
 65 70 75 80

Arg Ile Gly Trp Gly Pro Lys Ser Lys Ser Val Arg Glu Ile Ala Asp
 85 90 95

Arg Leu Asn Phe Ala Pro Ser Thr Leu Ala Phe Ile Asp Asp Gln Pro
 100 105 110

Phe Glu Arg Ala Glu Val Thr His Glu Leu Pro Glu Val Arg Thr Tyr
 115 120 125

Ala Ala Glu Gln Ala Thr Arg Leu Thr Gly Leu Pro Glu Phe Ser Pro
 130 135 140

Gly Thr Val Thr Val Asp Ser Thr Arg Arg Arg Ser Met Tyr Gln Ala
 145 150 155 160

Ser Phe Arg Arg Asp Ala Glu Arg Ser Asp Phe Thr Gly Pro Asp Ala
 165 170 175

Asp Phe Leu Arg Ser Leu Asp Ile Arg Met Arg Ile Ser Arg Ala Thr
 180 185 190

Pro Leu Glu Leu Ser Arg Val Glu Glu Leu Thr Leu Arg Thr Ser Gln
 195 200 205

Met Asn Ala Thr Gly Val His Tyr Ser Glu Asp Glu Leu Arg Ala Leu
 210 215 220

Ile Asp Asp Pro Asp His Glu Val Leu Val Thr Thr Val Thr Asp Arg
 225 230 235 240

Phe Gly Pro Tyr Gly Ala Val Gly Val Val Leu Leu Arg Arg Gly Pro
 245 250 255

Glu Ala Trp Arg Ile Lys Leu Leu Ala Thr Ser Cys Arg Val Val Ser

260

265

270

Leu Gly Ala Gly Thr Val Ile Leu Arg Trp Leu Thr Asp Gln Ala His
 275 280 285

Arg Ala Gly Val His Leu Gly Ala Asp Phe Arg Ala Thr Glu Arg Asn
 290 295 300

Arg Met Met Glu Val Ala Tyr Arg Phe Ala Gly Phe Thr Asp Asp Pro
 305 310 315 320

Cys Pro Cys Gln Asp Ala Ser Ala Ala Thr Gly Ala Ile Gly Arg Leu
 325 330 335

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

Ala Pro Asp Leu Ala Thr Gly Arg Arg Arg Pro Gly Gln Asp Pro Glu
 355 360 365

Arg Thr Pro
 370

<210> 80
 <211> 1122
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1122)

<400> 80
 gtg gct gac acc tcc cac gcc gac ggc acc gag gcc gag gag ctg ttc 48
 Val Ala Asp Thr Ser His Ala Asp Gly Thr Glu Ala Glu Glu Leu Phe
 1 5 10 15

acc gcg ctg gtc ggc gac cgg gcg gcc gag tgg gac cgg acg ggc gag 96

Thr	Ala	Leu	Val	Gly	Asp	Arg	Ala	Ala	Glu	Trp	Asp	Arg	Thr	Gly	Glu	
			20					25						30		
ctg	ccc	ctc	ggc	ctg	ctg	cac	gac	ctc	ggc	tcc	cgg	ggc	ctg	ctg	tgc	144
Leu	Pro	Leu	Gly	Leu	Leu	His	Asp	Leu	Gly	Ser	Arg	Gly	Leu	Leu	Cys	
		35					40					45				
gca	cag	gcg	ccc	gcg	gtc	cac	ggc	ggc	ctc	ggc	tgg	acc	agc	cgg	cgc	192
Ala	Gln	Ala	Pro	Ala	Val	His	Gly	Gly	Leu	Gly	Trp	Thr	Ser	Arg	Arg	
	50					55					60					
aac	ggc	gaa	ctc	acc	gcg	cac	gtc	ggc	gcc	ctc	tgc	agc	tcg	ttg	cgg	240
Asn	Gly	Glu	Leu	Thr		Ala	His	Val	Gly	Ala	Leu	Cys	Ser	Ser	Leu	Arg
65						70					75				80	
agc	gtc	atg	acc	tcg	cag	ggc	atg	gcc	gcc	tgg	acg	ctg	cgc	cgg	ctc	288
Ser	Val	Met	Thr	Ser	Gln	Gly	Met	Ala	Ala	Trp	Thr	Leu	Arg	Arg	Leu	
				85					90					95		
gcc	gga	gcg	gac	cag	cag	gcc	tcc	ctg	gtg	ccc	cgg	ttg	acc	ggc	ggc	336
Ala	Gly	Ala	Asp	Gln	Gln	Ala	Ser	Leu	Val	Pro	Arg	Leu	Thr	Gly	Gly	
			100					105					110			
gaa	ctg	gcc	gcc	gtg	gcg	ttc	acc	gag	gcc	ggc	gcc	ggc	agc	gac	ctg	384
Glu	Leu	Ala	Ala	Val	Ala	Phe	Thr	Glu	Ala	Gly	Ala	Gly	Ser	Asp	Leu	
		115					120					125				
tcc	gcc	ctg	cgc	acc	cgc	atc	acc	tcc	gac	ggc	gac	gaa	gtc	gtc	gtc	432
Ser	Ala	Leu	Arg	Thr	Arg	Ile	Thr	Ser	Asp	Gly	Asp	Glu	Val	Val	Val	
		130				135					140					
gac	ggc	gtc	aag	gtg	tgg	gcg	acc	aac	gcc	gcg	tac	gcg	gac	ctg	ctg	480
Asp	Gly	Val	Lys	Val	Trp	Ala	Thr	Asn	Ala	Ala	Tyr	Ala	Asp	Leu	Leu	
145					150				155					160		
gtg	gtc	ttc	ggc	cgc	acg	gag	cag	ggc	gcg	ggc	gcc	gtc	gtg	gtg	ccc	528
Val	Val	Phe	Gly	Arg	Thr	Glu	Gln	Gly	Ala	Gly	Ala	Val	Val	Val	Pro	
			165					170					175			
gcc	tcg	gcc	ccc	ggc	gta	cgc	gtc	gag	cgc	atc	acc	gac	gcg	cac	ggc	576
Ala	Ser	Ala	Pro	Gly	Val	Arg	Val	Glu	Arg	Ile	Thr	Asp	Ala	His	Gly	
			180					185				190				
tgc	cgc	gcg	gcc	gga	cac	gcc	gac	atc	cac	ctg	gac	ggg	gtc	cgg	ctg	624
Cys	Arg	Ala	Ala	Gly	His	Ala	Asp	Ile	His	Leu	Asp	Gly	Val	Arg	Leu	
		195					200					205				

ccg gca gac gcc ctg ctc cag ggc cac gac cgc acc ccc gcc ctg ctg	672
Pro Ala Asp Ala Leu Leu Gln Gly His Asp Arg Thr Pro Ala Leu Leu	
210 215 220	
gtc acc acc gcg ctg agc tac ggc cgg atg tcg gtg gcc tgg ggc tcc	720
Val Thr Thr Ala Leu Ser Tyr Gly Arg Met Ser Val Ala Trp Gly Ser	
225 230 235 240	
ctg ggc atc ctg cgc gcc tgc ctg gcc gcc gcc gca cgt cac gcc ggc	768
Leu Gly Ile Leu Arg Ala Cys Leu Ala Ala Ala Ala Arg His Ala Gly	
245 250 255	
gga cgg gag cag ttc ggc acg cgg ctc tcg gac cac cag ctc gtc gcc	816
Gly Arg Glu Gln Phe Gly Thr Arg Leu Ser Asp His Gln Leu Val Ala	
260 265 270	
cgg cac ctc gcc gaa ctg ttc gtc gcc gaa cag cac gcc gcc cgg gcg	864
Arg His Leu Ala Glu Leu Phe Val Ala Glu Gln His Ala Ala Arg Ala	
275 280 285	
tgc gag cac gcc agc gcc cag tgg gac gag ggc agc ccc gac atg gtg	912
Cys Glu His Ala Ser Ala Gln Trp Asp Glu Gly Ser Pro Asp Met Val	
290 295 300	
gtc gcc gcc gtc ctg gcc aag cac gtc gcg gcc acc ggc gcc gca cgc	960
Val Ala Ala Val Leu Ala Lys His Val Ala Ala Thr Gly Ala Ala Arg	
305 310 315 320	
ggg gcc gaa cgg gcc gtg cag gtg ctg gcg tcg gcc ggg gcg cgg gag	1008
Gly Ala Glu Arg Ala Val Gln Val Leu Ala Ser Ala Gly Ala Arg Glu	
325 330 335	
gga cac gtc gtc gca cgg gcg cac cgc gac gcc aag ctc atg gag atc	1056
Gly His Val Val Ala Arg Ala His Arg Asp Ala Lys Leu Met Glu Ile	
340 345 350	
atc gag ggc agc aac gag atc tgc gaa ctg gtc ctg gcc cgg cac gtg	1104
Ile Glu Gly Ser Asn Glu Ile Cys Glu Leu Val Leu Ala Arg His Val	
355 360 365	
atg tcc gcg gcc ggg tga	1122
Met Ser Ala Ala Gly	
370	

<210> 81

<211> 373

<212> PRT

<213> Streptomyces ambofaciens

<400> 81

Val	Ala	Asp	Thr	Ser	His	Ala	Asp	Gly	Thr	Glu	Ala	Glu	Glu	Leu	Phe
1				5					10					15	

Thr	Ala	Leu	Val	Gly	Asp	Arg	Ala	Ala	Glu	Trp	Asp	Arg	Thr	Gly	Glu
			20					25					30		

Leu	Pro	Leu	Gly	Leu	Leu	His	Asp	Leu	Gly	Ser	Arg	Gly	Leu	Leu	Cys
		35					40					45			

Ala	Gln	Ala	Pro	Ala	Val	His	Gly	Gly	Leu	Gly	Trp	Thr	Ser	Arg	Arg
	50					55					60				

Asn	Gly	Glu	Leu	Thr	Ala	His	Val	Gly	Ala	Leu	Cys	Ser	Ser	Leu	Arg
65						70				75					80

Ser	Val	Met	Thr	Ser	Gln	Gly	Met	Ala	Ala	Trp	Thr	Leu	Arg	Arg	Leu
				85					90					95	

Ala	Gly	Ala	Asp	Gln	Gln	Ala	Ser	Leu	Val	Pro	Arg	Leu	Thr	Gly	Gly
			100					105					110		

Glu	Leu	Ala	Ala	Val	Ala	Phe	Thr	Glu	Ala	Gly	Ala	Gly	Ser	Asp	Leu
		115					120					125			

Ser	Ala	Leu	Arg	Thr	Arg	Ile	Thr	Ser	Asp	Gly	Asp	Glu	Val	Val	Val
	130					135					140				

Asp	Gly	Val	Lys	Val	Trp	Ala	Thr	Asn	Ala	Ala	Tyr	Ala	Asp	Leu	Leu
145					150					155					160

Val	Val	Phe	Gly	Arg	Thr	Glu	Gln	Gly	Ala	Gly	Ala	Val	Val	Val	Pro
				165					170					175	

Ala Ser Ala Pro Gly Val Arg Val Glu Arg Ile Thr Asp Ala His Gly
 180 185 190

Cys Arg Ala Ala Gly His Ala Asp Ile His Leu Asp Gly Val Arg Leu
 195 200 205

Pro Ala Asp Ala Leu Leu Gln Gly His Asp Arg Thr Pro Ala Leu Leu
 210 215 220

Val Thr Thr Ala Leu Ser Tyr Gly Arg Met Ser Val Ala Trp Gly Ser
 225 230 235 240

Leu Gly Ile Leu Arg Ala Cys Leu Ala Ala Ala Arg His Ala Gly
 245 250 255

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

Arg His Leu Ala Glu Leu Phe Val Ala Glu Gln His Ala Ala Arg Ala
 275 280 285

Cys Glu His Ala Ser Ala Gln Trp Asp Glu Gly Ser Pro Asp Met Val
 290 295 300

Val Ala Ala Val Leu Ala Lys His Val Ala Ala Thr Gly Ala Ala Arg
 305 310 315 320

Gly Ala Glu Arg Ala Val Gln Val Leu Ala Ser Ala Gly Ala Arg Glu
 325 330 335

Gly His Val Val Ala Arg Ala His Arg Asp Ala Lys Leu Met Glu Ile
 340 345 350

Ile Glu Gly Ser Asn Glu Ile Cys Glu Leu Val Leu Ala Arg His Val

355

360

365

Met Ser Ala Ala Gly
370

<210> 82

<211> 312

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1)..(312)

<400> 82

atg acc acg acg tcc ggt cgg ccc gac ccc acc ggc acg acc ccc acc 48
Met Thr Thr Thr Ser Gly Arg Pro Asp Pro Thr Gly Thr Thr Pro Thr
1 5 10 15

gcc gac gag gtc gcc ggg gaa ctg ctc ggg ttc ctc gag gac cgc acc 96
Ala Asp Glu Val Ala Gly Glu Leu Leu Gly Phe Leu Glu Asp Arg Thr
20 25 30

aag acc acc tgg gag cgc gac cag gac ctg ttc gcc gtc ggc ggg atg 144
Lys Thr Thr Trp Glu Arg Asp Gln Asp Leu Phe Ala Val Gly Gly Met
35 40 45

tcc tcg ctg ttc gcc atg cag ctc gtc gtg cac ctg gag aag acc tac 192
Ser Ser Leu Phe Ala Met Gln Leu Val Val His Leu Glu Lys Thr Tyr
50 55 60

ggc atc gtc atc agc ggc gcc gac ctg atg ctc gac aac ttc cgc acc 240
Gly Ile Val Ile Ser Gly Ala Asp Leu Met Leu Asp Asn Phe Arg Thr
65 70 75 80

gtc gac gcg atg gtc cgg ctg gtc gga cgg ctg gcc gcg ccc ggc gag 288
Val Asp Ala Met Val Arg Leu Val Gly Arg Leu Ala Ala Pro Gly Glu
85 90 95

gcg acc ggg cac gca ggt ggc tga 312
Ala Thr Gly His Ala Gly Gly
100

<210> 83
 <211> 103
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 83

Met Thr Thr Thr Ser Gly Arg Pro Asp Pro Thr Gly Thr Thr Pro Thr
 1 5 10 15

Ala Asp Glu Val Ala Gly Glu Leu Leu Gly Phe Leu Glu Asp Arg Thr
 20 25 30

Lys Thr Thr Trp Glu Arg Asp Gln Asp Leu Phe Ala Val Gly Gly Met
 35 40 45

Ser Ser Leu Phe Ala Met Gln Leu Val Val His Leu Glu Lys Thr Tyr
 50 55 60

Gly Ile Val Ile Ser Gly Ala Asp Leu Met Leu Asp Asn Phe Arg Thr
 65 70 75 80

Val Asp Ala Met Val Arg Leu Val Gly Arg Leu Ala Ala Pro Gly Glu
 85 90 95

Ala Thr Gly His Ala Gly Gly
 100

<210> 84
 <211> 867
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(867)

<400> 84

gtg ccc gat tcc aag gaa aat tct ccg ctc gtc gtg ctc ggc gcg ggc

Val	Pro	Asp	Ser	Lys	Glu	Asn	Ser	Pro	Leu	Val	Val	Leu	Gly	Ala	Gly	
1				5					10					15		
gtc	atg	ggc	acg	gcc	atc	gcg	gca	ctc	gcc	gtc	gga	cac	ggg	cac	ccg	96
Val	Met	Gly	Thr	Ala	Ile	Ala	Ala	Leu	Ala	Val	Gly	His	Gly	His	Pro	
			20					25					30			
gtc	aca	ctg	gtc	gac	acc	tcc	gcc	ggg	gcg	cgc	gcg	gcc	gcc	ccc	gac	144
Val	Thr	Leu	Val	Asp	Thr	Ser	Ala	Gly	Ala	Arg	Ala	Ala	Ala	Pro	Asp	
			35					40					45			
gcg	gtc	gcg	ctg	cat	ctg	cgg	acg	gcc	cgg	ctg	atg	ggc	gcg	ctg	ccc	192
Ala	Val	Ala	Leu	His	Leu	Arg	Thr	Ala	Arg	Leu	Met	Gly	Ala	Leu	Pro	
			50					55				60				
cac	gac	cgc	ccg	ccc	ggg	gag	ctg	acc	gtc	gag	gag	gcg	ccg	gcc	gcc	240
His	Asp	Arg	Pro	Pro	Gly	Glu	Leu	Thr	Val	Glu	Glu	Ala	Pro	Ala	Ala	
65					70					75				80		
gtc	gcc	acc	gcg	acc	gcc	gtg	atc	gag	gcc	gtc	acc	gag	gac	ccc	gag	288
Val	Ala	Thr	Ala	Thr	Ala	Val	Ile	Glu	Ala	Val	Thr	Glu	Asp	Pro	Glu	
				85					90					95		
cgg	aag	gcc	gag	gtg	ctg	gcg	gac	ctg	gcg	tcc	gtg	gcg	cgc	ccg	ggg	336
Arg	Lys	Ala	Glu	Val	Leu	Ala	Asp	Leu	Ala	Ser	Val	Ala	Arg	Pro	Gly	
			100					105						110		
acg	ctg	ctc	gtc	agc	aac	acc	tcg	ggc	gtc	ccc	atc	gac	gag	ctg	gcc	384
Thr	Leu	Leu	Val	Ser	Asn	Thr	Ser	Gly	Val	Pro	Ile	Asp	Glu	Leu	Ala	
			115					120					125			
gac	gcc	gtc	ccc	cgc	ccc	gag	gac	ctc	gtc	ggc	gtg	cac	ttc	atg	aac	432
Asp	Ala	Val	Pro	Arg	Pro	Glu	Asp	Leu	Val	Gly	Val	His	Phe	Met	Asn	
			130					135					140			
ccc	gcg	tac	gtg	atc	ccc	acg	gtc	gag	gtg	gtc	ctc	gga	ccg	cgc	agc	480
Pro	Ala	Tyr	Val	Ile	Pro	Thr	Val	Glu	Val	Val	Leu	Gly	Pro	Arg	Ser	
145					150					155					160	
gga	gag	gcg	gcc	gcc	cgg	gcc	acc	cgg	gac	ctg	ctg	tcc	ggc	ctg	ggc	528
Gly	Glu	Ala	Ala	Ala	Arg	Ala	Thr	Arg	Asp	Leu	Leu	Ser	Gly	Leu	Gly	
				165						170				175		
cgc	cgg	ggc	atc	gtc	gtc	ggc	gac	ggc	gcc	ggc	ttc	gtg	acc	agc	cgc	576
Arg	Arg	Gly	Ile	Val	Val	Gly	Asp	Gly	Ala	Gly	Phe	Val	Thr	Ser	Arg	
			180					185						190		

ctg ctg cac cgg atg ctg aac gac gcc atc gcg gtg gtg cac gag ggc 624
 Leu Leu His Arg Met Leu Asn Asp Ala Ile Ala Val Val His Glu Gly
 195 200 205

cgg gcc acc ccg gag acc gtg gac gcg ctg atg cgc gac tgc atc ggc 672
 Arg Ala Thr Pro Glu Thr Val Asp Ala Leu Met Arg Asp Cys Ile Gly
 210 215 220

cac cgc acc gga ccc ctg gcc acg gcc gac ctg atc gga ctg gac aac 720
 His Arg Thr Gly Pro Leu Ala Thr Ala Asp Leu Ile Gly Leu Asp Asn
 225 230 235 240

ctg gcc gac tcg ctg cgg gtg atg cac gaa cgg acc ggc gat ccg gcg 768
 Leu Ala Asp Ser Leu Arg Val Met His Glu Arg Thr Gly Asp Pro Ala
 245 250 255

ctc cgc ccg agc gag ctg ctg ctg gac aag gtc cgc cag ggc ctg ctc 816
 Leu Arg Pro Ser Glu Leu Leu Leu Asp Lys Val Arg Gln Gly Leu Leu
 260 265 270

ggc cgc aag agc ggc cgg gga ttc tac gac tac cag gag gcc acg cga 864
 Gly Arg Lys Ser Gly Arg Gly Phe Tyr Asp Tyr Gln Glu Ala Thr Arg
 275 280 285

tga 867

<210> 85
 <211> 288
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 85

Val Pro Asp Ser Lys Glu Asn Ser Pro Leu Val Val Leu Gly Ala Gly
 1 5 10 15

Val Met Gly Thr Ala Ile Ala Ala Leu Ala Val Gly His Gly His Pro
 20 25 30

Val Thr Leu Val Asp Thr Ser Ala Gly Ala Arg Ala Ala Ala Pro Asp
 35 40 45

Ala Val Ala Leu His Leu Arg Thr Ala Arg Leu Met Gly Ala Leu Pro

210

50

55

60

His Asp Arg Pro Pro Gly Glu Leu Thr Val Glu Glu Ala Pro Ala Ala
 65 70 75 80

Val Ala Thr Ala Thr Ala Val Ile Glu Ala Val Thr Glu Asp Pro Glu
 85 90 95

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

Thr Leu Leu Val Ser Asn Thr Ser Gly Val Pro Ile Asp Glu Leu Ala
 115 120 125

Asp Ala Val Pro Arg Pro Glu Asp Leu Val Gly Val His Phe Met Asn
 130 135 140

Pro Ala Tyr Val Ile Pro Thr Val Glu Val Val Leu Gly Pro Arg Ser
 145 150 155 160

Gly Glu Ala Ala Ala Arg Ala Thr Arg Asp Leu Leu Ser Gly Leu Gly
 165 170 175

Arg Arg Gly Ile Val Val Gly Asp Gly Ala Gly Phe Val Thr Ser Arg
 180 185 190

Leu Leu His Arg Met Leu Asn Asp Ala Ile Ala Val Val His Glu Gly
 195 200 205

Arg Ala Thr Pro Glu Thr Val Asp Ala Leu Met Arg Asp Cys Ile Gly
 210 215 220

His Arg Thr Gly Pro Leu Ala Thr Ala Asp Leu Ile Gly Leu Asp Asn
 225 230 235 240

Leu Ala Asp Ser Leu Arg Val Met His Glu Arg Thr Gly Asp Pro Ala
 245 250 255

Leu Arg Pro Ser Glu Leu Leu Leu Asp Lys Val Arg Gln Gly Leu Leu
 260 265 270

Gly Arg Lys Ser Gly Arg Gly Phe Tyr Asp Tyr Gln Glu Ala Thr Arg
 275 280 285

<210> 86

<211> 1040

<212> DNA

<213> *Streptomyces fradiae*

<400> 86

ggcaggagcc acccgtgaca gggctgccgc ggcccgcggt ccgggtgccg ttccacgatc 60

tgcgggacgt gcacgcggcc acgggggttg agtcggagat cggcggcgcg ctgctgcgcg 120

tcgccgcgcg cgggcgctat ctgctgggtg ccgaactcgc cgcgttcgag gagcggttcg 180

ccgagtactg cggcaacgcc cactgtgtcg ccgtgggcag cgggctcgac gacgctcgtc 240

tggcgctgtg ggcgctcggg gtgggcgagg gcgacgaggt gatcgtgcc tgcacacgt 300

tcacgcgctc ctggctcgcg gtgtcggcga cgggtgccac cccggtgccg gtcgagcccg 360

gtgatcccgg cgagccgggg cccggggcgt tctgtctcga cccggaccgg ctggaggccg 420

cgtgacccc gcggaccagg gccgtgatgc ccgtgcattt ctacgggcac ccggtggatc 480

tggaccgggt cggggcgctt gcggagccgc acgggctggc cgtcgtggag gacgcggcgc 540

aggccacggc ccgttacgcg gggaggcgga tcggcagcgg ccaccgcacg gcgttcagct 600

tctacccggg gaagaacctg ggcgcgctcg gcgacggcgg tgccgtggtc acctcggacc 660

cggaactcgc ggaccggctg cggctgttgc gcaactacgg cggccgggag aagtaccggc 720

acgaggagcg gggcaccaac tcccgcctgg acgagctgca ggcggccgtg ctgtcgggtca 780

agctgccgta cctggacgcc tggaacaccc gccgcgggga gatcgccgcc cgctacggcg 840

aggcgctggc cggctctgcc ggcgtcaccg tgccggaggg ccgcgtggcg gagccggtct 900

ggcatcagta cgtgctgcgc agcccgtaacc gcgaccgggtt gcggcggcgg ctggccgagg 960
 cgggggtgga gaccctggtc cactatccgg tggccgtgca cgcgtcgggc gcgtacgcgg 1020
 gcgcggggcc gtgtcccgcc 1040

<210> 87
 <211> 388
 <212> PRT
 <213> *Streptomyces fradiae*

<400> 87

Met Thr Gly Leu Pro Arg Pro Ala Val Arg Val Pro Phe His Asp Leu
 1 5 10 15

Arg Asp Val His Ala Ala Thr Gly Val Glu Ser Glu Ile Gly Gly Ala
 20 25 30

Leu Leu Arg Val Ala Ala Arg Gly Arg Tyr Leu Leu Gly Ala Glu Leu
 35 40 45

Ala Ala Phe Glu Glu Arg Phe Ala Glu Tyr Cys Gly Asn Ala His Cys
 50 55 60

Val Ala Val Gly Ser Gly Leu Asp Asp Ala Arg Leu Ala Leu Trp Ala
 65 70 75 80

Leu Gly Val Gly Glu Gly Asp Glu Val Ile Val Pro Ser His Thr Phe
 85 90 95

Ile Ala Ser Trp Leu Ala Val Ser Ala Thr Gly Ala Thr Pro Val Pro
 100 105 110

Val Glu Pro Gly Asp Pro Gly Glu Pro Gly Pro Gly Ala Phe Leu Leu
 115 120 125

213

Asp Pro Asp Arg Leu Glu Ala Ala Leu Thr Pro Arg Thr Arg Ala Val
 130 135 140

Met Pro Val His Leu Tyr Gly His Pro Val Asp Leu Asp Pro Val Gly
 145 150 155 160

Ala Phe Ala Glu Pro His Gly Leu Ala Val Val Glu Asp Ala Ala Gln
 165 170 175

Ala Thr Ala Arg Tyr Arg Gly Arg Arg Ile Gly Ser Gly His Arg Thr
 180 185 190

Ala Phe Ser Phe Tyr Pro Gly Lys Asn Leu Gly Ala Leu Gly Asp Gly
 195 200 205

Gly Ala Val Val Thr Ser Asp Pro Glu Leu Ala Asp Arg Leu Arg Leu
 210 215 220

Leu Arg Asn Tyr Gly Ala Arg Glu Lys Tyr Arg His Glu Glu Arg Gly
 225 230 235 240

Thr Asn Ser Arg Leu Asp Glu Leu Gln Ala Ala Val Leu Ser Val Lys
 245 250 255

Leu Pro Tyr Leu Asp Ala Trp Asn Thr Arg Arg Arg Glu Ile Ala Ala
 260 265 270

Arg Tyr Gly Glu Ala Leu Ala Gly Leu Pro Gly Val Thr Val Pro Glu
 275 280 285

Gly Arg Val Ala Glu Pro Val Trp His Gln Tyr Val Leu Arg Ser Pro
 290 295 300

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

Leu Val His Tyr Pro Val Ala Val His Ala Ser Gly Ala Tyr Ala Gly
325 330 335

Ala Gly Pro Cys Pro Ala Gly Gly Leu Pro Arg Ala Glu Arg Leu Ala
340 345 350

Gly Glu Val Leu Ser Leu Pro Ile Gly Pro His Leu Pro Asp Glu Ala
355 360 365

Val Glu Val Val Ile Ala Ala Val Gln Ser Ala Ala Leu Asp Ser Trp
370 375 380

Glu Glu Gly Pro
385

<210>	88
<211>	271
<212>	PRT
<213>	Streptomyces mycarofaciens

<400> 88

Met Ser Pro Ile Ser Ala Ser Ala Pro Ala Ala Ser Arg Ser Thr Ala
1 5 10 15

Arg Arg Glu Leu Gly Gln Asn Phe Phe Arg Ser Ala Ala Ala Ala Cys
20 25 30

Arg Phe Ser Asp Gln Leu Asp Ala Phe Cys Ala Asp Leu Pro Gly Ser
35 40 45

Leu Ala Asp Val Leu Thr Val Glu Ile Gly Ala Gly Ser Gly Arg Val
50 55 60

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

Asp Ala Tyr Trp Ala Arg Arg Leu Thr Ala Glu Ser Leu Pro Asp Val
85 90 95

Thr Val Val Asn Glu Asp Phe Leu Asn Leu Gln Leu Pro Arg Gln Pro
100 105 110

Ile Arg Leu Ile Gly Asn Leu Pro Phe Val Ser Gly Thr Lys Ile Leu
115 120 125

Arg Arg Cys Leu Glu Leu Gly Pro Asn Arg Met Cys Gln Ala Val Phe
130 135 140

Leu Leu Gln Arg Glu Tyr Val Gly Lys Arg Thr Gly Ala Trp Gly Gly
145 150 155 160

Asn Leu Phe Asn Ala Gln Trp Glu Pro Trp Tyr Thr Phe Glu Gly Gly
165 170 175

Leu Ala Phe Ser Arg Asn Glu Phe Ser Pro Val Pro Arg Ala Asp Thr
180 185 190

Gln Thr Leu Val Val Met Pro Arg Arg Arg Pro Ser Val Pro Trp Arg
195 200 205

Glu Arg Thr Asp Tyr Gln Arg Phe Thr Gln Gln Ile Phe Asp Thr Gly
210 215 220

Gln Met Thr Ile Gly Glu Ala Ala Arg Lys Val Leu Arg Arg Gly His
225 230 235 240

Ala Gln Phe Val Arg Ser Ala Gly Val Arg Pro Ala Asp Arg Val Lys
245 250 255

Asp Leu Thr Val Arg Asp Trp Ala Ala Leu Phe Arg Ala Asn Pro
260 265 270

<210> 89
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo SRMR1

<400> 89
 ctgccagtcc tctcccagca gtacg 25

<210> 90
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo SRMR2

<400> 90
 tgaagctgga cgtctcctac gtcgg 25

<210> 91
 <211> 2755
 <212> DNA
 <213> Artificial

<220>
 <223> Cassete excisavel

<400> 91
 gatattctacc tcttcgtccc gaagcaactt ggggtgggccc tcgaccttgg gccagtgtgc 60
 ttctcacctg cggcgccagt ccatgggtgc tcgcggttgt ccgcgctgt gcgctggcgt 120
 tgtcacgcag ttagacaactc accttcccac gggcggtgtc gtagggctcg gctaagttgc 180
 ccgcatgact gacagggatc cctcgagaag ctttatgctt gttaaaccgtt ttgtgaaaaa 240
 atttttaaaa taaaaaaggg gacctctagg gtccccaatt aattagtaat ataattctatt 300
 aaaggtcatt caaaagggtca tccaccggat caattcccct gctcgcgcag gctgggtgcc 360

aagctctcgg gtaacatcaa ggcccgatcc ttggagccct tgccctcccg cacgatgatc	420
gtgccgtgat cgaaatccag atccttgacc cgcagttgca aaccctcact gatccgtaat	480
gtgagttagc tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg	540
ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac	600
gccaaagctgc ggtggcgtag accgtcgcct cggtcggccc gtagagattg gcgatcccga	660
ccgcagcacc accgagaacg tccccgacgt ggccgaccag cccgtcatcg tcaacgcctg	720
accgcggtgc ggacaggccg tgtcgcgacc ggccgtgcgg aattaagccg gcccgtagcc	780
tgtgaataga ggtccgctgt gacacaagaa tccctgttac ttctcgaccg tattgattcg	840
gatgattcct acgcgagcct gcggaacgac caggaattct gggagccgct ggcccgcga	900
gccctggagg agctcgggct gccggtgccg ccggtgctgc gggtgcccgg cgagagcacc	960
aaccccgtag tggtcggcga gcccgacccg gtgatcaagc tgttcggcga gactgggtgc	1020
ggtccggaga gcctcgcgtc ggagtcggag gcgtacgcgg tcctggcgga cgcgccggtg	1080
ccggtgcccc gcctcctcgg ccgcggcgag ctgcggcccc gcaccggagc ctggccgtgg	1140
ccctacctgg tgatgagccg gatgaccggc accacctggc ggtccgcgat ggacggcacg	1200
accgaccgga acgcgctgct cgccttgccc cgcgaactcg gccgggtgct cggccggctg	1260
cacagggtag cgctgaccgg gaacaccgtg ctaccccc attccgaggt cttcccggaa	1320
ctgctgcggg aacgccgcgc ggcgaccgtc gaggaccacc gcgggtgggg ctacctctcg	1380
ccccggctgc tggaccgcct ggaggactgg ctgccggacg tggacacgct gctggccggc	1440
cgcgaacccc ggttcgtcca cggcgacctg cacgggacca acatcttcgt ggacctggcc	1500
gcgaccgagg tcaccgggat cgtcgacttc accgacgtct atgcgggaga ctcccgtac	1560
agcctggtag aactgcatct caacgccttc cggggcgacc gcgagatcct ggccgcgctg	1620
ctcgacgggg cgcagtggaa gcggaccgag gacttcgccc gcgaactgct cgccttcacc	1680
ttcctgcacg acttcgaggt gttcgaggag acccgcgtgg atctctccgg cttcaccgat	1740
ccggaggaac tggcgcagtt cctctggggg ccgcgggaca ccgccccgg cgctgacgc	1800

cccggggccgc ccggcgccgc ccccgcccc cggcgccgc cggagcccc gcccgcgctc 1860
 gggagccccg ggcccgcgcc gaagcccgt gctgcgagcc cggagcgggc cggccgacgg 1920
 cggcgcgggc ccggcgcgcg acgctcagca gcggcgggcg tgaaaggccc tggcatcctc 1980
 gatcatctcc tccaggggtg tcggccgagc ttccatccca gctcggcaag gatcgatccg 2040
 cgcagatcag ttggaagaat ttgtccacta cgtgaaaggc gagatcacca aggtagtcgg 2100
 caaataatgt ctaacaattc gttcaagccg acgcccgttc gcggcgcggc ttaactcaag 2160
 cgtagatgc actaagcaca taattgctca cagccaaact atcaggtcaa gtctgctttt 2220
 attattttta agcgtgcata ataagcccta cacaaattgg gagatatatc atgaaaggct 2280
 ggctttttct tgttatcgca atagttggcg aagtaatcgc aacatccgca ttaaaatcta 2340
 gcgagggctt tactaagctg atccggtgga tgaccttttg aatgaccttt aatagattat 2400
 attactaatt aattggggac cctagaggtc ccctttttta ttttaaaaat tttttcacia 2460
 aacggtttac aagcataaag cttgttaaca gatctcccgg ctgcgcggac actccggtac 2520
 gcccgtagc gaggaggtct accggaagca gatccggccc gtcattcaga ccggcgctgt 2580
 ggtcatggac ggcatcttca agcgggggtcc ggcgcgatag tcacgcagat agacacgcac 2640
 agaaaacagg tgaggcagac cgtaacgtta cggctctgct cacctggtgt ttctctgtcg 2700
 gggcgcgggg atttgaacct acgacctctt cgtcccgaa gaagcgcgcg atatc 2755

<210> 92

<211> 2208

<212> DNA

<213> Artificial

<220>

<223> Cassete excisavel

<400> 92

gatatcgcg gcgcttcgtt cgggacgaag aggtcgtggg ttcaaattccc gccaccccgga 60

cagagaaaca ccaggtgagg cagaccgtaa cgttacgggc tgcctcacct gttttctgtg 120

cgtgtctatc tgcgtgacta tcgcgccgga ccccgcttga agatgccgtc catgaccaca	180
gcgccgggtct ggatgacggg ccggatctgc ttccggtaga cctcctcagt cacggccgta	240
ccggagtgtc cgacgagccg ggagatctgt taacaagctt tatgcttgta aaccgttttg	300
tgaaaaaatt tttaaaataa aaaaggggac ctctagggtc cccaattaat tagtaatata	360
atctattaaa ggtcattcaa aaggatcatcc accggatcaa ttcccttgct cgcgcaggct	420
gggtgccaaag ctctcgggta acatcaaggc ccgataccttg gagcccttgc cctcccgcac	480
gatgatcgtg ccgtgatcga aatccagatc cttgaccgcg agttgcaaac cctcactgat	540
ccggctcacg gtaactgatg ccgtatctgc agtaccagcg tacggccac agaatgatgt	600
cacgtgaaa atgccggcct ttgaatgggt tcatgtgcag ctccatcagc aaaaggggat	660
gataagttta tcaccaccga ctatctgcaa cagtgcggtt gatcgtgcta tgatcgactg	720
atgtcatcag cgggtggagtg caatgtcgtg caatacgaat ggcgaaaagc cgagctcatc	780
ggtcagcttc tcaaccttg ggttacccttc ggcgggtgtgc tgctgggtcca cagctccttc	840
cgtagcgtcc ggccctcga agatgggcca cttggactga tcgaggccct gcgtgctgcg	900
ctgggtccgg gagggacgct cgtcatgccc tcgtggtcag gtctggacga cgagccgttc	960
gatacctgcca cgtcgccctg tacaccggac cttggagtgt tctctgacac attctggcgc	1020
ctgccaaatg taaagcgag cgcccatcca ttgaccttg cggcagcggg gccacaggca	1080
gagcagatca tctctgatcc attgcccctg ccacctcact cgctgcaag cccggtcgcc	1140
cgtgtccatg aactcgatgg gcaggtactt ctctcggcg tgggacacga tgccaacacg	1200
acgctgcac ttgccgagtt gatggcaaag gttccctatg gggtgccgag aactgcacc	1260
attcttcagg atggcaagtt ggtacgcgtc gattatctcg agaataacca ctgctgtgag	1320
cgctttgcct tggcggacag gtggctcaag gagaagagcc ttcagaagga aggtccagtc	1380
ggcatgcct ttgctcgggt gatccgctcc cgcgacattg tggcgacagc cctgggtcaa	1440
ctgggccgag atccgttgat cttcctgcat ccgccagagg gcgggatgcg aagaatgcga	1500
tgccgctcgc cagtcgattg gctgagctca tgagcggaga acgagatgac gttggagggg	1560

caaggtcgcg ctgattgctg gggcaacacg tgaaaggcga gatcaccaag gtagtcggca 1620
 aataatgtct aacaattcgt tcaagccgac gccgcttcgc ggcgcggctt aactcaagcg 1680
 ttagatgcac taagcacata attgctcaca gccaaactat caggtcaagt ctgcttttat 1740
 tatttttaag cgtgcataat aagccctaca caaattggga gatatatcat gaaaggctgg 1800
 ctttttcttg ttatcgcaat agttggcgaa gtaatcgcaa catccgcatt aaaatctagc 1860
 gagggcttta ctaagctgat ccggtggatg accttttgaa tgacctttaa tagattatat 1920
 tactaattaa ttggggaccc tagaggtccc cttttttatt ttaaaaattt tttcacaaaa 1980
 cggtttacaa gcataaagct tctcgaggga tccctgtcag tcatgcgggc aacttagccg 2040
 agccctacga caccgcccgt gggaagggtga gtgtctaact gcgtgacaac gccagcgcac 2100
 agcggcggac aaccgcgagc acccatggac tggcgccgca ggtgagaagc aactggccc 2160
 aaggtcgagc gccacccaa gttgcttcgg gacgaagagg cagatatc 2208

<210> 93
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo ORF2A

<400> 93
 cccgcgcggc agcctctccg tgatcgagtc cggcgtgacc atcgcgcgcg cttcgttcgg 60

<210> 94
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo ORF2B

<400> 94
 gtcctgtgcg tcatgcagga aggtgtcgta gtcgcggtag atctgcctct tcgtcccgaa 60

<210> 95
 <211> 35
 <212> DNA
 <213> Artificial

<220>
 <223> Cicatriz att3

<400> 95
 atcgcgcgcg cttcggtcgg gacgaagagg tagat 35

<210> 96
 <211> 59
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR8

<400> 96
 cgggatgatc gcttggtccg cggccggatg cctagcctca tcgcgcgcgc ttcggtcgg 59

<210> 97
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR9

<400> 97
 cccgatccag aacgtctggt cggatgatcag gtcgctgttc atctgcctct tcgtcccga 60

<210> 98
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR3

<400> 98
 accggggcgg tcctcccctc cggggcggtca cggccgcgga atctgcctct tcgtcccga 60

<210> 99
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR4

<400> 99
 cacgcagcga gccgacgcac tgatggacga cacgatggcc atcgcgcgcg cttcgttcgg 60

<210> 100
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR5

<400> 100
 gggcgtgaag cgggcgagtg tggatgtcat gcgagtactc atcgcgcgcg cttcgttcgg 60

<210> 101
 <211> 60
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR6

<400> 101
 cgggaaacgg cgtcgcactc ctcggggggcc gcgtcagccc atctgcctct tcgtcccga 60

<210> 102
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo C9583

<400> 102

ctgcaggtgc tccagcgcgt cgatct

26

<210> 103

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo C9584

<400> 103

ctgcagacgg aggcggacct gcggt

26

<210> 104

<211> 33

<212> DNA

<213> Artificial

<220>

<223> Cicatriz att1

<400> 104

atcgcgcgct tcgttcggga cgaagaggta gat

33

<210> 105

<211> 34

<212> DNA

<213> Artificial

<220>

<223> Cicatriz att2

<400> 105

atcggcgcgc ttcgttcggg acgaagaggt agat

34

<210> 106

<211> 10325

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> misc_feature

<222> (3620)..(4069)

<223> n e a, c, g, ou t

<400> 106

ctgcagaggg cgccgacgtg cgcggtgagt tcgccgttgc gccggctggt ccagccgagg	60
ccgccgtgga ccgcgggcgc ctgtgcgcac agcaggcccc gggagccgag gtcgtgcagc	120
aggccgaggg gcagctcgcc cgtccggtcc cactcggccg cccggtcgcc gaccagcgcg	180
gtgaacagct cctcggcctc ggtgccgtcg gcgtgggagg tgtcagccac ctgcgtgccc	240
ggtcgcctcg ccgggcgcgcg ccagccgtcc gaccagccgg accatcgcggt cgacgggtgcg	300
gaagtgtgcg agcatcaggt cggcgccgct gatgacgatg ccgtaggtct tctccaggtg	360
cacgacgagc tgcattggca acagcgagga catcccgcgc acggcgaaca ggtcctggtc	420
gcgctcccag gtggtcttgg tgcggtcctc gaggaaccgc agcagttccc cggcgacctc	480
gtcggcggtg ggggtcgtgc cgggtggggtc gggccgaccg gacgtcgtgg tcatcgcggtg	540
gcctcctggt agtcgtagaa tccccggccg ctcttgccgc cgagcaggcc ctggcgacc	600
ttgtccagca gcagctcgct cgggcggagc gccggatcgc cgggtccgttc gtgcatcacc	660
cgcagcgagt cggccagggt gtccagtcgc atcaggtcgg ccgtggccag gggtcgggtg	720
cgggtggccga tgcagtcgcg catcagcgcg tccacggtct cgggggtggc ccggccctcg	780
tgcaccaccg cgatggcgtc gttcagcatc cgggtgcagca ggcggctggt cacgaagccg	840
gcgccgtcgc cgacgacgat gcccggcgcg ccagggccgg acagcaggtc ccgggtggcc	900
cgggcgggccg cctctccgct gcgcggtccg aggaccacct cgaccgtggg gatcacgtac	960
gcgggggttca tgaagtgcac gccgacgagg tcctcggggc gggggacggc gtcggccagc	1020
tcgtcgtatg ggacggccga ggtgttgctg acgagcagcg tccccgggcg cgccacggac	1080
gccagggtccg ccagcacctc ggccttccgc tcgggggtcct cggtgacggc ctcgatcacg	1140
gcggtcgcgg tggcgacggc ggccggcgcc tcctcgacgg tcagctcccc gggcgggcg	1200
tcgtggggca gcgcgcccac cagccggggc gtccgcagat gcagcgcgac cgcgtcgggg	1260
gcggccgcgc gcgccccggc ggaggtgtcg accagtgtga ccgggtgccc gtgtccgacg	1320

gcgagtgccg	cgatggccgt	gcccattgacg	cccgcgccga	gcacgacgag	cggagaattt	1380
tccttggaat	cgggcacgga	tcctccagaa	tcccgggggc	ggcggctgtg	cgaaatgctg	1440
tcccgaatcg	cttccgcgct	ccatcacccg	ggcgtgactt	tgtcacctag	caagcggtag	1500
gaggccgggc	ggaggctatg	tcattctcggc	cttgtccgct	tacgagttcc	gtccttaaag	1560
tgcgcccggc	gaacgcacgg	cggaattgac	tcgccgtccg	tcaaggacct	gcggttcgaa	1620
ggacctcgac	gcatgaagg	tttcacatgg	ctcatatcgc	attcttcac	cttccggtg	1680
ccgggcatgt	gaatccgacc	ctgggagtcg	ccgaggaact	ggtcgcgcgc	ggccaccggg	1740
tgacgttcgc	gctgtccgag	gacctcgcgc	agcgggcccg	gctgatcggc	gccgaggtgg	1800
tcacctatcc	ggtggacagg	caacggttcc	tggaccagat	ggtgccgcgg	caggacgcgg	1860
acgagtacac	ggacgaggac	gagttcgtcc	gggtcctgga	gtggctgctg	gacatgacgg	1920
tgcagaccat	ggaaccgctg	gagaggcaact	tcgccgggga	ccggcccgcac	gtcgtcgtca	1980
acgatccgtc	gtcgtctgtg	acgggacggc	tgttggcgga	ccggtggggc	atcccgttca	2040
tccgcagcac	tccgacctat	gccgccaacg	aacactggtc	gctgcatccg	ccggtcgact	2100
cggccgagcc	gccggacgac	cccagagctgc	acaagctgct	cgcgcggatc	gagcggctgc	2160
tggaggagca	gggcgtcgag	cacgacctgg	ccggcttcac	cggggtcctg	cacggcggtc	2220
cggccctgct	gtacatgccc	cgctcgttcc	agtacgcggg	cgagaccttc	gacgcacagc	2280
accacttcgt	cggccccctgc	ccgccccgca	ccgcgttoca	cggcgagtag	aggccggggg	2340
acgacgacgg	ccggccccctg	gtgctggtga	gtctcggaac	cctgtacaac	gaccggccgg	2400
acttcttccg	cacctgcctg	gaggcgttcg	gcgacgagcc	ctggaacgtg	cttctggtgc	2460
tgggcggcgg	ggtgcccgcg	gccgacctgg	gcccgttcc	cggcaatgtc	cgggtgaccg	2520
acttcgtgtc	gctgcgcgac	gtcctgcgc	acacggcggt	ggtggtgaac	cacggtggga	2580
tgagcaccgc	catggaggtg	ttctcgcacg	gtgtgccggg	ggtggcgatc	ccggtgatgc	2640
cggagccccg	ggccaccgcg	cggcggatcg	tcgaactggg	cctgggcgac	cagctgctga	2700

actcggagct gacggccgag tccctgcgtg ccacggtaag gcgggtgctg gcggactccg	2760
cgatcccggg gaacatgcgc gggttccggg agcagatcag ggcggccggc ggggcgcccg	2820
cggcggccga cgcgatcgag ggactgctgc cccgggtggg ctgagacgtc cgcgcccgac	2880
acgcgttcac cttccgaacg gcgggatcgc cccatgctca tcaccgaaac agcagtgcc	2940
gacgtgttcc gcatcgatcc ggaaccgctg ccggaccacc ggggcccgtt ctacgaagcg	3000
gtgcgcccg ggcgctgga ggcgcgcgtc gggcactcgg tcgaggtccg gcaggtccac	3060
tgcaccgtct ccgggcgcaa cgtactgcgc ggctgcacg ccaccaccct gccgccgggc	3120
caggccaaga tcctcacctg tgtgcggggg gcggcgctga ctatggtggg cgacatgagg	3180
gtcgggtcac cgggcttcgg acggtacgag gcgggtgcgc aggatgcccc gtcgggcacc	3240
gcgctctacc tgccggacgg catcggcctg ggctacgtgg ccctcgcgga cgacacatgc	3300
atgaactacc tgtgcaccga ggagtacgtc ccgggcatgg tcatcgacgt ggacgccctg	3360
gaccccgaac tcggcctgcc gtggggactg accggggatc ccgtccgctc cgcccgggac	3420
gcggcggcgc cgtccctgcg ggcggccgcg gcggcgggaa ttctgcccac gtacgaggac	3480
tgcttacggg tgcgcacgtc cgtgcccgcc cctcccgacc ggactcggct ctgacgcgac	3540
ggacacgacc gccgcgcac cgacgcgaat ccgacgcgaa tccgaactcg atttcccga	3600
ccgtggacgg agcgcgtccn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	4020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn cgccacaccc	4080
cccagcagcg tcagcgtccg cctcgggacc acaccgggca ggacgacggg gtgcggggcg	4140

acgaggctga	gggagccccg	caactccaag	gggcggctca	accgcttgag	ctgctccatc	4200
aactgctggg	ccgaggtggt	ccccgtgtgg	tgggccaccg	ggccgcggcc	gctcgccgag	4260
gccagccgga	ggttccccgc	caccgcacc	atgaagggca	ggccgcgag	accgagccct	4320
cgcaccagtc	gtggcagcac	ggccgcgcgt	gcgtccatca	ccacggggcg	ggccaaggag	4380
cggttggcct	gagcggcttc	ggccaccagg	cggacgatgt	tctcctcctc	accgggtgcg	4440
ccggggtgcc	ggggtgtgcc	ctccccgctc	cggccgcgcg	ccagcgtcag	gcgccagctg	4500
accggggcg	ccgtcgtgcc	cgaggccagc	cacagcccgt	gactctgctg	gcagctcacc	4560
acccggccca	ggtcggagac	gaagcgccgg	ctcaccocga	cggagcgcac	ccccgccttg	4620
gagaccacca	tcggccagat	caccagggcg	tcgggggtca	gcctgtcgtc	cacgtagcgg	4680
gccagcgcg	cccgcacctc	gccccagtcc	caggtcgacc	cggcgacgaa	gtggtgcagg	4740
ctctgttccg	ccgcgccctc	tccgacgaag	gcggcgaggt	tccgggcggg	cttgcgtccc	4800
tgtgccgtga	gcagcccgcg	tatgtactgc	tcacctcttc	tgcgctggtc	cgcccggcgg	4860
agcgaaccga	gtaactccgc	acatgcgtcg	gagacgagcg	agtcgaagtc	gtgccgcgcg	4920
gcgggaccca	cggggggcgc	cgcgccggcg	ggtcgaagcg	tgcgtaaggt	cataggagtc	4980
ctcgtggggg	cctcgtcatc	actgcgagtg	gactgcgacc	agcatcgcca	attcaccctg	5040
tgctccccag	gtacaccggg	cacactcggt	ccgcttgcgt	ccccggcg	cacggcccgc	5100
gtgtggagca	ctccccgctc	tccggacggg	gggggggaag	cccaccgcac	gcgcgccggg	5160
agcggtgccg	gggccccaca	acccctcacc	gacgggggtcc	gttcgaccac	atgccgtgat	5220
gatgccttcg	ccgcacggcc	tccttgaatg	acggaccgtc	agaaagacgt	cactccgcgc	5280
gcgcccgtc	ccgccccgac	ccgcccggc	cgcggtcccg	ccccgcgac	cctgtggggc	5340
gacgccctct	tgacggaccc	ggaatcccgg	gccgcgggac	gggacggcga	ggaactcgat	5400
gccgcaaagg	ttccgcgtcc	tcaccggcac	ctccgagatc	cagcgcaacg	gcacgcgcaa	5460
gctcctgggt	ttccaccact	gagccgcacc	ccctggcgag	cagatccggg	acaagacgcc	5520

accttcgcga cagtgttttag gaaaagttaa gtaaagaatt ccgcgagcgg attgccaggg	5580
agaacaaccc attgacgcgc accggtgcag cggccacatt gacggcacct gtcaagttca	5640
cccccaggag cttggaatcc catgcaggca attcggcacc acgtcatgct cgtcatggcc	5700
ttcgtgaccg tcgccacgac tttcctttctc tggccgtcga cgcaatccgc gcaagcgttt	5760
cccccgaccc cgaagcagac ggtactgaac cacctccggc ccatttccgg gaatcacatc	5820
gtctccggac agcacaacaa ggagcccggc tccgccccgg gccagtacac ccagcaggtc	5880
aaggacgtca ccgggcagta ccccggcctg tggggcggtg acctgatgtt cgccgcggcg	5940
gacgtggccg gccgccagcg cgtcgtcgac caggccagga ccgagtgggc gaacggatcg	6000
ctgggtctcg tcacctggca cgtctgcccg ccgaccggcg gcagcacctg tgcgttcgag	6060
ggcggcgctca agtccacgct gacgaacgcg cagttctcgc aggtcctcac ggagggcagt	6120
gccctgaaca gcgcatggaa gcggcgccctg gacgaggtcg tcccgtacct gcagcagctg	6180
gagaacgcgg gcgtccccgt cctcttccgg ccgctgcacg agatgaacga atcctggaac	6240
tgggtggggaa accggccccg agcgaacggc agcgcaacgcc tctaccagat caccgcgat	6300
cacctcgccg ggacgaaag gctggacaat ctgatctggg tctggaacgt ccaggacaat	6360
ccggcgggaa actggaacag ctactatccg ggagatcagt acgtggacgt cgtttcgctg	6420
gacgtctggt acaagagcca cccgagttcc gccgactacc agcagatgcg gagcatcgcg	6480
ggaacaaaac ccatggccct cgcggaagtg ggcaaaatgc cgaccgccgc gctgctggac	6540
agccagacgc ggtggacatg gttcatgatg tgggtccgagc atctgcgcgg gaacaattcc	6600
aacgccgaaa tacagacggc gtattttccac ccccggtgac tgaaccaggg ggaggtcgca	6660
ctgccctgac gctcggcgct gcccggtctc ctcacgcgcg ttctgacagg acgtcgcgga	6720
gagtgcgggg caagcggccg gtgagctggg cgcagtcgtc gcggaccgga tcccagcgct	6780
gttcgcgtac cgaggcgaac agggaggaga aggcgtagag ccaccagggg tccaggccgg	6840
tcgccgctgt ttcggcagag tacgtgtcgg ccgggatgtc cacgtaggcg accggtgtcc	6900
gccacacctc cgccgccgtg gaggcgatgc ctgccaggtc cgtcgactcg ggtcccgtga	6960

tgctcgtggtg	acggcccggtc	ggcccgccca	ccgcgagggc	ggcgagggcg	cgggccacgt	7020
cgtcgcgcgc	caccagggac	acccggccat	cggtcgccgg	cagcctcagc	agcccggtcg	7080
agcgcgcctg	ggtgagccag	cctaggaaga	actcgatgta	cagcgaggcc	ctggcgaaacg	7140
agcacggcac	gccggaggcg	agcagcagat	cctcgggtgag	ccggttgacg	acggcgtagc	7200
agaacgggga	ggccgagtcg	gcgtcgacgc	tgctcagtgc	cgcgacatgg	ccgacgcgct	7260
cggccaccac	ggcggcgacg	acgttgccgg	ggtgcagcag	cacccgtagc	tcggggcccg	7320
cgctggagac	gaggaccaga	gtgtccacgc	ccttcagcgc	cgcacgcaga	gcgggcgggt	7380
cggcgtagtc	ggcgacggcg	cactccaccc	ccggcggtag	tgccctggca	ggcagcttcc	7440
gcctggatcat	cgccacgacg	tccacgtcgg	cccggtcgcg	gagcagctgc	accaccgcgc	7500
tgcccagact	gcccgcgcgc	cctgtcaccc	cgatacgcgt	cgtcgcccc	gtcgccttgt	7560
cgtcgggtcgt	accaccgtag	ggggccaacc	gcgaccaggg	cttggaacga	gccggcccg	7620
cagggcacag	acgcgcgacg	ggtccgggtt	tcccgtagtc	ttttggaccg	ggacgccgga	7680
ccgcttcctt	tctacggtgg	agccgttccc	gcccagagcc	gcacgtcatc	gacgtgcggg	7740
gaagacagag	gtgataccga	tgctccgacg	ccgtctcgga	gggccgtagc	gccccctcgt	7800
cagtgccctg	tgccctggcg	cgatgccctt	cggcaccacc	gtcgacgaga	agacgtcctt	7860
cgccatcctc	gaccggttcg	tcgaggcccg	cggcagtcct	gtcgacaccg	ccgacaacta	7920
cgcgttctgg	gctcccggcg	ggaccgggga	cgagagcgag	aacaccgtag	ggcgctggct	7980
ggcgagccgc	cgccgcgcgc	acgaggtggt	gatctccacc	aaggtgggtg	cccgcgccac	8040
cgccccggc	agcggcctgg	agaccgccga	agggtgtcgc	gtccccgtca	tacggaaggc	8100
cgcgaggagc	agcctgcgac	gcctgggcac	cgatcgcatc	gatctgtact	ggaccacat	8160
cgaggaccgg	accgtccctc	tggaggagac	gctcggagct	ctcgacgagc	tggtcggcgg	8220
cggaaggtg	gcggtgctgg	gctgctcaa	ccacgcgggc	tggcgcatcg	aacgggccc	8280
cgcgctcgcc	cggaccaacg	gctggacggc	gtacacctgc	gtccagcagc	gctactccta	8340

ccttcagccg	cgcttcgacg	tcggactgcc	ggagagcgga	cacgtccacg	cgacctccga	8400
actcctcgac	cacgtacgca	gcgagcccga	cctgacgctg	ctggcctact	cgccccctgct	8460
ctcgggcgcg	tacacccggc	cggacaagac	cctgtccgcc	gcctacgacc	acccgggcac	8520
cgggcagcgg	ctgaccgtgc	tgcgggaggt	ggccgcccag	ctcgggtgcca	ccgccaacca	8580
ggtggtgctg	tcctggttgc	tcggaggtga	tccgcgggtg	atcccgatcg	tgggggtgag	8640
ttccgtcgag	cagctggacg	aggtgctggc	cgcgctcgag	ctggagctgc	cccgggagac	8700
gagggcacgg	ctggacctcg	ccgggcgggg	ctgacgggcc	acacaccgcc	cctcgcccgg	8760
gacgggctga	ccggccgcac	accgtccctc	agccgggacg	ggctgacggc	cacacaccgc	8820
cccctcagcc	cggctgggct	gggctgacgg	ccacacaccc	gctcctcagc	ccggctgggc	8880
tgggctgacg	gccacacacc	cgccccctag	cccggctggg	ctgggctgac	ggccacacgc	8940
cgccccctgc	ccgggtcggg	ctgaccagcc	acacgcgctc	cctcagtcgg	cccgtcccc	9000
ggcccgcgcc	ggtctgacgt	gccgccggtc	gtggcgccgg	gccgcgtcga	tggcctgtgc	9060
caggcgccgc	acggctccgg	gcacggcggtg	gtccgccaga	ttgccgtagc	cgaggaccag	9120
cgcggggccg	gcggcggcca	cggcccgcgc	cccgcgggag	tcgccggccc	gggcccgtcg	9180
gtggtgctcg	gcccgcgcca	tccggtaggc	gtcgagatcc	gccagcctca	cgtcacgccg	9240
cgcgcctcc	ctgacgacgg	tgggcgcccga	gcagtccgtc	agggacagca	gcagggtgaa	9300
gccggccgcc	gcaccggaca	cctggaaccc	cggcagcgac	gccgcgagtt	ctcccatgag	9360
gtgatcgcg	cgccgtttgt	agcgcagccg	tgcgcgccgc	aggtaccggt	cgtaggaccc	9420
gcgggcccag	aaccgggcga	acgcctcctg	gtcgatgacg	ggcgggtgga	cgcccccgac	9480
gtgctcccg	cccagggcct	cggctccagcg	cggcggggtc	accgcccagc	cgatccgcag	9540
ggccggcgag	agcgtcttgc	tgaccgaccc	cagcagcgcg	acgtgctccg	gtgccatgcc	9600
ctgcagggcg	cccacggggt	ggcggctgta	tcggaactcg	gcgtcgtagt	cgtcctccag	9660
caccaggccg	tcgacctcgc	gggcccaggc	gaccagctcg	ccgcgacggg	ccggggcgag	9720
gaccacaccg	gtcgggaact	ggtgggcggg	agcgacgata	accgaccgga	cctgaggcgc	9780

gcgccgcagc aggtcgaccc gcaggccgtc gccgtccacg ggaaccggga cgggcaccag 9840
 cccctccgcc ccgaccgtgg cgctcagccg cgaccagcca gggtcctcca gggccatgtg 9900
 ggtgtggccc tcggtccgca ggaccggca cagcctgcgc accgcgtcca gcgtgcccgc 9960
 gcagacgacg agctgccggg cttccagcga cgcgccggt ccccgcacga ggtaggccgc 10020
 gagccgctgc cggagccggg cgagaccggc cggatcgggg tagccgagct cggcccaggt 10080
 cagtgcgccg gtggcctccc gcaccgcgtc cgcccacgt ccgcggggaa aggcccgcag 10140
 gtcgggtatg ccgggcagca tgtcgaactc cggctccac ccggtaccga cgtccgcctc 10200
 gggcgtcacc ggagcgggga cggcctgcgc cctgaccgg gtggccgagc cgctgcgcgc 10260
 ctccaggtac ccctccgca cgagctgctc gtagggggga tcctctagag tcgacctgca 10320
 ggcct 10325

<210> 107
 <211> 1218
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1218)

<400> 107
 atg gct cat atc gca ttc ttc atc ctt ccg gtt gcc ggg cat gtg aat 48
 Met Ala His Ile Ala Phe Phe Ile Leu Pro Val Ala Gly His Val Asn
 1 5 10 15
 ccg acc ctg gga gtc gcc gag gaa ctg gtc gcg cgc ggc cac cgg gtg 96
 Pro Thr Leu Gly Val Ala Glu Glu Leu Val Ala Arg Gly His Arg Val
 20 25 30
 acg ttc gcg ctg tcc gag gac ctc gcc gag cgg gcc cgg ctg atc ggc 144
 Thr Phe Ala Leu Ser Glu Asp Leu Ala Glu Arg Ala Arg Leu Ile Gly
 35 40 45
 gcc gag gtg gtc acc tat ccg gtg gac agg caa cgg ttc ctg gac cag 192
 Ala Glu Val Val Thr Tyr Pro Val Asp Arg Gln Arg Phe Leu Asp Gln

50	55	60	
atg gtg ccg cgg cag gac gcg gac gag tac acg gac gag gac gag ttc			240
Met Val Pro Arg Gln Asp Ala Asp Glu Tyr Thr Asp Glu Asp Glu Phe			
65	70	75	80
gtc cgg gtc ctg gag tgg ctg ctg gac atg acg gtg cag acc atg gaa			288
Val Arg Val Leu Glu Trp Leu Leu Asp Met Thr Val Gln Thr Met Glu			
	85	90	95
ccg ctg gag agg cac ttc gcc ggg gac cgg ccc gac gtc gtc gtc aac			336
Pro Leu Glu Arg His Phe Ala Gly Asp Arg Pro Asp Val Val Val Asn			
	100	105	110
gat ccg tcg tcg ctg tgg acg gga cgg ctg ctg gcg gac cgg tgg ggc			384
Asp Pro Ser Ser Leu Trp Thr Gly Arg Leu Leu Ala Asp Arg Trp Gly			
	115	120	125
atc ccg gtc atc cgc agc act ccg acc tat gcc gcc aac gaa cac tgg			432
Ile Pro Val Ile Arg Ser Thr Pro Thr Tyr Ala Ala Asn Glu His Trp			
	130	135	140
tcg ctg cat ccg ccg gtc gac tcg gcc gag ccg ccg gac gac ccc gag			480
Ser Leu His Pro Pro Val Asp Ser Ala Glu Pro Pro Asp Asp Pro Glu			
145	150	155	160
ctg cac aag ctg ctc gcg cgg atc gag cgg ctg ctg gag gag cag ggc			528
Leu His Lys Leu Leu Ala Arg Ile Glu Arg Leu Leu Glu Glu Gln Gly			
	165	170	175
gtc gag cac gac ctg gcc ggc ttc acc ggg gtc ctg cac ggc ggt ccg			576
Val Glu His Asp Leu Ala Gly Phe Thr Gly Val Leu His Gly Gly Pro			
	180	185	190
gcc ctg ctg tac atg ccc cgc tcg ttc cag tac gcg ggc gag acc ttc			624
Ala Leu Leu Tyr Met Pro Arg Ser Phe Gln Tyr Ala Gly Glu Thr Phe			
	195	200	205
gac gca cag cac cac ttc gtc ggc ccc tgc ccg ccc cgc acc gcg ttc			672
Asp Ala Gln His His Phe Val Gly Pro Cys Pro Pro Arg Thr Ala Phe			
	210	215	220
cac ggc gag tgg agg ccg ggg gac gac gac ggc cgg ccc ctg gtg ctg			720
His Gly Glu Trp Arg Pro Gly Asp Asp Asp Gly Arg Pro Leu Val Leu			
225	230	235	240
gtg agt ctc gga acc ctg tac aac gac cgg ccg gac ttc ttc cgc acc			768

Val	Ser	Leu	Gly	Thr	Leu	Tyr	Asn	Asp	Arg	Pro	Asp	Phe	Phe	Arg	Thr		
				245					250					255			
tgc	ctg	gag	gcg	ttc	ggc	gac	gag	ccc	tgg	aac	gtg	ctt	ctg	gtg	ctg	816	
Cys	Leu	Glu	Ala	Phe	Gly	Asp	Glu	Pro	Trp	Asn	Val	Leu	Leu	Val	Leu		
			260					265					270				
ggc	ggc	ggg	gtg	ccc	gcg	gcc	gac	ctg	ggc	ccg	ctt	ccc	ggc	aat	gtc	864	
Gly	Gly	Gly	Val	Pro	Ala	Ala	Asp	Leu	Gly	Pro	Leu	Pro	Gly	Asn	Val		
			275					280					285				
cgg	gtg	acc	gac	ttc	gtg	tcg	ctg	cgc	gac	gtc	ctg	ccg	cac	acg	gcg	912	
Arg	Val	Thr	Asp	Phe	Val	Ser	Leu	Arg	Asp	Val	Leu	Pro	His	Thr	Ala		
		290					295					300					
gtg	gtg	gtg	aac	cac	ggt	ggg	atg	agc	acc	gcc	atg	gag	gtg	ttc	tcg	960	
Val	Val	Val	Asn	His	Gly	Gly	Met	Ser	Thr	Ala	Met	Glu	Val	Phe	Ser		
305						310					315				320		
cac	ggt	gtg	ccg	gtg	gtg	gcg	atc	ccg	gtg	atg	ccg	gag	ccc	cgg	gcc	1008	
His	Gly	Val	Pro	Val	Val	Ala	Ile	Pro	Val	Met	Pro	Glu	Pro	Arg	Ala		
				325						330				335			
acc	gcg	cgg	cgg	atc	gtc	gaa	ctg	ggc	ctg	ggc	gac	cag	ctg	ctg	aac	1056	
Thr	Ala	Arg	Arg	Ile	Val	Glu	Leu	Gly	Leu	Gly	Asp	Gln	Leu	Leu	Asn		
				340					345				350				
tcg	gag	ctg	acg	gcc	gag	tcc	ctg	cgt	gcc	acg	gta	cgg	cgg	gtg	ctg	1104	
Ser	Glu	Leu	Thr	Ala	Glu	Ser	Leu	Arg	Ala	Thr	Val	Arg	Arg	Val	Leu		
				355				360					365				
gcg	gac	tcc	gcg	atc	ccg	ggg	aac	atg	cgc	ggg	ttc	cgg	gag	cag	atc	1152	
Ala	Asp	Ser	Ala	Ile	Pro	Gly	Asn	Met	Arg	Gly	Phe	Arg	Glu	Gln	Ile		
			370				375					380					
agg	gcg	gcc	ggc	ggg	gcg	ccc	gcg	gcg	gcc	gac	gcg	atc	gag	gga	ctg	1200	
Arg	Ala	Ala	Gly	Gly	Ala	Pro	Ala	Ala	Ala	Asp	Ala	Ile	Glu	Gly	Leu		
385					390					395					400		
ctg	ccc	cgg	gtg	ggc	tga											1218	
Leu	Pro	Arg	Val	Gly													
				405													

<210> 108
 <211> 405
 <212> PRT

<213> Streptomyces ambofaciens

<400> 108

Met Ala His Ile Ala Phe Phe Ile Leu Pro Val Ala Gly His Val Asn
1 5 10 15

Pro Thr Leu Gly Val Ala Glu Glu Leu Val Ala Arg Gly His Arg Val
20 25 30

Thr Phe Ala Leu Ser Glu Asp Leu Ala Glu Arg Ala Arg Leu Ile Gly
35 40 45

Ala Glu Val Val Thr Tyr Pro Val Asp Arg Gln Arg Phe Leu Asp Gln
50 55 60

Met Val Pro Arg Gln Asp Ala Asp Glu Tyr Thr Asp Glu Asp Glu Phe
65 70 75 80

Val Arg Val Leu Glu Trp Leu Leu Asp Met Thr Val Gln Thr Met Glu
85 90 95

Pro Leu Glu Arg His Phe Ala Gly Asp Arg Pro Asp Val Val Val Asn
100 105 110

Asp Pro Ser Ser Leu Trp Thr Gly Arg Leu Leu Ala Asp Arg Trp Gly
115 120 125

Ile Pro Val Ile Arg Ser Thr Pro Thr Tyr Ala Ala Asn Glu His Trp
130 135 140

Ser Leu His Pro Pro Val Asp Ser Ala Glu Pro Pro Asp Asp Pro Glu
145 150 155 160

Leu His Lys Leu Leu Ala Arg Ile Glu Arg Leu Leu Glu Glu Gln Gly
165 170 175

Val Glu His Asp Leu Ala Gly Phe Thr Gly Val Leu His Gly Gly Pro
 180 185 190

Ala Leu Leu Tyr Met Pro Arg Ser Phe Gln Tyr Ala Gly Glu Thr Phe
 195 200 205

Asp Ala Gln His His Phe Val Gly Pro Cys Pro Pro Arg Thr Ala Phe
 210 215 220

His Gly Glu Trp Arg Pro Gly Asp Asp Asp Gly Arg Pro Leu Val Leu
 225 230 235 240

Val Ser Leu Gly Thr Leu Tyr Asn Asp Arg Pro Asp Phe Phe Arg Thr
 245 250 255

Cys Leu Glu Ala Phe Gly Asp Glu Pro Trp Asn Val Leu Leu Val Leu
 260 265 270

Gly Gly Gly Val Pro Ala Ala Asp Leu Gly Pro Leu Pro Gly Asn Val
 275 280 285

Arg Val Thr Asp Phe Val Ser Leu Arg Asp Val Leu Pro His Thr Ala
 290 295 300

Val Val Val Asn His Gly Gly Met Ser Thr Ala Met Glu Val Phe Ser
 305 310 315 320

His Gly Val Pro Val Val Ala Ile Pro Val Met Pro Glu Pro Arg Ala
 325 330 335

Thr Ala Arg Arg Ile Val Glu Leu Gly Leu Gly Asp Gln Leu Leu Asn
 340 345 350

Ser Glu Leu Thr Ala Glu Ser Leu Arg Ala Thr Val Arg Arg Val Leu
 355 360 365

Ala Asp Ser Ala Ile Pro Gly Asn Met Arg Gly Phe Arg Glu Gln Ile
 370 375 380

Arg Ala Ala Gly Gly Ala Pro Ala Ala Ala Asp Ala Ile Glu Gly Leu
 385 390 395 400

Leu Pro Arg Val Gly
 405

<210> 109

<211> 621

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1)..(621)

<400> 109

atg ctc atc acc gaa aca gca gtg ccc gac gtg ttc cgc atc gat ccg 48
 Met Leu Ile Thr Glu Thr Ala Val Pro Asp Val Phe Arg Ile Asp Pro
 1 5 10 15

gaa ccg ctg ccg gac cac cgg ggc cgg ttc tac gaa gcg gtg cgc cgc 96
 Glu Pro Leu Pro Asp His Arg Gly Arg Phe Tyr Glu Ala Val Arg Arg
 20 25 30

ggg ccg ctg gag gcc gcc gtc ggg cac tcg gtc gag gtc cgg cag gtc 144
 Gly Pro Leu Glu Ala Ala Val Gly His Ser Val Glu Val Arg Gln Val
 35 40 45

cac tgc acc gtc tcc ggg cgc aac gta ctg cgc ggc ctg cac gcc acc 192
 His Cys Thr Val Ser Gly Arg Asn Val Leu Arg Gly Leu His Ala Thr
 50 55 60

acc ctg ccg ccg ggc cag gcc aag atc ctc acc tgt gtg cgg ggt gcg 240
 Thr Leu Pro Pro Gly Gln Ala Lys Ile Leu Thr Cys Val Arg Gly Ala
 65 70 75 80

gcg ctg act atg gtg gtc gac atg agg gtc ggg tca ccg ggc ttc gga 288
 Ala Leu Thr Met Val Val Asp Met Arg Val Gly Ser Pro Gly Phe Gly

	85	90	95	
	cgg tac gag gcg gtg cgg cag gat gcc ccg tcg ggc acc gcg ctc tac			336
	Arg Tyr Glu Ala Val Arg Gln Asp Ala Pro Ser Gly Thr Ala Leu Tyr			
	100	105	110	
	ctg ccg gac ggc atc ggc ctg ggc tac gtg gcc ctc gcg gac gac aca			384
	Leu Pro Asp Gly Ile Gly Leu Gly Tyr Val Ala Leu Ala Asp Asp Thr			
	115	120	125	
	tgc atg aac tac ctg tgc acc gag gag tac gtc ccg ggc atg gtc atc			432
	Cys Met Asn Tyr Leu Cys Thr Glu Glu Tyr Val Pro Gly Met Val Ile			
	130	135	140	
	gac gtg gac gcc ctg gac ccc gaa ctc ggc ctg ccg tgg gga ctg acc			480
	Asp Val Asp Ala Leu Asp Pro Glu Leu Gly Leu Pro Trp Gly Leu Thr			
	145	150	155	160
	ggg gat ccc gtc cgc tcc gcc cgg gac gcg gcg gcg ccg tcc ctg cgg			528
	Gly Asp Pro Val Arg Ser Ala Arg Asp Ala Ala Ala Pro Ser Leu Arg			
	165	170	175	
	gcg gcc gcg gcg gcg gga att ctg ccc acg tac gag gac tgc tta cgg			576
	Ala Ala Ala Ala Ala Gly Ile Leu Pro Thr Tyr Glu Asp Cys Leu Arg			
	180	185	190	
	gtg cgc acg tcc gtg ccc gcc cct ccc gac cgg act cgg ctc tga			621
	Val Arg Thr Ser Val Pro Ala Pro Pro Asp Arg Thr Arg Leu			
	195	200	205	
<210>	110			
<211>	206			
<212>	PRT			
<213>	Streptomyces ambofaciens			
<400>	110			
Met Leu Ile Thr Glu Thr Ala Val Pro Asp Val Phe Arg Ile Asp Pro				
1	5	10	15	
Glu Pro Leu Pro Asp His Arg Gly Arg Phe Tyr Glu Ala Val Arg Arg				
20	25	30		
Gly Pro Leu Glu Ala Ala Val Gly His Ser Val Glu Val Arg Gln Val				

35

40

45

His Cys Thr Val Ser Gly Arg Asn Val Leu Arg Gly Leu His Ala Thr
 50 55 60

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

Ala Leu Thr Met Val Val Asp Met Arg Val Gly Ser Pro Gly Phe Gly
 85 90 95

Arg Tyr Glu Ala Val Arg Gln Asp Ala Pro Ser Gly Thr Ala Leu Tyr
 100 105 110

Leu Pro Asp Gly Ile Gly Leu Gly Tyr Val Ala Leu Ala Asp Asp Thr
 115 120 125

Cys Met Asn Tyr Leu Cys Thr Glu Glu Tyr Val Pro Gly Met Val Ile
 130 135 140

Asp Val Asp Ala Leu Asp Pro Glu Leu Gly Leu Pro Trp Gly Leu Thr
 145 150 155 160

Gly Asp Pro Val Arg Ser Ala Arg Asp Ala Ala Ala Pro Ser Leu Arg
 165 170 175

Ala Ala Ala Ala Ala Gly Ile Leu Pro Thr Tyr Glu Asp Cys Leu Arg
 180 185 190

Val Arg Thr Ser Val Pro Ala Pro Pro Asp Arg Thr Arg Leu
 195 200 205

<210> 111

<211> 960

<212> DNA

<213> Streptomyces ambofaciens

$\langle 220 \rangle$

<221> CDS

<222> (1) .. (903)

 $\langle 220 \rangle$

```
<221> misc_feature
```

<222> (905) .. (960)

<223> n e a, c, g, ou t

<400> 111

atg acc tta cgc acg ctt cga ccc gcc ggc gcg gcg ccc ccc gtg ggt 48
Met Thr Leu Arg Thr Leu Arg Pro Ala Gly Ala Ala Pro Pro Val Gly
1 5 10 15

```

ccc gcc gcg cgg cac gac ttc gac tgg ctc gtc tcc gac gca tgt gcg      96
Pro Ala Ala Arg His Asp Phe Asp Ser Leu Val Ser Asp Ala Cys Ala
          20                25                30

```

gag tta ctc ggt tgc ctc cgc cgg gcg gac cag cgc aga aga ggt gag 144
Glu Leu Leu Gly Ser Leu Arg Arg Ala Asp Gln Arg Arg Arg Gly Glu
35 40 45

cag tac ata cgc ggg ctg ctc acg gca cag gga cgc aag acc gcc cgg 192
Gln Tyr Ile Arg Gly Leu Leu Thr Ala Gln Gly Arg Lys Thr Ala Arg
50 55 60

aac ctc gcc gcc ttc gtc gga gag ggc gcg gcg gaa cag agc ctg cac 240
Asn Leu Ala Ala Phe Val Gly Glu Gly Ala Ala Glu Gln Ser Leu His
65 70 75 80

cac ttc gtc gcc ggg tcg acc tgg gac tgg ggc gag gtg cgg gcc gcg 288
His Phe Val Ala Gly Ser Thr Trp Asp Trp Gly Glu Val Arg Ala Ala
 85 90 95

ctg gcc cgc tac gtg gac gac agg ctg acc ccc gac gcc tgg gtg atc 336
Leu Ala Arg Tyr Val Asp Asp Arg Leu Thr Pro Asp Ala Trp Val Ile
100 105 110

tgg ccg atg gtg gtc tcc aag gcg ggg gtg cgc tcc gtc ggg gtg agc 384
Trp Pro Met Val Val Ser Lys Ala Gly Val Arg Ser Val Gly Val Ser
115 120 125

cgg cgc ttc gtc tcc gac ctg ggc cgg gtg gtg agc tgc cag cag agt 432
Arg Arg Phe Val Ser Asp Leu Gly Arg Val Val Ser Cys Gln Gln Ser
130 135 140

cac ggg ctg tgg ctg gcc tcg ggc acg acg gcc gcc ccg gtc agc tgg 480
 His Gly Leu Trp Leu Ala Ser Gly Thr Thr Ala Ala Pro Val Ser Trp
 145 150 155 160

cgc ctg acg ctg ggc ggc ggc cgg agc ggg gag ggc aca ccc cgg cac 528
 Arg Leu Thr Leu Gly Gly Gly Arg Ser Gly Glu Gly Thr Pro Arg His
 165 170 175

ccc ggc gca ccc ggt gag gag gag aac atc gtc cgc ctg gtg gcc gaa 576
Pro Gly Ala Pro Gly Glu Glu Glu Asn Ile Val Arg Leu Val Ala Glu
180 185 190

gcc gct cag gcc aac cgc tcc ttg gcc cgc ccc gtg gtg atg gac gca 624
Ala Ala Gln Ala Asn Arg Ser Leu Ala Arg Pro Val Val Met Asp Ala
195 200 205

cgc gcg gcc gtg ctg cca cga ctg gtg cga ggg ctc ggt ctc gcg ggc 672
 Arg Ala Ala Val Leu Pro Arg Leu Val Arg Gly Leu Gly Leu Ala Gly
 210 215 220

ctg ccc ttc atg gtg cgg gtg ggc ggg aac ctc cgg ctg gcc tcg gcg 720
Leu Pro Phe Met Val Arg Val Gly Gly Asn Leu Arg Leu Ala Ser Ala
225 230 235 240

agc ggc cgc ggc ccg gtg gac cac cac acg ggg acc acc tcg gcc cag 768
Ser Gly Arg Gly Pro Val Asp His His Thr Gly Thr Thr Ser Ala Gln
245 250 255

cag ttg atg gag cag ctc aag cgg ttg agc cgc ccc gtg gag ttg cgg 816
Gln Leu Met Glu Gln Leu Lys Arg Leu Ser Arg Pro Val Glu Leu Arg
260 265 270

ggc tcc ctc agc ctc gtc gcc ccg cac ccc gtc gtc ctg ccg ggt gtg 864
Gly Ser Leu Ser Leu Val Ala Pro His Pro Val Val Leu Pro Gly Val
275 280 285

```
gtc ccg agg cgg acg ctg acg ctg ctg ggg ggt gtg gcg gnnnnnnnnnn      913
Val Pro Arg Arg Thr Leu Thr Leu Leu Gly Gly Val Ala
      290              295              300
```

nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnn 960

 $\langle 210 \rangle$ 112

<211> 301

<212> PRT

<213> Streptomyces ambofaciens

<400> 112

Met Thr Leu Arg Thr Leu Arg Pro Ala Gly Ala Ala Pro Pro Val Gly
1 5 10 15

Pro Ala Ala Arg His Asp Phe Asp Ser Leu Val Ser Asp Ala Cys Ala
20 25 30

Glu Leu Leu Gly Ser Leu Arg Arg Ala Asp Gln Arg Arg Arg Gly Glu
35 40 45

Gln Tyr Ile Arg Gly Leu Leu Thr Ala Gln Gly Arg Lys Thr Ala Arg
50 55 60

Asn Leu Ala Ala Phe Val Gly Glu Gly Ala Ala Glu Gln Ser Leu His
65 70 75 80

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

Leu Ala Arg Tyr Val Asp Asp Arg Leu Thr Pro Asp Ala Trp Val Ile
100 105 110

Trp Pro Met Val Val Ser Lys Ala Gly Val Arg Ser Val Gly Val Ser
115 120 125

Arg Arg Phe Val Ser Asp Leu Gly Arg Val Val Ser Cys Gln Gln Ser
130 135 140

His Gly Leu Trp Leu Ala Ser Gly Thr Thr Ala Ala Pro Val Ser Trp
145 150 155 160

Arg Leu Thr Leu Gly Gly Gly Arg Ser Gly Glu Gly Thr Pro Arg His
165 170 175

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

Ala Ala Gln Ala Asn Arg Ser Leu Ala Arg Pro Val Val Met Asp Ala
 195 200 205

Arg Ala Ala Val Leu Pro Arg Leu Val Arg Gly Leu Gly Leu Ala Gly
 210 215 220

Leu Pro Phe Met Val Arg Val Gly Gly Asn Leu Arg Leu Ala Ser Ala
 225 230 235 240

Ser Gly Arg Gly Pro Val Asp His His Thr Gly Thr Thr Ser Ala Gln
 245 250 255

Gln Leu Met Glu Gln Leu Lys Arg Leu Ser Arg Pro Val Glu Leu Arg
 260 265 270

Gly Ser Leu Ser Leu Val Ala Pro His Pro Val Val Leu Pro Gly Val
 275 280 285

Val Pro Arg Arg Thr Leu Thr Leu Leu Gly Gly Val Ala
 290 295 300

<210> 113

<211> 1008

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1)..(1008)

<400> 113

atg cag gca att cgg cat cac gtc atg ctc gtc atg gcc ttc gtg acc
 Met Gln Ala Ile Arg His His Val Met Leu Val Met Ala Phe Val Thr
 1 5 10 15

gtc gcc acg act ttc ctt ctc tgg ccg tcg acg caa tcc gcg caa gcg	96
Val Ala Thr Thr Phe Leu Leu Trp Pro Ser Thr Gln Ser Ala Gln Ala	
20 25 30	
ttt ccc ccg acc ccg aag cag acg gta ctg aac cac ctc cgc gcc att	144
Phe Pro Pro Thr Pro Lys Gln Thr Val Leu Asn His Leu Arg Ala Ile	
35 40 45	
tcc ggg aat cac atc gtc tcc gga cag cac aac aag gag ccc gcc tcc	192
Ser Gly Asn His Ile Val Ser Gly Gln His Asn Lys Glu Pro Ala Ser	
50 55 60	
gcc ccg ggc cag tac acc cag cag gtc aag gac gtc acc ggg cag tac	240
Ala Pro Gly Gln Tyr Thr Gln Gln Val Lys Asp Val Thr Gly Gln Tyr	
65 70 75 80	
ccc ggc ctg tgg ggc ggt gac ctg atg ttc gcc gcg gcg gac gtg gcc	288
Pro Gly Leu Trp Gly Gly Asp Leu Met Phe Ala Ala Ala Asp Val Ala	
85 90 95	
ggc cgc cag cgc gtc gtc gac cag gcc agg acc gag tgg gcg aac gga	336
Gly Arg Gln Arg Val Val Asp Gln Ala Arg Thr Glu Trp Ala Asn Gly	
100 105 110	
tcg ctg gtc tcg ctc acc tgg cac gtc tgc ccg ccg acc ggc ggc agc	384
Ser Leu Val Ser Leu Thr Trp His Val Cys Pro Pro Thr Gly Gly Ser	
115 120 125	
acc tgt gcg ttc gag ggc ggc gtc aag tcc acg ctg acg aac gcg cag	432
Thr Cys Ala Phe Glu Gly Gly Val Lys Ser Thr Leu Thr Asn Ala Gln	
130 135 140	
ttc tcg cag gtc ctc acg gag ggc agt gcc ctg aac agc gca tgg aag	480
Phe Ser Gln Val Leu Thr Glu Gly Ser Ala Leu Asn Ser Ala Trp Lys	
145 150 155 160	
cgg cgc ctg gac gag gtc gtc ccg tac ctg cag cag ctg gag aac gcg	528
Arg Arg Leu Asp Glu Val Val Pro Tyr Leu Gln Gln Leu Glu Asn Ala	
165 170 175	
ggc gtc ccc gtc ctc ttc cgg ccg ctg cac gag atg aac gaa tcc tgg	576
Gly Val Pro Val Leu Phe Arg Pro Leu His Glu Met Asn Glu Ser Trp	
180 185 190	
aac tgg tgg gga aac cgg ccc gga gcg aac ggc agc gca cgc ctc tac	624
Asn Trp Trp Gly Asn Arg Pro Gly Ala Asn Gly Ser Ala Arg Leu Tyr	

195					200					205										
cag	atc	acc	cgc	gat	cac	ctc	gcc	ggg	acg	aaa	ggg	ctg	gac	aat	ctg	672				
Gln	Ile	Thr	Arg	Asp	His	Leu	Ala	Gly	Thr	Lys	Gly	Leu	Asp	Asn	Leu					
210					215					220										
atc	tgg	gtc	tgg	aac	gtc	cag	gac	aat	ccg	gcg	gga	aac	tgg	aac	agc	720				
Ile	Trp	Val	Trp	Asn	Val	Gln	Asp	Asn	Pro	Ala	Gly	Asn	Trp	Asn	Ser					
225					230					235					240					
tac	tat	ccg	gga	gat	cag	tac	gtg	gac	gtc	gtt	tcg	ctg	gac	gtc	tgg	768				
Tyr	Tyr	Pro	Gly	Asp	Gln	Tyr	Val	Asp	Val	Val	Ser	Leu	Asp	Val	Trp					
245					250					255										
tac	aag	agc	cac	ccg	agt	tcc	gcc	gac	tac	cag	cag	atg	cgg	agc	atc	816				
Tyr	Lys	Ser	His	Pro	Ser	Ser	Ala	Asp	Tyr	Gln	Gln	Met	Arg	Ser	Ile					
260					265					270										
gcg	gga	aca	aaa	ccc	atg	gcc	ctc	gcg	gag	ctg	ggc	aaa	atg	ccg	acc	864				
Ala	Gly	Thr	Lys	Pro	Met	Ala	Leu	Ala	Glu	Leu	Gly	Lys	Met	Pro	Thr					
275					280					285										
gcc	gcg	ctg	ctg	gac	agc	cag	acg	cgg	tgg	aca	tgg	ttc	atg	atg	tgg	912				
Ala	Ala	Leu	Leu	Asp	Ser	Gln	Thr	Arg	Trp	Thr	Trp	Phe	Met	Met	Trp					
290					295					300										
tcc	gag	cat	ctg	cgc	ggg	aac	aat	tcc	aac	gcc	gaa	ata	cag	acg	gcg	960				
Ser	Glu	His	Leu	Arg	Gly	Asn	Asn	Ser	Asn	Ala	Glu	Ile	Gln	Thr	Ala					
305					310					315					320					
tat	ttc	cac	ccc	cgt	gta	ctg	aac	cag	ggg	gag	gtc	gca	ctg	ccc	tga	1008				
Tyr	Phe	His	Pro	Arg	Val	Leu	Asn	Gln	Gly	Glu	Val	Ala	Leu	Pro						
325					330					335										

```
<210> 114
<211> 335
<212> PRT
<213> Streptomyces ambofaciens
```

<400> 114

Met Gln Ala Ile Arg His His Val Met Leu Val Met Ala Phe Val Thr
1 5 10 15

Val Ala Thr Thr Phe Leu Leu Trp Pro Ser Thr Gln Ser Ala Gln Ala

20

25

30

Phe Pro Pro Thr Pro Lys Gln Thr Val Leu Asn His Leu Arg Ala Ile
 35 40 45

Ser Gly Asn His Ile Val Ser Gly Gln His Asn Lys Glu Pro Ala Ser
 50 55 60

Ala Pro Gly Gln Tyr Thr Gln Gln Val Lys Asp Val Thr Gly Gln Tyr
 65 70 75 80

Pro Gly Leu Trp Gly Gly Asp Leu Met Phe Ala Ala Ala Asp Val Ala
 85 90 95

Gly Arg Gln Arg Val Val Asp Gln Ala Arg Thr Glu Trp Ala Asn Gly
 100 105 110

Ser Leu Val Ser Leu Thr Trp His Val Cys Pro Pro Thr Gly Gly Ser
 115 120 125

Thr Cys Ala Phe Glu Gly Gly Val Lys Ser Thr Leu Thr Asn Ala Gln
 130 135 140

Phe Ser Gln Val Leu Thr Glu Gly Ser Ala Leu Asn Ser Ala Trp Lys
 145 150 155 160

Arg Arg Leu Asp Glu Val Val Pro Tyr Leu Gln Gln Leu Glu Asn Ala
 165 170 175

Gly Val Pro Val Leu Phe Arg Pro Leu His Glu Met Asn Glu Ser Trp
 180 185 190

Asn Trp Trp Gly Asn Arg Pro Gly Ala Asn Gly Ser Ala Arg Leu Tyr
 195 200 205

246

Gln Ile Thr Arg Asp His Leu Ala Gly Thr Lys Gly Leu Asp Asn Leu
 210 215 220

Ile Trp Val Trp Asn Val Gln Asp Asn Pro Ala Gly Asn Trp Asn Ser
 225 230 235 240

Tyr Tyr Pro Gly Asp Gln Tyr Val Asp Val Val Ser Leu Asp Val Trp
 245 250 255

Tyr Lys Ser His Pro Ser Ser Ala Asp Tyr Gln Gln Met Arg Ser Ile
 260 265 270

Ala Gly Thr Lys Pro Met Ala Leu Ala Glu Leu Gly Lys Met Pro Thr
 275 280 285

Ala Ala Leu Leu Asp Ser Gln Thr Arg Trp Thr Trp Phe Met Met Trp
 290 295 300

Ser Glu His Leu Arg Gly Asn Asn Ser Asn Ala Glu Ile Gln Thr Ala
 305 310 315 320

Tyr Phe His Pro Arg Val Leu Asn Gln Gly Glu Val Ala Leu Pro
 325 330 335

<210> 115
 <211> 1038
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1) .. (1038)

<220>
 <221> CDS2
 <222> (190) .. (1038)

<400> 115

atg acg tgc ggg ctc ggg cgg gaa cgg ctc cac cgt aga aag gaa gcg	48
Met Thr Cys Gly Leu Gly Arg Glu Arg Leu His Arg Arg Lys Glu Ala	
1 5 10 15	
gtc cgg cgt ccc ggt cca aaa gag cac ggg aaa acc gga ccg atc gcg	96
Val Arg Arg Pro Gly Pro Lys Glu His Gly Lys Thr Gly Pro Ile Ala	
20 25 30	
cgt ctg tgc cct ggc ggg ccg gct cgt tcc aag ccc tgg tcg cgg ttg	144
Arg Leu Cys Pro Gly Gly Pro Ala Arg Ser Lys Pro Trp Ser Arg Leu	
35 40 45	
gcc ccc tac ggt ggt acg acc gac gac aag gcg acg ggg gcg acg atg	192
Ala Pro Tyr Gly Gly Thr Thr Asp Asp Lys Ala Thr Gly Ala Thr Met	
50 55 60	
cgt atc gcg gtg aca ggg gcg gcg ggc agt ctg ggc agg cgg gtg gtg	240
Arg Ile Ala Val Thr Gly Ala Ala Gly Ser Leu Gly Arg Arg Val Val	
65 70 75 80	
cag ctg ctc gcg gac cgg gcc gac gtg gac gtc gtg gcg atg acc agg	288
Gln Leu Leu Ala Asp Arg Ala Asp Val Asp Val Val Ala Met Thr Arg	
85 90 95	
cgg aag ctg cct gcc gag gca cta ccg ccg ggg gtg gag tgc gcc gtc	336
Arg Lys Leu Pro Ala Glu Ala Leu Pro Pro Gly Val Glu Cys Ala Val	
100 105 110	
gcc gac tac gcc gac ccg ccc gct ctg cgt gcg gcg ctg aag ggc gtg	384
Ala Asp Tyr Ala Asp Pro Pro Ala Leu Arg Ala Ala Leu Lys Gly Val	
115 120 125	
gac act ctg gtc ctc gtc tcc agc gac ggg ccc gac gca cgg gtg ctg	432
Asp Thr Leu Val Leu Val Ser Ser Asp Gly Pro Asp Ala Arg Val Leu	
130 135 140	
ctg cac cac cgc aac gtc gtc gcc gcc gtg gtg gcc gag cgc gtc ggc	480
Leu His His Arg Asn Val Val Ala Ala Val Val Ala Glu Arg Val Gly	
145 150 155 160	
cat gtc gcg gca ctg agc agc gtc gac gcc gac tcg gcc tcc ccg ttc	528
His Val Ala Ala Leu Ser Ser Val Asp Ala Asp Ser Ala Ser Pro Phe	
165 170 175	
tgc tac gcc gtc gtc aac cgg ctc acc gag gat ctg ctg ctc gcc tcc	576
Cys Tyr Ala Val Val Asn Arg Leu Thr Glu Asp Leu Leu Leu Ala Ser	
180 185 190	

ggc gtg ccg tgc tcg ttc gcc agg gcc tcg ctg tac atc gag ttc ttc 624
 Gly Val Pro Cys Ser Phe Ala Arg Ala Ser Leu Tyr Ile Glu Phe Phe
 195 200 205

cta ggc tgg ctc acc cag gcg cgc tcg acc ggg ctg ctg agg ctg ccg 672
 Leu Gly Trp Leu Thr Gln Ala Arg Ser Thr Gly Leu Leu Arg Leu Pro
 210 215 220

gca gcc gat ggc cgg gtg tcc ctg gtg gcg cgc gac gac gtg gcc cgc 720
 Ala Ala Asp Gly Arg Val Ser Leu Val Ala Arg Asp Asp Val Ala Arg
 225 230 235 240

gcc ctc gcc gcc ctc gcg gtg ggc ggg ccg acc ggc cgt cac cac gac 768
 Ala Leu Ala Ala Leu Ala Val Gly Gly Pro Thr Gly Arg His His Asp
 245 250 255

atc acg gga ccc gag tcg acg gac ctg gca ggc atc gcc tcc acg gcg 816
 Ile Thr Gly Pro Glu Ser Thr Asp Leu Ala Gly Ile Ala Ser Thr Ala
 260 265 270

gcg gag gtg tgg cgg aca ccg gtc gcc tac gtg gac atc ccg gcc gac 864
 Ala Glu Val Trp Arg Thr Pro Val Ala Tyr Val Asp Ile Pro Ala Asp
 275 280 285

acg tac tct gcc gaa aca gcg gcg acc ggc ctg gat ccc tgg tgg ctc 912
 Thr Tyr Ser Ala Glu Thr Ala Ala Thr Gly Leu Asp Pro Trp Trp Leu
 290 295 300

tac gcc ttc tcc tcc ctg ttc gcc tcg gta cgc gaa cag cgc tgg gat 960
 Tyr Ala Phe Ser Ser Leu Phe Ala Ser Val Arg Glu Gln Arg Trp Asp
 305 310 315 320

cgg gtc cgc gac gac tgc gcc cag ctc acc ggc cgc ttg ccc cgc act 1008
 Arg Val Arg Asp Asp Cys Ala Gln Leu Thr Gly Arg Leu Pro Arg Thr
 325 330 335

ctc cgc gac gtc ctg tca gaa cgc gcg tga 1038
 Leu Arg Asp Val Leu Ser Glu Arg Ala
 340 345

<210> 116

<211> 345

<212> PRT

<213> Streptomyces ambofaciens

<400> 116

Met Thr Cys Gly Leu Gly Arg Glu Arg Leu His Arg Arg Lys Glu Ala
 1 5 10 15

Val Arg Arg Pro Gly Pro Lys Glu His Gly Lys Thr Gly Pro Ile Ala
 20 25 30

Arg Leu Cys Pro Gly Gly Pro Ala Arg Ser Lys Pro Trp Ser Arg Leu
 35 40 45

Ala Pro Tyr Gly Gly Thr Thr Asp Asp Lys Ala Thr Gly Ala Thr Met
 50 55 60

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

Gln Leu Leu Ala Asp Arg Ala Asp Val Asp Val Val Ala Met Thr Arg
 85 90 95

Arg Lys Leu Pro Ala Glu Ala Leu Pro Pro Gly Val Glu Cys Ala Val
 100 105 110

Ala Asp Tyr Ala Asp Pro Pro Ala Leu Arg Ala Ala Leu Lys Gly Val
 115 120 125

Asp Thr Leu Val Leu Val Ser Ser Asp Gly Pro Asp Ala Arg Val Leu
 130 135 140

Leu His His Arg Asn Val Val Ala Ala Val Val Ala Glu Arg Val Gly
 145 150 155 160

His Val Ala Ala Leu Ser Ser Val Asp Ala Asp Ser Ala Ser Pro Phe
 165 170 175

Cys Tyr Ala Val Val Asn Arg Leu Thr Glu Asp Leu Leu Leu Ala Ser

180

185

190

Gly Val Pro Cys Ser Phe Ala Arg Ala Ser Leu Tyr Ile Glu Phe Phe
 195 200 205

Leu Gly Trp Leu Thr Gln Ala Arg Ser Thr Gly Leu Leu Arg Leu Pro
 210 215 220

Ala Ala Asp Gly Arg Val Ser Leu Val Ala Arg Asp Asp Val Ala Arg
 225 230 235 240

Ala Leu Ala Ala Leu Ala Val Gly Gly Pro Thr Gly Arg His His Asp
 245 250 255

Ile Thr Gly Pro Glu Ser Thr Asp Leu Ala Gly Ile Ala Ser Thr Ala
 260 265 270

Ala Glu Val Trp Arg Thr Pro Val Ala Tyr Val Asp Ile Pro Ala Asp
 275 280 285

Thr Tyr Ser Ala Glu Thr Ala Ala Thr Gly Leu Asp Pro Trp Trp Leu
 290 295 300

Tyr Ala Phe Ser Ser Leu Phe Ala Ser Val Arg Glu Gln Arg Trp Asp
 305 310 315 320

Arg Val Arg Asp Asp Cys Ala Gln Leu Thr Gly Arg Leu Pro Arg Thr
 325 330 335

Leu Arg Asp Val Leu Ser Glu Arg Ala
 340 345

<210> 117

<211> 282

<212> PRT

<213> Streptomyces ambofaciens

<400> 117

Met	Arg	Ile	Ala	Val	Thr	Gly	Ala	Ala	Gly	Ser	Leu	Gly	Arg	Arg	Val
1				5					10					15	

Val	Gln	Leu	Leu	Ala	Asp	Arg	Ala	Asp	Val	Asp	Val	Val	Ala	Met	Thr
			20					25					30		

Arg	Arg	Lys	Leu	Pro	Ala	Glu	Ala	Leu	Pro	Pro	Gly	Val	Glu	Cys	Ala
		35					40					45			

Val	Ala	Asp	Tyr	Ala	Asp	Pro	Pro	Ala	Leu	Arg	Ala	Ala	Leu	Lys	Gly
	50					55					60				

Val	Asp	Thr	Leu	Val	Leu	Val	Ser	Ser	Asp	Gly	Pro	Asp	Ala	Arg	Val
65					70					75					80

Leu	Leu	His	His	Arg	Asn	Val	Val	Ala	Ala	Val	Val	Ala	Glu	Arg	Val
				85					90					95	

Gly	His	Val	Ala	Ala	Leu	Ser	Ser	Val	Asp	Ala	Asp	Ser	Ala	Ser	Pro
			100					105					110		

Phe	Cys	Tyr	Ala	Val	Val	Asn	Arg	Leu	Thr	Glu	Asp	Leu	Leu	Leu	Ala
		115					120					125			

Ser	Gly	Val	Pro	Cys	Ser	Phe	Ala	Arg	Ala	Ser	Leu	Tyr	Ile	Glu	Phe
	130					135					140				

Phe	Leu	Gly	Trp	Leu	Thr	Gln	Ala	Arg	Ser	Thr	Gly	Leu	Leu	Arg	Leu
145					150					155					160

Pro	Ala	Ala	Asp	Gly	Arg	Val	Ser	Leu	Val	Ala	Arg	Asp	Asp	Val	Ala
				165					170					175	

Arg Ala Leu Ala Ala Leu Ala Val Gly Gly Pro Thr Gly Arg His His
 180 185 190

Asp Ile Thr Gly Pro Glu Ser Thr Asp Leu Ala Gly Ile Ala Ser Thr
 195 200 205

Ala Ala Glu Val Trp Arg Thr Pro Val Ala Tyr Val Asp Ile Pro Ala
 210 215 220

Asp Thr Tyr Ser Ala Glu Thr Ala Ala Thr Gly Leu Asp Pro Trp Trp
 225 230 235 240

Leu Tyr Ala Phe Ser Ser Leu Phe Ala Ser Val Arg Glu Gln Arg Trp
 245 250 255

Asp Arg Val Arg Asp Asp Cys Ala Gln Leu Thr Gly Arg Leu Pro Arg
 260 265 270

Thr Leu Arg Asp Val Leu Ser Glu Arg Ala
 275 280

<210> 118
 <211> 975
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(975)

<400> 118
 atg ctc cga cgc cgt ctc gga ggg ccg tca ggc ccc ctc gtc agt gcc 48
 Met Leu Arg Arg Arg Leu Gly Gly Pro Ser Gly Pro Leu Val Ser Ala
 1 5 10 15
 ctg tgc ctg ggc gcg atg ccc ttc ggc acc acc gtc gac gag aag acg 96
 Leu Cys Leu Gly Ala Met Pro Phe Gly Thr Thr Val Asp Glu Lys Thr
 20 25 30

tcc ttc gcc atc ctc gac cgg ttc gtc gag gcc ggc ggc agt ctc gtc	144
Ser Phe Ala Ile Leu Asp Arg Phe Val Glu Ala Gly Gly Ser Leu Val	
35 40 45	
gac acc gcc gac aac tac gcg ttc tgg gct ccc ggc ggc acc ggc gac	192
Asp Thr Ala Asp Asn Tyr Ala Phe Trp Ala Pro Gly Gly Thr Gly Asp	
50 55 60	
gag agc gag aac acc gtc ggg cgc tgg ctg gcg agc cgc cgc cgc cgc	240
Glu Ser Glu Asn Thr Val Gly Arg Trp Leu Ala Ser Arg Arg Arg Arg	
65 70 75 80	
gac gag gtg gtg atc tcc acc aag gtg ggt gcc cgc ccc acc gtc ccc	288
Asp Glu Val Val Ile Ser Thr Lys Val Gly Ala Arg Pro Thr Val Pro	
85 90 95	
ggc agc ggc ctg gag acc gcc gaa ggg ctg tcg gct ccc gtc ata cgg	336
Gly Ser Gly Leu Glu Thr Ala Glu Gly Leu Ser Ala Pro Val Ile Arg	
100 105 110	
aag gcc gcg gag gac agc ctg cga cgc ctg ggc acc gat cgc atc gat	384
Lys Ala Ala Glu Asp Ser Leu Arg Arg Leu Gly Thr Asp Arg Ile Asp	
115 120 125	
ctg tac tgg acc cac atc gag gac cgg acc gtc cct ctg gag gag acg	432
Leu Tyr Trp Thr His Ile Glu Asp Arg Thr Val Pro Leu Glu Glu Thr	
130 135 140	
ctc gga gct ctc gac gag ctg gtc ggc ggc ggc aag gtg gcg gtg ctg	480
Leu Gly Ala Leu Asp Glu Leu Val Gly Gly Gly Lys Val Ala Val Leu	
145 150 155 160	
ggc tgc tcc aac cac gcg ggc tgg cgc atc gaa cgg gcc cgc gcg ctc	528
Gly Cys Ser Asn His Ala Gly Trp Arg Ile Glu Arg Ala Arg Ala Leu	
165 170 175	
gcc cgg acc aac ggc tgg acg gcg tac acc tgc gtc cag cag cgc tac	576
Ala Arg Thr Asn Gly Trp Thr Ala Tyr Thr Cys Val Gln Gln Arg Tyr	
180 185 190	
tcc tac ctt cag ccg cgc ttc gac gtc gga ctg ccg gag agc gga cac	624
Ser Tyr Leu Gln Pro Arg Phe Asp Val Gly Leu Pro Glu Ser Gly His	
195 200 205	
gtc cac gcg acc tcc gaa ctc ctc gac cac gta cgc agc gag ccc gac	672
Val His Ala Thr Ser Glu Leu Leu Asp His Val Arg Ser Glu Pro Asp	
210 215 220	

ctg acg ctg ctg gcc tac tcg tcc ctg ctc tcg ggc gcg tac acc cgg 720
 Leu Thr Leu Leu Ala Tyr Ser Ser Leu Leu Ser Gly Ala Tyr Thr Arg
 225 230 235 240

ccg gac aag acc ctg tcc gcc gcc tac gac cac ccg ggc acc ggg cag 768
 Pro Asp Lys Thr Leu Ser Ala Ala Tyr Asp His Pro Gly Thr Gly Gln
 245 250 255

cgg ctg acc gtg ctg cgg gag gtg gcc gcc gag ctc ggt gcc acc gcc 816
 Arg Leu Thr Val Leu Arg Glu Val Ala Ala Glu Leu Gly Ala Thr Ala
 260 265 270

aac cag gtg gtg ctg tcc tgg ttg ctc gga ggt gat ccg ccg gtg atc 864
 Asn Gln Val Val Leu Ser Trp Leu Leu Gly Gly Asp Pro Pro Val Ile
 275 280 285

ccg atc gtg ggg gtg agt tcc gtc gag cag ctg gac gag gtg ctg gcc 912
 Pro Ile Val Gly Val Ser Ser Val Glu Gln Leu Asp Glu Val Leu Ala
 290 295 300

gcc gtc gag ctg gag ctg ccc cgg gag acg agg gca cgg ctg gac ctc 960
 Ala Val Glu Leu Glu Leu Pro Arg Glu Thr Arg Ala Arg Leu Asp Leu
 305 310 315 320

gcc ggg cgg ggc tga 975
 Ala Gly Arg Gly

<210> 119
 <211> 324
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 119

Met Leu Arg Arg Arg Leu Gly Gly Pro Ser Gly Pro Leu Val Ser Ala
 1 5 10 15

Leu Cys Leu Gly Ala Met Pro Phe Gly Thr Thr Val Asp Glu Lys Thr
 20 25 30

Ser Phe Ala Ile Leu Asp Arg Phe Val Glu Ala Gly Gly Ser Leu Val
 35 40 45

Asp Thr Ala Asp Asn Tyr Ala Phe Trp Ala Pro Gly Gly Thr Gly Asp
 50 55 60

Glu Ser Glu Asn Thr Val Gly Arg Trp Leu Ala Ser Arg Arg Arg Arg
 65 70 75 80

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

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

Lys Ala Ala Glu Asp Ser Leu Arg Arg Leu Gly Thr Asp Arg Ile Asp
 115 120 125

Leu Tyr Trp Thr His Ile Glu Asp Arg Thr Val Pro Leu Glu Glu Thr
 130 135 140

Leu Gly Ala Leu Asp Glu Leu Val Gly Gly Gly Lys Val Ala Val Leu
 145 150 155 160

Gly Cys Ser Asn His Ala Gly Trp Arg Ile Glu Arg Ala Arg Ala Leu
 165 170 175

Ala Arg Thr Asn Gly Trp Thr Ala Tyr Thr Cys Val Gln Gln Arg Tyr
 180 185 190

Ser Tyr Leu Gln Pro Arg Phe Asp Val Gly Leu Pro Glu Ser Gly His
 195 200 205

Val His Ala Thr Ser Glu Leu Leu Asp His Val Arg Ser Glu Pro Asp
 210 215 220

Leu Thr Leu Leu Ala Tyr Ser Ser Leu Leu Ser Gly Ala Tyr Thr Arg

225 230 235 240

Arg Leu Thr Val Leu Arg Glu Val Ala Ala Glu Leu Gly Ala Thr Ala
260 265 270

Asn Gln Val Val Leu Ser Trp Leu Leu Gly Gly Asp Pro Pro Val Ile
275 280 285

Pro Ile Val Gly Val Ser Ser Val Glu Gln Leu Asp Glu Val Leu Ala
290 295 300

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

Ala Gly Arg Gly

```
<210> 120
<211> 1227
<212> DNA
<213> Streptomyces ambofaciens
```

```
<220>
<221> CDS
<222> (1) .. (1227)
```

<400> 120
gtg acg ccc gag gcg gac gtc ggt acg cgg tgg gag ccg gag ttc gac 48
Val Thr Pro Glu Ala Asp Val Gly Thr Arg Trp Glu Pro Glu Phe Asp
1 5 10 15

atg ctg ccc ggc ata ccc gac ctg cgg gcc ttt ccc cgc gga cgg tgg 96
Met Leu Pro Gly Ile Pro Asp Leu Arg Ala Phe Pro Arg Gly Arg Trp
20 25 30

gcg gac gcg gtg cgg gag gcc acc ggc gca ctg acc tgg ggc gag ctc 144

Ala	Asp	Ala	Val	Arg	Glu	Ala	Thr	Gly	Ala	Leu	Thr	Trp	Gly	Glu	Leu	
		35					40					45				
ggc	tac	ccc	gat	ccg	gcc	ggt	ctc	gcc	cgg	ctc	cgg	cag	cgg	ctc	gcg	192
Gly	Tyr	Pro	Asp	Pro	Ala	Gly	Leu	Ala	Arg	Leu	Arg	Gln	Arg	Leu	Ala	
		50					55					60				
gcc	tac	ctc	gtg	cgg	gga	cgc	ggc	gcg	tcg	ctg	gaa	gcc	cgg	cag	ctc	240
Ala	Tyr	Leu	Val	Arg	Gly	Arg	Gly	Ala	Ser	Leu	Glu	Ala	Arg	Gln	Leu	
		65				70				75					80	
gtc	gtc	tgc	gcg	ggc	acg	ctg	gac	gcg	gtg	cgc	agg	ctg	tgc	cgg	gtc	288
Val	Val	Cys	Ala	Gly	Thr	Leu	Asp	Ala	Val	Arg	Arg	Leu	Cys	Arg	Val	
				85					90					95		
ctg	cgg	acc	gag	ggc	cac	acc	cac	atg	gcc	ctg	gag	gac	cct	ggc	tgg	336
Leu	Arg	Thr	Glu	Gly	His	Thr	His	Met	Ala	Leu	Glu	Asp	Pro	Gly	Trp	
			100						105					110		
tcg	cgg	ctg	agc	gcc	acg	gtc	ggg	gcg	gag	ggg	ctg	gtg	ccc	gtc	ccg	384
Ser	Arg	Leu	Ser	Ala	Thr	Val	Gly	Ala	Glu	Gly	Leu	Val	Pro	Val	Pro	
			115					120					125			
gtt	ccc	gtg	gac	ggc	gac	ggc	ctg	cgg	gtc	gac	ctg	ctg	cgg	cgc	gcg	432
Val	Pro	Val	Asp	Gly	Asp	Gly	Leu	Arg	Val	Asp	Leu	Leu	Arg	Arg	Ala	
		130					135						140			
cct	cag	gtc	cgg	tcg	gtg	atc	gtc	gct	ccc	gcc	cac	cag	ttc	ccg	acc	480
Pro	Gln	Val	Arg	Ser	Val	Ile	Val	Ala	Pro	Ala	His	Gln	Phe	Pro	Thr	
						150					155				160	
ggt	gtg	gtc	ctc	gcc	ccg	gcc	cgt	cgc	ggc	gag	ctg	gtc	gcc	tgg	gcc	528
Gly	Val	Val	Leu	Ala	Pro	Ala	Arg	Arg	Gly	Glu	Leu	Val	Ala	Trp	Ala	
					165					170				175		
cgc	gag	gtc	gac	ggc	ctg	gtg	ctg	gag	gac	gac	tac	gac	gcc	gag	ttc	576
Arg	Glu	Val	Asp	Gly	Leu	Val	Leu	Glu	Asp	Asp	Tyr	Asp	Ala	Glu	Phe	
				180					185					190		
cga	tac	gac	cgc	cac	ccc	gtg	ggc	gcc	ctg	cag	ggc	atg	gca	ccg	gag	624
Arg	Tyr	Asp	Arg	His	Pro	Val	Gly	Ala	Leu	Gln	Gly	Met	Ala	Pro	Glu	
			195					200				205				
cac	gtc	gcg	ctg	ctg	ggg	tcg	gtc	agc	aag	acg	ctc	tcg	ccg	gcc	ctg	672
His	Val	Ala	Leu	Leu	Gly	Ser	Val	Ser	Lys	Thr	Leu	Ser	Pro	Ala	Leu	
		210					215					220				

cgg atc ggc tgg gcg gtg acc ccg ccg cgc tgg acc gag gcc ctg ggc	720
Arg Ile Gly Trp Ala Val Thr Pro Pro Arg Trp Thr Glu Ala Leu Gly	
225 230 235 240	
gcg gag cac gtc ggg ggc gta cca ccg ccc gtc atc gac cag gag gcg	768
Ala Glu His Val Gly Gly Val Pro Pro Pro Val Ile Asp Gln Glu Ala	
245 250 255	
ttc gcc cgg ttc ctg gcc cgc ggg tcc tac gac cgg tac ctg cgg gcg	816
Phe Ala Arg Phe Leu Ala Arg Gly Ser Tyr Asp Arg Tyr Leu Arg Ala	
260 265 270	
gca cgg ctg cgc tac aaa cgg cgc cgc gat cac ctc atg gga gaa ctc	864
Ala Arg Leu Arg Tyr Lys Arg Arg Arg Asp His Leu Met Gly Glu Leu	
275 280 285	
gcg gcg tcg ctg ccg ggg ttc cag gtg tcc ggt gcg gcg gcc ggc ttc	912
Ala Ala Ser Leu Pro Gly Phe Gln Val Ser Gly Ala Ala Ala Gly Phe	
290 295 300	
cac ctg ctg ctg tcc ctg acg gac tgc tcg gcg ccc acc gtc gtc agg	960
His Leu Leu Leu Ser Leu Thr Asp Cys Ser Ala Pro Thr Val Val Arg	
305 310 315 320	
gag gcg gcg cgg cgt gac gtg agg ctg gcg gat ctc gac gcc tac cgg	1008
Glu Ala Ala Arg Arg Asp Val Arg Leu Ala Asp Leu Asp Ala Tyr Arg	
325 330 335	
atg gcg cgg gcc gag cac cac ccg acg gcc cgg gcc ggc gac tcc ggc	1056
Met Ala Arg Ala Glu His His Pro Thr Ala Arg Ala Gly Asp Ser Gly	
340 345 350	
ggg gcg cgg gcc gtg gcc gcc gcc ggc ccc gcg ctg gtc ctc ggc tac	1104
Gly Ala Arg Ala Val Ala Ala Ala Gly Pro Ala Leu Val Leu Gly Tyr	
355 360 365	
ggc aat ctg gcg gac cac gcc gtg ccc gga gcc gtg cgg cgc ctg gca	1152
Gly Asn Leu Ala Asp His Ala Val Pro Gly Ala Val Arg Arg Leu Ala	
370 375 380	
cag gcc atc gac gcg gcc cgg cgc cac gac cgg cgg cac gtc gac gcc	1200
Gln Ala Ile Asp Ala Ala Arg Arg His Asp Arg Arg His Val Asp Ala	
385 390 395 400	
ggg cgg gcc ggg gag cgg gcc gac tga	1227
Gly Arg Ala Gly Glu Arg Ala Asp	
405	

<210> 121
 <211> 408
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 121

Val Thr Pro Glu Ala Asp Val Gly Thr Arg Trp Glu Pro Glu Phe Asp
 1 5 10 15

Met Leu Pro Gly Ile Pro Asp Leu Arg Ala Phe Pro Arg Gly Arg Trp
 20 25 30

Ala Asp Ala Val Arg Glu Ala Thr Gly Ala Leu Thr Trp Gly Glu Leu
 35 40 45

Gly Tyr Pro Asp Pro Ala Gly Leu Ala Arg Leu Arg Gln Arg Leu Ala
 50 55 60

Ala Tyr Leu Val Arg Gly Arg Gly Ala Ser Leu Glu Ala Arg Gln Leu
 65 70 75 80

Val Val Cys Ala Gly Thr Leu Asp Ala Val Arg Arg Leu Cys Arg Val
 85 90 95

Leu Arg Thr Glu Gly His Thr His Met Ala Leu Glu Asp Pro Gly Trp
 100 105 110

Ser Arg Leu Ser Ala Thr Val Gly Ala Glu Gly Leu Val Pro Val Pro
 115 120 125

Val Pro Val Asp Gly Asp Gly Leu Arg Val Asp Leu Leu Arg Arg Ala
 130 135 140

Pro Gln Val Arg Ser Val Ile Val Ala Pro Ala His Gln Phe Pro Thr
 145 150 155 160

Gly Val Val Leu Ala Pro Ala Arg Arg Gly Glu Leu Val Ala Trp Ala
 165 170 175

Arg Glu Val Asp Gly Leu Val Leu Glu Asp Asp Tyr Asp Ala Glu Phe
 180 185 190

Arg Tyr Asp Arg His Pro Val Gly Ala Leu Gln Gly Met Ala Pro Glu
 195 200 205

His Val Ala Leu Leu Gly Ser Val Ser Lys Thr Leu Ser Pro Ala Leu
 210 215 220

Arg Ile Gly Trp Ala Val Thr Pro Pro Arg Trp Thr Glu Ala Leu Gly
 225 230 235 240

Ala Glu His Val Gly Gly Val Pro Pro Pro Val Ile Asp Gln Glu Ala
 245 250 255

Phe Ala Arg Phe Leu Ala Arg Gly Ser Tyr Asp Arg Tyr Leu Arg Ala
 260 265 270

Ala Arg Leu Arg Tyr Lys Arg Arg Arg Asp His Leu Met Gly Glu Leu
 275 280 285

Ala Ala Ser Leu Pro Gly Phe Gln Val Ser Gly Ala Ala Ala Gly Phe
 290 295 300

His Leu Leu Leu Ser Leu Thr Asp Cys Ser Ala Pro Thr Val Val Arg
 305 310 315 320

Glu Ala Ala Arg Arg Asp Val Arg Leu Ala Asp Leu Asp Ala Tyr Arg
 325 330 335

Met Ala Arg Ala Glu His His Pro Thr Ala Arg Ala Gly Asp Ser Gly

261

340

345

350

Gly Ala Arg Ala Val Ala Ala Ala Gly Pro Ala Leu Val Leu Gly Tyr
355 360 365

Gly Asn Leu Ala Asp His Ala Val Pro Gly Ala Val Arg Arg Leu Ala
370 375 380

Gln Ala Ile Asp Ala Ala Arg Arg His Asp Arg Arg His Val Asp Ala
385 390 395 400

Gly Arg Ala Gly Glu Arg Ala Asp
405

<210> 122

<211> 46

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR39

<400> 122

cccaagcttg agaagggagc ggacattcat ggcccgcgcc gaacgc

46

<210> 123

<211> 46

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR42

<400> 123

cgggatccgg ctgacatgg gagacgggag catgcgcgag ttcagc

46

<210> 124

<211> 46

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR40

<400> 124

cccaagcttg agaagggagc ggacattcaa tgctttggta aagcac

46

<210> 125

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR41

<400> 125

cccaagcttt caaggaacga cggggtgggc agtcaagt

38

<210> 126

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo A

<400> 126

ccagtagata tcccgccaac ccggagctgc ac

32

<210> 127

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo B

<400> 127

gaaaagatcc gtcattggggc cgtgcgctcc tt

32

<210> 128

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo C

<400> 128

cacgacccca tgacggatct tttccgctgc at

32

<210> 129

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo D

<400> 129

gagccggata tcatcggctc tgccttgctc gt

32

<210> 130

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo ORF23c

<400> 130

acgtgcgcgg tgagttcgcc gttgc

25

<210> 131

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo ORF25c

<400> 131

ctgaacgacg ccatcgcggt ggtgc

25

<210> 132

<211> 25

<212> DNA
 <213> Artificial

 <220>
 <223> Oligonucleotide ORF1*c

 <400> 132
 gaccacctcg aaccgtccgg cgtca 25

 <210> 133
 <211> 23
 <212> DNA
 <213> Artificial

 <220>
 <223> Oligonucleotide ORF2*c

 <400> 133
 ggcccgggtcc agcgtgccga agc 23

 <210> 134
 <211> 6174
 <212> DNA
 <213> Streptomyces ambofaciens

 <220>
 <221> misc_feature
 <222> (3620)..(4069)
 <223> n e a, c, g, ou t

 <400> 134
 ctgcagaggg cgccgacgtg cgcggtgagt tcgccgttgc gccggctggt ccagccgagg 60

 ccgccgtgga ccgcgggagc ctgtgcgcac agcaggcccc gggagccgag gtcgtgcagc 120

 aggccgaggg gcagctcgcc cgtccgggtcc cactcggccg cccggtcgcc gaccagcgcg 180

 gtgaacagct cctcggcctc ggtgccgtcg gcgtgggagg tgtcagccac ctgcgtgccc 240

 ggtcgcctcg ccgggagcgg ccagccgtcc gaccagccgg accatcgcg cgacgggtgcg 300

 gaagtgtgctg agcatcaggt cggcgccgct gatgacgatg ccgtaggtct tctccaggtg 360

 cacgacgagc tgcattggcg acagcgagga catcccgccg acggcgaaca ggtcctggtc 420

gcgctcccag	gtggtcttgg	tgcggtcctc	gaggaacccg	agcagttccc	cggcgacctc	480
gtcggcggtg	ggggtcgtgc	cggtggggtc	gggccgaccg	gacgtcgtgg	tcatcgcgtg	540
gcctcctggt	agtcgtagaa	tccccggccg	ctcttgccgc	cgagcaggcc	ctggcggacc	600
ttgtccagca	gcagctcgct	cgggcggagc	gccggatcgc	cggtccttcc	gtgcatcacc	660
cgcagcgagt	cggccagggt	gtccagtccg	atcaggtcgg	ccgtggccag	gggtccgggtg	720
cggtggccga	tgagtcgcg	catcagcgcg	tccacggtct	ccgggggtggc	ccggccctcg	780
tgaccaccgc	cgatggcgtc	gttcagcatc	cggtgcagca	ggcggctggt	cacgaagccg	840
gcgccgtcgc	cgacgacgat	gccccggcgg	cccaggccgg	acagcaggtc	ccgggtggcc	900
cgggcggccg	cctctccgct	gcgcggtccg	aggaccacct	cgaccgtggg	gatcacgtac	960
gcgggggttca	tgaagtgcac	gccgacgagg	tcctcggggc	gggggacggc	gtcggccagc	1020
tcgtcgatgg	ggacgcccga	ggtgttgctg	acgagcagcg	tccccgggcg	cgccacggac	1080
gccagggtccg	ccagcacctc	ggccttccgc	tcgggggtcct	cggtgacggc	ctcgatcacg	1140
gcggtcgcgg	tggcgacggc	ggccggcgcc	tcctcgacgg	tcagctcccc	gggcgggcgg	1200
tcgtggggca	gcgcgcccac	cagccggggc	gtccgcagat	gcagcgcgac	cgcgtcgggg	1260
gcggccgcgc	gcgccccggc	ggaggtgtcg	accagtgtga	ccgggtgccc	gtgtccgacg	1320
gcgagtgccg	cgatggccgt	gcccattgacg	cccgcgcoga	gcacgacgag	cggagaattt	1380
tccttggaat	cgggcacgga	tcctccagaa	tcccgggggc	ggcggctgtg	cgaaatgctg	1440
tcccgaatcg	cttccgcgct	ccatcacccg	ggcgtgactt	tgacacctag	caagcgggtg	1500
gaggccgggc	ggaggctatg	tcatctcggc	cttgtccgct	tacgagttcc	gtccttaaag	1560
tgcccccggc	gaacgcacgg	cgggaattgac	tcgccgtccg	tcaaggacct	gcggttcgaa	1620
ggacctcgac	gcatgaagg	tttcacatgg	ctcatatcgc	attcttcac	cttccggttg	1680
ccgggcatgt	gaatccgacc	ctgggagtcg	ccgagggaact	ggtcgcgcgc	ggccaccggg	1740
tgacgttcgc	gctgtccgag	gacctcgccg	agcggggccc	gctgatcggc	gccgaggtgg	1800

tcacctatcc	ggtggacagg	caacggttcc	tggaccagat	ggtgccgcgg	caggacgcgg	1860
acgagtacac	ggacgaggac	gagttcgtcc	gggtcctgga	gtggctgctg	gacatgacgg	1920
tgcagaccat	ggaaccgctg	gagaggcact	tcgccgggga	ccggcccgcac	gtcgtcgtca	1980
acgatccgtc	gtcgtctgtg	acgggacggc	tgctggcgga	ccggtggggc	atcccgggtca	2040
tccgcagcac	tccgacctat	gccgccaacg	aacactggtc	gctgcatccg	ccggtcgcact	2100
cggccgagcc	gccggacgac	cccagactgc	acaagctgct	cgcgcggatc	gagcggctgc	2160
tggaggagca	gggcgtcgag	cacgacctgg	ccggcttcac	cggggtcctg	cacggcggtc	2220
cggccctgct	gtacatgcc	cgctcgttcc	agtacgcggg	cgagaccttc	gacgcacagc	2280
accacttcgt	cggcccctgc	ccgccccgca	ccgcgttcca	cggcgagtgg	aggccggggg	2340
acgacgacgg	ccggcccctg	gtgctgggtga	gtctcggaac	cctgtacaac	gaccggccgg	2400
acttcttccg	cacctgcctg	gaggcgttcg	gcgacgagcc	ctggaacgtg	cttctggtgc	2460
tgggcggcgg	ggtgcccgcg	gccgacctgg	gcccgccttc	cggcaatgtc	cgggtgaccg	2520
acttcgtgtc	gctgcgcgac	gtcctgccgc	acacggcggt	ggtggtgaac	cacggtggga	2580
tgagcaccgc	catggagggtg	ttctcgcacg	gtgtgccggg	ggtggcgatc	ccggtgatgc	2640
cggagccccg	ggccaccgcg	cggcggatcg	tcgaactggg	cctgggcgac	cagctgctga	2700
actcgagact	gacggccgag	tccctgcgtg	ccacgggtacg	gcggtgctg	gcggactccg	2760
cgatccccgg	gaacatgcgc	gggttccggg	agcagatcag	ggcggccggc	ggggcgcccc	2820
cggcgccga	cgcgatcgag	ggactgctgc	cccgggtggg	ctgagacgtc	cgcgcccgcac	2880
acgcgttcac	cttccgaacg	gcgggatcgc	cccatgctca	tcaccgaaac	agcagtgtcc	2940
gacgtgttcc	gcatcgatcc	ggaaccgctg	ccggaccacc	ggggccgggtt	ctacgaagcg	3000
gtgcgccgcg	ggccgctgga	ggccgccgtc	gggcactcgg	tcgagggtccg	gcagggtccac	3060
tgcaccgtct	ccgggcgcaa	cgtactgcgc	ggcctgcacg	ccaccaccct	gccgccgggc	3120
caggccaaga	tcctcacctg	tgtgcggggg	gcggcgctga	ctatgggtgg	cgacatgagg	3180
gtcgggtcac	cgggcttcgg	acggtacgag	gcgggtgcggc	aggatgcccc	gtcgggcacc	3240

gcgctctacc	tgccggacgg	catcggcctg	ggctacgtgg	ccctcgcgga	cgacacatgc	3300
atgaactacc	tgtgcaccga	ggagtacgtc	ccgggcatgg	tcacgcacgt	ggacgccctg	3360
gaccccgaa	tcggcctgcc	gtggggactg	accggggatc	ccgtccgctc	cgcccgggac	3420
gcggcgggcg	cgtccctgcg	ggcggccgcg	gcggcgggaa	ttctgcccac	gtacgaggac	3480
tgcttacggg	tgcgcacgtc	cgtgcccgcc	cctcccgacc	ggactcggct	ctgacgcgac	3540
ggacacgacc	gccgcgcatc	cgacgcgaat	ccgacgcgaa	tccgaactcg	atttcccga	3600
ccgtggacgg	agcgcgtccn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4080
cccagcagcg	tcagcgtccg	cctcgggacc	acaccgggca	ggacgacggg	gtgcggggcg	4140
acgaggctga	gggagccccg	caactccacg	gggcgggtca	accgcttgag	ctgctccatc	4200
aactgctggg	ccgaggtggt	ccccgtgtgg	tggtccaccg	ggccgcggcc	gctcgccgag	4260
gccagccgga	ggttcccgcc	caccgcgacc	atgaagggca	ggccgcgag	accgagccct	4320
cgcaccagtc	gtggcagcac	ggccgcgcgt	gcgtccatca	ccacggggcg	ggccaaggag	4380
cggttggcct	gagcggcttc	ggccaccagg	cggacgatgt	tctcctctc	accgggtgcg	4440
ccgggggtgcc	ggggtgtgcc	ctccccgctc	cgccgcgcgc	ccagcgtcag	gcgccagctg	4500
accggggcg	ccgtcgtgcc	cgaggccagc	cacagcccg	gactctgctg	gcagctcacc	4560
accgggcca	ggtcggagac	gaagcgccgg	ctcaccgccga	cggagcgcac	ccccgccttg	4620

gagaccacca	tcggccagat	cacccaggcg	tcgggggtca	gcctgtcgtc	cacgtagcgg	4680
gccagcgcgg	cccgcacctc	gccccagtcc	caggctcgacc	cggcgacgaa	gtggtgcagg	4740
ctctgttccg	ccgcgccctc	tccgacgaag	gcggcgaggt	tccgggcggg	cttgcgtccc	4800
tgtgccgtga	gcagcccgcg	tatgtactgc	tcacctcttc	tgcgctggtc	cgcccggcgg	4860
agcgaaccga	gtaactccgc	acatgcgtcg	gagacgagcg	agtcgaagtc	gtgccgcgcg	4920
gcgggaccca	cggggggcgc	cgcgccggcg	ggtcgaagcg	tgcgtaaggt	cataggagtc	4980
ctcgtggggg	cctcgtcatc	actgcgagtg	gactgcgacc	agcatcgcca	attcaccctg	5040
tgctccccag	gtacaccggg	cacactcggt	ccgcttcgct	ccccggcggg	cacggcccgc	5100
gtgtggagca	ctccccgctc	tccggacggg	gggggggaag	cccaccgcac	gcgcgccggg	5160
agcgggtgcc	gggccccaca	accctcacc	gacgggggtcc	gttcgaccac	atgccgtgat	5220
gatgccttcg	ccgcacggcc	tccttgaatg	acggaccgtc	agaaagacgt	cactccgcgc	5280
gcgcccgcctc	ccgccccgac	ccgccccggc	cgccgtcccg	cccccgcgac	cctgtggggc	5340
gacgccctct	tgacggaccc	ggaatcccgg	gccgcgggac	gggacggcga	ggaactcgat	5400
gccgcaaagg	ttccgcgtcc	tcaccggcac	ctccgagatc	cagcgcaacg	gcacgcgcaa	5460
gctcctgggt	ttccaccact	gagccgcacc	ccctggcgag	cagatccggg	acaagacgcc	5520
accttcgcga	cagtgtttag	gaaaagttaa	gtaaagaatt	ccgcgagcgg	attgccaggg	5580
agaacaaccc	attgacgcgc	accggtgcag	cggccacatt	gacggcacct	gtcaagttca	5640
cccccaggag	cttggaatcc	catgcaggca	attcggcatc	acgtcatgct	cgtcatggcc	5700
ttcgtgaccg	tcgccacgac	tttccttctc	tggccgtcga	cgcaatccgc	gcaagcgttt	5760
cccccgaccc	cgaagcagac	ggtactgaac	cacctccgcg	ccatttccgg	gaatcacatc	5820
gtctccggac	agcacaacaa	ggagcccgcc	tccgccccgg	gccagtacac	ccagcaggtc	5880
aaggacgtca	ccgggcagta	ccccggcctg	tggggcggtg	acctgatgtt	cgccgcggcg	5940
gacgtggccg	gccgccagcg	cgtcgtcgac	caggccagga	ccgagtgggc	gaacggatcg	6000
ctgggtctcg	tcacctggca	cgtctgcccc	ccgaccggcg	gcagcacctg	tgcgttcgag	6060

ggcggcgctca agtccacgct gacgaacgog cagttctcgc aggtcctcac ggagggcagt 6120

gccctgaaca gcgcatggaa gcggcgccctg gacgaggtcg tcccgtacct gcag 6174

<210> 135

<211> 4770

<212> DNA

<213> *Streptomyces ambofaciens*

<400> 135

gaattccgcg agcggattgc cagggagaac aaccattga cgcgcaccgg tgcagcggcc 60

acattgacgg cacctgtcaa gttcaccccc aggagcttgg aatcccatgc aggcaattcg 120

gcatcacgtc atgctcgtca tggccttcgt gaccgtcgcc acgactttcc ttctctggcc 180

gtcgacgcaa tccgcgcaag cgtttccccc gaccccgaa cagacggtac tgaaccacct 240

ccgcgccatt tccgggaatc acatcgtctc cggacagcac aacaaggagc ccgcctccgc 300

cccgggccag tacaccagc aggtcaagga cgtcacccgg cagtaccccg gcctgtgggg 360

cggtgacctg atgttcgccg cggcggacgt ggccggccgc cagcgcgtcg tcgaccaggc 420

caggaccgag tgggcgaacg gatcgttgtt ctgcgtcacc tggcacgtct gcccgccgac 480

cggcggcagc acctgtgctt tcgagggcgg cgtcaagtcc acgctgacga acgcgcagtt 540

ctcgcaggtc ctcacggagg gcagtgcctt gaacagcgca tgggaagcgg gcctggacga 600

ggtcgtcccc tacctgcagc agctggagaa cgcgggcgtc cccgtcctct tccggccgct 660

gcacgagatg aacgaatcct ggaactggtg gggaaaccgg cccggagcga acggcagcgc 720

acgcctctac cagatcaccg gcgatcacct cgcggggacg aaagggctgg acaatctgat 780

ctgggtcttg aacgtccagg acaatccggc gggaaactgg aacagctact atccgggaga 840

tcagtacgtg gacgtcgttt cgctggacgt ctggtacaag agccaccga gttccgccga 900

ctaccagcag atgcggagca tcgcgggaac aaaacccatg gccctcgcgg agctgggcaa 960

aatgccgacc gccgcgtgct tggacagcca gacgcggtgg acatggttca tgatgtggtc 1020

cgagcatctg cgcgggaaca attccaacgc cgaaatacag acggcgattt tccacccccg 1080

tgtactgaac	caggggggagg	tcgcaactgcc	ctgacgctcg	gcgctgcccg	gctctctcac	1140
gcgcgttctg	acaggacgtc	gcggagagtg	cggggcaagc	ggccggtgag	ctgggcgcag	1200
tcgtcgcgga	cccgatccca	gcgctgttcg	cgtaccgagg	cgaacaggga	ggagaaggcg	1260
tagagccacc	agggatccag	gccggtcgcc	gctgttttcg	cagagtacgt	gtcggccggg	1320
atgtccacgt	aggcgaccgg	tgtccgccac	acctccgccc	ccgtggaggc	gatgcctgcc	1380
aggtccgtcg	actcgggtcc	cgtgatgtcg	tggtagcggc	cggtcggccc	gccaccgcg	1440
agggcggcga	gggcgcgggc	cacgtcgtcg	cgcgccacca	gggacacccg	gccatcggct	1500
gccggcagcc	tcagcagccc	ggtcgagcgc	gcctgggtga	gccagcctag	gaagaactcg	1560
atgtacagcg	aggccctggc	gaacgagcac	ggcacgccgg	aggcgagcag	cagatcctcg	1620
gtgagccggt	tgacgacggc	gtagcagaac	ggggaggccg	agtcggcgtc	gacgctgctc	1680
agtgccgcga	catggccgac	gcgctcggcc	accacggcgg	cgacgacgtt	gcggtggtgc	1740
agcagcacc	gtgcgtcggg	cccgtcgtcg	gagacgagga	ccagagtgtc	cacgcccttc	1800
agcgcgcac	gcagagcggg	cgggtcggcg	tagtcggcga	cggcgccactc	cacccccggc	1860
ggtagtgctt	cggcaggcag	cttcgcctcg	gtcatcgcca	cgacgtccac	gtcggcccgg	1920
tccgcgagca	gctgcaccac	ccgcctgccc	agactgcccg	ccgcccctgt	caccgcgata	1980
cgcacgtcg	ccccgctgc	cttgctcgtc	gtcgtaccac	cgtagggggc	caaccgcgac	2040
cagggcttgg	aacgagccgg	cccgccaggg	cacagacgcg	cgatcgggtc	ggttttcccg	2100
tgctcttttg	gaccgggacg	ccggaccgct	tcctttctac	ggtggagccg	ttcccggccc	2160
agcccgcacg	tcacgacgt	gcggggaaga	cagaggtgat	accgatgctc	cgacgccgtc	2220
tcggagggcc	gtcaggcccc	ctcgtcagtg	ccctgtgcct	gggcgcgatg	cccttcggca	2280
ccaccgtcga	cgagaagacg	tccttcgcca	tcctcgaccg	gttcgtcgag	gccggcggca	2340
gtctcgtcga	caccgccgac	aactacgcgt	tctgggctcc	cggcgggacc	ggggacgaga	2400
gcgagaacac	cgtcggggcg	tggctggcga	gccgccgccc	ccgcgacgag	gtggtgatct	2460

ccaccaaggt	gggtgcccgc	cccaccgtcc	ccggcagcgg	cctggagacc	gccgaagggc	2520
tgtcggctcc	cgtcatacgg	aaggccgcgg	aggacagcct	gcgacgcctg	ggcaccgatc	2580
gcatcgatct	gtactggacc	cacatcgagg	accggaccgt	ccctctggag	gagacgctcg	2640
gagctctcga	cgagctggtc	ggcggcggca	aggtggcggg	gctgggctgc	tccaaccacg	2700
cgggctggcg	catcgaacgg	gcccgcgcgc	tcgcccggac	caacggctgg	acggcgtaca	2760
cctgcgtcca	gcagcgctac	tcctaccttc	agccgcgctt	cgacgtcgga	ctgccggaga	2820
gcggacacgt	ccacgcgacc	tccgaactcc	tcgaccacgt	acgcagcgag	cccgacctga	2880
cgctgctggc	ctactcgtcc	ctgctctcgg	gcgcgtacac	ccggccggac	aagaccctgt	2940
ccgccgccta	cgaccacccg	ggcaccgggc	agcggctgac	cgtgctgcgg	gaggtggccg	3000
ccgagctcgg	tgccaccgcc	aaccagggtg	tgctgtcctg	gttgctcgga	ggtgatccgc	3060
cggtgatccc	gatcgtgggg	gtgagttccg	tcgagcagct	ggacgaggtg	ctggccgccg	3120
tcgagctgga	gctgccccgg	gagacgaggg	cacggctgga	cctcgccggg	cggggctgac	3180
gggccacaca	ccgcccctcg	cccgggacgg	gctgaccggc	cgcacaccgt	ccctcagccg	3240
ggacgggctg	acggccacac	acccgcccct	cagcccggct	gggctgggct	gacggccaca	3300
caccgcgtcc	tcagcccggc	tgggctgggc	tgacggccac	acaccgccc	ctcagcccgg	3360
ctgggctggg	ctgacggcca	cacgccgtcc	ctcgcccggg	tcgggctgac	cagccacacg	3420
ccgtccctca	gtcggccccg	tcccggcccc	gcccggcgct	gacgtgccgc	cggctcgtggc	3480
gccggggcgc	gtcgatggcc	tgtgccaggc	gccgcacggc	tccgggcacg	gcgtggtccg	3540
ccagattgcc	gtagccgagg	accagcgcgg	ggccggcggc	ggccacggcc	cgcgccccgc	3600
cggagtcgcc	ggcccggggc	gtcgggtggg	gctcggcccg	cgccatccgg	taggcgtcga	3660
gatccgccag	cctcacgtca	cgccgcgccg	cctccctgac	gacgggtggc	gccgagcagt	3720
ccgtcagggg	cagcagcagg	tggaagccgg	ccgccgcacc	ggacacctgg	aaccccggca	3780
gcgacgccgc	gagttctccc	atgaggtgat	cgcggcgccg	tttgtagcgc	agccgtgccg	3840
cccgcaggta	ccggtcgtag	gacccgcggg	ccaggaaccg	ggcgaacgcc	tcctggtcga	3900

tgacggggcgg tggtagcccc ccgacgtgct ccgcgcccag ggccctcggtc cagcgcgggcg 3960
 gggtcaccgc ccagccgata cgcagggccg gcgagagcgt cttgctgacc gaccccagca 4020
 gcgcgacgtg ctccggtgcc atgccctgca gggcgcccac ggggtggcgg tcgtatcgga 4080
 actcggcgctc gtagtcgtcc tccagcacca ggccgtcgac ctgcggggcc caggcgacca 4140
 gctcgcccgcg acgggcccgg gcgaggacca caccggtcgg gaactggtgg gcgggagcga 4200
 cgatcaccga ccggacctga ggcgcgcgcc gcagcaggtc gacccgcagg ccgtcgccgt 4260
 ccacgggaac cgggacgggc accagcccct ccgcccgcac cgtggcgctc agccgcgacc 4320
 agccagggtc ctccagggcc atgtgggtgt ggccctcggt ccgcaggacc cggcacagcc 4380
 tgcgcacccg gtccagcgtg cccgcgcaga cgacgagctg ccgggcttcc agcgacgcgc 4440
 cgctccccg cacgaggtag gccgcgagcc gctgccggag ccgggcgaga ccggccggat 4500
 cggggtagcc gagctcgccc caggtcagtg cgccggtggc ctcccgacc gcgtccgccc 4560
 accgtccgcg gggaaaggcc cgcaggtcgg gtatgccggg cagcatgtcg aactccggct 4620
 cccaccgcgt accgacgtcc gcctcgggcg tcaccggagc ggggacggcc tgcgcctga 4680
 cccgggtggc cgagccgctg cgcgcctcca ggtaccctc cgcgacgagc tgctcgtagg 4740
 ggggatcctc tagagtcgac ctgcaggcct 4770

<210> 136
 <211> 28
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo EDR31

<400> 136
 cccaagcttc tgcgcccgcg ggcgtgaa 28

<210> 137
 <211> 28
 <212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR37

<400> 137

gctctagaac cgtgtagccg cgccccgg

28

<210> 138

<211> 30

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo KF30

<400> 138

aagcttgtgt gcccgggtga cctggggagc

30

<210> 139

<211> 30

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo KF31

<400> 139

ggatcccgcg acggacacga ccgccgcgca

30

<210> 140

<211> 12134

<212> DNA

<213> Streptomyces ambofaciens

<400> 140

ctgcagaggg cgccgacgtg cgcggtgagt tcgccgttgc gccggctggt ccagccgagg

60

ccgccgtgga ccgcggggcgc ctgtgcgcac agcaggcccc gggagccgag gtcgtgcagc

120

aggccgaggg gcagctcgcc cgtccggtcc cactcggccg cccggtcgcc gaccagcgcg

180

gtgaacagct cctcggcctc ggtgccgtcg gcgtgggagg tgtcagccac ctgcgtgccc

240

ggtcgcctcg ccgggcgcgg ccagccgtcc gaccagccgg accatcgctg cgacgggtgcg	300
gaagttgtcg agcatcaggt cggcgccgct gatgacgatg ccgtaggtct tctccaggtg	360
cacgacgagc tgcattggcg acagcgagga catcccgccg acggcgaaca ggtcctggtc	420
gcgctcccag gtggtcttgg tgcggtcctc gaggaaccgg agcagttccc cggcgacctc	480
gtcggcgggtg ggggtcgtgc cgggtggggtc gggccgaccg gacgtcgtgg tcatcgctg	540
gcctcctggt agtcgtagaa tccccggccg ctcttgccgg cgagcaggcc ctggcggacc	600
ttgtccagca gcagctcgct cgggcggagc gccggatcgc cggtccttgc gtgcatcacc	660
cgcagcgagt cggccaggtt gtccagtccg atcaggtcgg ccgtggccag gggtcgggtg	720
cgggtggccga tgcagtcgcg catcagcgcg tccacggctc ccgggggtggc ccggccctcg	780
tgcaccaccg cgatggcgtc gttcagcatc cgggtgcagca ggcggctggt cacgaagccg	840
gcgccgtcgc cgacgacgat gccccggcgg cccaggccgg acagcaggtc ccgggtggcc	900
cgggcggccg cctctccgct gcgcgggtccg aggaccacct cgaccgtggg gatcacgtac	960
gcgggggttca tgaagtgcac gccgacgagg tcctcggggc gggggacggc gtcggccagc	1020
tcgtcgatgg ggacgccga ggtgttgctg acgagcagcg tccccgggcg cgccacggac	1080
gccagggtccg ccagcacctc ggccttccgc tcggggctct cggtgacggc ctcgatcacg	1140
gcggtcgcgg tggcgacggc ggccggcgcc tcctcgacgg tcagctcccc gggcgggcgg	1200
tcgtggggca gcgcgccccat cagccgggccc gtccgcagat gcagcgcgac cgcgtcgggg	1260
gcggccgcgc gcgccccggc ggaggtgtcg accagtgtga ccgggtgccg gtgtccgacg	1320
gcgagtgccg cgatggccgt gcccatgacg cccgcgccga gcacgacgag cggagaattt	1380
tccttggaat cgggcacgga tcctccagaa tcccgggggc ggcggctgtg cgaaatgctg	1440
tcccgaatcg cttccgcgct ccatcaccgg ggcgtgactt tgtcacctag caagcgggtg	1500
gaggccgggc ggaggctatg tcatctcggc cttgtccgct tacgagttcc gtccttaaag	1560
tgcgcccggc gaacgcacgg cggaattgac tcgccgtccg tcaaggacct gcggttcgaa	1620
ggacctcgac gcatgaaggg tttcacatgg ctcatatcgc attcttcac cttccggttg	1680

ccgggcatgt gaatccgacc ctgggagtcg ccgaggaact ggtcgcgcgc gccaccggg	1740
tgacgttcgc gctgtccgag gacctcgccg agcgggcccc gctgatcggc gccgaggtgg	1800
tcacctatcc ggtggacagg caacggttcc tggaccagat ggtgccgcgg caggacgcgg	1860
acgagtacac ggacgaggac gagttcgtcc gggtcctgga gtggctgctg gacatgacgg	1920
tgcagaccat ggaaccgctg gagaggcact tcgccgggga ccggcccgcac gtcgtcgtca	1980
acgatccgtc gtcgtgtgg acgggacggc tgctggcgga ccggtggggc atcccgtca	2040
tccgcagcac tccgacctat gccgccaacg aacactggtc gctgcatccg ccggtcgact	2100
cggccgagcc gccggacgac cccgagctgc acaagctgct cgcgcggatc gagcggctgc	2160
tggaggagca gggcgtcgag cacgacctgg ccggttcac cggggtcctg cacggcggtc	2220
cggccctgct gtacatgccc cgtcgttcc agtacgcggg cgagaccttc gacgcacagc	2280
accacttcgt cggccccctgc ccgccccgca ccgcgttcca cggcgagtgg aggccggggg	2340
acgacgacgg ccggccccctg gtgctggtga gtctcggaac cctgtacaac gaccggccgg	2400
acttcttccg cacctgcctg gaggcgttcg gcgacgagcc ctggaacgtg cttctggtgc	2460
tgggcggcgg ggtgcccgcg gccgacctgg gcccgcttcc cggcaatgtc cgggtgaccg	2520
acttcgtgtc gctgcgcgac gtcctgccgc acacggcggg ggtggtgaac cacggtggga	2580
tgagcaccgc catggaggtg ttctcgcacg gtgtgcggg ggtggcgatc ccggtgatgc	2640
cggagccccg gccaccgcg cggcggatcg tcgaactggg cctgggcgac cagctgctga	2700
actcgagct gacggccgag tccctgcgtg ccacggtagc gcgggtgctg gcggactccg	2760
cgatccccgg gaacatgcgc gggttccggg agcagatcag ggcggccggc ggggcgcccg	2820
cggcggccga cgcgatcgag ggactgctgc cccgggtggg ctgagacgtc cgcgccgcac	2880
acgcgttcac cttccgaacg gcgggatgc cccatgctca tcaccgaaac agcagtgccc	2940
gacgtgttcc gcatcgatcc ggaaccgctg ccggaccacc ggggcccgtt ctacgaagcg	3000
gtgcgccgcg ggccgctgga ggccgccgtc gggcactcgg tcgaggtccg gcaggtccac	3060

tgcaccgtct	ccgggcgcaa	cgtactgcgc	ggcctgcacg	ccaccaccct	gccgcccggc	3120
caggccaaga	tcctcacctg	tgtgcggggg	gcggcgctga	ctatgggtgg	cgacatgagg	3180
gtcgggtcac	cgggcttcgg	acggtacgag	gcggtgcggc	aggatgcccc	gtcgggcacc	3240
gcgctctacc	tgccggacgg	catcggcctg	ggctacgtgg	ccctcgcgga	cgacacatgc	3300
atgaactacc	tgtgcaccga	ggagtacgtc	ccgggcatgg	tcacgcacgt	ggacgccctg	3360
gaccccgaac	tcggcctgcc	gtggggactg	accggggatc	ccgtccgctc	cggccgggac	3420
gcggcggcgc	cgtccctgcg	ggcggccgcg	gcggcgggaa	ttctgcccac	gtacgaggac	3480
tgcttacggg	tgcgcacgtc	cgtgcccgcc	cctcccagac	ggactcggct	ctgacgcgac	3540
ggacacgacc	gccgcgcatc	cgacgcgaat	ccgacgcgaa	tccgaactcg	atttcccga	3600
ccgtggacgg	agcgcgccga	ggagcggaaa	ggtctgccgc	agaaacaggc	tgcgccctt	3660
tctttccgct	ttttccggtg	ccgggacggc	cgttccgatg	aaagcgaacg	tccgcatgcg	3720
acggccgtcc	ctcccgtacg	ctcctcccc	gtcgcggccg	gccggggcgg	cgggccggcc	3780
gaccgcgacg	gggcggggcg	gggtcacacc	gccgtcgccc	caccgaggtc	gccgcgcgcg	3840
gagggctgcg	ccagccgcag	cgtgtggggc	atggacgcca	gggtcacgtg	ccggtgccag	3900
ccctggaacg	accgtccttc	gaagtgcgcg	atgcccacgc	ccacgctcac	ccggtcgaag	3960
tcggcgctga	cctgttcggt	gagcatcgcc	agccgcagca	gggcgcgcgc	gtcccaggac	4020
gtgaggtcgg	tcagccacag	gtcggcgggg	cgcggcggtg	tgccccgcca	cacccccagc	4080
agcgtcagcg	tccgcctcgg	gaccacaccc	ggcaggacga	cggggtgcgg	ggcgacgagg	4140
ctgagggagc	cccgcaactc	cacggggcgg	ctcaaccgct	tgagctgctc	catcaactgc	4200
tgggccgagg	tgggtcccg	gtggtgggtc	accgggcgcg	ggccgctcgc	cgaggccagc	4260
cggaggttcc	cggccacccg	caccatgaag	ggcaggcccg	cgagaccgag	ccctcgcacc	4320
agtcgtggca	gcacggccgc	gcgtgcgtcc	atcaccacgg	ggcggggcaa	ggagcggttg	4380
gcctgagcgg	cttcggccac	caggcggacg	atgttctcct	cctcaccggg	tgcgccgggg	4440
tgccgggggtg	tgccctcccc	gctccggccg	ccgcccagcg	tcaggcgcca	gctgaccggg	4500

gcggccgtcg	tgcccagaggc	cagccacagc	ccgtgactct	gctggcagct	caccacccgg	4560
cccaggtcgg	agacgaagcg	ccggctcacc	ccgacggagc	gcacccccgc	cttgagagacc	4620
accatcggcc	agatcaccca	ggcgtcgggg	gtcagcctgt	cgtccacgta	gcggggccagc	4680
gcggccccga	cctcgcccca	gtcccaggtc	gacccggcga	cgaagtgggtg	caggctctgt	4740
tccgccgcgc	cctctccgac	gaaggcggcg	aggttccggg	cggctcttgcg	tccctgtgcc	4800
gtgagcagcc	cgcgtatgta	ctgctcacct	cttctgcgct	ggtccgcccg	gcggagcgaa	4860
ccgagtaact	ccgcacatgc	gtcggagacg	agcgagtoga	agtcgtgccg	cgcggcggga	4920
cccacggggg	gcgccgcgcc	ggcgggtcga	agcgtgcgta	aggtcatagg	agtcctcgtg	4980
ggggcctcgt	catcactgcg	agtggactgc	gaccagcadc	gccaatcac	ccgttgctcc	5040
ccaggtacac	cgggcacact	cgttccgctt	cgtccccccg	gcggcacggc	ccgcgtgtgg	5100
agcactcccc	gctctccgga	cgggtggggg	gaagcccacc	gcacgcgcgc	cggtagcggg	5160
gccggggccc	cacaaccctt	caccgacggg	gtccgttcga	ccacatgccg	tgatgatgcc	5220
ttcgccgcac	ggcctccttg	aatgacggac	cgtcagaaaag	acgtcactcc	gcgcgcgccc	5280
gctcccgccc	cgaccccgcc	cggccgcccgt	cccgcccccg	cgaccctgtg	ggccgacgcc	5340
ctcttgacgg	accggaatc	ccgggcccgc	ggacgggacg	gcgaggaact	cgatgccgca	5400
aaggttccgc	gtcctcaccg	gcacctccga	gatccagcgc	aacggcatcg	ccaagctcct	5460
ggctttccac	cactgagccg	caccccttgg	cgagcagatc	cgggacaaga	cgccaccttc	5520
gcgacagtgt	ttaggaaaaag	ttaagtaaag	aattccgcga	gcggattgcc	agggagaaca	5580
accattgac	gcgcaccggt	gcagcggcca	cattgacggc	acctgtcaag	ttcaccccca	5640
ggagcttgga	atcccatgca	ggcaattcgg	catcacgtca	tgctcgtcat	ggccttcgtg	5700
accgtcgcca	cgactttcct	tctctggcgg	tcgacgcaat	ccgcgcaagc	gtttcccccg	5760
acccgaagc	agacggtact	gaaccacctc	cgcgccatct	ccgggaatca	catcgtctcc	5820
ggacagcaca	acaaggagcc	cgcctccgcc	ccggggccagt	acaccagca	ggtcaaggac	5880

gtcaccgggc agtaccgccg cctgtggggc ggtgacctga tgttcgccgc ggcggacgtg	5940
gccggccgcc agcgcgtcgt cgaccaggcc aggaccgagt gggcgaacgg atcgctggtc	6000
tcgctcacct ggcacgtctg cccgccgacc ggcggcagca cctgtgcgtt cgagggcggc	6060
gtcaagtcca cgctgacgaa cgcgcagttc tcgcaggtec tcacggaggg cagtgccttg	6120
aacagcgcac ggaagcggcg cctggacgag gtcgtcccg acctgcagca gctggagaac	6180
gcggggcgcc ccgtcctctt ccggccgctg caccagatga acgaatcctg gaactggtgg	6240
ggaaaccggc ccggagcgaa cggcagcgca cgcctctacc agatcaccgc cgatcacctc	6300
gccgggacga aagggctgga caatctgac tgggtctgga acgtccagga caatccggcg	6360
ggaaactgga acagctacta tccgggagat cagtacgtgg acgtcgtttc gctggacgtc	6420
tggtacaaga gccaccgcag ttccgccgac taccagcaga tgcggagcat cgcgggaaca	6480
aaacccatgg ccctcgcgga gctgggcaaa atgccgaccg ccgcgctgct ggacagccag	6540
acgcggtgga catggttcat gatgtggtcc gagcatctgc gcgggaacaa ttccaacgcc	6600
gaaatacaga cggcgatattt ccacccccgt gtactgaacc agggggagggt cgcactgccc	6660
tgacgctcgg cgctgcccg ctctctcacg cgcgttctga caggacgtcg cggagagtgc	6720
ggggcaagcg gccggtgagc tgggcgcagt cgtcgcggac ccgatcccag cgctgttcgc	6780
gtaccgaggc gaacaggag gagaaaggct agagccacca gggatccagg ccggtcgccg	6840
ctgtttcggc agagtacgtg tcggccggga tgtccacgta ggcgaccggt gtccgccaca	6900
cctccgccgc cgtggaggcg atgcctgcca ggtccgtcga ctcgggtccc gtgatgtcgt	6960
ggtgacggcc ggtcggcccc cccaccgcga gggcggcgag ggcgcggggc acgtcgtcgc	7020
gcgccaccag ggacaccgcg ccacgcgctg ccggcagcct cagcagcccg gtcgagcgcg	7080
cctgggtgag ccagcctagg aagaactcga tgtacagcga ggccttggcg aacgagcacg	7140
gcacgccgga ggcgagcagc agatcctcgg tgagccgggt gacgacggcg tagcagaacg	7200
gggaggccga gtcggcgtcg acgtgtctca gtgccgcgac atggccgacg cgctcggcca	7260
ccacggcggc gacgacgttg cggtggtgca gcagcaccgc tgcgtcgggc ccgtcgtggtg	7320

agacgaggac	cagagtgtcc	acgcccttca	gcgccgcacg	cagagcgggc	gggtcggcgt	7380
agtcggcgac	ggcgactcc	acccccggcg	gtagtgcctc	ggcaggcagc	ttccgcctgg	7440
tcatcgccac	gacgtccacg	tcggcccggg	ccgcgagcag	ctgcaccacc	cgctgcccc	7500
gactgcccgc	cgccttctg	accgcgatac	gcacgtctgc	ccccgtcgcc	ttgtcgtcgg	7560
tcgtaccacc	gtagggggcc	aaccgcgacc	agggcttgga	acgagccggc	ccgccagggc	7620
acagacgcgc	gatcgggtcc	gttttcccgt	gctcttttgg	accgggacgc	cggaccgctt	7680
cctttctacg	gtggagccgt	tcccgcgccg	gccgcacgt	catcgacgtg	cggggaagac	7740
agaggtgata	ccgatgtctc	gacgccgtct	cggagggccg	tcaggcccc	tcgtcagtgc	7800
cctgtgcctg	ggcgcgatgc	ccttcggcac	caccgtcgac	gagaagacgt	ccttcgccat	7860
cctcgaccgg	ttcgtcgagg	ccggcggcag	tctcgtcgac	accgccgaca	actacgcgtt	7920
ctgggctccc	ggcgggaccg	gggacgagag	cgagaacacc	gtcggggcgt	ggctggcgag	7980
ccgccgccgc	cgcgacgagg	tggtgatctc	caccaagggtg	ggtgcccgcc	ccaccgtccc	8040
cggcagcggc	ctggagaccg	ccgaagggct	gtcggctccc	gtcatacgga	aggccgcgga	8100
ggacagcctg	cgacgcctgg	gcaccgatcg	catcgatctg	tactggaccc	acatcgagga	8160
ccggaccgtc	cctctggagg	agacgctcgg	agctctcgac	gagctggtcg	gcggcggcaa	8220
ggtggcggtg	ctgggctgct	ccaaccacgc	gggctggcgc	atcgaacggg	cccgcgcgct	8280
cgcccggacc	aacggctgga	cggcgtagac	ctgcgtccag	cagcgctact	cctaccttca	8340
gccgcgcttc	gacgtcggac	tgccggagag	cggacacgtc	cacgcgacct	ccgaactcct	8400
cgaccacgta	cgcagcgagc	ccgacctgac	gctgctggcc	tactcgtccc	tgctctcggg	8460
cgcgtacacc	cggccggaca	agaccctgtc	cgcgccttac	gaccaccggg	gcaccgggca	8520
gcggctgacc	gtgctgcggg	aggtggccgc	cagactcggg	gccaccgcca	accaggtggg	8580
gctgtcctgg	ttgctcggag	gtgatccgcc	ggtgatcccg	atcgtggggg	tgagttccgt	8640
cgagcagctg	gacgaggtgc	tggccgccgt	cagactggag	ctgccccggg	agacgagggc	8700

acggctggac	ctcgccgggc	ggggctgacg	ggccacacac	cgcccctcgc	ccgggacggg	8760
ctgaccggcc	gcacaccgtc	cctcagccgg	gacgggctga	cggccacaca	cccggcccctc	8820
agcccggctg	ggctgggctg	acggccacac	acccgctcct	cagcccggct	gggctgggct	8880
gacggccaca	cacccgcccc	tcagcccggc	tgggctgggc	tgacggccac	acgccgtccc	8940
tcgcccgggt	cgggctgacc	agccacacgc	cgccccctcag	tcggcccgcct	ccccggccccg	9000
cccggcgctc	acgtgccgcc	ggtcgtggcg	ccgggcccgcg	tcgatggcct	gtgccaggcg	9060
ccgcacggct	ccgggcacgg	cgtggtccgc	cagattgccg	tagccgagga	ccagcgccggg	9120
gccggcgggc	gccacggccc	gcgccccgcc	ggagtcgccg	gcccggggccg	tcgggtgggtg	9180
ctcgggccgc	gccatccggt	aggcgtcgag	atccgccagc	ctcacgtcac	gccgcgccgc	9240
ctccctgacg	acggtgggcg	ccgagcagtc	cgtcagggac	agcagcaggt	ggaagccggc	9300
cggcgaccg	gacacctgga	accccggcag	cgacgcgcgcg	agttctccca	tgaggtgatc	9360
gcggcgccgt	ttgtagcgca	gccgtgccgc	ccgcaggtac	cggtcgtagg	acccgcgggc	9420
caggaaccgg	gcgaacgcct	cctggtcgat	gacgggcggt	ggtacgcccc	cgacgtgctc	9480
cgcgccccagg	gcctcgggtc	agcgcggcgg	ggtcacccgc	cagccgatcc	gcagggccgg	9540
cgagagcgtc	ttgctgaccg	accccagcag	cgcgacgtgc	tccggtgcc	tgccctgcag	9600
ggcgcccacg	gggtggcggt	cgtatcgga	ctcggcgctc	tagtcgtcct	ccagcaccag	9660
gccgtcgacc	tcgcggggccc	aggcgaccag	ctcgccgcga	cgggcccggg	cgaggaccac	9720
accggtcggg	aactggtggg	cgggagcgac	gatcacccgac	cggacctgag	gcgcgcgccg	9780
cagcaggtcg	acccgcaggc	cgtcgccgtc	cacgggaacc	gggacgggca	ccagcccctc	9840
cgccccgacc	gtggcgctca	gccgcgacca	gccagggctc	tccagggcca	tgtgggtgtg	9900
gccctcggtc	cgcaggaccc	ggcacagcct	gcgcaccgcg	tccagcgtgc	ccgcgcagac	9960
gacgagctgc	cgggcttcca	gcgacgcgcc	gcgtccccgc	acgaggtagg	ccgcgagccg	10020
ctgccggagc	cgggcgagac	cggccggatc	ggggtagccg	agctcgcccc	aggtcagtgc	10080
gccggtggcc	tcccgcaccg	cgtccgcccc	ccgtccgcgg	ggaaaggccc	gcaggtcggg	10140

tatgccgggc	agcatgtcga	actccggctc	ccaccgcgta	ccgacgtccg	cctcggggcgt	10200
caccggagcg	gggacggcct	gcgccctgac	ccgggtggcc	gagccgctgc	gcgcctccag	10260
gtaccctcc	gcgacgagct	gcgcgtaggc	ctccgtcacc	accacccgcg	agcagcccag	10320
gtccgccgcc	agcgccctgc	tcggcggcag	cgcgctgccg	gagggcgatcc	gccacccggt	10380
caccgccgcc	cgcagggcac	ggctcagccg	gacgtgcagc	gggccgtccg	ccggcccgcc	10440
cagttcgaga	aaggttcccc	atgccgggtc	ggtccggctcg	tcgcccacgg	cagccgatgc	10500
taccaccgc	cccgcgacgg	ccgggcggcg	gcgtccacgg	gcccggtgca	caactggatgg	10560
cgtggcggag	acggaccggc	atctcgacca	gcggctggag	caggccatca	tggacctgct	10620
ggagcgacgg	gcggccactg	cctcgatctg	tccgtccgac	gccgcgcgtc	gggtccacga	10680
gggtgacgac	gagggctggc	gagccctgct	ggaaccgcgc	cgcgagcgg	cctggcgtct	10740
ggtgggggcc	ggcgaggtcg	aggtgaccca	ggccggccgg	cccgtcacgc	aggccgaggc	10800
ccgcggggcc	gtccgcattc	gtcgggcggg	ttcctgacgg	ctcgatccga	aggcggaact	10860
cccgtgcccc	ggcggcacgg	gagctcgacg	ccgggaggtc	agcgccgcag	ggtgagcaca	10920
cccggccgcc	acggcagctg	gtcgtagggg	ccgcccgcgg	tgggtgactt	gccctggtag	10980
aggaaccgca	ggttgacggg	gtcgacggtc	atggtctggt	cgggggtgtc	gcggaccagg	11040
tcaccgtggc	tgatgtcgtt	ggtccaggtg	gcccgcgtgt	tggccttgcc	cgcgaagggg	11100
ttgctctcgc	tcgcggcctg	cgggggtccac	gagccgctca	ggctggaggc	cgtgaaggaa	11160
cggaagtagc	gtccgttcgc	gcccacgcgc	tcgacgatca	tcaggtactg	gttctggccc	11220
tgaaccttgt	agacctgcac	gccctcgaac	aggttggcct	tcgtgtcgtc	catgatcgtc	11280
gtgtacgacg	agccgaagct	gcccgggaag	ttcccgatcg	gcatgctcgc	ccggtagatc	11340
ttgccgttgt	caccggcgaa	gaacaggtag	atgttctgtc	cgtcggcgat	cagggctctgg	11400
tcgatcgggc	cgggtgtcga	tcctgagatg	ctcccgggtga	acaacggctg	cggagccgac	11460
cagccgttgg	ggttgggccg	gtcgtcggac	gtgcggtaga	cgaagggcca	cgcgccccac	11520

```

tggtacgcca gcacccagac gttcttgggc gcgaagtaga acaggggtggg cgccaccgcc 11580
gcctgactca tcccgggtctg gccggccgac gccatgtccg accagttcgt gaagggggctg 11640
aacatcatcg agccgtagga cgacccccgac acgttcgacg cgtagaccag gtgcttgccg 11700
ttgtgggtca cgggtggtgaa gtccttgacc gccgcccacc cgttcgccgg ctgcgccagc 11760
acacccgtcg acgaccaccg gtacgccgac ggaagggcac acgccccgtc cgtcgggggc 11820
gtcccgggtca ggccggacca cttctgcccg ctgccaccga cgcacgtccc gatctgcacc 11880
gccgtgccgt tggccgtacc ggcgcccgcg gcctccaggc acagcccgga ctccacgccg 11940
acgaccgtgc cgtcggagtt caccgccac tgctggttcg caccgccgga acagctccag 12000
atctgcaccc gcgccccggg tgtggtggcg tgacccggaa cgtccaggca cttgttgccg 12060
tacacggtca gccgacggtc gtccgtcaac gtccactgct gattgctccc gcccagcag 12120
tcgtatatct gcag 12134

```

```

<210> 141
<211> 1164
<212> DNA
<213> Streptomyces ambofaciens

```

```

<220>
<221> CDS
<222> (1)..(1164)

```

```

<400> 141
atg acc tta cgc acg ctt cga ccc gcc ggc gcg gcg ccc ccc gtg ggt 48
Met Thr Leu Arg Thr Leu Arg Pro Ala Gly Ala Ala Pro Pro Val Gly
1          5          10          15

ccc gcc gcg cgg cac gac ttc gac tcg ctc gtc tcc gac gca tgt gcg 96
Pro Ala Ala Arg His Asp Phe Asp Ser Leu Val Ser Asp Ala Cys Ala
          20          25          30

gag tta ctc ggt tcg ctc cgc cgg gcg gac cag cgc aga aga ggt gag 144
Glu Leu Leu Gly Ser Leu Arg Arg Ala Asp Gln Arg Arg Arg Gly Glu
          35          40          45

cag tac ata cgc ggg ctg ctc acg gca cag gga cgc aag acc gcc cgg 192

```

Gln Tyr Ile Arg Gly Leu Leu Thr Ala Gln Gly Arg Lys Thr Ala Arg	
50 55 60	
aac ctc gcc gcc ttc gtc gga gag ggc gcg gcg gaa cag agc ctg cac	240
Asn Leu Ala Ala Phe Val Gly Glu Gly Ala Ala Glu Gln Ser Leu His	
65 70 75 80	
cac ttc gtc gcc ggg tcg acc tgg gac tgg ggc gag gtg cgg gcc gcg	288
His Phe Val Ala Gly Ser Thr Trp Asp Trp Gly Glu Val Arg Ala Ala	
85 90 95	
ctg gcc cgc tac gtg gac gac agg ctg acc ccc gac gcc tgg gtg atc	336
Leu Ala Arg Tyr Val Asp Asp Arg Leu Thr Pro Asp Ala Trp Val Ile	
100 105 110	
tgg ccg atg gtg gtc tcc aag gcg ggg gtg cgc tcc gtc ggg gtg agc	384
Trp Pro Met Val Val Ser Lys Ala Gly Val Arg Ser Val Gly Val Ser	
115 120 125	
cgg cgc ttc gtc tcc gac ctg ggc cgg gtg gtg agc tgc cag cag agt	432
Arg Arg Phe Val Ser Asp Leu Gly Arg Val Val Ser Cys Gln Gln Ser	
130 135 140	
cac ggg ctg tgg ctg gcc tcg ggc acg acg gcc gcc ccg gtc agc tgg	480
His Gly Leu Trp Leu Ala Ser Gly Thr Thr Ala Ala Pro Val Ser Trp	
145 150 155 160	
cgc ctg acg ctg ggc ggc ggc cgg agc ggg gag ggc aca ccc cgg cac	528
Arg Leu Thr Leu Gly Gly Gly Arg Ser Gly Glu Gly Thr Pro Arg His	
165 170 175	
ccc ggc gca ccc ggt gag gag gag aac atc gtc cgc ctg gtg gcc gaa	576
Pro Gly Ala Pro Gly Glu Glu Glu Asn Ile Val Arg Leu Val Ala Glu	
180 185 190	
gcc gct cag gcc aac cgc tcc ttg gcc cgc ccc gtg gtg atg gac gca	624
Ala Ala Gln Ala Asn Arg Ser Leu Ala Arg Pro Val Val Met Asp Ala	
195 200 205	
cgc gcg gcc gtg ctg cca cga ctg gtg cga ggg ctc ggt ctc gcg ggc	672
Arg Ala Ala Val Leu Pro Arg Leu Val Arg Gly Leu Gly Leu Ala Gly	
210 215 220	
ctg ccc ttc atg gtg cgg gtg ggc ggg aac ctc cgg ctg gcc tcg gcg	720
Leu Pro Phe Met Val Arg Val Gly Gly Asn Leu Arg Leu Ala Ser Ala	
225 230 235 240	

agc ggc cgc ggc ccg gtg gac cac cac acg ggg acc acc tcg gcc cag	768
Ser Gly Arg Gly Pro Val Asp His His Thr Gly Thr Thr Ser Ala Gln	
245 250 255	
cag ttg atg gag cag ctc aag cgg ttg agc cgc ccc gtg gag ttg cgg	816
Gln Leu Met Glu Gln Leu Lys Arg Leu Ser Arg Pro Val Glu Leu Arg	
260 265 270	
ggc tcc ctc agc ctc gtc gcc ccg cac ccc gtc gtc ctg ccg ggt gtg	864
Gly Ser Leu Ser Leu Val Ala Pro His Pro Val Val Leu Pro Gly Val	
275 280 285	
gtc ccg agg cgg acg ctg acg ctg ctg ggg gtg tgg cgg ggc aac cgc	912
Val Pro Arg Arg Thr Leu Thr Leu Leu Gly Val Trp Arg Gly Asn Arg	
290 295 300	
cgg cgc ccc gcc gac ctg tgg ctg acc gac ctc acg tcc tgg gac cgc	960
Arg Arg Pro Ala Asp Leu Trp Leu Thr Asp Leu Thr Ser Trp Asp Arg	
305 310 315 320	
ggc gcc ctg ctg cgg ctg gcg atg ctc acc gaa cag gtc gac gcc gac	1008
Gly Ala Leu Leu Arg Leu Ala Met Leu Thr Glu Gln Val Asp Ala Asp	
325 330 335	
ttc gac cgg gtg agc gtg ggc gtg ggc atg cgc gac ttc gaa gga cgg	1056
Phe Asp Arg Val Ser Val Gly Val Gly Met Arg Asp Phe Glu Gly Arg	
340 345 350	
tcg ttc cag ggc tgg cac cgg cac gtg acc ctg gcg tcc atc gcc cac	1104
Ser Phe Gln Gly Trp His Arg His Val Thr Leu Ala Ser Ile Ala His	
355 360 365	
acg ctg cgg ctg gcg cag ccc tcc gcg cgc ggc gac ctc ggt ggg gcg	1152
Thr Leu Arg Leu Ala Gln Pro Ser Ala Arg Gly Asp Leu Gly Gly Ala	
370 375 380	
acg gcg gtg tga	1164
Thr Ala Val	
385	

<210> 142

<211> 387

<212> PRT

<213> Streptomyces ambofaciens

<400> 142

Met Thr Leu Arg Thr Leu Arg Pro Ala Gly Ala Ala Pro Pro Val Gly
 1 5 10 15

Pro Ala Ala Arg His Asp Phe Asp Ser Leu Val Ser Asp Ala Cys Ala
 20 25 30

Glu Leu Leu Gly Ser Leu Arg Arg Ala Asp Gln Arg Arg Arg Gly Glu
 35 40 45

Gln Tyr Ile Arg Gly Leu Leu Thr Ala Gln Gly Arg Lys Thr Ala Arg
 50 55 60

Asn Leu Ala Ala Phe Val Gly Glu Gly Ala Ala Glu Gln Ser Leu His
 65 70 75 80

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

Leu Ala Arg Tyr Val Asp Asp Arg Leu Thr Pro Asp Ala Trp Val Ile
 100 105 110

Trp Pro Met Val Val Ser Lys Ala Gly Val Arg Ser Val Gly Val Ser
 115 120 125

Arg Arg Phe Val Ser Asp Leu Gly Arg Val Val Ser Cys Gln Gln Ser
 130 135 140

His Gly Leu Trp Leu Ala Ser Gly Thr Thr Ala Ala Pro Val Ser Trp
 145 150 155 160

Arg Leu Thr Leu Gly Gly Gly Arg Ser Gly Glu Gly Thr Pro Arg His
 165 170 175

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

Ala Ala Gln Ala Asn Arg Ser Leu Ala Arg Pro Val Val Met Asp Ala
 195 200 205

Arg Ala Ala Val Leu Pro Arg Leu Val Arg Gly Leu Gly Leu Ala Gly
 210 215 220

Leu Pro Phe Met Val Arg Val Gly Gly Asn Leu Arg Leu Ala Ser Ala
 225 230 235 240

Ser Gly Arg Gly Pro Val Asp His His Thr Gly Thr Thr Ser Ala Gln
 245 250 255

Gln Leu Met Glu Gln Leu Lys Arg Leu Ser Arg Pro Val Glu Leu Arg
 260 265 270

Gly Ser Leu Ser Leu Val Ala Pro His Pro Val Val Leu Pro Gly Val
 275 280 285

Val Pro Arg Arg Thr Leu Thr Leu Leu Gly Val Trp Arg Gly Asn Arg
 290 295 300

Arg Arg Pro Ala Asp Leu Trp Leu Thr Asp Leu Thr Ser Trp Asp Arg
 305 310 315 320

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

Phe Asp Arg Val Ser Val Gly Val Gly Met Arg Asp Phe Glu Gly Arg
 340 345 350

Ser Phe Gln Gly Trp His Arg His Val Thr Leu Ala Ser Ile Ala His
 355 360 365

Thr Leu Arg Leu Ala Gln Pro Ser Ala Arg Gly Asp Leu Gly Gly Ala

370

375

380

Thr Ala Val

385

<210> 143

<211> 849

<212> DNA

<213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1) .. (849)

<400> 143

atg	cgt	atc	gcg	gtg	aca	ggg	gcg	gcg	ggc	agt	ctg	ggc	agg	cgg	gtg	48
Met	Arg	Ile	Ala	Val	Thr	Gly	Ala	Ala	Gly	Ser	Leu	Gly	Arg	Arg	Val	
1				5					10					15		

gtg	cag	ctg	ctc	gcg	gac	cgg	gcc	gac	gtg	gac	gtc	gtg	gcg	atg	acc	96
Val	Gln	Leu	Leu	Ala	Asp	Arg	Ala	Asp	Val	Asp	Val	Val	Ala	Met	Thr	
			20					25					30			

agg	cgg	aag	ctg	cct	gcc	gag	gca	cta	ccg	ccg	ggg	gtg	gag	tgc	gcc	144
Arg	Arg	Lys	Leu	Pro	Ala	Glu	Ala	Leu	Pro	Pro	Gly	Val	Glu	Cys	Ala	
		35					40					45				

gtc	gcc	gac	tac	gcc	gac	ccg	ccc	gct	ctg	cgt	gcg	gcg	ctg	aag	ggc	192
Val	Ala	Asp	Tyr	Ala	Asp	Pro	Pro	Ala	Leu	Arg	Ala	Ala	Leu	Lys	Gly	
	50					55				60						

gtg	gac	act	ctg	gtc	ctc	gtc	tcc	agc	gac	ggg	ccc	gac	gca	cgg	gtg	240
Val	Asp	Thr	Leu	Val	Leu	Val	Ser	Ser	Asp	Gly	Pro	Asp	Ala	Arg	Val	
65					70					75				80		

ctg	ctg	cac	cac	cgc	aac	gtc	gtc	gcc	gcc	gtg	gtg	gcc	gag	cgc	gtc	288
Leu	Leu	His	His	Arg	Asn	Val	Val	Ala	Ala	Val	Val	Ala	Glu	Arg	Val	
				85				90					95			

ggc	cat	gtc	gcg	gca	ctg	agc	agc	gtc	gac	gcc	gac	tcg	gcc	tcc	ccg	336
Gly	His	Val	Ala	Ala	Leu	Ser	Ser	Val	Asp	Ala	Asp	Ser	Ala	Ser	Pro	
			100					105					110			

ttc	tgc	tac	gcc	gtc	gtc	aac	cgg	ctc	acc	gag	gat	ctg	ctg	ctc	gcc	384

Phe Cys Tyr Ala Val Val Asn Arg Leu Thr Glu Asp Leu Leu Leu Ala	
115 120 125	
tcc ggc gtg ccg tgc tcg ttc gcc agg gcc tcg ctg tac atc gag ttc	432
Ser Gly Val Pro Cys Ser Phe Ala Arg Ala Ser Leu Tyr Ile Glu Phe	
130 135 140	
ttc cta ggc tgg ctc acc cag gcg cgc tcg acc ggg ctg ctg agg ctg	480
Phe Leu Gly Trp Leu Thr Gln Ala Arg Ser Thr Gly Leu Leu Arg Leu	
145 150 155 160	
ccg gca gcc gat ggc cgg gtg tcc ctg gtg gcg cgc gac gac gtg gcc	528
Pro Ala Ala Asp Gly Arg Val Ser Leu Val Ala Arg Asp Asp Val Ala	
165 170 175	
cgc gcc ctc gcc gcc ctc gcg gtg ggc ggg ccg acc ggc cgt cac cac	576
Arg Ala Leu Ala Ala Leu Ala Val Gly Gly Pro Thr Gly Arg His His	
180 185 190	
gac atc acg gga ccc gag tcg acg gac ctg gca ggc atc gcc tcc acg	624
Asp Ile Thr Gly Pro Glu Ser Thr Asp Leu Ala Gly Ile Ala Ser Thr	
195 200 205	
gcg gcg gag gtg tgg cgg aca ccg gtc gcc tac gtg gac atc ccg gcc	672
Ala Ala Glu Val Trp Arg Thr Pro Val Ala Tyr Val Asp Ile Pro Ala	
210 215 220	
gac acg tac tct gcc gaa aca gcg gcg acc ggc ctg gat ccc tgg tgg	720
Asp Thr Tyr Ser Ala Glu Thr Ala Ala Thr Gly Leu Asp Pro Trp Trp	
225 230 235 240	
ctc tac gcc ttc tcc tcc ctg ttc gcc tcg gta cgc gaa cag cgc tgg	768
Leu Tyr Ala Phe Ser Ser Leu Phe Ala Ser Val Arg Glu Gln Arg Trp	
245 250 255	
gat cgg gtc cgc gac gac tgc gcc cag ctc acc ggc cgc ttg ccc cgc	816
Asp Arg Val Arg Asp Asp Cys Ala Gln Leu Thr Gly Arg Leu Pro Arg	
260 265 270	
act ctc cgc gac gtc ctg tca gaa cgc gcg tga	849
Thr Leu Arg Asp Val Leu Ser Glu Arg Ala	
275 280	
<210> 144	
<211> 282	
<212> PRT	

<213> Streptomyces ambofaciens

<400> 144

Met Arg Ile Ala Val Thr Gly Ala Ala Gly Ser Leu Gly Arg Arg Val
1 5 10 15

Val Gln Leu Leu Ala Asp Arg Ala Asp Val Asp Val Val Ala Met Thr
20 25 30

Arg Arg Lys Leu Pro Ala Glu Ala Leu Pro Pro Gly Val Glu Cys Ala
35 40 45

Val Ala Asp Tyr Ala Asp Pro Pro Ala Leu Arg Ala Ala Leu Lys Gly
50 55 60

Val Asp Thr Leu Val Leu Val Ser Ser Asp Gly Pro Asp Ala Arg Val
65 70 75 80

Leu Leu His His Arg Asn Val Val Ala Ala Val Val Ala Glu Arg Val
85 90 95

Gly His Val Ala Ala Leu Ser Ser Val Asp Ala Asp Ser Ala Ser Pro
100 105 110

Phe Cys Tyr Ala Val Val Asn Arg Leu Thr Glu Asp Leu Leu Leu Ala
115 120 125

Ser Gly Val Pro Cys Ser Phe Ala Arg Ala Ser Leu Tyr Ile Glu Phe
130 135 140

Phe Leu Gly Trp Leu Thr Gln Ala Arg Ser Thr Gly Leu Leu Arg Leu
145 150 155 160

Pro Ala Ala Asp Gly Arg Val Ser Leu Val Ala Arg Asp Asp Val Ala
165 170 175

Arg Ala Leu Ala Ala Leu Ala Val Gly Gly Pro Thr Gly Arg His His
 180 185 190

Asp Ile Thr Gly Pro Glu Ser Thr Asp Leu Ala Gly Ile Ala Ser Thr
 195 200 205

Ala Ala Glu Val Trp Arg Thr Pro Val Ala Tyr Val Asp Ile Pro Ala
 210 215 220

Asp Thr Tyr Ser Ala Glu Thr Ala Ala Thr Gly Leu Asp Pro Trp Trp
 225 230 235 240

Leu Tyr Ala Phe Ser Ser Leu Phe Ala Ser Val Arg Glu Gln Arg Trp
 245 250 255

Asp Arg Val Arg Asp Asp Cys Ala Gln Leu Thr Gly Arg Leu Pro Arg
 260 265 270

Thr Leu Arg Asp Val Leu Ser Glu Arg Ala
 275 280

<210> 145
 <211> 1512
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(1512)

<400> 145
 gtg ggc gac gac cgg acc gac ccg gca tgg gga acc ttt ctc gaa ctg 48
 Val Gly Asp Asp Arg Thr Asp Pro Ala Trp Gly Thr Phe Leu Glu Leu
 1 5 10 15

ggc ggg ccg gcg gac ggc ccg ctg cac gtc cgg ctg agc cgt gcc ctg 96
 Gly Gly Pro Ala Asp Gly Pro Leu His Val Arg Leu Ser Arg Ala Leu
 20 25 30

cgg gcg gcg gtg acc ggt ggg cgg atc gcc tcc ggc agc gcg ctg ccg Arg Ala Ala Val Thr Gly Gly Arg Ile Ala Ser Gly Ser Ala Leu Pro 35 40 45	144
ccg agc agg gcg ctg gcg gcg gac ctg ggc tgc tcg cgg tgg gtg gtg Pro Ser Arg Ala Leu Ala Ala Asp Leu Gly Cys Ser Arg Trp Val Val 50 55 60	192
acg gag gcc tac gcg cag ctc gtc gcg gag ggg tac ctg gag gcg cgc Thr Glu Ala Tyr Ala Gln Leu Val Ala Glu Gly Tyr Leu Glu Ala Arg 65 70 75 80	240
agc ggc tcg gcc acc cgg gtc agg gcg cag gcc gtc ccc gct ccg gtg Ser Gly Ser Ala Thr Arg Val Arg Ala Gln Ala Val Pro Ala Pro Val 85 90 95	288
acg ccc gag gcg gac gtc ggt acg cgg tgg gag ccg gag ttc gac atg Thr Pro Glu Ala Asp Val Gly Thr Arg Trp Glu Pro Glu Phe Asp Met 100 105 110	336
ctg ccc ggc ata ccc gac ctg cgg gcc ttt ccc cgc gga cgg tgg gcg Leu Pro Gly Ile Pro Asp Leu Arg Ala Phe Pro Arg Gly Arg Trp Ala 115 120 125	384
gac gcg gtg cgg gag gcc acc ggc gca ctg acc tgg ggc gag ctc ggc Asp Ala Val Arg Glu Ala Thr Gly Ala Leu Thr Trp Gly Glu Leu Gly 130 135 140	432
tac ccc gat ccg gcc ggt ctc gcc cgg ctc cgg cag cgg ctc gcg gcc Tyr Pro Asp Pro Ala Gly Leu Ala Arg Leu Arg Gln Arg Leu Ala Ala 145 150 155 160	480
tac ctc gtg cgg gga cgc ggc gcg tcg ctg gaa gcc cgg cag ctc gtc Tyr Leu Val Arg Gly Arg Gly Ala Ser Leu Glu Ala Arg Gln Leu Val 165 170 175	528
gtc tgc gcg ggc acg ctg gac gcg gtg cgc agg ctg tgc cgg gtc ctg Val Cys Ala Gly Thr Leu Asp Ala Val Arg Arg Leu Cys Arg Val Leu 180 185 190	576
cgg acc gag ggc cac acc cac atg gcc ctg gag gac cct ggc tgg tcg Arg Thr Glu Gly His Thr His Met Ala Leu Glu Asp Pro Gly Trp Ser 195 200 205	624
cgg ctg agc gcc acg gtc ggg gcg gag ggg ctg gtg ccc gtc ccg gtt Arg Leu Ser Ala Thr Val Gly Ala Glu Gly Leu Val Pro Val Pro Val 	672

210	215	220	
ccc gtg gac ggc gac ggc ctg cgg gtc gac ctg ctg cgg cgc gcg cct			720
Pro Val Asp Gly Asp Gly Leu Arg Val Asp Leu Leu Arg Arg Ala Pro			
225	230	235	240
cag gtc cgg tcg gtg atc gtc gct ccc gcc cac cag ttc ccg acc ggt			768
Gln Val Arg Ser Val Ile Val Ala Pro Ala His Gln Phe Pro Thr Gly			
	245	250	255
gtg gtc ctc gcc ccg gcc cgt cgc ggc gag ctg gtc gcc tgg gcc cgc			816
Val Val Leu Ala Pro Ala Arg Arg Gly Glu Leu Val Ala Trp Ala Arg			
	260	265	270
gag gtc gac ggc ctg gtg ctg gag gac gac tac gac gcc gag ttc cga			864
Glu Val Asp Gly Leu Val Leu Glu Asp Asp Tyr Asp Ala Glu Phe Arg			
	275	280	285
tac gac cgc cac ccc gtg ggc gcc ctg cag ggc atg gca ccg gag cac			912
Tyr Asp Arg His Pro Val Gly Ala Leu Gln Gly Met Ala Pro Glu His			
	290	295	300
gtc gcg ctg ctg ggg tcg gtc agc aag acg ctc tcg ccg gcc ctg cgg			960
Val Ala Leu Leu Gly Ser Val Ser Lys Thr Leu Ser Pro Ala Leu Arg			
305	310	315	320
atc ggc tgg gcg gtg acc ccg ccg cgc tgg acc gag gcc ctg ggc gcg			1008
Ile Gly Trp Ala Val Thr Pro Pro Arg Trp Thr Glu Ala Leu Gly Ala			
	325	330	335
gag cac gtc ggg ggc gta cca ccg ccc gtc atc gac cag gag gcg ttc			1056
Glu His Val Gly Gly Val Pro Pro Pro Val Ile Asp Gln Glu Ala Phe			
	340	345	350
gcc cgg ttc ctg gcc cgc ggg tcc tac gac cgg tac ctg cgg gcg gca			1104
Ala Arg Phe Leu Ala Arg Gly Ser Tyr Asp Arg Tyr Leu Arg Ala Ala			
	355	360	365
cgg ctg cgc tac aaa cgg cgc cgc gat cac ctc atg gga gaa ctc gcg			1152
Arg Leu Arg Tyr Lys Arg Arg Arg Asp His Leu Met Gly Glu Leu Ala			
	370	375	380
gcg tcg ctg ccg ggg ttc cag gtg tcc ggt gcg gcg gcc ggc ttc cac			1200
Ala Ser Leu Pro Gly Phe Gln Val Ser Gly Ala Ala Ala Gly Phe His			
385	390	395	400
ctg ctg ctg tcc ctg acg gac tgc tcg gcg ccc acc gtc gtc agg gag			1248

Leu Leu Leu Ser Leu Thr Asp Cys Ser Ala Pro Thr Val Val Arg Glu
 405 410 415
 gcg gcg cgg cgt gac gtg agg ctg gcg gat ctc gac gcc tac cgg atg 1296
 Ala Ala Arg Arg Asp Val Arg Leu Ala Asp Leu Asp Ala Tyr Arg Met
 420 425 430
 gcg cgg gcc gag cac cac ccg acg gcc cgg gcc ggc gac tcc ggc ggg 1344
 Ala Arg Ala Glu His His Pro Thr Ala Arg Ala Gly Asp Ser Gly Gly
 435 440 445
 gcg cgg gcc gtg gcc gcc gcc ggc ccc gcg ctg gtc ctc ggc tac ggc 1392
 Ala Arg Ala Val Ala Ala Ala Gly Pro Ala Leu Val Leu Gly Tyr Gly
 450 455 460
 aat ctg gcg gac cac gcc gtg ccc gga gcc gtg cgg cgc ctg gca cag 1440
 Asn Leu Ala Asp His Ala Val Pro Gly Ala Val Arg Arg Leu Ala Gln
 465 470 475 480
 gcc atc gac gcg gcc cgg cgc cac gac cgg cgg cac gtc gac gcc ggg 1488
 Ala Ile Asp Ala Ala Arg Arg His Asp Arg Arg His Val Asp Ala Gly
 485 490 495
 cgg gcc ggg gag cgg gcc gac tga 1512
 Arg Ala Gly Glu Arg Ala Asp
 500
 <210> 146
 <211> 503
 <212> PRT
 <213> Streptomyces ambofaciens
 <400> 146
 Val Gly Asp Asp Arg Thr Asp Pro Ala Trp Gly Thr Phe Leu Glu Leu
 1 5 10 15
 Gly Gly Pro Ala Asp Gly Pro Leu His Val Arg Leu Ser Arg Ala Leu
 20 25 30
 Arg Ala Ala Val Thr Gly Gly Arg Ile Ala Ser Gly Ser Ala Leu Pro
 35 40 45

Pro Ser Arg Ala Leu Ala Ala Asp Leu Gly Cys Ser Arg Trp Val Val
 50 55 60

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

Ser Gly Ser Ala Thr Arg Val Arg Ala Gln Ala Val Pro Ala Pro Val
 85 90 95

Thr Pro Glu Ala Asp Val Gly Thr Arg Trp Glu Pro Glu Phe Asp Met
 100 105 110

Leu Pro Gly Ile Pro Asp Leu Arg Ala Phe Pro Arg Gly Arg Trp Ala
 115 120 125

Asp Ala Val Arg Glu Ala Thr Gly Ala Leu Thr Trp Gly Glu Leu Gly
 130 135 140

Tyr Pro Asp Pro Ala Gly Leu Ala Arg Leu Arg Gln Arg Leu Ala Ala
 145 150 155 160

Tyr Leu Val Arg Gly Arg Gly Ala Ser Leu Glu Ala Arg Gln Leu Val
 165 170 175

Val Cys Ala Gly Thr Leu Asp Ala Val Arg Arg Leu Cys Arg Val Leu
 180 185 190

Arg Thr Glu Gly His Thr His Met Ala Leu Glu Asp Pro Gly Trp Ser
 195 200 205

Arg Leu Ser Ala Thr Val Gly Ala Glu Gly Leu Val Pro Val Pro Val
 210 215 220

Pro Val Asp Gly Asp Gly Leu Arg Val Asp Leu Leu Arg Arg Ala Pro
 225 230 235 240

Gln Val Arg Ser Val Ile Val Ala Pro Ala His Gln Phe Pro Thr Gly
 245 250 255

Val Val Leu Ala Pro Ala Arg Arg Gly Glu Leu Val Ala Trp Ala Arg
 260 265 270

Glu Val Asp Gly Leu Val Leu Glu Asp Asp Tyr Asp Ala Glu Phe Arg
 275 280 285

Tyr Asp Arg His Pro Val Gly Ala Leu Gln Gly Met Ala Pro Glu His
 290 295 300

Val Ala Leu Leu Gly Ser Val Ser Lys Thr Leu Ser Pro Ala Leu Arg
 305 310 315 320

Ile Gly Trp Ala Val Thr Pro Pro Arg Trp Thr Glu Ala Leu Gly Ala
 325 330 335

Glu His Val Gly Gly Val Pro Pro Pro Val Ile Asp Gln Glu Ala Phe
 340 345 350

Ala Arg Phe Leu Ala Arg Gly Ser Tyr Asp Arg Tyr Leu Arg Ala Ala
 355 360 365

Arg Leu Arg Tyr Lys Arg Arg Arg Asp His Leu Met Gly Glu Leu Ala
 370 375 380

Ala Ser Leu Pro Gly Phe Gln Val Ser Gly Ala Ala Ala Gly Phe His
 385 390 395 400

Leu Leu Leu Ser Leu Thr Asp Cys Ser Ala Pro Thr Val Val Arg Glu
 405 410 415

Ala Ala Arg Arg Asp Val Arg Leu Ala Asp Leu Asp Ala Tyr Arg Met
 420 425 430

Ala Arg Ala Glu His His Pro Thr Ala Arg Ala Gly Asp Ser Gly Gly
 435 440 445

Ala Arg Ala Val Ala Ala Ala Gly Pro Ala Leu Val Leu Gly Tyr Gly
 450 455 460

Asn Leu Ala Asp His Ala Val Pro Gly Ala Val Arg Arg Leu Ala Gln
 465 470 475 480

Ala Ile Asp Ala Ala Arg Arg His Asp Arg Arg His Val Asp Ala Gly
 485 490 495

Arg Ala Gly Glu Arg Ala Asp
 500

<210> 147
 <211> 276
 <212> DNA
 <213> Streptomyces ambofaciens

<220>
 <221> CDS
 <222> (1)..(276)

<400> 147
 gtg gcg gag acg gac cgg cat ctc gac cag cgg ctg gag cag gcc atc 48
 Val Ala Glu Thr Asp Arg His Leu Asp Gln Arg Leu Glu Gln Ala Ile
 1 5 10 15
 atg gac ctg ctg gag cga cgg gcg gcc act gcc tcg atc tgt ccg tcc 96
 Met Asp Leu Leu Glu Arg Arg Ala Ala Thr Ala Ser Ile Cys Pro Ser
 20 25 30
 gac gcc gcg cgt cgg gtc cac gag ggt gac gac gag ggc tgg cga gcc 144
 Asp Ala Ala Arg Arg Val His Glu Gly Asp Asp Glu Gly Trp Arg Ala
 35 40 45
 ctg ctg gaa ccc gcg cgc cga gcg gcc tgg cgt ctg gtg ggg tcc ggc 192
 Leu Leu Glu Pro Ala Arg Arg Ala Ala Trp Arg Leu Val Gly Ser Gly

50 55 60

gag gtc gag gtg acc cag gcc ggc cgg ccc gtc acg cag gcc gag gcc 240
 Glu Val Glu Val Thr Gln Ala Gly Arg Pro Val Thr Gln Ala Glu Ala
 65 70 75 80

cgc ggg ccc gtc cgc att cgt cgg gcg ggt tcc tga 276
 Arg Gly Pro Val Arg Ile Arg Arg Ala Gly Ser
 85 90

<210> 148
 <211> 91
 <212> PRT
 <213> Streptomyces ambofaciens

<400> 148

Val Ala Glu Thr Asp Arg His Leu Asp Gln Arg Leu Glu Gln Ala Ile
 1 5 10 15

Met Asp Leu Leu Glu Arg Arg Ala Ala Thr Ala Ser Ile Cys Pro Ser
 20 25 30

Asp Ala Ala Arg Arg Val His Glu Gly Asp Asp Glu Gly Trp Arg Ala
 35 40 45

Leu Leu Glu Pro Ala Arg Arg Ala Ala Trp Arg Leu Val Gly Ser Gly
 50 55 60

Glu Val Glu Val Thr Gln Ala Gly Arg Pro Val Thr Gln Ala Glu Ala
 65 70 75 80

Arg Gly Pro Val Arg Ile Arg Arg Ala Gly Ser
 85 90

<210> 149
 <211> 1236
 <212> DNA
 <213> Streptomyces ambofaciens

<220>

<221> CDS

<222> (1)..(1236)

<400> 149

ctg	cag	ata	tac	gac	tgc	tgg	ggc	ggg	agc	aat	cag	cag	tgg	acg	ttg	48
Leu	Gln	Ile	Tyr	Asp	Cys	Trp	Gly	Gly	Ser	Asn	Gln	Gln	Trp	Thr	Leu	
1				5					10					15		

acg	gac	gac	cgt	cgg	ctg	acc	gtg	tac	ggc	aac	aag	tgc	ctg	gac	gtt	96
Thr	Asp	Asp	Arg	Arg	Leu	Thr	Val	Tyr	Gly	Asn	Lys	Cys	Leu	Asp	Val	
			20					25					30			

ccg	ggt	cac	gcc	acc	aca	ccc	ggg	acg	cgg	gtg	cag	atc	tgg	agc	tgt	144
Pro	Gly	His	Ala	Thr	Thr	Pro	Gly	Thr	Arg	Val	Gln	Ile	Trp	Ser	Cys	
		35					40					45				

tcc	ggc	ggt	gcg	aac	cag	cag	tgg	cgg	gtg	aac	tcc	gac	ggc	acg	gtc	192
Ser	Gly	Gly	Ala	Asn	Gln	Gln	Trp	Arg	Val	Asn	Ser	Asp	Gly	Thr	Val	
	50					55				60						

gtc	ggc	gtg	gag	tcc	ggg	ctg	tgc	ctg	gag	gcc	gcg	ggc	gcc	ggt	acg	240
Val	Gly	Val	Glu	Ser	Gly	Leu	Cys	Leu	Glu	Ala	Ala	Gly	Ala	Gly	Thr	
65					70				75					80		

gcc	aac	ggc	acg	gcg	gtg	cag	atc	ggg	acg	tgc	gtc	ggt	ggc	agc	ggg	288
Ala	Asn	Gly	Thr	Ala	Val	Gln	Ile	Gly	Thr	Cys	Val	Gly	Gly	Ser	Gly	
			85					90					95			

cag	aag	tgg	tcc	ggc	ctg	acc	ggg	acg	ccc	ccg	acg	gac	ggg	gcg	tgt	336
Gln	Lys	Trp	Ser	Gly	Leu	Thr	Gly	Thr	Pro	Pro	Thr	Asp	Gly	Ala	Cys	
		100					105					110				

gcc	ctt	ccg	tcg	gcg	tac	cgg	tgg	tcg	tcg	acg	ggt	gtg	ctg	gcg	cag	384
Ala	Leu	Pro	Ser	Ala	Tyr	Arg	Trp	Ser	Ser	Thr	Gly	Val	Leu	Ala	Gln	
	115						120					125				

ccg	gcg	aac	ggg	tgg	gcg	gcg	gtc	aag	gac	ttc	acc	acc	gtg	acc	cac	432
Pro	Ala	Asn	Gly	Trp	Ala	Ala	Val	Lys	Asp	Phe	Thr	Thr	Val	Thr	His	
	130				135					140						

aac	ggc	aag	cac	ctg	gtc	tac	gcg	tcg	aac	gtg	tcg	ggg	tcg	tcc	tac	480
Asn	Gly	Lys	His	Leu	Val	Tyr	Ala	Ser	Asn	Val	Ser	Gly	Ser	Ser	Tyr	
145				150					155					160		

ggc	tcg	atg	atg	ttc	agc	ccc	ttc	acg	aac	tgg	tcg	gac	atg	gcg	tcg	528

Gly Ser Met Met Phe Ser Pro Phe Thr Asn Trp Ser Asp Met Ala Ser	
165 170 175	
gcc ggc cag acc ggg atg agt cag gcg gcg gtg gcg ccc acc ctg ttc	576
Ala Gly Gln Thr Gly Met Ser Gln Ala Ala Val Ala Pro Thr Leu Phe	
180 185 190	
tac ttc gcg ccc aag aac gtc tgg gtg ctg gcg tac cag tgg ggc gcg	624
Tyr Phe Ala Pro Lys Asn Val Trp Val Leu Ala Tyr Gln Trp Gly Ala	
195 200 205	
tgg ccc ttc gtc tac cgc acg tcc agc gac ccg gcc aac ccc aac ggc	672
Trp Pro Phe Val Tyr Arg Thr Ser Ser Asp Pro Ala Asn Pro Asn Gly	
210 215 220	
tgg tcg gct ccg cag ccg ttg ttc acc ggg agc atc tca gga tcc gac	720
Trp Ser Ala Pro Gln Pro Leu Phe Thr Gly Ser Ile Ser Gly Ser Asp	
225 230 235 240	
acc ggc ccg atc gac cag acc ctg atc gcc gac gga cag aac atg tac	768
Thr Gly Pro Ile Asp Gln Thr Leu Ile Ala Asp Gly Gln Asn Met Tyr	
245 250 255	
ctg ttc ttc gcc ggt gac aac ggc aag atc tac ccg gcg agc atg ccg	816
Leu Phe Phe Ala Gly Asp Asn Gly Lys Ile Tyr Arg Ala Ser Met Pro	
260 265 270	
atc ggg aac ttc ccg ggc agc ttc ggc tcg tcg tac acg acg atc atg	864
Ile Gly Asn Phe Pro Gly Ser Phe Gly Ser Ser Tyr Thr Thr Ile Met	
275 280 285	
agc gac acg aag gcc aac ctg ttc gag ggc gtg cag gtc tac aag gtt	912
Ser Asp Thr Lys Ala Asn Leu Phe Glu Gly Val Gln Val Tyr Lys Val	
290 295 300	
cag ggc cag aac cag tac ctg atg atc gtc gag gcg atg ggc gcg aac	960
Gln Gly Gln Asn Gln Tyr Leu Met Ile Val Glu Ala Met Gly Ala Asn	
305 310 315 320	
gga cgc tac ttc cgt tcc ttc acg gcc tcc agc ctg agc ggc tcg tgg	1008
Gly Arg Tyr Phe Arg Ser Phe Thr Ala Ser Ser Leu Ser Gly Ser Trp	
325 330 335	
acc ccg cag gcc gcg agc gag agc aac ccc ttc gcg ggc aag gcc aac	1056
Thr Pro Gln Ala Ala Ser Glu Ser Asn Pro Phe Ala Gly Lys Ala Asn	
340 345 350	

300

agc ggg gcc acc tgg acc aac gac atc agc cac ggt gac ctg gtc cgc 1104
Ser Gly Ala Thr Trp Thr Asn Asp Ile Ser His Gly Asp Leu Val Arg
355 360 365

gac aac ccc gac cag acc atg acc gtc gac ccc tgc aac ctg cgg ttc 1152
Asp Asn Pro Asp Gln Thr Met Thr Val Asp Pro Cys Asn Leu Arg Phe
370 375 380

ctc tac cag ggc aag tca ccc acc gcg ggc ggc ccc tac gac cag ctg 1200
Leu Tyr Gln Gly Lys Ser Pro Thr Ala Gly Gly Pro Tyr Asp Gln Leu
385 390 395 400

ccg tgg cgg ccg ggt gtg ctc acc ctg cgg cgc tga 1236
Pro Trp Arg Pro Gly Val Leu Thr Leu Arg Arg
405 410

<210>	150
<211>	411
<212>	PRT
<213>	Streptomyces ambofaciens

<400> 150

Leu Gln Ile Tyr Asp Cys Trp Gly Gly Ser Asn Gln Gln Trp Thr Leu
1 5 10 15

Thr Asp Asp Arg Arg Leu Thr Val Tyr Gly Asn Lys Cys Leu Asp Val
20 25 30

Pro Gly His Ala Thr Thr Pro Gly Thr Arg Val Gln Ile Trp Ser Cys
35 40 45

Ser Gly Gly Ala Asn Gln Gln Trp Arg Val Asn Ser Asp Gly Thr Val
50 55 60

Val Gly Val Glu Ser Gly Leu Cys Leu Glu Ala Ala Gly Ala Gly Thr
65 70 75 80

Ala Asn Gly Thr Ala Val Gln Ile Gly Thr Cys Val Gly Gly Ser Gly
85 90 95

Gln Lys Trp Ser Gly Leu Thr Gly Thr Pro Pro Thr Asp Gly Ala Cys
 100 105 110

Ala Leu Pro Ser Ala Tyr Arg Trp Ser Ser Thr Gly Val Leu Ala Gln
 115 120 125

Pro Ala Asn Gly Trp Ala Ala Val Lys Asp Phe Thr Thr Val Thr His
 130 135 140

Asn Gly Lys His Leu Val Tyr Ala Ser Asn Val Ser Gly Ser Ser Tyr
 145 150 155 160

Gly Ser Met Met Phe Ser Pro Phe Thr Asn Trp Ser Asp Met Ala Ser
 165 170 175

Ala Gly Gln Thr Gly Met Ser Gln Ala Ala Val Ala Pro Thr Leu Phe
 180 185 190

Tyr Phe Ala Pro Lys Asn Val Trp Val Leu Ala Tyr Gln Trp Gly Ala
 195 200 205

Trp Pro Phe Val Tyr Arg Thr Ser Ser Asp Pro Ala Asn Pro Asn Gly
 210 215 220

Trp Ser Ala Pro Gln Pro Leu Phe Thr Gly Ser Ile Ser Gly Ser Asp
 225 230 235 240

Thr Gly Pro Ile Asp Gln Thr Leu Ile Ala Asp Gly Gln Asn Met Tyr
 245 250 255

Leu Phe Phe Ala Gly Asp Asn Gly Lys Ile Tyr Arg Ala Ser Met Pro
 260 265 270

Ile Gly Asn Phe Pro Gly Ser Phe Gly Ser Ser Tyr Thr Thr Ile Met
 275 280 285

Ser Asp Thr Lys Ala Asn Leu Phe Glu Gly Val Gln Val Tyr Lys Val
 290 295 300

Gln Gly Gln Asn Gln Tyr Leu Met Ile Val Glu Ala Met Gly Ala Asn
 305 310 315 320

Gly Arg Tyr Phe Arg Ser Phe Thr Ala Ser Ser Leu Ser Gly Ser Trp
 325 330 335

Thr Pro Gln Ala Ala Ser Glu Ser Asn Pro Phe Ala Gly Lys Ala Asn
 340 345 350

Ser Gly Ala Thr Trp Thr Asn Asp Ile Ser His Gly Asp Leu Val Arg
 355 360 365

Asp Asn Pro Asp Gln Thr Met Thr Val Asp Pro Cys Asn Leu Arg Phe
 370 375 380

Leu Tyr Gln Gly Lys Ser Pro Thr Ala Gly Gly Pro Tyr Asp Gln Leu
 385 390 395 400

Pro Trp Arg Pro Gly Val Leu Thr Leu Arg Arg
 405 410

<210> 151

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo KF36

<400> 151

ttgccgtagc cgaggaccag cg

<210> 152

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

<220>
 <223> Oligonucleotideo KF37

<400> 152
 cacatggccc tggaggaccc tg 22

<210> 153
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo KF42

<400> 153
 aagcttgtac ggcccacaga atgatgtcac 30

<210> 154
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo KF43

<400> 154
 aagcttcgac taccttggtg atctgcctt 30

<210> 155
 <211> 65
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotideo KF32

<400> 155
 caaccgcttg agctgctcca tcaactgctg ggccgaggta tcgcgcgcgc ttcgttcggg 60

acgaa 65

<210> 156

<211> 65

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo KF33

<400> 156

tgggtcccgc cgcgcgccac gacttcgact cgctcgtcta tctgcctctt cgtcccgaag 60

caact 65

<210> 157

<211> 65

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR71

<400> 157

cgtcatcgac gtgcggggaa gacagaggtg ataccgatga tcgcgcgcgc ttcgttcggg 60

acgaa 65

<210> 158

<211> 65

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo EDR72

<400> 158

gccagcacct cgtccagctg ctgcacggaa ctcaccccca tctgcctctt cgtcccgaag 60

caact 65

<210> 159

<211> 66

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo KF52

<400> 159

gatccgccag cctcacgtca cgccgcgcgc cctccctgac atcgcgcgcg ctctcgttcgg 60

gacgaa 66

<210> 160

<211> 66

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotideo KF53

<400> 160

gaggcggacg tcggtacgcg gtgggagccg gagttcgaca atctgcctct tcgtcccga 60

gcaact 66

<210> 161

<211> 2540

<212> DNA

<213> Streptomyces ambofaciens

<400> 161

ctgcagggcg cccacgggggt ggcggtcgta tcggaactcg gcgtcgtagt cgtcctccag 60

caccaggccg tcgacctcgc gggcccaggc gaccagctcg ccgcgacggg ccggggcgag 120

gaccacaccg gtcgggaact ggtgggcggg agcgacgac accgaccgga cctgaggcgc 180

gcgccgcagc aggtcgaccc gcaggccgtc gccgtccacg ggaaccggga cgggcaccag 240

cccctccgcc ccgaccgtgg cgctcagccg cgaccagcca gggtcctcca gggccatgtg 300

ggtgtggccc tcggtccgca ggaccgggca cagcctgcgc accgcgtcca gcgtgcccgc 360

gcagacgacg agctgccggg cttccagcga cgcgccgcgt ccccgacga ggtaggccgc 420

gagccgctgc cggagccggg cgagaccggc cggatcgggg tagccgagct cggcccaggt 480

cagtgcgccg gtggcctccc gcaccgcgtc cgtcccaccgt ccgcggggaa aggcccgagc 540
 gtcgggtatg ccgggcagca tgtcgaactc cggctcccac cgcgtaccga cgtccgcctc 600
 gggcgtcacc ggagcgggga cggcctgcgc cctgaccgcg gtggccgagc cgtcgcgcgc 660
 ctccaggtac ccctccgcga cgagctgcgc gtaggcctcc gtcaccaccc accgcgagca 720
 gccaggtcc gccgccagcg ccctgctcgg cggcagcgcg ctgccggagg cgatccgccc 780
 accggtcacc gccgcccga gggcacggct cagccggacg tgcagcgggc cgtccgcccg 840
 cccgcccagt tcgagaaagg ttcccatgc cgggtcggtc cggctcgtgc ccacggcagc 900
 cgatgctacc caccgccccg cgacggccgg gcggcggcgt ccacggggcc ggtgcacact 960
 ggatggcgtg gcggagacgg accggcatct cgaccagcgg ctggagcagg ccatcatgga 1020
 cctgctggag cgacgggcgg ccactgcctc gatctgtccg tccgacgccg cgcgtcgggt 1080
 ccacgagggg gacgacgagg gctggcgagc cctgctggaa cccgcgcgcc gagcggcctg 1140
 gcgtctggtg gggtcggcg aggtcaggt gacccaggcc ggccggcccc tcacgcaggc 1200
 cgaggccccg gggcccgtcc gcattcgtcg ggcgggttcc tgacggctcg atccgaaggc 1260
 ggaactcccg tgccccggcg gcacgggagc tcgacgccgg gaggtcagcg ccgcagggtg 1320
 agcacacccg gccgccacgg cagctggctg taggggccgc ccgcggtggg tgacttgccc 1380
 tggtagagga accgcaggtt gcaggggtcg acggtcatgg tctggtcggg gttgtcgcg 1440
 accaggtcac cgtggctgat gtcgttggtc cagggtggccc cgtgtttggc cttgcccgcg 1500
 aaggggttgc tctcgtcgc ggctgcggg gtccacgagc cgtcaggct ggaggccgtg 1560
 aaggaacgga agtagcgtcc gttcgcgcc atgcctcga cgatcatcag gtactggttc 1620
 tggccctgaa cttttagac ctgcacgcc tcgaacaggc tggccttcgt gtcgtcatg 1680
 atcgtcgtgt acgacgagc gaagctgcc gggaagttcc cgatcggcat gctcggcccg 1740
 tagatcttgc cgttgtcacc ggcgaagaac aggtacatgt tctgtccgtc ggcgatcagg 1800
 gtctggtcga tcgggcgggt gtcggatcct gagatgctcc cgggtgaacaa cggctgcgga 1860
 gccgaccage cgttgggggt ggccgggtcg ctggacgtgc ggtagacgaa gggccacgcg 1920

ccccactggt acgccagcac ccagacgttc ttgggcgcga agtagaacag ggtgggcgcc	1980
accgccgcct gactcatccc ggtctggccg gccgacgcca tgtccgacca gttcgtgaag	2040
gggctgaaca tcatcgagcc gtaggacgac ccgacacgt togacgcgta gaccaggtgc	2100
ttgccgttgt gggtcacggt ggtgaagtcc ttgaccgcg cccaccogtt cgccggctgc	2160
gccagcacac ccgtcgacga ccaccggtac gccgacggaa gggcacacgc cccgtccgtc	2220
ggggggcgcc cggtcaggcc ggaccacttc tgcccgtgc caccgacga cgtcccgatc	2280
tgcaccgcg tgccgttggc cgtaccggcg ccgcggcct ccaggcacag cccggactcc	2340
acgccgacga ccgtgccgtc ggagttcacc cgccactgct ggttcgcacc gccggaacag	2400
ctccagatct gcaccgcgt cccgggtgtg gtggcgtgac ccggaacgtc caggcacttg	2460
ttgccgtaca cggtcagccg acggtcgtcc gtcaacgtcc actgctgatt gctcccgccc	2520
cagcagtcgt atatctgcag	2540

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

Código de Controle

Campo 1



C4CE4EA633555D66

Campo 2



78339BF4D573670D

Outras Informações:

- Nome do Arquivo: p131413 27797630_1.txt
- Data de Geração do Código: 29-02-2016
- Hora de Geração do Código: 09:03:52
- Código de Controle:
 - Campo 1: C4CE4EA633555D66
 - Campo 2: 78339BF4D573670D