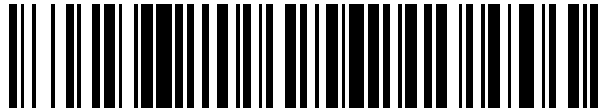


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 543 203**

51 Int. Cl.:

**C12R 1/465** (2006.01)  
**C07K 14/36** (2006.01)  
**C12N 15/52** (2006.01)  
**C12N 1/20** (2006.01)  
**C07G 11/00** (2006.01)  
**C12P 17/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.03.2009 E 09723397 (7)**

97 Fecha y número de publicación de la concesión europea: **06.05.2015 EP 2271666**

54 Título: **Grupo de genes NRPS-PKS, su manipulación y su utilidad**

30 Prioridad:

**20.03.2008 GB 0805363**

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

**17.08.2015**

73 Titular/es:

**SINTEF TTO AS (100.0%)  
Postboks 4764 Sluppen  
7465 Trondheim, NO**

72 Inventor/es:

**ZOTCHEV, SERGEY;  
JØRGENSEN, HANNE;  
SLETTA, HAVARD;  
ELLINGSEN, TROND ERLING;  
FJAERVIK, ESPEN;  
DEGNES, KRISTIN FLØGSTAD;  
KLINKENBERG, GEIR y  
BRUHEIM, PER**

74 Agente/Representante:

**ISERN JARA, Jorge**

**ES 2 543 203 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

## DESCRIPCIÓN

Grupo de genes NRPS-PKS, su manipulación y su utilidad

5 La presente invención se refiere a la clonación y secuenciación del grupo de genes que codifica la maquinaria biosintética para la síntesis de la macrolactama policétido BE-14106, que incluye tanto un dominio de adenilación de péptido sintetasas no ribosomales (NRPS) como una enzima biosintética o complejo enzimático policétido modular (PKS; enzima o complejo enzimático policétido sintasa). La maquinaria para la biosíntesis comprende por tanto un sistema enzimático NRPS-PKS híbrido. Por consiguiente, la invención se refiere a nuevos genes y a moléculas de  
10 ácido nucleico que codifican la maquinaria biosintética para la síntesis de la macrolactama BE-14106, que incluye una enzima biosintética o un sistema enzimático NRPS-policétido modular que interviene en la biosíntesis de BE-14106 y a la maquinaria biosintética que incluye el propio sistema o complejo enzimático NRPS-policétido sintasa modular (así como sus componentes). Adicionalmente, la invención se refiere al uso de estos genes, a moléculas de  
15 ácido nucleico, a la maquinaria, a enzimas y a sistemas enzimáticos o complejos de los mismos, que facilitan tanto la biosíntesis de BE-14106 como la síntesis de derivados de BE-14106 y nuevas estructuras macrolactama.

Los policétidos, o estructuras relacionadas o basadas en policétidos, son, forman la base de, productos naturales sintetizados por bacterias, hongos, plantas y animales, muchos de los cuales tienen la posibilidad de aplicarse como agentes farmacéuticos o como productos agrícolas o veterinarios, por ejemplo, como antibióticos, antifúngicos, citostáticos, antiolesterolémicos, antiparasitarios, coccidiostáticos, y promotores de crecimiento animal e insecticidas naturales.

Las bacterias *Streptomyces* gram positivas son las principales productoras de policétidos y de moléculas basadas en policétidos y la genética y la bioquímica de la biosíntesis de policétidos en estos organismos están relativamente bien caracterizadas (McDaniel R, *et al.*, Chem Rev. 2005 Feb; 105(2): 543-58).

Otros productores incluyen otros actinomicetos. Se conoce una serie de diferentes moléculas basadas en policétidos (o relacionadas con policétidos), cuyas macrolactamas representan una clase. Ogasawara Y. *et al.*; Chem Biol. Ene 2004; 11(1): 79-98 y Udvary *et al.*; Proc Natl Acad Sci US A, 19 jun 2007; 104(25): 10376-81, respectivamente, han descrito los grupos de genes biosintéticos para la síntesis de las macrolactamas vicenistatina y salinilactama.

BE-14106 (nombre alternativo GT-32A) es un antibiótico de macrolactama que tiene una estructura química como se expone en la Figura 1. Se ha aislado de una cepa de *Streptomyces spheroides* y se ha observado que tiene efectos citotóxicos en líneas de células de leucemia, así como actividad antimicrobiana contra una serie de organismos ensayados, actividad antiproliferativa contra una línea de células BALB3T3 transformada mediante H-ras y actividad inhibidora contra la reacción linfocitaria mixta (documento JP4001179, Kojiri *et al* 1992 Journal of Antibiotics, 868-74, Takahashi *et al* 1997, Journal of Antibiotics 186-8). También se ha aislado un análogo 8-desoxi (GT-32B) de una especie de *Streptomyces* no especificada y se observó que compartía muchas de las actividades de BE-14106 (Takahashi *et al*, citados anteriormente).

Los compuestos de macrolactama, tales como BE-14106, puede formarse mediante activación y cebado del sistema PKS con un aminoácido activado y la extensión del resto aminoacídico (cadena aminoacilo) por condensaciones repetidas de ácidos carboxílicos simples por policétido sintetasas (PKS) de manera similar a la biosíntesis de ácidos grasos. Por tanto, a diferencia del caso con una simple cadena de policétido en el que la "unidad iniciadora" es un resto de ácido carboxílico, en este caso, la unidad iniciadora para la PKS es un producto intermedio aminoacilo sintetizado a partir de un aminoácido y de una cadena acilo. Las PKS pueden organizarse como PKS repetitivas que re-utilizan dominios en una manera cíclica o como PKS modulares (de Tipo I) que contienen una secuencia de módulos (o unidades repetidas) distintos y no re-utilizan dominios. Cada módulo es responsable de un ciclo de condensación en la síntesis de la cadena del policétido y contiene diversos dominios enzimáticos. En el caso de BE-14106 la cadena "policetídica" es, estrictamente hablando, una cadena de aminoácidos - policetídica híbrida, o una cadena aminoacilo, pero en el presente documento se denomina "cadena policetídica". Por tanto, además de dominios para la condensación de los siguientes ácidos carboxílicos sobre la cadena del policétido en crecimiento, catalizada por el dominio de la  $\beta$ -cetoacil sintasa (KS), los módulos de la PKS de tipo I pueden contener dominios con actividades  $\beta$ -ceto-reductasa (KR), deshidratasa (DH) o enoil reductasa (ER), que determinan el estado reducido de unidades extendedoras incorporadas. Los dominios de la acetiltransferasa (AT) y de la proteína transportadora de acilos (ACP, *Acyl Carrier Protein*) presentes en cada módulo son responsables de la elección de la unidad extendedora y de la conservación de la cadena del policétido en crecimiento en la PKS, respectivamente. Después de finalizar la síntesis, la cadena del policétido se libera de las PKS mediante la acción de una tiosterasa (TE), que también está probablemente implicada en la ciclación del producto final. Por tanto, las PKS de tipo I representan una línea de ensamblaje para la biosíntesis de policétidos, que puede manipularse cambiando el número de módulos, sus especificidades hacia los ácidos carboxílicos e inactivando o insertando dominios con actividades reductoras (Weissman y Leadlay, Nat. Rev. Microbiol. 2005 Dec; 3(12): 925-36). Después de que la fracción policétido se sintetiza y cicla para formar un anillo de macrolactona (o macrolactama), esta puede modificarse mediante hidroxilación, glucosilación, metilación y/o acilación. Estas modificaciones pueden ser importantes para las actividades biológicas de determinados productos basados en policétidos. Como se describirá con más detalle más adelante, en el trabajo que conduce a la presente invención, los genes que codifican el sistema enzimático NRPS-

PKS de BE-14106 (el "grupo de genes" de BE-14106) se han clonado y secuenciado y se ha determinado que el sistema enzimático NRPS-PKS de BE-14106 contiene diversas PKS de tipo I, cada una de las cuales se organiza de manera modular y está constituida por unidades repetidas (módulos).

5 Los genes para la biosíntesis de policétidos en *Streptomyces* se organizan generalmente en grupos, y diversos de dichos grupos ya identificados son los responsables de la síntesis de diversos productos naturales. La clonación molecular y secuenciación completa de ADN de diversos grupos de genes antibióticos macrólidos de *Streptomyces* se ha descrito, incluyendo los de la avermectina, pikromicina y rapamicina (Ikeda H., Omura S. (2002). Biosynthesis, Regulation, and Genetics of Macrolide Production. In: Macrolide Antibiotics: Chemistry, Biology and Practice, 2ª Ed. (ed. S. Omura), págs. 286-326, Academic Press, Nueva York). Como se ha mencionado anteriormente, también se han notificado grupos de genes para los biosistemas de determinados antibióticos de macrolactama.

15 Como se ha indicado anteriormente y como se describe más adelante, la presente invención se basa en la identificación, clonación y secuenciación de un nuevo grupo de genes para la biosíntesis de BE-14106 que hasta ahora no estaba disponible. Adicionalmente, el análisis de los genes clonados ha permitido aclarar la ruta biosintética de BE-14106. Por consiguiente ahora se propone que el proceso normal de síntesis de BE-14106 se inicia a través de la síntesis de una unidad iniciadora (C17-C25), en la que una fracción acilo se sintetiza a partir de 1 unidad de propionato y 2 de acetato. La síntesis de la unidad iniciadora continúa con la activación de una molécula de glicina mediante un dominio de adenilación NRPS y cargando la glicina activada sobre una proteína transportadora peptídica. La desaminación oxidativa de la glicina libera amoniaco, lo que constituye un ataque nucleófilo en el carbonilo C-17 para formar un grupo imino C-17, que posteriormente se reduce a un grupo amino. La liberación de la cadena aminoacil de la proteína transportadora peptídica, produce la formación de un ácido carboxílico, que después se adenila y se liga con la coenzima A (CoA). La aminoacil-CoA activada resultante se transfiere al dominio ACP de una PKS mediante una aciltransferasa y se extiende y modifica por la acción secuencial de las enzimas en el sistema PKS como se describe adelante con más detalle. El dominio enzimático  $\beta$ -cetoacil sintasa (KS) en cada módulo cataliza la condensación del ácido carboxílico apropiado (por ejemplo, acetato o propionato), determinada por el módulo aciltransferasa (AT). Los dominios enzimáticos con actividad  $\beta$ -cetoalreductasa (KR) o deshidratasa (DH) determinan el estado reducido de las unidades extendedoras incorporadas.

20 La cadena lateral de hidratos de carbono C20-C25 de BE-14106 contiene parte de la unidad iniciadora y es el resultado de la ciclación del anillo de macrolactama. Finalmente, la modificación adicional del anillo de macrolactama se produce mediante hidroxilación.

25 El grupo de genes biosintético de BE-14106 también codifica o incluye diversos elementos reguladores y proteínas para el transporte de las moléculas sintetizadas.

30 Dado que la síntesis química de compuestos tales como éste es muy compleja, en la práctica se requiere usar una ruta biosintética y por consiguiente es deseable el aislamiento o la purificación de los compuestos de hospedadores apropiados. Como se ha reconocido en la técnica, esto ofrece la oportunidad de manipular genes del grupo de genes de PKS para cambiar la biosíntesis y por lo tanto dar como resultado la síntesis de nuevos policétidos o compuestos basados en policétidos modificados. Aunque se ha descrito que la modificación de diversos grupos de genes de PKS da como resultado la síntesis de diversos nuevos compuestos, aún continúa habiendo una necesidad y un deseo de aumentar el repertorio de compuestos disponibles, especialmente antibióticos, y/o de mejorar las propiedades (por ejemplo, eficacia, toxicidad, solubilidad en agua, etc.) de los fármacos existentes. La presente invención se dirige a estos objetivos, y se baja en la clonación y secuenciación de ADN del grupo de genes biosintético de BE-14106. Esto proporciona la primera secuencia para estos genes biosintéticos de antibióticos, así como una herramienta para la manipulación genética con objeto de modificar los niveles de expresión o las propiedades de BE-14106 y/o la producción de organismos, u obtener nuevos compuestos posiblemente útiles. A este respecto, aunque se conoce el antibiótico BE-14106, en el fondo de una pluralidad de moléculas basadas en policétido sintetizadas en *Streptomyces* y correspondientes a una pluralidad de grupos de genes biosintéticos, no fue una cuestión sencilla identificar y clonar el grupo de genes correcto para BE-14106; se ha requerido un considerable esfuerzo e ingenuidad en cuanto al análisis de secuencias y diseño o selección de sondas.

35 Los autores de la presente invención han aislado y purificado BE-14106 de una fuente previamente desconocida, el aislado bacteriano MP28-13, que se piensa que es una nueva cepa de *Streptomyces* (depositada con el número de depósito DSM21069, el 25 de enero de 2008, en la Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ)) que se aisló de sedimento de agua superficial en el fiordo de Trondheim, Noruega. El aislamiento de este nuevo microorganismo ha permitido a los inventores clonar y secuenciar todo el grupo de genes biosintético BE-14106. Este grupo contiene 22 genes que codifican proteínas que se supone que están implicadas en la biosíntesis de la molécula BE-14106 (véase la Tabla 1).

40 Para realizar esta clonación, se usaron cebadores oligonucleotídicos especialmente diseñados que representan secuencias que codifican partes del dominio ceto sintasa (KS) para la amplificación de regiones codificantes del dominio KS del aislado MP28-13. Una vez obtenidas y caracterizadas las secuencias amplificadas, basándose en análisis bioinformáticos complejos e intensos, se seleccionó una de las secuencias como una sonda. Esta sonda se usó para explorar la biblioteca genómica que se construyó para MP28-13. Los cósmidos que se identificaron de esta

manera se analizaron y se secuenciaron para proporcionar todo el grupo biosintético. La secuencia se había anotado por completo y se había aclarado una ruta de dos partes para la biosíntesis BE-14106, como se expone en las Figuras 2A y 2B. La primera parte de la ruta de la biosíntesis de la unidad aminoácido iniciadora se muestra en la Figura 2A y en la Figura 2B se representa la segunda parte, el alargamiento de la cadena aminoácido, su ciclación resultante en la formación del anillo de macrolactama y la post-modificación de PKS. Por tanto, se ha propuesto que el grupo de genes biosintético de BE-14106 codifica un primer sistema o complejo enzimático que comprende PKS y otras enzimas o proteínas para la síntesis de la cadena aminoácido y un sistema enzimático o complejo enzimático PKS adicional para la prolongación de dicha cadena de aminoácido, así como una enzima para la post-modificación de la molécula PKS, proteínas para la regulación de la ruta, y proteínas de resistencia/salida.

Basándose en el conocimiento de la secuencia, se desarrolló un método de manipulación genética de MP28-13 de especies de *Streptomyces*. De este modo fue posible demostrar que la nueva secuencia era de hecho responsable de la biosíntesis de BE-14106.

Además, como se describirá adelante con más detalle, la manipulación de secuencias de ADN funcionales dentro del nuevo grupo de genes biosintético que se ha identificado, puede conducir a la síntesis de nuevas estructuras moleculares, por ejemplo, derivados o análogos de BE-14106, con funciones o propiedades alteradas, por ejemplo, mejoradas. Como tal, el grupo de genes de BE-14106 puede manipularse para obtener no solo nuevos derivados o análogos de BE-14106 beneficiosos, sino también para mejorar y facilitar el proceso de producción biosintética (por ejemplo, mejorar el rendimiento, o las condiciones de producción o expandir el intervalo de células hospedadoras disponibles) o más preferentemente proporcionar nuevos compuestos con nuevas actividades y/o propiedades.

En la SEC ID N°: 1 se muestra la secuencia codificante completa (es decir la secuencia codificante de nucleótidos completa) del grupo de genes biosintético de BE-14106. Se ha observado que esta contiene diversos genes u ORF, que codifican las diversas proteínas y polipéptidos que son responsables de las actividades que se requieren para la biosíntesis de BE-14106.

El grupo de genes biosintético contiene genes y ORF que se piensa que codifican todas las proteínas y polipéptidos que se necesitan para la biosíntesis normal de BE-14106. Sin embargo, no todas las proteínas y polipéptidos codificados se les ha atribuido aún un papel en la biosíntesis y por eso puede que no todas las proteínas o polipéptidos codificados del grupo sean esenciales para la biosíntesis de BE-14106. Los diversos genes y ORF pueden codificar enzimas que catalizan una o más reacciones bioquímicas, o proteínas que no tienen actividad catalítica pero que, en su lugar, están implicadas en otros procesos, tales como la regulación de los procesos de la síntesis de BE-14106, o el transporte de BE-14106, por ejemplo.

Diversas enzimas son policétido sintetas (PKS) y es posible que varias de estas PKS puedan asociarse físicamente para formar un complejo enzimático, aunque esto aún no se ha establecido. En el presente documento dicho grupo o conjunto de enzimas recibe el nombre de sistema o complejo enzimático biosintético policétido o sistema o complejo enzimático PKS, aunque no es necesario que todas las enzimas/proteínas en el sistema/complejo sean policétido sintetas reales, es decir, que tengan actividad policétido sintasa; pudiendo tener otras actividades o papeles funcionales en la síntesis de BE-14106. Por ejemplo, un dominio de adenilación específico de péptido sintetas no ribosomales (NRPS) (BecL), junto con algunas otras proteínas accesorias (por ejemplo BecJ, BecS, BecU) codificadas por el grupo están implicados en la síntesis de la unidad iniciadora para la biosíntesis de BE-14106 activando un aminoácido (sustituyendo una glicina) y su carga en uno de los diversos módulos de la PKS BE-14106 para una prolongación adicional. Otras proteínas, tales como BecO, realizan la hidroxilación del anillo de macrolactama en C-8. Un grupo o conjunto de enzimas que comprende dicho dominio de NRPS y enzimas PKS pueden recibir el nombre de sistema enzimático o complejo enzimático NRPS-PKS híbrido. La agrupación de proteínas y polipéptidos que codifica el grupo de genes en su totalidad recibe conjuntamente el nombre de maquinaria biosintética para la biosíntesis de BE-14106.

Por tanto, en un aspecto, la presente invención proporciona una molécula de ácido nucleico que comprende:

- (a) una secuencia de nucleótidos como se muestra en la SEC ID N° 1; o
- (b) una secuencia de nucleótidos que es el complemento de la SEC ID N° 1; o
- (c) una secuencia de nucleótidos que tiene una identidad de secuencia de al menos 85 % (preferentemente una identidad de secuencia de al menos 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %) con la SEC ID N° 1; o
- (d) una parte de uno cualquiera de (a) a (c), en el que dicha parte comprende:

- (i) una secuencia de nucleótidos como se muestra en una cualquiera de las SEC ID Nos 2, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 o 44 o una secuencia de nucleótidos que es complementaria de la misma o que tiene una identidad de secuencia de al menos 85 %; o
- (ii) una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos seleccionadas de la SEC ID Nos 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45 o que tiene una identidad de secuencia de al menos 85 % (preferentemente al menos 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %) con las mismas,

en el que dicha molécula de ácido nucleico codifica o es complementaria a una molécula de ácido nucleico que codifica uno o más polipéptidos que tienen actividad funcional en la síntesis de BE-14106.

Una actividad funcional puede ser una actividad enzimática, por ejemplo, una actividad implicada en la síntesis o transporte o transferencia de BE-14106 (ésta puede ser la síntesis de BE-14106 o cualquier etapa que contribuya a la misma, o una modificación de la cadena BE-14106, etc.) y/o puede ser una actividad reguladora, por ejemplo regulación de la expresión de los genes (por ejemplo, un regulador transcripcional) o proteínas implicadas en la síntesis, o regulación del proceso sintético, y/o puede ser una "actividad transportadora". Por tanto, generalmente también se incluyen proteínas transportadoras implicadas en la transferencia o transporte de BE-14106, por ejemplo, en el transporte o salida de la molécula sintetizada dentro o fuera de la célula.

En el presente documento también se desvelan secuencias de nucleótidos que comprenden elementos genéticos funcionales tales como promotores, regiones operadoras promotoras, potenciadores, u otras secuencias reguladoras, etc. Por tanto, no es necesario que la molécula de ácido nucleico de la invención comprenda todo el grupo de genes de PKS sino que puede comprender una porción o parte de este, por ejemplo, una parte que codifique un polipéptido que tenga una función particular como se describe anteriormente y opcionalmente una secuencia reguladora. Esta puede comprender uno o más genes y opcionalmente secuencias reguladoras, uno o más módulos, dominios enzimáticos o elementos genéticos funcionales codificantes (por ejemplo elementos que controlen la expresión, la transcripción, la traducción, etc. de genes). Generalmente hablando, una molécula de ácido nucleico de la invención comprenderá diversos genes diferentes y opcionalmente secuencias reguladoras que conducen a la síntesis de BE-14106.

Un "gen biosintético u ORF de BE-14106" se define adicionalmente más adelante, pero en resumen en el contexto de la sección (e) anterior significa un gen u ORF que codifica una proteína o polipéptido que es funcional en el proceso biosintético de BE-14106. Como se ha indicado anteriormente, éste podría ser una enzima que estuvo implicada en la activación del aminoácido de partida, en la transferencia del aminoácido/cadena de aminoácido activada a una enzima PKS, la generación de la cadena policétido o modificación de la misma, o una proteína que es necesaria para la regulación o para el transporte o transferencia de la molécula en cualquier fase de su biosíntesis.

Una molécula de ácido nucleico de la invención puede ser una molécula de ácido nucleico aislada (en otras palabras aislada o separada de los componentes con los que normalmente se encuentra en la naturaleza) o puede ser una molécula de ácido nucleico recombinante o sintética.

Como se trata en cualquier parte del presente documento, el grupo de genes biosintético de BE-14106 es una gran molécula de ácido nucleico que contiene los diversos elementos genéticos o diferentes genes o las ORF que codifican las proteínas o los péptidos que se necesitan para la biosíntesis de la molécula BE-14106 o un derivado o análogo de BE-14106 o molécula relacionada con BE-14106. Cada gen biosintético u ORF de BE-14106 codifica una sola cadena polipeptídica (que, como alternativa, puede describirse como una proteína) que tiene o que se cree que tiene una función en la biosíntesis de la molécula BE-14106 o un derivado o análogo de BE-14106 o molécula relacionada con BE-14106. Se han identificado 22 de dichos genes u ORF (véase la Tabla 1). Como se observa en las Figuras 2A y 2B, a 14 de éstos se les atribuye un papel directo en la biosíntesis de BE-14106. Como se explica más adelante, también se piensa o se propone que algunos otros desempeñan un papel en la biosíntesis de BE-14106. Por tanto, por ejemplo, se piensa que *becH* y *M* codifican reguladores, se piensa que *BecL* está implicada en la activación de glicina, se piensa que *BecU* media la interacción proteica entre la ACP de *BecC* y se piensa que la PCP de *BecS*, *BecN* está implicada en la salida y/o resistencia, se piensa que *BecP* ayuda a la ciclación del anillo de macrolactama y *BecQ* en la liberación de la cadena aminoácido de inicio para el complejo *BecC-BecU-BecS*.

Determinadas proteínas tienen actividad enzimática y puede por tanto definirse como enzimas. Diversas de estas enzimas pueden describirse como policétido sintetas (PKS). Dichas enzimas contienen uno o más de un módulo y cada módulo puede contener de uno a seis, preferentemente dos, tres, cuatro o cinco dominios enzimáticos, cada uno de los cuales es responsable de una actividad diferente en la biosíntesis de la molécula BE-14106 o un derivado o análogo de BE-14106 o molécula relacionada con BE-14106. Como tal, en estas PKS puede haber sitios activos múltiples en un solo polipéptido o enzima.

Por ejemplo, la enzima *BecB* es una PKS y tiene tres módulos; el módulo 1 ("módulo de carga" en la Figura 2B) tiene un solo sitio activo o dominio (ACP), y cada uno de los módulos 2 y 3 (Módulos 1 y 2 en la Figura 2B) tiene cinco sitios activos o dominios que tienen actividades KS, AT, DH, KR y ACP. Otras PKS codificadas por el grupo de genes son *BecA*, *BecC*, *BecD*, *BecG*, *BecF* y *BecE*. Dichas PKS pueden contener numerosos dominios, poseyendo cada uno actividad catalítica para extender y/o alterar la estructura del policétido. El policétido pasa a lo largo de la proteína de tal manera que las diferentes actividades se realizan secuencialmente en la cadena del policétido en crecimiento. Como se ha indicado anteriormente, diversas de las PKS codificadas por el grupo de genes pueden asociarse para formar un complejo enzimático biosintético.

La molécula de ácido nucleico de la invención codifica (o comprende una secuencia de nucleótidos codificante) algunos, o más preferentemente todos, los polipéptidos o proteínas que están implicados en la biosíntesis de la

molécula BE-14106. Por ejemplo, la molécula de ácido nucleico puede contener cada uno de los 22 genes u ORF y por tanto codificar cada una de las proteínas que están implicadas en la biosíntesis de la molécula BE-14106 como se expone en la Tabla 1, o puede comprender una porción o parte de la secuencia de nucleótidos de SEC ID N°: 1 que es una secuencia que codifica una sola proteína o polipéptido codificado por un solo gen u ORF dentro del grupo de genes biosintético de BE-14106 como se define en la reivindicación 1. Se contemplan partes que comprenden, por ejemplo, al menos 11,12,13, 14, 15, 16, 17, 18, 19, 20, 21, o hasta 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, (por ejemplo 1-21, 2-20, 3-19, 4-18, 5-17,6-16, 7-15, 8-14, 9-13, 10-12) genes u ORF, en el que dichas partes son como se define en la reivindicación 1. Preferentemente la molécula de ácido nucleico de la invención codifica todas las proteínas que están implicadas en la biosíntesis de la molécula BE-14105 como se expone en la Tabla 1. Como alternativa, puede comprender todas las ORF/genos como se indica en la Tabla 1 excepto cualquiera de uno o más de *becR* y ORF6. Dado que la Tabla 1 expone todos los genes u ORF que se han caracterizado, una molécula de ácido nucleico que codifica todas las proteínas que están implicadas en la biosíntesis de la molécula BE-14106 como se expone en la Tabla 1, puede definirse como una secuencia que comprende el grupo de genes biosintético de BE-14106.

Por tanto, la molécula de ácido nucleico de la invención codifica uno o más polipéptidos como se define en la reivindicación 1 implicados en la biosíntesis de, o que tiene actividad funcional en, la síntesis de BE-14106. Como alternativa puede codificar una o más variantes funcionalmente equivalentes o sus equivalentes funcionales. Como se define anteriormente, las moléculas de ácido nucleico de la invención pueden comprender variantes funcionalmente equivalentes de SEC ID N° 1 y dichas variantes pueden incluir partes u homólogos definidos por un % de identidad de secuencia con la SEC ID N°: 1 como se define en la reivindicación 1. Dichas variantes funcionalmente equivalentes codifican proteínas/polipéptidos que tienen actividad funcional como se define anteriormente. Dicha actividad funcional puede ser una actividad enzimática, por ejemplo una actividad implicada en la síntesis o el transporte o la transferencia de BE-14106 (ésta puede ser la etapa de síntesis de la cadena o anillo o cualquier etapa que contribuya a la misma, o una modificación, etc. en cualquier fase de la biosíntesis, por ejemplo, *BecA*, *BecU*, *BecB*, *BecJ*, *BecK*, *BecS*, *BecO*, *BecD*, *BecG*, *BecF*, *BecE*, *BecT*, *BecQ*, *BecP*, *BecC*, *Becl*, *Becl*) y/o puede ser una actividad reguladora, por ejemplo, la regulación de la expresión de los genes o proteínas que están implicados en la síntesis, o regulación de los procesos sintéticos, por ejemplo, *BecH*, *BecM*, y/o puede ser una "actividad transportadora" o resistencia por ejemplo, *BecN*. Por tanto, generalmente también se incluyen proteínas transportadoras implicadas en la transferencia, transporte o salida de la molécula sintetizada dentro o fuera de la célula.

Dichas moléculas pueden tener al menos 300, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 5000, 10000, 15000, 20000, 30000, o 50000 bases. Por tanto, las longitudes representativas de los fragmentos incluyen fragmentos que tienen de 300 pb a 18000 pb de longitud, por ejemplo 300-3000 pb, 300-2500 pb, 2000-8000 pb, 3000-5000 pb, 4000-17000 pb, 7000-12000 pb u 8000-11000 pb de longitud. Como se ha mencionado anteriormente, dentro de la SEQ ID N°: 1 se han identificado diversos genes y ORF y, como se define en la reivindicación 1, partes o fragmentos que comprenden dichos genes u ORF representan "partes" de la SEC ID N°: 1. Éstos se clasifican en la Tabla 1 indicada a continuación:

Tabla 1

Nombre	Posición inicial en la SEC ID N°:1	Posición final en la SEC ID N°:1	Supuesta función de la proteína codificada	SEC ID N°: (ácido nucleico/proteína)
<i>becH</i>	458	3313	Regulador transcripcional de tipo LuxR	2/3
<i>becA</i>	3664	20412	PKS de tipo I, mod 1 de carga + mod 2 + mod 3 incompleto	4/5
<i>becl</i>	21832	20744 C	glicina oxidasa/oxidoreductasa dependiente de FAD	6/7
<i>becC</i>	23913	21829 C	PKS de tipo I, módulo 3 incompleto	8/9
<i>becU</i>	24508	23945 C	homólogo de <i>S. avermitilis</i> SAV_606, supuesta proteína accesoria NRPS	10/11
<i>becB</i>	35088	24505 C	PKS de tipo I, módulos 1, 2 y 3	12/13
<i>becK</i>	36752	35154 C	supuesta acil CoA sintasa/ligasa	14/15
<i>becK</i>	36947	37918	acil transferasa	16/17
<i>beeS</i>	38170	37934 C	proteína transportadora de peptidilo/acilo	18/19
<i>becL</i>	38288	39805	NRPS, dominio de adenilación	20/21
<i>becM</i>	40384	39788 C	regulador transcripcional de tipo TetR	22/23
<i>becN</i>	40486	42060	bomba de salida de tipo MFS	24/25
<i>becO</i>	43388	42153 C	P450 monooxigenasa	26/27
<i>becD</i>	53553	43435 C	PKS de tipo I, módulos 4 y 5	28/29
<i>becP</i>	54502	53561 C	L-aminoácido amidasa/prolina iminopeptidasa	30/31
<i>becG</i>	60565	54605 C	PKS de tipo I, módulo 9 + dominio TE	32/33
<i>becF</i>	70706	60573 C	PKS de tipo I, módulos 7 y 8	34/35

Nombre	Posición inicial en la SEC ID N°:1	Posición final en la SEC ID N°:1	Supuesta función de la proteína codificada	SEC ID N°: (ácido nucleico/proteína)
<i>becE</i>	75649	70754 C	PKS de tipo I, módulo 6	36/37
<i>becT</i>	76241	75954 C	proteína similar a SimX2, supuesta subunidad de la propionil CoA carboxilasa	38/39
<i>becQ</i>	76563	77336	tioesterasa, tipo II	40/41
<i>becR</i>	77489	78202	fosfolípido de tipo PlsC/glicerol aciltransferasa	42/43
<i>orf6</i>	79912	78302 C	Tripeptidil aminopeptidasa, secretada	44/45

En la Tabla anterior, "C" indica que la proteína está codificada por la cadena complementaria.

5 Las secuencias indicadas anteriormente representan por tanto genes biosintéticos u ORF de BE-14106. En otras palabras, dichos genes/ORF se encuentran dentro del grupo de genes biosintéticos de BE-14106 y codifican proteínas y polipéptidos que tienen, o que se ha propuesto que tienen, un papel en la biosíntesis de BE-14106 en *Streptomyces*. La expresión "gen biosintético de BE-14106" u "ORF biosintética BE-14106" también incluye genes y ORF que codifican proteínas que comparten actividad o función con las proteínas anteriores, y por ejemplo comparten altos niveles de identidad de secuencia, como se define en la reivindicación 1. Éstos pueden, como alternativa, describirse como "variantes funcionalmente equivalentes" o "equivalentes funcionales".

10 En general, el término "gen" incluye la ORF que codifica la proteína, junto con cualquiera de las secuencias reguladoras tales como promotores, mientras que el término "ORF" se refiere solamente a la parte del gen que es responsable de codificar la proteína.

15 Como se indica en el presente documento, las "variantes funcionalmente equivalentes" o "equivalentes funcionales" conservan al menos una función de la entidad con la que está relacionada (o de la que procede), por ejemplo, codifican una proteína con, sustancialmente, las mismas propiedades, o presentan sustancialmente las mismas propiedades o actividades reguladoras o funcionales distintas. Las propiedades o actividades pueden ensayarse para su uso en técnicas convencionales conocidas en la materia.

20 En el presente documento también se desvelan secuencias de nucleótidos que comprenden elementos genéticos funcionales tales como promotores, regiones operadoras promotoras, potenciadores, otras secuencias reguladoras, etc. Por tanto, la molécula de ácido nucleico de la invención no comprende necesariamente todo el grupo de genes sino que puede comprender una porción o parte de la que es una parte que codifica un polipéptido que tiene una función particular y opcionalmente una secuencia reguladora. Preferentemente ésta comprende uno o más genes y opcionalmente secuencias reguladoras. En el presente documento se desvelan secuencias que son más cortas que ésta y que codifican uno o más módulos, o dominios enzimáticos, o elementos genéticos funcionales no codificantes o codificantes (por ejemplo, elementos que controlan la expresión, la transcripción, la traducción, etc. de genes).

25 Por tanto, en la invención se extiende a una molécula de ácido nucleico que comprende una secuencia de nucleótidos seleccionada de SEC ID N°: 2, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 y 44 (identificada por referencia a las posiciones de nucleótidos de inicio y fin en la SEC ID N°: 1 como se muestra en la Tabla 1) o una secuencia de nucleótidos que es complementaria a las mismas.

30 También se proporcionan moléculas de ácido nucleico que comprenden secuencias de nucleótidos que presentan una identidad de secuencia de al menos 85 % (preferentemente al menos 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %) con una cualquiera de SEC ID N°: 2, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 y 44 o una secuencia de nucleótidos que es complementaria a las mismas.

35 La invención también se refiere a una molécula de ácido nucleico que comprende una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos seleccionadas de SEC ID N°: 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45 o una secuencia de nucleótidos que es complementaria a las mismas.

40 También se proporcionan moléculas de ácido nucleico que comprenden secuencias de nucleótidos que codifican una o más secuencias de aminoácidos (es decir polipéptidos) que presentan una identidad de secuencia de al menos 85 % (preferentemente al menos 87 %, 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %) con una cualquiera de SEC ID N°: 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45 o una secuencia de nucleótidos que es complementaria a las mismas.

45 En cada caso, la molécula de ácido nucleico es un gen biosintético o una ORF de BE-14106, como se define en el presente documento.

La identidad de secuencia de aminoácidos o de nucleótidos puede evaluarse mediante cualquier método conveniente. Sin embargo, para determinar el grado de identidad de secuencia entre secuencias, son útiles los programas informáticos que realizan alineamientos múltiples de secuencias, por ejemplo, Clustal W (Thompson, J. D et al., 1994, "CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix Choice". *Nucleic Acids Res* 22: 4673-4680). Para esta finalidad, también son útiles programas que comparan y alinean pares de secuencias, como ALIGN (E. Myers y W. Miller, 1988, "Optical Alignments in Linear Space", *CABIOS* 4: 11-17), FASTA (W.R. Pearson y D.J. Lipman, 1988, "Improved tools for biological sequence analysis", *PNAS* 85:2444-2448, y W.R. Pearson, 1990, "Rapid and sensitive sequence comparison with FASTP and FASTA" *Methods in Enzymology* 183:63-98) y BLAST con huecos (Altschul, S.F., et al., 1997, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs". *Nucleic Acids Res.* 25: 3389-3402). Además, el servidor Dali en el European Bioinformatics Institute ofrece alineamientos basados en estructuras de secuencias de proteínas (Holm, 1993, *J. of Mol. Biology*, 233: 123-38; Holm, 1995, *Trends in Biochemical Sciences*, 20: 478-480; Holm, 1998, *Nucleic Acid Research*, 26: 316-9).

Por ejemplo, la identidad de secuencias de nucleótidos puede determinarse usando el programa BestFit del Genetics Computer Group (GCG) Versión 10 del Paquete Informático de la Universidad de Wisconsin. El programa usa el algoritmo de homología local de Smith y Waterman con los valores por defecto: penalización por creación de huecos = 50, penalización por extensión de huecos = 3, promedio de emparejamiento = 10.000, promedio de emparejamiento erróneo = -9.000.

Las secuencias de nucleótidos de acuerdo con la invención pueden presentar una identidad de secuencia de al menos 85 %, 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % con la SEC ID N°: 1 y dichas secuencias codifican o son complementarias a una secuencia que codifica alguna o todas las proteínas que están implicadas en la biosíntesis de la molécula BE-14106. Las secuencias de nucleótidos que cumplen los criterios de % de identidad de secuencia definidos en el presente documento pueden considerarse como secuencias "sustancialmente idénticas".

Un aspecto adicional de la invención proporciona polipéptidos codificados por una molécula de ácido nucleico de la invención, como se define en el presente documento.

Como se ha indicado anteriormente, la SEC ID N°: 1 codifica diversas proteínas o polipéptidos y como tal este aspecto de la invención proporciona un polipéptido que comprende:

(a) todas las secuencias de aminoácidos como se muestra en una cualquiera o más de las SEC ID N°: 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45;

(b) todas las secuencias de aminoácidos que tienen una identidad de secuencia de al menos 85 % con una cualquiera o más de las SEC ID N°: 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45 en las que dicho polipéptido tiene una actividad funcional en la síntesis de BE-14106.

En particular la secuencia de aminoácidos puede presentar una identidad de secuencia de al menos 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % con el polipéptido de una cualquiera de las SEC ID Nos, 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45 en la que dicho polipéptido tiene una actividad funcional en la síntesis de BE-14106.

Las secuencias de aminoácidos (polipéptido) que cumplen con los criterios de % de identidad de secuencia o similitud del presente documento se consideran como "sustancialmente idénticas".

El polipéptido de la invención puede ser un polipéptido aislado, purificado o sintetizado. El término "polipéptido" se usa en el presente documento para incluir cualquier secuencia de aminoácidos de dos o más aminoácidos, es decir, se incluyen péptidos de longitudes tanto cortas como más largas (es decir, polipéptidos o proteínas).

Anteriormente se mencionan programas que determinan la identidad de secuencia de aminoácidos, por ejemplo, la identidad de secuencia o similitud de aminoácidos puede determinarse usando el programa BestFit del Genetics Computer Group (GCG) Versión 10, paquete informático de la Universidad de Wisconsin. El programa usa el algoritmo de homología local de Smith y Waterman con los valores por defecto: penalización por creación de hueco -8, penalización por extensión de hueco = 2, promedio de emparejamiento = 2,912, promedio de emparejamiento erróneo = -2,003. Una "parte" de la secuencia de aminoácidos de una cualquiera de las SEC ID Nos 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 o 45 (o una secuencia "sustancialmente idéntica" como se define anteriormente) puede comprender al menos 20, 30, 40, 50, 70, 100, 150, 200, 300, 400, 500, 1.000, 2.000, 5.000 o 10.000 aminoácidos contiguos.

El polipéptido es funcionalmente activo de acuerdo con las definiciones dadas anteriormente, por ejemplo, es enzimáticamente activo o tiene una actividad funcional reguladora o transportadora en la biosíntesis BE-14106 o un derivado de BE-14106. Una parte de un polipéptido puede corresponder a, o comprender, un dominio o sitio activo o un módulo como se indica anteriormente.



Las secuencias de nucleótidos y polipéptidos de la invención se han caracterizado y se han identificado diversas regiones funcionales en las mismas. En el presente documento se desvelan “partes” que corresponden a al menos un módulo o dominio enzimático, o elemento genético funcional no codificante o codificante. La siguiente Tabla 2 muestra las regiones funcionales identificadas con los productos de traducción de las ORF identificados en la SEC ID N° 1 que codifican enzimas PKS.

Tabla 2

Límites de dominio en las proteínas PKS BE-14016

BecA (SEC ID N° 5)

Molécula: Proteína BecA, 5582 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	17	436	KSq	Dominio similar a KS, módulo de carga
REGIÓN	543	858	AT0	Dominio AT, módulo de carga, propionato
REGIÓN	933	1004	ACP0	Dominio ACP, módulo de carga
REGIÓN	1027	1450	KS1	Dominio KS, módulo 1
REGIÓN	1561	1879	AT1	Dominio AT, acetato, módulo 1
REGIÓN	1892	2095	DH1	Dominio DH, módulo 1
REGIÓN	2415	2662	KR1	Dominio KR, módulo 1
REGIÓN	2698	2771	ACP1	Dominio ACP, módulo 1
REGIÓN	2795	3220	KS2	Dominio KS, módulo 2
REGIÓN	3331	3649	AT2	Dominio AT, acetato, módulo 2
REGIÓN	3663	3868	DH2	Dominio DH2, módulo 2
REGIÓN	4189	4435	KR2	Dominio KR, módulo 2
REGIÓN	4472	4545	ACP2	Dominio ACP, módulo 2
REGIÓN	4572	4995	KSx	Dominio KS, módulo 3
REGIÓN	5103	5416	ATx	Dominio AT, acetato, módulo 3

BecB (SEC ID N° 13)

Molécula: Proteína BecB, 3527 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	10	90	ACP1	Dominio ACP, módulo 1
REGIÓN	112	537	KS2	Dominio KS, módulo 2
REGIÓN	645	961	AT2	Dominio AT, acetato, módulo 2
REGIÓN	975	1169	DH2	Dominio DH, módulo 2
REGIÓN	1424	1672	KR2	Dominio KR, módulo 2
REGIÓN	1709	1782	ACP2	Dominio ACP, módulo 2
REGIÓN	1801	2225	KS3	Dominio KS, módulo 3
REGIÓN	2332	2651	AT3	Dominio AT, propionato, módulo 3
REGIÓN	2665	2865	DH3	Dominio DH, módulo 3
REGIÓN	3132	3374	KR3	Dominio KR, módulo 3
REGIÓN	3411	3485	ACP3	Dominio ACP, módulo 3

BecC (SEC ID N° 9)

Molécula: Proteína BecC, 694 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	346	600	KR	Dominio KR, módulo 3
REGIÓN	615	689	ACP	Dominio ACP, módulo 3

BecD (SEC ID N° 29)

Molécula: Proteína BecD, 3372 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	38	447	KS4	Dominio KS, módulo 4
REGIÓN	561	861	AT4	Dominio AT, acetato, módulo 4
REGIÓN	875	1070	DH4	Dominio DH, módulo 4
REGIÓN	1321	1570	KR4	Dominio KR, módulo 4
REGIÓN	1586	1659	ACP4	Dominio ACP, módulo 4
REGIÓN	1678	2089	KS5	Dominio KS, módulo 5
REGIÓN	2173	2466	AT5	Dominio AT, acetato, módulo 5
REGIÓN	2480	2672	DH5	Dominio DH, módulo 5
REGIÓN	2934	3178	KR5	Dominio KR, módulo 5
REGIÓN	3217	3290	ACP5	Dominio ACP, módulo 5

5

10

## BecE (SEC ID N° 37)

Molécula: Proteína BecE, 1631 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	34	445	KS6	Dominio KS, módulo 6
REGIÓN	529	822	AT6	Dominio AT, acetato, módulo 6
REGIÓN	836	984	DH6i	Dominio DH, módulo 6, delección C-term, probablemente inactiva
REGIÓN	1201	1448	KR6	Dominio KR, módulo 6
REGIÓN	1476	1550	ACP6	Dominio ACP, módulo 6

## BecF (SEC ID N° 35)

Molécula: Proteína BecF, 3377 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	37	447	KS7	Dominio KS, módulo 7
REGIÓN	558	864	AT7	Dominio AT, propionato, módulo 7
REGIÓN	878	1079	DH7	Dominio DH, módulo 7
REGIÓN	1341	1585	KR7	Dominio KR, módulo 7
REGIÓN	1618	1691	ACP7	Dominio ACP, módulo 7
REGIÓN	1710	2121	KS8	Dominio KS, módulo 8
REGIÓN	2203	2496	AT8	Dominio AT, acetato, módulo 8
REGIÓN	2510	2702	DH8	Dominio DH, módulo 8
REGIÓN	2944	3186	KR8	Dominio KR, módulo 8
REGIÓN	3221	3294	ACP8	Dominio ACP, módulo 8

## BecG (SEC ID N° 33)

Molécula: Proteína BecG, 1986 aa

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	35	445	KS9	Dominio KS, módulo 9
REGIÓN	553	847	AT9	Dominio AT, acetato, módulo 9
REGIÓN	861	1066	DH9	Dominio DH, módulo 9
REGIÓN	1341	1586	KR9	Dominio KR, módulo 9
REGIÓN	1616	1690	ACP9	Dominio ACP, módulo 9
REGIÓN	1767	1986	TE	Dominio TE

- 5 La ruta de la biosíntesis de BE-14106 se ha aclarado de la siguiente manera y se muestra en las Figuras 2A y 2B.

La biosíntesis de BE-14106 comienza con el ensamblaje de una fracción acilo C17-C25 por las proteínas BecA (que tiene la secuencia SEC ID N°: 5 y está codificada por la SEC ID N°: 4) y BecC (que tiene la secuencia SEC ID N°: 9 y se codifica por la SEC ID N°: 8) de 1 unidad de propionato y 2 de acetato (Figura 2A). El dominio KR en BecC es más probablemente inactivo, dejando un grupo carbonilo en el C19 de la cadena de polipéptido sintetizada. La biosíntesis de este aminoácido iniciador continúa con la activación de la glicina mediante el dominio de adenilación NRPS distinto BecL (que tiene la secuencia de SEC ID N°: 21 y está codificada por la SEC ID N°: 20) y la carga de la glicina activada sobre la proteína transportadora peptídica distinta BecS (que tiene la secuencia de SEC ID N°: 19 y está codificada por la SEC ID N°: 18). Se piensa que BecU (que tiene la secuencia de SEC ID N°: 11 y se codifica por la SEC ID N°: 10) media la interacción entre el ACP de BecC y BecS, acercando entre sí los sustratos. Presumiblemente, la D-aminoácido oxidasa BecI cataliza la desaminación oxidativa de la glicina, liberando amoníaco, que constituye un ataque nucleófilo en el carbonilo C-17 para dar un grupo C-17 imino, que posteriormente se reduce a un grupo amino. La última reducción supuestamente conduce a la oxidación de la cadena acilo y a la migración de dobles enlaces. La BecQ tioesterasa (que tiene la secuencia de SEC ID N°: 41 y está codificada por la SEC ID N°: 40) libera la cadena aminoácido del complejo BecC-BecU-BecS, dando como resultado la formación de un aminoácido-carboxílico. La supuesta acil-CoA ligasa, BecJ (que tiene la secuencia de SEC ID N°: 15 y está codificada por la SEC ID N°: 14), se supone que inactiva aminoácido-carboxílico resultante a través de la adenilación y posterior ligamiento con la CoA, haciendo que la aminoácido-CoA sea un sustrato aceptable para la aciltransferasa Beck (que tiene la secuencia de SEC ID N°: 17 y está codificada por la SEC ID N°: 16). Una aciltransferasa Beck distinta transfiere la cadena aminoácido activada al dominio ACP en el módulo 1 en BecB PKS, que tiene la secuencia de SEC ID N°: 13 y está codificada por la SEC ID N°: 12. La última PKS carece de todos los dominios en el módulo 1 excepto para la ACP (módulo de carga de la Figura 2B). Los módulos 2 y 3 (módulos 1 y 2 en la Figura 2B) de BecB alargan la fracción aminoácido de C17 con 1 unidad de acetato y 1 de propionato (C16-15 y C14-C13), respectivamente.

Después, la cadena en crecimiento pasa a BecD, que tiene la secuencia SEC ID N°: 29 y está codificada por la SEC ID N°: 28, para prolongación y modificación adicional, BecD que contiene los módulos 4 y 5 (módulos 3 y 4 en la Figura 2B), que alargan la cadena con 2 unidades de acetato (C12-11 y C10-9).

Después, la cadena pasa a BecE que tiene la secuencia SEC ID N°: 37 y está codificada por la SEC ID N°: 36. El módulo 6 (módulo 4 en la Figura 2B) de PKS BecE alarga la cadena con una unidad de acetato (C8-7). El hecho de que el dominio DH en este módulo esté inactivado es responsable de la aparición del grupo C-9 hidroxilo. (Se piensa que el dominio DH contiene una deleción en la región C-terminal que elimina la actividad).

5 Después, la cadena pasa a BecF que tiene la secuencia SEC ID N°: 35 y está codificada por la SEC ID N°: 34. Los módulos 7 y 8 (módulos 6 y 7 en la Figura 2B) de PKS BecF alargan la cadena con una unidad de propionato (C6-5) y una de acetato (C4-3), respectivamente.

10 Después, la cadena pasa a BecG que tiene la secuencia SEC ID N°: 33 y está codificada por la SEC ID N°: 32. El módulo 9 (módulo 8 en la Figura 2B) de PKS BecG extiende la cadena con una unidad de acetato y el dominio TE de BecG es responsable de la hidrólisis del enlace tioéster y de la liberación de la cadena del policétido amida completa de la PKS. Esto causa la ciclación del anillo de macrolactama, probablemente con la ayuda del supuesto homólogo de la prolina iminopeptidasa BecP (que tiene la secuencia SEC ID N°: 31 y está codificada por la SEC ID N°: 30). La biosíntesis se completa mediante la P450 monooxigenasa BecO (que tiene la secuencia SEC ID N°: 27 y está codificada por la SEC ID N°: 26), que hidroxila 8-desoxi BE-14106 en C-8.

20 En la biosíntesis de BE-14106 no puede definirse ningún papel claro para BecT (que tiene la secuencia SEC ID N°: 39 y está codificada por la SEC ID N°: 38). BecR (que tiene la secuencia SEC ID N°: 43 y está codificada por la SEC ID N°: 42) no está implicada en la biosíntesis de BE-14106, demostrado por el experimento de inactivación de genes (Ejemplo 10).

25 Las proteínas que se piensa que están implicadas en la regulación de la ruta incluyen la proteína BecH de tipo LuxR (que tiene la secuencia SEC ID N°: 3 y está codificada por la SEC ID N°: 2) y el regulador BecM de tipo TetR (que tiene la secuencia SEC ID N°: 23 y está codificada por la SEC ID N°: 22).

Se piensa que la bomba de salida BecN de tipo MFS (que tiene la secuencia SEC ID N°: 25 y está codificada por la SEC ID N°: 24) es la responsable de la salida/resistencia de BE-14106.

30 Las secuencias de nucleótidos de la presente invención proporcionan importantes herramientas e información que pueden utilizarse en diversas formas para manipular la biosíntesis de BE-14106, para sintetizar nuevos derivados o análogos de BE-14106 o nuevas moléculas o estructuras de macrolactama y proporcionar maquinaria biosintética de PKS nueva o modificada para la biosíntesis de dichas nuevas moléculas o estructuras de macrolactama. Por maquinaria biosintética PKS se entiende un grupo de proteínas (por ejemplo, codificadas por un grupo de genes) que comprende una o más PKS que pueden informar un complejo de proteínas, un conjunto o ensamblaje, que es funcional en la síntesis de policétidos, pero que no está necesariamente limitada solo a la presencia de enzimas o dominios enzimáticos PKS, y que pueden contener también otras actividades funcionales, por ejemplo, otras proteínas enzimáticas (por ejemplo, modificadoras) o funcionales transportadoras o reguladoras. Las proteínas codificadas por el grupo de genes puede por tanto considerarse como un "sistema enzimático" o "complejo enzimático" o "sistema de proteínas" o "complejo de proteínas" sin implicar necesariamente que las proteínas en el sistema estén físicamente asociadas de ningún modo. Están asociadas "funcionalmente" en el sentido de constituir la maquinaria biosintética para BE-14106. Como alternativa puede denominarse "sistema biosintético" o "complejo biosintético" o "ensamblaje".

45 Por tanto, por ejemplo, todo el grupo de genes biosintético o maquinaria biosintética o un constituyente enzimático o un complejo enzimático de BE-14106, como se proporciona en el presente documento, o una parte del mismo, puede someterse a modificación. La modificación se realiza modificando uno o más genes u ORF en el grupo de genes para producir la modificación de una o más proteínas o péptidos codificados (por ejemplo, enzimas o módulos, o dominios enzimáticos, o secuencias funcionales dentro de la proteína/péptido o enzima codificados). Por tanto, la actividad enzimática puede alterarse o inactivarse para dar como resultado la modificación en la molécula (por ejemplo, estructura de macrolactama) que se sintetiza. Dicha maquinaria biosintética NRPS-PKS modificada o derivatizada puede por tanto usarse para sintetizar fracciones de macrolactama nuevas o modificadas, como se describirá con más detalle a continuación. En esta situación, el grupo de genes biosintético o complejo enzimático o grupos de enzimas de BE-14106 proporcionados en el presente documento puede funcionar como sistema o secuencia "originario" o "molde" o "fuente" o para la modificación. Por tanto, la maquinaria biosintética de NRPS-PKS puede verse como un sistema biosintético NRPS-PKS o "sistema NRPS-PKS".

60 Como se describe con más detalle a continuación, en una realización la modificación del grupo de genes puede tener lugar *in situ*. En otras palabras, el grupo de genes endógeno que está contenido en un microorganismo que produce BE-14106 puede modificarse, por ejemplo por reemplazo de genes o inactivación de genes. Por tanto, el grupo de genes nativo, tal como se produce de manera natural en un microorganismo, puede modificarse. Aunque la expresión recombinante de una molécula de ácido nucleico de la invención es una posibilidad (es decir, la introducción de la molécula de ácido nucleico en una célula hospedadora (por ejemplo una célula hospedadora heteróloga) y el cultivo (o crecimiento) de esa célula hospedadora en condiciones que permitan expresar la molécula de ácido nucleico y que se produzca la molécula de macrolactama (es decir, condiciones que permitan la expresión del producto de la molécula de ácido nucleico para sintetizar la molécula de macrolactama), es menos preferida. En

dicho sistema de expresión recombinante, la molécula de ácido nucleico puede someterse a modificación antes de introducirse en la célula hospedadora y expresarse.

5 De acuerdo con la invención y como se describe adicionalmente más adelante, las partes no funcionales (por ejemplo, partes no biológicamente activas, por ejemplo, partes no codificantes) de dicho sistema (es decir, el grupo de genes o complejo de proteína o ensamblaje o maquinaria biosintética) pueden utilizarse como un “almazón”, y dejar las partes no modificadas y funcionales (por ejemplo, secuencias que codifican partes enzimáticas) que pueden modificarse para producir el sistema NRPS-PKS derivado o modificado. En realizaciones preferidas sólo puede modificarse una única región (por ejemplo, enzimática), módulo o dominio funcional seleccionado, o algunas, dejando la secuencia o estructura restante en gran parte intacta

15 Dentro del ámbito de la invención se incluyen policétido sintetas recombinantes o sintéticas u otras enzimas y complejos o sistemas que contienen dichas enzimas, u otras proteínas de la maquinaria biosintética, es decir, enzimas o proteínas o complejos o sistemas procedentes del armazón codificado por el grupo de genes biosintético de BE-14106 que se modifican para cambiar las propiedades de al menos una proteína codificada por el grupo de genes biosintético de BE-14106.

20 Por ejemplo, dicha modificación podría incluir secuencias que codifiquen una o más unidades funcionales (por ejemplo, módulos o dominios o incluso genes /ORF en su totalidad) procedentes de otras enzimas modulares. Como alternativa, dicha modificación podría introducir secuencias que codifiquen una o más unidades funcionales procedentes del grupo de genes biosintético BE-14106 pero que se encuentren en una localización diferente en la secuencia de origen natural. La modificación también puede ser una modificación que dé como resultado la inactivación o deleción de un dominio, módulo, enzima u otra unidad funcional codificada en la maquinaria biosintética BE-14106.

25 Dichas unidades funcionales pueden ser catalíticas o un dominio de proteína transportadora o reguladora. Para realizar dichas manipulaciones, las secuencias que se usan pueden proceder de la molécula de ácido nucleico de la invención, o las secuencias apropiadas que codifican dominios pueden proceder de moléculas de nucleótidos que codifican otras enzimas que sintetizan policétidos, enzimas que sintetizan péptidos, enzimas que sintetizan policétidos péptidos híbridos, enzimas que sintetizan ácidos grasos, u otros dominios enzimáticos conocidos en la técnica.

30 Por tanto, en un sentido muy general, la presente invención proporciona el uso de las moléculas de ácido nucleico de la invención como se define en el presente documento en la preparación de un grupo de genes biosintético de BE-14106 modificado, de la maquinaria biosintética codificada (o sistema proteico) y de las moléculas modificadas resultantes que se sintetizan de las mismas.

35 La secuencia de nucleótidos de las moléculas de ácido nucleico de la invención pueden utilizarse de este modo de acuerdo con la presente invención de una manera al azar o dirigida o diseñada, por ejemplo, para obtener y ensayar una estructura particular predeterminada o pre-diseñada, o para crear moléculas al azar, por ejemplo bibliotecas de estructuras de policétidos, por ejemplo, para explorar.

40 Por molécula de BE-14106 o derivado de BE-14106 modificado se entiende que la estructura química de BE-14106 se ha cambiado con respecto a la que se representa en la Figura 1. Dichas modificaciones pueden alterar las propiedades funcionales de BE-14106, de tal manera que la intensidad o la fuerza de la molécula (por ejemplo, efectos citotóxicos o actividad antiproliferativa) se potencia o reduce, por ejemplo, o pueden influir en la toxicidad, solubilidad, farmacocinética, afinidad selectiva con respecto a una diana (o dianas) celular u otras propiedades del compuesto.

45 Como alternativa, las modificaciones descritas en el presente documento pueden usarse para cambiar la manera en la que se transporta la molécula BE-14106 o se regula su biosíntesis.

50 En el presente documento se describe un grupo de genes biosintético nuevo o modificado, incluyendo la maquinaria biosintética codificada las proteínas constituyentes o el sistema de proteínas y las moléculas modificadas resultantes que se sintetizan a partir del mismo. En el presente documento también se describen células en las que se ha introducido un grupo de genes biosintético o células que contienen o comprenden dicho grupo de genes biosintético modificado o la maquinaria biosintética codificada modificada.

55 Los genes o elementos genéticos que pueden modificarse no solo incluyen los genes o las ORF de PKS reales (que codifican BecA, BecB, BecC, BecD, BecE, BecF, BecG) o los módulos individuales o sus dominios, sino también genes u ORF que codifican otras enzimas o proteínas funcionales implicadas en la biosíntesis de BE-14106 (por ejemplo que codifican BecU, BecI, BecP, BecJ, BecK, BecS, BecL, BecO, BecT, BecQ) y transporte (por ejemplo que codifican BecN) o regulación (por ejemplo que codifican BecH, BecM), todos los cuales se denominan en el presente documento en su conjunto “genes biosintéticos de BE-14106”.

60

Con relación a los genes de PKS reales, como se describirá con más detalle más adelante, éstos pueden modificarse para cambiar la naturaleza de un dominio de módulo de carga que determina la naturaleza de la unidad iniciadora, el número de módulos, la naturaleza del extendedor, así como las diversas actividades deshidratasa, reductasa y sintasa que determinan la estructura de la cadena del policétido.

Por tanto, por ejemplo, el número de módulos de una enzima PKS puede aumentarse o disminuirse, por ejemplo, pueden delecionarse uno o más módulos de una enzima PKS; los dominios KS, DH y/o KR de una enzima PKS pueden deleccionarse, inactivarse o introducirse (por ejemplo del mismo gen de PKS o de otro gen de PKS), un dominio AT de una enzima PKS puede modificarse para alterar su especificidad por una unidad iniciadora o extendedora; el dominio KS<sup>Q</sup> de BecA (SEC ID N° 4/5) puede modificarse (por ejemplo inactivarse) para alterar la naturaleza de la unidad iniciadora que constituirá parte de la cadena lateral.

El cambio de especificidad de la BecO hidroxilasa puede posibilitar la hidroxilación en los átomos de carbono de BE-14106 (o sus análogos) distintos de C-8.

Además de la modificación de las enzimas NRPS o PKS para cambiar la naturaleza de las moléculas sintetizadas, las moléculas de ácido nucleico de la invención también pueden utilizarse para manipular o facilitar el proceso biosintético, por ejemplo, extendiendo el intervalo del hospedador o aumentando la eficiencia del rendimiento o de la producción etc.

Para permitir la expresión recombinante de una molécula de ácido nucleico de la invención, la invención contempla un vector, por ejemplo un vector de clonación o expresión, que comprende el ácido nucleico de la invención y una célula hospedadora que contiene dicha molécula.

En la práctica, la modificación puede realizarse ventajosamente manipulando la secuencia del grupo de genes biosintético de BE-14106 *in situ* en una célula (que puede considerarse en el presente documento como "una célula hospedadora") por ejemplo para alterar la secuencia de nucleótidos que codifica una actividad enzimática, para inactivar o modificar esa actividad o para introducir una actividad. Pueden generarse construcciones genéticas apropiadas que contengan secuencias que tengan las modificaciones necesarias que van a introducirse en la secuencia del grupo de genes biosintético de BE-14106 (por ejemplo en la SEC ID N° 1) como las contenidas en una molécula de ácido nucleico endógena (en otras palabras una molécula de ácido nucleico contenida en la célula hospedadora). La introducción de un vector (por ejemplo, un plásmido) con esta secuencia en la célula hospedadora apropiada puede realizarse después. Finalmente, esto conduce a la integración de la secuencia modificada en el grupo de genes (por ejemplo dentro del genoma de la célula hospedadora) cerca de la parte correspondiente (por ejemplo endógena o de origen natural) del grupo de genes biosintético mediante recombinación homóloga. Después de un segmento de recombinación, la parte endógena del grupo de genes se reemplaza por la versión modificada y el vector se elimina.

La célula hospedadora modificada resultante contendrá por lo tanto un grupo de genes biosintético de BE-14106 modificado, que codifica un sistema enzimático BE-14106 modificado. La maquinaria biosintética BE-14106 modificada por tanto sintetiza una molécula BE-14106 modificada.

Como se ha indicado anteriormente, el grupo de genes que se modifica puede ser el grupo de genes nativo que está presente de manera natural en una célula hospedadora (microorganismo) que produce BE-14106 (y por tanto se modifica un ácido nucleico nativo o una molécula de ácido nucleico endógena). Menos preferentemente, pero aún dentro en la presente invención, la molécula de ácido nucleico de la invención puede introducirse en la célula hospedadora antes de la modificación. Como alternativa, la molécula de ácido nucleico se introduce en la célula hospedadora después de la modificación. Por tanto, puede modificarse una molécula de ácido nucleico exógena.

Dado que la invención proporciona la secuencia del grupo de genes de BE-14106 de longitud completa, esta estrategia de reemplazo de genes es de aplicación general para modificar el grupo de genes de BE-14106. La estrategia puede usarse, por ejemplo, para deleccionar una parte del grupo de genes, para introducir actividades o sustituir actividades encontradas dentro de las moléculas en la secuencia nativa o de tipo silvestre. La estrategia requiere el conocimiento de la secuencia del grupo de genes pero para realizar la manipulación no se requiere que todo el grupo de genes esté necesariamente aislado de una célula hospedadora.

Por tanto, en un aspecto adicional la presente invención proporciona un método para preparar una molécula de ácido nucleico que codifique una maquinaria biosintética BE-14106 NRPS-PKS modificada (o un sistema BE-14106 NRPS-PKS modificado), comprendiendo dicho método modificar una molécula de ácido nucleico de la invención que codifique dicha maquinaria biosintética BE-14106 NRPS-PKS (o un sistema BE-14106 NRPS-PKS modificado).

La molécula de ácido nucleico se modifica modificando su secuencia, y más particularmente la molécula de ácido nucleico puede modificarse introduciendo, mutando, deleccionando, reemplazando o inactivando una secuencia que codifique una o más actividades (o proteínas) codificadas por dicha molécula de ácido nucleico. Por tanto, pueden modificarse una o más secuencias que codifiquen actividades enzimáticas u otras actividades funcionales. Dicha modificación da como resultado una molécula de ácido nucleico que codifica una maquinaria biosintética BE-14106

5 NRPS-PKS (o un sistema BE-14106 NRPS-PKS) que tiene la función o la actividad o las propiedades alteradas en comparación con la maquinaria biosintética BE-14106 NRPS-PKS nativa o de tipo silvestre (o sistema BE-14106 NRPS-PKS). Por tanto, la maquinaria biosintética modificada (o sistema NRPS-PKS) puede tener una o más actividades enzimáticas alteradas o modificadas y puede dar como resultado la síntesis de una molécula (por ejemplo, una molécula de policétido o de macrolactama) que es distinta (o diferente a) de la molécula sintetizada a partir de la maquinaria biosintética (o sistema NRPS-PKS) nativa (es decir, no modificada). Como alternativa, como se ha indicado anteriormente, la modificación de la maquinaria biosintética puede dar como resultado una mejora en el proceso biosintético, por ejemplo, un rendimiento aumentado, etc.

10 El ácido nucleico que se modifica puede estar contenido dentro de una célula u organismo, que puede ser la célula u organismo usado para la producción de la molécula de macrolactama que sintetiza la maquinaria biosintética modificada.

15 Como se ha indicado anteriormente, la molécula de ácido nucleico que se modifica puede ser ventajosamente la molécula de ácido nucleico que está presente endógenamente (o naturalmente) en un organismo que produce BE-14106. Por tanto, el método de la invención puede implicar modificar *in situ* una molécula de ácido nucleico nativa (más particularmente modificando la secuencia de una molécula de ácido nucleico) dentro de una célula u organismo (generalmente una célula microbiana o un microorganismo) que produce BE-14106. La molécula de ácido nucleico es o representa el grupo de genes biosintético BE-14106 o una parte del mismo y puede por tanto verse como una molécula de ácido nucleico que codifique la maquinaria biosintética de BE-14106 o el sistema BE-14106 NRPS-PKS o una parte del mismo.

20 Diversos microorganismos diferentes pueden producir BE-14106 y adicionalmente se sabe que algunos microorganismos pueden producir derivados de origen natural de BE-14106 tales como derivado 8-desoxy (Takahashi *et al.*, citados anteriormente). La referencia a un grupo de genes BE-14106 o maquinaria biosintética (o sistema NRPS-PKS, etc.) pretende incluir grupos de genes o maquinaria biosintética, etc. que produzcan no solo la propia BE-14106 sino también derivados de origen natural tales como el derivado 8-desoxy (denominado GT32-B).

25 La invención proporciona adicionalmente un método para preparar una maquinaria biosintética BE-14106 NRPS-PKS modificada (o un sistema BE-14106 NRPS-PKS modificado), comprendiendo dicho método preparar una molécula de ácido nucleico como se define anteriormente y expresar dicha molécula de ácido nucleico en un microorganismo. Esto puede realizarse modificando simplemente la molécula de ácido nucleico contenida en la célula, como se ha descrito anteriormente, y permitiendo que la célula crezca en condiciones en las que se exprese la molécula de ácido nucleico modificada. Por tanto, por ejemplo, la molécula de ácido nucleico nativa en la célula puede modificarse *in situ* y se puede dejar que la célula crezca.

30 En el presente documento se describe una maquinaria biosintética BE-14106 NRPS-PKS modificada (o sistema BE-14106 NRPS-PKS modificado) obtenida mediante dicho método.

35 La invención también proporciona un método para preparar una molécula de macrolactama modificada, comprendiendo dicho método preparar una molécula de ácido nucleico modificada como se ha definido anteriormente, y expresar dicha molécula de ácido nucleico en un microorganismo.

40 Generalmente hablando, la molécula de ácido nucleico se expresará en una célula hospedadora en condiciones en las que la maquinaria biosintética modificada pueda expresarse. Como se ha indicado anteriormente, esto puede realizarse introduciendo la molécula de ácido nucleico en una célula hospedadora, pero generalmente la "célula hospedadora" será la célula u organismo en el que la molécula de ácido nucleico esté presente de manera natural o endógena y en la que se modifique la molécula de ácido nucleico. La célula hospedadora crecerá o se cultivará en condiciones que permitan que se exprese la molécula de ácido nucleico y la maquinaria biosintética modificada, y que se sintetice la molécula producida a partir de la maquinaria biosintética.

45 Por tanto, la molécula de ácido nucleico puede expresarse en cualquier célula hospedadora deseada, pero preferentemente se expresará en la célula o microorganismo del cual procede (o del cual puede proceder) y en el que la molécula esté presente de manera natural (no modificada).

50 El método de la invención para preparar una molécula de macrolactama modificada puede incluir la etapa adicional de recuperar (por ejemplo aislar o purificar) la molécula, por ejemplo, del medio de cultivo, en el que se ha desarrollado la célula hospedadora, o de la célula hospedadora. Por tanto, una molécula de macrolactama modificada obtenida por dicho método también se describe en el presente documento.

55 En el presente documento también se describe una célula o microorganismo que contiene una molécula de ácido nucleico que codifica una maquinaria biosintética BE-14106 NRPS-PKS (o sistema BE-14106 NRPS-PKS modificado) obtenida mediante un método como se define anteriormente en el presente documento. La célula o microorganismo puede contener una maquinaria biosintética BE-14106 NRPS-PKS modificada (o sistema BE-14106 NRPS-PKS modificado) obtenida mediante un método como se ha definido anteriormente.

60

65

La invención también puede proporcionar una célula hospedadora que contenga una molécula de ácido nucleico de la invención, en la que dicha molécula se ha introducido en una célula hospedadora heteróloga.

5 A modo de ejemplo representativo, se contempla que dichas manipulaciones del grupo de genes puedan realizarse con el objetivo de generar una molécula BE-14106 modificada en la que se introduce un grupo hidroxilo en una o más cualquiera de las posiciones C-2, 3, 4, 7, 11, 13, 15, 17, 21 y 23 de BE-14106 o una derivada o versión modificada del mismo. Esto puede realizarse inactivando o delecionando el dominio (o dominios) DH apropiado en el grupo de genes biosintético BE-14106. De acuerdo con la ruta biosintética proporcionada en el presente documento podrían realizarse las siguientes modificaciones:

10

Tabla 3

Posición en la que se introduce un grupo OH	Dominio DH a inactivar/delecionarse
3	dominio DH módulo 9 BecG
5	dominio DH módulo 8 BecF
7	dominio DH módulo 7 BecF
11	dominio DH módulo 5 BecD
13	dominio DH módulo 4 BecD
15	dominio DH módulo 3 BecB
17	dominio DH módulo 2 BecB
23	dominio DH módulo 1 BecA

15 Además, o como alternativa, se contempla que dichas manipulaciones del grupo de genes puedan realizarse con el objetivo de generar una molécula BE-14016 modificada en la que se introduce un grupo oxo (ceto) en una cualquiera o más de las posiciones C-2, 3, 4, 7, 9, 11, 13, 15, 17, 21 y 23 de BE-14106 una versión derivada o modificada del mismo. Esto puede realizarse inactivando o delecionando el dominio (o dominios) KR apropiado en el grupo de genes biosintético BE-14106. De acuerdo con la ruta biosintética propuesta en el presente documento podrán realizarse las siguientes modificaciones:

20

Tabla 4

Posición en la que se introduce un grupo oxo	Dominio KR a inactivar/delecionar
3	dominio KR módulo 9 BecG
5	dominio KR módulo 8 BecF
7	dominio KR módulo 7 BecF
11	dominio KR módulo 5 BecD
13	dominio KR módulo 4 BecD
15	dominio KR módulo 3 BecB
17	dominio KR módulo 2 BecB
23	dominio KR módulo 1 BecA

25 Por tanto, puede producirse un análogo de BE-14106 modificando los genes/proteína de la ruta biosintética de BE-14106. Un análogo de BE-14106 de la invención incluye una molécula que comprende una modificación seleccionada de una cualquiera o más del grupo que comprende 3-, 5-, 7-, 11-, 13-, 15-, 17- o 23-hidroxi BE-14106 o 3-, 5-, 7-, 9-, 11-, 13-, 15-, 17- o 23-oxo BE-14106 o combinaciones de las mismas o un análogo que comprende una combinación de un grupo 8-desoxi con una o más modificaciones seleccionadas de la introducción de un grupo hidroxilo u oxo en cualquiera de las posiciones 3, 5, 7, 11, 13, 15, 17 o 23 o un grupo oxo en la posición 9. En la tabla 5 se muestra una selección de las estructuras de dichos análogos y módulos representativos de BE-14106 que requieren la inactivación para generar dichos análogos.

30

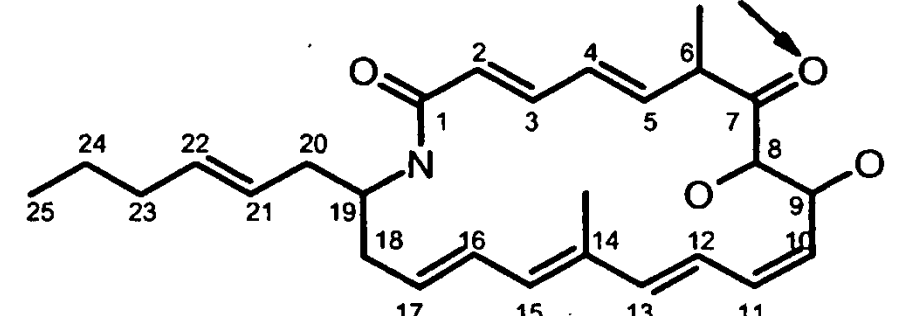
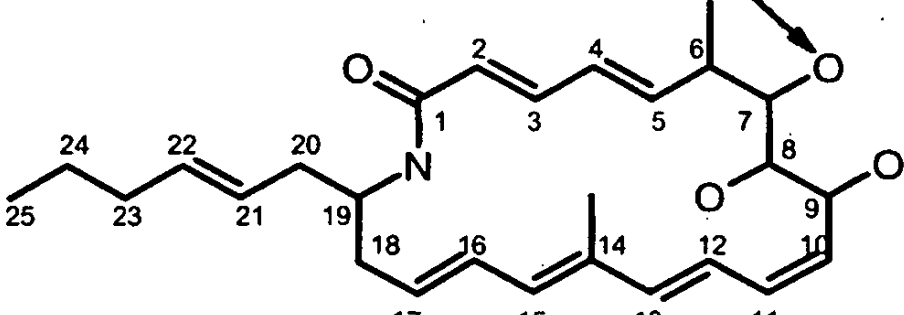
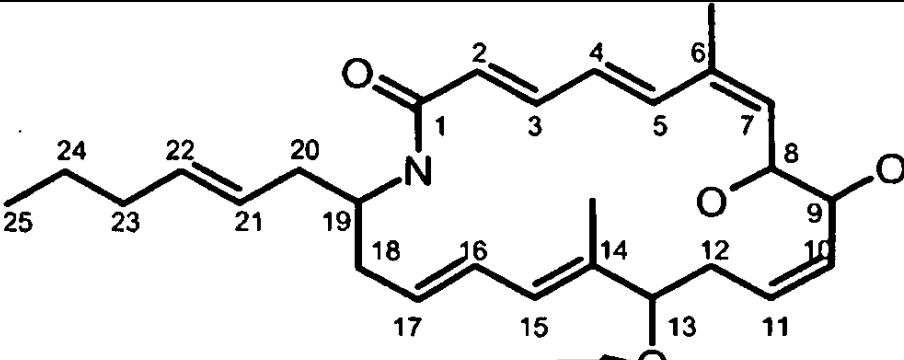
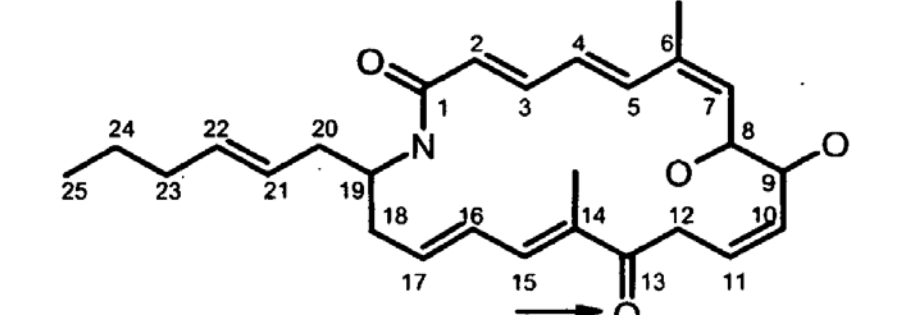
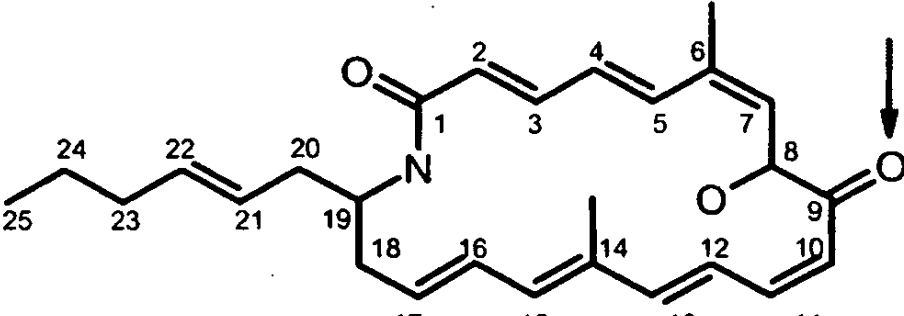
ES 2 543 203 T3

Tabla 5 – Análogos de BE-14106 que pueden producirse después de modificar por ingeniería genética los genes de la policétido sintasa

Nombre	Estructura Mutación	(s)
BE-14106		ninguna
8-desoxi BE-14106		BeCO
(1) 17-hidroxi BE-14106		DH2 (BecB)
(2) 15-hidroxi BE-14106		DH3 (BecB)
(3) 3-oxo BE-14106		KR9 (BecG)



Nombre	Estructura Mutación	(s)
(4) 3-hidroxi BE-14106		DH9 (BecG)
(5) 11-oxo BE-14106		KR5 (BecD)
(6) 11-hidroxi BE-14106		DHR (BecD)
(7) 17-oxo BE-14106		KR2 (BecB)
(8) 15-oxo BE-14106		KR3 (BecB)

Nombre	Estructura Mutación	(s)
(9) 7-oxo BE-14106		KR7 (BecF)
(10) 7- hidroxi BE- 14106		DH7 (BecF)
(11) 13- hidroxi BE- 14106		DH4 (BecD)
(12) 13- oxo BE- 14106		KR4 (BecD)
(13) 9-oxo BE-14106		KR6 (BecE)

Además de, o como una alternativa a las modificaciones anteriores, la cadena lateral C20-C25 puede acortarse por delección de las regiones de ADN en todo el módulo (o módulos) que codifican *becA* de la PKS *BecA*. Como alternativa, una o más moléculas de *BecA* pueden deleccionarse para cortar la cadena lateral. Por ejemplo, la delección de un módulo entero de *BecA* puede acortar la cadena lateral.

5 Además de, o como una alternativa a las modificaciones anteriores, uno o más de los módulos de PKS (por ejemplo de *BecA*, *BecB*, *BecD*, *BecE*, *BecF*, *BecG*) pueden deleccionarse para producir análogos que tengan a nillo de macrolactama que tengan menos miembros (por ejemplo anillos de 18-, 16-, 14-, 12-, 10- y 8- miembros).

10 Como se muestra en la Tabla 5, y además de o como una alternativa a otras modificaciones descritas o tabuladas anteriormente, *BecO* puede inactivarse o deleccionarse para eliminar el grupo hidroxilo en la posición 8.

15 Otras modificaciones que pueden realizarse además de o como alternativa a las modificaciones anteriores incluyen la sustitución del carbonilo C-1 de BE-14106 con tio-carbonilo (C=S) o carb oxamida (C=NH) para aumentar la estabilidad de la molécula (por ejemplo contra la degradación por proteasas a la que la BE-14106 puede ser susceptible debido a la similitud del enlace -N-C=O en BE-14106 con un enlace peptídico); la glucosilación de BE-14106 o de una molécula BE-14106 modificada, por ejemplo con fracciones de monosacáridos; la acilación de los grupos hidroxilo libres, y la reducción del doble enlace C21-C22 que está en la cadena lateral de BE-14106.

20 Las proteínas reguladoras pueden sobre expresarse para aumentar la producción de la molécula sintetizada, por ejemplo BE-14106 o un derivado o análogo de la misma.

25 Como tal, se proporciona un método para producir un grupo de genes biosintético de BE-14106 modificado, en el que en la molécula de ácido nucleico de la invención, como se define en el presente documento, se realizan una o más de las modificaciones expuestas anteriormente. Convenientemente, dichas modificaciones pueden realizarse llevando a cabo reemplazo de genes en una célula hospedadora que contenga endógenamente la molécula de ácido nucleico de la invención.

30 Por "reemplazo de genes" se entiende cualquier método en el que un gen o una ORF o parte de los mismos (por ejemplo un módulo, un dominio u otra unidad funcional), tal como se encuentra en la molécula de ácido nucleico de la invención, se reemplaza eficazmente con, o se sustituye por, una versión modificada de los mismos. La versión modificada de los mismos puede ser una en la que el gen o la ORF o porción o parte de los mismos (por ejemplo la secuencia que codifica la proteína o péptido o dominio o módulo de los mismos) se cambia o altera (por ejemplo se muta tal como por sustitución, delección o inserción de uno o más restos de nucleótidos). Dichos cambios darían como resultado cambios en la actividad del polipéptido codificado, pero también incluirían cambios que dan como resultado una o más proteínas o polipéptidos o dominios o módulos de los mismos que se deleccionan o inactivan. Preferentemente, la versión modificada es una en la que la secuencia que codifica dicha proteína o polipéptido o uno o más módulos o dominios de la misma se inactiva o deleciona.

40 Debe observarse que el método no requiere que todo el gen u ORF esté físicamente retirado y reemplazado, sino que alguna parte, o toda la secuencia, tal como se encuentra en la molécula de ácido nucleico de la invención, se reemplaza con una versión modificada de los mismos.

45 Los reemplazos de genes pueden realizarse de acuerdo con técnicas conocidas en la materia (véase, por ejemplo, Sekurova *et al.*, 1999 FEMS Microbiol. Lett., 177, 297-304).

50 Una célula hospedadora (por ejemplo un microorganismo) que contiene endógenamente la molécula de ácido nucleico de la invención se somete preferentemente a modificación de la molécula de ácido nucleico, por ejemplo, por reemplazo de genes, para cambiar o alterar el grupo de genes biosintético BE-14106 que está presente en dicha célula hospedadora. Una célula hospedadora contiene dógicamente la molécula de ácido nucleico de la invención si dicha molécula de ácido nucleico de la invención se encuentra normalmente en esa célula hospedadora o si se produce de manera natural en esta célula hospedadora (es decir, la molécula de ácido nucleico está presente en el material genético de la célula hospedadora). En otras palabras, la molécula de ácido nucleico de la invención no está presente en la célula hospedadora simplemente en virtud de haberse introducido (por ejemplo, transfectado o transferido) en la célula hospedadora, por ejemplo, en una molécula de ADN recombinante tal como un plásmido.

55 Si una célula hospedadora contiene o no la molécula de ácido nucleico de la presente invención puede averiguarse determinando si la célula hospedadora puede sintetizar BE-14106 o un derivado de la misma. Como alternativa o adicionalmente, puede realizarse técnicas genéticas para analizar las secuencias de las moléculas de ácido nucleico presentes en la célula hospedadora (por ejemplo PCR, transferencia de Southern y otras técnicas convencionales conocidas en la materia). Dichas técnicas genéticas pueden usarse para determinar si la molécula de ácido nucleico está presente endógenamente.

65 La célula hospedadora para su uso en los métodos de la invención puede ser cualquier célula u organismo, procarionota o eucariota, que se desee, pero generalmente será un microorganismo particularmente una bacteria. Más particularmente, la célula hospedadora será un actinomiceto.

Las células hospedadoras preferidas incluyen cepas de *Streptomyces*, que contienen de manera preferida, endógenamente la molécula de ácido nucleico de la invención. Un ejemplo adecuado es la cepa espi-A14106 de *Streptomyces* (FERMP-11378, como se denomina en el documento JP4001179). La nueva cepa aislada nombrada anteriormente, de la cual se secuenció el grupo de genes (aislado MP28-13), depositada en la DSMZ con el número de depósito DSM21069, es particularmente preferida.

Además, se describe un método para producir un polipéptido modificado codificado por el sistema enzimático biosintético BE-14106, en el que un grupo de genes biosintético BE-14106 modificado obtenible por los métodos descritos anteriormente, se expresa en una célula hospedadora. Una vez realizadas las modificaciones apropiadas en el grupo de genes biosintético BE-14106, por ejemplo mediante los métodos expuestos anteriormente, la célula hospedadora puede crecer o cultivarse en condiciones que permitan la expresión de los polipéptidos o proteínas del grupo de genes modificado para producir el polipéptido (o polipéptidos) biosintético BE-14106 modificado.

Como tal, la célula hospedadora modificada que contiene el grupo de genes biosintético BE-14106 modificado expresaba los polipéptidos y las proteínas codificados por el grupo de genes y esto permitirá la producción de un polipéptido o una proteína modificados, así como de las otras proteínas codificadas por el grupo de genes. Como resultado de la presencia de uno o más polipéptidos o proteínas modificados en la maquinaria biosintética, las propiedades de la maquinaria biosintética para BE-14106 cambiarán. Dependiendo de la naturaleza de la modificación (o modificaciones) que se ha realizado en el grupo de genes biosintético BE-14106, las propiedades de los polipéptidos, proteínas y enzimas codificados serán diferentes a las del tipo silvestre.

Para la mayoría de los propósitos, será suficiente la producción de la maquinaria biosintética BE-14106 dentro de la célula. Como alternativa, las proteínas que constituyen la maquinaria biosintética BE-14106 pueden purificarse de la célula en la que se expresan.

En el presente documento también se describe un polipéptido codificado por la molécula de ácido nucleico de la invención que se ha modificado como se describe anteriormente.

También se proporciona un método de producción de una molécula BE-14106 modificada. De acuerdo con este método, un grupo de genes biosintético BE-14106 modificado se prepara mediante los métodos descritos anteriormente y se expresa en una célula hospedadora. Esto puede realizarse desarrollando o cultivando una célula hospedadora en la que esté presente un grupo de genes biosintético BE-14106 modificado preparado por los métodos descritos anteriormente, en condiciones que permitan la expresión de los polipéptidos, proteínas y enzimas codificados por el grupo de genes biosintético BE-14106 modificado. La célula contendrá por tanto la maquinaria biosintética necesaria para la biosíntesis de la molécula BE-14106 modificada y se producirá la síntesis de esta molécula.

El método puede comprender adicionalmente la etapa de recuperar, por ejemplo, aislar o purificar la molécula BE-14106 modificada. Ésta puede aislarse o purificarse del medio de cultivo celular en el que se ha transportado o secretado si fuera apropiado, o de otra manera, de la célula hospedadora en la que se ha incluido. Por tanto, por ejemplo, las células del organismo de producción pueden recogerse, por ejemplo, por centrifugación, y pueden extraerse, por ejemplo, con uno o más disolventes orgánicos (por ejemplo, metanol u otros alcoholes). Las moléculas pueden recuperarse de dicho extracto, por ejemplo, por precipitación. La purificación adicional de un producto en bruto obtenido de esta manera puede incluir, por ejemplo, cromatografía, por ejemplo, HPLC (Cromatografía líquida de alto rendimiento, *High Performance Liquid Chromatography*).

Para posibilitar la realización práctica de la invención de acuerdo con los principios anteriores, la invención también proporciona una célula hospedadora heteróloga que contiene la molécula de ácido nucleico como se define en el presente documento. En el presente documento también se describe una célula que contiene una molécula de ácido nucleico de la invención modificada como se define anteriormente.

En general, los métodos de la invención pueden realizarse en cualquier célula hospedadora que contenga la molécula de ácido nucleico de la invención, preferentemente células hospedadoras que contengan endógenamente la molécula. Como se ha indicado anteriormente, las células hospedadoras preferidas incluyen células de *Streptomyces spp.* Las células más preferidas son células de *Streptomyces* que tienen las características de resistencia a antibióticos y sensibilidad expuestas en la Tabla 6.

En una realización altamente preferida, los métodos se realizan usando la nueva cepa de *Streptomyces* MP28-13 depositada con el número DSM21069 en la DSMZ) o un mutante o una cepa modificada de la misma que produce BE-14106 o un derivado del mismo.

También se describe un microorganismo, particularmente una bacteria y especialmente una *Streptomyces*, que puede obtenerse mediante los métodos descritos en el presente documento.

Un importante aspecto de la invención también proporciona una cepa de *Streptomyces* depositada con el número DSMZ 21069 en la DSMZ, o un mutante o una cepa modificada de la misma que produce BE-14106 o un derivado del mismo.

- 5 Puede observarse que los métodos de la invención dan como resultado la producción de moléculas de ácido nucleico derivadas, procedentes de las moléculas de ácido nucleico de la invención definidas anteriormente.

La molécula de ácido nucleico derivada puede formarse *in situ* en células hospedadoras o microorganismos en los que están incluidas las moléculas de ácido nucleico no modificadas. Como alternativa, pueden formarse en células hospedadoras en las que se introducen las moléculas de ácido nucleico. Dichos microorganismos “modificados” que contienen las moléculas de ácido nucleico derivadas o modificadas de acuerdo con la presente invención pueden usarse para formar bibliotecas, por ejemplo, bibliotecas de policétidos o moléculas basadas en policétidos o macrolactama en las que los miembros de la biblioteca se sintetizan por maquinarias biosintéticas BE-14106 modificadas o sistemas NRPS-PKS procedentes del sistema BE-14106 de origen natural proporcionados en el presente documento. Generalmente, muchos miembros de estas bibliotecas pueden ser, en sí mismos, nuevos compuestos. Por tanto, los métodos pueden dirigirse para la preparación de un compuesto individual. El compuesto puede ser nuevo o no, pero el método de preparación permite un método más conveniente para prepararlo. Los compuestos resultantes (por ejemplo policétidos) pueden modificarse adicionalmente, por ejemplo, para convertirlos en moléculas adicionalmente activas, por ejemplo antibióticos, por ejemplo por glucosilación. En el presente documento se desvelan métodos para recuperar nuevos compuestos con actividades deseadas explorando las bibliotecas.

Las bibliotecas o formas modificadas individuales, finalmente de moléculas basadas en policétidos, o macrolactama, pueden generarse modificando el grupo de genes BE-14106 NRPS-PKS de tal manera que las proteínas producidas por el grupo tengan actividades alteradas en uno o más aspectos, y produzcan por tanto compuestos distintos del producto natural del sistema NRPS-PKS codificado por el grupo de genes (es decir BE-14106). Por tanto, pueden prepararse nuevos compuestos, o en general prepararse compuestos más fácilmente, usando este método. Proporcionando una gran cantidad de genes o grupos de genes diferentes procedentes del grupo de genes BE-14106 de origen natural, cada uno de los cuales se ha modificado de una manera diferente del grupo nativo, puede producirse eficazmente una biblioteca combinatoria de compuestos como resultado de las múltiples variaciones de estas actividades. Las secuencias codificantes de NRPS-PKS modificadas y la maquinaria/sistema de biosíntesis desvelados en el presente documento representan por tanto secuencias codificantes y maquinaria o sistemas de enzimas/proteínas “procedentes de” una maquinaria o sistema biosintético BE-14106 NRPS-PKS de origen natural.

35 Por una maquinaria biosintética o sistema NRPS-PKS “procedente de” la maquinaria biosintética BE-14106 se entiende una maquinaria biosintética o sistema NRPS-PKS en el que al menos una actividad enzimática o funcional está mutada, delecionada, inactivada o reemplazada, para alterar la actividad. La alteración se produce cuando estas actividades están delecionadas o reemplazadas por una versión diferente de la actividad, o simplemente mutadas de tal manera que se obtiene un compuesto (por ejemplo un policétido o macrolactama) distinto del producto natural de estas actividades colectivas. Esto ocurre porque ha habido una alteración resultante de la unidad iniciadora y/o unidad extendedora y/o de la estereoquímica y/o del resultado de la longitud o ciclación de la cadena y/o del ciclo reductor o de deshidratación en una posición correspondiente en el compuesto del producto. Cuando una actividad delecionada se reemplaza, el origen de la actividad del reemplazo puede provenir de una actividad correspondiente en una NRPS o policétido sintasa diferente de origen natural o de una región diferente del mismo sistema/maquinaria NRPS-PKS.

La modificación o manipulación del NRPS-PKS modular puede implicar truncamiento, por ejemplo, deleción de genes o dominios o módulos o translocación, adición o inactivación de dominios/genes/módulos, que puede implicar inserción o deleción. Como alternativa, en la secuencia de nucleótidos de la parte seleccionada (por ejemplo en un gen/dominio/módulo, etc.) pueden realizarse modificaciones al azar o dirigidas (es decir mutaciones).

Ventajosamente, una maquinaria biosintética o sistema NRPS-PKS “procedente de” la maquinaria o sistema BE-14106 contiene al menos un dominio de adición de NRPS y al menos dos módulos de una enzima PKS y opcionalmente puede contener mutaciones, deleciones o reemplazos de una o más de las actividades de estos dominios o módulos funcionales de tal manera que se altere la naturaleza del compuesto resultante. Esta definición se aplica a niveles tanto proteicos como genéticos. Las realizaciones particulares preferidas incluyen aquellas en las que una KS, AT, KR o DH se ha inactivado o delecionado o reemplazado por una versión de la actividad de una maquinaria o sistema PKS/NRPS diferente o de otra localización dentro la misma maquinaria o sistema NRPS-PKS. También se prefieren derivados en los que al menos una actividad enzimática del ciclo de condensación (por ejemplo KR o DH) se ha delecionado o en la que cualquiera de estas actividades se ha mutado para cambiar el compuesto final sintetizado.

Por tanto, hay cinco grados de libertad para construir una maquinaria o sistema biosintético PKS en cuanto al compuesto que se producirá. En primer lugar, la longitud de la cadena del policétido se determinará mediante el número de módulos en la maquinaria o sistema. En segundo lugar, la naturaleza del esqueleto de carbono del policétido se determinará a través de las especificidades de las acil transferasas que determinan la naturaleza de las

unidades extendedoras en cada posición, por ejemplo, malonilo, metil malonilo o etil malonilo, etc. En tercer lugar, la especificidad del dominio de carga también tendrá un efecto sobre el esqueleto de carbono resultante del policétido. Por tanto, el dominio de carga puede usarse como una unidad iniciadora diferente, tal como acetilo, propionilo y similar. En cuarto lugar, el estado de oxidación en diversas posiciones del policétido lo determinará las porciones deshidratasa y reductasa de los módulos. Esto determinará la presencia y localización de acetona, alcohol, enlaces dobles o sencillos en el policétido.

Finalmente, la estereoquímica del policétido resultante es una función de diversos aspectos de la sintasa. El primer aspecto se relaciona con la especificidad AT/KS asociada con malonilos sustituidos como unidades extendedoras, que afecta a la estereoquímica solamente cuando el ciclo reductor desapa parece o cuando contiene solo una cetorreductasa dado que la deshidratasa podría anular la quiralidad. En segundo lugar, la especificidad de la cetorreductasa determinará la quiralidad de cualquier  $\beta$ -OH.

Modificando la PKS implicada en la biosíntesis de aminoácido "iniciador", el compuesto que se produce puede alterarse.

Por tanto, la maquinaria modificada del sistema NRPS-PKS puede permitir sintetizar una amplia variedad de compuestos.

El tamaño del producto sintetizado puede variar modificando el número de módulos.

Los productos de policétido/macrolactama de la maquinaria biosintética modificada pueden modificarse adicionalmente, por ejemplo, por glucosilación u otra derivatización, para presentar o mejorar la actividad, por ejemplo, actividad antibiótica. Los métodos para la glucosilación de policétidos son generalmente conocidos en la técnica; la glucosilación puede efectuarse intracelularmente proporcionando las enzimas de glucosilación apropiadas o puede efectuarse *in vitro* usando medios sintéticos químicos.

Para obtener moléculas de ácido nucleico que codifiquen una variedad de derivados (o análogos) del sistema BE-14106 NRPS-PKS de origen natural, y por tanto una variedad de compuestos basados en policétidos macrolactama, pueden obtenerse diversas construcciones diseñadas "mezclando y emparejando" porciones codificantes de la actividad enzimática y pueden introducirse modificaciones en la molécula de ácido nucleico hospedadora nativa/grupo de genes o porciones de los mismos.

Las mutaciones pueden realizarse en las secuencias nativas usando técnicas convencionales. Los sustratos para la mutación puede ser un grupo de genes entero o solamente uno o dos de los mismos; el sustrato para la mutación también puede ser porciones de uno o más de estos genes. Las técnicas para la mutación son muy conocidas en la materia y se describen en la bibliografía. Dichas técnicas incluyen la preparación de oligonucleótidos sintéticos que incluyen la mutación (o mutaciones) e inserción de la secuencia mutada en el gen usando digestión con endonucleasas de restricción. Como alternativa, las mutaciones pueden efectuarse usando un cebador con emparejamiento erróneo (generalmente de 15-30 nucleótidos de longitud) que se hibrida con la secuencia de nucleótidos nativa, a una temperatura por debajo de la temperatura de fusión, del dúplex desemparejado. El cebador puede realizarse específicamente manteniendo la longitud del cebador y la composición de bases dentro de límites relativamente estrechos y manteniendo la base mutante centralmente localizada. La extensión del cebador se efectúa usando ADN polimerasa, el producto clonado y los clones que contienen el ADN mutado, obtenidos por segregación de la cadena extendida del cebador, se leccionan. La técnica también es aplicable para generar múltiples mutaciones puntuales. La mutagénesis por PCR también encontrará uso para efectuar las mutaciones deseadas.

Los vectores usados para realizar las diversas operaciones para reemplazar la actividad enzimática en los genes u ORF hospedadores, o para soportar mutaciones en estas regiones de los genes u ORF hospedadores, pueden seleccionarse para que contengan secuencias de control uidas operativamente a las secuencias codificantes resultantes de manera que en el hospedador pueda efectuarse la expresión de las secuencias codificantes. Sin embargo, también pueden usarse vectores de clonación sencillos.

A continuación se describirá la invención con más detalle en los siguientes ejemplos no limitantes con referencia a los dibujos en los que:

La Figura 1 muestra la estructura química del antibiótico macrolactama BE-14106;

La Figura 2A muestra el inicio propuesto de la ruta biosintética para BE-14106;

La Figura 2B muestra el fin propuesto de la biosíntesis de BE-14106.

Ejemplo 1 – Caracterización del aislado MP28-13 (cepa de *Streptomyces* DSM 21069)

En medio de crecimiento agar ISP2 (Difco, Estados Unidos): el micelio sustrato es amarillo pálido, el mismo color que las placas de agar ISP2. El micelio aéreo es blanco y las esporas son casi blancas, solo ligeramente verduzcas.

En medio de crecimiento agar SFM (harina de soja, 20 g/l; manitol, 20 g/l; agar 20 g/l): El micelio sustrato es más beige que en ISP2, el micelio aéreo y las esporas son iguales que en ISP2.

5 El crecimiento en placas con ISP2 es visible después de 2 días, pero la esporulación tarda a aproximadamente 20 días. La esporulación es muy pobre en ambos medios.

Crecimiento en medios líquidos: crece bien en medio de crecimiento líquido TSB (Oxoid, RU), con agitación a 225 rpm y perlas de vidrio (3 mm). Se necesitan 2 días a 25 °C para obtener suficiente micelio.

10 La cepa crece a 20 °C, 25 °C y 30 °C, aunque la temperatura óptima es de aproximadamente 25 °C. A 30 °C la esporulación está afectada.

En la SEC ID N°: 46 se muestra la secuencia génica del ARN 16S de la cepa DSM 21069 (aislado MP28-13).

15 La Tabla 6 muestra las características de resistencia a antibióticos de la cepa DSM 21069.

Tabla 6 Resistencia a antibióticos

Antibiótico	5 µg/ml	10 µg/ml	20 µg/ml	50 µg/ml
Apramicina	sensible	sensible	sensible	sensible
Kanamicina	sensible	sensible	sensible	sensible
Neomicina	sensible	sensible	sensible	sensible
Rifamicina	resistente	resistente	resistente	sensible
Estreptomina	resistente	resistente	resistente	resistente
Thiostrepton	sensible	sensible	sensible	sensible

20 Ejemplo 2 – Generación de una sonda para el grupo de genes de biosíntesis BE-14106

Se aisló ADN total de DSM 21069 (MP28-13) usando el Kit DNeasy Blood & Tissue (QIAGEN). Los dominios β-cetoacil sintasa (KS) se amplificaron usando los cebadores degenerados KSMA-F (5'-TS GCS ATG GAC CCS CAG CAG-3' [SEC ID N° 47]) y KSMB-R (5'-CC SGT SCC GTG SGC C TC SAC-3' [SEC ID N° 48]) descritos por Izumikawa *et al.* ((2003) Bioorg. Med. Chem., 11, 340 1-3405). La mezcla de reacción de 50 µl contenía ADN total aislado de MP28-13 (10-20 ng), T ampón de Reacción 1 x T hermoPol (New England Biolabs), 400 nM de cada cebador, 200 µM de cada dNTP y 2,5 U de Taq ADN Polimerasa (New England Biolabs). La reacción se procesó a 95 °C durante 5 min, después 35 ciclos de 1 min a 95 °C, 1 min a 60 °C y 2 min a 72 °C, y después una extensión final de 5 min a 72 °C.

30 La mezcla de reacción de 50 µl se sometió a una electroforesis en gel y el fragmento de ADN resultante de aproximadamente 700 pb se purificó usando el QIAEX II Suspension (QIAGEN). El producto PCR purificado se clonó en el vector pDrive (QIAGEN) en células EZ de *E. coli* usando el Kit de Clonación por PCR de QIAGEN (QIAGEN). El ADN plasmídico de los transformantes se aisló usando el Sistema de Purificación de ADN Wizard® Plus SV Minipreps (Promega).

40 Se secuenciaron 8 plásmidos recombinantes usando los cebadores específicos de pDrive M13 directo (-20) (5' GTA AAA CGA CGG CCA GT 3' [SEC ID N° 49]) y M13 inverso (5' AAC AGC TAT GAC CAT G 3' [SEC ID N° 50]) descritos en el Manual de Clonación por PCR de QIAGEN (QIAGEN, 2001). La secuenciación se realizó usando el Kit de Secuenciación BigDye® Terminator v1.1 Cycle (Applied Biosystems). De las 8 secuencias obtenidas 5 fueron diferentes entre sí. La traducción en secuencias de proteína y las búsquedas BLAST dieron lugar a una coincidencia para la PKS de tipo I para todas las secuencias (véase Tabla 7).

45 Las secuencias más interesantes fueron la n° 1, 3, 6, 7 y 8. La secuencia n° 1 coincidió con una PKS implicada en la biosíntesis de meridamicina en *S. violaceusniger* (Sun *et al.*, 2006 Microbiol., 152, 3507-3515). Las secuencias n° 3, 7, 8 dieron una coincidencia con LnmJ implicada en la biosíntesis de leinamicina en *S. atroolivaceus* (Cheng *et al.*, 2003 Proc. Natl. Acad. Sci. USA, 100, 3149-3154). La secuencia n° 6 dio una fuerte coincidencia con Vin P1 implicada en la biosíntesis de la macro lactama vici-nistatina en *S. halstedii* (Ogasawara *et al.*, 2004 Chem. & Biol., 11, 79-86) y también con AVES 2 implicada en la biosíntesis del macrólido avermectina en *S. avermitilis* (Ikeda *et al.*, 1999 Proc. Natl. Acad. Sci. USA, 96, 9509-9514).

55 La secuencia n° 6 (SEC ID N°: 51) se seleccionó como una sonda para explorar la biblioteca genómica construida para DSM 21069 (MP28-13). Se generó una sonda marcada con digoxigenina (DIG) usando el Kit de Síntesis de Sondas PCR DIG (Roche Applied Science) y los cebadores M13 descritos anteriormente. El plásmido que contenía la secuencia n° 6 se usó como un molde. La reacción se procesó a 95 °C durante 3 min, después 30 ciclos de 45 s a 95 °C, 1 min a 55 °C y 3 min a 68 °C, y después una extensión final de 7 min a 68 °C. El producto PCR resultante se sometió a una electroforesis en gel y el fragmento de ADN se purificó con el QIAEX II Suspension (QIAGEN).

Tabla 7. Secuenciación de dominios KS amplificados por PCR de DSM 21069

Secuencia	Acierto inicial	Otros aciertos importantes
1	PKS de <i>S. aizunensis</i>	PKS de <i>S. violaceusniger</i> /biosíntesis del policétido meridamicina
2, 4	$\beta$ -cetoacil sintasa de <i>Clostridium</i> sp.	PKS de <i>Bacillus</i> sp.
3, 7, 8	PKS de <i>S. atroolivaceus</i> /biosíntesis de leinamicina	PKS de <i>Bacillus</i> sp.
5	PKS de <i>Amycolatopsis orientalis</i>	PKS de <i>S. violaceusniger</i> /biosíntesis del poliéter nigericina
6 (SEC ID N° 51)	PKS de <i>S. halstedii</i> /biosíntesis de halstoctacosanólido	PKS de <i>S. halstedii</i> /biosíntesis de la macrolactama vicensistatina y PKS de <i>S. avermitilis</i> / biosíntesis de macrolactama avermectina

## Ejemplo 3 – Optimización del procedimiento de conjugación

5 Para establecer un procedimiento para modificar genéticamente la cepa DSM 21069 (MP 28-13), se ensayó la conjugación con la cepa ET12567 de *E. coli* (pSOK804 + pUZ8002) (Sekurova *et al.*, 2004, J. Bacteriol., 186, 1345-1354) siguiendo el procedimiento descrito por Flett *et al.* (1997 FEMS Microbiol. Lett., 155, 223-229). Se realizaron pocas modificaciones del procedimiento. El donante de *E. coli* se desarrolló a una DO<sub>600</sub> de 0,4-0,5. Se usó solo una suspensión de esporas reciente de MP28-13 y la suspensión de esporas se hizo con 2xYT (Sambrook *et al.*, 2000 Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Nueva York, NY). Las esporas y las células donantes se mezclaron y se sedimentaron por centrifugación, y el sedimento se resuspendió en un volumen más pequeño y se diseminó sobre dos placas SFM. Las placas de conjugación se incubaron 24 h antes de la adición de antibióticos (ácido nalidixico 0,9 mg/ml y apramicina 1,5 mg/ml). La temperatura de incubación y la temperatura y tiempo del choque térmico se modificaron para encontrar el óptimo. Se obtuvieron los mejores resultados a una temperatura de incubación de 25 °C y con un choque térmico a 50 °C durante 5 min.

## Ejemplo 4 – Experimento de inactivación de genes

20 La secuencia PKS n° 6 de la cepa DSM 21069 (MP28-13) clonada en el vector pDrive se escindió del plásmido usando enzimas de restricción BamH I y Hind III, y se ligó con el fragmento BamH I/Hind III de 3,1 kb del vector pSOK201 (Zotchev *et al.*, 2000 Microbiol., 146, 611-619) y se transformó en DH5 $\alpha$  de *E. coli*. La nueva construcción se verificó por análisis de restricción y después se transformó en ET12576 (pUZ8002) de *E. coli*. La construcción se transfirió a DSM 21069 por conjugación siguiendo el procedimiento descrito anteriormente. El fragmento BamH I/Hind III de 3,1 kb del vector pSOK201 no contiene elementos genéticos necesarios para la replicación autónoma en *Streptomyces*. Por lo tanto, los transconjugantes solamente pueden obtenerse si esta parte del vector está ligada con un fragmento que tiene un alto nivel de homología con el fragmento de ADN cromosómico en DSM 21069. Dicha homología permite conducir la recombinación a la integración de todo el vector en la región cromosómica correspondiente. Si los fragmentos clonados no contienen codones de inicio o de terminación del gen, dicha integración conducirá a la alteración génica, inactivando eficazmente las copias cromosómicas del gen.

30 Se obtuvo un solo transconjugante y se analizó para la producción de BE-14106. No se observó producción de BE-14106, verificando que la secuencia n° 6 pertenecía al grupo de genes biosintético BE-14106.

## Ejemplo 5 – Construcción de la biblioteca genómica

35 La biblioteca genómica para DSM 21069 se construyó en el vector cósmido SuperCos 1 (Stratagene) de acuerdo con las instrucciones del fabricante (Stratagene, 2005). El ADN genómico se aisló de DSM 21069 siguiendo el procedimiento de mezcla de Kirby (Kieser *et al.*, 2000 Practical Streptomyces Genetics, The John Innes Foundation, Norwich, Inglaterra.), se digirió parcialmente con Mbo I y se desfosforiló antes del ligamiento con SuperCos 1 tratado con Xba I-, CIAP- y BamH I. Para la construcción de la biblioteca se usó XL1-Blue MR de *E. coli* (Stratagene) como hospedador.

## Ejemplo 6 – Exploración de la biblioteca genómica

45 La biblioteca se sembró en placas de agar de Luria-Bertani (LB) (Corning® Low Profile Square BioAssay Dish) que contenía 100  $\mu$ g/ml de ampicilina para dar ~2000 colonias por placa. Se seleccionaron 2304 colonias usando un Genetix QPixII Colony Picker y se transfirieron a 24 placas de 96 pocillos (Nunc) que contenían caldo LB y a 24 placas de 96 pocillos que contenían medio Hi+Ye Reducido (120  $\mu$ l en cada pocillo). Las placas con los pocillos se incubaron con agitación (900 rpm) a 30 °C durante una noche. Se añadió glicerol a las 24 placas LB para dar una concentración final de 15 % (v/v) y después se conservaron a -80 °C.

50 El cultivo de las 24 placas con Hi+Ye Reducido se transfirió a 6 placas de 384 pocillos (Nunc) usando un sistema de manipulación de líquidos robótico Tecan Genesis RSP 200 y después se estampó en un filtro (estampados de 4



copias) usando el Genetix QpixII Colony Picker. El proceso se repitió 4 veces para dar 4 copias de filtros. Los filtros se secaron durante 20 min en una campana estéril.

Los cultivos se lisaron colocando los filtros en un papel de filtro 3 MM Whatman saturado con SDS al 10 % durante 5 min. El ADN se desnaturalizó colocando los filtros en papel de filtro 3 MM saturado con tampón NaOH/Cloruro (NaCl 1,5 M, NaOH 0,5 M) durante 10 min y se neutralizó con Tampón Tris/NaCl (NaCl 3 M, Tris-Cl 1 M, pH 7,4) durante 10 min. Finalmente los filtros se sumergieron en Tampón 2X SSCP (2X SSC + Pirofosfato Sódico al 0,1 % (p/v) para eliminar los restos de colonias y se hornearon a 80 °C durante 2 horas. Los filtros se conservaron a 4 °C hasta que comenzó la hibridación. La hibridación se realizó como se describe para el Sistema DIG (Roche Applied Science) usando la sonda obtenida de DSM 21069. Exponiendo el filtro a una película de rayos X, se identificaron 3 cósmidos candidatos y sus huéspedes correspondientes pudieron volver a sembrarse en estrías de las placas LB conservadas a -80 °C.

El ADN cósmico se aisló de cultivos de una noche usando el Sistema de Purificación de ADN Wizard® Plus SV Minipreps (Promega) y se secuenció hasta el final usando cebadores diseñados para las regiones de los cósmidos flanqueando el sitio de inserto (SuperCos\_dir; GGC CGC AAT TAA CCC TCA C 3' [SEC ID N° 52] y SuperCos\_inv; 5' GGC CGC ATA CGA CTC AC 3' [SEC ID N° 53]). La secuenciación se realizó usando el Kit de Secuenciación por Ciclos BigDye® Terminator v1.1 (Applied Biosystems). En la Tabla 8 se dan los resultados.

Tabla 8: Secuenciación final del cósmido 1, 2 y 3

Cósmido	Cebador SuperCos_dir	Cebador SuperCos_inv
cósmido 1	Biotina sintasa	PKS (enlazador de dominio AT-DH)
cósmido 2	PKS (dominio KS)	PKS (dominio AT)
cósmido 3	Péptido deformilasa	PKS (dominio DH)

Los resultados indican que el cósmido 1 podría contener un extremo del grupo y el cósmido 3 el otro extremo. Los 3 cósmidos se ensayaron para observar si contenían cualquiera de las secuencias solapantes diseñando cebadores para las secuencias finales y usando aquellos cebadores para la secuenciación de los otros cósmidos. A partir de estos resultados se llegó a la conclusión de que el cósmido 2 y el cósmido 3 se solapaban, pero el cósmido 1 no tenía ningún solapamiento con el cósmido 2 y 3. Se diseñaron cebadores para la secuencia final del cebador directo del cósmido 2 para amplificar una nueva sonda para la parte ausente del grupo. Se generó una sonda marcada con digoxigenina (DIG) usando el Kit de Síntesis de Sondas por PCR DIG (Roche Applied Science) con el cósmido 2 como molde. La reacción se procesó a 95 °C durante 5 min, después 35 ciclos de 1 min a 95 °C, 1 min a 60 °C y 2 min a 72 °C, y después una extensión final de 5 min a 72 °C. El producto PCR resultante se sometió a una electroforesis en gel y el fragmento de ADN se purificó con el QIAEX II Suspension (QIAGEN). Para la hibridación con la nueva sonda se usó uno de los filtros copiados y se identificaron 2 nuevos cósmidos candidatos (cósmido 4 y 5). Este proceso con secuenciación final y secuenciación en cruzado con los otros cósmidos se repitió y se descubrió que el cósmido 4 se solapaba tanto con el cósmido 1 como con el cósmido 2.

Ejemplo 7 Verificación de la función de BeCA

Para verificar que el grupo de genes contenido en los cósmidos 1-4 era responsable de la producción de BE-14106, se realizó otro experimento de inactivación de genes. Se amplificaron fragmentos de PKS del cósmido 2 y 3 usando un cebador cósmido (SuperCos\_dir [SEC ID N° 52] o SuperCos\_inv [SEC ID N° 53]) y un cebador degenerado para el dominio KS (KSMA-D [SEC ID N° 47] o KSMB-I [SEC ID N° 48]). Se obtuvieron un fragmento de 1,2 kb para el cósmido 2 y un fragmento de 3,8 kb para el cósmido 3. Ambos fragmentos se clonaron en pDrive y pSOK201 como se ha descrito anteriormente y la construcción se transfirió a MP28-13 por conjugación. Para el fragmento del cósmido 2 solamente se obtuvo un trans conjugante. Para el fragmento del cósmido 3 se obtuvieron diversos transconjugantes y se seleccionaron 6 para análisis posterior.

Tabla 9 Análisis de producción de BE-14106 en mutantes knock-out

	Producción de BE-14106 en comparación con el TS (%)
MP28-13 (TS)	100
cósmido 2 mutante	0
cósmido 3 mutante 1	0,7
cósmido 3 mutante 2	0,9
cósmido 3 mutante 3	0,6
cósmido 3 mutante 4	0,6
cósmido 3 mutante 5	0,6
cósmido 3 mutante 6	0,7

Basándose en los resultados de la secuenciación en cruzado de los cósmidos y en los del experimento de inactivación de genes, se secuenciaron los cósmidos 1, 2, 3 y 4.

Ejemplo 8 – Producción, purificación e identificación de BE-14106

Cultivo de MP28-13

Preparación de inóculo convencional

- 5 Inóculo: Esporas de una placa de agar se transfirieron a un matraz agitador (250 ml, deflector) con 50 ml de medio TSB modificado complementado con glucosa (composición dada en la Tabla 9). Para aumentar las fuerzas de cizalla en el matraz agitador, se añadieron 3 g de perlas de vidrio de 3 mm.
- 10 Incubación: El cultivo se incubó a 25 °C durante 3 días a 225 rpm (incubadora agitadora de Infors Multitron, movimiento orbital, amplitud 2,5 cm).
- Conservación: Se añadió glicerol al cultivo a una concentración del 15 %. La mezcla se transfirió a crioviales y se conservó a -80 °C.

Preparación del pre-cultivo para la producción

- 15 Inóculo: 1,5 ml de inóculo convencional se transfirieron a un matraz agitador (250 ml, deflector) con 50 ml de medio TSB modificado complementado con glucosa (composición dada en la Tabla 10) y 3 g de perlas de vidrio de 3 mm.
- 20 Incubación: El cultivo se incubó a 25 °C durante 2 días a 200 rpm (incubadora agitadora de Infors Multitron, movimiento orbital, amplitud 2,5 cm).

Cultivo de producción

- 25 Inóculo: 3 ml de pre-cultivo para producción se transfirieron a un matraz agitador (500 ml, deflector) con 100 ml de medio BPS 0,3 x complementado con glucosa (composición dada en la Tabla 11) y 5 g de perlas de vidrio de 3 mm.
- Incubación: El cultivo se incubó a 25 °C durante 2 días a 200 rpm (incubadora agitadora de Infors Multitron, movimiento orbital, amplitud 2,5 cm).

30 Composición de medios usados para la producción

Tabla 10 Composición del medio TSB modificado complementado con glucosa

Compuesto	Conc	entración (g/l)
Caldo de soja tréptica		18,5
Glucosa <sup>a</sup>		20
<sup>a</sup> Esterilizado en autoclave por separado		

Tabla 11 Composición del medio 0,3 x PBS complementado con glucosa

Compuesto	Conc	entración (g/l)
Harina de avena		9,0
Extracto de malta		1,5
Extracto de levadura		0,9
MgSO <sub>4</sub> ·7H <sub>2</sub> O		0,12
NaCl		0,3
CaCO <sub>3</sub>		1,5
Almidón soluble		9,0
MOPS		11,1
Glucosa <sup>a</sup>		20
Solución rojo fenol (10 mg/ml) <sup>a, b</sup>		1,5
<sup>a</sup> Esterilizado en autoclave por separado.		
<sup>b</sup> Solución rojo fenol 10 mg/ml, pH ajustado a 8,2 con NaOH.		

- 35 Los componentes se añadieron a agua pre-calentada y los componentes se hincharon durante 10 minutos antes de esterilizar el medio e autoclave. El pH del medio se ajustó con HCl o NaOH después de la esterilización en autoclave hasta obtener un color naranja lo que se producía a un pH = 7 cuando se usaba rojo fenol como indicador de pH.

40 Purificación de BE-14106

Recogida y homogeneización de masa celular

- 45 La masa celular en el cultivo de producción se recogió por centrifugación y se liofilizó. El sedimento liofilizado se homogeneizó con perlas de hierro magnéticas hasta obtener un sedimento fino.

Purificación en bruto de BE-14106

El sedimento celular liofilizado se extrajo con metanol 240 ml/g durante 1 hora. Se añadieron perlas de vidrio para aumentar las fuerzas de cizalla. El sedimento celular se retiró por centrifugación seguido de filtración para eliminar toda la materia insoluble. Al sobrenadante transparente se añadió agua y se mantuvo en hielo durante aproximadamente 30 minutos para precipitar el BE-14106. El precipitado se recogió por centrifugación, se lavó con agua para retirar el metanol restante y se liofilizó. El producto liofilizado representa un producto bruto.

Purificación HPLC preparativa del producto bruto

El producto bruto se disolvió en DMSO y la purificación se realizó en una columna de fase inversa.

Método preparativo: BIOP PREP BE14106\_KFD.M

Sistema HPLC: HPLC preparativa Serie Agilent 1100 con sistema de recogida de fracción  
 Columna: PREP-C1 8, 10 µm, 50x250 mm (PN410910)  
 Temperatura de la columna: Ambiente  
 Fase móvil: acetato amónico 10 mM pH 4,0 (A) y metanol (B).

Tiempo %	B
0,00	85
8,50	85
8,60	100
10,60	100
10,70	85
13,00	85
Flujo eluyente:	85 ml/min

Las fracciones se añadieron al 1 % de una solución NH<sub>3</sub> 2 M y se conservaron a -20 °C.

Concentración del producto en las fracciones HPLC preparativas

La mayor parte del metanol en las fracciones se evaporó en un evaporador de vacío rotacional a 50 °C. La fase acuosa restante que contenía BE-14106 se congeló para aumentar el producto de precipitación y el precipitado se sedimentó por centrifugación. El sedimento se lavó con agua y se liofilizó para dar el producto final.

Análisis LC-DAD TOF de BE-14106 purificado por HPLC preparativa

Cálculo de pureza por LC-DAD-TOF

Después de la purificación por LC preparativa, se observó que BE-14106 constituía > 99 % de compuestos en la muestra que se absorbe a 291 nm como se determina por UV y datos TOF. Se supone que el coeficiente de extinción de los contaminantes es el mismo que para BE-14106.

Datos TOF-MS de BE-14106

El TOF-MS permite obtener una representación gráfica LCTOF de BE-14106 purificado. A partir de esto, la proporción m/z teórica exacta (ión negativo) de BE-14106 (C<sub>27</sub>H<sub>73</sub>NO<sub>3</sub>) es de 422,2701. Esta m/z se observó con precisión aceptable y el pico 422 se correlaciona con el pico UV del heptano.

Método LCTOF: BIOP BE14106 SE.M

Columna: Zorbax Bonus-RP 2,1x50 mm, 3,5 mm (Agilent Technologies).  
 Fase móvil A: acetato de amonio 10 mM (Riedel-de-Haën Cat N°: 34674),  
 Fase móvil B: acetonitrilo al 100 % superprado (Labscan UN1648)  
 Flujo: 0,3 ml/min  
 Temperatura de la columna: Ambiente

Tiempo %	B
0,00	40
10,00	70
10,10	90
12,00	90
12,10	40
17,00	40

Parámetros TOF-MS:

Ionización API-ES negativa

Gas de secado: 10 l/min  
 Presión del nebulizador: 40 psig  
 Temp. del gas de secado: 350 °C  
 Tensión capilar: 3000 V  
 Fragmentador: 200 V

#### Ejemplo 9 – Caracterización de la actividad antifúngica de la cepa DSM 21069

Se investigó la producción de la actividad antifúngica de la cepa DSM 21069, las condiciones de crecimiento de 25 °C en medio PM2 (Bredholt *et. al*, 2008, Marine Drugs, 6(1) págs. 12-24) durante 7 días, produjo una fuerte actividad antifúngica. Después de la incubación el medio se secó y después se extrajo con DMSO. Después de la filtración, los extractos de DMSO se usaron como muestras en un procedimiento de bioensayo robótico con las cepas CCUG3943 de *Candida albicans* y CCUG3942 de *C. glabrata* como organismos indicadores. La última cepa tenía un alto nivel de resistencia contra antibióticos de polieno, mientras que la cepa de *C. albicans* era sensible a polienos. El medio usado en el bioensayos fue AM 19(B) (peptona 9,4 g/l [Oxoid], extracto de levadura 4,7 g/l [Oxoid], extracto de carne 2,4 g/l [Difco], glucosa 10 g/l [BDH], agua destilada).

Las muestras de los extractos de DMSO con bioactividad interesante se fraccionaron usando un sistema de HPLC de la serie Agilent 1100 dotado con un detector de matriz de diodos (DAD) y un recolector de fracciones. Cada muestra se fraccionó en paralelo usando 2 tipos diferentes de columnas LC: Agilent ZORBAX Eclipse XDB-C18, 5 µm, 4,6 x 150 mm y Agilent SB-CN 3,5 µm, 4,6 x 76 mm. Para ambos tipos de columnas, se usó un flujo de 1 ml/min de una mezcla de ácido fórmico al 0,005 % en agua desionizada y como fase móvil se usó acetonitrilo. En ambos casos la concentración del acetonitrilo se mantuvo al 40 % los primeros minutos, aumentando después linealmente del 40 al 95 % durante los 9 minutos siguientes y manteniéndose a una concentración del 95 % para el resto del proceso. El recolector de fracciones se usó para recoger 12 fracciones del eluyente desde 1 minuto hasta 13 minutos de la inyección.

Las muestras se secaron en un instrumento SpeedVac (Thermo Scientific), se disolvieron en DMSO y se midió la bioactividad de las fracciones (ensayo descrito anteriormente).

Las fracciones con bioactividad se analizaron usando un sistema HPLC de la serie Agilent 1100 conectado a un detector de matriz de diodo (DAD) y a un espectrómetro de masas de tiempo de vuelo (TOF). En este análisis se usaron las mismas columnas y tampones que los descritos anteriormente para la etapa de fraccionamiento. Se realizó ionización por electrospray en el modo negativo (ESI-). Se usaron representaciones gráficas DAD para identificar el tiempo de retención aproximado de los compuestos bioactivos en los procesos de fraccionamiento y en el análisis LC-MS-TOF. Las masas moleculares que correspondían a picos significativos identificados en las muestras bioactivas de fraccionamientos en paralelo (columnas C18 y CN) se compararon y se identificaron las masas moleculares comunes a las fracciones de las columnas C18 y CN. Estas masas moleculares (ventana de 10 ppm) se sometieron a la versión online del Dictionary of Natural Products (<http://dnp.chemnetbase.com/>) para investigar los compuestos previamente caracterizados con bioactividad.

En el análisis LCMS de las fracciones bioactivas, se identificaron picos significativos con una masa molecular correspondiente al Antibiótico BE14106. La masa molecular observada en el análisis LC-MS-TOF estaba dentro de 1 ppm de la masa molecular dada en el DNP (masa exacta 423.277344). Además, los perfiles DAD de estos extractos se compararon con la información dada en el DNP sobre los espectros de absorbancia UV del Antibiótico BE14106. Se observó una buena correlación entre los datos dados en el DNP y el perfil DAD del compuesto identificado en los extractos.

#### Ejemplo 10 – Caracterización de las funciones de BecI, BecO, BecR, BecC y BecP

Para verificar las funciones de determinados genes en la biosíntesis de BE-14106, se realizó una serie de experimentos de inactivación de genes. Como se ha descrito anteriormente (Ejemplo 7), también se realizó un experimento de inactivación de genes usando fragmentos amplificados por PCR de los cósmidos 2 y 3. La secuenciación de estos fragmentos y la comparación con el grupo BE-14106 mostró que estos dos fragmentos formaban parte del *becA*, el fragmento de 1,1 kb que codifica las partes de los dominios KS y AT del módulo 2 y el fragmento de 3,7 kb que codifica los dominios KS, AT y DJ de l módulo 2. La producción de BE-14106 se vio claramente afectada en ambos mutantes (Tabla 9).

#### Construcción de vectores para experimentos de inactivación de genes

**Vector de reemplazo de *becI*:** El fragmento *Bgl* II-*Kpn* I de 3,63 kb del cósmido 2 se clonó en pGEM3Zf(-) se digirió con *Bam*H I-*Kpn*I, dando como resultado la construcción pBIR1. A partir de esta construcción un fragmento *Nru* I-*Fsp*A I de 0,8 kb se retiró y la construcción se volvió a ligar, dando como resultado la construcción pBIR2. A partir de la nueva construcción, un fragmento *Eco*R I-*Hind* III de 2,86 kb se escindió y se ligó con un fragmento *Eco*R I-*Hind* III de 3,11 kB de pSOK201, dando como resultado el vector de reemplazo de *becI*, pBIR3.

*Vector de reemplazo de becO*: El fragmento *Sph* I de 14,77 kb del cósmido 4 se clonó en pGEM3Zf(-) se digirió con *Sph* I, dando como resultado la construcción pBOR1A. A partir de pBOR1A un fragmento *Sph* I-*Xba* I de 3,71 kb se escindió y se ligó en el pGEM3Zf(-) digerido con *Sph* I-*Xba* I, dando la construcción pBOR1B. El fragmento *Eco*N I-*Age* I de 6,37 kb de pBOR1B se trató con Klenow para rellenar los extremos y se volvió a ligar como construcción pBOR2. Un fragmento *Eco*R I-*Hind* III de 3,23 kb se escindió de la construcción pBOR2 y se ligó con el fragmento *Eco*R I-*Hind* III de 3,11 kb de pSOK201, dando como resultado el vector de reemplazo de *becO*, pBOR3.

*Vector de reemplazo de becR*: El fragmento *Hind* III-*Bam*H I de 4,35 kb del cósmido 1 se clonó en pGEM3Zf(-) se digirió con *Hind* III-*Bam*H I, dando como resultado la construcción pBRR1. Un fragmento *Sna*B I-*Bsa*B I de 0,4 kb se retiró de pBRR1 y el religamiento dio como resultado la nueva construcción, pBRR2. El fragmento *Eco*R I-*Hind* III de 3,96 kb de la construcción pBRR2 se ligó con el fragmento *Eco*R I-*Hind* III de 3,11 kb de pSOK 201, dando como resultado el vector de reemplazo de *becR*, pBRR3.

*Vector de reemplazo de becC*: El fragmento *Bam*H I-*Sac* I de 5,24 kb del cósmido 3 se clonó en pGEM3Zf(-) se digirió con *Bam*H I-*Sac* I, produciendo la construcción pBCR1. El fragmento *Not* I-*Acc*65 I de 7,48 kb de pBCR1 se trató con Klenow para rellenar los extremos y se volvió a ligar como construcción pBCR2. Un fragmento *Eco*R I-*Hind* III de 4,33 se escindió de la construcción pBCR2 y se ligó con el fragmento *Eco*R I-*Hind* III de 3,11 kb de pSOK201, dando como resultado el vector de reemplazo de *becC*, pBCR3.

*Vector de reemplazo de becP*: El fragmento *Sac* I-*Sph* I de 6,49 kb del cósmido 4 se ligó en pGEM3Zf(-) se digirió con *Sac* I-*Sph* I, dando como resultado la construcción pBPR1A. Un fragmento *Hind* III-*Bcl* I de 3,75 kb se escindió de pBPR1A y se ligó en pLITMUS28 se digirió con *Hind* III-*Bam*H I, dando como resultado la construcción pBPR1B. El fragmento *Xmn* I-*Bbv*C I de 5,85 kb de pBPR1B se trató con Klenow para rellenar los extremos y volvió a ligarse como construcción pBPR2. El fragmento *Eco*R I-*Apa* I de 1,82 kb y el fragmento *Hind* III-*Apa* I 1,23 kb de pBPR2 se ligaron con el fragmento *Eco*R I-*Hind* III de 3,11 kb de pSOK201, dando como resultado el vector de reemplazo de *becP*, pBPR3.

Todos los vectores de reemplazo se introdujeron en ET 12567(pUZ8002) y después se usaron para la conjugación con *Streptomyces* sp. DSM 21069 siguiendo el procedimiento descrito por Flett *et al.*, 1997 (FEMS Microbiol. Lett., 155, págs., 223,-229), pero con las células donantes desarrolladas a una DO<sub>600</sub> de 0,4-0,5 y el tiempo de choque térmico reducido a 5 min. Los antibióticos se añadieron después de 24 h de incubación.

A *BecR* se le asignó inicialmente un supuesto papel de engarce junto con la cadena lateral acilo C20-C25 fabricada por *BecA* y el anillo de macrolactama. El mutante, verificado por un análisis de transferencia de Southern, se ensayó con respecto a la producción de BE-14106 mediante LC-MS de extractos de fermentación. La producción de BE-14106 no se vio afectada en el mutante  $\Delta$ *becR*, lo que implicaba que éste no participaba en la biosíntesis.

Además del papel de *BecR* había también cuestiones acerca de las funciones de *BecI*, *BecC* y *BecP* en la biosíntesis de BE-14106, y el papel sugerido de *BecO* como una hidroxilasa C-8 debía verificarse. Usando los vectores anteriores, se obtuvieron segundos mutantes cruzados para todos los genes y se verificó mediante análisis de transferencia de Southern. La producción de BE-14106 para cada cepa mutante se ensayó por LC-MS de extractos de fermentación. Para el mutante  $\Delta$ *becO*, la masa esperada correspondiente a la fórmula estequiométrica (C<sub>27</sub>H<sub>37</sub>NO<sub>2</sub>) de la BE-14106 se encontró con 1,0 ppm de diferencia de la masa teórica, confirmando por lo tanto el papel de *BecO* como una P450 monooxigenasa de hidroxilación BE-14106. El carbono C-8 representa la única diana probable de *BecO*, dado que el grupo hidroxilo en el C-9 parece deberse a la ausencia de la actividad del dominio DH en la PKS de *BecE* implicada en la biosíntesis del anillo de macrolactama.

Todos los análisis LC-MS de extractos de fermentación de los mutantes  $\Delta$ *becI*,  $\Delta$ *becC* y  $\Delta$ *becP* mostraron una ausencia completa de la producción de BE-14106, y no pudieron identificarse supuestos análogos/precursores de BE-14106. Esto podría indicar que las tres enzimas funcionan muy tempranamente en la biosíntesis, estando presumiblemente implicadas en la síntesis de la unidad amioacilo iniciadora.

Ejemplo 11 – Estudios de alimentación para determinar la incorporación de aminoácidos en BE-14106

Se realizaron estudios de alimentación de DSN21069 usando un medio de producción definido basado en el medio <sup>15</sup>N Silantes OD2 (Silantes pr. N° 103202). La composición de los medios era medio <sup>15</sup>N Silantes OD2 536 ml/l y en g de adición por l: MgSO<sub>4</sub>·7H<sub>2</sub>O, 0,4; CaCO<sub>3</sub>, 5,0; ( <sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,54; KH<sub>2</sub>PO<sub>4</sub>, 0,2 y g lucosa, 10. El medio se complementó con una solución de mineral de nutrientes traza, T MS1 (Borgos *et al.*, 2006, Arch. Microbiol., 185, págs. 165-171) 3 ml/l. La incorporación de aminoácidos no marcados se ensayó añadiendo D-asparagina 0,14 g/l o glicina 0,06 g/l o glutamato sódico 0,10 g/l o sin adición de aminoácido no marcado. Los cultivos de producción se inocularon con 3 % de un pre-cultivo de TSB 0,5 x cultivado como se ha descrito anteriormente, salvo que las células se lavaron con el medio de producción una vez antes de la inoculación para retirar componentes del pre-cultivo. Tanto el cultivo de producción como el pre-cultivo se cultivaron en matraces deflectores de agitación con perlas de vidrio como se ha descrito anteriormente.

Se realizaron análisis LC-MS cuantitativos y cualitativos de BE-14106 8-desoxi y BE-14106 en extractos de metanol de sedimentos de cultivo usando un sistema HPLC de la serie Agilent 1100 conectado a un detector de matriz de diodo (DAD) y a un espectrómetro de masas TOF. Se realizó ionización con electropulverización en el modo negativo (ESI-) o positivo (ESI+), esencialmente como se ha descrito anteriormente (Bruheim *et al.*, 2004, Antimicrob. Agents Chemother, 48, págs. 4120-4129), pero con las siguientes modificaciones: la separación LC se realizó en una columna de 2,1 x 50 mm Bonus-RP de Agilent ZORBAX. La concentración de acetonitrilo se aumentó linealmente del 40 al 70 % durante los 10 primeros min y después se mantuvo a una concentración del 90 % durante el resto del proceso. Las concentraciones de BE-14106 se determinaron mediante absorción de pico UV a 291 nm usando, como patrón, BE-14106 purificado mediante una HPLC preparativa.

Todos los nutrientes en los medios antes de la adición de los aminoácidos a ensayar para la incorporación en la biosíntesis se marcaron con el isótopo  $^{15}\text{N}$ . Los aminoácidos a ensayar se añadieron como  $^{14}\text{N}$ . Además de la D-asparagina y la glicina, la adición de L-glutamato y la no adición de aminoácidos extra se usaron como controles en el experimento. La concentración de aminoácidos  $^{15}\text{N}$  relevantes en los medios fue: asparagina 0  $\mu\text{M}$  (aspartato 54  $\mu\text{M}$ ), glicina 29  $\mu\text{M}$  (serina 11  $\mu\text{M}$ ) y glutamato 29  $\mu\text{M}$ . La adición de los aminoácidos no marcados añadidos fue a 20 x concentración de  $^{15}\text{N}$  aspartato, 27 x concentración de  $^{15}\text{N}$  glicina y 20 x concentración de  $^{15}\text{N}$  glutamato para D-asparagina, glicina y glutamato respectivamente. Los cultivos de producción se extrajeron en DMSO y los extractos se analizaron en LC-DAD-TOF. La molécula BE-14106 contenía un átomo de nitrógeno y por lo tanto se esperaba que si se incorporaba el N de uno de los aminoácidos  $^{14}\text{N}$  añadidos, se observaría una molécula de BE-14106 con una masa exacta (M-H) de 422,27 en un espectro LC-TOF mientras que la molécula BE-14106 marcada con  $^{15}\text{N}$  tendrá una masa exacta (M-H) de 423,28. Los cultivos de producción se extrajeron en DMSO y los extractos se analizaron en LC-DAD-TOF. Los espectros de masa TOF mostraron que la adición de glicina no marcada dio como resultado un 20 % de incorporación de  $^{14}\text{N}$ , mientras que la proporción de incorporación por adición de D-asparagina y glutamato fue del 3 %, que es la misma que para el control.

#### LISTADO DE SECUENCIAS

<110> Sinvent AS

<120> Nuevo grupo de genes NRPS-PKS, su manipulación y utilidad

<130> 27.20.96863/01

<150> GB0805363.9

<151> 20-03-2008

<160> 53

<170> PatentIn versión 3.3

<210> 1

<211> 80060

<212> ADN

<213> *Streptomyces* sp. MP28-13

<400> 1

cccccttggt taccgggcag ttaaggctgg tcttcaatat ggccgaaatc agccctgcgt  
 60  
 tgtgcgccga cgacggccga acccattggt aaccatgacc ggcgaccggg accgcatcac  
 120  
 cgtacaagtg gtctgtgaag cggccgagtc aggcactgtc acccccgcaa ttcgccccga  
 180  
 acggcagcag accccgctcg gcgcccttgc gaagttccca gccacggaga aatcgctcgc  
 240  
 aggttgctag gagatttcgc acaccctctc cgactacccc ttagccccgc accgattacc  
 300  
 ctatcgccga gtgggtgacc gaatgcggtc gctggcttct tgatgttatc gaatatggtc  
 360  
 gctagcttct tgatattgat gactaggttc cgcttctggc tgccggaaac cgacgagctc  
 420  
 cacgatcaaa aggtgtcaac aggggggaaa gtaagctgtg ttgaaggaaac gggacaagat  
 480  
 actcgaacag ctcgatgccc tactgatcca gtccacgcgg ggcagggggg ccatcgccgc  
 540  
 gatcagcggg tcgaccgcga tcggcaagac cacgacactc aacgccctgg ccgaacgggc  
 600  
 gacttcggcg gacatcacgg tgctcagcgt ggtgagttcg ccgcacgagc gcgaggttcc  
 660  
 ctacagcgc ctcgcccagc tgctgcattc gatcgaagcc cagtgcacca ccgcccagct  
 720  
 tccgcgcgcc gggggcgccc ggaccgtcca tgccggcagg caggcggacg aggccgccgc  
 780  
 cctcgccccg ccgttggccc aggacgacc gatgacagtc gcacggcggga cctaccagat  
 840  
 catcgccgag ctgaccgccc tgcggcctct gctgatcacg gtggacgaca tccagcacac  
 900

cgacacggcc accctcacgt gcctgcgcta cctggcgag cgtctgaccc agctctcact  
 960  
 cgcgctggg ttcacgcacg ggatctccgt cgacgagcag ccggctcgtg tcctggacga  
 1020  
 cctcctctac cgcacgagcg cgcgccactt ccacctggag ccgctctcgc gggtagacat  
 1080  
 catggggctg gccgcgaacc ggctcccggg cctcccgtcg gaccggctca tcgtcgagat  
 1140  
 ccaccggctg agcgggggca atccgctgct cgcccaggca ctcacgagg agcacaggct  
 1200  
 gcgtctcgcg tccgattccg tcgcggggcc gatgccgccc cagagcgacg gcacggtccg  
 1260  
 ccggtcgacc accgagggcg gggcgtcgat cggccccgcc ttctaccagg ccgtcctcgc  
 1320  
 ctacctgac cggctcggcc cgcgcgcggt ccgtctcgc cggcgcgctg ccctcctgga  
 1380  
 cgaggcgacg actccccctt tgttgagtcg gctcagcggg atcgacacgg aactgtcgaa  
 1440  
 acgttactta cggttgttca ccggcttggg cgtgctggag ggcgcgcggt tgcggcacgc  
 1500  
 gggcgtacgg caggccgtac tgggtgagat gcctcacggc gaggcgaccc agcagcgcta  
 1560  
 ccgcgccgct cgcctcctga acgagggagg agcggccgccc caagccgctg ccatgcatct  
 1620  
 gctcggcgtc ggtccgctcc acgacggatg ggtgctgccc gtactccagg aggccgccc  
 1680  
 ccacgccatg gaggacggg acgtaccgca gggcatccgt tatctggagc tggcctgcga  
 1740  
 atgctccctg gacgagggac agcggctctc ggcgaagtcg ctgtacgctt tcggccagtg  
 1800  
 gcagctcagg cccgccgagt ccgccccgca cttccgcgcg ctcaagggcc ccatccttga  
 1860  
 ggggaagctc acgggcaccg acgccctgtg ggtcgcggag ggcattgctt tccacctcga  
 1920  
 cttcgacgag gccgtcgagg ttgtcgacca tgtcaacagc ggcgaggcgg acatgtccac  
 1980  
 cgcgctgcac agcaccggga tgctcatggc cgcggaggtc ccgggcctgc tcgaacggct  
 2040  
 ggagcaccga ctgccggcca caaccgcccc ggcgacgtcc cattcggagt taagagcccc  
 2100  
 tcacgccctc gccctcatcc ttgagaacgg agcggacaag tacgccatcg ccctggcccga  
 2160  
 ccaggtgttc cagggcagcc agaactggcc gacgtcgaag ctcagcggcc tgccgaaggc  
 2220



gctgctggcc ctgtgctacg cggatcaact cgacaccgcc gccgcgtggt acgacctgct  
 2280  
 cgcgcccgag gtggaacgac acgacgcccc tggctggtgg gcccagatcg aaagcgtcgg  
 2340  
 cgcgctcctg gcgctgcggc gcgggcggct ggcggaacgcc gtacgccagg cggagacggc  
 2400  
 ccacgcccgg ctgtccgggc cgagggtgaa cgtcagcagc gcgcttgccg tgacctgct  
 2460  
 catcgaggcc cacaccgcca tgggcaacca tcagaccgcg gccaaagtac tggagacgtc  
 2520  
 cgagccgccg cccgcgctgt tcctcaccgc cgcggggctg cactacctgt acgcgcgcgg  
 2580  
 ccgtcatcac ctcgcgaccg gcaacaccta cctggctctg tccgacttcc aggagtgcgg  
 2640  
 cacgctgatg cgtcgtgga acatcgacac accctccctg gcgccctggc ggctgggaga  
 2700  
 ggcgagggtg tggctgcgcc tgggcgacca gaaacaggcg gcccggtcgc tcgagaagca  
 2760  
 gctggccaac ccggacgccg ggctcaccgc gtcgcgcggt atgaccctgc acgccctggc  
 2820  
 cctggctccag gcaaccgcca agcagcccc gatcctgcgg gacgcgttcc gtctgctgga  
 2880  
 ggccgcgggc gcgcgttacg aggctgcccg tgtcctggcc gatctcagcc gcgcctacca  
 2940  
 gcagttgggc gacaaacagg cccggccgac cgcacggcgg gcctggcggc tcgccaagag  
 3000  
 ctgccaggcg gagtcctct gccaggctct gctcccgaca tccatacccc agaacatgga  
 3060  
 cacaaagccg agcagggggt cctgcggcgc cgaccgcgcg ggccaggaca gttcggcac  
 3120  
 actcagtga tccgaacggc gcgtcgcac actggccgcc caggggtacg ccaaccgtga  
 3180  
 gatcgccgag aggctgttca tcacggtcag caccgtggag cagcatctga cccgcgtcta  
 3240  
 ccgcaagatg ggcatcagga accgcgagca gtcctgcag agagcccacg cggtcagcta  
 3300  
 cgagtcctgc tgatccgttc gacccccacc ccgtaagcgc cccggtccgc ctgcccgtcg  
 3360  
 accgacgggc aggcggtggc gttcggaggg cgagcaagga ctgcacgacc gtcctgcga  
 3420  
 gctgaggcgg tcccgaagc tcgggccggc ccggatcgga ccggtcctgg gcctgccgca  
 3480  
 aaggcctggc ccggaaagcc gccctggccc acctcggcgc gaccggcaaa ccacccatca  
 3540

cccgcgtcaa caaccctccg agccaataca cctagggggc tctacgggta gctgggggtg  
 3600  
 agcggcggag atgatgatcg ttccctcaag acgaacgcct gcccggcgtt ttcgaagcac  
 3660  
 gcgatgaggg aggcactcgc catcatgtcg agtgatcttg ttcacggcac ggacgcaatc  
 3720  
 gcaatcaccg gcatgtcatg ccgcctcccc caggcaccgg acgccaactc cttctgggag  
 3780  
 ttgctgcggt cgggccgtag tgcgatcacc gaggtgccgc cggaccgctg ggaccggac  
 3840  
 gaggtgctgc cggactcacc ggagcggcac cgcgcagcgc tgcgccacgg cgggttcctc  
 3900  
 gaccgagtgg accagttcga cgcggccttc ttcgggatct cggcgcgcga ggccgtggcg  
 3960  
 atcgaccgc agcagcggct cttcgccgag ctggcctggg aggcgctgga ggacgcggga  
 4020  
 atcgtccccg agacgctgcg gtccaccgcg accgccgtca tcgtgggagc gatcgccggg  
 4080  
 gactacgcgg cgctggcaca ccgcggcggg gcgatcacac agcactcgct gccggggctg  
 4140  
 aaccgcggtg tcatcgccaa ccgggtgtcc tacgcgctgg gactgaacgg cccgagcatg  
 4200  
 gcggtcgact cggcgcagtc gtcgtcgctc gtggcgggtg atctggccgt ggagagtctg  
 4260  
 cgcaagggtg agtgcaccct ggccctggcg ggcggtgtgg cgctcaactt cgcgcccgag  
 4320  
 agcggcgaag tcgccggaat gttcggcggg ttgtcgccgg acggacgctg cttcaccttc  
 4380  
 gacgcgcggg ccaacggcta cgtgcgcggt gagggcggcg gcgtcgtcgt actgaagccg  
 4440  
 ctggcacacg cgggtgcgca cggcgacacg gtgtacgggg tcatccgtgg caccgcggtc  
 4500  
 aacaacgacg gctccaccga cgggctcacg gtgccagcg ccgaggccca ggcgaccgtg  
 4560  
 ctgcgccagg cgtgcgagga cggcggcgtg gaccggccg aggtccgtta cgtggaactc  
 4620  
 cacggcacgg gcactccgac gggcgatccg ctggaggcgg cgggtgtggg cgccgcgtac  
 4680  
 ggcagcgcgc gtcctgcggg ctcaccggtg ctcgtcggct cggccaagac caacgtcggg  
 4740  
 cacctggagg gcgcggcggg catcgtcggg ctgatcaaga cggcgtgag catccggcac  
 4800  
 cgggagatcc cggccagcct caactacgag acgcccaacc cgcggatcga ccccgaggcg  
 4860

ctgaatctgc gcgtccagac cgcgtccggc ccgtggccgg acgcgccgct gctggccggg  
 4920  
 gtcagctcct tcggtgtggg cggaacgaac tgccatgtcg tactggccga ggcgcccag  
 4980  
 cgggccgcgt ccgaggagga cgcgccgcag ggcgacgagc cggagatccc gctggctccg  
 5040  
 tggctggtgt ccgggcgtac cgaggcggct ctgcgcgccc aggccgggcg gctgctggag  
 5100  
 cggcggacgg cggacgccga cgcgttcgac atcggccgct cgctcgcggg cacccgtacc  
 5160  
 cacttcgagc accgcgccgt cgcgctcggg ctcggacacg acgcgcagct tgaggcgttg  
 5220  
 cggctccggcg ccgacgtacc ggggctcgtc acgggggtca ccggcgacca cgggaagatc  
 5280  
 gcgctggtgt tcccggggca gggctcgcag tgggagggca tggcgcgcga gttgatgcg  
 5340  
 acgtcggcgg tcttccgcgc gtcgatcgag gcgtgccacg aagccctcgc gccgtacgtc  
 5400  
 gactggtcgc tgctggacac gctcaccgat gagtccggcg cgacgtccct cgaccgcgcg  
 5460  
 gacgtcgtcc agcccgtgct gttcgcggtg atggtctcgc tcgccagggg gtgggagtcg  
 5520  
 ctgggtgtac ggcccgaagc ggtcatcggg cactcgcagg gcgagatcgc ggcggcgcac  
 5580  
 atcgcgggag cgctcgacct ggcggacgcc gccaggatcg tggccctgcg cagccagacg  
 5640  
 atcatgacgc tggcgggtac cggggccatg gcgtcgggtc cgctggcggc cgaccgggtc  
 5700  
 accgagtaca tcgccccctt cggcgacggg ctgagcatcg ccgcgggtaa cgggccgacc  
 5760  
 accactgtcg tggccgaaa ccccgaagcc atcgcggagt tgctggcccc ctgcgaggcg  
 5820  
 gaggggattc gcgcgagggc cgtctcggcc gtggacttcg cctcccactc ttcgcacatg  
 5880  
 gaggcgggtca aggaccggtt gctggagcag ttcgccgagg tgacgccgcg ttcgtgcgac  
 5940  
 atcgcgttct actccacggt caccgcgagc gccatcgaca cggccggtct cgacgccggc  
 6000  
 tactggtact ccaaccttcg ccggccccgtc ctcttcgagg cgacgctcag ggccatggcg  
 6060  
 gaggacggct tcggcacggt cgctcagtcg agtccgcacc ccgttctcac gctcggattg  
 6120  
 cgggcgacgc tgccggacgc gctggtcgcg gactcgtctc gccgtaacga ggctccgtgg  
 6180

ccgcagctgc tgacctccct ggcggaactg cacgtatccg gcctgcccgt ggactggctc  
 6240  
 gccgtcttca aggggcgtac accgggtcgc gtggcgtcgc ccacgtacgc cttccagcgc  
 6300  
 gagcgtact ggcccagagt ctcgaccgcg ttcgagcccg ggacgcgcgg cgccgtccag  
 6360  
 caccaggaag ccgcgcgcga ggagatcccc gcggcgagct ccacctggtc ggatcggtg  
 6420  
 gccggactgc cggcggacga gcgctcccgt gaggcgtgg agctgggtgc gctgcgcacg  
 6480  
 gcgatcgtgc tcggtcatct gagcacggac ggggtcgtg tcggccaggc gttccgcgag  
 6540  
 ctgggcatgg actccacgat ggccgtccag ctccgtcaga acctcgtgga catcaccggg  
 6600  
 ctggcgtgc cggagaccgt cgtcttcgac tacgcgagcc cgtcccggct ggccagccgg  
 6660  
 ctctgcgaac tggccctggg cgaggacacg tcgtcggccg cgtccgcgt gtcgcggctg  
 6720  
 gcgtcggcgc tggacgccga cgatccgatc gtgatcgtcg gcatggcctg ccgttacccc  
 6780  
 ggccgcgcca gcaccccgga cgagctgtgg cagctcgtcg acgacggcgt gagcgcgatc  
 6840  
 tcgggcttcc ccaccgatcg cgggtgggac ctggacgccc tgtacgacct cgagcccggg  
 6900  
 gtgcgcggca agacctacac acggcacggc gggttcctcg acgaggccgc cgagttcgac  
 6960  
 accgagttct tcgggatcag cccgcgcgag gccaccgcga tggaccgcga gcagcggctg  
 7020  
 ctctccagg tcacctggga ggcgctgga cgcgccggga tcgaccgga cggcctccag  
 7080  
 ggcagcagca ccggtgtggt cgtgggcgcg atgtcgcagg agtacggccc ccggctgcac  
 7140  
 gagggcgacg acggactggg cggctatctg ctaccggca ccaccgcgag tgtcgtctcc  
 7200  
 gggcggatct cgtacacctt cggctctgag ggcccggcgg tcaccgtcga cacggcctgc  
 7260  
 tcgtcgtcgc tcgtcgcgat gcaccaggcg gcgcaggcgc tcgcgcgtcg ggagtgtcgc  
 7320  
 ctggccctgg cgggcggcgt cgcggtcatg gcgacgcccg gcatgttcgt ggagttcgga  
 7380  
 cagcagcgtg gtctggctcc cgacggcccg tgcaagtcgt tcgccggtgc ggcggacggc  
 7440  
 accatctggg ccgagggcgc gggcatggtc ctctggagc gtctgtccga cgcgaaggcc  
 7500

aacggacaca cggctcctggc cgtcatccgc ggctccgcgg tcaaccagga cggcgcagc  
 7560  
 aacggcctca ccgccccaa cgggccctcg cagcagcggg tcatcaccgc cgctctggcc  
 7620  
 ggcgcgggccc tcacgcccga ccaggctgac gcggctcgagg cacacggcac cggaaacccc  
 7680  
 ttgggcgacc cgatcgaggc ccaggcgctc ctggccacct acggccagaa ccgtgaagaa  
 7740  
 ccgctgtggc tcgggtcggt gaagtcgaac atcggccaca cgcaggccgc cgccggtatc  
 7800  
 ggcgggggtca tcaagatgat ccaggccatg cggcacggca ccctgccccg gaccctgac  
 7860  
 gtcgacgagc ccagccccga catcgactgg gactccggca acgtgaggct cctcaccgag  
 7920  
 gccccggcct ggccccgagac cgaccaccgc cgccgctccg ccgtctctc cttcggcatc  
 7980  
 agcggcacca acgcccacct catcctcgaa caggcccccg cgacgccgga gcccgtggac  
 8040  
 ggcgacgacg agcaggagac cccgcagggg gccctcgttc cctggttcct gtccgccaag  
 8100  
 agcgcccccg cgctccgcga ccaggcccag cgctcctcg accatgtcat cgcgcgcccc  
 8160  
 ggggtcgacc cggcgcacat cggccgtgcc ctgacagcca cccgcgcccc tttccagcac  
 8220  
 cgtgcgggtg tggtgggaga gggccgtgac gaactactcg cgggtcttcg ggcgctgagc  
 8280  
 aacgacgagt catcccgcgc ggtcgtcacc ggcacggcac ggaaggcac caccgcgttc  
 8340  
 ctgttcacgg gacagggcgc gcagcgggccc ggcattgggccc gcgagctcta cgacacgtac  
 8400  
 ccggtcttcc gggacagctt cgacgaggtc tgcgccaccc tcgaccggca tctgaacgcc  
 8460  
 gaacagccgg tcaaggacgt cgtcttcgcc gacgacgcca ccctgctcaa ccagaccgc  
 8520  
 tacaccagc ccgctctctt cgcgatcgag acctcgctct accgcctggt cgaatcattc  
 8580  
 gggatcacc cgcagcacct gaccggccac tccatcgggtg aactaccgc cgcccatatc  
 8640  
 gcgggcgtct tcaccctgaa cgacgcctgc cggctcgtcg ccgcgcgcgg ctcactgatg  
 8700  
 caggccctgc ccgccaacgg cgcgatgatc tccctgcgcg ccaccgagga gcagatcctt  
 8760  
 ccgttctcgc aaggccacga gcaccacgtc gccatcgcgg cgggtcaacgg gcccaactcg  
 8820

atcgtcatat cgggcgacca ggaagccacc accgccatcg cggaagccct ggccgagaca  
 8880  
 ggtgtcaaga cgcggcgccct cacggtctcc cacgcgttcc actcccccca catggacccc  
 8940  
 atgctggagg agttcgagcg tacggcggcg gacctgacgt accacgcccc gacgagccccg  
 9000  
 atcgtctcca acctcaccgg ccaactcgcc gaccaccgca tcaccacccc gcagtactgg  
 9060  
 gtccagcacg tccgggacgc cgtgcgcttc gccgacacca tcaccaccct cgatcagctc  
 9120  
 ggcacccggc actacctega actcggaccc gaccccgctc tcaccactct cgtcaacgag  
 9180  
 accctcggca agaccgggg caccatcccc accgccgtcc tgcgcaaggg gcaactccgaa  
 9240  
 gccgccacgc ttctcacggc gctcggcacc gcgtacaccg ccggcgcgcc cgtcaacctc  
 9300  
 agcagccacc tcccggcccc ccacaccac cccgacctgc ccacgtacc cttccagcgc  
 9360  
 cagcgttact ggcacgcggc caccaccgcc acgggagacg tgctcgtcggc cgggctgacc  
 9420  
 gccacgggac acccggtgct gaccacggcc gccgagcttc ccgatccggg cggggtgctg  
 9480  
 ctgaccggcc ggggtgaacgc ggcgtcacc gcctgggagg cggaccacgc cgtcttcggc  
 9540  
 accccggtca tgccgggcgt ggcgttcgct gacatgctgc tgcacgccgc ggcgctggtc  
 9600  
 gggcgccctc gtatcgagga actcaccac catgtcttcc tggccctgcc ggaacacggc  
 9660  
 gccctccagc tgagggtggt cgtccgcccg gcggacgact ccgggcggcg gtccttcgcc  
 9720  
 gttcaactcg gggccgagga cgccccgctg ggcgccgact ggacctgcca cgccaccggt  
 9780  
 gcgctgggcg tcgccccggc cgtaccgccg gcccttcgc ccgtggacgc ggcctggccc  
 9840  
 ccggcctccg ccgcggctct cgacaccgac ggcttctacc ggcggatcgc cgaggccggg  
 9900  
 ttcggctacg ggccggtctt ccaggggctc gccgccgcct gggaggacgg agacacgctg  
 9960  
 tacgccgagg tcgcccctgcc cgcggggacc tccccgggt cgtacggcgt ccaccccgcc  
 10020  
 ctgctcgact ccgcgctgca cccgatcgcc ctggccgcga ccggtgccga gacggacggc  
 10080  
 acactccatg tgccgttctc ctggagcggg gtgacgctcc acgcgtccgg cgcgcacacc  
 10140

ctgcgcgtcc ggctcgtgcg ctccacgccg gagacggtcg cgctgaccgt cacggacccc  
 10200  
 tccggcgcgc cggtgctgac cgtcgactcc ctcgccatgc gcgggggtccg tgccgagcag  
 10260  
 ctcgaagccg ccaggccccga ccgcgacggt gcgctgcacg acgtcgcctg gcgcgcggtg  
 10320  
 cccgcgccgc cacgcgccaa ctcgcgcccc gacggcgtcg gctgggcccgt ggtgggagac  
 10380  
 acccgtgacc cgcgagtcgc cgccgtgctg gcatcgcctc gtgcggcccgc cgagtcgtac  
 10440  
 ccggacgccg aagcgcctgcg cgcctcgttg cgggccggcg cggtcggcc ctcgacgatc  
 10500  
 gtcgcccgtc tcgccaccgg gaccgccgaa gacggcgcgc acccggtcgc ggcggcacac  
 10560  
 accgggacac gccacgccat gcacctgctc caggccgtcc tggccgacgg ggcaccggac  
 10620  
 tcccggctgg tgatcctcac cgagaacgcg agatacaccg gaaccggcga cgcggcggcc  
 10680  
 gacatggccg gtgcggcggc ctggggcctg atccgttcgg cccagtccga gcaccccac  
 10740  
 cggttcacgc tctcgcagct cgacggctcg cgggcctccg acgaggccgt cgtcgcggcg  
 10800  
 ctctccgccg gtgagcccca actggccgta cgcgaaggcc ggctgttcgt tccccgcctc  
 10860  
 gcccgctcga ccccgggcgc gattcccgcc acgttcgacg cgcggcgtac ggtgctcatc  
 10920  
 accggcggca ccggcgcgct gggctcgatt ctcgcgcgcc acctggtcac ccggcacggc  
 10980  
 gtcaggaagc tgctgctgac cagcaggagc ggtcgtacgg cggacagcac cgtcgtcgcg  
 11040  
 gaactcgccg gactcggcgc cgaggtcacc gtcgcggcct gcgacgtcac cgaccggctg  
 11100  
 tcgctggaga ccctgctcgc gggcctgccg acggagcatc cgctcggcgc ggtcgtgcac  
 11160  
 tgcgcgggcg tcttgacga cgggtgtggtc acggagctga ccgaggaccg gctggacgcg  
 11220  
 gtgctccgcc cgaagggtga cggcgcgtgg aacctgcacc aggtgacccg tgacatgggc  
 11280  
 ctggacctgg acgccttcgt gctgttctcg tccgtggctg gtgtcctcgg ctcgccccga  
 11340  
 caggccaact acgccgccgc caactcctc ctcgacgaac tcgccgagca ccgccgtacg  
 11400  
 gccgggctcc ccgccaagtc cctcgcctgg ggactgtggg agagcggcat ggccgacacc  
 11460

ctcgacgagc aggaccgggc gcggatgagc cgcggcggac tctcgccgat gcccgccgag  
11520

cgggcgctcg cgctcttcga ctccggcactc gcgacggcgc gggcggtgct cgtgcccgcc  
11580

ggggtcgatg tctcccacgc gcgcaecgag cgggcgtcga tgctctcccc tctgctcgcc  
11640

gaactgctcc cggcccaggc ggcgcccgcc gaacgaggcg gtgaagcggg cgacgagtcc  
11700

tcgctgcgcc agcagttggc cctcctgccc gagcccgaac agcgcgaact cctcctggag  
11760

gtcctgcgca agcatgtcgc ggcggttctg ggccacagct ctccgctcgt catcgacccc  
11820

gagagctcgt tcaaggacct cggtttcgac tcgctcgccg gcatcgaact cctcatggtg  
11880

ctgggcgagt ccatggggct gcacctgccg tccacgatgc tgttcgacca cccgacgcc  
11940

gagttgctga tcaccacct cagggacgaa ctggtcgacg acgaggccgt acctgtcacc  
12000

gccgcggaca ccgcggccgt cgccgtggca ccacgcgacg acgacgaacc catcgccgtg  
12060

atcggcatgg ggtgccgcta cccgggcggc gccaccacgc cggacgagct gtggcgactg  
12120

gtcaccgagg ggggtggacgc catcggtccc ttccccacca accgcggctg ggatctggag  
12180

gagctgttcg atcccgacct cgacatgcgc ggggaagacct atgcccgcaa gggcgggttc  
12240

ctctacgacg ccgaccgttt cgaccccgag ttcttcggca tcagcccccg cgaggccctg  
12300

gcgctcgacc cgcagcagcg actcctcctg gagaccacct gggagacgtt cgagaacgcg  
12360

ggcatccgcc ccgacacct gcgcggcaag cccgtcggcg tcttcgccgg cgtcgtcacc  
12420

caggagtacg gctccctcgt ccaccggggc accgagccgg tcgacggctt cctgctgacg  
12480

ggtacgacgg cgagtgtcgc ctccgggcga ctggcctaca cgctgggtct tgagggcccc  
12540

gcggtcaccg tcgacaccgc gtgctcctcc tcgctggtcg cgatgcacct ggcctgccag  
12600

tcgctgcgga acaacgagtc cacgatggcc ctggccggtg gcgccacggt gatggccaac  
12660

cccggaatgt tcctggagtt cagcccagcag cgcggtctgg ctcccgacgg ccggtgcaag  
12720

tcgctcgccg gtggcgccga cggcaccatc tgggccgagg gcgcgggcat ggtcctcctg  
12780



gagcgtctgt cggacgcca ggccaacgga cacaccgttc tcgccgtgat ccgcggctcg  
12840  
gccgtcaacc aggacggcgc cagcaacgga ctgaccgctc cgagcggccc gtcccagcag  
12900  
cgggtcatca ccgccgctct ggccggcgcg ggcctcacct ccgaccaggt cgacgcggtc  
12960  
gaaggacacg gcaccggaac cccgttgggc gacccgatcg aggcccaggc gtcctggcg  
13020  
acgtacggcc agggccgtga agccgaccag ccgctgtggc tcgggtcggt caagtccaac  
13080  
atcggccacg cgcaggccgc cgcgggtatc ggcgggggtca tcaagatgat ccaggccatg  
13140  
cgccacggca ctctccccg gaccctgcac gtcgacgagc ccagcccgca catcaactgg  
13200  
gcctcgggca acgtgaggct cctcaccgag gagcgcgcct ggcccagagc cgaccacccg  
13260  
cgccgctccg ccgtctctc ctccggcatc agcggcacca acgcccattg catcctcgaa  
13320  
caggcccccg cgacgccgga gcccgccgag gatgaccacg agcaggacgc gccgcaggcg  
13380  
accctggctc cgtgggtcct gtccggcaag acggagcagg cctcccgcga ccaggcgag  
13440  
cagctgcgca cgtacctcga actcaacccc gggctgcgca cggaccgagt cgctcacgcg  
13500  
ctcggcacca cccgcgcca gttccagtac cgggccgtgg tgctcggcac cgaccaccag  
13560  
gcgttcgacc gtgccctggg cacgctcacc ctccggcagc cgtccccggc gctggtacgc  
13620  
gggacgccgc accccggcaa gacagccttc atgttcacgg gacagggcgc gcagcgggccc  
13680  
ggcatgggccc gcgagctcta cgacacgtac ccggtcttcc gggacacggt cgacgaggtc  
13740  
tgcgccaccc tcgaccgga tctgaacgcc gaacagccgg tcaaggacgt cgtcttcgcc  
13800  
gacgacgcca tcctgctcaa ccagaccgc tacaccagg ccgctctctt cgcgatcgag  
13860  
acctcgtctt accgcctggg cgaatcattc gggatcacc cgcactacct gaccggccac  
13920  
tcgatcggtg agatcacggc cgcccacgtg gccgggatct tcaccctcga cgacgcctgc  
13980  
cgactggctc ccgcgcgcgg ctcaactgat caggctctgc ccgccaacgg cgtcatgatc  
14040  
tcgctgcggg ccaccgagga gcagatcgtc ccgttcctcg aaggccacga gcaccacgtc  
14100

agcgtcgcgg cgggtcaacgg gccagctcg atcgtcatct cgggcgacca ggaagccacc  
 14160  
 accgccatcg ccggcgctct cggcgagacg ggtgtcaaga cgcggcgctt cacggtctcc  
 14220  
 cacgcgttcc actcccccca catggacccc atgctggacg agttcgaact cgtggccgga  
 14280  
 gagctgacct accacgcccc gacgatcccc atcgtctcca acctcaccgg ccaactcgcc  
 14340  
 gaccactaca tcaccacccc gcagtactgg gtccagcacg tccgggaggc cgtccgcttc  
 14400  
 tccgacggca tcaccacct cgaccggctc ggcacccggc actacctga actcggaccc  
 14460  
 gaccccgctc tcaccacccat ggcgcaggac agcgtggccg atgacagcga cggcgcctc  
 14520  
 gtcgccaccc tgtaccgcga ccgggacgag aaccacagct tcctcaccgc cctggccacg  
 14580  
 gcacagccc atggcatcca ggtcggctgg acccccgctg tcggtgagac ctcggtcccc  
 14640  
 gccctggcc tccccaccta ccccttcag cggcagcact actggctgga ggcggcgaag  
 14700  
 cccacctcgg gtgccgacgg tctggggctg acggcgaccg accatccggt gctgaccacc  
 14760  
 ttggccgaac tcccggacgg aggcgggac ctgttcaccg gccgcgtctc cgggaacgat  
 14820  
 ccggactggg tggccgagca catcatctc gggacaatga tcgttccggg tgtggccttc  
 14880  
 gtcgacctcc tgctccacgc ggcacgccc atgtggactgc agcacatcga ggaactcacc  
 14940  
 caccacgtgt tcctcgccgt gccggagcgc gccgccctcc agctgcggct cctgatcgag  
 15000  
 ccggcggaca gctccggaag ccgggccttc gcgttctact cgcggccgga ggacgtcccc  
 15060  
 gtcgaaaccg actggaccct ccacgccacg ggcgcgctcg gagccgaacg cagggaaagtc  
 15120  
 cccgcgggcg ccgactcgt caggaacgag gtctggccgc ctgacatctc ggacaccatg  
 15180  
 gacgtggggg agttctaccg ccgggtcacg gacggtggct tcggctacgg accgctgttc  
 15240  
 cgagggtca agaaggcctg gcaggacggg aacacgacgt acgcggaagt ctctttgccc  
 15300  
 gccggtccg atcccggcga ctacggcatc caccgggtc tgctcgactc ggcgctccag  
 15360  
 ccggccgcgc tcatcatggg agagaccgag gcggccgact cgatccgggt gccgttctcc  
 15420

tgggccggtg tgtccctcca cgcgacgggg gccgactccc tgcgtatccg ccacacgtgg  
 15480  
 accacaccgg acaccgcgtc gctggtcata gccgaccaga cggggacacc ggatcatgacg  
 15540  
 atcgactccc tcgccatgcy gacggtcggc gccgaccaac tcgccgccac ccgtgcyggcy  
 15600  
 gacgcccggag agctgtacaa ggtcgactgg ttcgacgtcc agaccgtgga ggacaagacc  
 15660  
 cagggcgcgg gcaccgcaa gtgggcgggtg gtcgccgacc cggggaacac ccaggtcgcc  
 15720  
 gcggcactct ccccgctcgg cggcgcggtc gaggtggagc cggacgcgggt gacgctcccg  
 15780  
 acgacaccgg gggacgacac caccggccg gacgtgggtc tcacatggtg cgtctccgag  
 15840  
 cccggcgcgg acccggcaca ggccgcgcgc tccttcacc accgcgtgct cggcctcgtc  
 15900  
 caggcgggtc tctccgacga tcggccggac tcccgcctgg tgatcctcac caagggcgcg  
 15960  
 atgtccgccc gtagcggcgg cgcggccgac ctggccggag ccgcgggtctg ggggctgatc  
 16020  
 cgcaccgctc agaccgaaca ccccgatcgg ttcatacctga tggacctcga cggttcggat  
 16080  
 gcgtcgtcgc gggccgtggg cgtgcccctg aacgccggcg aaccccaact cggcgttcgc  
 16140  
 gacggcccggc tcctcgcccc tcgcctcggc cgtatcggca ccgccgactc cgagccgacc  
 16200  
 gccgcaccgg cgtcgttcga cccggacaag acggtgctca tcaccggcgg caccggcgcc  
 16260  
 ctgggcacgc tcctcgcccc tcacctgggtc acccgccacg gtgtgaagaa gctgctcctg  
 16320  
 accagcccggc gcggccgtcc ggcaggcagc accatcaccg cggaaactcgc cgaactcggc  
 16380  
 gccgagggtca ccatcgtggc ctgcgacgcg gcggaccggg agtcgctgga agccctgctc  
 16440  
 gcgagcctgc cggacgagca tccgctcggg gcggtggtgc actgcgcggg aacgctcgc  
 16500  
 gacggcatcg tcacggcgct gacacctgac cggttcgcgc aggtgctccg gccgaagggtg  
 16560  
 gacggcgcgt ggaacctgca cgagctgacc cgtgacctgg acctggacgc cttcgtgacg  
 16620  
 ttctcgtccg tggtcgggtg cctcggctcg ccgggacagt ccaactacgc cggcgcaac  
 16680  
 gtcttcctcg acgagctggc cgaacaccgc cgtacggccg gactgcccgc caagtccctc  
 16740

gcctggggac tgtgggagag cggcatggcc gacaccctcg acgagcagga ccagggcgcg  
 16800  
 atgaaccgcg gcggcctcct gccgatgccc gccgaacagg cactcgggca cttcgactcg  
 16860  
 gcgctcgcga ccgaccagac cgtcgtggtc ccggccaagc tcgacctcgc cgggctccgt  
 16920  
 gcccgcgcg cgacgggtccc ggtggcgccg atcttccgtg ggctgggtccg tacgccgctg  
 16980  
 cgcagcgcg cccagggcgg cggcgcggga gcggaggctg gagccctggg gcagtcgatc  
 17040  
 gcggggccc cggaggccga gcaggaccag atcatcctgg acttctcgcg caatcacgtg  
 17100  
 gccaccgtcc tcggacacgg ctcggcgaac gcgatcgacc ccgcgactc cttcaaggag  
 17160  
 ctgggcttcg actcgctcag ctcggtggaa ctgcgcaact cgctcaaaa ggcgtccggg  
 17220  
 atgcgactcc cgtccaccct gttgttcgac taccaccacc cctcgggtact ggccggctac  
 17280  
 atccgcaacc aactggcggg cggcaagcag gcggaggcag gcgcgcaagt gggccgccgc  
 17340  
 accgttcgcc cggcgtcctc gcggagcgcg gcggccgacc cgatcgtgat cgtgggcatg  
 17400  
 gggtgccgct tccccggtgg cgccgacacg cccgaggcgc tgtggaagct ggtcgcggac  
 17460  
 gagcgtgacg cgggtggggc ctteccccgac aaccgcggct gggacatcga gaacctcttc  
 17520  
 gacgacgacc ccgacgtacg ggggaagtcg tacgccagtg agggcgggtt cctctacgac  
 17580  
 gccgaccggt tcgacccccga gttcttcggc atcagccctc gcgaggccct ggcgctcgac  
 17640  
 ccgcagcagc ggctgctgct cgaaaccacc tgggaagcgt tcgagaacgc gggcatccgc  
 17700  
 cccgacactc tgcgcggcaa gcccgtcggc gtcttcgccc gcgtcgcggc cggggagtac  
 17760  
 gtctcgctca cccaccacgg cggcagcccc gtcgagggtt acctgctgac gggtagcagc  
 17820  
 gcgagtgtcg cctccgggcg catctcgtac acgctgggtc ttgagggccc cgcggtcacc  
 17880  
 gtcgacacgg cctgctcgtc gtcgctcgtc gcgatgcacc tggcgtgcca gtcactgcg  
 17940  
 aacaacgagt ccacgatggc cctggccggc ggcgccacga tcatgtcaa cgcgggcatg  
 18000  
 ttcattggagt tcagccgcca gcgtggctctg gctccccgaca gccgtgcca gtcctacgcg  
 18060

ggcgcccgcg acggcaccat ctgggccgag ggcgcgggca tggtcctcct ggagcgtctg  
 18120  
 tcggacgcca aggccaacgg acacacggtc ctggccgtca tccgcggctc ggccgtcaac  
 18180  
 caggacggcg ccagcaacgg cctcaccgcc cccaacgggc cctcgcagca gcgggtcatc  
 18240  
 aacacggcgc tcgccagcgc gggctctcacc cccgaccagg tcgacgccgt cgaaggacac  
 18300  
 ggcaccggaa cgccgctggg tgaccctgatc gaggccagg cctgtctctc cacctacggc  
 18360  
 cagaaccgtg aagagccgct gtggctcggg tcgttcaagt ccaacatcgg ccacgcgcag  
 18420  
 gccgcccgcg gcgtcggcgg ggtcatcaag atgatccagg ccatgcgcca cggcaccctg  
 18480  
 ccccggacgc tccacgtcga cgagcccagc ccgaacatcg actgggactc cggcaatgtg  
 18540  
 cggctcctga ccgaggcccg ggccctggccc gagaccgacc gcccgcgccc ctccgccgtc  
 18600  
 tcgtccttcg gcatcagcgg caccaacgcc cacctcatcc tggaggaagc gcccactccc  
 18660  
 acccaccctg agccggcccc cgagagcgc cgcagggcaa ccacgggtgc ctggatactc  
 18720  
 tcgggcaaga gtgaacaggc cgtgcgggac caggcccagc gcttgctcga ccacgtcagc  
 18780  
 gagtaccccg agctccagcc ggtcgacatc gcgtactcgc tggccaccgc ccgtacctcc  
 18840  
 ttcgagcgc aggccgtcgc gatcggcgc acccatgacg aactcgtcga ccacctccgc  
 18900  
 tcgctgacct aggaccccgg caccgccctc ctgcacggcc agtcccactc caagaagggtg  
 18960  
 gccctcctct tcaccggtca gggctcccag caccgggca tgggcccgtca gctctacgac  
 19020  
 acgcaccccg tctaccggga cgcgttcgac gaggtgaccg ccaccctgga ccagcacctc  
 19080  
 caggccgaac agccggtcaa ggacgtcgtc ttcgccgacg accccacct gctcaaccag  
 19140  
 acccggatac cccagcccgc catcttcgcc ctccagggtg ccctcaccg gctcctcgtc  
 19200  
 gacgagttcg gcgtctcccc caccatctc atcggccact cgatcggcga gatctcggcg  
 19260  
 gccacacgg cggggatcct caccctcgac gacgcctgcc ggctggctgc cgcccgcggc  
 19320  
 actctcatgc agaccctccc cgccaccggc gcgatgacgg ccgtcgaggc gaccgaggaa  
 19380

gaggtgctcc cgcacctcac ggagcgggtc ggtatcgccg cggtgaacgg cccgcgttcg  
 19440  
 gtggtcgtct ccggggacga agccgctgtc atcgccgtcg gcgaggagt ccccggtcag  
 19500  
 ggacgacgca tccgccgtct caccgtcagc cacgccttcc actcgcacca catggacccg  
 19560  
 atgctcggcg agctccacgc ggtggccgac acgctgacct accacgtgcc acgcaccccc  
 19620  
 ctcgtctcca ccgtcaccgg ccgcctggcc ggctccgaga tcaccagcgc cacctactgg  
 19680  
 agcgatcacg cccgcaacgc caccgcttc cacgacggcc tcaacacgct tcacgagcag  
 19740  
 ggcgtcacca cgtacatcga ggtcggcccc gacgccgtac tcgccgcgct gaccctgtaa  
 19800  
 gcgctgcccc acgccaccgc cgtaccgctg atccgggcca aggcctccga gccggccact  
 19860  
 ctgctcgacg ggctcgtccg ggctcacgtg tccggcgcca cggtcgactg ggcagggttc  
 19920  
 ctcgcacgac gcgggggacg gagcgtcgac ctccccacgt acgccttcca gcgtcggcgc  
 19980  
 cactggctgg agaccgccga ccccgctcgtt acggccgccc gcctcggctt ggagtccgcc  
 20040  
 tcccaccgcg tgctggccac gaccaccgaa ctcccggacg gaaccgcctt gttcacgggg  
 20100  
 cgggtgacgc tggccgacca cccgtggctc agcgaccaca ccgtcatggg aacggtcac  
 20160  
 ctcccgggca cggccttcgt ggagctcgcg ctccacgcgg ccgagacggt gggctctcag  
 20220  
 gagatagcgg aactcgtcct gcacgcgccg gtcaccttcg ggtcccagtc cgcggccctt  
 20280  
 ctccaggtga tcgtcggacc tgacgacccg tcggcgggcc ggaccctcac catcagggtcc  
 20340  
 cgttcggaa aggaccagtc ctggaccgag aacgcgaccg gcacgctcgg cgcactggtg  
 20400  
 ggggtctcct gacccccggc ggcagcgcga ggccctgaca agcgcccaag gccgcccga  
 20460  
 ccctctcagt ggaggggtgcg ggcggccttg ggcgttcggc gttgtcgggg gttcggccac  
 20520  
 cggcgcgggg cgctcggctt cacggacccg cgggcccggg cccagagccc gaccgggccc  
 20580  
 cgtgaacagc cacggccgtc ccactcccc ggctgcgagg gtatggacgg ccgtcgggcg  
 20640  
 acggggcggg tggacgacgg gcgggacgac cgcgggcccg ctcgcgacgc acgaagtccc  
 20700

cctcctcgcc ggcgccggcg ggaaggggga ctcgtgggac cgatcaggag aagcggagcg  
20760  
gggtgaacgg cgccgtcgcg gcgtgcggga cgggtgctggt caggtaggcc gccatggcca  
20820  
cggccgtgac gggagccagc tggactccca tccggaagtg accgctcgcc aggtagacat  
20880  
cgggaaccgt ggtcggaccc atgacgggga ggtcgtcggg cgaggcggga cggagccccg  
20940  
cggtgatctc ggcgaaacgc agggcgccca cctcgggaat gacctcggcg gcgcggcgca  
21000  
gcaggctgcc ggtgccctcg gcggtgacgg tctcgtcgta gccgcgctcc tcgtaggtgg  
21060  
cgccgacgac gagttcaccg tccaggcggg gagccaggta gagggacttg cccttgagaga  
21120  
tggcgcgggc cgtcacgttg agcagtggca cgtcggagcg caggcggagg atctggccct  
21180  
tgacgggacg gatctccggg atcgcgcccc cgggcagccc ctcgatctgg tgggtccagc  
21240  
agccggcggc caggacgaga cggtcgaagg gcaccttgtc accggtgtcg aggacgacgg  
21300  
tgtggtcctc gatccgggtg gctcccctgc gtacgacgct gccgcccagg gcttcgatgg  
21360  
cggtgatcag cgccgcgtcg agaacgcgcg ggtcgatggc gccgtcgtcc ggggacagca  
21420  
ggccgcccac cacggggggc aaggccggct ggagcttggc gcactcctcg gcggtcagcc  
21480  
gttcgggtcg cacaccgatg gacttctgga aggccagttg gcggtcaggg atcggcaggg  
21540  
attcctcgtc gaacgccgcg tccaggacgc cgtcgcgccc gaagccggcc ggagcgaact  
21600  
cctccagttc ctcgacgaag gacgggtaca gctcacggga ggccaggcac aggcgcagga  
21660  
gttcgggtg gtcgtacagc tggtcggtga tggccgggag cagtcccgcc gaggcgtggg  
21720  
acgccttgct gcccggggca gggtcgatca ccgtgacgga gacgccttcg cgcgcggccc  
21780  
gccatgcggt ggccaggccg atgacaccac cgccgacgac aacagttctc atgaattgct  
21840  
ctcctggagg agtcgacgca ggtgggaggc gagttctgcg ggggtgggat ggtcgaagac  
21900  
gagggtcgcc gggagccgca ggccgggtggc ggccgccagg tggcggctga gttcagagtt  
21960  
gccgatggag tcgagcccgg cggcgcgtgaa ctcccgtgc gggtcgatct cgtcgcgctc  
22020

ctcgtagccg aagaccaagg cggacaccga gcggacgaac tccagggcgg cggcgtcgcg  
22080

ttcggcctcg ggcagctcgg ccagccggct gaccaggggtg tcggtcacgg cgggcttgcc  
22140

ggtacgggtg ggcgccagct cctccaggac gggcgggatg tccccggtga ggccccctc  
22200

gttgagccgg acgggcgcca ggaccgcctc gtcgctgccg agcgtggcgt cgaagagagc  
22260

caggccctgg tcgtgccccca ggggggagaa gccggaccgc ttgatgcggt tgtggtcggc  
22320

ggcggacagg tcggagccca tgccgttctc gctctccac agaccccagg cgagcgagag  
22380

gccctgcagt ctgcgggcac ggcggtgggtg ggcgagcgcg tcgaggacgg cgttggcggc  
22440

tgcgtagttg ccctgccccg cgttgccgac cagccccgcg acggacgaga acaccacgaa  
22500

cgcgcagtcc gggtcctgga tcaggtcgtg caggtggagc gcggcgtcga ccttgggacg  
22560

cagaacctg tcgagtcgcg gtggcgtgag ggtgccgatc gcgccgtcgg acagcacacc  
22620

cgcggtgtgg atgacggcgg acggggcggg gatgccggcc agtgcgctct ccagcgacga  
22680

gcggtcggcg acgtcgcagg cgaccacgtc gacacgggcg ccgagcgcgg tgagttcggc  
22740

atccagctcg gcggcgccgg gcgagccggg accgcggcgg ctgagcagca cgaggtgccg  
22800

gacgtcatgg ccggtgacga ggtgccgggc gagcaggcgg cccagttctc cggtgccgcc  
22860

ggtgatcacg acggtacccg tgaggctcct gcgccgcggc ggcggtgggc tgcgtacgag  
22920

acgcggcctg agcagcgcgc cgttgccgac ggcgagttgg ggttcgccgg aggcgacggc  
22980

ggcggcgagg gagcggcccc agttcccggg gtcgtcgacg tcgaccagga cgaagcggtc  
23040

ggggtgttcg gactggacgc tgcggagcag tccccagacg gcggcgcccc acacgtcggg  
23100

caggtcctcc ccggcgcggg cggcgacggc tccgctactg accagcacca ggcgggtcct  
23160

ggcgaggcgg tggtcggcca ggcactcctg tacgaccatg agggcgcggt gcgccgccga  
23220

gcgtaccggc gagtcctcgt ccgtgcagga gacgacgatc acgtccggga cggctgtggc  
23280

ggccccgagc acggagtcga gttcgcgcag cgtggggtag gagtcgaaca aggggcgggt  
23340



ggccttcagc gcgccggtga ggccgatgcg gtcggttccg aggaatcccc agcgcctgctg  
 23400  
 gtcggagtgg tccgactgct gtacgacgga cttccatacg acgcgagca gttgggtgcg  
 23460  
 ctgcccgcg gcgtcgagct gctcggccgt gacgggcccg ctgacgacgg aaccgatcgt  
 23520  
 ggcgatcggc tcgccgtcgt cgccggtgag cgccagggag aagccgtcgt cgcccgggac  
 23580  
 ggtgtggatg cggacggcgc gggcgcccgc cttgtgcacc cgtatcccgc gcagcgcgaa  
 23640  
 cggcaggaac gtccggtcgt cgcggtcgt gtcggggacc agtgccgctt ggagcgcggt  
 23700  
 gtcgagcagg gccgggtgca ggaggtagtc gccgccggtc gcggtgctga gcgctgagtc  
 23760  
 ggcgtagacc tcctcgccgt gcgaccacac ggcgagcagg ccccgaagg ccggcccgta  
 23820  
 ctggagaccg cggtcggcga agcggctgta gaggtgctcg atgtcgacct cgcggactcc  
 23880  
 gtcaggaagc ccgggcccgt ctgctttact cattgtcgt tccttcgca agtccgggcg  
 23940  
 ggggtcagat cacgggcccgt ttggtgacgg cggcgggtgat ctcgccgtgg ctctgccac  
 24000  
 cttgctcgtc ggcgtagtgg cacttggtcc gcaggatcac cagggcgtcg cgcaggtcgt  
 24060  
 tggcgtccca ttcggcgatg tcgatgatct cgatctcgc ggtgaagcgg cggcccctgca  
 24120  
 tcgctgctgc gaagctgctc ttgaagtcgg tgatgaacac gtcggcgagc tgccgggtcc  
 24180  
 agaactgctc gatcgtccag ccggagaacg gtcgacgag tgactcccgt accgacatgg  
 24240  
 ccagcaggtg tagggccatc tggttgaagc agatgttgaa ctcgaccgag ttgaagtgcc  
 24300  
 cgggtgctgc gatgtagcag gactcgggga tctcgaaggt gcaggtggcg atgagtcgtc  
 24360  
 cgccgtcgcg ggggtcggc tccgaggtga ccgtggccga ggtgaggtac tcgcaccgct  
 24420  
 tggcccggta gggggtcagc acccgggtgca gcagttcgtc gtcgggtgggg aagacgcgcc  
 24480  
 ggtcgggtgc ttcgagaagc tgggtcatgt cgggtctcct gatcgtcgcg cgggtcggag  
 24540  
 aaggtcggcc aggtcatgt tctggatcgc ggcggtgcgg tccgcggcac cctcggccga  
 24600  
 ctcgggctcc ggttcgttct cgggtacgct gatcgccatt tcctcaagca ggtgcccagt  
 24660

gaggacgagc gaggtcgggt ggtcgaagac gagcgtcgcg ggcagggcgtt ccccgatgac  
24720

ctcgccgagg gcgttgcgga accgcacggc ggcgagggag tcgaagccca tgtcgcggaa  
24780

cgaccggctc gcgtcgcacct cgtcggcgtc gccgaagccg aggggtggtg cggcctgcgc  
24840

ccggacgatg tcgagcagtg tctgttcgcg gcgggacgga gtcagccgtg ccagccggtc  
24900

gcgcaggctc tcggccgggt cggcggcccc tgcggcggga tcgactacgc ggcgggacgg  
24960

gccgcgtacg agatcgcgca tcagataggg cacgtcggac cggctctgcc cgctcaggtc  
25020

gagccggacg ggcgccagga acgcccggtc gagcgcgccg gcggcgtcca gcatcgccag  
25080

gccctcgtcg gagccgatgg gcaggacacc gtcgcgcacc atgccggcca cgtcggcgtc  
25140

ggccagttcg ccggatcatg cgtcggcctc ttcccacagt ccccaggcca gggactgtgc  
25200

cgagagcccc ttcgcgcgcc ggtgcgcggc gaggccgtcg aggaaggtgt tggcggcggc  
25260

gtagttgccc tgtccggcgc cgccggtcac gcccgcggcg gacgagaaca tcacgaacgc  
25320

cgagaggtct tgaccggcgg tgagttcgtg caggttcagc gccgcgtcga ccttggggcg  
25380

catcaccgcg gccattctct ccgggtgtcag cgaggagagg atcccgtcgt cgaggacgcc  
25440

cgcggtgtgg acgattccgg tgaggggtgc gtcccaccagc acggccgcca gcgcgtcgcg  
25500

gtcggcggtg tcgcaggcga ccacctcgc acgggcgccg agcgcgggtga gttcggcgtc  
25560

cagttcggcg gcaccgggtg cggcggggcc gcggcggctg gtgagcagca ggtcggtgac  
25620

gccgtgggtg gtgaccaggt gcctggcgc gaggctgccc agggtgcccg tgccgccggt  
25680

gatcagtacg gtgccgggcc ggtcggccga gacgggtggcg gcggtgccgc tcaggtcccc  
25740

ggtgacgcgg gccagtcgcg ggcgggtggag ggtgccggag cgtatggcga tctgggcctc  
25800

gcccgtggcc accgccgcgt cgatgtcggc gtccgtgtgg tcggcgcggt cgtcgacggc  
25860

gtccaggtcg atcaggacga cgcgtccggg gttctccgtc tgagcgtgc gtacgagtcc  
25920

ccagacggcc gctccggcga ggtcggtgac gtcctcggcg ttcacggaca gggcggccgcg  
25980

ggtgacgacg gcgagcacgg cgtcgtcggc gccgctctcc agccaggact ggaggacggc  
 26040  
 gagtgtcgtg gcggtctcgg cgtgggccga cgcggcgggtg gtgccgcccg tgcagtgggtg  
 26100  
 caccgacggtc cggggggcgg tgcccaccagg tacgggtgcg gcgggtgccc agaccagctc  
 26160  
 gaacagcgag tcgtcgtacg cgccgggtgcg gatcggcccc cccggggcca ccggacgcag  
 26220  
 tacgagcgac tcgacggagg cgacggggcg gcccgcggtg tcggtgaggc gcagcgccac  
 26280  
 ctcgtcatca cccaggctgg tcatccgcag tctgagcgcg gcggctcccg aggcgtgcag  
 26340  
 gtcgacccgt gtccaggcga acggcagggg gaccctttgg gcgtcgggtg cgagcagtcc  
 26400  
 ggtcgcgtgg agagcggcgt cgagtgcggc cgggtgcagc ccgaaccgct gggcctcggc  
 26460  
 ggcggcctgc tcggggagcg tgatctccgc gaacacctcg tcgccggagc gccaggcggc  
 26520  
 acgcaaacc tggaacaccg gcccgtaggc gaggccgccc tcggcggccg gctcgtagaa  
 26580  
 gccggtgagg tcgacgggct cggcgccccg cggcggccag gccgccggct cgtgggccgg  
 26640  
 ttcgggcgcg gcgtcgggtg gcgtgccctg ggcgtggcgt gtccagaggg tgcggcggg  
 26700  
 ggcgttgtcg ggccgcgagt ggatcgacac ggtacgcaga tcggcgtcga ccacgacctg  
 26760  
 gatggccacg ccgccgcgct cggggagagc cagcgggtgct tcgagtgtca gttcttccac  
 26820  
 gcggcccccg ccgacctcgt caccggcccc gacggccagt tccacgaagc ctgacccggg  
 26880  
 gaacaggacg gtgccgccga cgacgtggtc ggcgagccat ggctggcccg agagcgagag  
 26940  
 gcgtccgggtg agcagtactc ccccggcccc ggcgaccggc acggccgccc cgagcagcgg  
 27000  
 gtggtcgggtg ggtatctggc cgatgcccac ggcgtcgccg gtgtgctcgg acaggtgcag  
 27060  
 ccagtaccgc ttgcgctgga aggggtaggt gggcagttcg acgtgctggg ccccggttcc  
 27120  
 gacgaactgg gtgtcccagt cgacggtcgc gccccgcacg tacgcctcgg ccagcgaggt  
 27180  
 gacgaactgg cgcgggctgc cggagtcgcg gcgcagtgtg ccgacgagtg tggcggccac  
 27240  
 cccggccgac tcggcgggtc cgtcaggtcc ggtcagcagg cccgggtgcg ggctggcctc  
 27300

gatgaaggtc tcgaagccgt cggcgagcag tgccccgctc gtgtcctcga accggacggt  
27360  
ctggcgagg ttggtgtacc agtactcggc gtcgagaccg gcggtgtcca gggttccccg  
27420  
ggtgacggtg gagtagaacg ggatggacga ggtacggggt gtcaggccgc tcagttcggg  
27480  
gagcagccgg tcgcggatct cctccacgaa catcgagtgg gaggcgtagt ccaccgggat  
27540  
acggcggggac tggacctcgt cgggtgtccag ccggccccgc agttcgtcga gggcctccgg  
27600  
gccgccgcac acgacgacgg agctcgggcc gttgacgacg gcgagttgca ggcggccgctc  
27660  
ccagtcggcc aggtactcac tcgcgcggtc ggcggacagg gcgaccgaca tcatgccgcc  
27720  
ccggccccgc aggtcctgcc ggatgacgcg gctgcgcagg gcgaccacgc gtgcgccgctc  
27780  
ctccagcgac agcgcgccgg cgacacaggc ggcggcgatc tcgccctggg agtggccgac  
27840  
gacggcggcc ggttcgacac cggccgcgcg ccaggcctcg gccagcgaca ccatcagcgc  
27900  
ccacagcgtg ggctggacga cgtcgacggc gtccagggcc ggtgcgccgg gcgcgccggtg  
27960  
gaccacgtcg agcaggttcc agtcgacgta cgattcgagt gccttggcgc attcgcccag  
28020  
gcgtgtcgcg aacgccggtg agtgctccag cagttcgacg cccatgccct gccactgcga  
28080  
gccctgtcct gggaagacca gcacggcgcg cgagtcgccg cgttgcaggc cctggaccac  
28140  
ggcggcggac gagttccccg cggccagcgc gtcgagtcgg gcgagcaggc cgtcgcggtc  
28200  
ggcgtccacg atcacggccc ggtggtccag ggcggcgcgg cctcgggcaa gcgtggcggc  
28260  
ggcgtccagt tcgtcgacgt cggcgttctc gcgcaggtgg gcggcgatgc gcgccgctg  
28320  
ggcgcgcac gcctcggggg tcgagcccga cacggtcagc gggacgaccg gaggccgctc  
28380  
gccgcgcggt tcgccccggc cggccggcgg cgctcgtcgc atgatcacgt gggcgttggt  
28440  
gccgctgacg ccgaaggagg agacggcggc gcggcgcggg cggccgctcg ccgtccagtc  
28500  
cctgggctcg gtgagcagtt cgacggcgcc ctcggctcag tcgatgtgcg gagagggggg  
28560  
ctcggcatgc aggggtgcgc gcagtacgcc gtggcgcac gccatcacca tcttgatgac  
28620

gcccggcgacg ccggcggcgg cctgggtgtg cccgaggttg gacttgaccg agccgagcca  
28680  
cagcggatcg tcgcggcgct cctggccgta ggtggccatc agcgcctgtg cctcgatcgg  
28740  
gtcgccgagg gtggtgccgg tgccgtggcc ctcgaccgcg tcgatgtcgg cgagtgtcag  
28800  
accggcggag gccacggcct gccggatcac ccgctgctgg gcggggccgt tgggggccg  
28860  
gatgccgttg gacgcgccgt cctggttgat cgccgagccg cggatcaccg cgaccaccgg  
28920  
gtgaccgttg cggcgggctg ccgacaggcg ttcgagcacc aggagtccgg cgccctcggc  
28980  
ccagccggtg ccgtcggcgc cggcggccta cgccttgtag cggccgtcgg cggcgaggcc  
29040  
gctctgcagg ctcatcccga cgaacgtctc gggcgtggcc atcacggtga cgccgccggc  
29100  
cagggcgagc gtgcactcgc cctgccgcag cgactgcacg gccagtgca tggccaccag  
29160  
ggaggaggag caggcgggtg cgatggtgac cgcagggccc tggagacca ggggttaggc  
29220  
gatgcgtccg gaggcgatgc tgccgaggcc gccccggcg gtgaagcccg cgaggtcctc  
29280  
ggggaaggcg ccggagcggg tcgtccagtc gtggttcatg acaccggcga acacaccggc  
29340  
ggcggtgccg tgcagggagt gcggatcgat gccgccctgt tcgagggcct cccaggcgac  
29400  
ttcgaggagc atgcgggtgct ggggggtccat cgcctgggcc tccttcgggc tgatgccgaa  
29460  
gaaggcgggg tcgaactcgg cgacgtcgtg caggaagccg ccctcacggg tgtacgtggt  
29520  
gccgtgcttg ccgggctcgg ggtcgtagat gtcgtcggtg tgccagtccc ggtcgtccgg  
29580  
gaactcgggt atggcgtcgg tgccgccggc gaccagccgc cacagttcct cggcggagggt  
29640  
cacgccgccg gggtagcggc agcccatcgc cacgatcgcg atgggctcgt cggtcgcgggt  
29700  
ccgcacgggt ggccgcggcg ccgcggtgac gggtagcggc ccgcgtacgg ttctgagcac  
29760  
gtggtcgacc agggcgcgcg gggtcgggtg gtcgaagatc agcgtggagg tgagccgcag  
29820  
cccgtgggcc gtaccgagcg cgttgccgag ttcgatggcg gccagtgagt cgaagcccag  
29880  
ctcgggtgaag gcgcggcggg ggtcgatcgc ggtgccgctg tcgtgtccga gtacggcggc  
29940

gatctgggtc cggaccaggt cgaggacgat gcgttcgctg tcggcgtccg gctttccggc  
30000  
gagccggctg gccagtgcgc cgccggggcc ggtcgcggcc gccgcggcgg tgttccggcg  
30060  
ggcggggcgg cggaccagcc cccgcagcat cggcgggatg ccttcggagc gggcccgcag  
30120  
cgcgcggggc tcgacgcgca gcgggacgag tgccggcgca ccggtgacga gcgcctggtc  
30180  
cagcagggcg aggttctcct cggcgtccag ggcggggagc ccggaacgct cggcccggcg  
30240  
gagggtgacc tcgtccaggc gggcgcccat tcccgcgtcg cccgcccaca ggttccacgc  
30300  
gagggaggtc gcggggaggc cctgggtgac gcggtgcgcg gcgagggcgt ccaggaaggc  
30360  
gttgccggcg gcgtagttgc cctgcccggc cgcctccatg gtcccggccg cggaggagaa  
30420  
caggacgaag gcggcgagcg ggaggcccg ggtcagctcg tgcagatgcc aggcgccgtc  
30480  
ggccttgccg tggaagacgg tgtggacctg ctcgggcgtg agcgatccgg cgagcccgtc  
30540  
gtcgacgact ccggcgctgt ggacgacacc cgtcagcggg tgctcgccgg ggaccgtctc  
30600  
cagcagcgcg gccagggcgg cgcggtcggc gacatcacag gcggcgaccg tggcgtgggc  
30660  
gcccaagccg gccagttcgg cgacgagttc ggcggcgccg gggctgtcgg ggccgcgtcg  
30720  
cccgtgagc agcaggtgcc gcacgccgtg ctgttgtacg aggtggcggg cgaccacgcc  
30780  
gccagaccg ccggtgccgc cggtgatcag caccgtgccg tccgggtccc acgcgacggg  
30840  
gcgcgggtcc tcgggtgcgg gcgtcctggc caggcggggc acgagtacct tgccgtcgcg  
30900  
cagggcgagt tggggttcgc cggtggccag ggcggcgcgc agttcgtgtt cgggcacgtc  
30960  
gtggtcggtg ccgacgagga cgaagcggcc ggggttctcg gcctggggcg agcgcacgag  
31020  
cgccacggg gacgcgaggg tgaagtcgta cggctcgccg cgggtgacga tcgcgagggc  
31080  
ggtgtcgtcg cgcagcgcgt cctggatgag ggtgagggcg tgcagtgtcg ccgcccgggc  
31140  
cgcgtcggcg ggttcgccgg ggaaggtctc ggggacgtac gccagttgga actcgtgcga  
31200  
cgcggcgtcc ggctcggacg gggcctggac ggcggtccat tcgacgcgca gcagcgagcc  
31260

gccgcccag gcgacggcc gtacgccgtt gatcggcacg gggcgctgga cgagcgagcg  
 31320  
 gacggtggcg acgggtgccc cgttcgcgtc ggcgagttcg accgtgacct cgtccggggc  
 31380  
 ggtctggacg acgtggacgc gtacggcggg ggcgccggtg gcgtggagtt cgacaccgcg  
 31440  
 ccaggcgaac ggcagccgca cctggccgtc gtcctccggg ctttcggcgg cgtgcagtga  
 31500  
 ggagtccagc agcgcggggg gcagtccgaa ccggtcgcct ccggcgctcg cgagggcgac  
 31560  
 ctcggcgaac acgtcgtccc cgcggcgcca cgcggcgcg cgcgctcgga acgcgggccc  
 31620  
 gtagccgtac ccctgcgcgg cctggcgctc gtagaggccg tcgacgtcga tcggctcggc  
 31680  
 gtcggccggc ggccactggg cgaacacggc cgggtcggcc gtcgcgccct cgccgtgctc  
 31740  
 ggtcagcacc ccgctggcgt ggcgggtcca gttgtcctcg ctccgcgcgg ggcgcgagta  
 31800  
 gatgtcgacc gagcggcggc ccgaggcgtc gagcgcgccg acgacgacct ggatctcgac  
 31860  
 tccggcgctg gtggtgaccg ccagcgggtg ctccagggtg agttcctcca ccgcggcga  
 31920  
 gccgcctcg tcgccggcgc ggacggccag ttcgacgagg gcggtgccgg gtacgaggac  
 31980  
 gacaccgcg atgacatggt cggccagcca gggctgtgcc cgggtggaca tccggcccgt  
 32040  
 caggatcact ccgtcggcgc ccgcgaggct gacggcgctg tccagcatcg ggtgcccgat  
 32100  
 cgtgccgggg tggctgctc cggcgtccag ccagtagcgc ctgcgctgga acgagtaggt  
 32160  
 gggcagcgc gtgcgggtgg cgcgcggcc ggcgaagaag cggtcccagt cgaccgcgag  
 32220  
 accgtgggcg tgcgccaggc cgagcgcggt gagactctgg cgctcctcgt cgtgcccgtc  
 32280  
 gcgcagcagg gcggcgaacg cggcgtcggc gccgccgctg gtcaggcagt cgcggcccat  
 32340  
 tgccgacagg acggcgtcgg ggcccagttc caggtaggtg gtgacgccct gggattcgag  
 32400  
 ccagctcatg gcgtcggcga accggaccgg ctccgcgacg tggcgtacct agtagtcggc  
 32460  
 gtcgaacgcg gcgacgggct cgccggtcag gttggacacc acggggatac gcggcggctg  
 32520  
 gtaggtcagg gtctcggcga cgcggcgga gtcgctcgagc atcggttcca tcagcggcga  
 32580

gtggaacgcg tgcgagacgc gcagccgccg cgtcttgctt ccctcctcgg tgaagcgggc  
 32640  
 ggcgacggcg agtacggcgt cggcttcacc ggacagcacg accgccagcg ggccgttgac  
 32700  
 ggcggcgatg cccaccttct cggtgagata cggggcgact tcggcctccg ccgcctggac  
 32760  
 ggcgaccatg gcgccgccgg cgggaagctc ctgcatcagc gtgccgcgcg cgcggaccag  
 32820  
 cagcgcggcg tcggccaggt tcatgacgcc ggcgacgtgc gcggcggcga tctcggcgat  
 32880  
 ggagtgtccg gtgaggagtc cgggcgtgac gccccacgac tccagcagcc ggaacatggc  
 32940  
 ggtctcgacg gcgaagaggg cggcctgcgc gtagcgggtc tgctggagca gttcggcctg  
 33000  
 cggcgaaccg ggctcggcga acagcacctc gcccaggggc tcgtcgaggt ggaggtcgag  
 33060  
 gtggtcggcg gcgtcctgga gcgcccgcgc gaacaccggg aaggcgcgca ccagttcgcg  
 33120  
 gcccattcca agacgctggc tgccctggcc ggtgaagagg aacgcggtac ggcccgcggc  
 33180  
 cggcgtcgta ccggtgagca gtccgggcgc gctctcgecc gcggcggagg cggtcagggc  
 33240  
 gcggcgcagt tcggcacggt cggcggccac gacggtggcg cggtcgcgca gcgcggcgcg  
 33300  
 cgtggtggcg gccgagtagg cgaggtccag cggcgaggcg tcgtcgaccg ccgtcaggag  
 33360  
 ccgttcggcc tggccgcgca gtgcgtcctg gccctgggcg gacaggacga ccggcacggg  
 33420  
 gcgctcggcc accgggggtt cgtcgctccc gtcggcggcg gcgacggggg cgggcggctc  
 33480  
 ctcgatgatg acgtgggcgt tggtgccact cgcgccgaac gaggagattc cggcacggcg  
 33540  
 cgggtgtgtc gcggcttccc aggggacggg ctcggtgagg agctggacgg tgccgtccga  
 33600  
 ccagtcgacg tgcgtggtcg gcgcgtccac gtgcaggggt gcgggcagcg tgctgtgacg  
 33660  
 catgcctgc accatcttga tgattccggc ggcgcccgcc gccgctccgg tgtggccgat  
 33720  
 gttggacttc accgaccca gccacagggg cttctcgggt gagcgtccgc gtccgtaggc  
 33780  
 ggcgaccagg gcctggccct cgatggggtc gccgagtgtg gtgccgggtc cgtgcgcctc  
 33840  
 gaccgcgtcg acctcgtcgg cggtcagacg ggcgtcgggt agcgcgcggc ggatgacgcg  
 33900



ctgctgcgcg aggccgttgg gggcggacag gccgttggtc gcgccgtcct ggttgatggc  
33960  
ggtgccgcgg atcacggcca gtacgggggtg cccatcgcgc agcgcgtcgg acaggcgggc  
34020  
gaggacgaac acgccgacgc cttcggcgaa cgagggtgccg tccgccgcgg cggcgaacgc  
34080  
cttgatgcgt ccgtcggggg ccagaccgcg ttggcggctg aagggtggtga acgtgccggg  
34140  
gctggatatc acggtgacgc cgccggccag cgccagcccg cactccccgc cctggagcga  
34200  
gcgcacggcc aggtggaggg ccgtgagcga gcccgagcac gcggtgtcga cggtgagtgt  
34260  
ggggccttcg aagccgaggg tgtaggcgac gcggccggac acgacgctgg gcaggttgcc  
34320  
gcccaggacg tagccgtcga gaccgtcggg ggccctcgtg gtccggggcc cgtactcgtg  
34380  
cggctccacg ccgatgaaca cgcccgaggg tgtgccgcgc aggcggcccc ggtcgatgcc  
34440  
cgcgtgttcg agcgttccc aggaggtctc cagcagcagc cgctgctggg ggtccatggc  
34500  
cagcgttcc ttgggactga tctggaagaa gtcggcgtcg aagtcggcgg cgtccgggag  
34560  
gaatccgccg cgccgcacat aactggcgcc ggtgacggac gggtcgggtg ttagatgcg  
34620  
gtcgagatcc cagccccggt cggagggaaa gtccgtgagc acatgggtgt cgtcggcgac  
34680  
gatccgccac aggtcctcgg gcgtgctcac accaccgggg aagcggcagc cgattccgac  
34740  
gatcgcgatc ggctcggcga acgccgacgg cagggcgggc gccgccgtct cgtcgggtgc  
34800  
gtccgctccg gcggtgccgt gcagcaggtc ggccaggtgg ccggcgaggg cgatcgggtg  
34860  
cgggtggtcg aaggcaaggg tgacgggcag gtccgcccc acctcggcgg cgagccgccg  
34920  
ctgcaaccgg accaggccca gcgagtccag gccccccgcc aggaacgggc ggtcggcggc  
34980  
caggtcgacg gcccttccc cgtaggcgtc cgggtcggcg gtccgcaaca ccgctgcggc  
35040  
ttgcgccagc accagttcca gcatcgcggc ccttgactgc tcggtcatgt gcatactccg  
35100  
caacgaagaa gaggcattga aacgaagaag acgcttgag tcgatcgagc gggtcaggac  
35160  
gccacgtggg cggccagggc ggtgcggctc accttgccgt tgacggtcac ggggaactcg  
35220

ttcaccacgc tcagaccgtg cgggaccatg taggcgggca gcgcggcgcg acagaacgcg  
 35280  
 aggacttctc tgggtgccgg cgcgggaccc gcggggggcg aggtgacgaa cgcgtgcagc  
 35340  
 agtgggtcgc cgtcctcgcg cggcaggacc atcgcgaccg cgccggtgac accggcgaac  
 35400  
 tggccgatac ggcgctccac ctcgccgagt tcgacgcggt tgccgcggat ctgcacctgg  
 35460  
 gagtgcacgc gaccgcggaa gtacagctcg ccctcgggccc ccatatgggc caggctgcgg  
 35520  
 gtcttgaaca ccacctgtcc cgatccgggg tgcagcgggt ccgggaggac gacggcgcgg  
 35580  
 gtggcctccg ggtcgcgcca gtagccggag aagagcgagg ggctgcgcag gtagatctcg  
 35640  
 ccgatgggtc cgggctcgtc gacggggctg ccgtcggacc cgaggagcat catctcggcg  
 35700  
 ccggcgcgag cgtaaccgat ggagaggcgt tcggtgccgg gcggcagcgg gttgggcacc  
 35760  
 tcggtgatgg acgcggccat cgtctcgggt gccccgtaac cgttggtgaa gcgcgccttc  
 35820  
 ggcagcagtt cctggaggcg ccgcagctcg tccaggggga agtcctcccc ggagaagggtg  
 35880  
 atgcggctga ggacgccctc ctggcccata tcggccagca ggtcggattc gtggcgcagt  
 35940  
 accgggcgcc agacggacgg gacaccgtcg acctgtgtga cccccgcata gcgcaggtac  
 36000  
 gaaaggaagc ggcggggcca gttgagccgg gcgcgcggca cgggcaccag cgtggcgcgg  
 36060  
 tggccgaggg cgaggccgat gtcgaacagc gcgaagtcga actggagcgg agaggtgttc  
 36120  
 gccaccgggt cctcggccgt gacgagccgt gcggcctcag cccccgcag gaacgcgatg  
 36180  
 accccgcggt ggctcatgac cacgcccttg ggccggccgg tcgtgcccga cgtgaagggtg  
 36240  
 atgtacgcgg tgtcgatggg cgtcaccacg cggcggcggc ggacgcgcgg cgcgggcgcc  
 36300  
 cgctccacgg tgaggccgcc ggggccgaag cgggcgggtg cgaccgtctc gggaatgccg  
 36360  
 gtgcgctgcc cgtggtcgga ctggatgtgc agcgcgggct ccgcgctgtc gatgatcgtc  
 36420  
 agcagccgct cgtccgggac ctcggggctg gtcgggggtga acggcagccc cagcgtggcg  
 36480  
 caggccagca gggtcgccac cgcgtcggcg ttgggtgtcgg attccaggat gacgcggctc  
 36540

cccacatcca ggccgagcgc gtcgagcgcg tcgacgaggc ggtcgacgcg ccgctccagt  
 36600  
 tcgccgtagg tgacggcttc cagaccaccg tccgccgccg cctggatcac ggcgggcccg  
 36660  
 tcgggggtccc tccggggcggc ggccagcaga tacgtccgca ggttgctccac cgacgcgtgc  
 36720  
 gcccggaccg gacgggctcc attcgggctc atgcgccacc ttcccgattc agcgtttccg  
 36780  
 agtaatgccc gccattttct aacaggtggg cttttcaact cgcaagaacc cctggccacc  
 36840  
 ggcgccccgaa ctagggggca ttaggggtta tgcggccagt agggacgcaa gaacactgac  
 36900  
 cgtacaacag gtgggatcga agtgccgggc tttcggaacg catccgatga gcccgacaaa  
 36960  
 tagggagagc gagaatatgt cggcgattat ctttcccgga atcggcccgg tccggctcgc  
 37020  
 cgactcggcg cggttcctgg tgacctatcc catagcccgc cgactcgtcg ctgagacgga  
 37080  
 ccgaatactg ggctattccc ttctcgacag ctatcgcgag gccgaagacc gcgacgacca  
 37140  
 gggggcgctc cccgagcccg cccggatcgc gttcctggtc cagtgcctgg cgctggcccga  
 37200  
 gtgggccgct aaggagaacg acctggaccg ggtcgtctgc gccggcgcca gcttcggcgg  
 37260  
 cacggcggca gcggtgcaact ccggcgcgct gtcgttcccc gaagccgtgg agatgaccgc  
 37320  
 cgcgtggggc cgccgagtcg acgactactt caccggtgag caccgcgaca tcgtcaccca  
 37380  
 gtcattcgcc cgcgtcgcgc ccgaccgct cgcggagatc caggccgagc tggacgcacg  
 37440  
 gggcgactgg aacgaggtgg cctgccaggt cgacaacgac ttccacatgc tgtcggtgcg  
 37500  
 cgaggacgtg gtcgagtggg tgcagggacg gtcgccgcgc gccggcggcc tgccgctgta  
 37560  
 cgtcatgcgg ccgccgatgc actcgacgct gttcggggcg ctgcccgaag agatcgcgaa  
 37620  
 cgggatcacc acggacatca cgttctccga tccccggatc cccgtggtgt ccgaccacga  
 37680  
 cgggtcgctg gtacggacgg gggccggggg gcggggagttg ctgctgaacg ccgtgacgca  
 37740  
 caccgtgcgg tggccggccg tcgtcgacac gatcaagggg ctcggcgtcg agcgggtgca  
 37800  
 tgtcaccggg caggacccc tgtggggacg ggtggatgct atgaccaacg cgttccaggt  
 37860

ggtggcggtg cgtccggaca cagctatgcg accgcgccgt cgcagcgcga tcgcatagcg  
 37920  
 gaaaagaatc cggtcagcgc atggcgtecca gggttttcca gagagccccg ggagtcgcga  
 37980  
 agttatccat gttgagcgca tcgtcgacga aacggactcc atatgctcc tccatcgagg  
 38040  
 acagcagtga gaccattccc atggagtcga gaccgcaatc gcgcaggctc aattcggcgg  
 38100  
 tcaactctc atccggttcc agcaaaggaa tctgtttgcg gagaagttgc tcgaatgatt  
 38160  
 cgtcccacat acgggctcct gtgttttccc gacgtttacc gaatagccgg cgtcgtgccc  
 38220  
 aacctaagcc ccgcgcccta ccggtcggca cccctactca cccctttccc tcttcggagc  
 38280  
 gaggacaatg aatcagacac ccgtccccgg acacggcctg cacgaacggt tcctgaccgg  
 38340  
 cctggcgctg tcgcccggcc ggaccgcgat ccgcgtgcac gccaccgaga gcctgacgta  
 38400  
 cgagcagatg cacgaactgg cgatgcgccg ggccgcggca ctgcgggcca tggctccgca  
 38460  
 agggccgcac aacgtcgccg tgctggcgga caagagcctg accgcttatg tcgggatcat  
 38520  
 cgccgcgctg tacgcgggcg ccaccgtcgt accgctcaac ccgcggttcc cggccgagcg  
 38580  
 caccgctcc atgctcatcg ccgccaacgt ctccaccgtc atcgtgatc cgatcggccc  
 38640  
 ctctcactc gcggagaccg agctggatct gccgctcctg gacgagggca ggacggggcc  
 38700  
 ctgctggac acgccggtgg ccgtcaacct ttccgatgtc gcgtacgtcc tgttcacctc  
 38760  
 gggctcgacg ggccgcccc aaggggtgcc gatcaccac ggggccaacc accactactt  
 38820  
 cgacctgctg gaccggcgt acgacttcag ccccgacgac gtgttctgcc agaacgtcgg  
 38880  
 actcaacttc gactgcgcca tgttcgagat gttctgcgcg tggggcaacg gggcgcaggt  
 38940  
 gcaccccgtc ccgcccgccg cccaccggga cctgccggcg ttcttggccg agcgggaagat  
 39000  
 gaccgtgtgg ttctccacc cagcggcat cacgttcac cggcggatgg gcggcctgac  
 39060  
 ccccggatcg atgcccacac tgcgctggac cttcttcgcc ggtgagggcg tgetgcacga  
 39120  
 ggacgcccgc gactggcacg tcgccgcacc ccagtcgaag atcgagaatc tgtacgggcc  
 39180

gaccgagctg accgtgacca tcaccgggca ccgctggtcg ccgaagacca ccgaggagca  
39240  
gaccgtgaac ggcggcgtgc cgatcgaaa ggtgcacccc ggccacgacc acctgctgct  
39300  
ggacgacgac ggcgagtcgg cggtggaggg cgaactgtgc gtcgccggac cgcagatgac  
39360  
acccggttac ctggacggcg acgacaaccg gggccgcttc ctcgagcacg ccggccgctcg  
39420  
ctggtaccgg accggcgacc gggtgccggc gctggacgac gacgagctga tctacctcgg  
39480  
ccggatggac gcccagggtc agatccaggg attccgggtc gaactggccg aggtcgacca  
39540  
tgtcgtccgg cagtgcaccg gtgtgcagaa cgcggccacc gtcacccggc cggcaccgaa  
39600  
cggcggactg gaactcgtcc tctactacac gggcgagcgc attccgtcgg cgacgctgcg  
39660  
ccgcgagctg gccgcgcacc tgcccgatcc gatggtgccc aagaccttc ggcacgtgcc  
39720  
ggagtcccg ctcaattcca accgcaaggt cgaccgggcg cagttggccc gggaggccgc  
39780  
cgcgctgtca gacggtcgtg cctgaccgga agtgacgggc gttctccgcg accgtccgta  
39840  
gcgctcggc gagcgactcg cccgtcgagg cgtgctccgg gtcgacgacc cactgcatca  
39900  
tcacgccgct gagcagcgcg tggtagaact gcccgaccgc ggtcgcggtg gggcccggca  
39960  
gctcttcgcc gccccacagg agccggacca ggccggtctg tgctgctgc tgcgccteta  
40020  
tgaagaagct gccgacctcg ggcacatggt cgcgctggga gatcgcgtcg aactgggccc  
40080  
cccacaccgg gcggtggcgc tcgaacagct cgatcacgcg cgtccaggcc acctcgaacc  
40140  
gcttgatcgg atcgtcgggg aggtctccca cgtctgccag agcactcttc agctcctggg  
40200  
cccactctc cagcgcctcc atgatggcg cgttgaggag tgcttcctg gtgccgtagt  
40260  
ggtagccgat cgaggcgaga ttcgtaccgg aagcctcgac gatgtcgcgc gcggtggtac  
40320  
gcgctaccc cttctcgtag aggcattgct tcgctccggc cagcagatcc tctcgggtgc  
40380  
ccatgccgag agtctagcca cccccctcag acatctgtct tgaacagatg tcctgaacag  
40440  
aatctgaatt agacgaacct cttatacaga tctagagctg aggccatgag tccgcaacgt  
40500

gcaaccttga gggactgggt cggcctcgcc gtccttgtcg tccctgtcct catgatgtcg  
 40560  
 atggacatga cggtgctgta ctccgcgctg ccgttctca gcgcgaccct ggaaccgagc  
 40620  
 gccaccgagc aactgtggat cgtggacatc tacgcgttca tgctcgccgg gctgctcatc  
 40680  
 gcgatgggca cactcgggta ccacatcggc cgccggcggc tgctgatcat cggcgcggtg  
 40740  
 gtgttcggcg cgtcgtcact ggcctccgcc tacgcgacca gcgccgagct tctgatcctc  
 40800  
 gcccgcgccg tgctcgggat gtccggcggc gtactcgcgc cgtccacgct ctcgctgatc  
 40860  
 cgcaacatgt tccaggatcc cggccagcgc cgtaccgcca tcgcggtatg gaccgcccgt  
 40920  
 ctctccggcg gcgccccct cggtcctgat gtgtcgggag tgctgctgga gcactactgg  
 40980  
 tggggctcgg tcttctgat caacatcccg gtgacgatcc tgatcgtggt gctcggcccc  
 41040  
 atcctcctgc cggagcaccg cgaccccag cccggccggt tcgacttctt cggtgccgtg  
 41100  
 ctgtcgtcgg ccgcgatgct tcccgtcatc tacggcatca aggaactcgc cgacgacggc  
 41160  
 ttcgactgga agtacgtggc ggtcaccgcc gccggcctgg tcatcggggg gctcttcgtc  
 41220  
 ctgcgccagc gcgcggcccc caatccgctg atcgacctga gcctcttccg cgaccggggg  
 41280  
 ttcaccgctt ccacatcggat caacctgggt gccctgttcg cgatgatcgg gttcctgctc  
 41340  
 ttcgcgaccc agtggatcca gctgggtccac gggctgaatc cgctggaggc gggcctctgg  
 41400  
 aactgcccc cgccgttggc ggtggcggtc acgacatcgg tcgccgtcgg gctggcgaag  
 41460  
 aagatccgcc ccggctacat catggccatc ggcattgtca tcgcgtcggc gggattcggc  
 41520  
 atcatgacgc aactgcgcgc cgattcagac ctggccatgg cggatgatcgg cgcgagcgtg  
 41580  
 ctgtcggccg gcgtcggcat ggcgatcccc ctgaccgccg acctgatcgt ctccgcggct  
 41640  
 ccggaggacc gcgtgggccc tgccgccgcg ctgcccgaga ccgccaacca gctcggcggg  
 41700  
 gcgctggggc tagcgtatcct cggcagcatc ggtgccgccg tgtacacccc tgacgtcggc  
 41760  
 gacgtgacga cggggctgcc acccgaggcc gcggaggcag cggagggttc gctcggcggc  
 41820

gcgacggaag tggcceaaca cctgcccggg gacacgggcg acgcccctcgt cacgtccgcc  
 41880  
 ggggaggcct tcacccgcgg catgaacctc agcgccgcgg tgggcggcgt cgtcatgctg  
 41940  
 ctcgggtgccg cgggcgcggc gctgctcctg cgccatgtca agactcccac cgtcacgtcc  
 42000  
 gcgcccggcg acgagacgaa gggcgagacg gcggacgagc cctcaccctg cccaagtag  
 42060  
 tgaccgcccgc ccggtagcgg cgcccaacca gaaggggtcc ccgcgcgaaag attcgcgcgg  
 42120  
 ggaccccttc tggcgtttct ggtacggggc tgtcagcggc cgagagtcac cggaagcgag  
 42180  
 gtgtaaccgg tcaggcccat gatggcgcgg cgcccggggg tgcccgccag ctcgatgttc  
 42240  
 gtgaacgtgt cgatcagctt cgggaagagg tcggcggcct ccatacgggc gaggctgccg  
 42300  
 ccgaagcaga agtggccgcc ggcgctgaag ctcaggtggg cgccgttgtc gcggctcagg  
 42360  
 tccagacggt gcgggtcggg gaagcgcgcc gggccccggt tggcggcgga gagcagggcc  
 42420  
 aggacgagaa cgccctcggg aatgtcgggt ccgccgatgg tgatggggcg cgtcgtcaga  
 42480  
 cggctcgacg ccgtggtgtg ggcggtgtga cgcagcagct cctcgacggc cgtcgggggtg  
 42540  
 atgctcttgt ccgcgcgcca gcgcttcagc tcgtcggggg gctcgatgag cgcgagcacg  
 42600  
 ccggtcgcga tgaggttggt ggtggccgcg aagcccgccg tgaacaggaa gaggatgagc  
 42660  
 gccatcagct cctcctcgtc gagcttcccg ttggccgcct cctggacgag ggcggacatc  
 42720  
 aggtcgtcct tgggttcggc gcggaccgcc ttgatcacgt cattgaagta cccggtcagc  
 42780  
 tcctcggcgg cggcgtcggc cgccgccagg tcctcatcgg tgtagacacc ggagaagacc  
 42840  
 cgtgaccagt cgtcggccag ctcccagggt cgcttgccgt cctcgtacgg aaggccgagc  
 42900  
 atgtcgctga tgaccgacgac ggggaagggc atggccagca gctcgacgat gtcgaccggc  
 42960  
 tcgcccggcg cggacctctc gacgagttcg ttgatcagct cgtcgggtccg cttctccacc  
 43020  
 gcgggctgca tcttcttgat cttgctgggg gtgaacaccc gggcggccag gccgcgaacc  
 43080  
 cgtccgtggt cgggcgcggt gagtgtgacc atcgagttca ggtacatacg cagggagatg  
 43140

tgctcggccc agtcctcgcg catctgagcg gcggcacggg cccactgtg gacatcgggc  
 43200  
 atcttcagga gctcggtcac ctctcgtac ccggagagcg cgtagatacc gagcgcggac  
 43260  
 ttgtggaccc ggttcaccga ctgcaaggtc tcgtagatcg ggaaggggtc gtcggggaag  
 43320  
 ggcggcgaca gcagcttcat cagcgccttcg tcggcctgct ggggtggggtc cgtcgccttcg  
 43380  
 gtggtcacgt ggtcgttctc ctcgttatgg gtggtgggtc cggtcggca ggttctacga  
 43440  
 cgcggaccgc gtatcgagct cgtcgtcagag gacgtcgaac agctcatcgg ccgagagcga  
 43500  
 cccgagttcg gcctcgtcgc tctcgtcccg cctgtgggtc tcggtccagc cggccgcccag  
 43560  
 ggcgcgcagc cggtcggcga cgcggccgaa cgtctcctcg tcgggcgtca ccgacgcca  
 43620  
 cgccgcttcc agacgggcca gttccgcctc cagcgggtgac ccggcgggct cctggtcggc  
 43680  
 ggtcagctgt tcgcgaggg actcggcgac ggcggaaggg gtgggatggt cgaagaccag  
 43740  
 ggtggccggc agcttcagcc ccgtggccga acgcagcgtg ttgcgagtt ccacggcgg  
 43800  
 gagcagtcg aatccgaggt ccttgaacgc ccggtcggag cggaccgcgt ccacgccgtc  
 43860  
 gtgaccgagc accgccgcca cgtgggtacg gaccaactcg accaggatgc tgtcgcgcgc  
 43920  
 ggcttcgtcg gccgccgccc ccagggtccg tcgcagtccc gcgccgccgt ccccgcccgc  
 43980  
 ggcgcgcgcc gcgccccgg cgcgacggcg cgatgtgggc accagaccac gcagcagcga  
 44040  
 cggcagttgg tcgacggcgc ggggtccgag gggccctgtg tccaggcggga acggcgcccag  
 44100  
 cgcgggctcc cccgtacgca acgcgtcgtc cagcagcggc aggccctcct tctcgcagcag  
 44160  
 ggcgggcagg cccatgcgga gcatgcgctt gcggtcggcg tcggtcagct caccgagccc  
 44220  
 ggtctcggct tcccacagac cgaacgcag tgaggtcacc ggcaggccgg cctggtggcg  
 44280  
 gtggtgggcg agcgcgtcca ggaaaacgtt cgccgcggcg tagttgccct ggcccggcg  
 44340  
 cagcagcagg ccgcccacgc aggagaacag gacgaacgcg gacaggtcgc ggtcggcagt  
 44400  
 cagctcgtgc agatgccagg cgccgtccac cttgggcccgc aggaccgtgt ccatacgtc  
 44460



cggcgtcagc ccgcggaacca gcccgttgct caccacaccg gccgcgtgga cgaccgcccc  
 44520  
 gacgggatgc cggtcgagca ccgcggccag cgcggcgcgg tcggccacgt cgcacgcctc  
 44580  
 gatgtccacc cgcgcgccgt ggccggtcag ttcggcgcgc agttccgccg cgcccgggtg  
 44640  
 gctccccccc cgacggctgg tcaggaccag atgccggatg ccgtgctccg tcaccagggtg  
 44700  
 gcgggcgacc agggcaccga gaccgctggt gccgccggtg atcagcacgg tcggcgacga  
 44760  
 ctcccacggg tcgcgtccgg cgcccaccgc cgagggcgggt acgcgggtga gggcgggtac  
 44820  
 gagaagtctc ccgcctcgaa ccgccacctc ggggtgcgtc acgaagggcg gaacgggtggc  
 44880  
 cccgtcggcc aggtcgagca gaaggaaccg gccgggattc tccgccgccg ccgcccggac  
 44940  
 cagccccac acgggtgcct ggctcagatc gacgtcctc ccctcgatc gcaccgcgcg  
 45000  
 acgggtgacg acggcgagct tctcatcgt cctgcccttg tccgccagcc attcctggac  
 45060  
 gcgcgccagc acctcgtcgg cgaccgcgtg cgcggcctc gatgtgtcac ctccggcgcg  
 45120  
 cgggacctc tacaccacga cgggtgtctc cagtacgggc acgtcgccgg ccttggtcca  
 45180  
 ggtgagggcg aacagcgact cgcggtggtc gccgtcggcc cgtagctgct cggccgacac  
 45240  
 gggccgcgag gtcaggctc cgacggagag cacgggcgcg ccggttccgt cggcgaccag  
 45300  
 aacctccgt ccgtcgcgc cctccgggtt cgacagacgg acgcgcagcg cggaggcggc  
 45360  
 ggcgcggtgg agcgtgacac cgttccagga gaacggcagc tccggcgcgt cgggtggcggc  
 45420  
 gccctcctc atgagacca cgtgcatcgc cgcgtccagc aacgccggat gcaggccgaa  
 45480  
 cctggcagcc tccgttcct ccgggagcga gacctggcg aacacgacgc cgtcaccgtc  
 45540  
 ccgccaggcc gcttcagcc cctggaacgt ggggccgtag tcgtaaccgc gggcgaggag  
 45600  
 ccgctcgtag gcgccgtcca ccgggagcgg cgtggcggcc atcggcggcc actggatcag  
 45660  
 gtcggacgag ggagatacgg ccgacgggag gagggcaccg gccgcgttgc gggccagat  
 45720  
 ctcgtcgtcc agggacgagt agatctcgac ggtgcgtgac tcggagtcgt cggggccgcc  
 45780

cacgagaacg cgcacggcga ccgcgccctt ctcgggaacg acgagcggcg cctccagggt  
 45840  
 cagctcgtcg accgtgccgc agtccacctg tgcggcggct tgcagcgcga gttcagaccg  
 45900  
 cccggtgccca ggcagcagca gggtagccag tacgtcgtgg tcggcgagcc aggggtgggt  
 45960  
 gtcgagggag agacgtccgg tgaacaccac accgccggtg tcgggcagcc cgatgcccga  
 46020  
 ggtgagcagg gggtagatcga ccgcgtcggc accggctgcc gccgtcact gctcagatcag  
 46080  
 ccagtaacgc ttgcgctgga agggatacgt gggcagatcg acccgccgcg caccacgccc  
 46140  
 gtcgaacacc gcgtcccagt ccaccgccac accggccgcg aacaaccggc ccacaccggc  
 46200  
 gaacacactc tccacctcgg ggcggtcacg ccgcagcgtg ggcgccaggg tgccgtcggc  
 46260  
 actctgtccc gccatcgccg tcagcacacc atccggaccg atctccagga accgcgtcac  
 46320  
 accctcgtcc tgcaaatagc gcacatcgtc cgcgaaacgc accgcgtccc gcacatgccc  
 46380  
 caccagtac tcggccgaac cgacgtcctt ggtagccgg atggtcggct cccggtagg  
 46440  
 ggcgtctctc gcgaccttgc ggaagtcgtc cagcatcggg tccatcagcg gcgagtggaa  
 46500  
 cgcgtgcgac accttcagcc gagtcgtctt gcggtcggtg aaccgctccg cgaccgccgt  
 46560  
 caccgccttc tcctcgcccg aaatcaccac cgaagaaagg ctgttgaccg cggcaacact  
 46620  
 caccaggccg ctgagatgcg gaaccacctc ctctccgtg gcctggatcg cgaccatcgc  
 46680  
 cccacccgcc ggcaacgcct gcatcaaccg cccacgcgcc gagatcaacc ggccccatc  
 46740  
 ctccagaccg aacacccccg ccacatgcgc cggcccaac tcaccgatcg aatggcccac  
 46800  
 caggaagtcc ggcttgacc cccacgactc caccaaccgg aacaacgcca cctcaagagc  
 46860  
 gaagatcgcg ggctgggtga actccgtacg gccagagcc tcctcatcac cccacatcac  
 46920  
 ctcacgcaca gcgggatcga gcacgccaca cacctcgtcg aacgccgatg cgaaggcggg  
 46980  
 gaacgtctcg tacaactccc gccccatacc aaggcgttga ctcccctgcc cggatgaacag  
 47040  
 gaacgccacc ttgcccacac cggctctcgc cgtgaccgtc tccgaccgga tgagcaccgc  
 47100

gcggtgttcc agcgcggccc ggcctgtcgc ggccgagtac gcgacgtcca gcgggtcggc  
 47160  
 gttcgcggcc agttcgccga agcggccgat ctgcgctcc agcggcccg gggctttccc  
 47220  
 cgacaccacc accggcgcca ccggcaactc ccgccgctcc accaccactt cagcaaccgg  
 47280  
 cacgacttcc tcgacgatca cgtgggcggt cgtgccactg attccgaagg aggacacggc  
 47340  
 tgcccggcgc ggacggccct cgctcggcca ctacgcgcc tccgtcagca gccgcacctc  
 47400  
 acccgcgtcc cagtccacct gcttcgtcgg ctcatccaca tgcagcgtct tgggcagcct  
 47460  
 gccgtggcgc atcgctcga ccatcttgat gatccccgcc acaccgccc ccgcctgcgt  
 47520  
 atgaccgatg ttcgacttga tcgaaccag ccacaacggc cgcccttcgg ggcgatcctg  
 47580  
 accgtacgtc tccaacaggg cctgggcctc gatcgggtcg cccagcgtcg tgcccgtgcc  
 47640  
 atgcgcctcg accgcgtcca catccgccgt cgacagacc gccttcgcca ggcctgctt  
 47700  
 gatcacgcga cgctgggagg ggcggttcgg cgcggtgatg ccgttgctgg cgccgtcctg  
 47760  
 gttgagcgcg ctcccgcgca ccaccgcgag caccggatgc cccagccgac ggcctccga  
 47820  
 cagacgctcc accaccagga cgcccgcgcc ctcgccccac ccggtgccgt cgggtggagga  
 47880  
 cgagaaggac ttgcagcggc cgtccgcggc caggccgcgc tgctcgtga agtcgatgaa  
 47940  
 cgaccgcggt gttcccatga cggtcacgcc gccgacgagg gcgagcgagc actcgcggga  
 48000  
 acgcagtgcc tgcgccccc agtgcagggc caccaggac gacgagcatg ccgtgtccac  
 48060  
 gctcaccgcg gggccttcga gaccgagggt gtaggccacc cggcccgaca cgaggctgcc  
 48120  
 gccgccgggt ccgccggggt agtcgtggta catcaccctg gcgaacacac cggtcgggct  
 48180  
 gcccttcagc gtggtgggtg cgatcccggc acgctccagc gcctcccacg aggtctccag  
 48240  
 cagcaaccgc tgctgcgggt ccatgtccag ggcctcgcgc gggctgatgc cgaagaaatc  
 48300  
 ggcgtcgaac tgtgtggcgt cgtgcaggaa tccgccgtcg cgcacatagg tcttccccgg  
 48360  
 aattccgggc tcggggtcgt agatgtcctc cagccccag ccacggtcgg ccgggaactc  
 48420

cgatatcgcg tcgacaccct cgtcgcgag cgcaccacaac ccctcgggcg agtccacgcc  
 48480  
 acccgggtag cggcacgcca tcgagacaat ggcgatcggc tcgtcgtccg ccgggcgcac  
 48540  
 gacggacgcg accggagccg actccacggc tccggagagc tcccgcagca ggtggcccgc  
 48600  
 caggacgacg ggggtggggt agtcgaagac gagggtggcg ggcaggcgca gtccggtggc  
 48660  
 ggcacccagc aggttgcgca gttcgatcgc cgtgagggag tcgaatccga ggtcgcggaa  
 48720  
 ggcccgtcc gggtcgaccg cctcggctcc ggcattggcg aggaccatcg ccgcctggga  
 48780  
 gcgtacgaga gcgaggagt cgtcggcgcg ttccgcgtcg ccgagtcccg ccaggcgctg  
 48840  
 ccgcaggacg ccggtcgchg acgcgtcgtt gtcgaggaca cggcgcgacc ggccgcggac  
 48900  
 gaggccgcgc atgatgggcg gcggcgagtc gaaggcagcg aggtcgaacc ggacgggcac  
 48960  
 cagtacgggc tcccaccagc cggcggcggc gtcgagcagg gcgaggccct cctccggtgc  
 49020  
 cagggagcgc atgccgaccg aggcgagacc gtccgccatg ccctctccgg cccacggacc  
 49080  
 ccaggccagt gacagcggcg gcagtcggcg ggcgtgccga tgccgcgca ggccgtcgag  
 49140  
 cagggtgctg gcggcggcgt agttgccctg tccgggggca ccgagtacgc ccgcgacgga  
 49200  
 cgagaacagc acgaaagcgg tcagcccat gtcgcgggtg agttcgtgca gatgccaggc  
 49260  
 cgcgtccgcc ttcggacgga gcacgtggtc gacgcgctcg ggcgtcatcg agaggatcac  
 49320  
 gccgtcatcc aggacggccg cggcgtgtac gaccccgccg atggtccggc cgtcgcgagc  
 49380  
 cgtggccagg gcgtcgcggt cggcggcgtc gcacgcggcc acctcgacct cggcggcggg  
 49440  
 cccggtcagt tcctcggcga gtcggcggc tcccggcgcc tgcgggcccc gtcggtggt  
 49500  
 cagcagcagc cgccgtacct cgtggcgggt gacgaggtgc cgggcgacgg cggccccgag  
 49560  
 ggcaccggtc ccgcccgtga tcaggacggt gcgttcggtg tcccagggtg aggcggtcga  
 49620  
 atcgcgcggt acgcccacca gccggggcac ccgtgtctcg ccgccgtca cgcgcagttc  
 49680  
 gggttcggcg atcgcgcgca cgcgggcggg gtcgaccggg gcgtcgggtg cgacgaggaa  
 49740

gaagcggccg g gatgctcgc cctcggcggc gcgtaccagc ccccaggcgg cggcgtggcc  
 49800  
 gaggtcggag ccgtcgggtg cgccgtcggg gacgacgacg agggcgggtg cggcgtcgac  
 49860  
 tgcgccctgg acggcggcca gtacgccggg ggtcaccgcg cgtacggcgg ccggtgtgcc  
 49920  
 gccggtcgtc ggcgggcagt ggtgaacgct cacctcggtg ttctcagccg gtgaccggac  
 49980  
 ctgcgccggc gggaccagc ccaccggaa gagggcggtc gcgaccggg cgcggccgg  
 50040  
 cgcgaggccg tccgccgtga cggggcgcag cgtcagcgat ccgaccgagg cgaccggccg  
 50100  
 tcccatggcg tcggccagct ccagggacac ggacttgtct ccccgacac gcaggctcac  
 50160  
 ccgcgcggtc gtggcgccct ccgcgtgcag ggtgacgtcc gaccagggtg acggaagtac  
 50220  
 cttcgcgtcc tcgccggcca ggtcgcgac gtgcaacgcc gcgtcgaaca gggcgggatg  
 50280  
 gaggccgaag ccgccggcgt ccgcaccctc gggcagagct gtctcggcga acagctcgtc  
 50340  
 gccgcgccgc catcccgcgc gcagtccttg gaacgtcggc ccgtactcca gccgttcgta  
 50400  
 gagtccttcg accgcgaggg gctcggcgtc gcgcggaggc cacaccgtga ggtcgggtggc  
 50460  
 cggctcaccg gcgtcggggc ccaggaagcc ctcggcgtgc aggatccatt cgcggtcgag  
 50520  
 cggggcgtcg tcggctcggg agaacaccct cacagggcgg agtcccgcgc cgtcggggcgc  
 50580  
 gtcgacggcg acctggatgt ggacgccgcc gtgctcgggc aggaacagcg gcgccgcgac  
 50640  
 gttgagttcc tcgacgcggg tccatcccac ctggtcccc gcgcgcaccg cgagttccac  
 50700  
 gtacgcggtg ccgggcagca ggacggagcc catcacggtg tggtcggaca gccaggggtg  
 50760  
 ggccccggtg gacagccgtc cgggtgaacac cacaccgtcg gagccgggca ggtgcaccat  
 50820  
 cgcaccgagc agcgggtggt cgggccggtc gagaccggcg gaggtgacgt cggcggccag  
 50880  
 gccgcgcttg tcgaaccagt aacgggcgcg ctggaagggg taggtgggca ggtcgcgcg  
 50940  
 ccgggcggcg cggccgtcga agacaccgc ccagtcgact ttcaggcccc cgggtgtcg  
 51000  
 gccgccgagg gcggtgagga cggcgacggg tccggtttcg ttacggcgca gggccgggac  
 51060

gagtacggcg gtgtcggcgg tggtcgtgag gcaactggcgg gccatcgcgg tgaggatgcc  
51120

gtcggggccc agttcgacgt accgggtgac gccttcggat tccagccggg tcacggcgtc  
51180

ggcgaagcgg acggtgtcac ggacatgccc caccagtac gcggcgggtg tgacatcgcc  
51240

gttggccacg acgggaaggc cgggctccgc gtaggcgacg cgctccgcga cacggcggaa  
51300

gtcgtcagac atcgggtcca tcagcgtcga gtggaacgcg tgcgacacct tcagccggtt  
51360

ggtcctgcgc gcgccgagct gttcgacgac cgcgtccacg gcctcctcgg tgcccgagag  
51420

caccacggag gcggggccgt tgacggcggc gattcccacg ccgtcccga gacgcgggac  
51480

gacctcctcc tcggcggcct cgaccgccac catcgccccg cccgggggca gcgcctgcat  
51540

cagccgtgcc cgcgcggtga tcagcgaggc ggcgtcgggc agggagaaca cgccggccac  
51600

atgggcggcg gcgagttcgc cgatggagtg tcccgcgacg aagtcgggccc gtacacccca  
51660

cgattcgagc agccggaaca gggcgatctg gaggtcgaag atcgcgggct gggatgaactc  
51720

ggtacgccgc agggcctctt cgtcgcccca catcacctcg cgcacggcgg ggtcgcgcgc  
51780

ggagcacacc tcgtcgaagg cgcgtgcgaa cacggggaac gccgcgtgca ggtcgcggcc  
51840

catgccgagg cgctggctgc cctggccggt gaagaggaag gcggtcagcc cggtcgaccg  
51900

cgcggcggcg ccgaggggag cgccgtccgc caacgcggtc agcgcgccga ggatctcgtc  
51960

gcggtcgtgg ccgaccacga cggcgcggcg ggtgagcgcg gcgcgggcgg tggcgagggg  
52020

gtacgccatg tccacagggt cgaggtcagg cgtggtccgc aagtggtcgg ccaactgccg  
52080

ggcctgggcg cgcacgcggg tctcgggtgcc ggccgacagc ggcagcggca gcggcacggg  
52140

caggggtacg tcggtggacg ggcgcgggtc ggcgggctcc ggctcggagt cggacgcggg  
52200

gtcgggctgc gcctgctcga tgatcacatg ggcgttggtg ccgctgacgc cgaacgagga  
52260

caccgcggcg cggcgcggcc ggtcggcgtc gggccaggca cgcgcctcgg tgagcaggcg  
52320

gacgtttccg gcctcccagt ccacctgggg cgacgcctcg tcgacgtgca gtgtcttcgg  
52380

cagcgtgccg cgtcggatgg cctcgaccat cttgatgacg cccgcgacac cggcggcggc  
 52440  
 ctgcgtgtga ccgatgttgg acttgatcga acccagccac agcggttctt cgcggccctg  
 52500  
 tccgtacgtg gcgagcagtg cctgcgcctc gatggggctg ccgagcgagg tgccgggtgcc  
 52560  
 gtggccctcc atgacgtcga catcggcggg cgtcagtcgg gccgcgggta gcgcctgctg  
 52620  
 gatgacacgg cgctgcgagg ggccgttggg ggcgctgaag ccgttggacg cgccgtcctg  
 52680  
 gttgaccgcc gagccgcgca cgacggccag tacgggggtg ccgttgcggc gggcgtcgga  
 52740  
 cagcttctcc agcagcagga cgccggcgcc ctcggcccag cccgtcccgt cggcggaggc  
 52800  
 ggcgaacgag cggcagcggc cgtcggccga gagtccgcgc tgctcgtga actcgatgaa  
 52860  
 cgtctcgggc gtggccatga cgctgacgcc gccggccagc gccagggtgc actctcccga  
 52920  
 gcgcagggcc tgcgacgccc actggagtgc gaccagcgc gacgagcagg ccgtgtcgac  
 52980  
 ggtgaccgcc gggccttcca gcccgagggt gtacgcgacg cgtccggaga cgaggctgcc  
 53040  
 gtcgctgctg gtgatgccgt agtcgtggta catcgcgccc gcgaagacac cgggtgcggct  
 53100  
 gccgcgcagc gacgccgggt cgagtcgggc gcgctccatc gactcccagg tgatctccag  
 53160  
 caggagccgc tgctgggggt ccatggtgag ggccctcgcgc gggctgatcc cgaagaactc  
 53220  
 ggggtcgaac tgcgcggcgt cgtgcaggaa tccgccctcg cgcgcgtacg tcttgccggg  
 53280  
 ctttccgggc tcggggctgt acagcgcgt catgtcccag ccgcggctgt ccgggaacgg  
 53340  
 gctgatcgcg tcacggcctt cctcgaccag cgcccacagc tcctcggggg atccgacgcc  
 53400  
 accggggtac cggcaggcca tgccgatgat cgcgatgggt tcggtgggta cgctcgcggc  
 53460  
 ggcgcgcagg ctgtcgttgt cctgccgcag cctctcgttc tcgaccagcg agccacgcag  
 53520  
 tgcctcgacg atctcctcga cttcggcgtt catcgtcgtc tcagcctccg ctcacggctg  
 53580  
 tggctccgct gtcgctcacg cgtaccgagt caaggaatcc gccgagcacc tcgtagaaca  
 53640  
 gctccggctc ctccagggtg cgggttgggc tggagttctc aaggatttcc cagcggggcg  
 53700

ccggaatgag ttcttggtag gggcgcaccg tgaccggggt ggcctcgtcg tggcggccccg  
 53760  
 acatgatgag ggtgggcgcg ctgatgtcgg gcaggcagtc gatcaccgac cagtcgcgga  
 53820  
 tgctgccgat gacatggaac tcgttgggac cgttcatcgt gcggtagacc gtcgggtcgg  
 53880  
 tgacggcttc caggtaggag gccatgagtt cgctgggcca cggctcgacg cggcagacgt  
 53940  
 ggccgctgta gaagaccagc atcgcctcca ggtactcgtc gctgtcggtg gtgccggcgg  
 54000  
 cctcgtgccg ccgcagtgtc tcgtcgacgc cgggcggcag ttgggcgcgc aggacgtcca  
 54060  
 tctccgacag ccacagaggg taggaggccg gtgcgttggc gatgaccagg ccgcgcagcc  
 54120  
 cggcgggttc ggccgaggcg tgccaggcgg cgagcagtc gccccacgac tgtccgaaca  
 54180  
 ggacgtagtc gtcggcgatg tcgagccggc gcagcaggtt ctccagctcg tcgcggaaga  
 54240  
 gctgggggggt ccagaagccg gggtcggcgt cgggaaggtg ggtggagccg ccgttcccga  
 54300  
 tctggtcgta gtgcaccacc gaccagccct gttcggcgtg gacggacagc cctgtcaggt  
 54360  
 agtcgtgggt ggagccgggg cctccgtgca cgacgacgag ggccgggcgg ccctcagcgg  
 54420  
 gctgcccggg gacgcggtac caggctcttg actccccgaa gggaacagtt cctttggccg  
 54480  
 tggccgtggg cgtggacgcc atttctcaaa ccacctaagt tcgggtcgtt ctcagcagcg  
 54540  
 gttgccgct cccccgcac cggcgtgcac ttcttccgag gggaaacctg gcgtcggcca  
 54600  
 cggctcagtc gagcgtctcc agccattcct cgatgactcg ggcggtcgcc ggggcgtggt  
 54660  
 cctgcccgag cgagaaatgg ttgccttcga cggcgcgag ggtgtgctcg gagtcccacg  
 54720  
 ggccggcccc catctccgcc acgtcgacct cctcgggggg ctgaacgaac ggctcggacg  
 54780  
 cctggacgaa cagggtcggg gtgtccagcc gtacggggtc gaagcccgcc agtacctgga  
 54840  
 agtagtgcgg catcgcggac agtcgcgccg cgtcgtagtt gccgagggtc gtctccaccg  
 54900  
 tcagcagttc gcccatgagg tggtcgaacc cgacgttcat cggcgtgtcc tcgaccctga  
 54960  
 aggtgtcgat caggacgagt ccggcggggc ggaccttgag cgtctccttc aggtgacggg  
 55020



cgatgatgtg gccgatgatg ccgccggagg agtagccgag cagtacgaac ggctcaccgt  
 55080  
 ccgccgccgc cagcacggcg tcgcccagca cctgtgtcag cacctcgacg gagtcgggca  
 55140  
 gtggctcgtc ccggtggaat ccgggcagtg ccaccgccga cacgtgccgc acgtcccgga  
 55200  
 attcggaaac gagccgggcg tgctggtgca cgccgccgcc cgccatgggt gtggccaggc  
 55260  
 agatcagccg ggggcggccg gggccgtccg ccaaccgcac cgtcttcggg gtcctcgca  
 55320  
 ggtcggcggg tgtgacgaac cggggtcgca gcgccgcgac cgccgacatc aggccagtg  
 55380  
 cccctgtcgt gtcaccggcc cggaccgcc gccggaacat ctcggtcacc gtctcgtcgt  
 55440  
 cctcttccga gggctcggcg gcggactcgg cgccggcggc ggccgtcccg gactccatct  
 55500  
 cgtcggcgag gagacgggcc aggttgggcg gggctctgct gtcgaacacc gccatcgtgg  
 55560  
 gcagcttcgc cccggtggcg gcgatcagcg cggtgccgag ttccatcgcg gtgagcgagt  
 55620  
 cgaaccggga ctccaggaag tcgcggtggt ggtcgacggc ttcggcgtcg gcgtgtccga  
 55680  
 gcacggaggc ggcgaggctc agcaccagat cgctgagcgc ccgctccccg tgctgcccgg  
 55740  
 gcacgcggc cagctctcgc cgcagcggcg acgggtcggg ggtcgcggag cggcggccta  
 55800  
 cggcggggac caggccgcgc aggacgacgg ggagttcgc tccggccccg ttccgacgcg  
 55860  
 cccgcaggtc caggctcacc ggcacgagcg ccggctcggg cctcacggac gcggcgtcga  
 55920  
 acagcgcgag gccgtcctgg gacgacagcg cgggcatgcc ctggcggcgc agccgccgca  
 55980  
 ggtctgcctc gctcaggtgg cccgccattc cgccggtgct cgcccacagc ccccaggcga  
 56040  
 gggactgcgc gggcagtccc tcggcgtggc gccgcgcggc gagggcgtcg aggaaggtgt  
 56100  
 tggcggcggc gtagttgcc tggccgggcg agccgagcac gcccgccgcg gacgagaaca  
 56160  
 ggacgaaggc gccagctcc tggtcgcggg tgaggctcgt caggttgagt gcgccgtcca  
 56220  
 ccttggggcg caggacctgg tcgaggccgt gcgccgtcag gttggcgatc atgctgtcgg  
 56280  
 ccagtacccc ggcggtgtgg acgacggagc cgagggagcg gccggccagc agcgcctcga  
 56340

cggcgtcccg gtcggccacg tcgcaggcgg cgatctcgac cgtcgcaccg agggcgggtca  
 56400  
 gctcctggtg gaggtcggcc gcgccctgcg cgtcgatgcc gcgacggctg gtcagcagca  
 56460  
 ggtcccggac gccgcgttcg gcgaccaagt ggccggcgac gagcgcgccc agaccgcctg  
 56520  
 tgccgcccgt gatcagtacg gctccggggc ggtcccaggg cgagcgcgag ggtgctcct  
 56580  
 cctccgggtg gcccgggacg cgcgcgagcc gggggaccag gatctcctgg ccccgtagc  
 56640  
 gcaggtcggg ctcgcccacg gcgacggccc ggccgatcgt ctccgggtcg tcgccgtcgg  
 56700  
 tgtcgagcag cgcgaagcgc tccgggtcct ccgaacgggc ggccgctacc aggccccagg  
 56760  
 cggagggctg tgccaggctg tcaccgctgg tgaccaccag gagcctgctg ccggcgaacc  
 56820  
 gttcgtccgt gagccaggtc tggagggctc ggagcgtctg ggccgtgacc gtcctgacgt  
 56880  
 cgtcggcgat ccggccgggtg cctgcccggg gtgtccaggc gaccgtcgag ggcaccgggc  
 56940  
 cggagagttc ggccagttcg gtgaccgtcg tccagtccgc ctccggcggc gtggcgggtg  
 57000  
 cggcgggtgt ccacgcgagg tggaaagagg agtcgcccgg cccggcgggc gccgggtgca  
 57060  
 gctgctcggc ggagacctc cgggagatca gcgactccac gtacgcgacg ggacggccct  
 57120  
 gggggtcggc gacgcggatc gtcgtgccgc cctcggcgtt gggggtgaac cggacgcgcg  
 57180  
 cggccctggc gccgaaggcg tgcagccgca caccctcca ggcgaaaggc agtgagggtg  
 57240  
 ctctgtcctc gcccgcgccg agtatcaccg cgtgcagagc ggagtcgagc agcgcgggt  
 57300  
 gcagtacgaa cctctcggcg tcggccacgt cgtcggcgag cgccacctg gcgaatgtct  
 57360  
 cgtcaccgac gcgccacgcc gccttgagcg cctggaacgc ggggccgtag acgtatccga  
 57420  
 ggtccgcgaa ctttcgtac gcgccctcgg tggatcctc cgtgcctcgc tcgggcggcc  
 57480  
 agcgcgacag gtcgaaggag gtcgcctcgt tggactcctc ggccaccagg gcgccttcgg  
 57540  
 cgtgcaggac ccatggcgcg tcgtcgtcgg cgtcctcggc gagcagtggt atgctcaggg  
 57600  
 ggcgccggcc ggtgtcctcg accggttcac cgaccgtcag ccgcagttgt acgcccccgc  
 57660

cctcgggcag gaccagcggc gcgcgcagcg tcagttcttc gaggacgccg tggccgacct  
 57720  
 ggtcgcaccac gcgtaccgcc agttcgacga acgccgtgcc gggcagcagc gtcgcgcccc  
 57780  
 gtacctcgtg gtcggcgagc caggggtggg tgtcggtcga cagacggccc gtgcagatca  
 57840  
 ccgtccgtga gccgggaacg gcgatctcgg cgctcagcag ggggtggctg agcgcgctga  
 57900  
 ggcccatgga cgccgcgctg ccgcctcctg cgtccaccgg ctctggagc cagtaacggg  
 57960  
 tgcgctggaa ggagtaggtg ggcagatcga cccggtgtgc ccgccggccc gcgaagaagg  
 58020  
 cgtccagtc gggagagacg cccgtggtgt gcagatgggc gacggcggtc agcagtgctg  
 58080  
 tggcctcggg ccggtcgcgg cgcagggcgg cggcgggtgt ggcctccggc gcggtctggc  
 58140  
 gggccatggc ggcgagggcc ccgtcgggtc cgagctccag gaaccgggtg acgccctcgg  
 58200  
 cctccaggcg tcgtacgctg tcggcgaacc gcaccgcgtc gcgcacatgc cgtacccagt  
 58260  
 agtcggcgga cgccatgtcc ttcaccagcc ggatgcgcgg ccgttcgtag gtgagggact  
 58320  
 cggcgacctt gcggaagtcc tccagcatcg gttccatcag cggcgagtgg aacgcgtgag  
 58380  
 agacggtcag ccgtcgcgtc ctgcggtcgg cgaagtgctc ggtgaccgcc tccacagcgt  
 58440  
 cctcgtgcc cgaaacgacc acggaggacg ggctgttgag ggcggcgatc cccacctcct  
 58500  
 ccgtgagcag cggcgcgact tcctcctcgg tggcctcgat ggccgtcatg gccccgcccc  
 58560  
 ccgggagcgc ctgcatcagc cgcccgcggt cggcgaccag ccgcgcggcg tcctcaaggc  
 58620  
 cgagaacgcc cgcgacatgg gccgccgcca gctcgcgat ggagtgcccc gcgaggtagt  
 58680  
 cgggcttgat tccccaggac tccaccagcc ggaacagggc gacttccaga gcgaagatcg  
 58740  
 cgggctgggc gtactcgggt cggtgcaacg ccgactcgtc gccccacacc acgtccttga  
 58800  
 gcgacaggcc cgtggcctcg cacacctcgt cgagcgcggc ggtgaagacc ggggaaggct  
 58860  
 cgtacaactc ccgtcccatg ccgaggcgtt ggctgccctg gccggtgaag aggaacgcca  
 58920  
 ccttgccctc gcgccgcttg ccggtgacga cggacggcga gggggttccc gcggccagcg  
 58980

cggtgagccc cgcgagaagg ccctgacggt cgtcggcgac gatcgccgca cggtgttcga  
 59040  
 gagccgcgcg gcccctcgcc agggacaggc ccacgtccgc aggcgtcagg tcgggcccgt  
 59100  
 cgcgcagatg ggagtgaagg ctttcggcct gcgcggacag cgcctgctgg gtcctgcccg  
 59160  
 acagggtcca cagcaccggg cccccggtgg tcgccgcgac cgggggcgcg tgctcctcgg  
 59220  
 cgggcggtgc ctctcgatg atgacgtggg cgttggtccc gctgatgccg aaggacgaca  
 59280  
 cccccgcgcg gcgcgggtgc tcttggtcgg gccaccgccg cgcctcgggtg agcagccgga  
 59340  
 cgtcgccggc ctcccagtc acctgggggtg tcggggcatc gacgtgcagc gtgggcccga  
 59400  
 tgacaccgtg ccgatggcc tcgaccacct tgatgatgcc cgccacgccc gccgcggcct  
 59460  
 ggggtgacc gatgttcgac ttgatcgaac ccagccacag cggtcggtcc ccggggcggg  
 59520  
 cctgcccgta ggcggcgagc agggcctgcg cctcgatcgg gtcgccgagc gtcgtgcccg  
 59580  
 tgccgtggcc ctcgatcagg tcgacccccgt cggcggacac ccgggcgttg gccagcgcct  
 59640  
 gcctgatcac ccgctgctgg gccgggcccgt tgggggctgt gatgccgttg ctggcgcctg  
 59700  
 cctggttgat cgccgtaccg cgcacgatcc ccagcaccgg gtggtcgttg cggcgggccc  
 59760  
 ccgacagccc ctccaccagg atcatgccga cgccctcacc ccagccggtg ccgtcggccc  
 59820  
 ccgcggcgta cgacttgag cggccgctcg tcgccagccc gcgctggtgg ctgaactcga  
 59880  
 tgaaggcttc ggggtgctgc atcacggtga caccgccggc gagggcgagc gagcactccc  
 59940  
 ccgaccgcag ggcctggacc gcccagtgca gggcgaccag cgaggacgag caggcgggtg  
 60000  
 cgatcgtcac cgcggggccc tcaagacca ggggtgtaggc gaccggcccg gaggccatgg  
 60060  
 cgcccgtgct gctgtgtgac gtgtagtcgt ggtacatcat cccggcgaac acaccggtcg  
 60120  
 ggctgccctt gagagtcgtc gggtcgatgc ccgcccgctc gagcacttcc cacgacgcct  
 60180  
 ccagcagcaa tcgctgctga gggccatca ccagcgcctc gttcggcgcg atccccgaaga  
 60240  
 aaccgggatc gaactcggcc gcgtcgtaca ggaacccgcc ttcgcgcgag tacgtcttgc  
 60300

cgggctttcc cggctcgggg tcgtagatgc cctcctcgtc ccagccgagg tcggcgggga  
 60360  
 accgcgagac ggcgtccgtc ccgtcggcga ccagccgcca cagctcctcg ggagaagtcg  
 60420  
 cggcgggta gcggcagctc atcgccacga tcgcatcggg ttcgaggag gcggcctgca  
 60480  
 gggcgcgggt gcgcgtacgc aggcctcctg actccttgag cgacgcgcgc agcggcccca  
 60540  
 cgagtttctg gtcgggtgctg gccatcgtgt cctcagactt cccgcgtcgc ttcgtccgga  
 60600  
 tcggtgtcct gcagggccat gctgatcagg gcttcggcgt ccatcgcgtc gatcgagtc  
 60660  
 tcctcgggca cggctgccgc catgccggag gcggcggcgg atccggcgag ttcgagcagg  
 60720  
 gcctccatca gccccgcgtc gtgcagccgg gtgagcggaa tcgacccgag caggcggcgg  
 60780  
 accgtctcct cgtcggcggc ggcgccggcg tcaccgctcg gtgccagctc cgcgccgatg  
 60840  
 tgctcggcga gcacccgggc ggtcggatgg tcgaagatca tgggtggcga cagtcgcagc  
 60900  
 ccggccgcgg cgttgagggt gttgcggaac tcgacagcgc ccagcgagtc gaagcccaga  
 60960  
 tcgccgaagg cccgctccgg ttcgacggcc tcgggacccg cgtgaccag taccgccgcc  
 61020  
 gcgtgggtac ggacgaggtc gaggacttcg tcgtaccgat cgtcggcggg cagtgccgcc  
 61080  
 aggcgcttgc gcagcgcggc cccgccccc gtggattgcg ccgacacggc acgccgcgag  
 61140  
 gagccgcgta ccagtcgcg cagcatcaaa ggcacgctcg ccgcgttcag cgtcctggtg  
 61200  
 tccaggttca tgggcaccag cggcgggtg gccagcgcgc ccgccacgctc gaggagttcg  
 61260  
 agaccctgct cgggcgagag tcccacgagg cccgcgcggg cgatccgctg ccggtcgggtg  
 61320  
 tccgccaggt caccggccat gccggcgtcg gtcgtccaca ggccccaggc cagggactgg  
 61380  
 gcgggcaggc cgtcggcgcg ccgggtgcgcg gccagtgctt ccagaagcgc gttggccgcg  
 61440  
 gcgtagttgc cctgccccgg tgtgccgatc acgccggcca ccgaggagaa gactacgaac  
 61500  
 gcggtcaggt ccatgtcgcg tgtcagctcg tgcagatgca gggcggccac cgccttggtg  
 61560  
 gtgacgacct tgtccaggcg ttcgggggtc agcgcgcga tcacgccgctc gtccagaacg  
 61620

cccgccgct gcaccacacc ggtgagcgaa cgcccggcca gcagcgcgc gagcgcctca  
61680  
cggtcgccga cgtcgcaggc ggcgacctcg acctcgcccc ccagcccggc cagtcctcc  
61740  
accaactccg ccgcgccggg cgccgccagg ccccggcggc tggtcagcag cagtcgccgt  
61800  
acgccgtgcc cggtgacgag atggcgcgcg accagtcgc ccagggcgcc ggacgcgccc  
61860  
gtgatcagta cctcgtcgc gaagaccgag gacggctcgg actccgcgac cgaggccgcc  
61920  
ctcagcctcg gtacgtagac cttcccgtcc cggacggcca cctggggctc tcccgccgcc  
61980  
agcggcaggg cgatgtccgc cttctcgtcc tcaccggcca cgtcgaccag gacgaaccgg  
62040  
cccggatcct ccgactgggc gctgcgcacc agtccccacg ccgagggcgc gccaggtcg  
62100  
ctacccgct caccggccac ggacaccgcg ccgcgcgtga cgaccgcgag ggtctcggtc  
62160  
tcgccttga tcgccttga cagcggatgc agtgtggaga gcacgtcgtc gccctcggtc  
62220  
cgccacacct tcgcgtcacc ggccgcgtcg tcgcgtacgg cgacgggcgt cactcgacc  
62280  
tggaagagcg agtcgacgcg cgtgagcgca ccggcggcca tcggacgcag ggtcagggcg  
62340  
tccacggaga cgaccggctg cccggcgccg tcgacggcgt ggagggccac ccggtcctgg  
62400  
ccgcggagcg tcatccggac acggagcgcg gtggcggcgg aggcgtgcag ttccactccc  
62460  
gcccaggaga acggcaggac gacgcggctg tcgcccgaca gcaacgggac cgtgtgcagg  
62520  
gccgcgtcca ggagtgcgg gtggatcccg aagcgggtccg ccgccccgga caggacgacg  
62580  
tcggcgtaga ccgtcccctc ttcgcgccag gcggacttga gtccctggaa cgccggccccg  
62640  
tactcgactc cggttccgc caggctctcg tacagccccg ccagttccac gggctctgca  
62700  
ccgggcggtg gccactgggt catgcctcgc gccgcggggc cgccgggtggc cggggccagg  
62760  
gttccggcgg cgtggcgggt ccagggcaga tggggatcgt ccgcgcctcg ggcgtacacc  
62820  
tcgacagccc ggtggcccgc tccgtcgtcc accccgacca cgacctgtac ggccgtggcg  
62880  
gtgtgttcgg caaggaccag cgggtgcctcg atcgtcagtt cctcgatccg gccgcagccg  
62940

acctcgtcgc ccgcgcgggac ggccagttcg acgaagcccc taccggggaa gagcaggggtg  
 63000  
 ccggcgaccg cgtgttcggc gagccagggc tgggtaccga gcgacagacg cccggtcagc  
 63060  
 acggtccggt ccgccccgc gacagcgacc gtctgggtcca gcagcgggtg gtcggacgcg  
 63120  
 tccgcgcccgc gccccgactc gatccagaac ggctgccgct ggaaggggta cgtcggcagc  
 63180  
 tcgacccgcc gcgccccgcg cccgtcgaac acggcgttcc agtcgacccg cacaccggcg  
 63240  
 gcgaacagcc ggccgacgcc ggtgaagagg gtttcacgc ccggacggtc gtttcgctgg  
 63300  
 gtggccgcga ccgtcccgtc ggcggtctgc cgaccatcg cggtgagcac actgtccggt  
 63360  
 ccgacctcca tgaaccgggt gatgccccgg tcctgcaggt ggcgtacgtc gtcggcgaac  
 63420  
 cgcaccgcgt cgcgcacgtg ccgtaccag tactccgccg acgccacgtc cttggtcagc  
 63480  
 tggatgaccg gctcgcggta ggtgacgctc tcggcgacct tccggaactc ttcgagcatc  
 63540  
 ggggtccatca gcggcgagtg gaacgcgtgc gacacctca gccgagtcgt cttgcggctg  
 63600  
 gcgaaccgct cggcgaccgc ggtgacggcc tcctcggcgc ccgagatcac caccgagccg  
 63660  
 ggggtgttga ccgcccgcac gctcacctgc tcggtgaggt gcggcgtgac ctctctctcg  
 63720  
 gtggcccgga tcgccaccat cgcgccgccg ggccggagtt cctggatcag ccgtccgcgc  
 63780  
 gccgtgatca gccgggcggc atcagccaga tcgaacaccc cggccgtatg ggcggcggcg  
 63840  
 agttccccga tggagtggcc ggtcatgagg tccggcttga tccccacga ctccaccagc  
 63900  
 cggaacaggg cgacctggag agcgaagatc gcgggctggg tgaactcggg gcggcccagg  
 63960  
 gcctctctgt cgccccacat cacctcgcgc agcgcgggggt cgagcacggc gcacacctg  
 64020  
 tcgaacgccg tggcgaagc ggggaacgtc tcgtacagct cctggcccat cccggccac  
 64080  
 tgactgccct gtccggtgaa caggaacgcc agccccctt cggtgaccga gtcgatcag  
 64140  
 gtctcggagc cgatccgcac cgcgcgggtg tccagggcgg cccggccggt cgcggccgag  
 64200  
 tacgccacgt ccagctcgtc cgcgtcgacg gaggtgatcc ggtcgagctg ggcctggagc  
 64260

gcggtacggg tccgggccga cagcaccagg gggacgacgg gcaactcccc ccgctccacc  
 64320  
 ggcgcttcct cgaccgggac ggcctcctcg atgatgacgt gggcgttggt gccgctgatg  
 64380  
 ccgaaggagg acacgcccgc gcggcgcggg cggccgctgt tcggccactc cctcgtctcg  
 64440  
 gtgagcagct ggacctggcc ggcctcccag tccacctggg gtgtggggcg gtccacgtgc  
 64500  
 agcgtccgtg gcagcgtgcc ctgccgcatc gcctcgacca tcttgatgat gcccgccaca  
 64560  
 cccgcggcgg cctgggtgtg accgatgttc gacttgatcg accccagcca cagcggtcgg  
 64620  
 tcctcggggc ggccttggcc gtaggtggcc agcagcgcct gcgcctcgat cggatcacc  
 64680  
 aggggtggtg ccgtgccgtg cgcctcgacc gcgtccacat cggcgccccg caggccccgc  
 64740  
 ttggccagcg cctgcttgat caccgcctgc tgggacggcc cgttgggggc ggtcaggccg  
 64800  
 ttgctcgcgc cgtcctggtt ggtcgcctgt ccccgtacga gggccagcac cgggtggccg  
 64860  
 ttgcggcggg cgtccgacag ccgctccacc aggagcatgc cgacaccctc actccagccg  
 64920  
 gcgccgtcgg cggcggccgc gaaggacttg cagcggccgt cggtcgccag accccgctgc  
 64980  
 tcgctgaact cgatgaagt gtccgccgcg gccatcacgg tgacgccgcc ggccaggggc  
 65040  
 agcgagcact cccccgaccg cagggcctgc gccgccaggt gcagggcgac cagcgaggac  
 65100  
 gagcaggcgg tgtcgacggt caccgcgggg ccttcgagcc ccagggtgta ggagacgcgg  
 65160  
 ccggaggcga tggcgcccgt gctgctgttg tgggtgtagt cgtggtacat catccccggc  
 65220  
 aagacaccgg tcaggctgcc cttgagagtc gtcgggtcga tgccccgcgc ctcaagaacc  
 65280  
 tcccacgacg cctccagcag cagccgctgc tgagggtcca tcaccagcgc ctcgttcggc  
 65340  
 gcgatccccg agaagccggg atcgaactgg gccgcgtcgt aaaggaatcc gcccttgtcg  
 65400  
 acgtagctgg tgcgggggcg ggtggccgtg gggtcgtaga tccgctccag gtcccagccg  
 65460  
 cggtcggtgg ggaagtgtga gatggcgtcc gtgccgctgt cgaccagccg ccacaggtcc  
 65520  
 tccggcgagg acacgcctcc cgggtagcgg cacgccatcg cgacgatcgc gatcgggtcg  
 65580



tcgccgaccg gggccgcgac gggggtgaga cggccctcgt gcaccgttcc cgagacctcg  
 65640  
 tccagcagat gacgcgcgag aacgggtgggg ttcgggtagt cgaacaccag cgtggccgga  
 65700  
 agccgcaggc cggtcgcgcc gccgagaccg ttgctcagtt ccatcgccgc cagcgagtcg  
 65760  
 acaccagat cgcggaacgc gcgctccggg tcgacggccc ccggtccggc gtagccgagc  
 65820  
 gtggtggcgg cctgcgcgcg gaccaggttc agcagcatgt cgaagcgtg gtcgtcggac  
 65880  
 atgcccacga gccgctcgcg gagcccgtcc gcgctcggcg ggggtcgggc ggtgccacgg  
 65940  
 gtgacgaccg ggaccaggcc gcgcagcagc tccggcaccg cgccgccggc gcggacggcc  
 66000  
 gggaggtcga gcttgacggg gacgagtacg gcgggtccgt ccgccgccgt cgcagcgtcg  
 66060  
 aagagcicca gcccctcgtt gtgggacagc gacaggatgc cgccgcgttc catacgggac  
 66120  
 cggtcggtgt ccgtcagttc actggccata ccggtgctgg tccgccacag gccccacgcc  
 66180  
 agggactggg cgggcaggcc acccgcacgg cggcgtcgg ccagggcgtc cagataggcg  
 66240  
 ttggccgccg cgtagttgcc ctgccccggc gagccgatca cgccggccgc ggaggagaag  
 66300  
 agtacgaacg cggtcaggtc catgtcgcgt gtcagctcgt gcagatacag ggcggcgtcc  
 66360  
 acctggggc gcatcaccag gtccaccgc tcgggcgtca gggaccgat cacaccgtcg  
 66420  
 tccaggacac ccgctcgcgt caccacaccg gtgagcgaac gcccggccag cagcgcgcg  
 66480  
 agcgcctcac ggtcgcggac gtcgcaggcg gcgacctcga cctcggcccc cagccccgtc  
 66540  
 aactcctcca ccaactccgc cgcgccgggc gtgtccacgc cccggcgacc ggtcagcagg  
 66600  
 agccgtgaca ccccgtactc ggtgacgaga tgccgggcca gcagagcacc caggacaccc  
 66660  
 actccaccgg tgatcagcac ctcgtcgccg aacgccggcg ccaggtcgtc cgtcgtctcc  
 66720  
 ggcaccgcgg acacgcgggc cagccggggc acgtgggcga caccgtcgcg taccaccacc  
 66780  
 cggggctcac cggtcgacag cgcaggcgcg aggtcggcgt tgctcgcgt cggcgcctcg  
 66840  
 tcggtgtcac cgtccaggtc gatcaggaag aaccggccgg ggtcttcggt ctgggcggtg  
 66900

cgcaccagac cccagacggc cgccgcggcc aggtcgtcga cgtccccgcc gttcaccgag  
 66960  
 accgcgcccgc gcgtgacgac caccagacgg gagccggcgg actgcagtgc ctccagcggc  
 67020  
 aggttcaccg ccgcgcgtac gtccagtccg ccggggaggc ggaacacctc ctcgtcggtg  
 67080  
 ggcggtgccc ccccggtgga ggcggtaccg gcggcgaccg gggccagggc gacgtggtac  
 67140  
 agcggctccg tacgggcctt ggtcgccatg tccgtgaggg gccgcaggac caacgagtcg  
 67200  
 accgtggcga cgggccggcc ggtcgcgctc gcgatggtca gggccgccac gccgtcctgt  
 67260  
 acgggcgtga cgcgcacccg cagcgcgccc gcgccggagg cgtgcagttc cactcccgac  
 67320  
 caggcgaacg gcagcatcgc cacatcgccc gttcccggcg gggagagtcc gatggcgtgc  
 67380  
 agggccgctg cgaagagggc cgggtgcaga ccaaaggcgt ccgccaccgc gttgtccggc  
 67440  
 agggcgatct cggcgaacac ctcgtcgccc gcccgccagg cggcccgaag cccccggaag  
 67500  
 gtcggccccg acgccaacc cgtgccgacc aactcctcgt agaggggtgc gacatcgagg  
 67560  
 tccagcggct cggcgcccgg cgggggccac tcggccagtt cccctcccc gccggaggtg  
 67620  
 gcggtggcga gcagaccggt ggcgtgccgg ttccagggca ggtcggtcgc gtcctggtcg  
 67680  
 cgggagtaca cctggacctc gcggcgcccc tcttcgtccg ccgctccgac gacgacctgg  
 67740  
 acggcgaccc cgccgcgttc ggcgaggacc agcggcgcct cgatcgtcag ctctcgaca  
 67800  
 cggccgcagc cgacctcgtc gccggccccg atcaccagct ccacgaatcc ggtgccgggg  
 67860  
 aagaggatcg agccgccgat gacatggtcg gtgagccagg gcagcgtccc ggtcgacagc  
 67920  
 cgtccgggtga gcacgacctc ctccgagccc gcgagcatga ccatcgcccc gagcagcggg  
 67980  
 tggcccagcg agcccaggcc catggaggcc gcgtcggcgt tcgcccgtctc gtcgttcagc  
 68040  
 cagtagcggg tgtgctgga ggcgtacgtg ggcagttccg tctgccccgc cccggtcgcg  
 68100  
 tcgtacacct tctcccagtc gacgtccacg ccgcgggtgt gtgcctgggc cagcacgctc  
 68160  
 aggaaacggt cgaggccgcc gtcggtgcgg cgcagcgtgc cgatcgtggt ctgttccatg  
 68220

ctccgcccga gcaccgggtg cgggctggcc tcgatgaaca cccccacgcc ctgctcggtc  
 68280  
 agccggcgga tgggccggtc gaactccacg gtctggcgca gattccggtg ccagtagccg  
 68340  
 gcgtcgagag cggtggtgtc cagcagcccc cgggtcacgg tggagtagaa ggggatccgc  
 68400  
 gcggcgcggg gcttgatggg cggcagcacg tcgagcagtt cccgctctat gcgctccacc  
 68460  
 tgtgccgagt gcgaggcgta gtccacctgg atccggcgcg cccgtacgcc gtcggcctcg  
 68520  
 caggaggcca tcagtccgtc cagcgcgctg acctcgcccc agaccacggt ggcggcggtg  
 68580  
 ccgttgacca ccgcatgcc cgtccgggtg ccccagcgtt cgatcagccg ctccggtctcc  
 68640  
 tcccggggga gggcgaccga caccatcccc ccctggccgg agagggccag cagcgccttg  
 68700  
 ctgcccaggg cgaccacccg ggcgcccgtc tgcaacgaca gcgctccggc cacgctcggc  
 68760  
 gcggcgatct cggcctggga gtggcccacc acgcccagcg gctcgactcc gtagtgccgc  
 68820  
 cacagaccgg ccagtgacac catgaccgcc cacaggacgg gctgcacgac gtctacagct  
 68880  
 tccagcaggg cctcgtcacc gagcgcctcg ctcaacgacc agtccacgaa cggcgccagt  
 68940  
 gcctcctcgc acgcgacat ctgctccgtg aacacggcgg acgacgccat cagctcggtc  
 69000  
 gccatgcccc cccactggga gccctgtccg gggaacacca tgaccgcgcc ggcgcccggc  
 69060  
 cgggcccgtg tcaccggtgc ttccccctgc accagagcgg cgagcccctg ggagaagccc  
 69120  
 tcttcgctcg cggccagtac ggcggcccgg tagtcgtacg acgcccgccg cgtggccagg  
 69180  
 gaccatcccc cgtccaccgg ccgtaggtcg tgcgtggcga cggcctggag ccgctccgcg  
 69240  
 taggcgcgta cggcggcctc ggtcttcccc gagatcagcc agggcaccac cggaaagtcc  
 69300  
 cgggtctcac gcggctccgg cgactccggc gactcttcgg ccggtacgtc ctccgctcgc  
 69360  
 ggagcctcct ccaggacgat gtgggcggtg gtcccgtgg ccccgaacga ggacacggcg  
 69420  
 gcgcccgcgg ggcccctc cggccaggcg cgggcctcgg tgagcagccg gacgcttccc  
 69480  
 gccgtccagt ccacctgcgg ggacggttcg tccacgtgca gtgtcttggg cagggagccg  
 69540

tggcgtatcg cctggaccat cttgatgatg cccgccacac cggcggcggc ctgtgtgtgc  
 69600  
 ccgatgttcg acttgaccga cccgagccag agcggggcgg cctcggggcg gtccctgccc  
 69660  
 taggtggcca gcagcgcctg cgctcgate gggtcgccc a ggggtggccc ggtaccgtgg  
 69720  
 gcctccacga cgtccacatc ggaccccacc aggcccgcgt tggccagggc ctgctggatc  
 69780  
 acccggcgcct gggacgggcc gttcggggcg gtgatgccgt tgctggcggc gtccctggtg  
 69840  
 acggcgcctgc cgcgcacgac cgccagaatc ggggtggccgt tgcggcgggc gtccgacagc  
 69900  
 cgctccacca ccagcatgcc gacaccctca ccccaccggg tgccgtcggc cgccgagggc  
 69960  
 aacgccttgc agcggccgtc cgtcgcacaga ccccgcctgcc ggctgaactc cacgaaggtg  
 70020  
 accggcgtcg acatgatcgt cacgccggcg gccagggcga gggaacattc cccactgcgc  
 70080  
 agggccttga ccgccatgtc cagtgcgacc agcagggacg agcagggcgg gtccgacggc  
 70140  
 accgccggac cctccagccc cagcgtgtag gccaccggc cggaaccag cgcaccggc  
 70200  
 aactgctgc tcggatagtc gtggtacacg agcccggcga agactccggg ggcgggtccc  
 70260  
 ttcagcgcag gcgggtcgat cccggcccgc tccacggcct cccaggaggt ttccagcagc  
 70320  
 agccgctgct gcgggtccat cgccatcggc tcaagggggc tgatgccgaa gaactcgggg  
 70380  
 tcgaagtcgc ccgcgccgtc caggaagctg cccttggtga cgtagctggt gttctcggc  
 70440  
 gtcccgtcag ggtcgtacag cgagtcacag ttccagcccc ggtcgggtgg gaactcggcc  
 70500  
 accgtttcca gaccgtcggc gacgaggtcc cagagtcctt ccggcggagga cgcgccaccg  
 70560  
 gggaaagcggc atgccattcc cacgatcgcg atcggctccc gctcccgatc ctccgctca  
 70620  
 cgcagacggc gccgagtctg ctgcaactca ccggtgga a gccgcagata ctgcggaagc  
 70680  
 cgttcgtcgt tggagtcctt cgacatcgtc atcaaccaat ccgtgaaagt tcacaaagca  
 70740  
 gaggcgggag ccgtcaggac agtccgagtt ccttggtcag caccgcgaac aactcgtcat  
 70800  
 cggattcggg cacgagggcg ccgacgggct cgtccgctac gccgctccc gcgtctcgcg  
 70860

cccgtcgcga catggtttcg aggcgggccc tgaccttggc gtgggcttcg cggtcgcccc  
70920  
atgccagctc ggtcagcacg gcttccagcc ggttcagttc ggccagcagc gcggccgccc  
70980  
cgtcggcctc ctccgggccc agcccgtcgt gcacgagctc ggcgatggcc aggggggtcg  
71040  
ggtagtcgaa ggccagggtc gggggcagtg acagccctgt ctcggtgttg agggcgttgc  
71100  
gcagttccac cgcgcccagc gaggtgaggc cgagttcgtt gaaagcgcgg tccgggggca  
71160  
cctcgtcggc gccgcccgtg ccgagcacct gggcgacgtg gccccgtacc aactccagga  
71220  
ggaacttctc gcgttccggc acggtcagtc cggacagccg ctcccacgac gcccggcctg  
71280  
ccggcgcgcg acggttcttg ccgcgtacca gtccgccgaa tatcgcgggc agtgtccccg  
71340  
tctcggcgag ggaccgcac accgccaggt cgaaccggac cggcagcgtc atcggcggcc  
71400  
ggccgcgcag ggtcgcggca tccatcagcg ccagcccctc gtccggggac agcggcggca  
71460  
ttcccagcgc cctgagccgc gccagctcgg cgtcgtccag acggtcggcc atgccccga  
71520  
cctcgggtcca gggccccag ccgaggggtca gcgcgggcag ccccttggtc gcgcgatgcc  
71580  
tggcgagcgc gtccagccag gcgttcgcgg cggcgtagtt gccctgtccc gcgcccgcga  
71640  
gcgtgccggc gacggaggag aagatcacga acgacgagag atcgtggtcc agggtgagtt  
71700  
cgtgcagatg ccaggcggcg tcgaccttcg gcctcatgac cgtgtccagc cgctcgggcg  
71760  
tcaacgtccc catgagtccg tcgtccagca caccggcccga gtgcacgagg cccgtcaggg  
71820  
gccgctcggc cagcagggcg gccagcgcgt cccggtcggt gacgtcgcag acgacgacct  
71880  
ccacctcggc gcccgatccg gtcagctcgg ccaccagttc ggccgcgccc ggcgtgtcca  
71940  
tgccccggcg getggtcagc agcaggctct tcgcgccgtg ctccgcggcc aggtgcccgg  
72000  
cggccagggc gccgagcccc ttcaggccac cggtgatcag gatcgtgccc tgcgtgcccc  
72060  
acgactggac gcggccttcg accgcgggga cgcgggcccag cctcgcacc cggatctcgc  
72120  
cgtcccgtac ggcgatctcg ggctcggcga ggcgcagcac ctcgtcgggg agttcggctt  
72180

cggcgtccag gtccaccagg acgaaccggc ccgggtgctc cgactccgcc gaacgcacca  
 72240  
 ggccccagac cgccgcgtgc ccgaggtcgc ttccgtcggg cgcgccgcgc gtgacgacga  
 72300  
 cgagcttggc ggatgccgcc tgctcgtcgt ccagccaccg ctggatgacg gggagcacct  
 72360  
 gatgggtgat cgtccggaag ctttcgggcg cgggctcgtc cagtggcggg cactcgtaga  
 72420  
 cgacatgagc accggtctcg gagccgacgg gtgcgggggc tgcgaccagc tcgacgtgga  
 72480  
 acagcgattc acggccgccc gtgtgtgccc gtacctcgtc cgtggtgatc gagcgggtct  
 72540  
 cgaccgaggc caccgacaga acgggctgtc cgtcctggcc gacgctcacc gaggcggcgc  
 72600  
 gccacacatg ggcgagcgtc ggctcggcgt cggggtggag tccggcggtg tgcgcggcgg  
 72660  
 cttcgagcac cgtggcgtac gtctcgtcgt cacgtgtcca cgaggactcg gtggcctcgg  
 72720  
 aggtcaggac gccggtggcg tgcttgacc caggcccggt ccggtcgcca cggcgcgcgt  
 72780  
 acacgctgaa cgccccggcc tcgtcgacca tcgcgcgcag tgctcgtcgc tcctcgggtc  
 72840  
 cggccatgac cagcggcgcg tggaggtcga gcgcttcgac gcgcccgcga cccgtctgct  
 72900  
 ccgctgccgt cagggcgagt tccaggaaca cttccccggg cacgaccacg gcgcccgaagg  
 72960  
 cgtcgtggtc ggcgagccag ggggtgggtgt cgagggagag acgtccggtg aacaccacac  
 73020  
 cgccggtgtc cgggagcacc atgccggagg tcagcagggg gtggtcgacg gcgtccgcgc  
 73080  
 cggccgtgga cttggactgc tcgatcagcc agtaacgctt gcgctggaag gcgtacgtgg  
 73140  
 gcagatccac ccgccgcgca ccacgcccgt cgaacaccgc gtcccagtcg accggaacac  
 73200  
 cggccgcgaa caaccggccc accccggcga acacactctc cacctcgggg cggtcgcgtc  
 73260  
 gcagcgtggg cgccagggtg ccgtcggcac tctgtcccgc catcgccgtc agcacaccat  
 73320  
 ccggaccgat ctccaggaac cgcgtcacac cctcgtcctg caaatgccgg acatcatcgg  
 73380  
 cgaaccgcac cgcgtcgcgt acgtgcccga cccagtactc cgccgaaccg acgtccttgg  
 73440  
 tcagccggat ggccggttcg ctgtacgtga cgctctcggc gatctcgcgg aagtcgtcca  
 73500

gcatcgggtc catcagcggg gagtggaagg cgtgagagac cgtcagacgg ttgcgcttgc  
 73560  
 ggtcggtgaa ccgctccgcg accgccgtca ccgccttctc ctcgcccga atcaccaccg  
 73620  
 aggacgggct gttgaccgcc gcgacactca cctcgtccgt gagcagcggg agtacttctt  
 73680  
 cttcgggtggc ctggatcgcg accatcgccc caccgcggcg caacgcctgc atcaaccgcc  
 73740  
 cacgcgccga gatcaaccgg cccgcactct ccagaccgaa ccccccgcc acatgcgccg  
 73800  
 ccgccagctc accgatcgaa tgccccacca ggaagtccgg cttgaccccc cacgactcca  
 73860  
 ccaaccggaa caacgccacc tcaagagcga agatcgcagg ctgggtgaac tccgtacggc  
 73920  
 tcagtacttc ttcgctgccc cacatcacct cacgcacagc ggggtcagac accgcacaca  
 73980  
 cctcgtcgaa cgccgacgcg aaggcgggga acgtctcgta caactcccgt cccataccca  
 74040  
 ggcggtgact cccctgcccc gtgaacagga acgccacctt gccctcggcg accgagccgg  
 74100  
 tcacggtctc ggggccacc agcaccgcc ggtgctccag aacagcgcgc cctgtcggcg  
 74160  
 ccgagtacgc cacgtcagac gcgttcccgt tcgaggccag ttcgccgaag cggccgatct  
 74220  
 gcgcctccag cgccgccggg gtcttccccg acaccaccac cggcgcacc ggcaactccc  
 74280  
 gccgtccac caccacttcc tcgacgggga cggcctcttc gacgatgacg tgggcgttgg  
 74340  
 ttccgctgag cccgaacgac gacactcccg cgcggcggcg acggccttcg ctcggccact  
 74400  
 cacgcgcctc cgtcagcagc cgcacctcac ccgcgtccca gtccacctgc ttcgtcggct  
 74460  
 catccacatg cagcgtcttg ggcagcctgc cgtggcgcac cgcctcgacc atcttgatga  
 74520  
 tccccgccac acccgccgcc gcctgcgtat gaccgatgtt cgacttgatc gaaccagcc  
 74580  
 acaacggccg cccctccgga cggcctgac cgtacgtctc cagcagcgc tgcgcctcga  
 74640  
 tcgggtcacc cagcgtcgta cccgtgccgt gtgcctcgac cgcgtccaca tccgccgtcg  
 74700  
 acagaccgc cttcgccagc gcctgcttga tcaccggcg ctgcgccggg ccggtggggg  
 74760  
 cggtcagacc gttgctggcc ccgtcctggt tcagcgcgct cccgcgcacc accgccagca  
 74820

ccggatgcc caggcgacgc gcgtccgaca gccgctccag aaggagtact ccgacacctt  
 74880  
 cggagcagct catgccgttg gcggcgctgc tgaacgactt gcagcggccg tcgatcgaca  
 74940  
 ggccgcgctg ctcgctgaag tagaggaaca tctcgggctg ggacatcacc gtcacaccac  
 75000  
 ccgcgagagc gagcgagcac tcaccggaac gcagcgactg gatcgccgtg tgcagcgcga  
 75060  
 ccagcgagga cgagcacgcg gtgtccatgg tgaccgcggg acccaccagg cccaacgtgt  
 75120  
 aggcgacgcg cccggagacg atgctgccgc cgctgctgcc gccggcgtag tcgtggtaca  
 75180  
 tcaccccgcc gaacacaccg gtcgggctgc cttcagcga ggccgggctg atcccgccgc  
 75240  
 gctccagcac ctcccaggac gcctccagca gcagccgctg ctgcgggctg gtctccaggg  
 75300  
 cctcgcgcgg actgataccg aagaactcgg cgtcgaactc cgctgcgctg tgcaggaatc  
 75360  
 cgccctcgcg ggacgttgtc tttccgggct tgccgacctc ggggtcgtag atgtcctcca  
 75420  
 ctccccagcc gcggtctgcc gggaaactcgg agatcgcgct gacgccctcg tcgacgagcc  
 75480  
 gccacaagtc ctcgggcgag ttcacaccgc ccgggtagcg gcacgccatc gagacaatgg  
 75540  
 cgatcggctc ctggtcacgc tgctcaagct cggcgacgcg tctgcgcgct gtgcgcagat  
 75600  
 ccgtggctgc gcgcttgagg tagtcaagaa gcttctctg gtcgctcacc aaagccaccc  
 75660  
 cgggtcgaaa gggcgctcac acttacgacg cgcgcttgcc gggaacgtta cgcaggggtga  
 75720  
 aaagaccccc caaccctga tcgcccctaa cgtggcccga ccccttcgcc cggcgcgacc  
 75780  
 atcggtttcc gaccccggcg tccggaccac gtgcagccca caccgaccgc gcccgatcgt  
 75840  
 ggcaagggtt cgccctgtca cgcgcggtag tcgggtcttg gggacgtacg gggtcggctc  
 75900  
 gccgcccgcg ggcgacgcac ccgggcaccg gtggttcggc gggacggcgc gggtcagctc  
 75960  
 cgccactctc tgtcctggcc gcccgctccag gcgcccggcg gtgtgtacca ggcaccgggt  
 76020  
 gaggcgtaac ggaccgcgcc ccaggcggag ttgcgccgct gtcgggccgc ggaggacgcg  
 76080  
 acggcctgct cccgggtacc ggtcagggcg ctcaacaccg cgagcagcgc cgccagttcc  
 76140



atctctcccg gagcgcccct gacgacgcgc acgaggtgct cggcaggtgt gacggcgacg  
 76200  
 ggttcccgtg ttcgcgggcc ggccacgtag tcagggtgca tgatttcctt ccggtactct  
 76260  
 tcggtgctct tcgaaagact ccgcccgtgc tttctccgtg ctttctctcc ttcattgaaa  
 76320  
 ggtccctccg ccccgcaact ccagcggatc gaccccctaa ttccccgctt cgtcccccta  
 76380  
 ttcgaaagcc gtggtagtcc gtcctcacia caacgggccg ccattccgga ccgagcgcga  
 76440  
 cccgcgaagc gggtagcgtt cgataggggt catgaggggt cggcacgcgc ccgtacacac  
 76500  
 atgaacgctg aacacagcct gccgcagcgg cttcagccgg atcgatccaa ggaacttgat  
 76560  
 tgatgaccgc catatcgagc gacaacgcgt ccaattggat tcgagaattt catccggcgg  
 76620  
 accggacatc cccgaggatg atctgcttc cccacgcggg cgggtcggcg agctactact  
 76680  
 tccccgtctc ccgggcgctg gccgggaaga tcgaagtcct cgccatccag taccgggggc  
 76740  
 gccaggaccg ttacacggaa ccggccatcg gcaacgtcga ggccctcgcc gccgcggtct  
 76800  
 tccgtgagct tccgacggag gacctggacc ggacctggtt cttcgggcac agcatggggg  
 76860  
 ccgccgtcgc cttcgaggtg gcccggtga tggaacggga gttgaaccag tcgcctgtcg  
 76920  
 ggatcatcct ctccggccgg cgcgcaccgt cccggttcg tcccagacc ctccacctgc  
 76980  
 agggcgacgc ggcgatcatc gccaacgtgc agtcgctcag cggtaaccgac gcgatcctct  
 77040  
 tcgaggacct cgacaccag cggctgatca tgccggcgct gcgagccgac taccgggcca  
 77100  
 tcgagacctc ccggccgccc ggcactccac gcgtcgcgtg cccgatccac accttcgtgg  
 77160  
 gcgacgccga cccggttgcc acgctggacg aggtcggcag ctggcgcgac cacacctcgg  
 77220  
 ccgagtacac cctgcgcgtt ttcccgggtg accacttcta tctgacggcg cgtgccgtcg  
 77280  
 aggtcatctc cgcgatctcc cagctgatcg tggagcccac ccagaccgc ggctgatcgc  
 77340  
 gccgcgggtg ccgccaccgc cgccggaggc ggtgggatcg gcgtcgggcc ggaacgcgac  
 77400  
 cgggcccacg gatcgactgg actctcgcct gtgtgtttgt atcaggcgaa cagcgaaggca  
 77460

gccgggccgg cagagtggga gaaacgccgt gttctactac gtactgaagt acgtgctgtt  
 77520  
 ggggcccgtg ctgCGgttgc tcttccggcc ccggatcgag gggctcgaac acatccccggc  
 77580  
 ggacggcgcc gcgatcgctg cgggcaatca tctctccttc tccgaccact tctgatgcc  
 77640  
 cgcgatcatc cggcggcgga tcacgtttct cgcgaaggcc gagtacttca ccggtccccg  
 77700  
 cgtcaagggg cgcctcaccg cctccttctt ccgCGcgctc ggccagatcc cggtcgaccg  
 77760  
 gtccggcaag gaggccggga aggcccgcat ccgggaaggg ctCGgggtgc tcggcaaggg  
 77820  
 tgagttgctg gggatctacc cggagggcac gcgctcgcac gacggacggc tctacaaggg  
 77880  
 caaggtcggg gtggcgggtga tggccatcag ggcgcaggtc ccggtggtgc cgtgCGcgat  
 77940  
 ggtgggtacg ttcgagatcc agccgccccg tcagaagatc ccgaacatcc ggcggggtcac  
 78000  
 gatccggttc ggtgagccgc tggacttctc gcgctacgCG ggtctggaga accagaaggc  
 78060  
 ggcgggtccgc gcggtcaccg acgagatcat gtacgCGatc ctCGgtctgt ccgggcagga  
 78120  
 gtacgtggac cggtagcCCg ccgaggtgaa ggccgaggag gcgCagcagg cGCCgaagaa  
 78180  
 gttccccgCG ctgCGacgct gagcaccgCG gggccgcccc gcagccggac ggcgaaggag  
 78240  
 gggcggctgc cgtctccggc cgcCGccct cccccggttc accggtgctg cccgtgcctt  
 78300  
 acggcttggg tgtggcgtgc ggtgCGcacg tcacgtcCGg gcggtccagc tcgCCgctga  
 78360  
 gcagatacga ctCGaccCGg tcgttgatgc acgCGttCGc gacattgggtg atgCCgtCGg  
 78420  
 aaccggcgtc ccgctcgggtg atcaggcggtg agcccttgaa gcgcttgtgc agctcGacgg  
 78480  
 cccccccgta cgggggtggcg gcgtcacgCG tggactgCGc gatcagtagc ggcggcaggc  
 78540  
 cgcggccccgt accgacctcg atCGgtgtct gctgctctac gccccaggtc gaacagggca  
 78600  
 ggttcatcca ggcgttggcc cagggtgagga acgggtggtc gcggtggagc cgggtgttgt  
 78660  
 cccgggtcca ggtgCGccag ctCGtgggCC acttggcCGt ggcgCactCG acggcgggtg  
 78720  
 agaccgCGtt gctgttctCG gcgCGgggtgt tgcccaccgt gtcggacagg tccggcCGcg  
 78780

cggcgtcgac gagcgctgg gtgtctccgg ccaggtactt gtcaccaggtg tcggcgaccg  
 78840  
 gcacccagct ggagtcgtag tagggcgcg cctggaacag cccgatgagt tcggccggtc  
 78900  
 ccacgacgcc gccgatcggg ttcttcttgg cggtggcgag gagcttgtcc cactgcttct  
 78960  
 cgaccttggc gacggtgtcg ccgatgtgga acgccgcgct gttctcggcg acccacttct  
 79020  
 tccagtcgtc gaagcgtgtc tcgaaggcga cgtcctggtc caggttggcc tggtagcaga  
 79080  
 tcttctcctt cgacgggttg accacgctgt ccagcaccat gcgccgtaca tgggacggga  
 79140  
 agagcgtgcc gtagacggcg cccaggtagg tgccgtagga gacaccacg tagttgagct  
 79200  
 tcttgtcgcc gagcgcgcc cgcaggacgt ccaggtcgcg cgcgctggtg ggcgtcgta  
 79260  
 tgtgcggcag catccagccg ctgcgctcct tgcagccgct cgcgtactcg gccgcgagct  
 79320  
 tgcgctgggc gcgcttgtcg gcctcgggtg cggggacggg gtcggccttg ggagccttga  
 79380  
 cgaactcctg cgggtcgacg caggagatgg gcgctcagcg cccgacgccg cgcgggtcga  
 79440  
 agccgacgaa gtcgtacgcc ttggcggcgt ccgccagat ggcgttcttg gtcacgacgc  
 79500  
 ggcgcgggaa cgccatgccg gacgcgccgg ggcaccggg gttgtagacg agggagccct  
 79560  
 gacgctcacc cgccgtcccg gtggtgccga tccggctcgc ggctatcttg atctgcttgc  
 79620  
 cgttcggacg ggcgtagtcg agcgggacac tgatgtagcc gactggatc ggcttctcca  
 79680  
 gggcccagtc ggccgggacg tccgccagc ccatgccctt cttcgcggcc cgtcggcgg  
 79740  
 cgatctcggc gccggccgcc tgggcgttca gcgtcttctg ttccttgccc gacgcggccc  
 79800  
 ggccactgtc ggcgcggccc tcggcgctcg ccgacgaggc ggcgagcgtg gtcgctatga  
 79860  
 gcgtgcccgc gagcagagtg ccggccgagc caagcactgc tgtgcgtctc aagtggtagc  
 79920  
 tccccctggc tgacgcgccg cacgctgcgg cgctgttcgt agggttaccg agaggatcct  
 79980  
 tatctctgtg ggccgggttga gaacaggcca tacggcatat tctttgccga accgatagcc  
 80040  
 ggtgagtcgg gttccgctca  
 80060

<210> 2  
 <211> 2856  
 <212> ADN

ES 2 543 203 T3

<213> *Streptomyces* sp. MP28-13

<400> 2

gtggtgaagg aacgggacaa gatactcgaa cagctcgatg ccctactgat ccagtccacg  
60  
cggggcaggg gggccatcgc cgcgatcagc ggttcgaccg cgatcggcaa gaccacgaca  
120  
ctcaacgcc tggccgaacg ggcgacttcg gcggacatca cggtgctcag cgtggtgagt  
180  
tcgccgcacg agcgcgaggt tccctacagc gccctcgccc agctgctgca ttcgatcgaa  
240  
gccagtgca ccaccgccga cgttccgcgc gccgggggcg ccgggaccgt ccatgccggc  
300  
aggcaggcgg acgaggccgc cgccctcgcc ccgccgttgg cccaggacga cccgatgaca  
360  
gtcgcacggc ggacctacca gatcatcgcc gagctgaccg ccctgcggcc tctgctgatc  
420  
acggtggacg acatccagca caccgacagc gccaccctca cgtgcctgcg ctacctggcg  
480  
cagcgtctga cccagctctc actcgcgctg gtgttcacgc acgggatctc cgtcgcagag  
540  
cagccggctc gtgtcctgga cgacctctc taccgcacga gcgcgcgcca ctccacctg  
600  
gagccgctct cgcgggtgga catcatgggg ctggccgcga accggctccc ggtcctcccg  
660  
tcggaccggc tcatcgtcga gatccaccgg ctgagcgggg gcaatccgct gctcgcaccg  
720  
gcactcatcg aggagcacag gctgcgtctc gcgtccgatt ccgtcgcggg gccgatgccg  
780  
ccgcagagcg acggcatcgg ccgccggctc accaccgagg gcggggcgctc gatcggcccc  
840  
gccttctacc aggccgtcct cgcctacctg caccggctcg gcccgcgcgc ggtccgtctc  
900  
gcccggtgcg tcgccctcct ggacgaggcg acgactcccc ttctgttgag tcggctcagc  
960  
ggtatcgaca cggaactgtc gaaacgttac ttacggttgt tcaccggctt gggcgtgctg  
1020  
gagggcgcgc ggttgccgca cgcgggcgta cggcaggccg tactgggtga gatgcctcac  
1080  
ggcgaggcga cccagcagcg ctaccgcgcc gctcgcctcc tgaacgaggg aggagcggcg  
1140  
ccgcaagccg tcgccatgca tctgctcggc gtcggtccgc tccacgacgg atgggtgctg  
1200

cccgtactcc aggaggccgc ggcccacgcc atggaggacg gtgacgtacc gcagggcatc  
 1260  
 cgttatctgg agctggcctg cgaatgctcc ctggacgagg gacagcggct ctcggcgaag  
 1320  
 tcgctgtacg ccttcggcca gtggcagctc aggcccgccg agtccgcccc gcacttcgcg  
 1380  
 gcgctcaagg gccccatcct tgaggggaag ctcacgggca ccgacgcctt gtgggtcgcc  
 1440  
 gagggcatgc tcttccacct cgacttcgac gagggcgtcg aggttgtcga ccatgtcaac  
 1500  
 agcggcgagg cggacatgtc caccgcgctg cacagcaccg ggatgctcat ggccgcggag  
 1560  
 gtccccggcc tgctcgaacg gctggagcac ccaactgccg ccacaaccgc cccggcgacg  
 1620  
 tcccattcgg agttaagagc ccgtcacgcc ctcgccctca tccttgagaa cggagcggac  
 1680  
 aagtacgcca tcgccctggc cgaccaggtg ttccagggca gccagaactg gccgacgtcg  
 1740  
 aagctcagcg gcctgccgaa ggcgctgctg gccctgtget acgcggatca actcgacacc  
 1800  
 gccgcccgcg ggtacgacct gctcgcggcc gaggtggaac gacacgacgc ccctggctgg  
 1860  
 tgggcccaga tcgaaagcgt cggcgcgctc ctggcgctgc ggcgcgggcg gctggcggac  
 1920  
 gccgtacgcc aggcggagac ggcccacgcc cggtgttccg ggccgaggtg gaacgtcagc  
 1980  
 agcgcgcttg cgctgaccgt gctcatcgag gcccacaccg ccatgggcaa ccatcagacc  
 2040  
 gcggccaagt atctggagac gtccgagccg ccgcccgcgc tgttcctcac ccgcgcgggg  
 2100  
 ctgcactacc tgtacgcgcg cggccgtcat cacctcgcga ccggcaaacac ctacctggct  
 2160  
 ctgtccgact tccaggagtg cggcacgctg atgcgtcget ggaacatcga cacaccctcc  
 2220  
 ctggcgccct ggcggctggg agaggcggag gtgtggctgc gcctgggcga ccagaaacag  
 2280  
 gcggccccggc tcgtcgagaa gcagctggcc aaccggacg ccgggctcac ccggtcgcgc  
 2340  
 ggtatgacct tgcacgcctt ggccctggtc caggcaaccg ccaagcagcc cccgatcctg  
 2400  
 cgggacgcgt tccgtctgct ggaggccgcg ggcgcgcggt acgaggctgc ccgtgtcctg  
 2460  
 gccgatctca gccgcgccta ccagcagttg ggcgacaaac aggcccggcc gaccgcacgg  
 2520

cgggcctggc ggctcgccaa gagctgccag gcggagtccc tctgccaggc tctgctcccg  
2580

acatccatac cccagaacat ggacacaaag ccgagcgagg ggtcctgcgg cgccgaccgc  
2640

gcgggccagg acagcttcgg cacactcagt gaatccgaac ggcgcgctgc cacactggcc  
2700

gcccaggggt acgccaaccg tgagatcgcc gagaggctgt tcatcacggt cagcaccgtg  
2760

gagcagcatc tgaccgcgt ctaccgcaag atgggcatca ggaaccgca gcagctcctg  
2820

cagagagccc acgcggtcag ctacgagtcc gtctga  
2856

<210> 3  
<211> 951  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 3

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

Ile Gln Ser Thr Arg Gly Arg Gly Ala Ile Ala Ala Ile Ser Gly Ser  
20 25 30

Thr Ala Ile Gly Lys Thr Thr Thr Leu Asn Ala Leu Ala Glu Arg Ala  
35 40 45

Thr Ser Ala Asp Ile Thr Val Leu Ser Val Val Ser Ser Pro His Glu  
50 55 60

Arg Glu Val Pro Tyr Ser Ala Leu Ala Gln Leu Leu His Ser Ile Glu  
65 70 75 80

Ala Gln Cys Thr Thr Ala Asp Val Pro Arg Ala Gly Gly Ala Gly Thr  
85 90 95

Val His Ala Gly Arg Gln Ala Asp Glu Ala Ala Ala Leu Ala Pro Pro  
100 105 110

Leu Ala Gln Asp Asp Pro Met Thr Val Ala Arg Arg Thr Tyr Gln Ile  
115 120 125

Ile Ala Glu Leu Thr Ala Leu Arg Pro Leu Leu Ile Thr Val Asp Asp  
130 135 140

Ile Gln His Thr Asp Thr Ala Thr Leu Thr Cys Leu Arg Tyr Leu Ala  
145 150 155 160

10

Gln Arg Leu Thr Gln Leu Ser Leu Ala Leu Val Phe Thr His Gly Ile  
 165 170 175  
 Ser Val Asp Glu Gln Pro Ala Arg Val Leu Asp Asp Leu Leu Tyr Arg  
 180 185 190  
 Thr Ser Ala Arg His Phe His Leu Glu Pro Leu Ser Arg Val Asp Ile  
 195 200 205  
 Met Gly Leu Ala Ala Asn Arg Leu Pro Val Leu Pro Ser Asp Arg Leu  
 210 215 220  
 Ile Val Glu Ile His Arg Leu Ser Gly Gly Asn Pro Leu Leu Ala Gln  
 225 230 235 240  
 Ala Leu Ile Glu Glu His Arg Leu Arg Leu Ala Ser Asp Ser Val Ala  
 245 250 255  
 Gly Pro Met Pro Pro Gln Ser Asp Gly Ile Gly Arg Arg Ser Thr Thr  
 260 265 270  
 Glu Gly Gly Ala Ser Ile Gly Pro Ala Phe Tyr Gln Ala Val Leu Ala  
 275 280 285  
 Tyr Leu His Arg Leu Gly Pro Arg Ala Val Arg Leu Ala Arg Cys Val  
 290 295 300  
 Ala Leu Leu Asp Glu Ala Thr Thr Pro Leu Leu Leu Ser Arg Leu Ser  
 305 310 315 320  
 Gly Ile Asp Thr Glu Leu Ser Lys Arg Tyr Leu Arg Leu Phe Thr Gly  
 325 330 335  
 Leu Gly Val Leu Glu Gly Ala Arg Leu Arg His Ala Gly Val Arg Gln  
 340 345 350  
 Ala Val Leu Gly Glu Met Pro His Gly Glu Ala Thr Gln Gln Arg Tyr  
 355 360 365  
 Arg Ala Ala Arg Leu Leu Asn Glu Gly Gly Ala Pro Pro Gln Ala Val  
 370 375 380  
 Ala Met His Leu Leu Gly Val Gly Pro Leu His Asp Gly Trp Val Leu  
 385 390 395 400  
 Pro Val Leu Gln Glu Ala Ala Ala His Ala Met Glu Asp Gly Asp Val  
 405 410 415  
 Pro Gln Gly Ile Arg Tyr Leu Glu Leu Ala Cys Glu Cys Ser Leu Asp

420					425					430					
Glu	Gly	Gln	Arg	Leu	Ser	Ala	Lys	Ser	Leu	Tyr	Ala	Phe	Gly	Gln	Trp
		435					440					445			
Gln	Leu	Arg	Pro	Ala	Glu	Ser	Ala	Pro	His	Phe	Arg	Ala	Leu	Lys	Gly
	450					455					460				
Pro	Ile	Leu	Glu	Gly	Lys	Leu	Thr	Gly	Thr	Asp	Ala	Leu	Trp	Val	Ala
465					470					475					480
Glu	Gly	Met	Leu	Phe	His	Leu	Asp	Phe	Asp	Glu	Ala	Val	Glu	Val	Val
				485					490					495	
Asp	His	Val	Asn	Ser	Gly	Glu	Ala	Asp	Met	Ser	Thr	Ala	Leu	His	Ser
			500					505					510		
Thr	Arg	Met	Leu	Met	Ala	Ala	Glu	Val	Pro	Gly	Leu	Leu	Glu	Arg	Leu
		515					520					525			
Glu	His	Pro	Leu	Pro	Ala	Thr	Thr	Ala	Pro	Ala	Thr	Ser	His	Ser	Glu
	530					535					540				
Leu	Arg	Ala	Arg	His	Ala	Leu	Ala	Leu	Ile	Leu	Glu	Asn	Gly	Ala	Asp
545					550					555					560
Lys	Tyr	Ala	Ile	Ala	Leu	Ala	Asp	Gln	Val	Phe	Gln	Gly	Ser	Gln	Asn
				565					570					575	
Trp	Pro	Thr	Ser	Lys	Leu	Ser	Gly	Leu	Pro	Lys	Ala	Leu	Leu	Ala	Leu
			580					585					590		
Cys	Tyr	Ala	Asp	Gln	Leu	Asp	Thr	Ala	Ala	Ala	Trp	Tyr	Asp	Leu	Leu
		595					600					605			
Ala	Ala	Glu	Val	Glu	Arg	His	Asp	Ala	Pro	Gly	Trp	Trp	Ala	Gln	Ile
	610					615					620				
Glu	Ser	Val	Gly	Ala	Leu	Leu	Ala	Leu	Arg	Arg	Gly	Arg	Leu	Ala	Asp
625					630					635					640
Ala	Val	Arg	Gln	Ala	Glu	Thr	Ala	His	Ala	Arg	Leu	Ser	Gly	Pro	Arg
				645					650					655	
Trp	Asn	Val	Ser	Ser	Ala	Leu	Ala	Leu	Thr	Val	Leu	Ile	Glu	Ala	His
			660					665					670		
Thr	Ala	Met	Gly	Asn	His	Gln	Thr	Ala	Ala	Lys	Tyr	Leu	Glu	Thr	Ser
		675					680					685			



Glu Pro Pro Pro Ala Leu Phe Leu Thr Arg Ala Gly Leu His Tyr Leu  
 690 695 700  
 Tyr Ala Arg Gly Arg His His Leu Ala Thr Gly Asn Thr Tyr Leu Ala  
 705 710 715  
 Leu Ser Asp Phe Gln Glu Cys Gly Thr Leu Met Arg Arg Trp Asn Ile  
 725 730 735  
 Asp Thr Pro Ser Leu Ala Pro Trp Arg Leu Gly Glu Ala Glu Val Trp  
 740 745 750  
 Leu Arg Leu Gly Asp Gln Lys Gln Ala Ala Arg Leu Val Glu Lys Gln  
 755 760 765  
 Leu Ala Asn Pro Asp Ala Gly Leu Thr Arg Ser Arg Gly Met Thr Leu  
 770 775 780  
 His Ala Leu Ala Leu Val Gln Ala Thr Ala Lys Gln Pro Pro Ile Leu  
 785 790 795  
 Arg Asp Ala Phe Arg Leu Leu Glu Ala Ala Gly Ala Arg Tyr Glu Ala  
 805 810 815  
 Ala Arg Val Leu Ala Asp Leu Ser Arg Ala Tyr Gln Gln Leu Gly Asp  
 820 825 830  
 Lys Gln Ala Arg Pro Thr Ala Arg Arg Ala Trp Arg Leu Ala Lys Ser  
 835 840 845  
 Cys Gln Ala Glu Ser Leu Cys Gln Ala Leu Leu Pro Thr Ser Ile Pro  
 850 855 860  
 Gln Asn Met Asp Thr Lys Pro Ser Glu Gly Ser Cys Gly Ala Asp Arg  
 865 870 875 880  
 Ala Gly Gln Asp Ser Phe Gly Thr Leu Ser Glu Ser Glu Arg Arg Val  
 885 890 895  
 Ala Thr Leu Ala Ala Gln Gly Tyr Ala Asn Arg Glu Ile Ala Glu Arg  
 900 905 910  
 Leu Phe Ile Thr Val Ser Thr Val Glu Gln His Leu Thr Arg Val Tyr  
 915 920 925  
 Arg Lys Met Gly Ile Arg Asn Arg Glu Gln Leu Leu Gln Arg Ala His  
 930 935 940  
 Ala Val Ser Tyr Glu Ser Val  
 945 950

ES 2 543 203 T3

<210> 4  
 <211> 16749  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

5

<400> 4

atgagggagg cactcgccat catgtcgagt gatcttgttc acggcacgga cgcaatcgca  
 60  
 atcaccggca tgtcatgccg cctccccag gcacccgacg ccaactcctt ctgggagttg  
 120  
 ctgcggtcgg gccgtagtgc gatcaccgag gtgccgcccgg accgctggga cccggacgag  
 180  
 gtgctgcccgg actcaccgga gcggcaccgc gcagcgtgc gccacggcgg gttcctcgac  
 240  
 cgagtggacc agttcgacgc ggccttcttc gggatctcgc cgcgcgaggc cgtggcgatc  
 300  
 gacccgcagc agcggctctt cgccgagctg gcctgggagg cgctggagga cgcgggaatc  
 360  
 gtccccgaga cgctgcggtc caccgcgacc gccgtcatcg tgggagcgat cgccggggac  
 420  
 tacgcgggcg tggcacaccg cggcgggtgc atcacacagc actcgtgcc ggggctgaac  
 480  
 cgcggtgtca tcgccaaccg ggtgtcctac gcgctgggac tgaacggccc gagcatggcg  
 540  
 gtcgactcgg cgcagtcgtc gtcgctcgtg gcggtgcatc tggccgtgga gagtctgcgc  
 600  
 aaggggtgagt gcaccctggc cctggcgggc ggtgtggcgc tcaacttcgc gcccgagagc  
 660  
 gccgaagtgc ccggaatggt cggcgggttg tcgccggacg gacgctgctt caccttcgac  
 720  
 gcgcgggcca acggctacgt gcgcggtgag ggcggcggcg tcgtcgtact gaagccgctg  
 780  
 gcacacgcgg tgcgcgacgg cgacacggtg tacggggtca tccgtggcac cgcggtcaac  
 840  
 aacgacggct ccaccgacgg gctcacggtg cccagcgcgg aggccaggc gaccgtgctg  
 900  
 cgccaggcgt gcgaggacgc cggcgtggac ccggccgagg tccgttacgt ggaactccac  
 960  
 ggcaaggcca ctccgacggg cgatccgctg gaggcggcgg gtgtgggccc cgcgtacggc  
 1020  
 agcgcgcgtc ctgcgggctc accggtactc gtcggctcgg ccaagaccaa cgtcgggcac  
 1080  
 ctggagggcg cggcgggcat cgtcgggctg atcaagacgg cgctgagcat ccggcaccgg  
 1140

gagatcccgg ccagcctcaa ctacgagacg cccaacccgc ggatcgaccc cgaggcgctg  
 1200  
 aatctgcgcg tccagaccgc gtccggcccc tggccggacg cgccgctgct ggccggggtc  
 1260  
 agctccttcg gtgtgggagg aacgaactgc catgtcgtac tggccgaggc gcccgagcgg  
 1320  
 gccgcgtccg aggaggacgc gccgcagggc gacgagccgg agatcccgct ggctccgtgg  
 1380  
 ctggtgtccg ggcgtaccga ggcggctctg cgcgcccagg ccgggaggct gctggagcgg  
 1440  
 cggacggcgg acgcccacgc gttcgacatc ggccgctcgc tcgaggcac ccgtaccac  
 1500  
 ttcgagcacc gcgccgtcgc gctcgggctc ggacacgacg cgcagcttga ggcgttcgg  
 1560  
 tccggcggccg acgtaccggg gctcgtcacg ggggtcaccg gcgaccacgg gaagatcgcg  
 1620  
 ctggtgttcc cggggcaggg ctcgcagtgg gaggcatgg cgcgcgagtt gatgagcagc  
 1680  
 tcggcggctt tccgcgcgtc gatcgaggcg tgccacgaag ccctcgcgcc gtacgtcgac  
 1740  
 tggtcgctgc tggacacgct caccgatgag tccggcgcga cgtccctcga ccgcgcggac  
 1800  
 gtcgtccagc ccgtgctggt cgcgggtgatg gtctcgtcgc ccagggtgtg ggagtcgctg  
 1860  
 ggtgtacggc ccgacgcggg catcgggcac tcgcagggcg agatcgcggc ggcgcacatc  
 1920  
 gcgggagcgc tcgacctggc ggacgccgcc aggatcgtgg ccctgcgcag ccagacgatc  
 1980  
 atgacgctgg cgggtaccgg ggccatggcg tcggtgccgc tggcggccga ccgggtcacc  
 2040  
 gagtacatcg ccccttcggt cgacgggctg agcatcgccg cgggtcaacgg gccgaccacc  
 2100  
 actgtcgtgg ccggaaaccc cgacgccatc gccgagttgc tggcccgtg cgaggcggag  
 2160  
 gggattcgcg cgagggccgt ctcggccgtg gacttcgcct ccactcttc gcacatggag  
 2220  
 gcggtcaagg accggttgct ggagcagttc gccgaggtga cgccgcgttc gtgcgacatc  
 2280  
 gcgttctact ccacggtcac cgcgagcggc atcgacacgg ccggtctcga cgccggctac  
 2340  
 tggtaactca accttcgccg gcccgtcctc ttcgaggcga cgctcagggc catggcggag  
 2400  
 gacggcttcg gcacgttcgt cgagtcgagt ccgcaccccc ttctcacgct cggattgcgg  
 2460

gcgacgctgc cggacgcgct ggtcgcggac tcgctgcgcc gtaacgaggc tccgtggccg  
 2520  
 cagctgctga cctccctggc ggaactgcac gtatccggcc tgcccgtgga ctggtccgcc  
 2580  
 gtcttcaagg ggcgtacacc gggcgcgctg gcgctgcccc cgtacgcctt ccagcgcgag  
 2640  
 cgctactggc ccgaggtctc gaccgcgttc gagccccggga cgcgcggcgc cgtccagcac  
 2700  
 caggaagccg cgcgcgagga gatccccgcg gcgagctcca cctggtcggg tcggtggcc  
 2760  
 ggactgccgg cggacgagcg ctcccgtgag gcgctggagc tggcgcggct gcgcacggcg  
 2820  
 atcgtgctcg gtcactgag cacggacggg gtcgatgtcg gccaggcgtt ccgcgagctg  
 2880  
 ggcatggact ccacgatggc cgtccagctc cgtcagaacc tcgtggacat caccgggctg  
 2940  
 gcgctgccgg agaccgctgt cttegactac gcgagcccgt cccggctggc cagccggctc  
 3000  
 tgcgaactgg ccctgggcca ggacacgctg tcggccgcgt ccgcgctgtc gcggtcggcg  
 3060  
 tcggtgctgg acgccgacga tccgatcgtg atcgtcggca tggcctgccc ttacccccggc  
 3120  
 ggcgccagca cccccgacga gctgtggcag ctcgtcgacg acggcgtgag cgcgatctcg  
 3180  
 ggcttcccc a ccgacgcgg gtgggacctg gacgccctgt acgacccccg gccccggggtg  
 3240  
 cgcggcaaga cctacacacg gcacggcggg ttcctcgacg aggccgccga gttcgacacc  
 3300  
 gagttcttcg ggatcagccc gcgcgaggcc accgcgatgg acccgcagca gcggtgctc  
 3360  
 ctccaggtca cctgggaggc gctggaacgc gccgggatcg acccggacgg cctccagggc  
 3420  
 agcagcaccg gtgtgttcgt gggcgcgatg tcgcaggagt acggcccccg gctgcacgag  
 3480  
 ggcgacgacg gactgggccc ctatctgctc accggcacca ccgcgagtgt cgtctccggg  
 3540  
 cggatctcgt acaccttcgg tcttgagggc ccggcggcga ccgtcgacac ggcctgctcg  
 3600  
 tcgtcgtcgc tcgcgatgca ccaggcggcg caggcgtgc gcgtcgggga gtgctcgtg  
 3660  
 gccctggcgg gcggcgtcgc ggtcatggcg acgccccgca tgttcgtgga gttcggacag  
 3720  
 cagcgtggtc tggctccccg cggccggctg aagtcgttcg ccggtgcggc ggacggcacc  
 3780

atctgggccg agggcgcggg catggtcctc ctggagcgtc tgtccgacgc gaaggccaac  
 3840  
 ggacacacgg tcctggccgt catccgcggc tccgcggtca accaggacgg cgccagcaac  
 3900  
 ggccctcaccg cccccaacgg gccctcgag cagcgggtca tcaccgccgc tctggccggc  
 3960  
 gcgggcctca cgcccgaacca ggtcgacgcg gtcgaggcac acggcaccgg aaccccgttg  
 4020  
 ggcgacccga tcgaggcca ggcgctcctg gccacctacg gccagaaccg tgaagaaccg  
 4080  
 ctgtggctcg ggtcgttgaa gtcgaacatc ggccacacgc aggccgccgc cggtatcggc  
 4140  
 ggggtcatca agatgatcca ggccatgcgc cacggcacc tgccccggac cctgcacgtc  
 4200  
 gacgagccca gcccgacat cgactgggac tccggcaacg tgcggctcct caccgaggcc  
 4260  
 cgggcctggc ccgagaccga ccacccgcgc cgctccgccg tctctcctt cggcatcagc  
 4320  
 ggcaccaacg cccacctcat cctcgaacag gccccgcga cgccggagcc cgtggacggc  
 4380  
 gacgacgagc aggagacccc gcagggggcc ctcgttcctt ggttctctgtc cgccaagagc  
 4440  
 gccccgcgc tccgcgacca ggcccagcgc ctctcgcacc atgtcatcgc gcgccccggg  
 4500  
 gtcgacccgg cgcacatcgg ccgtgccttg acagccaccc gcgcccgttt ccagcaccgt  
 4560  
 gcggtgggtg tgggagaggg ccgtgacgaa ctactcgcgg gtcttcgggc gctgagcaac  
 4620  
 gacgagtcac cccgcgcggt cgtcaccggc acggcacggg aaggcaccac cgcgttcttg  
 4680  
 ttcacgggac agggcgcgca gcgggcccggc atgggcccgc agctctacga cacgtacccg  
 4740  
 gtcttccggg acagcttcga cgaggtctgc gccaccctcg accggcatct gaacgccgaa  
 4800  
 cagccgggtca aggacgtcgt ctctcgcgac gacgccaccc tgctcaacca gaccgcctac  
 4860  
 acccaggccg ctctcttcgc gatcgagacc tcgctctacc gcctggtcga atcattcggg  
 4920  
 atcacccccg agcacctgac cggccactcc atcgggtgaac tcaccgccgc ccatatcgcg  
 4980  
 ggcgtcttca ccctgaacga cgctgcggc ctcgtcgccg cgcgcggctc actgatgcag  
 5040  
 gccctgcccg ccaacggcgc gatgatctcc ctgcgcgcca ccgaggagca gatccttccg  
 5100

ttctcgaag gccacgagca ccacgtcgcc atcgcggcgg tcaacgggccc caactcgate  
 5160  
 gtcatatcgg gcgaccagga agccaccacc gccatcgcgg aagccctggc cgagacaggt  
 5220  
 gtcaagacgc ggcgcctcac ggtctccac gcgttccact cccccacat ggaccccatg  
 5280  
 ctggaggagt tcgagcgtac ggcggcggac ctgacgtacc acgccccgac gagccccgatc  
 5340  
 gtctccaacc tcaccggcca actcgcggac caccgcatca ccaccccgca gtactgggtc  
 5400  
 cagcacgtcc gggacgccgt gcgcttcgcc gacaccatca ccaccctcga tcagctcggc  
 5460  
 accggcact acctcgaact cggacccgac cccgtcctca ccactctcgt caacgagacc  
 5520  
 ctcggaaga cccggggcac catccccacc gccgtcctgc gcaaggggca ctccgaagcc  
 5580  
 gccacgttc tcacggcgct cgccaccgcy tacaccgccg gcgcgcccgt caacctcagc  
 5640  
 agccacctc cggcccccca caccacccc gacctgccc cgtaccctt ccagcgccag  
 5700  
 cgttactggc acgcggccac caccgccacg ggagacgtgt cgtcggccgg gctgaccgcc  
 5760  
 acgggacacc cggtgctgac cacggccgcc gagcttccc atccgggccc gttgctgctg  
 5820  
 accggccggg tgaacgcggc gtcacccgcc tgggcggcgg accacgccgt ctccggcacc  
 5880  
 ccggtcatgc cgggcgtggc gttcgtcgac atgctgctgc acgcccggc gctggtcggg  
 5940  
 cgccctcgt tcgaggaact caccacccat gtcttctcgg ccctgcccga acacggcgcc  
 6000  
 ctccagctga gggtggtcgt ccgcccggc gacgactccg ggcggcggtc ctccgccgtc  
 6060  
 cactcgcggc ccgaggacgc cccgctgggc gccgactgga cctgccacgc caccggtgcg  
 6120  
 ctgggcgtc ccccggccgt accgccggc ctccgcccg tggacgcggc ctggcccccg  
 6180  
 gcctccgcc cggtcctcga caccgacggc ttctaccggc ggatcgccga ggccgggttc  
 6240  
 ggctacggc cggtcttcca ggggctcgcc gccgcctggg aggacggaga cacgctgtac  
 6300  
 gccgaggtc ccctgcccgc ggggacctc ccgggctcgt acggcgtcca ccccggcctg  
 6360  
 ctgactccg cgctgcaccc gatcgccctg gccgcgaccg gtgccgagac ggacggcaca  
 6420

ctccatgtgc cgttctcctg gagcgggtgtg acgctccacg cgtccggcgc gcacaccctg  
 6480  
 cgcgtccggc tcgtgcgctc cacgcccggag acggtcgcgc tgaccgtcac ggaccctgcc  
 6540  
 ggcgcgccgg tgctgaccgt cgactccctc gccatgcgcg gggtcctgtc cgagcagctc  
 6600  
 gaagccgcca ggccccgaccg cgacgggtgcg ctgcacgacg tcgcctggcg cgcggtgccc  
 6660  
 gcgcccggccac gcgccaaactc gcgccccgac ggcgtcggct gggccgtggt gggagacacc  
 6720  
 cgtgaccgcg gagtcgcccgc cgtgctggca tcgctcgggtg cggccgccga gtcgtaccgc  
 6780  
 gacgccgaag cgctgcgcgc ctcgttgcgg gccggcgcgg tccggccctc gacgatcgtc  
 6840  
 gcccgtttcg ccaccgggac cgccgaagac ggcgcccgacc cggtcgcggc ggcacacacc  
 6900  
 gggacacgcc acgccatgca cctgctccag gccgtcctgg ccgacggggc accggactcc  
 6960  
 cggctggtga tcctcaccga gaacgcgaga tacaccgga ccggcgacgc ggcggccgac  
 7020  
 atggccgggtg cggcgggtctg gggcctgatc cgttcggccc agtccgagca ccccgaccgg  
 7080  
 ttcacgtcc tcgacgtcga cggctcgcgg gcctccgacg aggccgtcgt cgcggcgtc  
 7140  
 tccgccgggtg agccccaaact ggccgtacgc gaaggccggc tgttcgttcc ccgcctcggc  
 7200  
 cgtctgacct cgggcgcgat tcccgccacg ttcgacgcgc ggcgtacggt gctcatcacc  
 7260  
 ggcggcaccg gcgcgctggg ctcgattctc gcgcgccacc tggtcacccg gcacggcgtc  
 7320  
 aggaagctgc tgctgaccag caggagcggc cgtacggcgg acagcaccgt cgtcgcggaa  
 7380  
 ctcgccggac tcggcgcga ggtcaccgtc gcggcctgcg acgtcaccga ccggctgtcg  
 7440  
 ctggagacct tgctcgcggg cctgccgacg gagcatccgc tcggcgcggc cgtgcactgc  
 7500  
 gcgggcgtcc tggacgacgg tgtggtcacg gagctgaccg aggaccggct ggacgcgggtg  
 7560  
 ctccgcccga aggtggacgg cgcgtggaac ctgcaccagg tgaccctga catgggcctg  
 7620  
 gacctggacg ccttcgtgct gttctcgtcc gtggtcgggtg tcctcggctc gcccgacag  
 7680  
 gccaaactac ccgccgcaa ctcttctc gacgaactcg ccgagcaccg ccgtacggcc  
 7740

gggctccccg ccaagtcctt cgcctgggga ctgtgggaga gcggcatggc cgacacctc  
 7800  
 gacgagcagg accgggcgcg gatgagccgc ggcggactct cgccgatgcc cgccgagcgg  
 7860  
 gcgctcgcgc tcttcgactc ggcactcgcg acggcgcggg cggtgctcgt gcccgccggg  
 7920  
 gtcgatgtct cccacgcgcg cacgcagcgg gcgtcgatgc tctccccctt gctcgccgaa  
 7980  
 ctgctccccg cccaggcggc gcccgccgaa cgaggcgggtg aagcggtcga cgagtcctcg  
 8040  
 ctgcccagc agttggcctt cctgcccag cccgaacagc gcgaactcct cctggaggtc  
 8100  
 ctgcccagc atgtcgcggc ggttctgggc cacagctctc cgctcgtcat cgaccccag  
 8160  
 agctcgttca aggacctcgg tttcgactcg ctcgccgga tcgaactcct catggtgctg  
 8220  
 ggcgagtcca tggggctgca cctgccgtcc acgatgctgt tcgaccacc gacgcccag  
 8280  
 ttgctgatca cccacctcag ggacgaactg gtcgacgacg aggccgtacc tgtcaccgcc  
 8340  
 gcggacaccg cggccgtcgc cgtggcacca cgcgacgacg acgaacccat cgccgtgatc  
 8400  
 ggcgatggggg gccgctacc gggcggcgc accacgcccg acgagctgtg gcgactggtc  
 8460  
 accgagggggg tggacgccat cggctcctt cccaccaacc gcggctggga tctggaggag  
 8520  
 ctgttcgatc ccgaccccga catgcgcggg aagacctatg cccgcaaggg cgggttcctc  
 8580  
 tacgacgccg accgtttcga ccccgagttc ttcggcatca gccccgcga ggccctggcg  
 8640  
 ctgacccgc agcagcgact cctcctggag accacctggg agacgttcga gaacgcgggc  
 8700  
 atccgccccg acaccctgcg cggcaagccc gtcggcgtct tcgcccggct cgtcaccag  
 8760  
 gagtacggct ccctcgtcca ccggggcacc gagccggtcg acggcttctt gctgacgggt  
 8820  
 acgacggcga gtgtcgcctc cgggcgactg gcctacacgc tgggtcttga gggcccggcg  
 8880  
 gtcaccgctg acaccgcgtg ctctctctcg ctggtcgcga tgcacctggc ctgcccagtcg  
 8940  
 ctgcccgaaca acgagtcac gatggccctg gccgggtggcg ccacggtgat ggccaacccc  
 9000  
 ggaatgttcc tggagttcag ccgccagcgc ggtctggctc ccgacggccg gtgcaagtcg  
 9060



ttcgccggtg gcgccgacgg caccatctgg gccgagggcg cgggcatggt cctcctggag  
 9120  
 cgtctgtcgg acgccaaggc caacggacac accgttctcg ccgtgatccg cggtcggcc  
 9180  
 gtcaaccagg acggcgccag caacggactg accgctccga gcggcccgtc ccagcagcgg  
 9240  
 gtcacaccg ccgtctggc cggcgcgggc ctcacctccg accaggtcga cgcggtcgaa  
 9300  
 ggacacggca ccggaacccc gttgggagac ccgatcgagg cccaggcgtc cctggcgacg  
 9360  
 tacggccagg gccgtgaagc cgaccagccg ctgtggctcg ggtcgttcaa gtcgaacatc  
 9420  
 ggccacgcgc aggccgcccg cggtatcggc ggggtcatca agatgatcca ggccatgcgc  
 9480  
 cacggcactc tccccggac cctgcacgtc gacgagccca gcccgcacat caactgggccc  
 9540  
 tcgggcaacg tgcggctcct caccgaggag cgcgcctggc ccgagaccga ccaccgcgc  
 9600  
 cgctccgccc tctctcctt cggcatcagc ggcaaccaacg cccatgtcat cctcgaacag  
 9660  
 gccccgcga cgccggagcc cgccgaggat gaccacgagc aggacgcgcc gcaggcgacc  
 9720  
 ctggtcccgt gggctctgtc cggcaagacg gagcaggccc tccgcgacca ggcgagcag  
 9780  
 ctgcgcacgt acctcgaact caacccccggg ctgcgcacgg accgagtcgc tcacgcgctc  
 9840  
 gccaccacc gcgcccagtt ccagtaccgg gccgtggtgc tcggcaccga ccaccaggcg  
 9900  
 ttcgaccgtg ccctgggcac gtcaccctc ggcgagccgt ccccgcgct ggtacgcggg  
 9960  
 acgccgcacc ccggcaagac agccttcatg ttcacgggac agggcgcgca gcgggccggc  
 10020  
 atgggccgcg agctctacga cacgtacccg gtcttccggg acacgttcga cgaggtctgc  
 10080  
 gccaccctcg accggcatct gaacgccgaa cagccggtca aggacgtcgt cttcggccgac  
 10140  
 gacgccatcc tgctcaacca gaccgcctac acccaggccg ctctcttcgc gatcgagacc  
 10200  
 tcgctctacc gcctggtcga atcattcggg atcaccgcc actacctgac cggccactcg  
 10260  
 atcggtgaga tcacggccgc ccacgtggcc gggatcttca cctcgcgca cgctgcccga  
 10320  
 ctggtcggcg cgcgcggctc actgatgcag gctctgcccc ccaacggcgt catgatctcg  
 10380

ctgcgggcca ccgaggagca gatcgtcccc ttcctcgaag gccacgagca ccacgtcagc  
 10440  
 gtcgcgggcg tcaacgggcc cagctcgatc gtcattctcg gcgaccagga agccaccacc  
 10500  
 gccatcgccg gcgctctcgc cgagacgggt gtcaagacgc ggcgcctcac ggtctcccac  
 10560  
 gcgttccact cccccacat ggaccccatg ctggacgagt tcgaactcgt ggccggagag  
 10620  
 ctgacctacc acgccccgac gatccccgat gttccaacc tcaccggcca actcgcggac  
 10680  
 cactacatca ccaccccgca gtactgggtc cagcacgtcc gggaggccgt ccgcttctcc  
 10740  
 gacggcatca ccaccctcga ccggctcggc acccggcact acctcgaact cggacccgac  
 10800  
 cccgtcctca ccaccatggc gcaggacagc gtggccgatg acagcgacgc cgccctcgtc  
 10860  
 gccaccctgt accgcgaccg ggacgagaac cacagcttcc tcaccgccct ggccacggca  
 10920  
 cacgccccatg gcatccaggt cggctggacc cccgtggtcg gtgagacctc ggtccccgcc  
 10980  
 ctcggcctcc ccacctacc ctccagcgc cagcactact ggctggaggc ggcgaagccc  
 11040  
 acctcgggtg ccgacggtct ggggctgacg gcgaccgacc atccggtgct gaccacctg  
 11100  
 gccgaactcc cggacggagg cgggcacctg ttcaccggcc gcgtctccgg gaacgatccg  
 11160  
 gactgggtg ccgagcacat catcttcggg acaatgatc ttcgggtgt ggcttcgtc  
 11220  
 gacctctgc tccacgcggc acgccatgtg gactgcgagc acatcgagga actcaccac  
 11280  
 cacgtgttcc tcgccgtgcc ggagcgcgcc gccctccagc tgcggctcct gatcgagccg  
 11340  
 gcggacagct ccggaagccg ggccttcgcg ttctactcgc ggccggagga cgtcccggtc  
 11400  
 gaaaccgact ggaccctcca cgccacgggc gcgctcggag ccgaacgcag ggaagtcccc  
 11460  
 gcgggcgccc actcgctcag gaacgaggtc tggccgcctg acatctcgga caccatggac  
 11520  
 gtgggggagt tctaccgccg ggtcacggac ggtggcttcg gctacggacc gctgttccga  
 11580  
 gggctcaaga aggcctggca ggacgggaac acgacgtacg cgggaagtctc cttgccccgc  
 11640  
 ggctccgatc ccggcgacta cggcatccac cccggtctgc tcgactcggc gctccagccg  
 11700

gccgcgctca tcatgggaga gaccgaggcg gccgactcga tccgggtgcc gttctcctgg  
 11760  
 gccggtgtgt ccctccacgc gacggggggcc gactccctgc gtatccgcca cacgtggacc  
 11820  
 acaccggaca ccgcgtcgtt ggtcatcgcg gaccagacgg ggacaccggg catgacgatc  
 11880  
 gactccctcg ccatgcggac ggtcggcgcc gaccaactcg ccgccacccg tgcggcggac  
 11940  
 gccggagagc tgtacaaggt cgactggttc gacgtccaga ccgtggagga caagaccag  
 12000  
 ggcgcgggca ccgccaagtg ggcggtggtc gccgaccgga ggaacaccca ggtcgcgcg  
 12060  
 gcactctccc cgctcggcgc cgcggtcagag gtggagccgg acgcggtgac gctcccgacg  
 12120  
 acaccggggg acgacaccac ccggccggac gtggtcttca catggtgcgt ctccgagccc  
 12180  
 ggcgccgacc cggcacaggc cgcgcgctcc ttcaccacc gcgtgctcgg cctcgtccag  
 12240  
 gcggtcctct ccgacgatcg gccggactcc cgcctggtga tcctcaccaa gggcgcgatg  
 12300  
 tccgccggtg gcggcggcgc ggccgacctg gccggagccg cggctctgggg gctgatccgc  
 12360  
 accgctcaga ccgaacaccc cgatcggttc atcctgatgg acctcgacgg ttcggatgcg  
 12420  
 tcgctgcggg ccgtgggcgc tgccctgaac gccggcgaac cccaactcgc cgttcgcgac  
 12480  
 ggccggctcc tcgcccctcg cctcggcccgt atcggcaccg ccgactccga gccgaccgcc  
 12540  
 gcaccggcgt cgttcgacct ggacaagacg gtgctcatca ccggcggcac cggcgcctg  
 12600  
 ggcacgctcc tcgcccgtca cctgggtcacc cgccacgggt tgaagaagct gctcctgacc  
 12660  
 agccggcgcg gccgtccggc aggcagcacc atcaccgagg aactcggcga actcggcgcc  
 12720  
 gaggtcacca tcgtggcctg cgacgcggcg gaccgggagt cgctggaagc cctgctcgcg  
 12780  
 agcctgccgg acgagcatcc gctcggagcg gtggtgcact gcgcgggaac gctcgcgac  
 12840  
 ggcacgtcga ccggcgtgac acctgaccgg ttcgacgagg tgctccggcc gaaggtggac  
 12900  
 ggcgcgtgga acctgcacga gctgaccctg gacctggacc tggacgcctt cgtgacgttc  
 12960  
 tcgtccggtg tcggtgtcct cggctcggcg ggacagtcca actacgccgc cgcgaacgct  
 13020

ttcctcgacg agctggccga acaccgccgt acggccggac tgccccccaa gtccctcgcc  
 13080  
 tggggactgt gggagagcgg catggccgac accctcgacg agcaggacca ggcgcggatg  
 13140  
 aaccgcggcg gcctcctgcc gatgcccgcc gaacaggcac tcgggcactt cgactcggcg  
 13200  
 ctgcgaccg accagaccgt cgtggtcccc gccaaagctcg acctcgccgg gctccgtgcc  
 13260  
 cgcgccgcga cggccccggt ggcgccgatc ttccgtgggc tggccgtac gccgctgcgc  
 13320  
 agcgcggccc aggcggggcg cgcgggagcg gaggtcggag ccctggggca gtcgatcgcg  
 13380  
 ggccgcccgg aggccgagca ggaccagatc atcctggact tcctgcgcaa tcacgtggcc  
 13440  
 accgtcctcg gacacggctc ggcgaacgcg atcgaccccc cgcactcctt caaggagctg  
 13500  
 ggcttcgact cgctcagctc ggtggaactg cgcaactcgc tcaacaaggc gtccgggatg  
 13560  
 cgactcccgt ccacctggt gttcgactac cccaccccct cggactggc cggctacatc  
 13620  
 cgcaaccaac tggcggggcg caagcaggcg gaggcaggcg cgcaagtggc ccgccgcacc  
 13680  
 gttcggccgg cgtcctcgcg gagcgacgcg gccgacccga tcgtgatcgt gggcatgggg  
 13740  
 tgccgcttc ccggtggcgc cgacacgccc gaggcgctgt ggaagctggt cgcggacgag  
 13800  
 cgtgacgcgg tgggggcctt ccccgacaac cgcggctggg acatcgagaa cctcttcgac  
 13860  
 gacgaccccc acgtacgggg gaagtcgtac gccagtgagg gcgggttctt ctacgacgcc  
 13920  
 gaccgtttcg accccgagtt cttcggcatc agccctcgcg aggccctggc gctcgacccg  
 13980  
 cagcagcggc tgctgctcga aaccacctgg gaagcgttcg agaacgcggg catccgcccc  
 14040  
 gacactctgc gcggcaagcc cgtcggcgtc ttcgccggcg tcgcggccgg ggagtacgtc  
 14100  
 tcgctacccc accacggcgg cgagcccgtc gagggttacc tgctgacggg tacgacggcg  
 14160  
 agtgctgcct ccgggcgcat ctcgtacacg ctgggtcttg agggccccgc ggtcaccgtc  
 14220  
 gacacggcct gctcgtcgtc gctcgtcgcg atgcacctgg cgtgccagtc actgcggaac  
 14280  
 aacgagtcca cgatggccct ggccggcggc gccacgatca tgtccaacgc gggcatgttc  
 14340

atggagttca gccgccagcg tggctcggct cccgacagcc gtgccaagtc ctacgcgggc  
14400  
gccgccgacg gcaccatctg ggccgagggc gcgggcatgg tcctcctgga gcgtctgtcg  
14460  
gacgccaaag ccaacggaca cacggtcctg gccgtcatcc gcggctcggc cgtaaccag  
14520  
gacggcgcca gcaacggcct caccgcccc aacgggccc cgcagcagcg ggtcatcaac  
14580  
acggcgctcg ccagcgcggg tctcaccccc gaccaggtcg acgccgtcga aggacacggc  
14640  
accggaacgc cgctgggtga cccgatcgag gcccaggccc tgctctccac ctacggccag  
14700  
aacgtgaag agccgctgtg gctcggatcg ttcaagtcca acatcgcca cgcgcaggcc  
14760  
gccgccggcg tcggcggggc catcaagatg atccaggcca tgcgccacgg caccctgccc  
14820  
cggacgctcc acgtcgacga gccagcccc aacatcgact gggactccgg caatgtgcgg  
14880  
ctcctgaccg aggccccggc ctggccccgag accgaccgcc cgcgccgctc cgccgtctcg  
14940  
tccttcggca tcagcggcac caacgcccc ctcatcctgg aggaagcggc cactcccacc  
15000  
caccctgacg cggccccga gagcgcaccg caggcaacca cggtgccctg gatactctcg  
15060  
ggcaagagtg aacaggccgt gcgggaccag gcccagcgt tgctcgacca cgtcagcgag  
15120  
taccctgacg tccagccggc cgacatcgcg tactcgtctg ccaccgcccg tacctccttc  
15180  
gagcgcagg ccgtcgcgat cggcgcacc catgacgaac tcgtcgacca cctccgctcg  
15240  
ctgaccagg accccggcac cgccctctg cacggccagt cccactcaa gaaggtggcc  
15300  
ctctcttca ccggtcaggg ctcccagcac ccgggcatgg gccgtcagct ctacgacag  
15360  
caccctgtct accgggacgc gttcgacgag gtgaccgcca ccctggacca gcacctccag  
15420  
gccgaacagc cggtaagga cgtcgtcttc gccgacgacc ccaccctgct caaccagacc  
15480  
cggtaacacc agcccccat cttcgccctc cagggtggccc tcaccggct cctcgtcgac  
15540  
gagttcggcg tctccccac ccattctatc ggccactcga tcggcgagat ctcggcggcc  
15600  
cacacggcgg ggatcctcac cctcgacgac gctgcccggc tggtcgccgc ccgcggcact  
15660

ctcatgcaga ccctccccgc caccggcgcg atgacggccg tcgaggcgac cgaggaagag  
15720

gtgctccccgc acctcacgga gcgggtcggt atcgccgcgg tgaacggccc gcgttcggtg  
15780

gtcgtctccg gggacgaagc cgctgtcatc gccgtcggcg aggagtctgc cggtcagggg  
15840

cgacgcatcc gccgtctcac cgtcagccac gccttccact cgcaccacat ggacccgatg  
15900

ctcggcgagc tccacgcggt ggccgacacg ctgacctacc acgtgccacg cccccgctc  
15960

gtctccaccg tcaccggccg cctggccggc tccgagatca ccagcgccac ctactggagc  
16020

gatcacgccc gcaacgccac ccgcttccac gacggcctca acacgcttca cgagcagggc  
16080

gtcaccacgt acatcgaggt cggccccgac gccgtactcg ccgcgctgac ccgtgaagcg  
16140

ctgccccgacg ccaccgccgt accgctgatc cgggccaagg cctccgagcc ggccactctg  
16200

ctcgacgggc tcgtccgggc tcacgtgtcc ggcgccacgg tcgactgggc agggttcctc  
16260

gcacgacgcg ggggcaggag cgtcgacctt cccacgtacg ccttccagcg tcggcgccac  
16320

tggctggaga ccgccgacct cgtcggtacg gccgccggcc tcggtctgga gtccgcctcc  
16380

caccgctgc tggccacgac caccgaactc ccggacggaa ccgccctgtt cacggggcg  
16440

gtgacgctgg ccgaccacc gtggctcagc gaccacaccg tcatgggaac ggtcatcctc  
16500

ccgggcacgg ccttcgtgga gctcgcgctc cacgcggccg agacgggagg tctcgacgag  
16560

atagcggaac tcgtcctgca cgcgccggtc accttcgggt cccagtccgc ggcccttctc  
16620

caggtgatcg tcggacctga cgacccgctg gcgggcccga ccctcaccat caggtccctg  
16680

tcggaagagg accagtcctg gaccgagaac gcgaccggca cgctcggcgc actggtgggg  
16740

gtctcctga  
16749

<210> 5  
<211> 5582  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

<400> 5

Met Arg Glu Ala Leu Ala Ile Met Ser Ser Asp Leu Val His Gly Thr

5

10

ES 2 543 203 T3

1				5					10					15	
Asp	Ala	Ile	Ala	Ile	Thr	Gly	Met	Ser	Cys	Arg	Leu	Pro	Gln	Ala	Pro
			20					25					30		
Asp	Ala	Asn	Ser	Phe	Trp	Glu	Leu	Leu	Arg	Ser	Gly	Arg	Ser	Ala	Ile
		35				40					45				
Thr	Glu	Val	Pro	Pro	Asp	Arg	Trp	Asp	Pro	Asp	Glu	Val	Leu	Pro	Asp
	50					55					60				
Ser	Pro	Glu	Arg	His	Arg	Ala	Ala	Leu	Arg	His	Gly	Gly	Phe	Leu	Asp
65					70					75					80
Arg	Val	Asp	Gln	Phe	Asp	Ala	Ala	Phe	Phe	Gly	Ile	Ser	Pro	Arg	Glu
				85					90					95	
Ala	Val	Ala	Ile	Asp	Pro	Gln	Gln	Arg	Leu	Phe	Ala	Glu	Leu	Ala	Trp
			100					105					110		
Glu	Ala	Leu	Glu	Asp	Ala	Gly	Ile	Val	Pro	Glu	Thr	Leu	Arg	Ser	Thr
		115					120					125			
Ala	Thr	Ala	Val	Ile	Val	Gly	Ala	Ile	Ala	Gly	Asp	Tyr	Ala	Ala	Leu
	130					135					140				
Ala	His	Arg	Gly	Gly	Ala	Ile	Thr	Gln	His	Ser	Leu	Pro	Gly	Leu	Asn
145					150					155					160
Arg	Gly	Val	Ile	Ala	Asn	Arg	Val	Ser	Tyr	Ala	Leu	Gly	Leu	Asn	Gly
				165					170					175	
Pro	Ser	Met	Ala	Val	Asp	Ser	Ala	Gln	Ser	Ser	Ser	Leu	Val	Ala	Val
			180					185					190		
His	Leu	Ala	Val	Glu	Ser	Leu	Arg	Lys	Gly	Glu	Cys	Thr	Leu	Ala	Leu
		195					200					205			
Ala	Gly	Gly	Val	Ala	Leu	Asn	Phe	Ala	Pro	Glu	Ser	Ala	Glu	Val	Ala
	210					215					220				
Gly	Met	Phe	Gly	Gly	Leu	Ser	Pro	Asp	Gly	Arg	Cys	Phe	Thr	Phe	Asp
225					230					235					240
Ala	Arg	Ala	Asn	Gly	Tyr	Val	Arg	Gly	Glu	Gly	Gly	Gly	Val	Val	Val
				245					250					255	
Leu	Lys	Pro	Leu	Ala	His	Ala	Val	Arg	Asp	Gly	Asp	Thr	Val	Tyr	Gly
			260					265					270		

ES 2 543 203 T3

Val Ile Arg Gly Thr Ala Val Asn Asn Asp Gly Ser Thr Asp Gly Leu  
 275 280 285  
 Thr Val Pro Ser Ala Glu Ala Gln Ala Thr Val Leu Arg Gln Ala Cys  
 290 295 300  
 Glu Asp Ala Gly Val Asp Pro Ala Glu Val Arg Tyr Val Glu Leu His  
 305 310 315 320  
 Gly Thr Gly Thr Pro Thr Gly Asp Pro Leu Glu Ala Ala Gly Val Gly  
 325 330 335  
 Ala Ala Tyr Gly Ser Ala Arg Pro Ala Gly Ser Pro Val Leu Val Gly  
 340 345  
 Ser Ala Lys Thr Asn Val Gly His Leu Glu Gly Ala Ala Gly Ile Val  
 355 360 365  
 Gly Leu Ile Lys Thr Ala Leu Ser Ile Arg His Arg Glu Ile Pro Ala  
 370 375 380  
 Ser Leu Asn Tyr Glu Thr Pro Asn Pro Arg Ile Asp Pro Glu Ala Leu  
 385 390 395 400  
 Asn Leu Arg Val Gln Thr Ala Ser Gly Pro Trp Pro Asp Ala Pro Leu  
 405 410 415  
 Leu Ala Gly Val Ser Ser Phe Gly Val Gly Gly Thr Asn Cys His Val  
 420 425 430  
 Val Leu Ala Glu Ala Pro Glu Arg Ala Ala Ser Glu Glu Asp Ala Pro  
 435 440 445  
 Gln Gly Asp Glu Pro Glu Ile Pro Leu Ala Pro Trp Leu Val Ser Gly  
 450 455 460  
 Arg Thr Glu Ala Ala Leu Arg Ala Gln Ala Gly Arg Leu Leu Glu Arg  
 465 470 475 480  
 Arg Thr Ala Asp Ala Asp Ala Phe Asp Ile Gly Arg Ser Leu Ala Gly  
 485 490 495  
 Thr Arg Thr His Phe Glu His Arg Ala Val Ala Leu Gly Leu Gly His  
 500 505 510  
 Asp Ala Gln Leu Glu Ala Leu Arg Ser Gly Ala Asp Val Pro Gly Leu  
 515 520 525  
 Val Thr Gly Val Thr Gly Asp His Gly Lys Ile Ala Leu Val Phe Pro



530					535					540					
Gly 545	Gln	Gly	Ser	Gln	Trp 550	Glu	Gly	Met	Ala	Arg 555	Glu	Leu	Met	Arg	Thr 560
Ser	Ala	Val	Phe	Arg 565	Ala	Ser	Ile	Glu	Ala 570	Cys	His	Glu	Ala	Leu 575	Ala
Pro	Tyr	Val	Asp 580	Trp	Ser	Leu	Leu	Asp 585	Thr	Leu	Thr	Asp	Glu 590	Ser	Gly
Ala	Thr	Ser 595	Leu	Asp	Arg	Ala	Asp 600	Val	Val	Gln	Pro	Val 605	Leu	Phe	Ala
Val	Met 610	Val	Ser	Leu	Ala	Arg 615	Val	Trp	Glu	Ser	Leu 620	Gly	Val	Arg	Pro
Asp 625	Ala	Val	Ile	Gly	His 630	Ser	Gln	Gly	Glu	Ile 635	Ala	Ala	Ala	His	Ile 640
Ala	Gly	Ala	Leu	Asp 645	Leu	Ala	Asp	Ala	Ala 650	Arg	Ile	Val	Ala	Leu 655	Arg
Ser	Gln	Thr	Ile 660	Met	Thr	Leu	Ala	Gly 665	Thr	Gly	Ala	Met	Ala 670	Ser	Val
Pro	Leu	Ala 675	Ala	Asp	Arg	Val	Thr 680	Glu	Tyr	Ile	Ala	Pro 685	Phe	Gly	Asp
Gly	Leu 690	Ser	Ile	Ala	Ala	Val 695	Asn	Gly	Pro	Thr	Thr 700	Thr	Val	Val	Ala
Gly 705	Asn	Pro	Asp	Ala	Ile 710	Ala	Glu	Leu	Leu	Ala 715	Arg	Cys	Glu	Ala	Glu 720
Gly	Ile	Arg	Ala	Arg 725	Ala	Val	Ser	Ala	Val 730	Asp	Phe	Ala	Ser	His 735	Ser
Ser	His	Met	Glu 740	Ala	Val	Lys	Asp 745	Arg	Leu	Leu	Glu	Gln 750	Phe	Ala	Glu
Val	Thr	Pro 755	Arg	Ser	Cys	Asp	Ile 760	Ala	Phe	Tyr	Ser	Thr 765	Val	Thr	Ala
Ser	Ala 770	Ile	Asp	Thr	Ala	Gly 775	Leu	Asp	Ala	Gly	Tyr 780	Trp	Tyr	Ser	Asn
Leu 785	Arg	Arg	Pro	Val	Leu 790	Phe	Glu	Ala	Thr	Leu 795	Arg	Ala	Met	Ala	Glu 800

Asp Gly Phe Gly Thr Phe Val Glu Ser Ser Pro His Pro Val Leu Thr  
 805 810 815  
 Leu Gly Leu Arg Ala Thr Leu Pro Asp Ala Leu Val Ala Asp Ser Leu  
 820 825 830  
 Arg Arg Asn Glu Ala Pro Trp Pro Gln Leu Leu Thr Ser Leu Ala Glu  
 835 840 845  
 Leu His Val Ser Gly Leu Pro Val Asp Trp Ser Ala Val Phe Lys Gly  
 850 855 860  
 Arg Thr Pro Gly Arg Val Ala Leu Pro Thr Tyr Ala Phe Gln Arg Glu  
 865 870 875  
 Arg Tyr Trp Pro Glu Val Ser Thr Ala Phe Glu Pro Gly Thr Arg Gly  
 885 890 895  
 Ala Val Gln His Gln Glu Ala Ala Arg Glu Glu Ile Pro Ala Ala Ser  
 900 905 910  
 Ser Thr Trp Ser Asp Arg Leu Ala Gly Leu Pro Ala Asp Glu Arg Ser  
 915 920 925  
 Arg Glu Ala Leu Glu Leu Val Arg Leu Arg Thr Ala Ile Val Leu Gly  
 930 935 940  
 His Leu Ser Thr Asp Gly Val Asp Val Gly Gln Ala Phe Arg Glu Leu  
 945 950 955 960  
 Gly Met Asp Ser Thr Met Ala Val Gln Leu Arg Gln Asn Leu Val Asp  
 965 970 975  
 Ile Thr Gly Leu Ala Leu Pro Glu Thr Val Val Phe Asp Tyr Ala Ser  
 980 985 990  
 Pro Ser Arg Leu Ala Ser Arg Leu Cys Glu Leu Ala Leu Gly Glu Asp  
 995 1000 1005  
 Thr Ser Ser Ala Ala Ser Ala Leu Ser Arg Ser Ala Ser Val Leu  
 1010 1015 1020  
 Asp Ala Asp Asp Pro Ile Val Ile Val Gly Met Ala Cys Arg Tyr  
 1025 1030 1035  
 Pro Gly Gly Ala Ser Thr Pro Asp Glu Leu Trp Gln Leu Val Asp  
 1040 1045 1050  
 Asp Gly Val Ser Ala Ile Ser Gly Phe Pro Thr Asp Arg Gly Trp

1055					1060					1065				
Asp	Leu	Asp	Ala	Leu	Tyr	Asp	Pro	Glu	Pro	Gly	Val	Arg	Gly	Lys
	1070					1075					1080			
Thr	Tyr	Thr	Arg	His	Gly	Gly	Phe	Leu	Asp	Glu	Ala	Ala	Glu	Phe
	1085					1090					1095			
Asp	Thr	Glu	Phe	Phe	Gly	Ile	Ser	Pro	Arg	Glu	Ala	Thr	Ala	Met
	1100					1105					1110			
Asp	Pro	Gln	Gln	Arg	Leu	Leu	Leu	Gln	Val	Thr	Trp	Glu	Ala	Leu
	1115					1120					1125			
Glu	Arg	Ala	Gly	Ile	Asp	Pro	Asp	Gly	Leu	Gln	Gly	Ser	Ser	Thr
	1130					1135					1140			
Gly	Val	Phe	Val	Gly	Ala	Met	Ser	Gln	Glu	Tyr	Gly	Pro	Arg	Leu
	1145					1150					1155			
His	Glu	Gly	Asp	Asp	Gly	Leu	Gly	Gly	Tyr	Leu	Leu	Thr	Gly	Thr
	1160					1165					1170			
Thr	Ala	Ser	Val	Val	Ser	Gly	Arg	Ile	Ser	Tyr	Thr	Phe	Gly	Leu
	1175					1180					1185			
Glu	Gly	Pro	Ala	Val	Thr	Val	Asp	Thr	Ala	Cys	Ser	Ser	Ser	Leu
	1190					1195					1200			
Val	Ala	Met	His	Gln	Ala	Ala	Gln	Ala	Leu	Arg	Val	Gly	Glu	Cys
	1205					1210					1215			
Ser	Leu	Ala	Leu	Ala	Gly	Gly	Val	Ala	Val	Met	Ala	Thr	Pro	Gly
	1220					1225					1230			
Met	Phe	Val	Glu	Phe	Gly	Gln	Gln	Arg	Gly	Leu	Ala	Pro	Asp	Gly
	1235					1240					1245			
Arg	Cys	Lys	Ser	Phe	Ala	Gly	Ala	Ala	Asp	Gly	Thr	Ile	Trp	Ala
	1250					1255					1260			
Glu	Gly	Ala	Gly	Met	Val	Leu	Leu	Glu	Arg	Leu	Ser	Asp	Ala	Lys
	1265					1270					1275			
Ala	Asn	Gly	His	Thr	Val	Leu	Ala	Val	Ile	Arg	Gly	Ser	Ala	Val
	1280					1285					1290			
Asn	Gln	Asp	Gly	Ala	Ser	Asn	Gly	Leu	Thr	Ala	Pro	Asn	Gly	Pro
	1295					1300					1305			

Ser Gln Gln Arg Val Ile Thr Ala Ala Leu Ala Gly Ala Gly Leu  
 1310 1315 1320

Thr Pro Asp Gln Val Asp Ala Val Glu Ala His Gly Thr Gly Thr  
 1325 1330 1335

Pro Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr  
 1340 1345 1350

Gly Gln Asn Arg Glu Glu Pro Leu Trp Leu Gly Ser Leu Lys Ser  
 1355 1360 1365

Asn Ile Gly His Thr Gln Ala Ala Ala Gly Ile Gly Gly Val Ile  
 1370 1375 1380

Lys Met Ile Gln Ala Met Arg His Gly Thr Leu Pro Arg Thr Leu  
 1385 1390 1395

His Val Asp Glu Pro Ser Pro His Ile Asp Trp Asp Ser Gly Asn  
 1400 1405 1410

Val Arg Leu Leu Thr Glu Ala Arg Ala Trp Pro Glu Thr Asp His  
 1415 1420 1425

Pro Arg Arg Ser Ala Val Ser Ser Phe Gly Ile Ser Gly Thr Asn  
 1430 1435 1440

Ala His Leu Ile Leu Glu Gln Ala Pro Ala Thr Pro Glu Pro Val  
 1445 1450 1455

Asp Gly Asp Asp Glu Gln Glu Thr Pro Gln Gly Ala Leu Val Pro  
 1460 1465 1470

Trp Phe Leu Ser Ala Lys Ser Ala Pro Ala Leu Arg Asp Gln Ala  
 1475 1480 1485

Gln Arg Leu Leu Asp His Val Ile Ala Arg Pro Gly Val Asp Pro  
 1490 1495 1500

Ala His Ile Gly Arg Ala Leu Thr Ala Thr Arg Ala Arg Phe Gln  
 1505 1510 1515

His Arg Ala Val Val Val Gly Glu Gly Arg Asp Glu Leu Leu Ala  
 1520 1525 1530

Gly Leu Arg Ala Leu Ser Asn Asp Glu Ser Ser Arg Ala Val Val  
 1535 1540 1545

Thr Gly Thr Ala Arg Glu Gly Thr Thr Ala Phe Leu Phe Thr Gly

1550						1555					1560			
Gln	Gly	Ala	Gln	Arg	Ala	Gly	Met	Gly	Arg	Glu	Leu	Tyr	Asp	Thr
	1565					1570					1575			
Tyr	Pro	Val	Phe	Arg	Asp	Ser	Phe	Asp	Glu	Val	Cys	Ala	Thr	Leu
	1580					1585					1590			
Asp	Arg	His	Leu	Asn	Ala	Glu	Gln	Pro	Val	Lys	Asp	Val	Val	Phe
	1595					1600					1605			
Ala	Asp	Asp	Ala	Thr	Leu	Leu	Asn	Gln	Thr	Arg	Tyr	Thr	Gln	Ala
	1610					1615					1620			
Ala	Leu	Phe	Ala	Ile	Glu	Thr	Ser	Leu	Tyr	Arg	Leu	Val	Glu	Ser
	1625					1630					1635			
Phe	Gly	Ile	Thr	Pro	Gln	His	Leu	Thr	Gly	His	Ser	Ile	Gly	Glu
	1640					1645					1650			
Leu	Thr	Ala	Ala	His	Ile	Ala	Gly	Val	Phe	Thr	Leu	Asn	Asp	Ala
	1655					1660					1665			
Cys	Arg	Leu	Val	Ala	Ala	Arg	Gly	Ser	Leu	Met	Gln	Ala	Leu	Pro
	1670					1675					1680			
Ala	Asn	Gly	Ala	Met	Ile	Ser	Leu	Arg	Ala	Thr	Glu	Glu	Gln	Ile
	1685					1690					1695			
Leu	Pro	Phe	Leu	Glu	Gly	His	Glu	His	His	Val	Ala	Ile	Ala	Ala
	1700					1705					1710			
Val	Asn	Gly	Pro	Asn	Ser	Ile	Val	Ile	Ser	Gly	Asp	Gln	Glu	Ala
	1715					1720					1725			
Thr	Thr	Ala	Ile	Ala	Glu	Ala	Leu	Ala	Glu	Thr	Gly	Val	Lys	Thr
	1730					1735					1740			
Arg	Arg	Leu	Thr	Val	Ser	His	Ala	Phe	His	Ser	Pro	His	Met	Asp
	1745					1750					1755			
Pro	Met	Leu	Glu	Glu	Phe	Glu	Arg	Thr	Ala	Ala	Asp	Leu	Thr	Tyr
	1760					1765					1770			
His	Ala	Pro	Thr	Ser	Pro	Ile	Val	Ser	Asn	Leu	Thr	Gly	Gln	Leu
	1775					1780					1785			
Ala	Asp	His	Arg	Ile	Thr	Thr	Pro	Gln	Tyr	Trp	Val	Gln	His	Val
	1790					1795					1800			

ES 2 543 203 T3

Arg Asp Ala Val Arg Phe Ala Asp Thr Ile Thr Thr Leu Asp Gln  
 1805 1810 1815  
 Leu Gly Thr Arg His Tyr Leu Glu Leu Gly Pro Asp Pro Val Leu  
 1820 1825 1830  
 Thr Thr Leu Val Asn Glu Thr Leu Gly Lys Thr Arg Gly Thr Ile  
 1835 1840 1845  
 Pro Thr Ala Val Leu Arg Lys Gly His Ser Glu Ala Ala Thr Leu  
 1850 1855 1860  
 Leu Thr Ala Leu Ala Thr Ala Tyr Thr Ala Gly Ala Pro Val Asn  
 1865 1870 1875  
 Leu Ser Ser His Leu Pro Ala Pro His Thr His Pro Asp Leu Pro  
 1880 1885 1890  
 Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp His Ala Ala Thr Thr  
 1895 1900 1905  
 Ala Thr Gly Asp Val Ser Ser Ala Gly Leu Thr Ala Thr Gly His  
 1910 1915 1920  
 Pro Val Leu Thr Thr Ala Ala Glu Leu Pro Asp Pro Gly Gly Leu  
 1925 1930 1935  
 Leu Leu Thr Gly Arg Val Asn Ala Ala Ser Pro Ala Trp Ala Ala  
 1940 1945 1950  
 Asp His Ala Val Phe Gly Thr Pro Val Met Pro Gly Val Ala Phe  
 1955 1960 1965  
 Val Asp Met Leu Leu His Ala Ala Ala Leu Val Gly Arg Pro Arg  
 1970 1975 1980  
 Ile Glu Glu Leu Thr His His Val Phe Leu Ala Leu Pro Glu His  
 1985 1990 1995  
 Gly Ala Leu Gln Leu Arg Val Val Val Arg Pro Ala Asp Asp Ser  
 2000 2005 2010  
 Gly Arg Arg Ser Phe Ala Val His Ser Arg Pro Glu Asp Ala Pro  
 2015 2020 2025  
 Leu Gly Ala Asp Trp Thr Cys His Ala Thr Gly Ala Leu Gly Val  
 2030 2035 2040  
 Ala Pro Ala Val Pro Pro Ala Leu Pro Pro Val Asp Ala Ala Trp

ES 2 543 203 T3

2045			2050			2055								
Pro	Pro	Ala	Ser	Ala	Ala	Val	Leu	Asp	Thr	Asp	Gly	Phe	Tyr	Arg
	2060					2065					2070			
Arg	Ile	Ala	Glu	Ala	Gly	Phe	Gly	Tyr	Gly	Pro	Val	Phe	Gln	Gly
	2075					2080					2085			
Leu	Ala	Ala	Ala	Trp	Glu	Asp	Gly	Asp	Thr	Leu	Tyr	Ala	Glu	Val
	2090					2095					2100			
Ala	Leu	Pro	Ala	Gly	Thr	Ser	Pro	Gly	Ser	Tyr	Gly	Val	His	Pro
	2105					2110					2115			
Gly	Leu	Leu	Asp	Ser	Ala	Leu	His	Pro	Ile	Ala	Leu	Ala	Ala	Thr
	2120					2125					2130			
Gly	Ala	Glu	Thr	Asp	Gly	Thr	Leu	His	Val	Pro	Phe	Ser	Trp	Ser
	2135					2140					2145			
Gly	Val	Thr	Leu	His	Ala	Ser	Gly	Ala	His	Thr	Leu	Arg	Val	Arg
	2150					2155					2160			
Leu	Val	Arg	Ser	Thr	Pro	Glu	Thr	Val	Ala	Leu	Thr	Val	Thr	Asp
	2165					2170					2175			
Pro	Ser	Gly	Ala	Pro	Val	Leu	Thr	Val	Asp	Ser	Leu	Ala	Met	Arg
	2180					2185					2190			
Gly	Val	Arg	Ala	Glu	Gln	Leu	Glu	Ala	Ala	Arg	Pro	Asp	Arg	Asp
	2195					2200					2205			
Gly	Ala	Leu	His	Asp	Val	Ala	Trp	Arg	Ala	Val	Pro	Ala	Pro	Pro
	2210					2215					2220			
Arg	Ala	Asn	Ser	Arg	Pro	Asp	Gly	Val	Gly	Trp	Ala	Val	Val	Gly
	2225					2230					2235			
Asp	Thr	Arg	Asp	Pro	Arg	Val	Ala	Ala	Val	Leu	Ala	Ser	Leu	Gly
	2240					2245					2250			
Ala	Ala	Ala	Glu	Ser	Tyr	Pro	Asp	Ala	Glu	Ala	Leu	Arg	Ala	Ser
	2255					2260					2265			
Leu	Arg	Ala	Gly	Ala	Val	Arg	Pro	Ser	Thr	Ile	Val	Ala	Arg	Phe
	2270					2275					2280			
Ala	Thr	Gly	Thr	Ala	Glu	Asp	Gly	Ala	Asp	Pro	Val	Ala	Ala	Ala
	2285					2290					2295			

His Thr Gly Thr Arg His Ala Met His Leu Leu Gln Ala Val Leu  
 2300 2305 2310  
 Ala Asp Gly Ala Pro Asp Ser Arg Leu Val Ile Leu Thr Glu Asn  
 2315 2320 2325  
 Ala Arg Tyr Thr Gly Thr Gly Asp Ala Ala Ala Asp Met Ala Gly  
 2330 2335 2340  
 Ala Ala Val Trp Gly Leu Ile Arg Ser Ala Gln Ser Glu His Pro  
 2345 2350 2355  
 Asp Arg Phe Thr Leu Leu Asp Val Asp Gly Ser Arg Ala Ser Asp  
 2360 2365 2370  
 Glu Ala Val Val Ala Ala Leu Ser Ala Gly Glu Pro Gln Leu Ala  
 2375 2380 2385  
 Val Arg Glu Gly Arg Leu Phe Val Pro Arg Leu Ala Arg Leu Thr  
 2390 2395 2400  
 Pro Gly Ala Ile Pro Ala Thr Phe Asp Ala Arg Arg Thr Val Leu  
 2405 2410 2415  
 Ile Thr Gly Gly Thr Gly Ala Leu Gly Ser Ile Leu Ala Arg His  
 2420 2425 2430  
 Leu Val Thr Arg His Gly Val Arg Lys Leu Leu Leu Thr Ser Arg  
 2435 2440 2445  
 Ser Gly Arg Thr Ala Asp Ser Thr Val Val Ala Glu Leu Ala Gly  
 2450 2455 2460  
 Leu Gly Ala Glu Val Thr Val Ala Ala Cys Asp Val Thr Asp Arg  
 2465 2470 2475  
 Leu Ser Leu Glu Thr Leu Leu Ala Gly Leu Pro Thr Glu His Pro  
 2480 2485 2490  
 Leu Gly Ala Val Val His Cys Ala Gly Val Leu Asp Asp Gly Val  
 2495 2500 2505  
 Val Thr Glu Leu Thr Glu Asp Arg Leu Asp Ala Val Leu Arg Pro  
 2510 2515 2520  
 Lys Val Asp Gly Ala Trp Asn Leu His Gln Val Thr Arg Asp Met  
 2525 2530 2535  
 Gly Leu Asp Leu Asp Ala Phe Val Leu Phe Ser Ser Val Val Gly



ES 2 543 203 T3

	2540					2545					2550			
Val	Leu 2555	Gly	Ser	Pro	Gly	Gln 2560	Ala	Asn	Tyr	Ala	Ala 2565	Ala	Asn	Ser
Phe	Leu 2570	Asp	Glu	Leu	Ala	Glu 2575	His	Arg	Arg	Thr	Ala 2580	Gly	Leu	Pro
Ala	Lys 2585	Ser	Leu	Ala	Trp	Gly 2590	Leu	Trp	Glu	Ser	Gly 2595	Met	Ala	Asp
Thr	Leu 2600	Asp	Glu	Gln	Asp	Arg 2605	Ala	Arg	Met	Ser	Arg 2610	Gly	Gly	Leu
Ser	Pro 2615	Met	Pro	Ala	Glu	Arg 2620	Ala	Leu	Ala	Leu	Phe 2625	Asp	Ser	Ala
Leu	Ala 2630	Thr	Ala	Arg	Ala	Val 2635	Leu	Val	Pro	Ala	Gly 2640	Val	Asp	Val
Ser	His 2645	Ala	Arg	Thr	Gln	Arg 2650	Ala	Ser	Met	Leu	Ser 2655	Pro	Leu	Leu
Ala	Glu 2660	Leu	Leu	Pro	Ala	Gln 2665	Ala	Ala	Pro	Ala	Glu 2670	Arg	Gly	Gly
Glu	Ala 2675	Val	Asp	Glu	Ser	Ser 2680	Leu	Arg	Gln	Gln	Leu 2685	Ala	Leu	Leu
Pro	Glu 2690	Pro	Glu	Gln	Arg	Glu 2695	Leu	Leu	Leu	Glu	Val 2700	Leu	Arg	Lys
His	Val 2705	Ala	Ala	Val	Leu	Gly 2710	His	Ser	Ser	Pro	Leu 2715	Val	Ile	Asp
Pro	Glu 2720	Ser	Ser	Phe	Lys	Asp 2725	Leu	Gly	Phe	Asp	Ser 2730	Leu	Ala	Gly
Ile	Glu 2735	Leu	Leu	Met	Val	Leu 2740	Gly	Glu	Ser	Met	Gly 2745	Leu	His	Leu
Pro	Ser 2750	Thr	Met	Leu	Phe	Asp 2755	His	Pro	Thr	Pro	Glu 2760	Leu	Leu	Ile
Thr	His 2765	Leu	Arg	Asp	Glu	Leu 2770	Val	Asp	Asp	Glu	Ala 2775	Val	Pro	Val
Thr	Ala 2780	Ala	Asp	Thr	Ala	Ala 2785	Val	Ala	Val	Ala	Pro 2790	Arg	Asp	Asp

Asp Glu Pro Ile Ala Val Ile Gly Met Gly Cys Arg Tyr Pro Gly  
 2795 2800 2805

Gly Ala Thr Thr Pro Asp Glu Leu Trp Arg Leu Val Thr Glu Gly  
 2810 2815 2820

Val Asp Ala Ile Gly Ser Phe Pro Thr Asn Arg Gly Trp Asp Leu  
 2825 2830 2835

Glu Glu Leu Phe Asp Pro Asp Pro Asp Met Arg Gly Lys Thr Tyr  
 2840 2845 2850

Ala Arg Lys Gly Gly Phe Leu Tyr Asp Ala Asp Arg Phe Asp Pro  
 2855 2860 2865

Glu Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Leu Asp Pro  
 2870 2875 2880

Gln Gln Arg Leu Leu Leu Glu Thr Thr Trp Glu Thr Phe Glu Asn  
 2885 2890 2895

Ala Gly Ile Arg Pro Asp Thr Leu Arg Gly Lys Pro Val Gly Val  
 2900 2905 2910

Phe Ala Gly Val Val Thr Gln Glu Tyr Gly Ser Leu Val His Arg  
 2915 2920 2925

Gly Thr Glu Pro Val Asp Gly Phe Leu Leu Thr Gly Thr Thr Ala  
 2930 2935 2940

Ser Val Ala Ser Gly Arg Leu Ala Tyr Thr Leu Gly Leu Glu Gly  
 2945 2950 2955

Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala  
 2960 2965 2970

Met His Leu Ala Cys Gln Ser Leu Arg Asn Asn Glu Ser Thr Met  
 2975 2980 2985

Ala Leu Ala Gly Gly Ala Thr Val Met Ala Asn Pro Gly Met Phe  
 2990 2995 3000

Leu Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys  
 3005 3010 3015

Lys Ser Phe Ala Gly Gly Ala Asp Gly Thr Ile Trp Ala Glu Gly  
 3020 3025 3030

Ala Gly Met Val Leu Leu Glu Arg Leu Ser Asp Ala Lys Ala Asn

ES 2 543 203 T3

3035						3040						3045		
Gly	His	Thr	Val	Leu	Ala	Val	Ile	Arg	Gly	Ser	Ala	Val	Asn	Gln
3050						3055					3060			
Asp	Gly	Ala	Ser	Asn	Gly	Leu	Thr	Ala	Pro	Ser	Gly	Pro	Ser	Gln
3065						3070					3075			
Gln	Arg	Val	Ile	Thr	Ala	Ala	Leu	Ala	Gly	Ala	Gly	Leu	Thr	Ser
3080						3085					3090			
Asp	Gln	Val	Asp	Ala	Val	Glu	Gly	His	Gly	Thr	Gly	Thr	Pro	Leu
3095						3100					3105			
Gly	Asp	Pro	Ile	Glu	Ala	Gln	Ala	Leu	Leu	Ala	Thr	Tyr	Gly	Gln
3110						3115					3120			
Gly	Arg	Glu	Ala	Asp	Gln	Pro	Leu	Trp	Leu	Gly	Ser	Phe	Lys	Ser
3125						3130					3135			
Asn	Ile	Gly	His	Ala	Gln	Ala	Ala	Ala	Gly	Ile	Gly	Gly	Val	Ile
3140						3145					3150			
Lys	Met	Ile	Gln	Ala	Met	Arg	His	Gly	Thr	Leu	Pro	Arg	Thr	Leu
3155						3160					3165			
His	Val	Asp	Glu	Pro	Ser	Pro	His	Ile	Asn	Trp	Ala	Ser	Gly	Asn
3170						3175					3180			
Val	Arg	Leu	Leu	Thr	Glu	Glu	Arg	Ala	Trp	Pro	Glu	Thr	Asp	His
3185						3190					3195			
Pro	Arg	Arg	Ser	Ala	Val	Ser	Ser	Phe	Gly	Ile	Ser	Gly	Thr	Asn
3200						3205					3210			
Ala	His	Val	Ile	Leu	Glu	Gln	Ala	Pro	Ala	Thr	Pro	Glu	Pro	Ala
3215						3220					3225			
Glu	Asp	Asp	His	Glu	Gln	Asp	Ala	Pro	Gln	Ala	Thr	Leu	Val	Pro
3230						3235					3240			
Trp	Val	Leu	Ser	Gly	Lys	Thr	Glu	Gln	Ala	Leu	Arg	Asp	Gln	Ala
3245						3250					3255			
Gln	Gln	Leu	Arg	Thr	Tyr	Leu	Glu	Leu	Asn	Pro	Gly	Leu	Arg	Thr
3260						3265					3270			
Asp	Arg	Val	Ala	His	Ala	Leu	Ala	Thr	Thr	Arg	Ala	Gln	Phe	Gln
3275						3280					3285			

Tyr Arg Ala Val Val Leu Gly Thr Asp His Gln Ala Phe Asp Arg  
 3290 3295 3300

Ala Leu Gly Thr Leu Thr Leu Gly Glu Pro Ser Pro Ala Leu Val  
 3305 3310 3315

Arg Gly Thr Pro His Pro Gly Lys Thr Ala Phe Met Phe Thr Gly  
 3320 3325 3330

Gln Gly Ala Gln Arg Ala Gly Met Gly Arg Glu Leu Tyr Asp Thr  
 3335 3340 3345

Tyr Pro Val Phe Arg Asp Thr Phe Asp Glu Val Cys Ala Thr Leu  
 3350 3355 3360

Asp Arg His Leu Asn Ala Glu Gln Pro Val Lys Asp Val Val Phe  
 3365 3370 3375

Ala Asp Asp Ala Ile Leu Leu Asn Gln Thr Arg Tyr Thr Gln Ala  
 3380 3385 3390

Ala Leu Phe Ala Ile Glu Thr Ser Leu Tyr Arg Leu Val Glu Ser  
 3395 3400 3405

Phe Gly Ile Thr Pro His Tyr Leu Thr Gly His Ser Ile Gly Glu  
 3410 3415 3420

Ile Thr Ala Ala His Val Ala Gly Ile Phe Thr Leu Asp Asp Ala  
 3425 3430 3435

Cys Arg Leu Val Ala Ala Arg Gly Ser Leu Met Gln Ala Leu Pro  
 3440 3445 3450

Ala Asn Gly Val Met Ile Ser Leu Arg Ala Thr Glu Glu Gln Ile  
 3455 3460 3465

Val Pro Phe Leu Glu Gly His Glu His His Val Ser Val Ala Ala  
 3470 3475 3480

Val Asn Gly Pro Ser Ser Ile Val Ile Ser Gly Asp Gln Glu Ala  
 3485 3490 3495

Thr Thr Ala Ile Ala Gly Ala Leu Ala Glu Thr Gly Val Lys Thr  
 3500 3505 3510

Arg Arg Leu Thr Val Ser His Ala Phe His Ser Pro His Met Asp  
 3515 3520 3525

Pro Met Leu Asp Glu Phe Glu Leu Val Ala Gly Glu Leu Thr Tyr

3530						3535						3540		
His	Ala	Pro	Thr	Ile	Pro	Ile	Val	Ser	Asn	Leu	Thr	Gly	Gln	Leu
	3545					3550					3555			
Ala	Asp	His	Tyr	Ile	Thr	Thr	Pro	Gln	Tyr	Trp	Val	Gln	His	Val
	3560					3565					3570			
Arg	Glu	Ala	Val	Arg	Phe	Ser	Asp	Gly	Ile	Thr	Thr	Leu	Asp	Arg
	3575					3580					3585			
Leu	Gly	Thr	Arg	His	Tyr	Leu	Glu	Leu	Gly	Pro	Asp	Pro	Val	Leu
	3590					3595					3600			
Thr	Thr	Met	Ala	Gln	Asp	Ser	Val	Ala	Asp	Asp	Ser	Asp	Ala	Ala
	3605					3610					3615			
Leu	Val	Ala	Thr	Leu	Tyr	Arg	Asp	Arg	Asp	Glu	Asn	His	Ser	Phe
	3620					3625					3630			
Leu	Thr	Ala	Leu	Ala	Thr	Ala	His	Ala	His	Gly	Ile	Gln	Val	Gly
	3635					3640					3645			
Trp	Thr	Pro	Val	Val	Gly	Glu	Thr	Ser	Val	Pro	Ala	Leu	Gly	Leu
	3650					3655					3660			
Pro	Thr	Tyr	Pro	Phe	Gln	Arg	Gln	His	Tyr	Trp	Leu	Glu	Ala	Ala
	3665					3670					3675			
Lys	Pro	Thr	Ser	Gly	Ala	Asp	Gly	Leu	Gly	Leu	Thr	Ala	Thr	Asp
	3680					3685					3690			
His	Pro	Val	Leu	Thr	Thr	Leu	Ala	Glu	Leu	Pro	Asp	Gly	Gly	Gly
	3695					3700					3705			
His	Leu	Phe	Thr	Gly	Arg	Val	Ser	Gly	Asn	Asp	Pro	Asp	Trp	Val
	3710					3715					3720			
Ala	Glu	His	Ile	Ile	Phe	Gly	Thr	Met	Ile	Val	Pro	Gly	Val	Ala
	3725					3730					3735			
Phe	Val	Asp	Leu	Leu	Leu	His	Ala	Ala	Arg	His	Val	Asp	Cys	Glu
	3740					3745					3750			
His	Ile	Glu	Glu	Leu	Thr	His	His	Val	Phe	Leu	Ala	Val	Pro	Glu
	3755					3760					3765			
Arg	Ala	Ala	Leu	Gln	Leu	Arg	Leu	Leu	Ile	Glu	Pro	Ala	Asp	Ser
	3770					3775					3780			

ES 2 543 203 T3

Ser Gly Ser Arg Ala Phe Ala Phe Tyr Ser Arg Pro Glu Asp Val  
 3785 3790 3795

Pro Val Glu Thr Asp Trp Thr Leu His Ala Thr Gly Ala Leu Gly  
 3800 3805 3810

Ala Glu Arg Arg Glu Val Pro Ala Gly Ala Asp Ser Leu Arg Asn  
 3815 3820 3825

Glu Val Trp Pro Pro Asp Ile Ser Asp Thr Met Asp Val Gly Glu  
 3830 3835 3840

Phe Tyr Arg Arg Val Thr Asp Gly Gly Phe Gly Tyr Gly Pro Leu  
 3845 3850 3855

Phe Arg Gly Leu Lys Lys Ala Trp Gln Asp Gly Asn Thr Thr Tyr  
 3860 3865 3870

Ala Glu Val Ser Leu Pro Ala Gly Ser Asp Pro Gly Asp Tyr Gly  
 3875 3880 3885

Ile His Pro Gly Leu Leu Asp Ser Ala Leu Gln Pro Ala Ala Leu  
 3890 3895 3900

Ile Met Gly Glu Thr Glu Ala Ala Asp Ser Ile Arg Val Pro Phe  
 3905 3910 3915

Ser Trp Ala Gly Val Ser Leu His Ala Thr Gly Ala Asp Ser Leu  
 3920 3925 3930

Arg Ile Arg His Thr Trp Thr Thr Pro Asp Thr Ala Ser Leu Val  
 3935 3940 3945

Ile Ala Asp Gln Thr Gly Thr Pro Val Met Thr Ile Asp Ser Leu  
 3950 3955 3960

Ala Met Arg Thr Val Gly Ala Asp Gln Leu Ala Ala Thr Arg Ala  
 3965 3970 3975

Ala Asp Ala Gly Glu Leu Tyr Lys Val Asp Trp Phe Asp Val Gln  
 3980 3985 3990

Thr Val Glu Asp Lys Thr Gln Gly Ala Gly Thr Ala Lys Trp Ala  
 3995 4000 4005

Val Val Ala Asp Pro Gly Asn Thr Gln Val Ala Ala Ala Leu Ser  
 4010 4015 4020

Pro Leu Gly Ala Ala Val Glu Val Glu Pro Asp Ala Val Thr Leu

ES 2 543 203 T3

4025						4030										4035
Pro	Thr	Thr	Pro	Gly	Asp	Asp	Thr	Thr	Arg	Pro	Asp	Val	Val	Phe		
4040						4045					4050					
Thr	Trp	Cys	Val	Ser	Glu	Pro	Gly	Ala	Asp	Pro	Ala	Gln	Ala	Ala		
4055						4060					4065					
Arg	Ser	Phe	Thr	His	Arg	Val	Leu	Gly	Leu	Val	Gln	Ala	Val	Leu		
4070						4075					4080					
Ser	Asp	Asp	Arg	Pro	Asp	Ser	Arg	Leu	Val	Ile	Leu	Thr	Lys	Gly		
4085						4090					4095					
Ala	Met	Ser	Ala	Gly	Ser	Gly	Gly	Ala	Ala	Asp	Leu	Ala	Gly	Ala		
4100						4105					4110					
Ala	Val	Trp	Gly	Leu	Ile	Arg	Thr	Ala	Gln	Thr	Glu	His	Pro	Asp		
4115						4120					4125					
Arg	Phe	Ile	Leu	Met	Asp	Leu	Asp	Gly	Ser	Asp	Ala	Ser	Leu	Arg		
4130						4135					4140					
Ala	Val	Gly	Ala	Ala	Leu	Asn	Ala	Gly	Glu	Pro	Gln	Leu	Ala	Val		
4145						4150					4155					
Arg	Asp	Gly	Arg	Leu	Leu	Ala	Pro	Arg	Leu	Ala	Arg	Ile	Gly	Thr		
4160						4165					4170					
Ala	Asp	Ser	Glu	Pro	Thr	Ala	Ala	Pro	Ala	Ser	Phe	Asp	Pro	Asp		
4175						4180					4185					
Lys	Thr	Val	Leu	Ile	Thr	Gly	Gly	Thr	Gly	Ala	Leu	Gly	Thr	Leu		
4190						4195					4200					
Leu	Ala	Arg	His	Leu	Val	Thr	Arg	His	Gly	Val	Lys	Lys	Leu	Leu		
4205						4210					4215					
Leu	Thr	Ser	Arg	Arg	Gly	Arg	Pro	Ala	Gly	Ser	Thr	Ile	Thr	Ala		
4220						4225					4230					
Glu	Leu	Ala	Glu	Leu	Gly	Ala	Glu	Val	Thr	Ile	Val	Ala	Cys	Asp		
4235						4240					4245					
Ala	Ala	Asp	Arg	Glu	Ser	Leu	Glu	Ala	Leu	Leu	Ala	Ser	Leu	Pro		
4250						4255					4260					
Asp	Glu	His	Pro	Leu	Gly	Ala	Val	Val	His	Cys	Ala	Gly	Thr	Leu		
4265						4270					4275					

Asp Asp Gly Ile Val Thr Ala Leu Thr Pro Asp Arg Phe Asp Glu  
 4280 4285 4290  
 Val Leu Arg Pro Lys Val Asp Gly Ala Trp Asn Leu His Glu Leu  
 4295 4300 4305  
 Thr Arg Asp Leu Asp Leu Asp Ala Phe Val Thr Phe Ser Ser Val  
 4310 4315 4320  
 Val Gly Val Leu Gly Ser Pro Gly Gln Ser Asn Tyr Ala Ala Ala  
 4325 4330 4335  
 Asn Val Phe Leu Asp Glu Leu Ala Glu His Arg Arg Thr Ala Gly  
 4340 4345 4350  
 Leu Pro Ala Lys Ser Leu Ala Trp Gly Leu Trp Glu Ser Gly Met  
 4355 4360 4365  
 Ala Asp Thr Leu Asp Glu Gln Asp Gln Ala Arg Met Asn Arg Gly  
 4370 4375 4380  
 Gly Leu Leu Pro Met Pro Ala Glu Gln Ala Leu Gly His Phe Asp  
 4385 4390 4395  
 Ser Ala Leu Ala Thr Asp Gln Thr Val Val Val Pro Ala Lys Leu  
 4400 4405 4410  
 Asp Leu Ala Gly Leu Arg Ala Arg Ala Ala Thr Val Pro Val Ala  
 4415 4420 4425  
 Pro Ile Phe Arg Gly Leu Val Arg Thr Pro Leu Arg Ser Ala Ala  
 4430 4435 4440  
 Gln Ala Gly Gly Ala Gly Ala Glu Val Gly Ala Leu Gly Gln Ser  
 4445 4450 4455  
 Ile Ala Gly Arg Pro Glu Ala Glu Gln Asp Gln Ile Ile Leu Asp  
 4460 4465 4470  
 Phe Leu Arg Asn His Val Ala Thr Val Leu Gly His Gly Ser Ala  
 4475 4480 4485  
 Asn Ala Ile Asp Pro Ala His Ser Phe Lys Glu Leu Gly Phe Asp  
 4490 4495 4500  
 Ser Leu Ser Ser Val Glu Leu Arg Asn Ser Leu Asn Lys Ala Ser  
 4505 4510 4515  
 Gly Met Arg Leu Pro Ser Thr Leu Leu Phe Asp Tyr Pro Thr Pro



	4520					4525								4530
Ser	Val 4535	Leu	Ala	Gly	Tyr	Ile 4540	Arg	Asn	Gln	Leu	Ala 4545	Gly	Gly	Lys
Gln	Ala 4550	Glu	Ala	Gly	Ala	Gln 4555	Val	Ala	Arg	Arg	Thr 4560	Val	Arg	Pro
Ala	Ser 4565	Ser	Arg	Ser	Asp	Ala 4570	Ala	Asp	Pro	Ile	Val 4575	Ile	Val	Gly
Met	Gly 4580	Cys	Arg	Phe	Pro	Gly 4585	Gly	Ala	Asp	Thr	Pro 4590	Glu	Ala	Leu
Trp	Lys 4595	Leu	Val	Ala	Asp	Glu 4600	Arg	Asp	Ala	Val	Gly 4605	Ala	Phe	Pro
Asp	Asn 4610	Arg	Gly	Trp	Asp	Ile 4615	Glu	Asn	Leu	Phe	Asp 4620	Asp	Asp	Pro
Asp	Val 4625	Arg	Gly	Lys	Ser	Tyr 4630	Ala	Ser	Glu	Gly	Gly 4635	Phe	Leu	Tyr
Asp	Ala 4640	Asp	Arg	Phe	Asp	Pro 4645	Glu	Phe	Phe	Gly	Ile 4650	Ser	Pro	Arg
Glu	Ala 4655	Leu	Ala	Leu	Asp	Pro 4660	Gln	Gln	Arg	Leu	Leu 4665	Leu	Glu	Thr
Thr	Trp 4670	Glu	Ala	Phe	Glu	Asn 4675	Ala	Gly	Ile	Arg	Pro 4680	Asp	Thr	Leu
Arg	Gly 4685	Lys	Pro	Val	Gly	Val 4690	Phe	Ala	Gly	Val	Ala 4695	Ala	Gly	Glu
Tyr	Val 4700	Ser	Leu	Thr	His	His 4705	Gly	Gly	Glu	Pro	Val 4710	Glu	Gly	Tyr
Leu	Leu 4715	Thr	Gly	Thr	Thr	Ala 4720	Ser	Val	Ala	Ser	Gly 4725	Arg	Ile	Ser
Tyr	Thr 4730	Leu	Gly	Leu	Glu	Gly 4735	Pro	Ala	Val	Thr	Val 4740	Asp	Thr	Ala
Cys	Ser 4745	Ser	Ser	Leu	Val	Ala 4750	Met	His	Leu	Ala	Cys 4755	Gln	Ser	Leu
Arg	Asn 4760	Asn	Glu	Ser	Thr	Met 4765	Ala	Leu	Ala	Gly	Gly 4770	Ala	Thr	Ile

Met Ser Asn Ala Gly Met Phe Met Glu Phe Ser Arg Gln Arg Gly  
 4775 4780 4785

Leu Ala Pro Asp Ser Arg Ala Lys Ser Tyr Ala Gly Ala Ala Asp  
 4790 4795 4800

Gly Thr Ile Trp Ala Glu Gly Ala Gly Met Val Leu Leu Glu Arg  
 4805 4810 4815

Leu Ser Asp Ala Lys Ala Asn Gly His Thr Val Leu Ala Val Ile  
 4820 4825 4830

Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr  
 4835 4840 4845

Ala Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Asn Thr Ala Leu  
 4850 4855 4860

Ala Ser Ala Gly Leu Thr Pro Asp Gln Val Asp Ala Val Glu Gly  
 4865 4870 4875

His Gly Thr Gly Thr Pro Leu Gly Asp Pro Ile Glu Ala Gln Ala  
 4880 4885 4890

Leu Leu Ser Thr Tyr Gly Gln Asn Arg Glu Glu Pro Leu Trp Leu  
 4895 4900 4905

Gly Ser Phe Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly  
 4910 4915 4920

Val Gly Gly Val Ile Lys Met Ile Gln Ala Met Arg His Gly Thr  
 4925 4930 4935

Leu Pro Arg Thr Leu His Val Asp Glu Pro Ser Pro Asn Ile Asp  
 4940 4945 4950

Trp Asp Ser Gly Asn Val Arg Leu Leu Thr Glu Ala Arg Ala Trp  
 4955 4960 4965

Pro Glu Thr Asp Arg Pro Arg Arg Ser Ala Val Ser Ser Phe Gly  
 4970 4975 4980

Ile Ser Gly Thr Asn Ala His Leu Ile Leu Glu Glu Ala Pro Thr  
 4985 4990 4995

Pro Thr His Pro Glu Pro Ala Pro Glu Ser Ala Pro Gln Ala Thr  
 5000 5005 5010

Thr Val Pro Trp Ile Leu Ser Gly Lys Ser Glu Gln Ala Val Arg

ES 2 543 203 T3

5015						5020						5025		
Asp	Gln	Ala	Gln	Arg	Leu	Leu	Asp	His	Val	Ser	Glu	Tyr	Pro	Glu
5030						5035					5040			
Leu	Gln	Pro	Val	Asp	Ile	Ala	Tyr	Ser	Leu	Ala	Thr	Ala	Arg	Thr
5045						5050					5055			
Ser	Phe	Glu	Arg	Gln	Ala	Val	Ala	Ile	Gly	Ala	Thr	His	Asp	Glu
5060						5065					5070			
Leu	Val	Asp	His	Leu	Arg	Ser	Leu	Thr	Gln	Asp	Pro	Gly	Thr	Ala
5075						5080					5085			
Leu	Leu	His	Gly	Gln	Ser	His	Ser	Lys	Lys	Val	Ala	Leu	Leu	Phe
5090						5095					5100			
Thr	Gly	Gln	Gly	Ser	Gln	His	Pro	Gly	Met	Gly	Arg	Gln	Leu	Tyr
5105						5110					5115			
Asp	Thr	His	Pro	Val	Tyr	Arg	Asp	Ala	Phe	Asp	Glu	Val	Thr	Ala
5120						5125					5130			
Thr	Leu	Asp	Gln	His	Leu	Gln	Ala	Glu	Gln	Pro	Val	Lys	Asp	Val
5135						5140					5145			
Val	Phe	Ala	Asp	Asp	Pro	Thr	Leu	Leu	Asn	Gln	Thr	Arg	Tyr	Thr
5150						5155					5160			
Gln	Pro	Ala	Ile	Phe	Ala	Leu	Gln	Val	Ala	Leu	Thr	Arg	Leu	Leu
5165						5170					5175			
Val	Asp	Glu	Phe	Gly	Val	Ser	Pro	Thr	His	Leu	Ile	Gly	His	Ser
5180						5185					5190			
Ile	Gly	Glu	Ile	Ser	Ala	Ala	His	Thr	Ala	Gly	Ile	Leu	Thr	Leu
5195						5200					5205			
Asp	Asp	Ala	Cys	Arg	Leu	Val	Ala	Ala	Arg	Gly	Thr	Leu	Met	Gln
5210						5215					5220			
Thr	Leu	Pro	Ala	Thr	Gly	Ala	Met	Thr	Ala	Val	Glu	Ala	Thr	Glu
5225						5230					5235			
Glu	Glu	Val	Leu	Pro	His	Leu	Thr	Glu	Arg	Val	Gly	Ile	Ala	Ala
5240						5245					5250			
Val	Asn	Gly	Pro	Arg	Ser	Val	Val	Val	Ser	Gly	Asp	Glu	Ala	Ala
5255						5260					5265			

Val Ile Ala Val Gly Glu Glu Phe Ala Gly Gln Gly Arg Arg Ile  
 5270 5275 5280

Arg Arg Leu Thr Val Ser His Ala Phe His Ser His His Met Asp  
 5285 5290 5295

Pro Met Leu Gly Glu Leu His Ala Val Ala Asp Thr Leu Thr Tyr  
 5300 5305 5310

His Val Pro Arg Thr Pro Leu Val Ser Thr Val Thr Gly Arg Leu  
 5315 5320 5325

Ala Gly Ser Glu Ile Thr Ser Ala Thr Tyr Trp Ser Asp His Ala  
 5330 5335 5340

Arg Asn Ala Thr Arg Phe His Asp Gly Leu Asn Thr Leu His Glu  
 5345 5350 5355

Gln Gly Val Thr Thr Tyr Ile Glu Val Gly Pro Asp Ala Val Leu  
 5360 5365 5370

Ala Ala Leu Thr Arg Glu Ala Leu Pro Asp Ala Thr Ala Val Pro  
 5375 5380 5385

Leu Ile Arg Ala Lys Ala Ser Glu Pro Ala Thr Leu Leu Asp Gly  
 5390 5395 5400

Leu Val Arg Ala His Val Ser Gly Ala Thr Val Asp Trp Ala Gly  
 5405 5410 5415

Phe Leu Ala Arg Arg Gly Gly Arg Ser Val Asp Leu Pro Thr Tyr  
 5420 5425 5430

Ala Phe Gln Arg Arg Arg His Trp Leu Glu Thr Ala Asp Pro Val  
 5435 5440 5445

Gly Thr Ala Ala Gly Leu Gly Leu Glu Ser Ala Ser His Pro Leu  
 5450 5455 5460

Leu Ala Thr Thr Thr Glu Leu Pro Asp Gly Thr Ala Leu Phe Thr  
 5465 5470 5475

Gly Arg Val Thr Leu Ala Asp His Pro Trp Leu Ser Asp His Thr  
 5480 5485 5490

Val Met Gly Thr Val Ile Leu Pro Gly Thr Ala Phe Val Glu Leu  
 5495 5500 5505

Ala Leu His Ala Ala Glu Thr Val Gly Leu Asp Glu Ile Ala Glu

5510						5515						5520		
Leu	Val	Leu	His	Ala	Pro	Val	Thr	Phe	Gly	Ser	Gln	Ser	Ala	Ala
5525						5530					5535			
Leu	Leu	Gln	Val	Ile	Val	Gly	Pro	Asp	Asp	Pro	Ser	Ala	Gly	Arg
5540						5545					5550			
Thr	Leu	Thr	Ile	Arg	Ser	Arg	Ser	Glu	Glu	Asp	Gln	Ser	Trp	Thr
5555						5560					5565			
Glu	Asn	Ala	Thr	Gly	Thr	Leu	Gly	Ala	Leu	Val	Gly	Val	Ser	
5570						5575					5580			

<210> 6  
 <211> 1089  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13  
 <400> 6

5

```

atgagaactg ttgtcgtcgg cggtggtgtc atcggcctgg ccaccgcatg gcgggcccgcg
60
cgcgaaggcg tctccgtcac ggtgatcgac cctgccccgg gcagcaaggc gtcccacgcc
120
tcggcgggac tgctcccggc catcaacgac cagctgtacg accagccgga actcctgcgc
180
ctgtgcctgg cctcccgtga gctgtaccg tcttctcgtc aggaactgga ggagttcgct
240
ccggccggct tccggcgcga cggcgtcctg gacgcggcgt tcgacgagga atccctgccg
300
atcctcgacc gccaaactggc cttccagaag tccatcggtg tgcgcaccga acggctgacc
360
gccgaggagt gcgccaaagt ccagccggcc ttcgcccccg tggtcggcgg cctgctgtcc
420
ccggacgacg gcgccatcga cccgcgcggt ctcgacgcgg cgctgatcac cgccatcgaa
480
gccctgggcg gcagcgtcgt acgcagggga gccacccgga tcgaggacca caccgtcgtc
540
ctcgacaccg gtgacaaggt gcccttcgac cgtctcgtcc tggccgcccg ctgctggacc
600
caccagatcg aggggctgcc cgcgggdcgc atcccggaga tccgtcccgt caagggccag
660
atcctccgcc tgcgctccga cgtgccactg ctcaacgtga cggcccgcgc catctccaag
720
ggcaagtccc tctacctggc tccccgcctg gacgggtaac tcgtcgtcgg cgccacctac
780
gaggagcgcg gctacgacga gaccgtcacc gccgagggca ccggcagcct gctgcgccgc
840
  
```

10

gcccgcgagg tcattcccga ggtgggcgcc ctgcgtttcg ccgagatcac cgcggggctc  
900

cgccccgcct cgcgccgacga cctccccgtc atgggtccga ccacggttcc cgatgtctac  
960

ctggcgagcg gtcacttccg gatgggagtc cagctggctc ccgtcacggc cgtggccatg  
1020

gcggcctacc tgaccgacac cgccccgcac gccgcgacgg cgccgttcac cccgctccgc  
1080

ttctcctga  
1089

<210> 7  
<211> 362  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 7

Met Arg Thr Val Val Val Gly Gly Gly Val Ile Gly Leu Ala Thr Ala  
1 5 10 15

Trp Arg Ala Ala Arg Glu Gly Val Ser Val Thr Val Ile Asp Pro Ala  
20 25 30

Pro Gly Ser Lys Ala Ser His Ala Ser Ala Gly Leu Leu Pro Ala Ile  
35 40 45

Asn Asp Gln Leu Tyr Asp Gln Pro Glu Leu Leu Arg Leu Cys Leu Ala  
50 55 60

Ser Arg Glu Leu Tyr Pro Ser Phe Val Glu Glu Leu Glu Glu Phe Ala  
65 70 75 80

Pro Ala Gly Phe Arg Arg Asp Gly Val Leu Asp Ala Ala Phe Asp Glu  
85 90 95

Glu Ser Leu Pro Ile Leu Asp Arg Gln Leu Ala Phe Gln Lys Ser Ile  
100 105 110

Gly Val Arg Thr Glu Arg Leu Thr Ala Glu Glu Cys Ala Lys Leu Gln  
115 120 125

Pro Ala Phe Ala Pro Val Val Gly Gly Leu Leu Ser Pro Asp Asp Gly  
130 135 140

Ala Ile Asp Pro Arg Val Leu Asp Ala Ala Leu Ile Thr Ala Ile Glu  
145 150 155 160

Ala Leu Gly Gly Ser Val Val Arg Arg Gly Ala Thr Arg Ile Glu Asp  
165 170 175

10

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

<210> 8  
 <211> 2085  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

<400> 8

atgagtaaag cagaacggcc cgggcttctt gacggagtcc gcgaggtcga catcgagcac  
 60  
 ctctacagcc gcttcgccga ccgcgggtctc cagtacgggc cggccttccg gggcctgcgc  
 120  
 gccgtgtggt cgcacggcga ggaggtctac gccgactcag cgctcgacac cgcgaccggc  
 180

5

10

ggcgactacc tcctgcaccc ggccctgctc gacaccgcgc tccaagcggc actgggtcccc  
 240  
 gacatcgacc gcgacgaccg gacgttcctg ccgttcgcgc tgcgcgggat acgggtgcac  
 300  
 aagggcggcg cccgcgccgt ccgcatccac accgtccccg gcgacgacgg cttctccctg  
 360  
 gcgctcaccg gcgacgacgg cgagccgatc gccacgatcg gttccgctgt cagccggccc  
 420  
 gtcacggccc agcagctcga cgccgcggcg cagcgcaccc aactgctgcg cgtcgtatgg  
 480  
 aagtccgctg tacagcagtc ggacaactcc gaccagcagc gctggggatt cctcggaaac  
 540  
 gaccgcatcg gcctcaccgg cgcgctgaag gccacccgcc ccttgttcga ctccatcccc  
 600  
 acgctgcgcg aactcgactc cgtgctccgg gccgccacag ccgtcccggg cgtgatcgtc  
 660  
 gtctcctgca cggacgagga ctgcgccgta cgctcggcgg cgcaacgcgc cctcatggtc  
 720  
 gtacaggagt gcctggccga ccaccgcctc gccaaagacc gcctggtgct ggtcagtagc  
 780  
 ggagccgctg ccgcccgcgc cggggaggac ctgtccgacg tgtcgggcgc cgccgtctgg  
 840  
 ggactgctcc gcagcgtcca gtccgaacac cccgaccgct tcgtcctggt cgacgtcgac  
 900  
 gacccccgga actcggggccg ctccctcgcc gccgccgctg cctccggcga accccaactc  
 960  
 gccgtgcgca acggcgcgct gctcaggccg cgtctcgtac gcagcccacc gccgccgcgg  
 1020  
 cgcaggagcc tcacgggtac cgtcgtgatc accggcggca ccggagaact gggccgcctg  
 1080  
 ctgcgccggc acctcgtcac cggccatgac gtccggcacc tcgtgctgct cagccgccgc  
 1140  
 ggtcccggct cgcccggcgc cgccgagctg gatgccgaac tcaccgcgct cggcgcccgt  
 1200  
 gtcgacgtgg tcgcctgcga cgtcgcggac cgctcgtcgc tggagagcgc actggccggc  
 1260  
 atccccgccc cgtccgccgt catccacacc gcgggtgtgc tgtccgacgg cgcgatcggc  
 1320  
 accctcacgc cacgcggact cgacaagggt ctgctccca aggtcgcgc cgcgctccac  
 1380  
 ctgcacgacc tgatccagga cccggactgc gcgttcgtgg tgttctcgtc cgtcgcgggg  
 1440  
 ctggtcggca acgcggggca gggcaactac gcagccgcca acgcccctct cgacgcgctc  
 1500



gcccaccacc gccgtgcccg cagactgcag ggcctctcgc tcgcctgggg tctgtgggag  
 1560  
 agcgagaacg gcatgggctc cgacctgtcc gccgccgacc acaaccgcat caagcggctc  
 1620  
 ggcttcgccc ccctggggca cgaccagggc ctggctctct tcgacgccac gctcggcagc  
 1680  
 gacgaggcgg tcctggcgcc cgtccggctc aacgaggcgg gcctcaccgg ggacatcccc  
 1740  
 cccgtcctgg aggagctggc gcccacccgt accggcaagc ccgccgtgac cgacaccctg  
 1800  
 gtcagccggc tggccgagct gcccgaggcc gaacgcgacg ccgccgccct ggagtctgtc  
 1860  
 cgctcgggtg ccgccttggc cttcggctac gagagcggcg acgagatcga cccgcagcgg  
 1920  
 gagttcagcg ccgccgggct cgactccatc ggcaacctcg aactcagccg ccacctggcg  
 1980  
 gccgccaccg gcctgcggct cccggcgacc ctgctcttcg accatcccac ccccgagaa  
 2040  
 ctcgcctccc acctgcgtcg actcctccag gagagcaatt catga  
 2085

<210> 9  
 <211> 694  
 <212> PRT  
 <213> *Streptomyces* sp. MP28-13  
 <400> 9

5

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

10

Pro Gly Asp Asp Gly Phe Ser Leu Ala Leu Thr Gly Asp Asp Gly Glu  
 115 120 125  
 Pro Ile Ala Thr Ile Gly Ser Val Val Ser Arg Pro Val Thr Ala Glu  
 130 135 140  
 Gln Leu Asp Ala Ala Ala Gln Arg Thr Gln Leu Leu Arg Val Val Trp  
 145 150 155 160  
 Lys Ser Val Val Gln Gln Ser Asp Asn Ser Asp Gln Gln Arg Trp Gly  
 165 170 175  
 Phe Leu Gly Thr Asp Arg Ile Gly Leu Thr Gly Ala Leu Lys Ala Thr  
 180 185 190  
 Arg Pro Leu Phe Asp Ser Tyr Pro Thr Leu Arg Glu Leu Asp Ser Val  
 195 200 205  
 Leu Arg Ala Ala Thr Ala Val Pro Asp Val Ile Val Val Ser Cys Thr  
 210 215 220  
 Asp Glu Asp Ser Pro Val Arg Ser Ala Ala Gln Arg Ala Leu Met Val  
 225 230 235 240  
 Val Gln Glu Cys Leu Ala Asp His Arg Leu Ala Lys Thr Arg Leu Val  
 245 250 255  
 Leu Val Ser Ser Gly Ala Val Ala Ala Arg Ala Gly Glu Asp Leu Ser  
 260 265 270  
 Asp Val Ser Gly Ala Ala Val Trp Gly Leu Leu Arg Ser Val Gln Ser  
 275 280 285  
 Glu His Pro Asp Arg Phe Val Leu Val Asp Val Asp Asp Pro Gly Asn  
 290 295 300  
 Ser Gly Arg Ser Leu Ala Ala Ala Val Ala Ser Gly Glu Pro Gln Leu  
 305 310 315 320  
 Ala Val Arg Asn Gly Ala Leu Leu Arg Pro Arg Leu Val Arg Ser Pro  
 325 330 335  
 Pro Pro Pro Arg Arg Arg Ser Leu Thr Gly Thr Val Val Ile Thr Gly  
 340 345 350  
 Gly Thr Gly Glu Leu Gly Arg Leu Leu Ala Arg His Leu Val Thr Gly  
 355 360 365  
 His Asp Val Arg His Leu Val Leu Leu Ser Arg Arg Gly Pro Gly Ser  
 370 375 380

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

Glu Phe Ser Ala Ala Gly Leu Asp Ser Ile Gly Asn Leu Glu Leu Ser  
 645 650 655

Arg His Leu Ala Ala Ala Thr Gly Leu Arg Leu Pro Ala Thr Leu Val  
 660 665 670

Phe Asp His Pro Thr Pro Ala Glu Leu Ala Ser His Leu Arg Arg Leu  
 675 680 685

Leu Gln Glu Ser Asn Ser  
 690

5

<210> 10  
 <211> 564  
 <212> ADN  
 <213> *Streptomyces sp.* MP28-13  
 <400> 10

atgaccagc ttctcgaaga caccgaccgg cgcgtcttcc ccaccgacga cgaactgctg  
 60

caccgggtgc tgacccccta ccgggccaag cgggtgcgagt acctcacctc ggccacggtc  
 120

acctcggagg gcgacccccg cgacggcgga cgactcatcg ccacctgcac cttcgagatc  
 180

cccgagtcct gctacatcga cgacaccggg cacttcaact cggtcgagtt caacatctgc  
 240

ttcaaccaga tggcctaacta cctgctggcc atgtcggtae gggagtcact cgtcagagccg  
 300

ttctccggct ggacgatcga gcagttctgg acccggcagc tcgccgacgt gttcatcacc  
 360

gacttcaaga gcagcttccg cagcgcgatg cagggccgcc gcttcaccgg cgagatcgag  
 420

atcatcgaca tcgccgaatg ggacgccaac gacctgcgcg acgccctggt gatcctgcgg  
 480

accaagtgcc actacgccga cgagcaaggt ggcgagagcc acggcgagat caccgccgcc  
 540

gtcaccaacc cgcccgtgat ctga  
 564

10

15

<210> 11  
 <211> 187  
 <212> PRT  
 <213> *Streptomyces sp.* MP28-13  
 <400> 11

ES 2 543 203 T3

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

<210> 12  
 <211> 10584  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13  
 <400> 12

5

ES 2 543 203 T3

atgaccgagc agtcaagggc cgcgatgctg gaactggtgc tggcgcaagc cgcagcggtg  
60

ttgcggaccg ccgacccgga cgcctacggg gaaggggccc tcgacctggc cgccgaccgc  
120

ccgttcctgg cggggggcct ggactcgctg ggcctggtcc ggttgcagcg gcggctcgcc  
180

gccgaggtgg gggcggacct gcccgtcacc cttgccttcg accacccgac accgatcgcc  
240

ctcggccggc acctggccga cctgctgcac ggcaccgccc gagcggacgc acccgacgag  
300

acggcggcgc ccgccctgcc gtcggcgttc ggcgagccga tcgcgatcgt cggaatcggc  
360

tgccgcttcc ccggtggtgt gagcacgccc gaggacctgt ggcggatcgt cgccgacgac  
 420  
 acccatgtgc tcacggactt tccctccgac cggggctggg atctcgaccg catctacaac  
 480  
 ccggaccegt ccgtcaccgg cgccagttat gtgcggcgcg gcggattcct cccggacgcc  
 540  
 gccgacttcg acgccgactt cttccagatc agtcccaagg aagcgtggc catggacccc  
 600  
 cagcagcggc tgctgctgga gacctcctgg gaagcgtctg aacacgcggg catcgacccc  
 660  
 ggccgcctgc gcggcacacc ctcgggcgtg ttcacggtcg tggagccgca cgagtacggg  
 720  
 ccccgacgc acgaggcccc cgacggtctc gacggctacg tcctgggcgg caacctgccc  
 780  
 agcgtcgtgt ccggccgctg cgcctacacc ctcggcttcg aaggccccac actcaccgtc  
 840  
 gacaccgcgt gctcgggctc gtcacggcc ctccacctgg ccgtgcgctc gctccagggc  
 900  
 ggggagtgcg ggctggcgtt ggccggcggc gtcaccgtga tatccagccc cggcacgttc  
 960  
 accaccttca gccgccaacg cggctctggcc cccgacggac gcatcaaggc gttcgccgcc  
 1020  
 gcggcggacg gcacctcgtt cgccgaaggc gtcggcgtgt tcgtcctcgc ccgctgtcc  
 1080  
 gacgcgtgc gcgatgggca ccccgactg gccgtgatcc gcggcaccgc catcaaccag  
 1140  
 gacggcgcga ccaacggcct gtccgcccc aacggcctcg cgcagcagcg cgtcatccgc  
 1200  
 cgcgcgctca ccgacgcccg tctgaccgcc gacgaggtcg acgcggtcga ggcgcacggc  
 1260  
 accggcacca cactcggcga ccccatcgag ggccaggccc tggtcgccgc ctacggacgc  
 1320  
 ggacgctcac ccgagaagcc cctgtggctg gggtcgggtga agtccaacat cggccacacc  
 1380  
 ggagcggcgg cgggcgcccg cggaatcatc aagatggtgc aggcgatgcg tcacagcacg  
 1440  
 ctgccgcgca ccctgcacgt ggacgcgccc accacgcacg tcgactggtc ggacggcacc  
 1500  
 gtccagctcc tcaccgagcc cgtcccctgg gaagccgcg acacaccgcg ccgtgccgga  
 1560  
 atctcctcgt tcggcgcgag tggcaccaac gccacgtca tcacgagga gccgcccgcc  
 1620  
 cccgtcggc ccgcccacgg gagegacgaa cccccggtgg ccgagcggcc cgtgccggtc  
 1680

gtcctgtccg cccagggcca ggacgcactg cgcggccagg ccgaacggct cctgacggcg  
 1740  
 gtcgacgacg cctcgccgct ggacctcgcc tactcggccg ccaccacgcg cgccgcgctg  
 1800  
 cgcgaccgcg ccaccgtcgt ggccgcccac cgtgccgaac tgcgcccgcg cctgaccgcc  
 1860  
 ctcgcccggg gcgagagcgc gcccggactg ctcaccggta cgacgcccggc cgcgggcccgt  
 1920  
 accgcgttcc tcttcaccgg ccagggcagc cagcgtcttg gcatggggccg cgaactggtg  
 1980  
 cgcgccttcc cggtgttcgc gcgggcccgc caggacgccg ccgaccacct cgacctccac  
 2040  
 ctcgacgagc ccctggggca ggtgctgttc gccgagcccg gttcggcgca ggccgaactg  
 2100  
 ctccagcaga cccgctacgc gcaggcccgc ctcttcggcg tcgagaccgc catgttccgg  
 2160  
 ctgctggagt cgtggggcgt cacgcccgga ctctcaccg gacactccat cggcgagatc  
 2220  
 gccgcccgcg acgtcgcccg cgtcatgaac ctggcccagc cggcgtgct ggtcggcgcg  
 2280  
 cgcggcacgc tgatgcagga gcttcccgcc ggccggccca tggtcgccgt ccaggcggcg  
 2340  
 gaggccgaag tcgccccgta tctcaccgag aagggtgggca tcgccgccgt caacggccccg  
 2400  
 ctggcggctg tgctgtccgg tgagaccgac gccgtactcg ccgtcgcccg ccgcttcacc  
 2460  
 gagggaggac gcaagacgcg gcggctgcgc gtctcgcacg cgttccactc gccgctgatg  
 2520  
 gaaccgatgc tcgacgactt ccgccgcgtc gccgagacce tgacctacca gccgccgctg  
 2580  
 atccccgtgg tgtccaacct gaccggcgag cccgtcgccc cgttcgacgc cgactactgg  
 2640  
 gtacgccacg tgcgcgagcc ggtccgggtc gccgacgcca tgagctggct cgaatcccag  
 2700  
 ggcgtcacca cctacctgga actgggcccc gacgccgtcc tgtcggcaat gggccgcgac  
 2760  
 tgcctgaccg acggcggcgc cgacgccgcg ttcgccgccc tgctgcgcga cgggcacgac  
 2820  
 gagggagcgc agagtctcac cgcgctcggc ctggcgcacg cccacggctc cgcggtcgac  
 2880  
 tgggaccgct tcttcgccgg ccgcggcgcc caccgcaccg cgctgcccac ctactcgttc  
 2940  
 cagcgcaggc gctactggct ggacgccgga agcagccacc ccggcacgat cgggcacccc  
 3000



atgctggaca gcgccgtcag cctcgcgggc gccgacggag tgatcctgac gggccggatg  
 3060  
 tccaccggg cacagccctg gctggccgac catgtcatcg cgggtgtcgt cctcgtaccc  
 3120  
 ggcaccgccc tcgtcgaact ggccgtccgc gccggcgacg aggcgggctg cgccgcggtg  
 3180  
 gaggaactca ccctggagac accgctggcg gtcaccaccg acgccggagt cgagatccag  
 3240  
 gtcgtcgtcg gcgcgctcga cgcctcgggc cgccgctcgg tcgacatcta ctcgcccc  
 3300  
 gcgcggagcg aggacaactg gacccgccac gccagcgggg tgctgaccga gcacggcgag  
 3360  
 ggcgcgacgg ccgaccggc cgtgttcgcc cagtggccgc cggccgacgc cgagccgatc  
 3420  
 gacgtcagcg gcctctacga gcgccaggcc gcgcaggggt acggctacgg gcccgcgttc  
 3480  
 cgacgcgtgc gcgccgcgtg gcgccgcggg gacgacgtgt tcgccgaggt cgccctcgac  
 3540  
 gacgccggag gcgaccggtt cggactgcac cccgcgctgc tggactcctc actgcacgcc  
 3600  
 gccgaaggcc cggaggacga cggccagggt cggtgccgt tcgcctggcg cgggtgcgaa  
 3660  
 ctccacgcca ccggcgccac cgccgtacgc gtccacgtcg tccagaccgc cccggacgag  
 3720  
 gtcacggctc aactcgccga cgcgaacggg gcacccgtcg ccaccgtccg ctcgctcgtc  
 3780  
 cagcgccccg tgccgatcaa cggcgtacgg ccgtccgcct cgggcggcgg ctcgctgctg  
 3840  
 cggtcgaat ggaccgccgt ccaggccccg tccgagccgg acgccgcgtc gcacgagttc  
 3900  
 caactggcgt acgtccccga gacctcccc ggcgaaccg ccgacgcggc ccgggcggcg  
 3960  
 aactgcacg ccctaccct catccaggac gcgctgcgcg acgacaccg cctcgcgatc  
 4020  
 gtcaccgcg gcgagccgta cgacttcacc ctcgctccc cgtgggcgct cgtgcgtcc  
 4080  
 gcccaggccg agaaccggc ccgcttcgtc ctcgctggca ccgaccacga cgtgcccga  
 4140  
 cacgaactgc gcgccccct ggccaaccg gaacccaac tcgccctgcg cgacggcaag  
 4200  
 gtactcgtgc cccgcctggc caggacgccc gcacccgagg acccgcgccc cgtcgcgtgg  
 4260  
 gaccggacg gcacggtgct gatcaccggc ggcaccggcg gtctgggcgg cgtggtcggc  
 4320

cgccacctcg tacaacagca cggcgtgcgg cacctgctgc tcaccgggcg acgcggcccc  
 4380  
 gacagccccg gcgccgccga actcgtcgcc gaactggccg gcttgggccc ccacgccacg  
 4440  
 gtcgccgcct gtgatgtcgc cgaccgcgcc gccctggccg cgctgctgga gacggtcccc  
 4500  
 ggcgagcacc cgctgacggg tgtcgtccac agcgcgggag tcgtcgacga cgggctcgcc  
 4560  
 ggatcgctca cgcccagca ggtccacacc gtcttccacg gcaaggccga cggcgcctgg  
 4620  
 catctgcacg agctgacctg cggcctcccg ctcgccgcct tcgtcctggt ctcctccgcg  
 4680  
 gccgggacca tggaggcggc cgggcagggc aactacgccg ccgccaacgc cttcctggac  
 4740  
 gccctcgccg cgcaccgcgt caccagggc ctccccgca cctccctcgc gtggaacctg  
 4800  
 tgggcgggcg acgcgggaat gggcgcgccg ctggacgagg tcaccctccg ccgggcccag  
 4860  
 cgttccgggc tccccgcctt ggacgccgag gagaacctcg ccctgctgga ccaggcgtc  
 4920  
 gtcaccggtg cgccggcact cgtcccgtg cgcgtcgacg cccgcgcgct gcgggcccgc  
 4980  
 tccgaaggca tcccgccgat gctgcggggg ctggtccgcc cgcccgcctg ccggaacacc  
 5040  
 gccgcggcgg ccgcgaccgg ccccggcggc gacttggccg accggetcgc cggaaagccc  
 5100  
 gacgccgaac gcgaacgcat cgtcctcgac ctgggtccgga cccagatcgc cgccgtactc  
 5160  
 ggacacgaca gcggcaccgc gatcgacccc cgcgcgcctt tcaccgagct gggcttcgac  
 5220  
 tcactggccg ccatcgaact gcgcaacgcg ctcggtacgg ccaccgggct gcggctcacc  
 5280  
 tccacgctga tcttcgacca cccgaccccg cgcgccctgg tcgaccacgt gctcgaaacc  
 5340  
 gtacgcggag ccgtacctgt caccgcggcg ccgcggccca ccgtgcggac cgcgaccgac  
 5400  
 gagcccatcg cgatcgtggc gatgggctgc cgctaccccg gcggcgtgac ctccgccgag  
 5460  
 gaactgtggc ggctggctgc cggcggcacc gacgccatca ccgagttccc ggacgaccgg  
 5520  
 gactggcaca ccgacgacat ctacgacccc gagcccggca agcacggcac cacgtacacc  
 5580  
 cgtgagggcg gcttcctgca cgacgtcgcc gagttcgacc ccgccttctt cggcatcagc  
 5640

ccgaaggagg cccagggcat ggacccccag caccgcatgc tcctcgaagt cgcttgggag  
 5700  
 gccctcgaac agggcgcat cgatccgcac tcctgcacg gcaccgccgc cgggtgtgttc  
 5760  
 gccggtgtca tgaaccacga ctggacgacc cgctccggcg ccgtccccga ggacctcgcg  
 5820  
 ggcttcaccg ccggggggcg cctcggcagc atcgctccg gacgcatcgc ctacaccctg  
 5880  
 ggtctccagg gccctgcggt caccatcgac accgcctgct cctcctccct ggtggccatg  
 5940  
 cactgggcga tgcagtcgct gcggcagggc gagtgcacgc tcgccctggc cggcggcgtc  
 6000  
 accgtgatgg ccacgcccga gacgttcgtc gggatgagcc tgcagagcgg cctcgccgcc  
 6060  
 gacggccgct gcaagggcga cggcgccggc gccgacggca ccggctgggg cgagggcgcc  
 6120  
 ggactcctgg tgctcgaacg cctgtcggac gcccggcgca acggtcaccc ggtggtcgcg  
 6180  
 gtgatccgcg gctcggcgat caaccaggac ggcgcgtcca acggcatcgc ggcccccaac  
 6240  
 ggccccgcc agcagcgggt gatccggcag gccgtggcct ccgccggtct gacactcgcc  
 6300  
 gacatcgacg cggtcgaggg ccacggcacc ggcaccacc tcggcgacc gatcgaggca  
 6360  
 caggcgctga tggccaccta cggccaggag cggcgcgacg atccgctgtg gctcggctcg  
 6420  
 gtcaagtcca acctcgggca caccaggcc gccgccggcg tcgccggcgt catcaagatg  
 6480  
 gtgatggcga tgcgccacgg cgtactgccg cgcaccctgc atgccgagac cccctctccg  
 6540  
 cacatcgact ggaccgaggg cgccgtcgaa ctgctcaccg agcccagggga ctggacggcc  
 6600  
 gacggccgcc cgcgccgcgc cgccgtctcc tccttcggcg tcagcggcac caacgcccac  
 6660  
 gtgatcatcg agcaggcgcc gccggccggc ccgggcgaac cgcgcggcga gcggcctccg  
 6720  
 gtcgtcccgc tgaccgtgtc gggctcgacc cccgaggcga tgcgcgcca ggcggcgcgc  
 6780  
 atcgccgcc acctgcgcga gaacggcgac gtcgacgaac tggacgccgc cgccacgctt  
 6840  
 gcccgaggcc gcgccgccct ggaacaccgg gccgtgatcg tggacgccga ccgcgacggc  
 6900  
 ctgctcggcg gactcgacgc gctggccgcc gggaaactcgt ccgccgccgt ggtccagggc  
 6960

ctgcaacgcg gcgactcgcg cgccgtgctg gtcttcccag gacagggctc gcagtggcag  
 7020  
 ggcattggcg tcgaactgct ggagcactca ccggcgttcg cgacacgcct gggcgaatgc  
 7080  
 gccaaaggcac tcgaatcgta cgtcgactgg aacctgctcg acgtgggtcca cggcgcgccc  
 7140  
 ggcgcaccgg ccctggacgc cgtcgacgtc gtccagccca cgctgtgggc gctgatggtg  
 7200  
 tcgctggccg aggcctggcg cgcggccggt gtcgaaccgg ccgccgtcgt cggcactcc  
 7260  
 cagggcgaga tcgccgccgc ctgtgtcgcc ggcgcgctgt cgctggagga cggcgcacgc  
 7320  
 gtggtcgcgc tgcgcagccg cgtcatccgg caggacctgg cgggccgggg cggcatgatg  
 7380  
 tcggtcgcgc tgtccgccga ccgcgcgagt gactacctgg ccgactggga cggccgcctg  
 7440  
 caactcgcgc tcgtcaacgg cccgagctcc gtcgtcgtgt gcggcggccc ggaggccctc  
 7500  
 gacgaactgc ggggccggct ggacaccgac gaggtccagt cccgccgtat cccggtggac  
 7560  
 tacgctccc actcgatggt cgtggaggag atccgcgacc ggctgctcac cgaactgagc  
 7620  
 ggcctgacac cccgtacctc gtccatcccg ttctactcca ccgtcaccgc gggaaacctg  
 7680  
 gacaccgccc gtctcgacgc cgagtactgg tacaccaacc tgcgccagac cgtccggttc  
 7740  
 gaggacacga cgcgggcact gctcggccgac ggettcgaga ccttcacga ggcagccccg  
 7800  
 caccggggcc tgctgaccgg actcgacgag accgccgagt cggccggggg ggcgccaca  
 7860  
 ctcgctggca cactgcgccg cgactccggc agcccgcgcc agttcgtcac ctcgctggcc  
 7920  
 gaggcgtacg tgcggggcgc gaccgtcgac tgggacacc agttcgtcgg aaccggggcc  
 7980  
 cagcacgtcg aactgeccac ctacccttc cagcgcaagc ggtactggct gcacctgtcc  
 8040  
 gagcacaccg gcgacgccgt gggcatcggc cagatacca ccgaccacc gctgctcggg  
 8100  
 gcggccgtgc cggtcgccgg ggccggggga gtactgctca ccggacgcct ctgctctcc  
 8160  
 ggccagccat ggctcgccga ccacgtcgtc ggcggcaccg tcctgttccc cgggtcaggc  
 8220  
 ttcgtggaac tggccgtccg ggccgggtgac gaggtcggcc ggggccgcgt ggaagaactg  
 8280

acactcgaag caccgctggc tctccccgag cgcggcggcg tggccatcca ggtcgtggtc  
 8340  
 gacgccgatc tgcgtaccgt gtcgatccac tcgcggccccg acaacgcccc cgccgatacc  
 8400  
 ctctggacac gccacgcccc gggcacgctc accgacgccc cgccccgaacc ggccccacgag  
 8460  
 ccggcggcct ggccgcccgc gggcggcggag cccgtcgacc tcaccggctt ctacgagccg  
 8520  
 gccgccgacg gcggcctcgc ctacggggccg gtgttccagg gtttgctgac cgcctggcgc  
 8580  
 tccggcgacg aggtgttcgc ggagatcacg ctccccgagc aggccgccgc cgaggcccag  
 8640  
 cggttcgggc tgcacccggc cgcactcgac gccgctctcc acgcgaccgg actgctcgcc  
 8700  
 accgacgccc aaagggtcac cctgccgttc gcctggacac gggtcgacct gcacgcctcg  
 8760  
 ggagcccggc cgctcagact gcggatgacc agcctgggtg atgacgaggt ggcgctgcgc  
 8820  
 ctaccgaca ccgcggggccg ccccgtcgcc tccgtcgagt cgctcgact gcgtccgggt  
 8880  
 gccccgggcg ggccgatccg caccggcgcg tacgacgact cgctgttcga gctgggtctgg  
 8940  
 gcacccgccc caccggtacc tggcggcacc gccccccgga ccgtcgtgca ccaactgcacg  
 9000  
 ggccggcaca ccgcccgcgc ggcccacgcc gagaccgcca cgacactcgc cgtcctccag  
 9060  
 tcctggctgg agagcggcgc cgacgacgcc gtgctcgccc tcgtcaccgg cggcggcctg  
 9120  
 tccgtgaacg gcgaggacgt caccgacctc gccggagcgg ccgtctgggg actcgtacgc  
 9180  
 agcgttcaga cggagaacct cggacgcgct gtcttgatcg acctggacgc cgtcgacgac  
 9240  
 cgcgccgacc acacggacgc cgacatcgac gcggcgggtg ccacgggcca ggcccagatc  
 9300  
 gccatacgct ccggcaccct ccaccgcccg cgactggccc gcgtcaccgg ggacctgagc  
 9360  
 ggcaccgccc ccaccgtctt cggcgaccgg cccggcaccg tactgatcac cggcggcacc  
 9420  
 ggcaccctgg gcagcctcgt cgccaggcac ctggtcacca cccacggcgt caccgacctg  
 9480  
 ctgctcacca gccgcccggg ccccggcga cccgggtgccc ccgaactgga cgccgaactc  
 9540  
 accgcgctcg gcgcccgtgt cgaggtggtc gcctgcgaca ccgccgaccg cgacgcgctg  
 9600

gcggccgtgc tggcggaccg caccctcacc ggaatcgtcc acaccgcggg cgtcctcagc  
9660

gacgggatcc tctcctcgtc gacaccggag agaatggccg cggatgatgc ccccaaggtc  
9720

gacgcggcgc tgaacctgca cgaactcacc gccggtcaag acctctcggc gttcgtgatg  
9780

ttctcgtccg ccgcgggcgt gaccggcggc gccggacagg gcaactacgc cgccgccaac  
9840

accttctcgc acggcctcgc cgcgcaccgg cgcgcgaacg ggctctcggc acagtccctg  
9900

gcctggggac tgtgggaaga ggccagcggc atgaccggcg aactggccga cgccgacgtg  
9960

gccggcatgg tgcgcgacgg tgtcctgccc atcggctccg acgagggcct ggcgatgctg  
10020

gacgccgccc gcgcgctcga ccgggcgttc ctggcgcccg tccggctcga cctgagcggg  
10080

cagagccggt ccgacgtgcc ctatctgatg cgcgatctcg tacgcggccc gtcccgccgc  
10140

gtagtcatc ccgccgcacg ggccgcccga ccggccgaga gcctgcgcga ccggctggca  
10200

cggtgactc cgtcccgccg cgaacagaca ctgctcgaca tcgtccgggc gcaggccgcc  
10260

accaccctcg gcttcggcga cgccgacgag gtcgacgccg accggtcggt ccgcgacatg  
10320

ggcttcgact ccctcggcgc cgtgcggttc cgcaacgccc tcggcgaggt catcggggaa  
10380

cgctgcccg cgacgctcgt cttcgaccac ccgacctcgc tcgtctcac tcggcacctg  
10440

cttgaggaaa tggcgatcga cgtacccgag aacgaaccgg agcccagatc ggccgagggg  
10500

gccgcggacc gcaccgccgc gatccagaac atgagcctgg ccgaccttct ccgcaccgcg  
10560

cgacgatcag gagacccgac atga  
10584

<210> 13  
<211> 3527  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 13

Met Thr Glu Gln Ser Arg Ala Ala Met Leu Glu Leu Val Leu Ala Gln  
1 5 10 15

Ala Ala Ala Val Leu Arg Thr Ala Asp Pro Asp Ala Tyr Gly Glu Gly  
20 25 30

10

ES 2 543 203 T3

Ala Val Asp Leu Ala Ala Asp Arg Pro Phe Leu Ala Gly Gly Leu Asp  
35 40 45

Ser Leu Gly Leu Val Arg Leu Gln Arg Arg Leu Ala Ala Glu Val Gly  
50 55 60

Ala Asp Leu Pro Val Thr Leu Ala Phe Asp His Pro Thr Pro Ile Ala  
65 70 75 80

Leu Ala Gly His Leu Ala Asp Leu Leu His Gly Thr Ala Gly Ala Asp  
85 90 95

Ala Pro Asp Glu Thr Ala Ala Pro Ala Leu Pro Ser Ala Phe Gly Glu  
100 105 110

Pro Ile Ala Ile Val Gly Ile Gly Cys Arg Phe Pro Gly Gly Val Ser  
115 120 125

Thr Pro Glu Asp Leu Trp Arg Ile Val Ala Asp Asp Thr His Val Leu  
130 135 140

Thr Asp Phe Pro Ser Asp Arg Gly Trp Asp Leu Asp Arg Ile Tyr Asn  
145 150 155 160

Pro Asp Pro Ser Val Thr Gly Ala Ser Tyr Val Arg Arg Gly Gly Phe  
165 170 175

Leu Pro Asp Ala Ala Asp Phe Asp Ala Asp Phe Phe Gln Ile Ser Pro  
180 185 190

Lys Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr  
195 200 205

Ser Trp Glu Ala Leu Glu His Ala Gly Ile Asp Pro Gly Arg Leu Arg  
210 215 220

Gly Thr Pro Ser Gly Val Phe Ile Gly Val Glu Pro His Glu Tyr Gly  
225 230 235 240

Pro Arg Thr His Glu Ala Pro Asp Gly Leu Asp Gly Tyr Val Leu Gly  
245 250 255

Gly Asn Leu Pro Ser Val Val Ser Gly Arg Val Ala Tyr Thr Leu Gly  
260 265 270

Phe Glu Gly Pro Thr Leu Thr Val Asp Thr Ala Cys Ser Gly Ser Leu  
275 280 285

Thr Ala Leu His Leu Ala Val Arg Ser Leu Gln Gly Gly Glu Cys Gly

290			295			300									
Leu 305	Ala	Leu	Ala	Gly	Gly 310	Val	Thr	Val	Ile	Ser 315	Ser	Pro	Gly	Thr	Phe 320
Thr	Thr	Phe	Ser	Arg 325	Gln	Arg	Gly	Leu	Ala 330	Pro	Asp	Gly	Arg	Ile 335	Lys
Ala	Phe	Ala	Ala 340	Ala	Ala	Asp	Gly	Thr 345	Ser	Phe	Ala	Glu	Gly 350	Val	Gly
Val	Phe	Val 355	Leu	Ala	Arg	Leu	Ser 360	Asp	Ala	Leu	Arg	Asp 365	Gly	His	Pro
Val	Leu 370	Ala	Val	Ile	Arg	Gly 375	Thr	Ala	Ile	Asn	Gln 380	Asp	Gly	Ala	Thr
Asn 385	Gly	Leu	Ser	Ala	Pro 390	Asn	Gly	Leu	Ala	Gln 395	Gln	Arg	Val	Ile	Arg 400
Arg	Ala	Leu	Thr	Asp 405	Ala	Arg	Leu	Thr	Ala 410	Asp	Glu	Val	Asp	Ala 415	Val
Glu	Ala	His	Gly 420	Thr	Gly	Thr	Thr	Leu 425	Gly	Asp	Pro	Ile	Glu 430	Gly	Gln
Ala	Leu	Val 435	Ala	Ala	Tyr	Gly	Arg 440	Gly	Arg	Ser	Pro	Glu 445	Lys	Pro	Leu
Trp	Leu 450	Gly	Ser	Val	Lys	Ser 455	Asn	Ile	Gly	His	Thr 460	Gly	Ala	Ala	Ala
Gly 465	Ala	Ala	Gly	Ile	Ile 470	Lys	Met	Val	Gln	Ala 475	Met	Arg	His	Ser	Thr 480
Leu	Pro	Arg	Thr	Leu 485	His	Val	Asp	Ala	Pro 490	Thr	Thr	His	Val	Asp 495	Trp
Ser	Asp	Gly	Thr 500	Val	Gln	Leu	Leu	Thr 505	Glu	Pro	Val	Pro	Trp 510	Glu	Ala
Ala	Asp	Thr 515	Pro	Arg	Arg	Ala	Gly 520	Ile	Ser	Ser	Phe	Gly 525	Ala	Ser	Gly
Thr	Asn 530	Ala	His	Val	Ile	Ile 535	Glu	Glu	Pro	Pro	Ala 540	Pro	Val	Ala	Ala
Ala 545	Asp	Gly	Ser	Asp	Glu 550	Pro	Pro	Val	Ala	Glu 555	Arg	Pro	Val	Pro	Val 560



Val Leu Ser Ala Gln Gly Gln Asp Ala Leu Arg Gly Gln Ala Glu Arg  
 565 570 575  
 Leu Leu Thr Ala Val Asp Asp Ala Ser Pro Leu Asp Leu Ala Tyr Ser  
 580 585 590  
 Ala Ala Thr Thr Arg Ala Ala Leu Arg Asp Arg Ala Thr Val Val Ala  
 595 600 605  
 Ala Asp Arg Ala Glu Leu Arg Arg Ala Leu Thr Ala Leu Ala Ala Gly  
 610 615 620  
 Glu Ser Ala Pro Gly Leu Leu Thr Gly Thr Thr Pro Ala Ala Gly Arg  
 625 630 635 640  
 Thr Ala Phe Leu Phe Thr Gly Gln Gly Ser Gln Arg Leu Gly Met Gly  
 645 650 655  
 Arg Glu Leu Val Arg Ala Phe Pro Val Phe Ala Arg Ala Leu Gln Asp  
 660 665 670  
 Ala Ala Asp His Leu Asp Leu His Leu Asp Glu Pro Leu Gly Glu Val  
 675 680 685  
 Leu Phe Ala Glu Pro Gly Ser Ala Gln Ala Glu Leu Leu Gln Gln Thr  
 690 695 700  
 Arg Tyr Ala Gln Ala Ala Leu Phe Ala Val Glu Thr Ala Met Phe Arg  
 705 710 715 720  
 Leu Leu Glu Ser Trp Gly Val Thr Pro Gly Leu Leu Thr Gly His Ser  
 725 730 735  
 Ile Gly Glu Ile Ala Ala Ala His Val Ala Gly Val Met Asn Leu Ala  
 740 745 750  
 Asp Ala Ala Leu Leu Val Gly Ala Arg Gly Thr Leu Met Gln Glu Leu  
 755 760 765  
 Pro Ala Gly Gly Ala Met Val Ala Val Gln Ala Ala Glu Ala Glu Val  
 770 775 780  
 Ala Pro Tyr Leu Thr Glu Lys Val Gly Ile Ala Ala Val Asn Gly Pro  
 785 790 795 800  
 Leu Ala Val Val Leu Ser Gly Glu Thr Asp Ala Val Leu Ala Val Ala  
 805 810 815  
 Ala Arg Phe Thr Glu Glu Gly Arg Lys Thr Arg Arg Leu Arg Val Ser

ES 2 543 203 T3

820				825				830							
His	Ala	Phe	His	Ser	Pro	Leu	Met	Glu	Pro	Met	Leu	Asp	Asp	Phe	Arg
		835					840					845			
Arg	Val	Ala	Glu	Thr	Leu	Thr	Tyr	Gln	Pro	Pro	Arg	Ile	Pro	Val	Val
	850					855					860				
Ser	Asn	Leu	Thr	Gly	Glu	Pro	Val	Ala	Ala	Phe	Asp	Ala	Asp	Tyr	Trp
865					870					875					880
Val	Arg	His	Val	Arg	Glu	Pro	Val	Arg	Phe	Ala	Asp	Ala	Met	Ser	Trp
				885					890					895	
Leu	Glu	Ser	Gln	Gly	Val	Thr	Thr	Tyr	Leu	Glu	Leu	Gly	Pro	Asp	Ala
			900					905					910		
Val	Leu	Ser	Ala	Met	Gly	Arg	Asp	Cys	Leu	Thr	Asp	Gly	Gly	Ala	Asp
		915					920					925			
Ala	Ala	Phe	Ala	Ala	Leu	Leu	Arg	Asp	Gly	His	Asp	Glu	Glu	Arg	Gln
	930					935					940				
Ser	Leu	Thr	Ala	Leu	Gly	Leu	Ala	His	Ala	His	Gly	Leu	Ala	Val	Asp
945					950					955					960
Trp	Asp	Arg	Phe	Phe	Ala	Gly	Arg	Gly	Ala	His	Arg	Thr	Ala	Leu	Pro
				965					970					975	
Thr	Tyr	Ser	Phe	Gln	Arg	Arg	Arg	Tyr	Trp	Leu	Asp	Ala	Gly	Ser	Ser
			980					985					990		
His	Pro	Gly	Thr	Ile	Gly	His	Pro	Met	Leu	Asp	Ser	Ala	Val	Ser	Leu
		995					1000					1005			
Ala	Gly	Ala	Asp	Gly	Val	Ile	Leu	Thr	Gly	Arg	Met	Ser	Thr	Arg	
	1010					1015					1020				
Ala	Gln	Pro	Trp	Leu	Ala	Asp	His	Val	Ile	Ala	Gly	Val	Val	Leu	
	1025					1030					1035				
Val	Pro	Gly	Thr	Ala	Leu	Val	Glu	Leu	Ala	Val	Arg	Ala	Gly	Asp	
	1040					1045					1050				
Glu	Ala	Gly	Cys	Ala	Ala	Val	Glu	Glu	Leu	Thr	Leu	Glu	Thr	Pro	
	1055					1060					1065				
Leu	Ala	Val	Thr	Thr	Asp	Ala	Gly	Val	Glu	Ile	Gln	Val	Val	Val	
	1070					1075					1080				

Gly Ala Leu Asp Ala Ser Gly Arg Arg Ser Val Asp Ile Tyr Ser  
 1085 1090 1095  
 Arg Pro Ala Arg Ser Glu Asp Asn Trp Thr Arg His Ala Ser Gly  
 1100 1105 1110  
 Val Leu Thr Glu His Gly Glu Gly Ala Thr Ala Asp Pro Ala Val  
 1115 1120 1125  
 Phe Ala Gln Trp Pro Pro Ala Asp Ala Glu Pro Ile Asp Val Asp  
 1130 1135 1140  
 Gly Leu Tyr Glu Arg Gln Ala Ala Gln Gly Tyr Gly Tyr Gly Pro  
 1145 1150 1155  
 Ala Phe Arg Arg Val Arg Ala Ala Trp Arg Arg Gly Asp Asp Val  
 1160 1165 1170  
 Phe Ala Glu Val Ala Leu Asp Asp Ala Gly Gly Asp Arg Phe Gly  
 1175 1180 1185  
 Leu His Pro Ala Leu Leu Asp Ser Ser Leu His Ala Ala Glu Gly  
 1190 1195 1200  
 Pro Glu Asp Asp Gly Gln Val Arg Leu Pro Phe Ala Trp Arg Gly  
 1205 1210 1215  
 Val Glu Leu His Ala Thr Gly Ala Thr Ala Val Arg Val His Val  
 1220 1225 1230  
 Val Gln Thr Ala Pro Asp Glu Val Thr Val Glu Leu Ala Asp Ala  
 1235 1240 1245  
 Asn Gly Ala Pro Val Ala Thr Val Arg Ser Leu Val Gln Arg Pro  
 1250 1255 1260  
 Val Pro Ile Asn Gly Val Arg Pro Ser Ala Ser Gly Gly Gly Ser  
 1265 1270 1275  
 Leu Leu Arg Val Glu Trp Thr Ala Val Gln Ala Pro Ser Glu Pro  
 1280 1285 1290  
 Asp Ala Ala Ser His Glu Phe Gln Leu Ala Tyr Val Pro Glu Thr  
 1295 1300 1305  
 Phe Pro Gly Glu Pro Ala Asp Ala Ala Arg Ala Ala Thr Leu His  
 1310 1315 1320  
 Ala Leu Thr Leu Ile Gln Asp Ala Leu Arg Asp Asp Thr Arg Leu

1325						1330						1335		
Ala	Ile	Val	Thr	Arg	Gly	Glu	Pro	Tyr	Asp	Phe	Thr	Leu	Ala	Ser
1340						1345					1350			
Pro	Trp	Ala	Leu	Val	Arg	Ser	Ala	Gln	Ala	Glu	Asn	Pro	Gly	Arg
1355						1360					1365			
Phe	Val	Leu	Val	Gly	Thr	Asp	His	Asp	Val	Pro	Glu	His	Glu	Leu
1370						1375					1380			
Arg	Ala	Ala	Leu	Ala	Thr	Gly	Glu	Pro	Gln	Leu	Ala	Leu	Arg	Asp
1385						1390					1395			
Gly	Lys	Val	Leu	Val	Pro	Arg	Leu	Ala	Arg	Thr	Pro	Ala	Pro	Glu
1400						1405					1410			
Asp	Pro	Arg	Pro	Val	Ala	Trp	Asp	Pro	Asp	Gly	Thr	Val	Leu	Ile
1415						1420					1425			
Thr	Gly	Gly	Thr	Gly	Gly	Leu	Gly	Gly	Val	Val	Ala	Arg	His	Leu
1430						1435					1440			
Val	Gln	Gln	His	Gly	Val	Arg	His	Leu	Leu	Leu	Thr	Gly	Arg	Arg
1445						1450					1455			
Gly	Pro	Asp	Ser	Pro	Gly	Ala	Ala	Glu	Leu	Val	Ala	Glu	Leu	Ala
1460						1465					1470			
Gly	Leu	Gly	Ala	His	Ala	Thr	Val	Ala	Ala	Cys	Asp	Val	Ala	Asp
1475						1480					1485			
Arg	Ala	Ala	Leu	Ala	Ala	Leu	Leu	Glu	Thr	Val	Pro	Gly	Glu	His
1490						1495					1500			
Pro	Leu	Thr	Gly	Val	Val	His	Ser	Ala	Gly	Val	Val	Asp	Asp	Gly
1505						1510					1515			
Leu	Ala	Gly	Ser	Leu	Thr	Pro	Glu	Gln	Val	His	Thr	Val	Phe	His
1520						1525					1530			
Gly	Lys	Ala	Asp	Gly	Ala	Trp	His	Leu	His	Glu	Leu	Thr	Arg	Gly
1535						1540					1545			
Leu	Pro	Leu	Ala	Ala	Phe	Val	Leu	Phe	Ser	Ser	Ala	Ala	Gly	Thr
1550						1555					1560			
Met	Glu	Ala	Ala	Gly	Gln	Gly	Asn	Tyr	Ala	Ala	Ala	Asn	Ala	Phe
1565						1570					1575			

Leu Asp Ala Leu Ala Ala His Arg Val Thr Gln Gly Leu Pro Ala  
 1580 1585 1590

Thr Ser Leu Ala Trp Asn Leu Trp Ala Gly Asp Ala Gly Met Gly  
 1595 1600 1605

Ala Arg Leu Asp Glu Val Thr Leu Arg Arg Ala Glu Arg Ser Gly  
 1610 1615 1620

Leu Pro Ala Leu Asp Ala Glu Glu Asn Leu Ala Leu Leu Asp Gln  
 1625 1630 1635

Ala Leu Val Thr Gly Ala Pro Ala Leu Val Pro Leu Arg Val Asp  
 1640 1645 1650

Ala Arg Ala Leu Arg Ala Arg Ser Glu Gly Ile Pro Pro Met Leu  
 1655 1660 1665

Arg Gly Leu Val Arg Pro Pro Ala Arg Arg Asn Thr Ala Ala Ala  
 1670 1675 1680

Ala Ala Thr Gly Pro Gly Gly Ala Leu Ala Asp Arg Leu Ala Gly  
 1685 1690 1695

Lys Pro Asp Ala Glu Arg Glu Arg Ile Val Leu Asp Leu Val Arg  
 1700 1705 1710

Thr Gln Ile Ala Ala Val Leu Gly His Asp Ser Gly Thr Ala Ile  
 1715 1720 1725

Asp Pro Arg Arg Ala Phe Thr Glu Leu Gly Phe Asp Ser Leu Ala  
 1730 1735 1740

Ala Ile Glu Leu Arg Asn Ala Leu Gly Thr Ala Thr Gly Leu Arg  
 1745 1750 1755

Leu Thr Ser Thr Leu Ile Phe Asp His Pro Thr Pro Arg Ala Leu  
 1760 1765 1770

Val Asp His Val Leu Glu Thr Val Arg Gly Ala Val Pro Val Thr  
 1775 1780 1785

Ala Ala Pro Arg Pro Thr Val Arg Thr Ala Thr Asp Glu Pro Ile  
 1790 1795 1800

Ala Ile Val Ala Met Gly Cys Arg Tyr Pro Gly Gly Val Thr Ser  
 1805 1810 1815

Ala Glu Glu Leu Trp Arg Leu Val Ala Gly Gly Thr Asp Ala Ile

ES 2 543 203 T3

1820						1825						1830			
Thr	Glu	Phe	Pro	Asp	Asp	Arg	Asp	Trp	His	Thr	Asp	Asp	Ile	Tyr	
	1835					1840					1845				
Asp	Pro	Glu	Pro	Gly	Lys	His	Gly	Thr	Thr	Tyr	Thr	Arg	Glu	Gly	
	1850					1855					1860				
Gly	Phe	Leu	His	Asp	Val	Ala	Glu	Phe	Asp	Pro	Ala	Phe	Phe	Gly	
	1865					1870					1875				
Ile	Ser	Pro	Lys	Glu	Ala	Gln	Ala	Met	Asp	Pro	Gln	His	Arg	Met	
	1880					1885					1890				
Leu	Leu	Glu	Val	Ala	Trp	Glu	Ala	Leu	Glu	Gln	Gly	Gly	Ile	Asp	
	1895					1900					1905				
Pro	His	Ser	Leu	His	Gly	Thr	Ala	Ala	Gly	Val	Phe	Ala	Gly	Val	
	1910					1915					1920				
Met	Asn	His	Asp	Trp	Thr	Thr	Arg	Ser	Gly	Ala	Val	Pro	Glu	Asp	
	1925					1930					1935				
Leu	Ala	Gly	Phe	Thr	Ala	Gly	Gly	Gly	Leu	Gly	Ser	Ile	Ala	Ser	
	1940					1945					1950				
Gly	Arg	Ile	Ala	Tyr	Thr	Leu	Gly	Leu	Gln	Gly	Pro	Ala	Val	Thr	
	1955					1960					1965				
Ile	Asp	Thr	Ala	Cys	Ser	Ser	Ser	Leu	Val	Ala	Met	His	Trp	Ala	
	1970					1975					1980				
Met	Gln	Ser	Leu	Arg	Gln	Gly	Glu	Cys	Thr	Leu	Ala	Leu	Ala	Gly	
	1985					1990					1995				
Gly	Val	Thr	Val	Met	Ala	Thr	Pro	Glu	Thr	Phe	Val	Gly	Met	Ser	
	2000					2005					2010				
Leu	Gln	Ser	Gly	Leu	Ala	Ala	Asp	Gly	Arg	Cys	Lys	Ala	Tyr	Gly	
	2015					2020					2025				
Ala	Gly	Ala	Asp	Gly	Thr	Gly	Trp	Gly	Glu	Gly	Ala	Gly	Leu	Leu	
	2030					2035					2040				
Val	Leu	Glu	Arg	Leu	Ser	Asp	Ala	Arg	Arg	Asn	Gly	His	Pro	Val	
	2045					2050					2055				
Val	Ala	Val	Ile	Arg	Gly	Ser	Ala	Ile	Asn	Gln	Asp	Gly	Ala	Ser	
	2060					2065					2070				

Asn Gly Ile Ala Ala Pro Asn Gly Pro Ala Gln Gln Arg Val Ile  
 2075 2080 2085  
 Arg Gln Ala Val Ala Ser Ala Gly Leu Thr Leu Ala Asp Ile Asp  
 2090 2095 2100  
 Ala Val Glu Gly His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile  
 2105 2110 2115  
 Glu Ala Gln Ala Leu Met Ala Thr Tyr Gly Gln Glu Arg Arg Asp  
 2120 2125 2130  
 Asp Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Leu Gly His Thr  
 2135 2140 2145  
 Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala  
 2150 2155 2160  
 Met Arg His Gly Val Leu Pro Arg Thr Leu His Ala Glu Thr Pro  
 2165 2170 2175  
 Ser Pro His Ile Asp Trp Thr Glu Gly Ala Val Glu Leu Leu Thr  
 2180 2185 2190  
 Glu Pro Arg Asp Trp Thr Ala Asp Gly Arg Pro Arg Arg Ala Ala  
 2195 2200 2205  
 Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Ile Ile  
 2210 2215 2220  
 Glu Gln Ala Pro Pro Ala Gly Pro Gly Glu Pro Arg Gly Glu Arg  
 2225 2230 2235  
 Pro Pro Val Val Pro Leu Thr Val Ser Gly Ser Thr Pro Glu Ala  
 2240 2245 2250  
 Met Arg Ala Gln Ala Ala Arg Ile Ala Ala His Leu Arg Glu Asn  
 2255 2260 2265  
 Gly Asp Val Asp Glu Leu Asp Ala Ala Ala Thr Leu Ala Arg Gly  
 2270 2275 2280  
 Arg Ala Ala Leu Glu His Arg Ala Val Ile Val Asp Ala Asp Arg  
 2285 2290 2295  
 Asp Gly Leu Leu Ala Gly Leu Asp Ala Leu Ala Ala Gly Asn Ser  
 2300 2305 2310  
 Ser Ala Ala Val Val Gln Gly Leu Gln Arg Gly Asp Ser Arg Ala

ES 2 543 203 T3

	2315					2320						2325		
Val	Leu 2330	Val	Phe	Pro	Gly	Gln 2335	Gly	Ser	Gln	Trp	Gln 2340	Gly	Met	Gly
Val	Glu 2345	Leu	Leu	Glu	His	Ser 2350	Pro	Ala	Phe	Ala	Thr 2355	Arg	Leu	Gly
Glu	Cys 2360	Ala	Lys	Ala	Leu	Glu 2365	Ser	Tyr	Val	Asp	Trp 2370	Asn	Leu	Leu
Asp	Val 2375	Val	His	Gly	Ala	Pro 2380	Gly	Ala	Pro	Ala	Leu 2385	Asp	Ala	Val
Asp	Val 2390	Val	Gln	Pro	Thr	Leu 2395	Trp	Ala	Leu	Met	Val 2400	Ser	Leu	Ala
Glu	Ala 2405	Trp	Arg	Ala	Ala	Gly 2410	Val	Glu	Pro	Ala	Ala 2415	Val	Val	Gly
His	Ser 2420	Gln	Gly	Glu	Ile	Ala 2425	Ala	Ala	Cys	Val	Ala 2430	Gly	Ala	Leu
Ser	Leu 2435	Glu	Asp	Gly	Ala	Arg 2440	Val	Val	Ala	Leu	Arg 2445	Ser	Arg	Val
Ile	Arg 2450	Gln	Asp	Leu	Ala	Gly 2455	Arg	Gly	Gly	Met	Met 2460	Ser	Val	Ala
Leu	Ser 2465	Ala	Asp	Arg	Ala	Ser 2470	Glu	Tyr	Leu	Ala	Asp 2475	Trp	Asp	Gly
Arg	Leu 2480	Gln	Leu	Ala	Val	Val 2485	Asn	Gly	Pro	Ser	Ser 2490	Val	Val	Val
Cys	Gly 2495	Gly	Pro	Glu	Ala	Leu 2500	Asp	Glu	Leu	Arg	Gly 2505	Arg	Leu	Asp
Thr	Asp 2510	Glu	Val	Gln	Ser	Arg 2515	Arg	Ile	Pro	Val	Asp 2520	Tyr	Ala	Ser
His	Ser 2525	Met	Phe	Val	Glu	Glu 2530	Ile	Arg	Asp	Arg	Leu 2535	Leu	Thr	Glu
Leu	Ser 2540	Gly	Leu	Thr	Pro	Arg 2545	Thr	Ser	Ser	Ile	Pro 2550	Phe	Tyr	Ser
Thr	Val 2555	Thr	Ala	Gly	Thr	Leu 2560	Asp	Thr	Ala	Gly	Leu 2565	Asp	Ala	Glu



Tyr Trp Tyr Thr Asn Leu Arg Gln Thr Val Arg Phe Glu Asp Thr  
 2570 2575 2580

Thr Arg Ala Leu Leu Ala Asp Gly Phe Glu Thr Phe Ile Glu Ala  
 2585 2590 2595

Ser Pro His Pro Gly Leu Leu Thr Gly Leu Asp Glu Thr Ala Glu  
 2600 2605 2610

Ser Ala Gly Val Ala Ala Thr Leu Val Gly Thr Leu Arg Arg Asp  
 2615 2620 2625

Ser Gly Ser Pro Arg Gln Phe Val Thr Ser Leu Ala Glu Ala Tyr  
 2630 2635 2640

Val Arg Gly Ala Thr Val Asp Trp Asp Thr Gln Phe Val Gly Thr  
 2645 2650 2655

Gly Ala Gln His Val Glu Leu Pro Thr Tyr Pro Phe Gln Arg Lys  
 2660 2665 2670

Arg Tyr Trp Leu His Leu Ser Glu His Thr Gly Asp Ala Val Gly  
 2675 2680 2685

Ile Gly Gln Ile Pro Thr Asp His Pro Leu Leu Gly Ala Ala Val  
 2690 2695 2700

Pro Val Ala Gly Ala Gly Gly Val Leu Leu Thr Gly Arg Leu Ser  
 2705 2710 2715

Leu Ser Gly Gln Pro Trp Leu Ala Asp His Val Val Gly Gly Thr  
 2720 2725 2730

Val Leu Phe Pro Gly Ser Gly Phe Val Glu Leu Ala Val Arg Ala  
 2735 2740 2745

Gly Asp Glu Val Gly Arg Gly Arg Val Glu Glu Leu Thr Leu Glu  
 2750 2755 2760

Ala Pro Leu Ala Leu Pro Glu Arg Gly Gly Val Ala Ile Gln Val  
 2765 2770 2775

Val Val Asp Ala Asp Leu Arg Thr Val Ser Ile His Ser Arg Pro  
 2780 2785 2790

Asp Asn Ala Pro Ala Asp Thr Leu Trp Thr Arg His Ala Gln Gly  
 2795 2800 2805

Thr Leu Thr Asp Ala Ala Pro Glu Pro Ala His Glu Pro Ala Ala

ES 2 543 203 T3

2810						2815					2820			
Trp	Pro	Pro	Pro	Gly	Ala	Glu	Pro	Val	Asp	Leu	Thr	Gly	Phe	Tyr
	2825					2830					2835			
Glu	Pro	Ala	Ala	Asp	Gly	Gly	Leu	Ala	Tyr	Gly	Pro	Val	Phe	Gln
	2840					2845					2850			
Gly	Leu	Arg	Ala	Ala	Trp	Arg	Ser	Gly	Asp	Glu	Val	Phe	Ala	Glu
	2855					2860					2865			
Ile	Thr	Leu	Pro	Glu	Gln	Ala	Ala	Ala	Glu	Ala	Gln	Arg	Phe	Gly
	2870					2875					2880			
Leu	His	Pro	Ala	Ala	Leu	Asp	Ala	Ala	Leu	His	Ala	Thr	Gly	Leu
	2885					2890					2895			
Leu	Ala	Thr	Asp	Ala	Gln	Arg	Val	Thr	Leu	Pro	Phe	Ala	Trp	Thr
	2900					2905					2910			
Arg	Val	Asp	Leu	His	Ala	Ser	Gly	Ala	Ala	Ala	Leu	Arg	Leu	Arg
	2915					2920					2925			
Met	Thr	Ser	Leu	Gly	Asp	Asp	Glu	Val	Ala	Leu	Arg	Leu	Thr	Asp
	2930					2935					2940			
Thr	Ala	Gly	Arg	Pro	Val	Ala	Ser	Val	Glu	Ser	Leu	Val	Leu	Arg
	2945					2950					2955			
Pro	Val	Ala	Pro	Gly	Gly	Pro	Ile	Arg	Thr	Gly	Ala	Tyr	Asp	Asp
	2960					2965					2970			
Ser	Leu	Phe	Glu	Leu	Val	Trp	Ala	Pro	Ala	Ala	Pro	Val	Pro	Gly
	2975					2980					2985			
Gly	Thr	Ala	Pro	Arg	Thr	Val	Val	His	His	Cys	Thr	Gly	Gly	Thr
	2990					2995					3000			
Thr	Ala	Ala	Ser	Ala	His	Ala	Glu	Thr	Ala	Thr	Thr	Leu	Ala	Val
	3005					3010					3015			
Leu	Gln	Ser	Trp	Leu	Glu	Ser	Gly	Ala	Asp	Asp	Ala	Val	Leu	Ala
	3020					3025					3030			
Val	Val	Thr	Arg	Gly	Ala	Leu	Ser	Val	Asn	Gly	Glu	Asp	Val	Thr
	3035					3040					3045			
Asp	Leu	Ala	Gly	Ala	Ala	Val	Trp	Gly	Leu	Val	Arg	Ser	Ala	Gln
	3050					3055					3060			

Thr Glu Asn Pro Gly Arg Val Val Leu Ile Asp Leu Asp Ala Val  
 3065 3070 3075

Asp Asp Arg Ala Asp His Thr Asp Ala Asp Ile Asp Ala Ala Val  
 3080 3085 3090

Ala Thr Gly Glu Ala Gln Ile Ala Ile Arg Ser Gly Thr Leu His  
 3095 3100 3105

Arg Pro Arg Leu Ala Arg Val Thr Gly Asp Leu Ser Gly Thr Ala  
 3110 3115 3120

Ala Thr Val Phe Gly Asp Arg Pro Gly Thr Val Leu Ile Thr Gly  
 3125 3130 3135

Gly Thr Gly Thr Leu Gly Ser Leu Val Ala Arg His Leu Val Thr  
 3140 3145 3150

Thr His Gly Val Thr Asp Leu Leu Leu Thr Ser Arg Arg Gly Pro  
 3155 3160 3165

Ala Ala Pro Gly Ala Ala Glu Leu Asp Ala Glu Leu Thr Ala Leu  
 3170 3175 3180

Gly Ala Arg Val Glu Val Val Ala Cys Asp Thr Ala Asp Arg Asp  
 3185 3190 3195

Ala Leu Ala Ala Val Leu Ala Asp Arg Thr Leu Thr Gly Ile Val  
 3200 3205 3210

His Thr Ala Gly Val Leu Asp Asp Gly Ile Leu Ser Ser Leu Thr  
 3215 3220 3225

Pro Glu Arg Met Ala Ala Val Met Arg Pro Lys Val Asp Ala Ala  
 3230 3235 3240

Leu Asn Leu His Glu Leu Thr Ala Gly Gln Asp Leu Ser Ala Phe  
 3245 3250 3255

Val Met Phe Ser Ser Ala Ala Gly Val Thr Gly Gly Ala Gly Gln  
 3260 3265 3270

Gly Asn Tyr Ala Ala Ala Asn Thr Phe Leu Asp Gly Leu Ala Ala  
 3275 3280 3285

His Arg Arg Ala Asn Gly Leu Ser Ala Gln Ser Leu Ala Trp Gly  
 3290 3295 3300

Leu Trp Glu Glu Ala Ser Gly Met Thr Gly Glu Leu Ala Asp Ala

3305						3310						3315
Asp Val	Ala Gly	Met Val	Arg	Asp Gly	Val Leu	Pro	Ile Gly	Ser				
3320			3325			3330						
Asp Glu	Gly Leu	Ala Met	Leu	Asp Ala	Ala Gly	Ala	Leu Asp	Arg				
3335			3340			3345						
Ala Phe	Leu Ala	Pro Val	Arg	Leu Asp	Leu Ser	Gly	Gln Ser	Arg				
3350			3355			3360						
Ser Asp	Val Pro	Tyr Leu	Met	Arg Asp	Leu Val	Arg	Gly Pro	Ser				
3365			3370			3375						
Arg Arg	Val Val	Asp Pro	Ala	Ala Arg	Ala Ala	Glu	Pro Ala	Glu				
3380			3385			3390						
Ser Leu	Arg Asp	Arg Leu	Ala	Arg Leu	Thr Pro	Ser	Arg Arg	Glu				
3395			3400			3405						
Gln Thr	Leu Leu	Asp Ile	Val	Arg Ala	Gln Ala	Ala	Thr Thr	Leu				
3410			3415			3420						
Gly Phe	Gly Asp	Ala Asp	Glu	Val Asp	Ala Asp	Arg	Ser Phe	Arg				
3425			3430			3435						
Asp Met	Gly Phe	Asp Ser	Leu	Ala Ala	Val Arg	Phe	Arg Asn	Ala				
3440			3445			3450						
Leu Gly	Glu Val	Ile Gly	Glu	Arg Leu	Pro Ala	Thr	Leu Val	Phe				
3455			3460			3465						
Asp His	Pro Thr	Ser Leu	Val	Leu Thr	Arg His	Leu	Leu Glu	Glu				
3470			3475			3480						
Met Ala	Ile Asp	Val Pro	Glu	Asn Glu	Pro Glu	Pro	Glu Ser	Ala				
3485			3490			3495						
Glu Gly	Ala Ala	Asp Arg	Thr	Ala Ala	Ile Gln	Asn	Met Ser	Leu				
3500			3505			3510						
Ala Asp	Leu Leu	Arg Thr	Ala	Arg Arg	Ser Gly	Asp	Pro Thr					
3515			3520			3525						

<210> 14  
 <211> 1599  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

5

<400> 14

atgagcccga atggagcccc tccggtccgg gcgcacgcgt cggaggacaa cctgcggacg  
60

tatctgctgg ccgccgcccc gagggacccc gaccggcccc ccgtgatcca ggcggcggcg  
120

gacgggtggtc tggagaccgt cacctacggc gaactggagc ggcgcgtcga ccgcctcgtc  
180

gacgcgctcg acgcgctcgg cctggatgtg ggcgaccgcg tcatcctgga atccgacacc  
240

aacgccgacg cggtagcgac cctgctggcc tgcgccacgc tggggctgcc gttcaccgcc  
300

accagccccg aggtccccga cgagcggctg ctgacgatca tcgacagcgc ggagccccgc  
360

ctgcacatcc agtccgacca cgggcagcgc accggcattc ccgagacggt cggcaccgcc  
420

cgcttcggcc ccggcggcct caccgtggag cgggcgcccc cgccgcgcgt ccgccgccgc  
480

cgctggtgga cgcccatcga caccgcgtac atcaccttca cgtcgggcac gaccggcccg  
540

cccaagggcg tggatcatgag ccaccgcggg gtcacgcgt tcctgcgggg cgctgaggcc  
600

gcacggctcg tcacggccga ggaccgggtg gcgaacacct ctccgctcca gttcgacttc  
660

gcgctgttcg acatcggcct cgccctcggc cacggcgcca cgctggtgcc cgtgccgcgc  
720

gcccggctca actggccccg ccgcttcctt tcgtacctgc gcgatgcggg ggtcacacag  
780

gtcgacggtg tcccgctcgt ctggcgcccc gtactgcgcc acgaatccga cctgctggcc  
840

gagatgggccc aggagggcgt cctcagccgc atcaccttct ccggggagga cttccccctg  
900

gacgagctgc ggcgcctcca ggaactgctg ccgaaggcgc gcttcaccaa cggttacggg  
960

gccaccgaga cgatggccgc gtccatcacc gaggtgcccc acccgctgcc gcccggcacc  
1020

gaacgcctct ccatcggtta cgccgctgcc ggcgccgaga tgatgctcct cgggtccgac  
1080

ggcagccccg tcgacgagcc cggcaccatc ggcgagatct acctgcgcag cccctcgtc  
1140

ttctccggct actggcgcga cccggaggcc acccgcgccg tcgtcctccc ggacccgctg  
1200

caccccgat cgggacaggt ggtgttcaag accggcgacc tggcccatat ggggccccgag  
1260

ggcgagctgt acttccgcgg tcgcgtcgac tcccagggtc agatccgcgg caaccgcgtc  
1320

gaactcggcg aggtggagcg ccgtatcggc cagttcggcg gtgtcaccgg cgcggtcgcg  
1380  
atggtcctgc cgcgcgagga cggcgaccca ctgctgcacg cgttcgtcac ctgcgcccc  
1440  
gctgggtccc cgcgggacac cagagaagtc ctcgcgttct gtcgcgccgc gctgcccgcc  
1500  
tacatgggtcc cgcacggctt gagcgtgggtg aacgagttcc ccgtgaccgt caacggcaag  
1560  
gtcgaccgca ccgccctggc cgcccacgtg gctcctga  
1599

<210> 15  
<211> 532  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 15

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

10

Thr Thr Gly Arg Pro Lys Gly Val Val Met Ser His Arg Gly Val Ile  
 180 185 190  
 Ala Phe Leu Arg Gly Ala Glu Ala Ala Arg Leu Val Thr Ala Glu Asp  
 195 200 205  
 Arg Val Ala Asn Thr Ser Pro Leu Gln Phe Asp Phe Ala Leu Phe Asp  
 210 215 220  
 Ile Gly Leu Ala Leu Gly His Gly Ala Thr Leu Val Pro Val Pro Arg  
 225 230 235 240  
 Ala Arg Leu Asn Trp Pro Arg Arg Phe Leu Ser Tyr Leu Arg Asp Ala  
 245 250 255  
 Gly Val Thr Gln Val Asp Gly Val Pro Ser Val Trp Arg Pro Val Leu  
 260 265  
 Arg His Glu Ser Asp Leu Leu Ala Glu Met Gly Gln Glu Gly Val Leu  
 275 280 285  
 Ser Arg Ile Thr Phe Ser Gly Glu Asp Phe Pro Leu Asp Glu Leu Arg  
 290 295 300  
 Arg Leu Gln Glu Leu Leu Pro Lys Ala Arg Phe Thr Asn Gly Tyr Gly  
 305 310 315  
 Ala Thr Glu Thr Met Ala Ala Ser Ile Thr Glu Val Pro Asn Pro Leu  
 325 330 335  
 Pro Pro Gly Thr Glu Arg Leu Ser Ile Gly Tyr Ala Val Ala Gly Ala  
 340 345 350  
 Glu Met Met Leu Leu Gly Ser Asp Gly Ser Pro Val Asp Glu Pro Gly  
 355 360 365  
 Thr Ile Gly Glu Ile Tyr Leu Arg Ser Pro Ser Leu Phe Ser Gly Tyr  
 370 375 380  
 Trp Arg Asp Pro Glu Ala Thr Arg Ala Val Val Leu Pro Asp Pro Leu  
 385 390 395 400  
 His Pro Gly Ser Gly Gln Val Val Phe Lys Thr Gly Asp Leu Ala His  
 405 410 415  
 Met Gly Pro Glu Gly Glu Leu Tyr Phe Arg Gly Arg Val Asp Ser Gln  
 420 425 430  
 Val Gln Ile Arg Gly Asn Arg Val Glu Leu Gly Glu Val Glu Arg Arg





gaagagatcg cgaacgggat caccacggac atcacgttct ccgatccccg gatccccgtg  
720  
gtgtccgacc acgacgggtc gctggtacgg acgggggccc ggggtgcggga gttgctgctg  
780  
aacgccgtga cgcacaccgt gcggtggccc gccgtcgtcg acacgatcaa ggggctcggc  
840  
gtcgagcggg tgcattgtcac cgggcaggac gccctgtggg gacgggtgga tgcattgacc  
900  
aacgcgttcc aggtggtggc ggtgcgtccg gacacagcta tgcgaccgcg ccgtcgcagc  
960  
gcgatcgc at ag  
972

<210> 17  
<211> 323  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 17

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

10

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

<210> 18  
 <211> 237  
 <212> ADN  
 <213> *Streptomyces sp.* MP28-13

5

<400> 18

atgtgggacg aatcattcga gcaacttctc cgcaaacaga ttcctttgct ggaaccggat  
 60  
 gaggagttga ccgccgaatt gagcctgctc gattgcggtc tcgactccat gggaatggtc  
 120  
 tcactgctgt cctcgatgga ggacgcatat ggagtcggtt tcgtcgacga tgcgctcaac  
 180  
 atggataact tcgcgactcc cggggctctc tggaaaaccc tggacgccat gcgctga  
 237

10

<210> 19  
 <211> 78  
 <212> PRT  
 <213> *Streptomyces sp.* MP28-13

15

<400> 19

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

5 <210> 20  
 <211> 1518  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

10 <400> 20

atgaatcaga caccctgccc cggacacggc ctgcacgaac ggttcctgac cggcctggcg  
 60  
 ctgtcgcccc gccggaccgc gatccgcgtg cacgccaccg agagcctgac gtacgagcag  
 120  
 atgcacgaac tggcgatgcg ccgggcccgc gcactgcggg ccatggctcc gcaagggccg  
 180  
 cacaacgtcg ccgtgctggc ggacaagagc ctgaccgctt atgtcgggat catcgccgcg  
 240  
 ctgtacgcgg gcgccaccgt cgtaccgctc aaccgcgggt tcccggccga gcgcaccgcg  
 300  
 tccatgctca tcgccgcaa cgtctccacc gtcacgctg atccgatcgg ccgctcctca  
 360  
 ctgcgaggaga ccgagctgga tctgcccgtc ctggacgagg gcaggacggg gccctcgctg  
 420  
 gacacgcccg tggccgtcaa cccttcgat gtcgcgtacg tcctgttcac ctcgggctcg  
 480  
 acgggcccgc ccaagggggt gccgatcacc cacggggcca accaccacta cttegacctg  
 540  
 ctggaccggc gctacgactt cagccccgac gacgtgttct gccagaacgt cggactcaac  
 600  
 ttgactgcg ccatgttcga gatgttctgc gcgtggggca acggggcgca ggtgcacccc  
 660  
 gtcccggccc ccgcccaccg ggacctgccg gcgttcttgg ccgagcggaa gatgaccgtg  
 720

tggttctcca ccccgagcgg catcacgttc atccggcggg tgggcggcct gacccccgga  
780

tccgatgcca cactgcgctg gaccttcttc gccgggtgagg cgctgctgca cgaggacgcc  
840

gccgactggc acgtcgccgc accccagtcg aagatcgaga atctgtacgg gccgaccgag  
900

ctgaccgtga ccatcaccgg gcaccgctgg tcgccgaaga ccaccgagga gcagaccgtg  
960

aacggcggcg tgccgatcgg aaaggtgcac cccggccacg accacctgct gctggacgac  
1020

gacggcgagt cggcgggtgga gggcgaactg tgcgtcgccg gaccgcagat gacacccggt  
1080

tacctggacg gcgacgaca cgggggcccgc ttcctcgagc acgccggccg tcgctggtac  
1140

cggaccggcg accgggtgcg gcggctggac gacgacgagc tgatctacct cggccggatg  
1200

gacgcccagg tgcagatcca gggattccgg gtcgaactgg ccgaggtcga ccatgtcgtc  
1260

cggcagtgca ccggtgtgca gaacgcggcc accgtcaccg gcccggcacc gaacggcggg  
1320

ctggaactcg tccttacta cacgggcgag cgcattccgt cggcgacgct gcgccgag  
1380

ctggccgccc acctgcccga tccgatggtg cccaagacct tccggcacgt gccggagtcc  
1440

ccgctcaatt ccaaccgcaa ggtcgaccgg gcgcagttgg cccgggaggc cgccgcgctg  
1500

tcagacggtc gtgcctga  
1518

<210> 21  
<211> 505  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 21

Met Asn Gln Thr Pro Val Pro Gly His Gly Leu His Glu Arg Phe Leu  
1 5 10 15

Thr Gly Leu Ala Leu Ser Pro Gly Arg Thr Ala Ile Arg Val His Ala  
20 25 30

Thr Glu Ser Leu Thr Tyr Glu Gln Met His Glu Leu Ala Met Arg Arg  
35 40 45

Ala Ala Ala Leu Arg Ala Met Ala Pro Gln Gly Pro His Asn Val Ala  
50 55 60

10

Val 65 Leu Ala Asp Lys 70 Ser Leu Thr Ala Tyr 75 Val Gly Ile Ile Ala 80 Ala  
 Leu Tyr Ala Gly 85 Ala Thr Val Val Pro 90 Leu Asn Pro Arg Phe Pro Ala 95  
 Glu Arg Thr 100 Arg Ser Met Leu Ile 105 Ala Ala Asn Val Ser Thr Val Ile  
 Ala Asp Pro 115 Ile Gly Arg Ser 120 Ser Leu Ala Glu Thr 125 Glu Leu Asp Leu  
 Pro Val 130 Leu Asp Glu Gly 135 Arg Thr Gly Pro Ser 140 Leu Asp Thr Pro Val  
 Ala Val 145 Asn Pro Ser 150 Asp Val Ala Tyr Val 155 Leu Phe Thr Ser Gly Ser 160  
 Thr Gly Arg Pro 165 Lys Gly Val Pro Ile 170 Thr His Gly Ala Asn His His 175  
 Tyr Phe Asp 180 Leu Leu Asp Arg Arg 185 Tyr Asp Phe Ser Pro 190 Asp Asp Val  
 Phe Cys 195 Gln Asn Val Gly Leu 200 Asn Phe Asp Cys Ala 205 Met Phe Glu Met  
 Phe 210 Cys Ala Trp Gly Asn 215 Gly Ala Gln Val His 220 Pro Val Pro Pro Ala  
 Ala 225 His Arg Asp Leu 230 Pro Ala Phe Leu Ala 235 Glu Arg Lys Met Thr Val 240  
 Trp Phe Ser Thr 245 Pro Ser Gly Ile Thr 250 Phe Ile Arg Arg Met Gly Gly 255  
 Leu Thr Pro 260 Gly Ser Met Pro Thr 265 Leu Arg Trp Thr Phe 270 Phe Ala Gly  
 Glu Ala 275 Leu Leu His Glu Asp 280 Ala Ala Asp Trp His 285 Val Ala Ala Pro  
 Gln 290 Ser Lys Ile Glu Asn 295 Leu Tyr Gly Pro Thr 300 Glu Leu Thr Val Thr  
 Ile 305 Thr Gly His Arg 310 Trp Ser Pro Lys Thr 315 Thr Glu Glu Gln Thr Val 320  
 Asn Gly Gly Val 325 Pro Ile Gly Lys Val 330 His Pro Gly His Asp 335 His Leu

Leu Leu Asp Asp Asp Gly Glu Ser Ala Val Glu Gly Glu Leu Cys Val  
 340 345 350  
 Ala Gly Pro Gln Met Thr Pro Gly Tyr Leu Asp Gly Asp Asp Asn Arg  
 355 360 365  
 Gly Arg Phe Leu Glu His Ala Gly Arg Arg Trp Tyr Arg Thr Gly Asp  
 370 375 380  
 Arg Val Arg Arg Leu Asp Asp Asp Glu Leu Ile Tyr Leu Gly Arg Met  
 385 390 395 400  
 Asp Ala Gln Val Gln Ile Gln Gly Phe Arg Val Glu Leu Ala Glu Val  
 405 410 415  
 Asp His Val Val Arg Gln Cys Thr Gly Val Gln Asn Ala Ala Thr Val  
 420 425 430  
 Thr Arg Pro Ala Pro Asn Gly Gly Leu Glu Leu Val Leu Tyr Tyr Thr  
 435 440 445  
 Gly Glu Arg Ile Pro Ser Ala Thr Leu Arg Arg Glu Leu Ala Ala His  
 450 455 460  
 Leu Pro Asp Pro Met Val Pro Lys Thr Phe Arg His Val Pro Glu Phe  
 465 470 475 480  
 Pro Leu Asn Ser Asn Arg Lys Val Asp Arg Ala Gln Leu Ala Arg Glu  
 485 490 495  
 Ala Ala Ala Leu Ser Asp Gly Arg Ala  
 500 505

<210> 22  
 <211> 597  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

<400> 22

atgggacacc gagaggatct gctggccgga gcgaagcaat gcctctacga gaaggggtac  
60

gcgcgtacca ccgcgcgca catcgtcgag gcttccggta cgaatctcgc ctcgatcggc  
120

taccactacg gcaccaagga agcactcctc aacgccgcca tcatggaggc gctggaggag  
180

tgggcccagg agctgaagag tgctctggca gacgtgggag acctccccga cgatccgatc  
240

aagcggttcg aggtggcctg gacgcgcgtg atcgagctgt tcgagcgcca ccgcccgggtg  
300

5

10

tggggggccc agttcgacgc gatctcccag cgcgaccatg tgccccgaggt cggcagcttc  
 360  
 ttcatagagg cgcagcagca ggcacagaac ggcctgggtcc ggctcctgtg gggcggcgaa  
 420  
 gagctgccgg gccccaccgc gaccgcggtc gggcagttct accacgcgct gctcagcggc  
 480  
 gtgatgatgc agtgggtcgt cgacccggag cacgcctcga cgggcgagtc gctcggccgag  
 540  
 gcgctacgga cggtcgcgga gaacgccctg cacttcgggt caggcacgac cgtctga  
 597

<210> 23  
 <211> 198  
 <212> PRT  
 <213> *Streptomyces sp.* MP28-13  
 <400> 23

5

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

10

Ser Leu Ala Glu Ala Leu Arg Thr Val Ala Glu Asn Ala Arg His Phe  
 180 185 190

Gly Ser Gly Thr Thr Val  
 195

5

<210> 24  
 <211> 1575  
 <212> ADN  
 <213> *Streptomyces sp.* MP28-13  
 <400> 24

atgagtcgc aacgtgcaac cttgagggac tgggtcggcc tcgccgtcct tgtcgtccct  
 60  
 gtcctcatga tgtcgatgga catgacgggtg ctgtacttcg cgctgccgtt cctcagcgcg  
 120  
 accctggaac cgagcgcacc cgagcaactg tggatcgtgg acatctacgc gttcatgctc  
 180  
 gccgggctgc tcatcgcgat gggcacactc ggtgaccaca tcggccgccc gccggctgctg  
 240  
 atcatcggcg cgggtggtgtt cggcgcgctc tctactggcct ccgcctacgc gaccagcgc  
 300  
 gagcttctga tcctcgcccc cgccgtgctc ggtatgtccg gcgccgtact cgcgccgtcc  
 360  
 acgctctcgc tgatccgcaa catgttccag gatcccggcc agcgccttac cgccatcgcg  
 420  
 gtatggaccg ccggtctctc cggcggcgcc gccctcggtc cgatcgtgtc gggagtgctg  
 480  
 ctggagcact actggtgggg ctcggtcttc ctgatcaaca tcccggtgac gatcctgatc  
 540  
 gtggtgctcg gccccatcct cctgccggag caccgcgacc ccgagcccgg ccgtttcgac  
 600  
 ttctcgggtg ccgtgctgtc gctggccgcg atgcttcccg tcatctacgg catcaaggaa  
 660  
 ctgccgacg acggcttcga ctggaagtac gtggcggta ccgccgcccg cctggtcatc  
 720  
 ggggtgctct tcgtcctgcg ccagcgcgcg gccccaatc cgctgatcga cctgagcctc  
 780  
 ttccgcgacc gggggttcac cgcgtccatc ggagtcaacc tggtgccct gttcgcgatg  
 840  
 atcgggttcc tgctcttcgc gaccagtggt atccagctgg tccacgggct gaatccgctg  
 900  
 gaggcgggcc tctggacact gcccgcgccc ttggcgggtg cggtcacgac atcgggtcgc  
 960  
 gtcgggctg cgaagaagat ccgccccggc tacatcatgg ccatcggcat ggtcatcgcg  
 1020

10



tcggcgggat tcgccatcat gacgcaactg cgcgccgatt cgagcctggc catggcggtg  
1080

atcggcgcga gcgtgctgtc ggccggcgtc ggcattggcga tccccctgac cgccgacctg  
1140

atcgtctccg cggctccgga ggaccgcgtg ggcgctgccg ccgcgctgcc cgagaccgcc  
1200

aaccagctcg gcggagcgct gggcgtagcg atcctcggca gcatcgggtc cgccgtgtac  
1260

acccgtgacg tcgccgacgt gacgacgggg ctgccacccg aggccgcgga ggcagcggag  
1320

ggttcgctcg gcggcgcgac ggaagtggcc aaacacctgc ccggtgacac gggcgcgccc  
1380

ctcgtcacgt ccgccgggga ggccttcacc cgcggcatga acctcagcgc cgcggtgggc  
1440

ggcgtcgtca tgctgctcgg tgccgcgggc gcggcgctgc tcctgcgcca tgtcaagact  
1500

cccaccgtca cgtccgcgcc ggcggacgag acgaagggcg agacggcgga cgagccctca  
1560

cccgtcccca agtag  
1575

<210> 25

<211> 524

<212> PRT

<213> *Streptomyces* sp. MP28-13

<400> 25

Met Ser Pro Gln Arg Ala Thr Leu Arg Asp Trp Val Gly Leu Ala Val  
1 5 10 15

Leu Val Val Pro Val Leu Met Met Ser Met Asp Met Thr Val Leu Tyr  
20 25 30

Phe Ala Leu Pro Phe Leu Ser Ala Thr Leu Glu Pro Ser Ala Thr Glu  
35 40 45

Gln Leu Trp Ile Val Asp Ile Tyr Ala Phe Met Leu Ala Gly Leu Leu  
50 55 60

Ile Ala Met Gly Thr Leu Gly Asp His Ile Gly Arg Arg Arg Leu Leu  
65 70 75 80

Ile Ile Gly Ala Val Val Phe Gly Ala Ser Ser Leu Ala Ser Ala Tyr  
85 90 95

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

5

10

Ser Gly Ala Val Leu Ala Pro Ser Thr Leu Ser Leu Ile Arg Asn Met  
 115 120 125  
 Phe Gln Asp Pro Gly Gln Arg Arg Thr Ala Ile Ala Val Trp Thr Ala  
 130 135 140  
 Gly Leu Ser Gly Gly Ala Ala Leu Gly Pro Ile Val Ser Gly Val Leu  
 145 150 155 160  
 Leu Glu His Tyr Trp Trp Gly Ser Val Phe Leu Ile Asn Ile Pro Val  
 165 170 175  
 Thr Ile Leu Ile Val Val Leu Gly Pro Ile Leu Leu Pro Glu His Arg  
 180 185 190  
 Asp Pro Glu Pro Gly Arg Phe Asp Phe Leu Gly Ala Val Leu Ser Leu  
 195 200 205  
 Ala Ala Met Leu Pro Val Ile Tyr Gly Ile Lys Glu Leu Ala Asp Asp  
 210 215 220  
 Gly Phe Asp Trp Lys Tyr Val Ala Val Thr Ala Ala Gly Leu Val Ile  
 225 230 235 240  
 Gly Val Leu Phe Val Leu Arg Gln Arg Ala Ala Pro Asn Pro Leu Ile  
 245 250 255  
 Asp Leu Ser Leu Phe Arg Asp Arg Gly Phe Thr Ala Ser Ile Gly Val  
 260 265 270  
 Asn Leu Val Ala Leu Phe Ala Met Ile Gly Phe Leu Leu Phe Ala Thr  
 275 280 285  
 Gln Trp Ile Gln Leu Val His Gly Leu Asn Pro Leu Glu Ala Gly Leu  
 290 295 300  
 Trp Thr Leu Pro Ala Pro Leu Ala Val Ala Val Thr Thr Ser Val Ala  
 305 310 315 320  
 Val Gly Leu Ala Lys Lys Ile Arg Pro Gly Tyr Ile Met Ala Ile Gly  
 325 330 335  
 Met Val Ile Ala Ser Ala Gly Phe Ala Ile Met Thr Gln Leu Arg Ala  
 340 345 350  
 Asp Ser Ser Leu Ala Met Ala Val Ile Gly Ala Ser Val Leu Ser Ala  
 355 360 365  
 Gly Val Gly Met Ala Ile Pro Leu Thr Ala Asp Leu Ile Val Ser Ala

370                                      375                                      380  
 Ala Pro Glu Asp Arg Val Gly Ala Ala Ala Ala Leu Pro Glu Thr Ala  
 385                                      390                                      395                                      400  
 Asn Gln Leu Gly Gly Ala Leu Gly Val Ala Ile Leu Gly Ser Ile Gly  
                                     405                                      410  
 Ala Ala Val Tyr Thr Arg Asp Val Ala Asp Val Thr Thr Gly Leu Pro  
                                     420                                      425                                      430  
 Pro Glu Ala Ala Glu Ala Ala Glu Gly Ser Leu Gly Gly Ala Thr Glu  
                                     435                                      440                                      445  
 Val Ala Lys His Leu Pro Gly Asp Thr Gly Asp Ala Leu Val Thr Ser  
                                     450                                      455                                      460  
 Ala Gly Glu Ala Phe Thr Arg Gly Met Asn Leu Ser Ala Ala Val Gly  
 465                                      470                                      475  
 Gly Val Val Met Leu Leu Gly Ala Ala Gly Ala Ala Leu Leu Leu Arg  
                                     485                                      490                                      495  
 His Val Lys Thr Pro Thr Val Thr Ser Ala Pro Ala Asp Glu Thr Lys  
                                     500                                      505                                      510  
 Gly Glu Thr Ala Asp Glu Pro Ser Pro Val Pro Lys  
                                     515                                      520

<210> 26  
 <211> 1236  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13  
 <400> 26

gtgaccaccg aagcgacgcc caccaccag caggccgacg aagcgctgat gaagctgctg  
 60  
 tgcgcccect tccccgacga ccccttcccg atctacgaga ccttgcagtc ggtgaaccgg  
 120  
 gtccacaagt ccgcgctcgg tatctacgcg ctctccgggt acgaggaggt gaccgagctc  
 180  
 ctgaagatgc ccgatgtcca cagtggggcc cgtgccgccc ctcagatgcg cgaggactgg  
 240  
 gccgagcaca tctccctgcg tatgtacctg aactcgatgg tcacactcaa cgcgcccgcg  
 300  
 cacggacggg ttcgcgccct ggcccggcgg gtgttcaccc ccagcaagat caagaagatg  
 360  
 cagcccgcgg tggagaagcg gaccgacgag ctgatcaacg aactcgtcga gaggtccgcc  
 420

5

10

ggcggcgagc cggtcgacat cgtcgagctg ctggccatgc ccttccccgt cgcggtcac  
480

agcgacatgc tcggccttcc gtacgaggac ggcaagcgca cctgggagct ggccgacgac  
540

tggtcacggg tcttctccgg tgtctacacc gatgaggacc tggcggcggc cgacgccgcc  
600

gccgaggagc tgaccgggta ctccaatgac gtgatcaagg cggctccgcgc cgaacccaag  
660

gacgacctga tgtccgccct cgtccaggag gcggccaacg ggaagctcga cgaggaggag  
720

ctgatggcgc tcatacctctt cctgttcacg gcgggcttcg cggccaccac caacctcatc  
780

gcgaccggcg tgctcgcgct catcgagcac cccgacgagc tgaagcgctg gcgcgcggac  
840

aagagcatca ccccgacggc cgtcgaggag ctgctgcgtc acaccgcca caccacggcg  
900

tcgagccgtc tgacgacgcg ccccatcacc atcggcggca ccgacattcc cgagggcggt  
960

ctcgtcctgg ccttctctc cgccgccaac cgggaccggy cgcgcttccc cgaccgac  
1020

cgtctggacc tgagccgca caacggcgcc cacctgagct tcagcgccgg cggccacttc  
1080

tgcttcggcg gcagcctcgc ccgtatggag gccgccgacc tcttccgaa gctgatcgac  
1140

acgttcacga acatcgagct ggcgggcacc cccgggcgcc gcgccatcat gggcctgacc  
1200

ggttacacct cgcttccggt gactctcggc cgctga  
1236

<210> 27

<211> 411

<212> PRT

<213> *Streptomyces* sp. MP28-13

<400> 27

Met Thr Thr Glu Ala Thr Pro Thr Thr Gln Gln Ala Asp Glu Ala Leu  
1 5 10 15

Met Lys Leu Leu Ser Pro Pro Phe Pro Asp Asp Pro Phe Pro Ile Tyr  
20 25 30

Glu Thr Leu Gln Ser Val Asn Arg Val His Lys Ser Ala Leu Gly Ile  
35 40 45

Tyr Ala Leu Ser Gly Tyr Glu Glu Val Thr Glu Leu Leu Lys Met Pro  
50 55 60

5

10

ES 2 543 203 T3

Asp Val His Ser Gly Ala Arg Ala Ala Ala Gln Met Arg Glu Asp Trp  
 65 70 75 80  
 Ala Glu His Ile Ser Leu Arg Met Tyr Leu Asn Ser Met Val Thr Leu  
 85 90 95  
 Asn Ala Pro Asp His Gly Arg Val Arg Gly Leu Ala Ala Arg Val Phe  
 100 105 110  
 Thr Pro Ser Lys Ile Lys Lys Met Gln Pro Ala Val Glu Lys Arg Thr  
 115 120 125  
 Asp Glu Leu Ile Asn Glu Leu Val Glu Arg Ser Ala Gly Gly Glu Pro  
 130 135 140  
 Val Asp Ile Val Glu Leu Leu Ala Met Pro Phe Pro Val Ala Val Ile  
 145 150 155 160  
 Ser Asp Met Leu Gly Leu Pro Tyr Glu Asp Gly Lys Arg Thr Trp Glu  
 165 170 175  
 Leu Ala Asp Asp Trp Ser Arg Val Phe Ser Gly Val Tyr Thr Asp Glu  
 180 185 190  
 Asp Leu Ala Ala Ala Asp Ala Ala Ala Glu Glu Leu Thr Gly Tyr Phe  
 195 200 205  
 Asn Asp Val Ile Lys Ala Val Arg Ala Glu Pro Lys Asp Asp Leu Met  
 210 215 220  
 Ser Ala Leu Val Gln Glu Ala Ala Asn Gly Lys Leu Asp Glu Glu Glu  
 225 230 235 240  
 Leu Met Ala Leu Ile Leu Phe Leu Phe Thr Ala Gly Phe Ala Ala Thr  
 245 250 255  
 Thr Asn Leu Ile Ala Thr Gly Val Leu Ala Leu Ile Glu His Pro Asp  
 260 265 270  
 Glu Leu Lys Arg Trp Arg Ala Asp Lys Ser Ile Thr Pro Thr Ala Val  
 275 280 285  
 Glu Glu Leu Leu Arg His Thr Ala His Thr Thr Ala Ser Ser Arg Leu  
 290 295 300  
 Thr Thr Arg Pro Ile Thr Ile Gly Gly Thr Asp Ile Pro Glu Gly Val  
 305 310 315 320  
 Leu Val Leu Ala Leu Leu Ser Ala Ala Asn Arg Asp Pro Ala Arg Phe  
 325 330 335

Pro Asp Pro His Arg Leu Asp Leu Ser Arg Asp Asn Gly Ala His Leu  
 340 345 350

Ser Phe Ser Ala Gly Gly His Phe Cys Phe Gly Gly Ser Leu Ala Arg  
 355 360 365

Met Glu Ala Ala Asp Leu Phe Pro Lys Leu Ile Asp Thr Phe Thr Asn  
 370 375 380

Ile Glu Leu Ala Gly Thr Pro Gly Arg Arg Ala Ile Met Gly Leu Thr  
 385 390 395 400

Gly Tyr Thr Ser Leu Pro Val Thr Leu Gly Arg  
 405 410

<210> 28  
 <211> 10119  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

5

<400> 28

atgaacgccg aagtcgagga gatcgtcgag gcactgcgtg gctcgtctggt cgagaacgag  
 60

aggctgcggc aggacaacga cagcctgcgc gccgccgca gcgtaaccac cgaaccatc  
 120

gcgatcatcg gcatggcctg ccggtacccc ggtggcgtcg gatccccga ggagctgtgg  
 180

gcgctggtcg aggaaggccg tgacgcgatc agcccgttcc cggacgaccg cggctgggac  
 240

atgagcgcgc tgtacgacc cagacccgga aagcccggca agacgtacgc gcgcgagggc  
 300

ggattcctgc acgacgccgc gcagttcgac cccgagttct tcgggatcag cccgcgcgag  
 360

gccctcacca tggaccccca gcagcggctc ctgctggaga tcacctggga gtcgatggag  
 420

cgcgccggac tcgaccggc gtcgctgcgc ggcagccgca ccggtgtctt cgccggcgcg  
 480

atgtaccacg actacggcat caccagcagc gacggcagcc tcgtctccgg acgcgtcgcg  
 540

tacaccctcg ggctggaagg cccggcggtc accgtcgaca cggcctgctc gtcgctcgtg  
 600

gtcgcactcc agtgggcgtc gcaggccctg cgctcgggag agtgcaccct ggcgctggcc  
 660

ggcggcgtca gcgtcatggc cacgcccag acgttcatcg agttcagcga gcagcgcgga  
 720

ctctcggccc acggccgctg ccgctcgttc gccgcctccg ccgacgggac gggctgggccc  
 780

10

gagggcgccg gcgtcctgct gctggagaag ctgtccgacg cccgccgcaa cgggcacccc  
840  
gtactggccg tcgtgcgcgg ctcggcggtc aaccaggacg gcgctccaa cggcttcagc  
900  
gcccccaacg gcccctcgca gcgccgtgtc atccagcagg cgctcaccgc ggccggactg  
960  
acgaccgccc atgtcgacgt catggagggc cacggcaccg gcacctcgct cggcgacccc  
1020  
atcgaggcgc aggcactgct cggcacgtac ggacagggcc gcgaagaacc gctgtggctg  
1080  
ggttcgatca agtccaacat cggtcacacg caggcccgcc ccggtgtcgc gggcgatc  
1140  
aagatggctg aggccatccg acgcggcacg ctgccgaaga cactgcacgt cgacgaggcg  
1200  
tcgccccagg tggactggga ggccggaaac gtccgcctgc tcaccgaggc gcgtgcctgg  
1260  
cccgatgccg accggccgcg ccgcgccgcg gtgtcctcgt tcggcgtcag cggcaccaac  
1320  
gcccattgta tcatcgagca ggcgcagccc gaccccgct ccgactccga gccggagccc  
1380  
gccgcaccgc gcccgctccac cgacgtacc ctgaccgtgc cgctgccgct gccgctgtcg  
1440  
gccggcaccg agaccgcgat gcgcgcccag gcccggcagt tggccgacca cttgaggacc  
1500  
acgcctgacc tcgaccctgt ggacatggcg tactcctcgc ccaccgcccg cggcgcgctc  
1560  
accgcccgcg ccgtcgtggt cggccacgac cgcgacgaga tcctcggcgc gctgaccgcg  
1620  
ttggcggacg gcgtccccct cggcggcgcc gcgcggtcga ccgggctgac cgccttcctc  
1680  
ttaccggcc agggcagcca gcgcctcggc atgggcccgc acctgcacgc ggcgttcccc  
1740  
gtgttcgcac gcgccttcga cgaggtgtgc tccgcgctcg accccgccgt gcgcgaggtg  
1800  
atgtggggcg acgaagaggc cctgcggcgt accgagttca cccagcccgc gatcttcgca  
1860  
ctccagatcg ccctgttccg gctgctcga tcgtgggggtg tacggcccga ctctgctcgc  
1920  
ggacactcca tcggcgaact cggcccgcc catgtggccg gcgtgttctc cctgcccgac  
1980  
gccgcctcgc tgatcaccgc gcgggcacgg ctgatgcagg cgctgcccc gggcggggcg  
2040  
atggtggcgg tcgaggccgc cgaggaggag gtcgtcccgc tgcttcggga cggcgaggga  
2100

atcgccgccc tcaacggccc cgcctccgtg gtgctctcgg gcaccgagga ggccgtggac  
 2160  
 gcggtcgtcg aacagctcgg cgcgcgcagg accaaccggc tgaagggtgc gcacgcgttc  
 2220  
 cactcgacgc tgatggaccc gatgctcgac gacttccgcc gtgtcgcgga gcgcgtcggc  
 2280  
 tacgcgggagc ccggccttcc cgtcgtggcc aacggcgatg tcaccaccgc cgcgtactgg  
 2340  
 gtggggcatg tccgtgacac cgtccgcttc gccgacgccg tgaccgggct ggaatccgaa  
 2400  
 ggctcacc ccgtacgtcga actgggcccc gacggcatcc tcaccgcgat ggcccgccag  
 2460  
 tgcctcacga ccaccgccga caccgccgta ctcgtccccg ccctgcgccg taacgaaacc  
 2520  
 ggaccgctc cegtctcac cggcctcggc ggctgcaca ccgcgggct gaaagtcgac  
 2580  
 tgggcgggtg tcttcgacgg ccgcggcgcc cggcgcgtcg acctgccac ctacccttc  
 2640  
 cagcgcgcc gttactggtt cgacaagcgc ggctgggcg gcgacgtcac ctccgccgt  
 2700  
 ctcgaccggc ccgaccacc gctgctcgtt gcgatggtgc acctgcccg ctccgacgt  
 2760  
 gtggtgttca ccggacggct gtccaccggg gccaccctt ggctgtccga ccacaccgtg  
 2820  
 atgggctccg tctgctgcc cggcaccgcg tacgtggaac tcgcggtgcg cgcgggggac  
 2880  
 caggtgggat ggaaccgcgt cgaggaactc aacgtcgcgg cggcgtgtt cctgcccgag  
 2940  
 cacggcggcg tccacatcca ggtcgcctc gacgcgccc acgcgtcggg actccgccct  
 3000  
 gtgaggggtg tctcccgagc cgacgacgcc ccgctcgacc gcgaatggat cctgcacgcc  
 3060  
 gagggcttcc tggcgcgccga cggcggtag ccggccaccg acctcacggt gtggcctccg  
 3120  
 cgcgacgcc agcccctcgc ggtcgaagga ctctacgaac ggctggagta cgggccgacg  
 3180  
 ttccagggac tgcgcgcggg atggcggcgc ggcgacgagc tgttcgccga gacagctctg  
 3240  
 cccgaggggtg cggacgccgg cggcttcggc ctccatccc ccctgttcga cgcggcgttg  
 3300  
 cacgtgctcg acctggccgg cgaggacgcg aaggctcttc cgttcacctg gtcggacgtc  
 3360  
 accctgcacg cggagggcgc cacgaccgcg cgggtgagcc tgcgtgtccg gggagacaag  
 3420



tccgtgtccc tggagctggc cgacgccatg ggacggcccgg tcgcctcggg cggatcgctg  
 3480  
 acgctgcgcc ccgtcacggc ggacggcctc gcgccggccg cggcccgggt cgcgaacgcc  
 3540  
 ctcttccggg tggactgggt cccggccggc gaggtccggg caccggctga gaacaccgag  
 3600  
 gtgagcgttc accactgccc gccgacgacc ggcggcacac cggccgccgt acgcgcgggtg  
 3660  
 accaccggcg tactggccgc cgtccagggc gcagtcgacg gcggcaccgc cctcgtcgtc  
 3720  
 gtcaccgacg gcgccaccga cggctccgac ctcggccacg ccgccgctg ggggctggta  
 3780  
 cgcgccggcc agggcgagca tcccggccgc ttcttcctcg tcgacaccga cgtcccggtc  
 3840  
 gaccccggcc gcgtcgtcgc gatcggcgaa cccgaactgc gcgtgagcgg cggcgagaca  
 3900  
 cgggtgcccc ggctggtcgg cgtaccgctc gattcgaccg cctccacctg ggacaccgaa  
 3960  
 cgcaccgtcc tgatcaccgg cgggaccggg gccctcgggg ccgccgctgc ccggcacctc  
 4020  
 gtcacccgcc acgaggtacg gcggctgctg ctgaccagcc gacggggccc gcaggcgccc  
 4080  
 ggagccggcc agctcggcca ggaactgacc gggctcggcg ccgaggtcga ggtggccgcg  
 4140  
 tgcgacgccc ccgaccgca cgccctggcc acgctgctcg acggccggac catcggcggg  
 4200  
 gtcgtacacg ccgcgggcgt cctggatgac ggcgtgatcc tctcgatgac gcccgagcgc  
 4260  
 gtcgaccacg tgctccgtcc gaaggcggac gcggcctggc atctgcacga actcaccggc  
 4320  
 gacatggggc tgaccgcttt cgtgctgttc tcgtccgctg cgggcgtact cggtgcccc  
 4380  
 ggacagggca actacgccgc cgccagcacc ctgctcgacg gcctcgcgcg gcacggcac  
 4440  
 gccgccggac tgccggcgct gtcactggcc tggggctcgt gggccggaga gggcatggcg  
 4500  
 gacggctctg cctcggtcgg catgcgctcc ctggcaccgg aggagggcct cgccctgctc  
 4560  
 gacgccggcc ccggcgtggc ggagcccgta ctggtgcccg tccggttcga cctcgtcgc  
 4620  
 ttcgactcgc cgccgccat catgcgcggc ctcgtccgcg gccggtcgcg ccgtgtcctc  
 4680  
 gacaacgacg cgtccgcgac cggcgtcctg cggcagcgcc tggcgggact cggcgacgcg  
 4740

gaacgcgccg acgaactcct cgctctcgta cgctcccagg cggcgatggt cctgcgccat  
 4800  
 gccggagccg aggcggtcga cccggagcgg gccttccgcg acctcggatt cgactccctc  
 4860  
 acggcgatcg aactgcgcaa cctgctgggt gccgccaccg gactgcgcct gcccgccacc  
 4920  
 ctcgtcttcg actaccccac ccccgtcgtc ctggcgggcc acctgctgcg ggagctctcc  
 4980  
 ggagccgtgg agtcggctcc ggtcgcgtec gtcgtgcgcc cggcggacga cgagccgatc  
 5040  
 gccattgtct cgatggcgctg ccgctacccg ggtggcgctgg actcgcgccga ggggttgctg  
 5100  
 cggctcgtcg acgaggggtg cgacgcgata tcggagtcc cggccgaccg tggctggggc  
 5160  
 gtggaggaca tctacgacct cgagcccgga attccgggga agacctatgt gcgcgacggc  
 5220  
 ggattcctgc acgacgccac acagttcgac gccgatttct tcggcatcag cccgcgcgag  
 5280  
 gccctggaca tggacctgca gcagcggctg ctgctggaga cctcgtggga ggcgctggag  
 5340  
 cgtgccggga tcgcacctac cacgctgaag ggcagcccga ccggtgtgtt cgccgggggtg  
 5400  
 atgtaccacg actacccccg cggcaccggc ggcggcagcc tcgtgtcggg ccgggtggcc  
 5460  
 tacaccctcg gtctcgaagg ccccgcggtg agcgtggaca cggcatgctc gtcgtccctg  
 5520  
 gtggccctgc actgggcggc gcaggcactg cgttccggcg agtgctcgtc cgccctcgtc  
 5580  
 ggcggcgtga ccgcatggg aacaccgcg tcgttcatcg acttcagcga gcagcgcggc  
 5640  
 ctggcccgcg acggccgctg caagtccttc tcgtctcca ccgacggcac cgggtggggc  
 5700  
 gagggcgcgg gcgtcctggt ggtggagcgt ctgtcggagg cgcgtcggct ggggcatccg  
 5760  
 gtgctcgcgg tggcgcgcgg gagcgcgctc aaccaggacg gcgccagcaa cggcatcacc  
 5820  
 gcgccgaacg gcccctcca gcgtcgcgtg atcaagcagg cgctggcgaa ggcgggtctg  
 5880  
 tcgacggcgg atgtggacgc ggtcgaggcg catggcacgg gcacgacgct gggcgacctg  
 5940  
 atcgaggccc aggccctggt ggagacgtac ggtcaggatc gccccgaagg gcggccgctg  
 6000  
 tggctgggtt cgatcaagtc gaacatcgtt catacgcagg cggcggcggg tgtggcggg  
 6060

atcatcaaga tggtcgaggc gatgcgccac ggcaggctgc ccaagacgct gcatgtggat  
6120  
gagccgacga agcaggtgga ctgggacgcg ggtgaggtgc ggctgctgac ggaggcgcgt  
6180  
gagtggccga gcgagggccg tccgcgccgg gcagccgtgt cctccttcgg aatcagtggc  
6240  
acgaacgccc acgtgatcgt cgaggaagtc gtgccggtg ctgaagtggg ggtggagcgg  
6300  
cgggagttgc cggtggcgcc ggtggtggtg tcggggaaga ccccgccggc gctggaggcg  
6360  
cagatcggcc gcttcggcga actggccgcg aacggcgacc cgctggacgt cgcgtactcg  
6420  
gccgcgacag gccgggcccgc gctggaacac cgcgcggtgc tcatcgggtc ggagacggtc  
6480  
acgggcgaga ccggtgtggg caaggtggcg ttcctgttca ccgggcaggg gagtcaacgc  
6540  
cttggtatgg ggcgggagtt gtacgagacg ttccccgcct tcgcatcggc gttcgacgag  
6600  
gtgtgtggcg tgctcgatcc cgctgtgcgt gaggtgatgt ggggtgatga ggaggctctg  
6660  
ggccgtacgg agttcaccca gcccgcgatc ttcgctcttg aggtggcgtt gttccggtg  
6720  
gtggagtcgt ggggggtcaa gccggacttc ctggtgggcc attcgatcgg tgagttggcg  
6780  
gcggcgcgatg tggcgggggt gttcggctctg gaggatgcgg gccggttgat ctcggcgcgt  
6840  
gggcggttga tgcaggcgtt gccggcgggt ggggcgatgg tcgcatcca ggccacggag  
6900  
gaggaggtgg ttccgcatct cagcggcctg gtgagtgtg ccgcggtcaa cagcctttct  
6960  
tcggtggtga tttcgggcca ggagaaggcg gtgacggcgg tcgaggagcg gttcaccgac  
7020  
cgcaagacga ctcggctgaa ggtgtcgcac gcgttccact cgccgctgat ggacccgatg  
7080  
ctggacgact tccgcaaggt cgcggagagc gccacctacc gggagccgac catccggctg  
7140  
accaaggacg tcggttcggc cgagtactgg gtggggcatg tgcgggacgc ggtgcgtttc  
7200  
gcggacgatg tgcgctattt gcaggacgag ggtgtgacgc ggttcctgga gatcggctccg  
7260  
gatggtgtgc tgacggcgat ggcgggacag agtgccgacg gcaccctggc gccacgctg  
7320  
cggcgtgacc gccccgaggt ggagagtgtg ttcgccggtg tgggcccgtt gttcgcggcc  
7380

ggtgtggcgg tggactggga cgcggtgttc gacgggcgtg gtgcgcggcg ggtcgatctg  
 7440  
 cccacgtatc ccttccagcg caagcgttac tggctgatcg agcagtcgac ggcggcagcc  
 7500  
 ggtgccgacg cggtcgatca cccccctgtc acctcgggca tcgggctgcc cgacaccggc  
 7560  
 ggtgtggtgt tcaccggacg tctctccctc gacaccacc cctggctcgc cgaccacgac  
 7620  
 gtactgggta ccctgctgct gcctggcacc gggctggctg aactcgcgct gcaagccgcc  
 7680  
 gcacaggtgg actgcggcac ggtcgacgag ctgaccctgg aggcgccgct cgtcgttccc  
 7740  
 gagaagggcg cggtcgccgt gcgcgttctc gtgggcggcc ccgacgactc cgagtcacgc  
 7800  
 accgtcgaga tctactcgtc cctggacgac gagatctgga cccgcaacgc ggccggtgcc  
 7860  
 ctctcccgt cggccgtatc tccctcgtcc gacctgatcc agtggccgcc gatcggcgcc  
 7920  
 acgccgctcc cggtgacgg cgccctacgag cggctcctcg cccgcggtta cgactacggc  
 7980  
 cccacgttcc aggggctgaa agcggcctgg cgggacggtg acggcgtcgt gttcgcgag  
 8040  
 gtctcgtccc cggagggaaac ggaggctgcc aggttcggcc tgcattccggc gttgctggac  
 8100  
 gcggcgatgc acgtgggtct catcgaggaa ggcgccgcca ccgacgcgcc ggagctgccg  
 8160  
 ttctcctgga acggtgtcac gctccaccgc gccggcgcct ccgcgctgcg cgtccgtctg  
 8220  
 tcgaacccgg agggcggcga cggcacggag gttctggctg ccgacggaac cggcgcgccc  
 8280  
 gtgctctccg tcgcgagcct gacctcgcgg cccgtgtcgg ccgagcagct acgggcccac  
 8340  
 ggcgaccacc gcgagtcgct gttcgcctc acctggacca aggccggcga cgtgcccgta  
 8400  
 ctggagacac ccgtcgtggt gtacgaggtc ccgcgcgccg aaggtgacac atccgaggcc  
 8460  
 gcgcacgcgg tcgccgacga ggtgctggcg cgcgtccagg aatggctggc ggacaagggc  
 8520  
 aggagcgatg agaagctcgc cgtcgtcacc cgtcgcgcgg tgccgatcga gggcggagac  
 8580  
 gtcgatctga gccaggcacc cgtgtggggg ctggtccggg cggcggcggc ggagaatccc  
 8640  
 ggccggttcc ttctgctcga cctgggcgac ggggccaccg ttccgccctt cgtcgacgca  
 8700

cccgaggtgg cggttcgagg cggcgaactt ctcgtacccg ccctcaccgg cgtacccgcc  
8760  
tcggcggtgg acgccggacg cgacccgtgg gagtcgtcgc cgaccgtgct gatcaccggc  
8820  
ggcaccagcg gtctcgggtg cctgggtcgc cgccacctgg tgacggagca cggcatccgg  
8880  
catctggtcc tgaccagccg tcggggcggg agcgcaccgg gcgcggcgga actgcgcgcc  
8940  
gaactgaccg gccacggcgc gcgggtggac atcgaggcgt gcgacgtggc cgaccgcgcc  
9000  
gcgctggccg cgggtgctcga ccggcatccc gtcggggcgg tcgtccacgc ggccggtgtc  
9060  
gtggacaacg ggctgggtcc cgggctgacg ccggagcgtg tggacacggt cctgcggccc  
9120  
aaggtggacg gcgcctggca tctgcacgag ctgactgccg accgcgacct gtccgcgttc  
9180  
gtcctgttct cctcgatggg cggcctgctg ctggcgccgg gccagggcaa ctacgccggg  
9240  
gcgaacgttt tcctggacgc gctcgcacc caccgccacc aggccggcct gccggtgacc  
9300  
tcactggcgt tcggtctgtg ggaagccgag accgggctcg gtgagctgac cgacgccgac  
9360  
cgcaagcgca tgctccgcat gggcctgccc gccctgtcgc agaaggaggg cctggcgctg  
9420  
ctggacgacg cgttgcgtac gggggagccc gcgctggcgc cgttccgcct ggacacaggg  
9480  
gccctgcgga cccgcgccgt cgaccaactg ccgtcgtcgc tgcgtggtct ggtgccaca  
9540  
tcgcgccgtc gcgccggggg cgcgggcgcg gccggcgcg gggacggcgg cgcgggactg  
9600  
cgacggaccc tggcggcggc ggccgacgaa gccgcgcgcg acagcatcct ggtcgagttg  
9660  
gtccgtaccc acgtggcggc ggtgctcggg cacgacggcg tggacgcggt ccgctccgac  
9720  
cgggcgttca aggacctcg attcgactcg ctcaccgccg tggaaactgc caacacgctg  
9780  
cgttcggcca cggggctgaa gctgccggcc accctggtct tcgaccatcc cacccttcc  
9840  
gccgtcgcgg agtccctgcg cgaacagctg accggcgacc aggagcccgc cgggtcaccg  
9900  
ctggaggcgg aactggcccg tctggaagcg gcgttggcgt cggtgacgcc cgacgaggag  
9960  
acgttcggcc gcgtcggcga ccggtgcgc gccctggcgg ccggctggac cgagaccac  
10020

aggcgggacg agagcgacga ggccgaactc gggtcgctct cggccgatga gctgttcgac  
10080

gtcctcgacg acgagctcga tacgcgggtcc gcgtcgtag  
10119

<210> 29  
<211> 3372  
<212> PRT  
<213> *Streptomyces* sp. MP28-13  
  
<400> 29

5

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

10

Ala Leu Arg Ser Gly Glu Cys Thr Leu Ala Leu Ala Gly Gly Val Ser  
 210 215 220

Val Met Ala Thr Pro Glu Thr Phe Ile Glu Phe Ser Glu Gln Arg Gly  
 225 230 235 240

Leu Ser Ala Asp Gly Arg Cys Arg Ser Phe Ala Ala Ser Ala Asp Gly  
 245 250 255

Thr Gly Trp Ala Glu Gly Ala Gly Val Leu Leu Leu Glu Lys Leu Ser  
 260 265 270

Asp Ala Arg Arg Asn Gly His Pro Val Leu Ala Val Val Arg Gly Ser  
 275 280 285

Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Phe Ser Ala Pro Asn Gly  
 290 295 300

Pro Ser Gln Arg Arg Val Ile Gln Gln Ala Leu Thr Ala Ala Gly Leu  
 305 310 315 320

Thr Thr Ala Asp Val Asp Val Met Glu Gly His Gly Thr Gly Thr Ser  
 325 330 335

Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln  
 340 345 350

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

His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Glu  
 370 375 380

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

Ser Pro Gln Val Asp Trp Glu Ala Gly Asn Val Arg Leu Leu Thr Glu  
 405 410 415

Ala Arg Ala Trp Pro Asp Ala Asp Arg Pro Arg Arg Ala Ala Val Ser  
 420 425 430

Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Ile Ile Glu Gln Ala  
 435 440 445

Gln Pro Asp Pro Ala Ser Asp Ser Glu Pro Glu Pro Ala Ala Pro Arg  
 450 455 460

Pro Ser Thr Asp Val Pro Leu Thr Val Pro Leu Pro Leu Pro Leu Ser





Ser His Ala Phe His Ser Thr Leu Met Asp Pro Met Leu Asp Asp Phe  
 740 745 750  
 Arg Arg Val Ala Glu Arg Val Ala Tyr Ala Glu Pro Gly Leu Pro Val  
 755 760 765  
 Val Ala Asn Gly Asp Val Thr Thr Ala Ala Tyr Trp Val Gly His Val  
 770 775 780  
 Arg Asp Thr Val Arg Phe Ala Asp Ala Val Thr Arg Leu Glu Ser Glu  
 785 790 795 800  
 Gly Val Thr Arg Tyr Val Glu Leu Gly Pro Asp Gly Ile Leu Thr Ala  
 805 810 815  
 Met Ala Arg Gln Cys Leu Thr Thr Thr Ala Asp Thr Ala Val Leu Val  
 820 825 830 835  
 Pro Ala Leu Arg Arg Asn Glu Thr Gly Pro Val Ala Val Leu Thr Ala  
 835 840 845  
 Leu Gly Gly Leu His Thr Ala Gly Leu Lys Val Asp Trp Ala Gly Val  
 850 855 860  
 Phe Asp Gly Arg Gly Ala Arg Arg Val Asp Leu Pro Thr Tyr Pro Phe  
 865 870 875 880  
 Gln Arg Ala Arg Tyr Trp Phe Asp Lys Arg Gly Leu Gly Gly Asp Val  
 885 890 895  
 Thr Ser Ala Gly Leu Asp Arg Pro Asp His Pro Leu Leu Gly Ala Met  
 900 905 910  
 Val His Leu Pro Gly Ser Asp Gly Val Val Phe Thr Gly Arg Leu Ser  
 915 920 925  
 Thr Gly Ala His Pro Trp Leu Ser Asp His Thr Val Met Gly Ser Val  
 930 935 940  
 Leu Leu Pro Gly Thr Ala Tyr Val Glu Leu Ala Val Arg Ala Gly Asp  
 945 950 955 960  
 Gln Val Gly Trp Asn Arg Val Glu Glu Leu Asn Val Ala Ala Pro Leu  
 965 970 975  
 Phe Leu Pro Glu His Gly Gly Val His Ile Gln Val Ala Val Asp Ala  
 980 985 990  
 Pro Asp Ala Ser Gly Leu Arg Pro Val Arg Val Phe Ser Arg Ala Asp

995					1000					1005				
Asp	Ala	Pro	Leu	Asp	Arg	Glu	Trp	Ile	Leu	His	Ala	Glu	Gly	Phe
	1010					1015					1020			
Leu	Ala	Pro	Asp	Ala	Gly	Glu	Pro	Ala	Thr	Asp	Leu	Thr	Val	Trp
	1025					1030					1035			
Pro	Pro	Arg	Asp	Ala	Glu	Pro	Leu	Ala	Val	Glu	Gly	Leu	Tyr	Glu
	1040					1045					1050			
Arg	Leu	Glu	Tyr	Gly	Pro	Thr	Phe	Gln	Gly	Leu	Arg	Ala	Gly	Trp
	1055					1060					1065			
Arg	Arg	Gly	Asp	Glu	Leu	Phe	Ala	Glu	Thr	Ala	Leu	Pro	Glu	Gly
	1070					1075					1080			
Ala	Asp	Ala	Gly	Gly	Phe	Gly	Leu	His	Pro	Ala	Leu	Phe	Asp	Ala
	1085					1090					1095			
Ala	Leu	His	Val	Leu	Asp	Leu	Ala	Gly	Glu	Asp	Ala	Lys	Val	Leu
	1100					1105					1110			
Pro	Phe	Thr	Trp	Ser	Asp	Val	Thr	Leu	His	Ala	Glu	Gly	Ala	Thr
	1115					1120					1125			
Thr	Ala	Arg	Val	Ser	Leu	Arg	Val	Arg	Gly	Asp	Lys	Ser	Val	Ser
	1130					1135					1140			
Leu	Glu	Leu	Ala	Asp	Ala	Met	Gly	Arg	Pro	Val	Ala	Ser	Val	Gly
	1145					1150					1155			
Ser	Leu	Thr	Leu	Arg	Pro	Val	Thr	Ala	Asp	Gly	Leu	Ala	Pro	Ala
	1160					1165					1170			
Ala	Ala	Arg	Val	Ala	Asn	Ala	Leu	Phe	Arg	Val	Asp	Trp	Val	Pro
	1175					1180					1185			
Ala	Gly	Glu	Val	Arg	Ser	Pro	Ala	Glu	Asn	Thr	Glu	Val	Ser	Val
	1190					1195					1200			
His	His	Cys	Pro	Pro	Thr	Thr	Gly	Gly	Thr	Pro	Ala	Ala	Val	Arg
	1205					1210					1215			
Ala	Val	Thr	Thr	Gly	Val	Leu	Ala	Ala	Val	Gln	Gly	Ala	Val	Asp
	1220					1225					1230			
Gly	Gly	Thr	Ala	Leu	Val	Val	Val	Thr	Asp	Gly	Ala	Thr	Asp	Gly
	1235					1240					1245			

Ser Asp Leu Gly His Ala Ala Ala Trp Gly Leu Val Arg Ala Ala  
 1250 1255 1260

Glu Gly Glu His Pro Gly Arg Phe Phe Leu Val Asp Thr Asp Ala  
 1265 1270 1275

Pro Val Asp Pro Ala Arg Val Val Ala Ile Gly Glu Pro Glu Leu  
 1280 1285 1290

Arg Val Ser Gly Gly Glu Thr Arg Val Pro Arg Leu Val Gly Val  
 1295 1300 1305

Pro Leu Asp Ser Thr Ala Ser Thr Trp Asp Thr Glu Arg Thr Val  
 1310 1315 1320

Leu Ile Thr Gly Gly Thr Gly Ala Leu Gly Ala Ala Val Ala Arg  
 1325 1330 1335

His Leu Val Thr Arg His Glu Val Arg Arg Leu Leu Leu Thr Ser  
 1340 1345 1350

Arg Arg Gly Pro Gln Ala Pro Gly Ala Ala Glu Leu Ala Glu Glu  
 1355 1360 1365

Leu Thr Gly Leu Gly Ala Glu Val Glu Val Ala Ala Cys Asp Ala  
 1370 1375 1380

Ala Asp Arg Asp Ala Leu Ala Thr Leu Leu Asp Gly Arg Thr Ile  
 1385 1390 1395

Gly Gly Val Val His Ala Ala Gly Val Leu Asp Asp Gly Val Ile  
 1400 1405 1410

Leu Ser Met Thr Pro Glu Arg Val Asp His Val Leu Arg Pro Lys  
 1415 1420 1425

Ala Asp Ala Ala Trp His Leu His Glu Leu Thr Arg Asp Met Gly  
 1430 1435 1440

Leu Thr Ala Phe Val Leu Phe Ser Ser Val Ala Gly Val Leu Gly  
 1445 1450 1455

Ala Pro Gly Gln Gly Asn Tyr Ala Ala Ala Ser Thr Leu Leu Asp  
 1460 1465 1470

Gly Leu Ala Arg His Arg His Ala Ala Gly Leu Pro Ala Leu Ser  
 1475 1480 1485

Leu Ala Trp Gly Pro Trp Ala Gly Glu Gly Met Ala Asp Gly Leu

1490						1495						1500			
Ala	Ser	Val	Gly	Met	Arg	Ser	Leu	Ala	Pro	Glu	Glu	Gly	Leu	Ala	
1505						1510					1515				
Leu	Leu	Asp	Ala	Ala	Ala	Gly	Val	Ala	Glu	Pro	Val	Leu	Val	Pro	
1520						1525					1530				
Val	Arg	Phe	Asp	Leu	Ala	Ala	Phe	Asp	Ser	Pro	Pro	Pro	Ile	Met	
1535						1540					1545				
Arg	Gly	Leu	Val	Arg	Gly	Arg	Ser	Arg	Arg	Val	Leu	Asp	Asn	Asp	
1550						1555					1560				
Ala	Ser	Ala	Thr	Gly	Val	Leu	Arg	Gln	Arg	Leu	Ala	Gly	Leu	Gly	
1565						1570					1575				
Asp	Ala	Glu	Arg	Ala	Asp	Glu	Leu	Leu	Ala	Leu	Val	Arg	Ser	Gln	
1580						1585					1590				
Ala	Ala	Met	Val	Leu	Arg	His	Ala	Gly	Ala	Glu	Ala	Val	Asp	Pro	
1595						1600					1605				
Glu	Arg	Ala	Phe	Arg	Asp	Leu	Gly	Phe	Asp	Ser	Leu	Thr	Ala	Ile	
1610						1615					1620				
Glu	Leu	Arg	Asn	Leu	Leu	Gly	Ala	Ala	Thr	Gly	Leu	Arg	Leu	Pro	
1625						1630					1635				
Ala	Thr	Leu	Val	Phe	Asp	Tyr	Pro	Thr	Pro	Val	Val	Leu	Ala	Gly	
1640						1645					1650				
His	Leu	Leu	Arg	Glu	Leu	Ser	Gly	Ala	Val	Glu	Ser	Ala	Pro	Val	
1655						1660					1665				
Ala	Ser	Val	Val	Arg	Pro	Ala	Asp	Asp	Glu	Pro	Ile	Ala	Ile	Val	
1670						1675					1680				
Ser	Met	Ala	Cys	Arg	Tyr	Pro	Gly	Gly	Val	Asp	Ser	Pro	Glu	Gly	
1685						1690					1695				
Leu	Trp	Arg	Leu	Val	Asp	Glu	Gly	Val	Asp	Ala	Ile	Ser	Glu	Phe	
1700						1705					1710				
Pro	Ala	Asp	Arg	Gly	Trp	Gly	Val	Glu	Asp	Ile	Tyr	Asp	Pro	Glu	
1715						1720					1725				
Pro	Gly	Ile	Pro	Gly	Lys	Thr	Tyr	Val	Arg	Asp	Gly	Gly	Phe	Leu	
1730						1735					1740				

His Asp Ala Thr Gln Phe Asp Ala Asp Phe Phe Gly Ile Ser Pro  
 1745 1750 1755  
 Arg Glu Ala Leu Asp Met Asp Pro Gln Gln Arg Leu Leu Leu Glu  
 1760 1765 1770  
 Thr Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Ala Pro Thr Thr  
 1775 1780 1785  
 Leu Lys Gly Ser Pro Thr Gly Val Phe Ala Gly Val Met Tyr His  
 1790 1795 1800  
 Asp Tyr Pro Gly Gly Thr Gly Gly Gly Ser Leu Val Ser Gly Arg  
 1805 1810 1815  
 Val Ala Tyr Thr Leu Gly Leu Glu Gly Pro Ala Val Ser Val Asp  
 1820 1825 1830  
 Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Trp Ala Ala Gln  
 1835 1840 1845  
 Ala Leu Arg Ser Gly Glu Cys Ser Leu Ala Leu Val Gly Gly Val  
 1850 1855 1860  
 Thr Val Met Gly Thr Pro Arg Ser Phe Ile Asp Phe Ser Glu Gln  
 1865 1870 1875  
 Arg Gly Leu Ala Ala Asp Gly Arg Cys Lys Ser Phe Ser Ser Ser  
 1880 1885 1890  
 Thr Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Val Leu Val Val  
 1895 1900 1905  
 Glu Arg Leu Ser Glu Ala Arg Arg Leu Gly His Pro Val Leu Ala  
 1910 1915 1920  
 Val Val Arg Gly Ser Ala Leu Asn Gln Asp Gly Ala Ser Asn Gly  
 1925 1930 1935  
 Ile Thr Ala Pro Asn Gly Pro Ser Gln Arg Arg Val Ile Lys Gln  
 1940 1945 1950  
 Ala Leu Ala Lys Ala Gly Leu Ser Thr Ala Asp Val Asp Ala Val  
 1955 1960 1965  
 Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Ala  
 1970 1975 1980  
 Gln Ala Leu Leu Glu Thr Tyr Gly Gln Asp Arg Pro Glu Gly Arg

ES 2 543 203 T3

1985					1990					1995				
Pro	Leu	Trp	Leu	Gly	Ser	Ile	Lys	Ser	Asn	Ile	Gly	His	Thr	Gln
	2000					2005					2010			
Ala	Ala	Ala	Gly	Val	Ala	Gly	Ile	Ile	Lys	Met	Val	Glu	Ala	Met
	2015					2020					2025			
Arg	His	Gly	Arg	Leu	Pro	Lys	Thr	Leu	His	Val	Asp	Glu	Pro	Thr
	2030					2035					2040			
Lys	Gln	Val	Asp	Trp	Asp	Ala	Gly	Glu	Val	Arg	Leu	Leu	Thr	Glu
	2045					2050					2055			
Ala	Arg	Glu	Trp	Pro	Ser	Glu	Gly	Arg	Pro	Arg	Arg	Ala	Ala	Val
	2060					2065					2070			
Ser	Ser	Phe	Gly	Ile	Ser	Gly	Thr	Asn	Ala	His	Val	Ile	Val	Glu
	2075					2080					2085			
Glu	Val	Val	Pro	Val	Ala	Glu	Val	Val	Val	Glu	Arg	Arg	Glu	Leu
	2090					2095					2100			
Pro	Val	Ala	Pro	Val	Val	Val	Ser	Gly	Lys	Thr	Pro	Ala	Ala	Leu
	2105					2110					2115			
Glu	Ala	Gln	Ile	Gly	Arg	Phe	Gly	Glu	Leu	Ala	Ala	Asn	Gly	Asp
	2120					2125					2130			
Pro	Leu	Asp	Val	Ala	Tyr	Ser	Ala	Ala	Thr	Gly	Arg	Ala	Ala	Leu
	2135					2140					2145			
Glu	His	Arg	Ala	Val	Leu	Ile	Gly	Ser	Glu	Thr	Val	Thr	Gly	Glu
	2150					2155					2160			
Thr	Gly	Val	Gly	Lys	Val	Ala	Phe	Leu	Phe	Thr	Gly	Gln	Gly	Ser
	2165					2170					2175			
Gln	Arg	Leu	Gly	Met	Gly	Arg	Glu	Leu	Tyr	Glu	Thr	Phe	Pro	Ala
	2180					2185					2190			
Phe	Ala	Ser	Ala	Phe	Asp	Glu	Val	Cys	Gly	Val	Leu	Asp	Pro	Ala
	2195					2200					2205			
Val	Arg	Glu	Val	Met	Trp	Gly	Asp	Glu	Glu	Ala	Leu	Gly	Arg	Thr
	2210					2215					2220			
Glu	Phe	Thr	Gln	Pro	Ala	Ile	Phe	Ala	Leu	Glu	Val	Ala	Leu	Phe
	2225					2230					2235			

Arg Leu Val Glu Ser Trp Gly Val Lys Pro Asp Phe Leu Val Gly  
 2240 2245 2250  
 His Ser Ile Gly Glu Leu Ala Ala Ala His Val Ala Gly Val Phe  
 2255 2260 2265  
 Gly Leu Glu Asp Ala Gly Arg Leu Ile Ser Ala Arg Gly Arg Leu  
 2270 2275 2280  
 Met Gln Ala Leu Pro Ala Gly Gly Ala Met Val Ala Ile Gln Ala  
 2285 2290 2295  
 Thr Glu Glu Glu Val Val Pro His Leu Ser Gly Leu Val Ser Val  
 2300 2305 2310  
 Ala Ala Val Asn Ser Leu Ser Ser Val Val Ile Ser Gly Glu Glu  
 2315 2320 2325  
 Lys Ala Val Thr Ala Val Ala Glu Arg Phe Thr Asp Arg Lys Thr  
 2330 2335 2340  
 Thr Arg Leu Lys Val Ser His Ala Phe His Ser Pro Leu Met Asp  
 2345 2350 2355  
 Pro Met Leu Asp Asp Phe Arg Lys Val Ala Glu Ser Ala Thr Tyr  
 2360 2365 2370  
 Arg Glu Pro Thr Ile Arg Leu Thr Lys Asp Val Gly Ser Ala Glu  
 2375 2380 2385  
 Tyr Trp Val Gly His Val Arg Asp Ala Val Arg Phe Ala Asp Asp  
 2390 2395 2400  
 Val Arg Tyr Leu Gln Asp Glu Gly Val Thr Arg Phe Leu Glu Ile  
 2405 2410 2415  
 Gly Pro Asp Gly Val Leu Thr Ala Met Ala Gly Gln Ser Ala Asp  
 2420 2425 2430  
 Gly Thr Leu Ala Pro Thr Leu Arg Arg Asp Arg Pro Glu Val Glu  
 2435 2440 2445  
 Ser Val Phe Ala Gly Val Gly Arg Leu Phe Ala Ala Gly Val Ala  
 2450 2455 2460  
 Val Asp Trp Asp Ala Val Phe Asp Gly Arg Gly Ala Arg Arg Val  
 2465 2470 2475  
 Asp Leu Pro Thr Tyr Pro Phe Gln Arg Lys Arg Tyr Trp Leu Ile

ES 2 543 203 T3

2480						2485					2490			
Glu	Gln	Ser	Thr	Ala	Ala	Ala	Gly	Ala	Asp	Ala	Val	Asp	His	Pro
2495						2500					2505			
Leu	Leu	Thr	Ser	Gly	Ile	Gly	Leu	Pro	Asp	Thr	Gly	Gly	Val	Val
2510						2515					2520			
Phe	Thr	Gly	Arg	Leu	Ser	Leu	Asp	Thr	His	Pro	Trp	Leu	Ala	Asp
2525						2530					2535			
His	Asp	Val	Leu	Gly	Thr	Leu	Leu	Leu	Pro	Gly	Thr	Gly	Leu	Val
2540						2545					2550			
Glu	Leu	Ala	Leu	Gln	Ala	Ala	Ala	Gln	Val	Asp	Cys	Gly	Thr	Val
2555						2560					2565			
Asp	Glu	Leu	Thr	Leu	Glu	Ala	Pro	Leu	Val	Val	Pro	Glu	Lys	Gly
2570						2575					2580			
Ala	Val	Ala	Val	Arg	Val	Leu	Val	Gly	Gly	Pro	Asp	Asp	Ser	Glu
2585						2590					2595			
Ser	Arg	Thr	Val	Glu	Ile	Tyr	Ser	Ser	Leu	Asp	Asp	Glu	Ile	Trp
2600						2605					2610			
Thr	Arg	Asn	Ala	Ala	Gly	Ala	Leu	Leu	Pro	Ser	Ala	Val	Ser	Pro
2615						2620					2625			
Ser	Ser	Asp	Leu	Ile	Gln	Trp	Pro	Pro	Ile	Gly	Ala	Thr	Pro	Leu
2630						2635					2640			
Pro	Val	Asp	Gly	Ala	Tyr	Glu	Arg	Leu	Leu	Ala	Arg	Gly	Tyr	Asp
2645						2650					2655			
Tyr	Gly	Pro	Thr	Phe	Gln	Gly	Leu	Lys	Ala	Ala	Trp	Arg	Asp	Gly
2660						2665					2670			
Asp	Gly	Val	Val	Phe	Ala	Glu	Val	Ser	Leu	Pro	Glu	Gly	Thr	Glu
2675						2680					2685			
Ala	Ala	Arg	Phe	Gly	Leu	His	Pro	Ala	Leu	Leu	Asp	Ala	Ala	Met
2690						2695					2700			
His	Val	Gly	Leu	Ile	Glu	Glu	Gly	Ala	Ala	Thr	Asp	Ala	Pro	Glu
2705						2710					2715			
Leu	Pro	Phe	Ser	Trp	Asn	Gly	Val	Thr	Leu	His	Arg	Ala	Gly	Ala
2720						2725					2730			



Ser Ala Leu Arg Val Arg Leu Ser Asn Pro Glu Gly Gly Asp Gly  
 2735 2740 2745  
 Thr Glu Val Leu Val Ala Asp Gly Thr Gly Ala Pro Val Leu Ser  
 2750 2755 2760  
 Val Ala Ser Leu Thr Ser Arg Pro Val Ser Ala Glu Gln Leu Arg  
 2765 2770 2775  
 Ala Asp Gly Asp His Arg Glu Ser Leu Phe Ala Leu Thr Trp Thr  
 2780 2785 2790  
 Lys Ala Gly Asp Val Pro Val Leu Glu Thr Pro Val Val Val Tyr  
 2795 2800 2805  
 Glu Val Pro Arg Ala Glu Gly Asp Thr Ser Glu Ala Ala His Ala  
 2810 2815 2820  
 Val Ala Asp Glu Val Leu Ala Arg Val Gln Glu Trp Leu Ala Asp  
 2825 2830 2835  
 Lys Gly Arg Ser Asp Glu Lys Leu Ala Val Val Thr Arg Arg Ala  
 2840 2845 2850  
 Val Pro Ile Glu Gly Glu Asp Val Asp Leu Ser Gln Ala Pro Val  
 2855 2860 2865  
 Trp Gly Leu Val Arg Ala Ala Ala Ala Glu Asn Pro Gly Arg Phe  
 2870 2875 2880  
 Leu Leu Leu Asp Leu Gly Asp Gly Ala Thr Val Pro Pro Phe Val  
 2885 2890 2895  
 Asp Ala Pro Glu Val Ala Val Arg Gly Gly Glu Leu Leu Val Pro  
 2900 2905 2910  
 Ala Leu Thr Arg Val Pro Ala Ser Ala Val Asp Ala Gly Arg Asp  
 2915 2920 2925  
 Pro Trp Glu Ser Ser Pro Thr Val Leu Ile Thr Gly Gly Thr Ser  
 2930 2935 2940  
 Gly Leu Gly Ala Leu Val Ala Arg His Leu Val Thr Glu His Gly  
 2945 2950 2955  
 Ile Arg His Leu Val Leu Thr Ser Arg Arg Gly Gly Ser Ala Pro  
 2960 2965 2970  
 Gly Ala Ala Glu Leu Arg Ala Glu Leu Thr Gly His Gly Ala Arg

ES 2 543 203 T3

2975						2980						2985		
Val	Asp	Ile	Glu	Ala	Cys	Asp	Val	Ala	Asp	Arg	Ala	Ala	Leu	Ala
2990						2995					3000			
Ala	Val	Leu	Asp	Arg	His	Pro	Val	Gly	Ala	Val	Val	His	Ala	Ala
3005						3010					3015			
Gly	Val	Val	Asp	Asn	Gly	Leu	Val	Arg	Gly	Leu	Thr	Pro	Glu	Arg
3020						3025					3030			
Met	Asp	Thr	Val	Leu	Arg	Pro	Lys	Val	Asp	Gly	Ala	Trp	His	Leu
3035						3040					3045			
His	Glu	Leu	Thr	Ala	Asp	Arg	Asp	Leu	Ser	Ala	Phe	Val	Leu	Phe
3050						3055					3060			
Ser	Ser	Met	Gly	Gly	Leu	Leu	Leu	Ala	Ala	Gly	Gln	Gly	Asn	Tyr
3065						3070					3075			
Ala	Ala	Ala	Asn	Val	Phe	Leu	Asp	Ala	Leu	Ala	His	His	Arg	His
3080						3085					3090			
Gln	Ala	Gly	Leu	Pro	Val	Thr	Ser	Leu	Ala	Phe	Gly	Leu	Trp	Glu
3095						3100					3105			
Ala	Glu	Thr	Gly	Leu	Gly	Glu	Leu	Thr	Asp	Ala	Asp	Arg	Lys	Arg
3110						3115					3120			
Met	Leu	Arg	Met	Gly	Leu	Pro	Ala	Leu	Ser	Gln	Lys	Glu	Gly	Leu
3125						3130					3135			
Ala	Leu	Leu	Asp	Asp	Ala	Leu	Arg	Thr	Gly	Glu	Pro	Ala	Leu	Ala
3140						3145					3150			
Pro	Phe	Arg	Leu	Asp	Thr	Gly	Ala	Leu	Arg	Thr	Arg	Ala	Val	Asp
3155						3160					3165			
Gln	Leu	Pro	Ser	Leu	Leu	Arg	Gly	Leu	Val	Pro	Thr	Ser	Arg	Arg
3170						3175					3180			
Arg	Ala	Gly	Gly	Ala	Gly	Ala	Ala	Gly	Gly	Gly	Asp	Gly	Gly	Ala
3185						3190					3195			
Gly	Leu	Arg	Arg	Thr	Leu	Ala	Ala	Ala	Ala	Asp	Glu	Ala	Ala	Arg
3200						3205					3210			
Asp	Ser	Ile	Leu	Val	Glu	Leu	Val	Arg	Thr	His	Val	Ala	Ala	Val
3215						3220					3225			

Leu Gly His Asp Gly Val Asp Ala Val Arg Ser Asp Arg Ala Phe  
 3230 3235 3240  
 Lys Asp Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn  
 3245 3250 3255  
 Thr Leu Arg Ser Ala Thr Gly Leu Lys Leu Pro Ala Thr Leu Val  
 3260 3265 3270  
 Phe Asp His Pro Thr Pro Ser Ala Val Ala Glu Ser Leu Arg Glu  
 3275 3280 3285  
 Gln Leu Thr Gly Asp Gln Glu Pro Ala Gly Ser Pro Leu Glu Ala  
 3290 3295 3300  
 Glu Leu Ala Arg Leu Glu Ala Ala Leu Ala Ser Val Thr Pro Asp  
 3305 3310 3315  
 Glu Glu Thr Phe Gly Arg Val Ala Asp Arg Leu Arg Ala Leu Ala  
 3320 3325 3330  
 Ala Gly Trp Thr Glu Thr His Arg Arg Asp Glu Ser Asp Glu Ala  
 3335 3340 3345  
 Glu Leu Gly Ser Leu Ser Ala Asp Glu Leu Phe Asp Val Leu Asp  
 3350 3355 3360  
 Asp Glu Leu Asp Thr Arg Ser Ala Ser  
 3365 3370

<210> 30  
 <211> 942  
 <212> ADN  
 <213> *Streptomyces sp.* MP28-13  
 <400> 30

atggcggtcca cgcccacggc cacggccaaa ggaactgttc ccttcgggga gtacaagacc  
 60  
 tgggtaccgcg tcaccgggca gcccgctgag ggccgcccgg ccctcgtcgt cgtgcacgga  
 120  
 ggccccggct ccaccacga ctacctgaca gggctgtccg tctacgccga acagggctgg  
 180  
 tcggtggtgc actacgacca gatcgggaac ggcgggtcca cccacctcc cgacgccgac  
 240  
 cccggcttct ggacccccca gctcttccgc gacgagctgg agaacctgct gcgccggctc  
 300  
 gacatcgccg acgactacgt cctgttcgga cagtcgtggg gcggactgct cgccgcctgg  
 360

5

10

cacgcctcgg ccgaaccgc cgggctgcgc ggcctggtca tcgccaacgc accggcctcc  
420

taccctctgt ggctgtcggg gatggacgtc ctgcgcgccc aactgccgcc cggcgtcgcg  
480

gagacactgc ggcggcacga ggccgcccgc accaccgaca gcgacgagta cctggaggcg  
540

atgctggtct tctacagccg ccacgtctgc cgcgtcgcgc cgtggcccag cgaactcatg  
600

gcctcctacc tggaagccgt caccgaccgc acggtctacc gcacgatgaa cggctccaac  
660

gagttccatg tcatcggcag catccgcgac tggtcggtga tcgactgcct gcccgacatc  
720

agcgcgcccc ccctcatcat gtcggggccgc cacgacgagg ccacccccgt cacggtgcgc  
780

ccctaccagg aactcattcc gggcgccccgc tgggaaatcc ttgagaactc cagccacaac  
840

ccgcacctgg aggagccgga gctgttctac gaggtgctcg gcggattcct tgactcggta  
900

cgcgtgagcg acacgcggac cacgaccgtg agcggaggct ga  
942

<210> 31  
<211> 313  
<212> PRT  
<213> *Streptomyces* sp. MP28-13  
  
<400> 31

5

Met Ala Ser Thr Pro Thr Ala Thr Ala Lys Gly Thr Val Pro Phe Gly  
1 5 10 15

Glu Tyr Lys Thr Trp Tyr Arg Val Thr Gly Gln Pro Ala Glu Gly Arg  
20 25 30

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

Leu Thr Gly Leu Ser Val Tyr Ala Glu Gln Gly Trp Ser Val Val His  
50 55 60

Tyr Asp Gln Ile Gly Asn Gly Gly Ser Thr His Leu Pro Asp Ala Asp  
65 70 75 80

Pro Gly Phe Trp Thr Pro Gln Leu Phe Arg Asp Glu Leu Glu Asn Leu  
85 90 95

Leu Arg Arg Leu Asp Ile Ala Asp Asp Tyr Val Leu Phe Gly Gln Ser  
100 105 110

10

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

<210> 32  
 <211> 5961  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

<400> 32

atggccgaca ccgaccagaa actcgtggcg gcgctgcgcg cgctcgtcaa ggagtccgag  
60

agcctgcgta cgcgcaaccg cgccctgcag gccgcctccc gcgaaccgat cgcgatcgtg  
120

gcgatgagct gccgctaccc cggcgcgact tctcccagg agctgtggcg gctggtcgcc  
180

5

10

gacgggacgg acgccgtctc gcggttcccc gccgaccgcg gctgggacga ggagggcatc  
 240  
 tacgaccccc agccgggaaa gcccggcaag acgtactcgc gcgaaggcgg gttcctgtac  
 300  
 gacgcggccg agttc gatcc cggtttcttc gggatcgcgc cgaacgaggc gctggtgatg  
 360  
 gaccctcagc agcgattgct gctggaggcg tcgtgggaag tgctcgagcg ggcgggcatc  
 420  
 gacccgacga ctctcaaggg cagccccgacc ggtgtgttcg ccgggatgat gtaccacgac  
 480  
 tacacgtaca acagcagcac gggcgccatg gcctccggcc gggtcgccta caccctgggt  
 540  
 cttgagggcc ccgcggtgac gatcgacacc gcctgctcgt cctcgctggt cgcctgcac  
 600  
 tgggcggtcc aggccctgcg gtcgggggag tgctcgctcg ccctcgccgg cgggtgcacc  
 660  
 gtgatggcga cccccgagac cttcatcgag ttcagccacc agcgcgggct ggcgaccgac  
 720  
 ggccgctgca agtcgtacgc cgcggcggcc gacggcaccg gctggggtga gggcgtcggc  
 780  
 atgatcctgg tggagcggct gtcggacgcc cgccgcaacg accacccggt gctggggatc  
 840  
 gtgcgcggtgta cggcgatcaa ccaggacggc gccagcaacg gcatcacagc ccccaacggc  
 900  
 ccggcccagc agcgggtgat caggcaggcg ctggccaacg cccgggtgtc cgccgacggg  
 960  
 gtcgacctga tcgagggcca cggcaccggc acgacgctcg gcgacccgat cgagggcgag  
 1020  
 gccctgctcg ccgcctacgg gcaggaccgc cccggggacc gaccgctgtg gctgggttcg  
 1080  
 atcaagtcca acatcgggtca caccagggc gcggcgggcg tggcgggcat catcaagggtg  
 1140  
 gtcgaggcca tccggcacgg tgtcatgccg cccacgctgc acgtcgatgc cccgacacc  
 1200  
 caggtggact gggaggccgg cgacgtccgg ctgctcaccg aggcgcggcg gtggccccgac  
 1260  
 caggagcacc cgcccgcgcc gggggtgtcg tccttcggca tcagcgggac caacgcccac  
 1320  
 gtcatcatcg aggaggcacc gcccgccgag gagcacgcgc ccccggctgc ggcgaccacc  
 1380  
 gggggccccg tgctgtggac cctgtccggc aggacccagc aggcgctgtc cgcgcaggcc  
 1440  
 gaaagccttc actcccatct gcgcgagcgg cccgacctga cgcctgcgga cgtgggcctg  
 1500

tccctggcga ggggccgcgc ggctctcgaa caccgtgcgg cgatcgtcgc cgacgaccgt  
 1560  
 cagggccttc tcgcggggct caccgcgctg gccgcgggaa cccctcgcgc gtccgtcgtc  
 1620  
 accggcaagc ggcgcgaggg caaggtggcg ttcctcttca ccggccaggg cagccagcgc  
 1680  
 ctcggcatgg gacgggagtt gtacgagacc ttcccggctc tcaccgccgc gctcgacgag  
 1740  
 gtgtgcgagg ccacgggcct gtcgctcaag gacgtggtgt ggggcgacga gtcggcgttg  
 1800  
 caccgcaccg agtacgccc a gcccgcgatc ttcgctctgg aagtcgcctt gttccggctg  
 1860  
 gtggagtctt ggggaatcaa gcccgactac ctcgccgggc actccatcgg cgagctggcg  
 1920  
 gcggcccatg tcgcgggcgt tctcggcctt gaggacgccg cgcggctggt cgccgaacgc  
 1980  
 gggcggctga tgcaggcgct cccggcgggc ggggccatga cggccatcga ggccaccgag  
 2040  
 gaggaagtcg cgccgctgct cacggaggag gtggggatcg ccgccctcaa cagcccgtcc  
 2100  
 tccgtggtcg tttcgggcag cgaggacgct gtggaggcgg tcaccgagca cttcgccgac  
 2160  
 cgcaggacgc gacggctgac cgtctctcac gcgttccact cgccgctgat ggaaccgatg  
 2220  
 ctggaggact tccgcaaggt cgccgagtc ctcacctacg aacggccgcg catccggctg  
 2280  
 gtgaaggaca tggcgtccgc cgactactgg gtacggcatg tgcgcgacgc ggtgcggttc  
 2340  
 gccgacgacg tacgacgcct ggaggccgag ggcgtcacc cggttcctgga gctcggaccc  
 2400  
 gacggggccc tcgccccat ggcccgcag accgcgccgg aggccaccac cgccgccgc  
 2460  
 ctgcgccgcg accggcccga ggccacgaca ctgctgaccg ccgtcgcca tctgcacacc  
 2520  
 acgggcgtct ctcccactg gaccgccttc ttcgcgggcc ggcgggcaca ccgggtcgat  
 2580  
 ctgccacct actccttcca gcgcaccctg tactggctcc aggagccggt ggacgcagga  
 2640  
 ggcggcagcg cggcgtccat gggcctcagc gcgctcgacc accccctgct gagcgcagag  
 2700  
 atcgccttcc ccggctcacg gacgggtgat tgcacgggcc gtctgtcgac cgacaccac  
 2760  
 ccctggctcg ccgaccacga ggtactgggc gcgacgctgc tgcccggcac ggcgttcgct  
 2820

gaactggcgg tacgcgtggg cgaccaggtc ggccacggcg tcctcgaaga actgacgctg  
 2880  
 cgcgcgccgc tggtcctgcc cgagggcggc ggcgtacaac tgcggctgac ggtcggtgaa  
 2940  
 ccggtcaggg acaccggccg ggcccccctg agcatccact cgctcgccga ggacgccgac  
 3000  
 gacgacgcg catgggtcct gcacgccgaa ggcgccctgg tggccgagga gtccaacgag  
 3060  
 gcgacctcct tcgacctgtc gcgctggccg cccgacgagg cgacgaggat caccaccgag  
 3120  
 ggcgcgtacg aaaggttcgc ggacctcgga tacgtctacg gccccgcgtt ccaggcgctc  
 3180  
 aaggcggcgt ggcgcgtcgg tgacgagaca ttcgccgagg tggcgctcgc cgacgacgtg  
 3240  
 gccgacgccg agaggttcgt actgcacccg gcgctgctcg actccgctct gcacgcggtg  
 3300  
 atactcggcg cgggcgagga cgaagccacc tcaactgcctt tcgcctggaa ggggtgtcgg  
 3360  
 ctgcacgcct tcggcgccag ggccgcgcgc gtccggttca cccccaacgc cgagggcggc  
 3420  
 acgacgatcc gcgtcgccga cccccagggc cgtcccgtcg cgtacgtgga gtcgctgatc  
 3480  
 tccccgggagg tctccgccga gcagctcgca ccggcgcccc cggggccggg cgactccctc  
 3540  
 ttccacctcg cgtggacacc cgccgccacc gccaccgctg ccgaggcgga ctggacgacg  
 3600  
 gtcaccgaac tggccgaact ctccggcccc gtgccctcga cggtcgcctg gacacccccg  
 3660  
 gcaggcaccg gccggatcgc cgacgacgtc aggacggtca ccgccagac gctccggacc  
 3720  
 ctccagacct ggctcacgga cgaacggttc gccggcagca ggctcctggt ggtcaccgcg  
 3780  
 ggtgacgacc tggcacacgc ctccgcctgg ggcctggtac gcgccgcccg ttcggaggac  
 3840  
 ccggagcgct tcgcgtgct cgacaccgac ggcgacgacc cggagacgat cggccgggcc  
 3900  
 gtcgcgtcgg gcgagcccga cctgcgcgta cggggccagg agatcctggt cccccggctc  
 3960  
 gcgcgcgtcc cggccgcacc ggaggaggac gcaccctcgc gctcgccctg ggaccggccc  
 4020  
 ggagccgtac tgatcaccgg cggcacaggc ggtctcggcg cgctcgctgc ccgccacttg  
 4080  
 gtcgccgaac gcggcgtccg ggacctgctg ctgaccagcc gtcgcggcat cgacgcgacg  
 4140



ggcgcggccg acctccacca ggagctgacc gccctcggtg cgacggtcga gatcgccgcc  
 4200  
 tgcgacgtgg ccgaccggga cgccgtcgag gcgctgctgg ccggccgctc cctcggctcc  
 4260  
 gtcgtccaca ccgccggggt actggccgac agcatgatcg ccaacctgac ggcgcacggc  
 4320  
 ctcgaccagg tcctgcgccc caaggtggac ggcgcactca acctgcacga cctcaccgca  
 4380  
 gaccaggagc tgggcgcctt cgtcctgttc tcgtccgagg cgggcgtgct cggtcgcgcc  
 4440  
 ggccagggca actacgccgc cgccaacacc ttcctcgacg ccctcgcggc gcggcgccac  
 4500  
 gccgagggac tgcccgcgca gtcctcggcc tgggggctgt gggcgacac cggcggaatg  
 4560  
 gcgggcccacc tgagcgaggc agacctgagg cggctgcgcc gccaggcat gcccgcgctg  
 4620  
 tcgtcccagg acggcctcgc gctgttcgac gccgcgtccg tgaggcccga gccggcgctc  
 4680  
 gtgccgatga gcctggacct gcgggcgctg cggaacgggg ccggaggcga actccccgtc  
 4740  
 gtctcgcgag gcctgggtccc cgccgtacgg cgccgctccg cgaccaccga cccgtcggcg  
 4800  
 ctgcggcgag agctggccgc gatgcccgcg cagcagcggg agcgggcgct cagcgatctg  
 4860  
 gtgctgagcc tcgccgcctc cgtgctcgga cacgccgacg ccgaagccgt cgaccccagc  
 4920  
 cgcgacttcc tggagtccgg gttcgactcg ctcaccgca tggaaactgc caccgcgctg  
 4980  
 atcgcgcgca ccggggcgaa gctgcccacg atggcggtgt tcgacagcaa gaccccggcc  
 5040  
 aacctggccc gtctcctcgc cgacgagatg gagtccggga cggccgcccgc cggcgccgag  
 5100  
 tccgcccggc agcccctcga agaggacgac gagacggtga ccgagatggt ccggcgggag  
 5160  
 gtccggggccg gtgacacgac aggggcaactg ggcctgatgt cggcggtcgc ggcgctgcga  
 5220  
 ccccggttcg tcacaccgca cgacctcgcg aggaccccga agacggtgca gttggcggac  
 5280  
 ggccccggcc gccccggct gatctgctg gccacacca tggcgggagg cggcgtgcac  
 5340  
 cagcacgccc ggctcggttc cgaattccgg gacgtgcggc acgtgtcggc ggtggcactg  
 5400  
 cccggattcc accgggacga gccactgccc gactccgctg aggtgctgac acaggtgctg  
 5460

ggcgacgccg tgctggcggc ggcggacggt gagccgttcg tactgctcgg ctactcctcc  
5520  
ggcggcatca tcggccacat catcgcccgt cacctgaagg agacgctcaa ggtcccgcgcc  
5580  
gccggactcg tcctgatcga caccttcagg gtcgaggaca cggcgatgaa cgtcggggtc  
5640  
gaccacctca tgggcgaact gctgacggtg gagacgaccc tcggcaacta cgacgcggcg  
5700  
cgactgtccg cgatgccgca ctacttccag gtactggcgg gcttcgaccc cgtacggctg  
5760  
gacacaccga ccctgttcgt ccaggcgtcc gagccgttcg ttcagcccc cgagggggtc  
5820  
gacgtggcgg agatgcgggc ccgcccgtgg gactccgagc acaccctgcg caccgtcgaa  
5880  
ggcaaccatt tctcgctcgg gcaggaccac gccccggcga ccgcccgagt catcgaggaa  
5940  
tggctggaga cgctcgactg a  
5961

<210> 33  
<211> 1986  
<212> PRT  
<213> *Streptomyces* sp. MP28-13  
  
<400> 33

5

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

10

Glu Ala Ser Trp Glu Val Leu Glu Arg Ala Gly Ile Asp Pro Thr Thr  
 130 135 140  
 Leu Lys Gly Ser Pro Thr Gly Val Phe Ala Gly Met Met Tyr His Asp  
 145 150 155  
 Tyr Thr Tyr Asn Ser Ser Thr Gly Ala Met Ala Ser Gly Arg Val Ala  
 165 170 175  
 Tyr Thr Leu Gly Leu Glu Gly Pro Ala Val Thr Ile Asp Thr Ala Cys  
 180 185 190  
 Ser Ser Ser Leu Val Ala Leu His Trp Ala Val Gln Ala Leu Arg Ser  
 195 200 205  
 Gly Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr  
 210 215 220  
 Pro Glu Thr Phe Ile Glu Phe Ser His Gln Arg Gly Leu Ala Thr Asp  
 225 230 235 240  
 Gly Arg Cys Lys Ser Tyr Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly  
 245 250 255  
 Glu Gly Val Gly Met Ile Leu Val Glu Arg Leu Ser Asp Ala Arg Arg  
 260 265 270  
 Asn Asp His Pro Val Leu Gly Ile Val Arg Gly Thr Ala Ile Asn Gln  
 275 280 285  
 Asp Gly Ala Ser Asn Gly Ile Thr Ala Pro Asn Gly Pro Ala Gln Gln  
 290 295 300  
 Arg Val Ile Arg Gln Ala Leu Ala Asn Ala Arg Val Ser Ala Asp Gly  
 305 310 315 320  
 Val Asp Leu Ile Glu Gly His Gly Thr Gly Thr Thr Leu Gly Asp Pro  
 325 330 335  
 Ile Glu Ala Gln Ala Leu Leu Ala Ala Tyr Gly Gln Asp Arg Pro Gly  
 340 345 350  
 Asp Arg Pro Leu Trp Leu Gly Ser Ile Lys Ser Asn Ile Gly His Thr  
 355 360 365  
 Gln Ala Ala Ala Gly Val Ala Gly Ile Ile Lys Val Val Glu Ala Ile  
 370 375 380

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

Val Ala Glu Arg Gly Arg Leu Met Gln Ala Leu Pro Ala Gly Gly Ala  
 660 665 670  
 Met Thr Ala Ile Glu Ala Thr Glu Glu Glu Val Ala Pro Leu Leu Thr  
 675 680 685  
 Glu Glu Val Gly Ile Ala Ala Leu Asn Ser Pro Ser Ser Val Val Val  
 690 695 700  
 Ser Gly Ser Glu Asp Ala Val Glu Ala Val Thr Glu His Phe Ala Asp  
 705 710 715 720  
 Arg Arg Thr Arg Arg Leu Thr Val Ser His Ala Phe His Ser Pro Leu  
 725 730 735  
 Met Glu Pro Met Leu Glu Asp Phe Arg Lys Val Ala Glu Ser Leu Thr  
 740 745 750  
 Tyr Glu Arg Pro Arg Ile Arg Leu Val Lys Asp Met Ala Ser Ala Asp  
 755 760 765  
 Tyr Trp Val Arg His Val Arg Asp Ala Val Arg Phe Ala Asp Asp Val  
 770 775 780  
 Arg Arg Leu Glu Ala Glu Gly Val Thr Arg Phe Leu Glu Leu Gly Pro  
 785 790 795 800  
 Asp Gly Ala Leu Ala Ala Met Ala Arg Gln Thr Ala Pro Glu Ala Thr  
 805 810 815  
 Thr Ala Ala Ala Leu Arg Arg Asp Arg Pro Glu Ala Thr Thr Leu Leu  
 820 825 830  
 Thr Ala Val Ala His Leu His Thr Thr Gly Val Ser Pro Asp Trp Thr  
 835 840 845  
 Ala Phe Phe Ala Gly Arg Arg Ala His Arg Val Asp Leu Pro Thr Tyr  
 850 855 860  
 Ser Phe Gln Arg Thr Arg Tyr Trp Leu Gln Glu Pro Val Asp Ala Gly  
 865 870 875 880  
 Gly Gly Ser Ala Ala Ser Met Gly Leu Ser Ala Leu Asp His Pro Leu  
 885 890 895  
 Leu Ser Ala Glu Ile Ala Val Pro Gly Ser Arg Thr Val Ile Cys Thr  
 900 905 910

Gly Arg Leu Ser Thr Asp Thr His Pro Trp Leu Ala Asp His Glu Val  
 915 920 925

Leu Gly Ala Thr Leu Leu Pro Gly Thr Ala Phe Val Glu Leu Ala Val  
 930 935 940

Arg Val Gly Asp Gln Val Gly His Gly Val Leu Glu Glu Leu Thr Leu  
 945 950 955 960

Arg Ala Pro Leu Val Leu Pro Glu Gly Gly Gly Val Gln Leu Arg Leu  
 965 970 975

Thr Val Gly Glu Pro Val Glu Asp Thr Gly Arg Arg Pro Leu Ser Ile  
 980 985 990

His Ser Leu Ala Glu Asp Ala Asp Asp Asp Ala Pro Trp Val Leu His  
 995 1000 1005

Ala Glu Gly Ala Leu Val Ala Glu Glu Ser Asn Glu Ala Thr Ser  
 1010 1015 1020

Phe Asp Leu Ser Arg Trp Pro Pro Asp Glu Ala Thr Arg Ile Thr  
 1025 1030 1035

Thr Glu Gly Ala Tyr Glu Arg Phe Ala Asp Leu Gly Tyr Val Tyr  
 1040 1045 1050

Gly Pro Ala Phe Gln Ala Leu Lys Ala Ala Trp Arg Val Gly Asp  
 1055 1060 1065

Glu Thr Phe Ala Glu Val Ala Leu Ala Asp Asp Val Ala Asp Ala  
 1070 1075 1080

Glu Arg Phe Val Leu His Pro Ala Leu Leu Asp Ser Ala Leu His  
 1085 1090 1095

Ala Val Ile Leu Gly Ala Gly Glu Asp Glu Ala Thr Ser Leu Pro  
 1100 1105 1110

Phe Ala Trp Lys Gly Val Arg Leu His Ala Phe Gly Ala Arg Ala  
 1115 1120 1125

Ala Arg Val Arg Phe Thr Pro Asn Ala Glu Gly Gly Thr Thr Ile  
 1130 1135 1140

Arg Val Ala Asp Pro Gln Gly Arg Pro Val Ala Tyr Val Glu Ser  
 1145 1150 1155

Leu Ile Ser Arg Glu Val Ser Ala Glu Gln Leu Ala Pro Ala Pro  
 1160 1165 1170

Ala Gly Pro Gly Asp Ser Leu Phe His Leu Ala Trp Thr Pro Ala  
 1175 1180 1185

Ala Thr Ala Thr Ala Ala Glu Ala Asp Trp Thr Thr Val Thr Glu  
 1190 1195 1200

Leu Ala Glu Leu Ser Gly Pro Val Pro Ser Thr Val Ala Trp Thr  
 1205 1210 1215

Pro Pro Ala Gly Thr Gly Arg Ile Ala Asp Asp Val Arg Thr Val  
 1220 1225 1230

Thr Ala Gln Thr Leu Arg Thr Leu Gln Thr Trp Leu Thr Asp Glu  
 1235 1240 1245

Arg Phe Ala Gly Ser Arg Leu Leu Val Val Thr Arg Gly Asp Asp  
 1250 1255 1260

Leu Ala His Ala Ser Ala Trp Gly Leu Val Arg Ala Ala Arg Ser  
 1265 1270 1275

Glu Asp Pro Glu Arg Phe Ala Leu Leu Asp Thr Asp Gly Asp Asp  
 1280 1285 1290

Pro Glu Thr Ile Gly Arg Ala Val Ala Ser Gly Glu Pro Asp Leu  
 1295 1300 1305

Arg Val Arg Gly Gln Glu Ile Leu Val Pro Arg Leu Ala Arg Val  
 1310 1315 1320

Pro Ala Ala Pro Glu Glu Asp Ala Pro Ser Arg Ser Pro Trp Asp  
 1325 1330 1335

Arg Pro Gly Ala Val Leu Ile Thr Gly Gly Thr Gly Gly Leu Gly  
 1340 1345 1350

Ala Leu Val Ala Arg His Leu Val Ala Glu Arg Gly Val Arg Asp  
 1355 1360 1365

Leu Leu Leu Thr Ser Arg Arg Gly Ile Asp Ala Gln Gly Ala Ala  
 1370 1375 1380

Asp Leu His Gln Glu Leu Thr Ala Leu Gly Ala Thr Val Glu Ile  
 1385 1390 1395

Ala Ala Cys Asp Val Ala Asp Arg Asp Ala Val Glu Ala Leu Leu  
 1400 1405 1410

Ala Gly Arg Ser Leu Gly Ser Val Val His Thr Ala Gly Val Leu  
. 1415 1420 1425

Ala Asp Ser Met Ile Ala Asn Leu Thr Ala His Gly Leu Asp Gln  
1430 1435 1440

Val Leu Arg Pro Lys Val Asp Gly Ala Leu Asn Leu His Asp Leu  
1445 1450 1455

Thr Arg Asp Gln Glu Leu Gly Ala Phe Val Leu Phe Ser Ser Ala  
1460 1465 1470

Ala Gly Val Leu Gly Ser Pro Gly Gln Gly Asn Tyr Ala Ala Ala  
1475 1480 1485

Asn Thr Phe Leu Asp Ala Leu Ala Ala Arg Arg His Ala Glu Gly  
1490 1495 1500

Leu Pro Ala Gln Ser Leu Ala Trp Gly Leu Trp Ala Asp Thr Gly  
1505 1510 1515

Gly Met Ala Gly His Leu Ser Glu Ala Asp Leu Arg Arg Leu Arg  
1520 1525 1530

Arg Gln Gly Met Pro Ala Leu Ser Ser Gln Asp Gly Leu Ala Leu  
1535 1540 1545

Phe Asp Ala Ala Ser Val Arg Pro Glu Pro Ala Leu Val Pro Met  
1550 1555 1560

Ser Leu Asp Leu Arg Ala Leu Arg Asn Gly Ala Gly Gly Glu Leu  
1565 1570 1575

Pro Val Val Leu Arg Gly Leu Val Pro Ala Val Arg Arg Arg Ser  
1580 1585 1590

Ala Thr Thr Asp Pro Ser Ala Leu Arg Arg Glu Leu Ala Ala Met  
1595 1600 1605

Pro Ala Gln Gln Arg Glu Arg Ala Leu Ser Asp Leu Val Leu Ser  
1610 1615 1620

Leu Ala Ala Ser Val Leu Gly His Ala Asp Ala Glu Ala Val Asp  
1625 1630 1635

Pro Ser Arg Asp Phe Leu Glu Ser Gly Phe Asp Ser Leu Thr Ala  
1640 1645 1650

Met Glu Leu Arg Thr Ala Leu Ile Ala Ala Thr Gly Ala Lys Leu  
1655 1660 1665



Pro Thr Met Ala Val Phe Asp Ser Lys Thr Pro Ala Asn Leu Ala  
 1670 1675 1680  
 Arg Leu Leu Ala Asp Glu Met Glu Ser Gly Thr Ala Ala Ala Gly  
 1685 1690  
 Ala Glu Ser Ala Ala Glu Pro Ser Glu Glu Asp Asp Glu Thr Val  
 1700 1705 1710  
 Thr Glu Met Phe Arg Arg Ala Val Arg Ala Gly Asp Thr Thr Gly  
 1715 1720 1725  
 Ala Leu Gly Leu Met Ser Ala Val Ala Ala Leu Arg Pro Arg Phe  
 1730 1735 1740  
 Val Thr Pro Ala Asp Leu Ala Arg Thr Pro Lys Thr Val Arg Leu  
 1745 1750 1755  
 Ala Asp Gly Pro Gly Arg Pro Arg Leu Ile Cys Leu Ala Thr Pro  
 1760 1765 1770  
 Met Ala Gly Gly Gly Val His Gln His Ala Arg Leu Gly Ser Glu  
 1775 1780 1785  
 Phe Arg Asp Val Arg His Val Ser Ala Val Ala Leu Pro Gly Phe  
 1790 1795 1800  
 His Arg Asp Glu Pro Leu Pro Asp Ser Val Glu Val Leu Thr Gln  
 1805 1810 1815  
 Val Leu Gly Asp Ala Val Leu Ala Ala Ala Asp Gly Glu Pro Phe  
 1820 1825 1830  
 Val Leu Leu Gly Tyr Ser Ser Gly Gly Ile Ile Gly His Ile Ile  
 1835 1840 1845  
 Ala Arg His Leu Lys Glu Thr Leu Lys Val Pro Pro Ala Gly Leu  
 1850 1855 1860  
 Val Leu Ile Asp Thr Phe Arg Val Glu Asp Thr Ala Met Asn Val  
 1865 1870 1875  
 Gly Phe Asp His Leu Met Gly Glu Leu Leu Thr Val Glu Thr Thr  
 1880 1885 1890  
 Leu Gly Asn Tyr Asp Ala Ala Arg Leu Ser Ala Met Pro His Tyr  
 1895 1900 1905

Phe Gln Val Leu Ala Gly Phe Asp Pro Val Arg Leu Asp Thr Pro  
 1910 1915 1920  
 Thr Leu Phe Val Gln Ala Ser Glu Pro Phe Val Gln Pro Pro Glu  
 1925 1930 1935  
 Gly Val Asp Val Ala Glu Met Arg Ala Arg Pro Trp Asp Ser Glu  
 1940 1945 1950  
 His Thr Leu Arg Thr Val Glu Gly Asn His Phe Ser Leu Gly Gln  
 1955 1960 1965  
 Asp His Ala Pro Ala Thr Ala Arg Val Ile Glu Glu Trp Leu Glu  
 1970 1975 1980  
 Thr Leu Asp  
 1985

<210> 34  
 <211> 10134  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13  
 <400> 34

atgtcgaagg actccaacga cgaacggctt cgcgagtatc tgcggcttgc caccggtgag  
 60  
 ttgcagcaga ctcggcgccg tctgcgtgag gcggaggatc gggagcggga gccgatcgcg  
 120  
 atcgtgggaa tggcatgccg ctteccccgt ggcgcgtcct cgccggaagg actctgggac  
 180  
 ctcgtcgccg acggtctgga aacggtgggc gagttcccca ccgaccgggg ctggaacctg  
 240  
 gactcgctgt acgaccctga cgggacgggc gagaacacca gctacgtcaa caagggcagc  
 300  
 ttcttgagcg gcgcgggcga cttcgacccc gagttcttcg gcatcagccc ccttgaggcg  
 360  
 atggcgatgg acccgagca gcggctgctg ctggaaacct cctgggaggc cgtggagcgg  
 420  
 gccgggatcg acccgccgtc gctgaagggc accgccaccg gagtcttcgc cgggctcgtg  
 480  
 taccagact atccgagcag cagtgtgacc ggtgcgctgg tttccggccg ggtggcctac  
 540  
 acgctggggc tggagggctc ggcggtgacc gtcgacaccg cctgctcgtc ctcgctggtc  
 600  
 gcactggaca tggcgggtcaa ggccctgcgc agtggggaat gtccctcgc cctggccggc  
 660  
 ggcgtgacga tcatgtcgac gccggtcacc ttcgtggagt tcagccggca gcggggctctg  
 720

5

10

tcgacggacg gccgctgcaa ggcgttcgcc tcggcggccg acggcaccgg gtgggggtgag  
 780  
 ggtgtcggca tgctggtggt ggagcggctg tcggacgccc gccgcaacgg ccacccgatt  
 840  
 ctggcggctg tgcgcggcag cgccgtcaac caggacggcg ccagcaacgg catcaccgcc  
 900  
 ccgaacggcc cgtcccagcg ccgggtgatc cagcaggccc tggccaacgc gggcctggtg  
 960  
 gggctccgatg tggacgtcgt ggaggccac ggtaccggga ccaccctggg cgacccgatc  
 1020  
 gaggcgcagg cgctgctggc cacctacggg caggaccgcc ccgaggaccg cccgctctgg  
 1080  
 ctcgggtcgg tcaagtcgaa catcgggcac acacaggccg ccgccggtgt ggcgggcatc  
 1140  
 atcaagatgg tccaggcgat acgccacggc tccctgcca agacactgca cgtggacgaa  
 1200  
 ccgtccccgc aggtggactg gacggcggga agcgtccggc tgctcaccga gggccgcgcc  
 1260  
 tggccggagg gcggccccgc ccgcgccgcc gtgtcctcgt tcggggccag cgggaccaac  
 1320  
 gccacatcg tcctggagga ggctccgcag gccgaggacg taccggccga agagtcgccg  
 1380  
 gagtcgccgg agccgcgtga gcaccgggaa cttccggtgg tgccctggct gatctccggg  
 1440  
 aagaccgagg ccgccgtacg cgcctacgcg gagcggctcc aggccgtcgc cacgcacgac  
 1500  
 ctacggccgg tggacgtggg atggtccctg gccacgcggc gggcgtcgtg cgactaccgg  
 1560  
 gccgccgtac tggccgccga cgaagagggc ttctcccagg ggctcgcgc tctggtgcag  
 1620  
 ggggaagcac cggtgacagc ggccccggcc ggcgccggcg cggtcatggt gttccccgga  
 1680  
 cagggtccc agtgggtggg catggcgacc gagctgatgg cgtcgtccgc cgtgttcacg  
 1740  
 gagcagatgt ccgcgtgcga ggaggcactg gcgccgttcg tggactggtc gttgagcgag  
 1800  
 gcgctcggtg acgaggccct gctggaacgt gtagacgtcg tgcagcccgt cctgtgggcg  
 1860  
 gtcatggtgt cactggccgg tctgtggcgg cactacggag tcgagcccgt cggcgtggtg  
 1920  
 ggccactccc agggcgagat cgccgcggcg agcgtggccg gagcgtctgc gttgcaggac  
 1980  
 ggcgccggg tggtcgccct gcgcagcaag gcgctgctgg ccctctccgg ccagggcggg  
 2040

atggtgtcgg tcgccctccc ccgggaggag accgagcggc tgatcgagcg ctggggcacc  
2100  
cggacgggca tcgcggtggt caacggcaac gccgccaccg tggctctcggg cgaggtcgac  
2160  
gcgctggacg aactgatggc ctcttgcgag gccgacggcg tacgggcgcg ccggatccag  
2220  
gtggactacg cctcgactc ggcacaggtg gagcgcatag agcgggaact gctcgacgtg  
2280  
ctggcgccca tcaagccccg cgccgcgcbg atcccccttct actccaccgt gaccggcggg  
2340  
ctgctggaca ccaccgctct cgacgccggc tactggtacc ggaatctgcb ccagaccgtg  
2400  
gagttcgacc ggaccatccg ccggctgacc gagcagggcg tgggggtggt catcgaggcc  
2460  
agccccgacc cggtgctggc gccgagcatg gaacagacca cgatcggcac gctgcgccgc  
2520  
aacgacggcg gcctcgaccg tttcctgagc gtgctggccc aggcacacac ccgcbggcbg  
2580  
gacgtcgact gggagaaggt gtacgacgbg accggggcbg ggcagacgga actgccccacg  
2640  
tacgccttcc agcacaaccg ctactggctg aacgacgaga cggcgaacgc cgacgbggcc  
2700  
tccatgggcb tgggctcgbt gggccacccb ctgctcgggg cbatgggtcat gctcgbggg  
2760  
tcggaggagg tcgtgctcac cggacggctg tcgaccggga cbgtgcbctg gctcaccgac  
2820  
catgtcatcb gcggctcgbt cctcttcccc ggcaccggat tcgtggagct ggtgatccgb  
2880  
gccggcbgac aggtcggctg cbggcbgtgcb gaggagctga cbatcbgagc gccgbtggcb  
2940  
ctcgbccgaa gcggcgggggt cbccgbtccag gtcgtcgbtcb gagcbggcbga cgaagagggc  
3000  
cbccgbgagb tccaggtgta ctccccgbac caggacgbga cbgacctgcb ctggaaccgb  
3060  
cacgbccacc gctgtcgbcb caccgbccacc tccgbccggg gaggggaact ggcbgagtg  
3120  
cccccbcccg gcgcbgagcb gctggacctc gatgtcbgaca cbctctacga ggagttggcb  
3180  
ggcacggggt tggcgtacgb gccgaccttc cbggggcttc gggcbcbctg gcgggcbggc  
3240  
gacgbaggtg tcgcbgagat cbccctgcbg gacaacgbcb tggcbgagcb ctttgbtcbt  
3300  
cacccggccc tcttcbgacgb gggcbtgbac gccatcbgac tctccccgbcb gggaaaccgb  
3360

gatgtggcga tgctgccgtt cgcctggtcg ggagtggaac tgcacgcctc cggcgcgggc  
3420  
gcgctgcggg tgcgcgtcac gcccgtagag gacggcgtgg cggccctgac catcgccgac  
3480  
gcgaccggcc ggcccgtcgc cacggtcgac tcgttggtcc tgcggcccct cacggacatg  
3540  
gcgaccaagg cccgtacgga gccgctgtac cacgtcgccc tggccccggt cgccgcccgt  
3600  
accgcctcca ccggggggca gccgcccac gacgaggagg tgttccgcct ccccggcgga  
3660  
ctggacgtac gcgcggcggg gaacctggcg ctggaggcac tgcagtccgc cggctcccgt  
3720  
ctggtggctc tcacgcgcgg cgcggtctcg gtgaacggcg gggacgtcga cgacctggcc  
3780  
gcggcggccg tctgggtct ggtgcgtacc gccagaccg aagaccccgg ccggttcttc  
3840  
ctgatcgacc tggacggtga caccgacgag gcgccgagcg cggacaacgc cgacctcgcg  
3900  
cctgcgctgt cgaccggtga gccccgggtg gtggtacgcg acggtgtcgc ccacgtgccc  
3960  
cggctggccc gcgtgtccgc ggtgccggag acgacggacg acctggcgcc ggcgttcggc  
4020  
gacgaggtgc tgatcaccgg tggagtgggt gtcctgggtg ctctgctggc ccggcatctc  
4080  
gtcaccgagt acgggggtgc acggctcctg ctgaccggtc gccggggcgt ggacacgccc  
4140  
ggcgcggcgg agttggtgga ggagttgagc gggctggggg ccgaggtcga ggtcgcccgc  
4200  
tgcgacgtcg gcgaccgtga ggcgctcgc gcgctgctgg ccgggcgttc gtcaccggt  
4260  
gtggtgcacg cggcgggtgt cctggacgac ggtgtgatcg ggtccctgac gcccgagcgg  
4320  
gtggacctgg tgatgcgccc caaggtggac gccgccctgt atctgcacga gctgacacgc  
4380  
gacatggacc tgaccgcgtt cgtactcttc tcctccgcgg ccggcgtgat cggctcgccc  
4440  
gggcagggca actacgcggc ggccaacgcc tatctggacg ccctggccga gcgccgccgt  
4500  
gcgggtggcc tgcccccca gtccctggcg tggggcctgt ggcggaccag caccggtatg  
4560  
gccagtgaac tgacggacac cgaccggtcc cgtatggaac gcggcggcat cctgtcgtcg  
4620  
tcccacagcg aggggctggc gctcttcgac gctgcgacgg cggcggacgg acccgccgta  
4680

ctgctccccg tcaagctcga cctccccggc gtccgcgccg gcggcgcggt gccggagctg  
 4740  
 ctgcgcggcc tggctcccggt cgtcaccctgt ggcaccgccc gcaccgcgcg cgacgcggac  
 4800  
 gggctccgcg agcggctcgt gggcatgtcc gacgaccagc gcttcgacat gctgctgaac  
 4860  
 ctggtccgcg cgcaggccgc caccacgctc ggctacgccg gaccgggggc cgtcgaccgg  
 4920  
 gagcgcgcgt tccgcgatct ggggtgctgac tcgctggcgg cgatggaact gcgcaacggt  
 4980  
 ctcggcggcg cgaccggcct gcggcttccg gccacgctgg tgttcgacta cccgaacccc  
 5040  
 accgttctcg cgcgtcatct gctggacgag gtctcgggaa cgggtgcacga gggccgtctc  
 5100  
 acccccgtcg cggccccggt cggcgacgac ccgatcgcga tcgtcgcgat ggcgtgccgc  
 5160  
 taccggggag gcgtgtcctc gccggaggac ctgtggcggc tggtcgacag cggcacggac  
 5220  
 gccatctcac acttccccac cgaccgcggc tgggacctgg agcggatcta cgaccccacg  
 5280  
 gccacccgcc cccgcaccag ctacgtcgac aagggcggat tcctttacga cgcgccccag  
 5340  
 ttcgatcccg gcttcttcgg gatcgcgccg aacgaggcgc tggatgatgga ccctcagcag  
 5400  
 cggctgctgc tggaggcgtc gtgggaggtt cttgagcgcg cgggcatcga cccgacgact  
 5460  
 ctcaagggca gcctgaccgg tgtcttcgcc gggatgatgt accacgacta caccacaac  
 5520  
 agcagcacgg gcgccatcgc ctccggccgc gtctcctaca ccctggggct cgaaggcccc  
 5580  
 gcggtgaccg tcgacaccgc ctgctcgtcc tcgctggctg ccctgcacct ggcggcgacg  
 5640  
 gccctgcggt cgggggagtg ctgctcgtcc ctggcccggc gcgtcaccgt gatggccgcg  
 5700  
 gcggacaact tcatcgagtt cagcgagcag cggggctctg cgaccgacgg ccgctgcaag  
 5760  
 tccttcgcgg ccgccgccga cggcgccggc tggagtgagg gtgtcggcat gctcctggtg  
 5820  
 gagcggctgt cggacgcccg ccgcaacggc caccgggtgc tggccctcgt acggggcacg  
 5880  
 gcgaccaacc aggacggcgc gagcaacggc ctgaccgccc ccaacgggccc gtcccagcag  
 5940  
 cgggtgatca agcaggcgtc ggccaacgcg ggcctggcgg gcgccgatgt ggacgcggtc  
 6000

gaggcgcacg gcacgggcac caccctgggt gatccgatcg aggcgcaggc gctgctggcc  
6060  
acctacggcc agggccgccc cgaggaccga ccgctgtggc tggggtcgat caagtcgaac  
6120  
atcggtcaca cccaggccgc cgcgggtgtg gcgggcatca tcaagatggt cgaggcgatg  
6180  
cggcagggca cgctgccacg gacgctgcac gtggacgcgc ccacacccca ggtggactgg  
6240  
gaggccggcc aggtccagct gctcaccgag acgagggagt ggccgaacga cggccgcccg  
6300  
cgccgcgcg gctgtcctc ctccggcatc agcggcacca acgcccacgt catcatcgag  
6360  
gaggccgtcc cggtcgagga agcggcggg gagcggcggg agttgcccgt cgtccccctg  
6420  
gtgctgtcgg cccggacccg taccgcgctc caggcccagc tcgaccggat cacctccgtc  
6480  
gacgcggacg agctggacgt ggcgtactcg gccgcgaccg gccgggcccgc cctggaacac  
6540  
cgcgcgggtg ggatcggctc cgagaccgtg atcgactcgg tcaccgaggg cgggctggcg  
6600  
ttcctgttca ccggacaggg cagtcagtgg gccgggatgg gccaggagct gtacgagacg  
6660  
ttccccgctt tcgccacggc gttcgacgag gtgtgcgccg tgctcgaccc cgcgctgcgc  
6720  
gagggtgatg ggggcgacga ggaggccctg ggccgcaccg agttcaccca gcccgcgatc  
6780  
ttcgtctcc aggtcgccct gttccggctg gtggagtcgt gggggatcaa gccggacctc  
6840  
atgaccggcc actccatcgg ggaactcgcc gccgccata cggccggggg gttcgatctg  
6900  
gctgatgccg cccggctgat cacggcgcgc ggacggctga tccaggaact cccgcccggc  
6960  
ggggcgatgg tggcgatccg ggccaccgag gaggaggtca cggcgacct caccgagcag  
7020  
gtgagcgtcg cggcggtaa cccccggc tcggtggtga tctcgggccc cgaggaggcc  
7080  
gtcaccgcgg tcgccgagcg gttcgccgac cgcaagacga ctcggctgaa ggtgtcgcac  
7140  
gcgttccact cgccgctgat ggaccgatg ctcgaagagt tccggaaggc cgccgagagc  
7200  
gtcacctacc gcgagccggt catccagctg accaaggacg tggcgtcggc ggagtactgg  
7260  
gtacggcacg tgcgcgacgc ggtgcggttc gccgacgacg tacgccacct gcaggaccgg  
7320

ggcatcaccg ggttcatgga ggtcggaccg gacagtgtgc tcaccgcat ggtgcggcag  
 7380  
 accgccgacg ggacggtcgc ggccaccag cgaaacgacc gtccgggcgt ggaaaccctc  
 7440  
 ttaccggcg tcggccggct gttcggccc ggtgtgcggg tcgactgga cggcgtgttc  
 7500  
 gacgggcccg gggcgggcg ggtcggactg ccgacgtacc cttccagcg gcagccgttc  
 7560  
 tggatcgagt cggggcgcg cgccgacgcg tccgaccacc cgctgctgga ccagacggtc  
 7620  
 gctgtcggcg gggcggaccg gaccgtgctg accgggcgtc tgtcgtcgg taccagccc  
 7680  
 tggctcggcg aacacggcgt cgccggcacc ctgctcttcc ccggtacggg cttcgtcgaa  
 7740  
 ctggccgtcc gcgcgggca cgaggtcggc tgcggccgga tcgaggaact gacgatcgag  
 7800  
 gcaccgctgg tccttgccga acacaccgcc acggccgtac aggtcgtggt cggggtggac  
 7860  
 gacggagcgg gccaccgggc tgtcggggtg tacgcccgag gcgcggacga tccccatctg  
 7920  
 ccctggacc gccacggcgc cggaaccctg gccccggcca ccggcggccc cgcggccgag  
 7980  
 gcgatgacc agtggccacc gcccgggtca gagcccgtgg aactggcggg gctgtacgag  
 8040  
 agcctggcgg aagccggagt cgagtacggg ccggcgttcc agggactcaa gtccgcctgg  
 8100  
 cgcgaagagg ggacggctca cgccgacgtc gtcctgtccg gggcggcggga ccgcttcggg  
 8160  
 atccaccgg cactcctgga cgcgccctg cacacggctc cgttgctgtc gggcgacgac  
 8220  
 cgcgtcgtcc tgccgttctc ctgggcggga gtggaactgc acgcctccgg cgccaccgcg  
 8280  
 ctccgtgtcc ggatgacgct ccgcggccag gaccgggtgg ccctccacgc cgtcgacggc  
 8340  
 gccgggacg cggtcgtctc cgtggacgcc ctgaccctgc gtccgatggc cggcgggtgc  
 8400  
 ctacgcgcg tcgactcgt cttccaggtc gagtggacgc ccgtcgccgt acgcgacgac  
 8460  
 gcggccgggt acgcgaagg gtggcggacc gagggcgacg acgtgctctc cacactgcat  
 8520  
 ccgctgctca aggcgatcca ggcggagacc gagaccctcg cggtcgtcac gcgcggcggc  
 8580  
 gtgtccgtgg ccggtgagcg ggtgagcgac ctggccggcg cctcggcgtg gggactggtg  
 8640



cgcagcgcgg agtcggagga tccgggcccg ttcgtcctgg tcgacgtggc cggtgaggac  
 8700  
 gagaaggcgg acatcgcctt ggcgctggcg gcgggagagc cccaggtggc cgtccgggac  
 8760  
 ggggaaggctt acgtaccgag gctgagggcg gcctcggctc cggagtccga gccgtcctcg  
 8820  
 gtcttcggcg acgaggtact gatcacgggc gcgtccggcg ccctgggccc actggtcgcg  
 8880  
 cgccatctcg tcaccgggca cggcgtacgg cgactgctgc tgaccagccg ccggggcctg  
 8940  
 gcggcgcccc gcgcggcggg gttggtggag gagctggccc ggctgggggc cgaggtcgag  
 9000  
 gtcgcccctt gcgacgtcgg cgaccgtgag gcgctcggcg cgctgctggc cgggcgttcg  
 9060  
 ctaccgggtg tgggtcacgc ggcgggctt ctggacgacg gcgtgatcgc gtcgctgacc  
 9120  
 cccgaacgcc tggacaaggc cgtcacaccc aaggcgggtg ccgccctgca tctgcacgag  
 9180  
 ctgacacgcg acatggacct gaccgcttc gtactcttct cctcgggtggc cggcgtgatc  
 9240  
 ggcacaccgg ggcagggcaa ctacgcccg gcacaacgcgc ttctggacgc actggcccgc  
 9300  
 caccggcgcg ccgacggcct gcccgcccag tccctggcct ggggcctgtg gacgaccgac  
 9360  
 gccggcatgg ccggtgacct ggcggacacc gaccggcagc ggatcgaccg cgcgggcctc  
 9420  
 gtgggactct cggccgagca gggctctgaa ctctcgcagc tggcggggcg gctggccaca  
 9480  
 ccggcgctgg tgcccatgaa cctggacacc aggacgctga acgcggcccga cgtgccttg  
 9540  
 atgctgcgcg gactggtacg cggtcctcgc cggcgtgccg tgctggcgca atccacgggc  
 9600  
 ggcggggccc cgctgcgcaa gcgcctggcg gcaactgccc ccgacgatcg gtacgacgaa  
 9660  
 gtcctcgacc tcgtccgtac ccacgcggcg gcggtactgg gtcacgcggg tcccgaggcc  
 9720  
 gtcgaaccgg agcgggcctt cggcgtctg ggcctcact cgctggggcg tgctgagttc  
 9780  
 cgcaacaccc tcaacgccgc ggcgggctg cgactgtccg ccaccatgat ctccgaccat  
 9840  
 ccgaccgccc ggggtgctcg cgagcacatc ggcgcggagc tggcaccgca cggtgaccgc  
 9900  
 ggcggcccgc ccgacgagga gacggctcgc cgcctgctcg ggtcgattcc gctcaccg  
 9960

ctgcacgacg cggggctgat ggaggccctg ctcgaactcg ccggatccgg cgccgcctcc  
10020

ggcatggcgg cagccgtgcc cgaggaggac tcgatcgacg cgatggacgc cgaagccctg  
10080

atcagcatgg ccctgcagga caccgatccg gacgaagcga cgcgggaagt ctga  
10134

<210> 35  
<211> 3377  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 35

Met Ser Lys Asp Ser Asn Asp Glu Arg Leu Arg Glu Tyr Leu Arg Leu  
1 5 10 15

Ala Thr Gly Glu Leu Gln Gln Thr Arg Arg Arg Leu Arg Glu Ala Glu  
20 25 30

Asp Arg Glu Arg Glu Pro Ile Ala Ile Val Gly Met Ala Cys Arg Phe  
35 40 45

Pro Gly Gly Ala Ser Ser Pro Glu Gly Leu Trp Asp Leu Val Ala Asp  
50 55 60

Gly Leu Glu Thr Val Gly Glu Phe Pro Thr Asp Arg Gly Trp Asn Leu  
65 70 75 80

Asp Ser Leu Tyr Asp Pro Asp Gly Thr Gly Glu Asn Thr Ser Tyr Val  
85 90 95

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

Phe Gly Ile Ser Pro Leu Glu Ala Met Ala Met Asp Pro Gln Gln Arg  
115 120 125

Leu Leu Leu Glu Thr Ser Trp Glu Ala Val Glu Arg Ala Gly Ile Asp  
130 135 140

Pro Pro Ser Leu Lys Gly Thr Ala Thr Gly Val Phe Ala Gly Leu Val  
145 150 155 160

Tyr His Asp Tyr Pro Ser Ser Ser Val Thr Gly Ala Leu Val Ser Gly  
165 170 175

Arg Val Ala Tyr Thr Leu Gly Leu Glu Gly Pro Ala Val Thr Val Asp  
180 185 190

Thr Ala Cys Ser Ser Ser Leu Val Ala Leu Asp Met Ala Val Lys Ala

10

195					200					205					
Leu	Arg	Ser	Gly	Glu	Cys	Ser	Leu	Ala	Leu	Ala	Gly	Gly	Val	Thr	Ile
	210					215					220				
Met	Ser	Thr	Pro	Val	Thr	Phe	Val	Glu	Phe	Ser	Arg	Gln	Arg	Gly	Leu
225					230					235					240
Ser	Thr	Asp	Gly	Arg	Cys	Lys	Ala	Phe	Ala	Ser	Ala	Ala	Asp	Gly	Thr
				245					250					255	
Gly	Trp	Gly	Glu	Gly	Val	Gly	Met	Leu	Val	Val	Glu	Arg	Leu	Ser	Asp
			260					265					270		
Ala	Arg	Arg	Asn	Gly	His	Pro	Ile	Leu	Ala	Val	Val	Arg	Gly	Ser	Ala
		275					280					285			
Val	Asn	Gln	Asp	Gly	Ala	Ser	Asn	Gly	Ile	Thr	Ala	Pro	Asn	Gly	Pro
	290					295					300				
Ser	Gln	Arg	Arg	Val	Ile	Gln	Gln	Ala	Leu	Ala	Asn	Ala	Gly	Leu	Val
305					310					315					320
Gly	Ser	Asp	Val	Asp	Val	Val	Glu	Ala	His	Gly	Thr	Gly	Thr	Thr	Leu
				325					330						335
Gly	Asp	Pro	Ile	Glu	Ala	Gln	Ala	Leu	Leu	Ala	Thr	Tyr	Gly	Gln	Asp
			340					345					350		
Arg	Pro	Glu	Asp	Arg	Pro	Leu	Trp	Leu	Gly	Ser	Val	Lys	Ser	Asn	Ile
		355					360					365			
Gly	His	Thr	Gln	Ala	Ala	Ala	Gly	Val	Ala	Gly	Ile	Ile	Lys	Met	Val
	370					375					380				
Gln	Ala	Ile	Arg	His	Gly	Ser	Leu	Pro	Lys	Thr	Leu	His	Val	Asp	Glu
385					390					395					400
Pro	Ser	Pro	Gln	Val	Asp	Trp	Thr	Ala	Gly	Ser	Val	Arg	Leu	Leu	Thr
				405					410					415	
Glu	Ala	Arg	Ala	Trp	Pro	Glu	Gly	Gly	Pro	Arg	Arg	Ala	Ala	Val	Ser
			420					425					430		
Ser	Phe	Gly	Ala	Ser	Gly	Thr	Asn	Ala	His	Ile	Val	Leu	Glu	Glu	Ala
		435					440					445			
Pro	Gln	Ala	Glu	Asp	Val	Pro	Ala	Glu	Glu	Ser	Pro	Glu	Ser	Pro	Glu
	450					455					460				

Pro Arg Glu His Arg Glu Leu Pro Val Val Pro Trp Leu Ile Ser Gly  
 465 470 475 480  
 Lys Thr Glu Ala Ala Val Arg Ala Tyr Ala Glu Arg Leu Gln Ala Val  
 485 490 495  
 Ala Thr His Asp Leu Arg Pro Val Asp Val Gly Trp Ser Leu Ala Thr  
 500 505 510  
 Arg Arg Ala Ser Tyr Asp Tyr Arg Ala Ala Val Leu Ala Ala Asp Glu  
 515 520 525  
 Glu Gly Phe Ser Gln Gly Leu Ala Ala Leu Val Gln Gly Glu Ala Pro  
 530 535 540  
 Val Thr Ala Ala Arg Pro Gly Ala Gly Ala Val Met Val Phe Pro Gly  
 545 550 555 560 565  
 Gln Gly Ser Gln Trp Val Gly Met Ala Thr Glu Leu Met Ala Ser Ser  
 565 570 575  
 Ala Val Phe Thr Glu Gln Met Ser Ala Cys Glu Glu Ala Leu Ala Pro  
 580 585 590  
 Phe Val Asp Trp Ser Leu Ser Glu Ala Leu Gly Asp Glu Ala Leu Leu  
 595 600 605  
 Glu Arg Val Asp Val Val Gln Pro Val Leu Trp Ala Val Met Val Ser  
 610 615 620  
 Leu Ala Gly Leu Trp Arg His Tyr Gly Val Glu Pro Val Gly Val Val  
 625 630 635 640  
 Gly His Ser Gln Gly Glu Ile Ala Ala Ala Ser Val Ala Gly Ala Leu  
 645 650 655  
 Ser Leu Gln Asp Gly Ala Arg Val Val Ala Leu Arg Ser Lys Ala Leu  
 660 665 670  
 Leu Ala Leu Ser Gly Gln Gly Gly Met Val Ser Val Ala Leu Pro Arg  
 675 680 685  
 Glu Glu Thr Glu Arg Leu Ile Glu Arg Trp Gly Thr Arg Thr Gly Ile  
 690 695 700  
 Ala Val Val Asn Gly Asn Ala Ala Thr Val Val Ser Gly Glu Val Asp  
 705 710 715 720  
 Ala Leu Asp Glu Leu Met Ala Ser Cys Glu Ala Asp Gly Val Arg Ala

				725						730					735			
Arg	Arg	Ile	Gln	Val	Asp	Tyr	Ala	Ser	His	Ser	Ala	Gln	Val	Glu	Arg			
			740					745					750					
Ile	Glu	Arg	Glu	Leu	Leu	Asp	Val	Leu	Ala	Pro	Ile	Lys	Pro	Arg	Ala			
		755					760					765						
Ala	Arg	Ile	Pro	Phe	Tyr	Ser	Thr	Val	Thr	Gly	Gly	Leu	Leu	Asp	Thr			
	770					775					780							
Thr	Ala	Leu	Asp	Ala	Gly	Tyr	Trp	Tyr	Arg	Asn	Leu	Arg	Gln	Thr	Val			
	785				790					795					800			
Glu	Phe	Asp	Arg	Thr	Ile	Arg	Arg	Leu	Thr	Glu	Gln	Gly	Val	Gly	Val			
				805					810					815				
Phe	Ile	Glu	Ala	Ser	Pro	His	Pro	Val	Leu	Ala	Pro	Ser	Met	Glu	Gln			
			820					825					830					
Thr	Thr	Ile	Gly	Thr	Leu	Arg	Arg	Asn	Asp	Gly	Gly	Leu	Asp	Arg	Phe			
		835						840				845						
Leu	Ser	Val	Leu	Ala	Gln	Ala	His	Thr	Arg	Gly	Val	Asp	Val	Asp	Trp			
	850					855					860							
Glu	Lys	Val	Tyr	Asp	Ala	Thr	Gly	Ala	Gly	Gln	Thr	Glu	Leu	Pro	Thr			
	865				870					875					880			
Tyr	Ala	Phe	Gln	His	Asn	Arg	Tyr	Trp	Leu	Asn	Asp	Glu	Thr	Ala	Asn			
				885					890					895				
Ala	Asp	Ala	Ala	Ser	Met	Gly	Leu	Gly	Ser	Leu	Gly	His	Pro	Leu	Leu			
			900					905					910					
Gly	Ala	Met	Val	Met	Leu	Ala	Gly	Ser	Glu	Glu	Val	Val	Leu	Thr	Gly			
		915					920					925						
Arg	Leu	Ser	Thr	Gly	Thr	Leu	Pro	Trp	Leu	Thr	Asp	His	Val	Ile	Gly			
	930					935					940							
Gly	Ser	Ile	Leu	Phe	Pro	Gly	Thr	Gly	Phe	Val	Glu	Leu	Val	Ile	Arg			
	945				950					955					960			
Ala	Gly	Asp	Glu	Val	Gly	Cys	Gly	Arg	Val	Glu	Glu	Leu	Thr	Ile	Glu			
				965					970					975				
Ala	Pro	Leu	Val	Leu	Ala	Glu	Arg	Gly	Gly	Val	Ala	Val	Gln	Val	Val			
			980					985					990					

val Gly Ala Ala Asp Glu Glu Gly Arg Arg Glu val Gln Val Tyr Ser  
 995 1000 1005  
 Arg Asp Gln Asp Ala Thr Asp Leu Pro Trp Asn Arg His Ala Thr  
 1010 1015 1020  
 Gly Leu Leu Ala Thr Ala Thr Ser Ala Gly Gly Gly Glu Leu Ala  
 1025 1030 1035  
 Glu Trp Pro Pro Pro Gly Ala Glu Pro Leu Asp Leu Asp Val Asp  
 1040 1045 1050  
 Thr Leu Tyr Glu Glu Leu Val Gly Thr Gly Leu Ala Tyr Gly Pro  
 1055 1060 1065  
 Thr Phe Arg Gly Leu Arg Ala Ala Trp Arg Ala Gly Asp Glu Val  
 1070 1075 1080  
 Phe Ala Glu Ile Ala Leu Pro Asp Asn Ala Val Ala Asp Ala Phe  
 1085 1090 1095  
 Gly Leu His Pro Ala Leu Phe Asp Ala Gly Leu His Ala Ile Gly  
 1100 1105 1110  
 Leu Ser Pro Ala Gly Thr Gly Asp Val Ala Met Leu Pro Phe Ala  
 1115 1120 1125  
 Trp Ser Gly Val Glu Leu His Ala Ser Gly Ala Gly Ala Leu Arg  
 1130 1135 1140  
 Val Arg Val Thr Pro Val Gln Asp Gly Val Ala Ala Leu Thr Ile  
 1145 1150 1155  
 Ala Asp Ala Thr Gly Arg Pro Val Ala Thr Val Asp Ser Leu Val  
 1160 1165 1170  
 Leu Arg Pro Leu Thr Asp Met Ala Thr Lys Ala Arg Thr Glu Pro  
 1175 1180 1185  
 Leu Tyr His Val Ala Leu Ala Pro Val Ala Ala Gly Thr Ala Ser  
 1190 1195 1200  
 Thr Gly Gly Gln Pro Pro Asn Asp Glu Glu Val Phe Arg Leu Pro  
 1205 1210 1215  
 Gly Gly Leu Asp Val Arg Ala Ala Val Asn Leu Ala Leu Glu Ala  
 1220 1225 1230  
 Leu Gln ser Ala Gly Ser Arg Leu val val val Thr Arg Gly Ala

1235						1240						1245		
Val	Ser	Val	Asn	Gly	Gly	Asp	Val	Asp	Asp	Leu	Ala	Ala	Ala	Ala
	1250					1255					1260			
Val	Trp	Gly	Leu	Val	Arg	Thr	Ala	Gln	Thr	Glu	Asp	Pro	Gly	Arg
	1265					1270					1275			
Phe	Phe	Leu	Ile	Asp	Leu	Asp	Gly	Asp	Thr	Asp	Glu	Ala	Pro	Ser
	1280					1285					1290			
Ala	Asp	Asn	Ala	Asp	Leu	Ala	Pro	Ala	Leu	Ser	Thr	Gly	Glu	Pro
	1295					1300					1305			
Arg	Val	Val	Val	Arg	Asp	Gly	Val	Ala	His	Val	Pro	Arg	Leu	Ala
	1310					1315					1320			
Arg	Val	Ser	Ala	Val	Pro	Glu	Thr	Thr	Asp	Asp	Leu	Ala	Pro	Ala
	1325					1330					1335			
Phe	Gly	Asp	Glu	Val	Leu	Ile	Thr	Gly	Gly	Val	Gly	Val	Leu	Gly
	1340					1345					1350			
Ala	Leu	Leu	Ala	Arg	His	Leu	Val	Thr	Glu	Tyr	Gly	Val	Ser	Arg
	1355					1360					1365			
Leu	Leu	Leu	Thr	Gly	Arg	Arg	Gly	Val	Asp	Thr	Pro	Gly	Ala	Ala
	1370					1375					1380			
Glu	Leu	Val	Glu	Glu	Leu	Ser	Gly	Leu	Gly	Ala	Glu	Val	Glu	Val
	1385					1390					1395			
Ala	Ala	Cys	Asp	Val	Gly	Asp	Arg	Glu	Ala	Leu	Ala	Ala	Leu	Leu
	1400					1405					1410			
Ala	Gly	Arg	Ser	Leu	Thr	Gly	Val	Val	His	Ala	Ala	Gly	Val	Leu
	1415					1420					1425			
Asp	Asp	Gly	Val	Ile	Gly	Ser	Leu	Thr	Pro	Glu	Arg	Val	Asp	Leu
	1430					1435					1440			
Val	Met	Arg	Pro	Lys	Val	Asp	Ala	Ala	Leu	Tyr	Leu	His	Glu	Leu
	1445					1450					1455			
Thr	Arg	Asp	Met	Asp	Leu	Thr	Ala	Phe	Val	Leu	Phe	Ser	Ser	Ala
	1460					1465					1470			
Ala	Gly	Val	Ile	Gly	Ser	Pro	Gly	Gln	Gly	Asn	Tyr	Ala	Ala	Ala
	1475					1480					1485			

Asn Ala Tyr Leu Asp Ala Leu Ala Glu Arg Arg Arg Ala Gly Gly  
 1490 1495 1500  
 Leu Pro Ala Gln Ser Leu Ala Trp Gly Leu Trp Arg Thr Ser Thr  
 1505 1510 1515  
 Gly Met Ala Ser Glu Leu Thr Asp Thr Asp Arg Ser Arg Met Glu  
 1520 1525 1530  
 Arg Gly Gly Ile Leu Ser Leu Ser His Ser Glu Gly Leu Ala Leu  
 1535 1540 1545  
 Phe Asp Ala Ala Thr Ala Ala Asp Gly Pro Ala Val Leu Val Pro  
 1550 1555 1560  
 Val Lys Leu Asp Leu Pro Ala Val Arg Ala Gly Gly Ala Val Pro  
 1565 1570 1575  
 Glu Leu Leu Arg Gly Leu Val Pro Val Val Thr Arg Gly Thr Ala  
 1580 1585 1590  
 Arg Thr Arg Ala Asp Ala Asp Gly Leu Arg Glu Arg Leu Val Gly  
 1595 1600 1605  
 Met Ser Asp Asp Gln Arg Phe Asp Met Leu Leu Asn Leu Val Arg  
 1610 1615 1620  
 Ala Gln Ala Ala Thr Thr Leu Gly Tyr Ala Gly Pro Gly Ala Val  
 1625 1630 1635  
 Asp Pro Glu Arg Ala Phe Arg Asp Leu Gly Val Asp Ser Leu Ala  
 1640 1645 1650  
 Ala Met Glu Leu Arg Asn Gly Leu Gly Gly Ala Thr Gly Leu Arg  
 1655 1660 1665  
 Leu Pro Ala Thr Leu Val Phe Asp Tyr Pro Asn Pro Thr Val Leu  
 1670 1675 1680  
 Ala Arg His Leu Leu Asp Glu Val Ser Gly Thr Val His Glu Gly  
 1685 1690 1695  
 Arg Leu Thr Pro Val Ala Ala Pro Val Gly Asp Asp Pro Ile Ala  
 1700 1705 1710  
 Ile Val Ala Met Ala Cys Arg Tyr Pro Gly Gly Val Ser Ser Pro  
 1715 1720 1725  
 Glu Asp Leu Trp Arg Leu Val Asp Ser Gly Thr Asp Ala Ile Ser



1730						1735						1740
His	Phe 1745	Pro	Thr	Asp	Arg	Gly 1750	Trp	Asp	Leu	Glu	Arg 1755	Ile Tyr Asp
Pro	Thr 1760	Ala	Thr	Arg	Pro	Arg 1765	Thr	Ser	Tyr	Val	Asp 1770	Lys Gly Gly
Phe	Leu 1775	Tyr	Asp	Ala	Ala	Gln 1780	Phe	Asp	Pro	Gly	Phe 1785	Phe Gly Ile
Ala	Pro 1790	Asn	Glu	Ala	Leu	Val 1795	Met	Asp	Pro	Gln	Gln 1800	Arg Leu Leu
Leu	Glu 1805	Ala	Ser	Trp	Glu	Val 1810	Leu	Glu	Arg	Ala	Gly 1815	Ile Asp Pro
Thr	Thr 1820	Leu	Lys	Gly	Ser	Leu 1825	Thr	Gly	Val	Phe	Ala 1830	Gly Met Met
Tyr	His 1835	Asp	Tyr	Thr	His	Asn 1840	Ser	Ser	Thr	Gly	Ala 1845	Ile Ala Ser
Gly	Arg 1850	Val	Ser	Tyr	Thr	Leu 1855	Gly	Leu	Glu	Gly	Pro 1860	Ala Val Thr
Val	Asp 1865	Thr	Ala	Cys	Ser	Ser 1870	Ser	Leu	Val	Ala	Leu 1875	His Leu Ala
Ala	Gln 1880	Ala	Leu	Arg	Ser	Gly 1885	Glu	Cys	Ser	Leu	Ala 1890	Leu Ala Gly
Gly	Val 1895	Thr	Val	Met	Ala	Ala 1900	Ala	Asp	Asn	Phe	Ile 1905	Glu Phe Ser
Glu	Gln 1910	Arg	Gly	Leu	Ala	Thr 1915	Asp	Gly	Arg	Cys	Lys 1920	Ser Phe Ala
Ala	Ala 1925	Ala	Asp	Gly	Ala	Gly 1930	Trp	Ser	Glu	Gly	Val 1935	Gly Met Leu
Leu	Val 1940	Glu	Arg	Leu	Ser	Asp 1945	Ala	Arg	Arg	Asn	Gly 1950	His Pro Val
Leu	Ala 1955	Leu	Val	Arg	Gly	Thr 1960	Ala	Thr	Asn	Gln	Asp 1965	Gly Ala Ser
Asn	Gly 1970	Leu	Thr	Ala	Pro	Asn 1975	Gly	Pro	Ser	Gln	Gln 1980	Arg Val Ile

Lys Gln Ala Leu Ala Asn Ala Gly Leu Ala Gly Ala Asp Val Asp  
 1985 1990 1995  
 Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile  
 2000 2005 2010  
 Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Gly Arg Pro Glu  
 2015 2020 2025  
 Asp Arg Pro Leu Trp Leu Gly Ser Ile Lys Ser Asn Ile Gly His  
 2030 2035 2040  
 Thr Gln Ala Ala Ala Gly Val Ala Gly Ile Ile Lys Met Val Glu  
 2045 2050 2055  
 Ala Met Arg Gln Gly Thr Leu Pro Arg Thr Leu His Val Asp Ala  
 2060 2065 2070  
 Pro Thr Pro Gln Val Asp Trp Glu Ala Gly Gln Val Gln Leu Leu  
 2075 2080 2085  
 Thr Glu Thr Arg Glu Trp Pro Asn Asp Gly Arg Pro Arg Arg Ala  
 2090 2095 2100  
 Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile  
 2105 2110 2115  
 Ile Glu Glu Ala Val Pro Val Glu Glu Ala Pro Val Glu Arg Arg  
 2120 2125 2130  
 Glu Leu Pro Val Val Pro Leu Val Leu Ser Ala Arg Thr Arg Thr  
 2135 2140 2145  
 Ala Leu Gln Ala Gln Leu Asp Arg Ile Thr Ser Val Asp Ala Asp  
 2150 2155 2160  
 Glu Leu Asp Val Ala Tyr Ser Ala Ala Thr Gly Arg Ala Ala Leu  
 2165 2170 2175  
 Glu His Arg Ala Val Arg Ile Gly Ser Glu Thr Val Ile Asp Ser  
 2180 2185 2190  
 Val Thr Glu Gly Gly Leu Ala Phe Leu Phe Thr Gly Gln Gly Ser  
 2195 2200 2205  
 Gln Trp Ala Gly Met Gly Gln Glu Leu Tyr Glu Thr Phe Pro Ala  
 2210 2215 2220  
 Phe Ala Thr Ala Phe Asp Glu Val Cys Ala Val Leu Asp Pro Ala

ES 2 543 203 T3

2225						2230						2235		
Leu	Arg	Glu	Val	Met	Trp	Gly	Asp	Glu	Glu	Ala	Leu	Gly	Arg	Thr
	2240					2245					2250			
Glu	Phe	Thr	Gln	Pro	Ala	Ile	Phe	Ala	Leu	Gln	Val	Ala	Leu	Phe
	2255					2260					2265			
Arg	Leu	Val	Glu	Ser	Trp	Gly	Ile	Lys	Pro	Asp	Leu	Met	Thr	Gly
	2270					2275					2280			
His	Ser	Ile	Gly	Glu	Leu	Ala	Ala	Ala	His	Thr	Ala	Gly	Val	Phe
	2285					2290					2295			
Asp	Leu	Ala	Asp	Ala	Ala	Arg	Leu	Ile	Thr	Ala	Arg	Gly	Arg	Leu
	2300					2305					2310			
Ile	Gln	Glu	Leu	Pro	Pro	Gly	Gly	Ala	Met	Val	Ala	Ile	Arg	Ala
	2315					2320					2325			
Thr	Glu	Glu	Glu	Val	Thr	Pro	His	Leu	Thr	Glu	Gln	Val	Ser	Val
	2330					2335					2340			
Ala	Ala	Val	Asn	Thr	Pro	Gly	Ser	Val	Val	Ile	Ser	Gly	Ala	Glu
	2345					2350					2355			
Glu	Ala	Val	Thr	Ala	Val	Ala	Glu	Arg	Phe	Ala	Asp	Arg	Lys	Thr
	2360					2365					2370			
Thr	Arg	Leu	Lys	Val	Ser	His	Ala	Phe	His	Ser	Pro	Leu	Met	Asp
	2375					2380					2385			
Pro	Met	Leu	Glu	Glu	Phe	Arg	Lys	Val	Ala	Glu	Ser	Val	Thr	Tyr
	2390					2395					2400			
Arg	Glu	Pro	Val	Ile	Gln	Leu	Thr	Lys	Asp	Val	Ala	Ser	Ala	Glu
	2405					2410					2415			
Tyr	Trp	Val	Arg	His	Val	Arg	Asp	Ala	Val	Arg	Phe	Ala	Asp	Asp
	2420					2425					2430			
Val	Arg	His	Leu	Gln	Asp	Arg	Gly	Ile	Thr	Arg	Phe	Met	Glu	Val
	2435					2440					2445			
Gly	Pro	Asp	Ser	Val	Leu	Thr	Ala	Met	Val	Arg	Gln	Thr	Ala	Asp
	2450					2455					2460			
Gly	Thr	Val	Ala	Ala	Thr	Gln	Arg	Asn	Asp	Arg	Pro	Gly	Val	Glu
	2465					2470					2475			

Thr Leu Phe Thr Gly Val Gly Arg Leu Phe Ala Ala Gly Val Arg  
 2480 2485 2490  
 Val Asp Trp Asn Ala Val Phe Asp Gly Arg Gly Ala Arg Arg Val  
 2495 2500 2505  
 Glu Leu Pro Thr Tyr Pro Phe Gln Arg Gln Pro Phe Trp Ile Glu  
 2510 2515 2520  
 Ser Gly Arg Gly Ala Asp Ala Ser Asp His Pro Leu Leu Asp Gln  
 2525 2530 2535  
 Thr Val Ala Val Ala Gly Ala Asp Arg Thr Val Leu Thr Gly Arg  
 2540 2545 2550  
 Leu Ser Leu Gly Thr Gln Pro Trp Leu Ala Glu His Ala Val Ala  
 2555 2560 2565  
 Gly Thr Leu Leu Phe Pro Gly Thr Gly Phe Val Glu Leu Ala Val  
 2570 2575 2580  
 Arg Ala Gly Asp Glu Val Gly Cys Gly Arg Ile Glu Glu Leu Thr  
 2585 2590 2595  
 Ile Glu Ala Pro Leu Val Leu Ala Glu His Thr Ala Thr Ala Val  
 2600 2605 2610  
 Gln Val Val Val Gly Val Asp Asp Gly Ala Gly His Arg Ala Val  
 2615 2620 2625  
 Glu Val Tyr Ala Arg Gly Ala Asp Asp Pro His Leu Pro Trp Thr  
 2630 2635 2640  
 Arg His Ala Ala Gly Thr Leu Ala Pro Ala Thr Gly Gly Pro Ala  
 2645 2650 2655  
 Ala Glu Ala Met Thr Gln Trp Pro Pro Pro Gly Ala Glu Pro Val  
 2660 2665 2670  
 Glu Leu Ala Gly Leu Tyr Glu Ser Leu Ala Glu Ala Gly Val Glu  
 2675 2680 2685  
 Tyr Gly Pro Ala Phe Gln Gly Leu Lys Ser Ala Trp Arg Glu Glu  
 2690 2695 2700  
 Gly Thr Val Tyr Ala Asp Val Val Leu Ser Gly Ala Ala Asp Arg  
 2705 2710 2715  
 Phe Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu His Thr Val

ES 2 543 203 T3

2720						2725						2730		
Pro	Leu	Leu	Ser	Gly	Asp	Asp	Arg	Val	Val	Leu	Pro	Phe	Ser	Trp
	2735					2740					2745			
Ala	Gly	Val	Glu	Leu	His	Ala	Ser	Gly	Ala	Thr	Ala	Leu	Arg	Val
	2750					2755					2760			
Arg	Met	Thr	Leu	Arg	Gly	Gln	Asp	Arg	Val	Ala	Leu	His	Ala	Val
	2765					2770					2775			
Asp	Gly	Ala	Gly	Gln	Pro	Val	Val	Ser	Val	Asp	Ala	Leu	Thr	Leu
	2780					2785					2790			
Arg	Pro	Met	Ala	Ala	Gly	Ala	Leu	Thr	Arg	Val	Asp	Ser	Leu	Phe
	2795					2800					2805			
Gln	Val	Glu	Trp	Thr	Pro	Val	Ala	Val	Arg	Asp	Asp	Ala	Ala	Gly
	2810					2815					2820			
Asp	Ala	Lys	Val	Trp	Arg	Thr	Glu	Gly	Asp	Asp	Val	Leu	Ser	Thr
	2825					2830					2835			
Leu	His	Pro	Leu	Leu	Lys	Ala	Ile	Gln	Ala	Glu	Thr	Glu	Thr	Leu
	2840					2845					2850			
Ala	Val	Val	Thr	Arg	Gly	Ala	Val	Ser	Val	Ala	Gly	Glu	Arg	Val
	2855					2860					2865			
Ser	Asp	Leu	Ala	Gly	Ala	Ser	Ala	Trp	Gly	Leu	Val	Arg	Ser	Ala
	2870					2875					2880			
Gln	Ser	Glu	Asp	Pro	Gly	Arg	Phe	Val	Leu	Val	Asp	Val	Ala	Gly
	2885					2890					2895			
Glu	Asp	Glu	Lys	Ala	Asp	Ile	Ala	Leu	Ala	Leu	Ala	Ala	Gly	Glu
	2900					2905					2910			
Pro	Gln	Val	Ala	Val	Arg	Asp	Gly	Lys	Val	Tyr	Val	Pro	Arg	Leu
	2915					2920					2925			
Arg	Ala	Ala	Ser	Val	Ala	Glu	Ser	Glu	Pro	Ser	Ser	Val	Phe	Gly
	2930					2935					2940			
Asp	Glu	Val	Leu	Ile	Thr	Gly	Ala	Ser	Gly	Ala	Leu	Gly	Gly	Leu
	2945					2950					2955			
Val	Ala	Arg	His	Leu	Val	Thr	Gly	His	Gly	Val	Arg	Arg	Leu	Leu
	2960					2965					2970			

Leu Thr Ser Arg Arg Gly Leu Ala Ala Pro Gly Ala Ala Glu Leu  
 2975 2980 2985  
 Val Glu Glu Leu Ala Gly Leu Gly Ala Glu Val Glu Val Ala Ala  
 2990 3000  
 Cys Asp Val Gly Asp Arg Glu Ala Leu Ala Ala Leu Leu Ala Gly  
 3005 3010 3015  
 Arg Ser Leu Thr Gly Val Val His Ala Ala Gly Val Leu Asp Asp  
 3020 3025 3030  
 Gly Val Ile Ala Ser Leu Thr Pro Glu Arg Leu Asp Lys Val Val  
 3035 3040 3045  
 Thr Pro Lys Ala Val Ala Ala Leu His Leu His Glu Leu Thr Arg  
 3050 3055 3060  
 Asp Met Asp Leu Thr Ala Phe Val Leu Phe Ser Ser Val Ala Gly  
 3065 3070 3075  
 Val Ile Gly Thr Pro Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala  
 3080 3085 3090  
 Leu Leu Asp Ala Leu Ala Ala His Arg Arg Ala Asp Gly Leu Pro  
 3095 3100 3105  
 Ala Gln Ser Leu Ala Trp Gly Leu Trp Thr Thr Asp Ala Gly Met  
 3110 3115 3120  
 Ala Gly Asp Leu Ala Asp Thr Asp Arg Gln Arg Ile Asp Arg Ala  
 3125 3130 3135  
 Gly Leu Val Gly Leu Ser Pro Glu Gln Gly Leu Glu Leu Leu Asp  
 3140 3145 3150  
 Val Ala Gly Ala Leu Ala Thr Pro Ala Leu Val Pro Met Asn Leu  
 3155 3160 3165  
 Asp Thr Arg Thr Leu Asn Ala Ala Asp Val Pro Leu Met Leu Arg  
 3170 3175 3180  
 Gly Leu Val Arg Gly Ser Ser Arg Arg Ala Val Ser Ala Gln Ser  
 3185 3190 3195  
 Thr Gly Gly Gly Ala Ala Leu Arg Lys Arg Leu Ala Ala Leu Pro  
 3200 3205 3210  
 Ala Asp Asp Arg Tyr Asp Glu Val Leu Asp Leu Val Arg Thr His

3215		3220		3225
Ala Ala	Ala Val	Leu Gly	His	Ala Gly
3230			3235	Pro Glu
				Ala Val
				Glu Pro
Glu Arg	Ala Phe	Gly Asp	Leu	Gly Phe
3245			3250	Asp Ser
				Leu
				Gly Ala
				Val
Glu Phe	Arg Asn	Thr Leu	Asn	Ala Ala
3260			3265	Ala Gly
				Leu
				Arg Leu
				Ser
Ala Thr	Met Ile	Phe Asp	His	Pro Thr
3275			3280	Ala Arg
				Val
				Leu Ala
				Glu
His Ile	Gly Ala	Glu Leu	Ala	Pro Asp
3290			3295	Gly Asp
				Ala
				Gly Ala
				Ala
Ala Asp	Glu Glu	Thr Val	Arg	Arg Leu
3305			3310	Leu Gly
				Ser
				Ile Pro
				Leu
Thr Arg	Leu His	Asp Ala	Gly	Leu Met
3320			3325	Glu Ala
				Leu
				Leu Glu
				Leu
Ala Gly	Ser Gly	Ala Ala	Ser	Gly Met
3335			3340	Ala Ala
				Ala
				Val Pro
				Glu
Glu Asp	Ser Ile	Asp Ala	Met	Asp Ala
3350			3355	Glu Ala
				Leu
				Ile Ser
				Met
Ala Leu	Gln Asp	Thr Asp	Pro	Asp Glu
3365			3370	Ala Thr
				Arg
				Glu Val

<210> 36  
 <211> 4896  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

5

<400> 36

gtgagcgacc aggagaagct tcttgactac ctcaagcgcg cgaccacgga tctgcgca  
 60

gcgcgagac gcgtcgccga gcttgagcag cgtgaccagg agccgatcgc cattgtctcg  
 120

atggcggtgcc gctacccggg cggtgtgaac tcgcccgagg acttgtggcg gctcgtcgac  
 180

gagggcgctc acgcatctc cgagttcccc gcagaccgcg gctggggagt ggaggacatc  
 240

tacgaccccc aggtcggcaa gcccggaaag acaacgtccc gcgagggcgg attcctgcac  
 300

gacgcagcgg agttcgacgc cgagttcttc ggtatcagtc cgcgcgaggc cctggagacc  
 360

10

gacccgcagc agcggctgct gctggaggcg tcctgggagg tgctggagcg cgccgggata  
420  
gacccgcct cgctgaaggg cagcccgacc ggtgtgttcg ccggggtgat gtaccacgac  
480  
tacgccggcg gcagcagcgg cggcagcatc gtctccgggc gcgtcgccta cacgttgggc  
540  
ctggtgggtc ccgcggtcac catggacacc gcgtgctcgt cctcgtggt cgcgctgcac  
600  
acggcgatcc agtcgctgcg ttccggtgag tgctcgtcgc ctctcgcggg tggtgtgacg  
660  
gtgatgtcca cgcccagat gttcctctac ttcagcagac agcgcggcct gtcgatcgac  
720  
ggccgctgca agtcgttcag cgacgccgcc aacggcatga gctgctccga aggtgtcggg  
780  
gtactccttc tggagcggct gtcggacgcg cgtcgcctgg ggcacccggg gctggcggtg  
840  
gtgcgcggga gcgcgctgaa ccaggacggg gccagcaacg gtctgaccgc cccaacggc  
900  
ccggcgcagc gccgggtgat caagcaggcg ctggcgaagg cgggtctgtc gacggcggat  
960  
gtggacgcgg tcgaggcaca cggcacgggt acgacgctgg gtgaccgat cgaggcgcag  
1020  
gcgctgctgg agacgtacgg tcagggccgt ccggaggggc ggccgttgtg gctgggttcg  
1080  
atcaagtcga acatcggta tacgcaggcg gcggcgggtg tggcggggat catcaagatg  
1140  
gtcgaggcga tgcgccacgg caggctgcc aagacgctgc atgtggatga gccgacgaag  
1200  
caggtggact gggacgcggg tgaggtgcgg ctgctgacgg aggcgcgtga gtggccgagc  
1260  
gaaggccgtc cgcgccgcgc gggagtgtcg tcgttcgggc tcagcggaac caacgcccac  
1320  
gtcatcgtcg aagaggccgt ccccgctcag gaagtgggtg tggagcggcg ggagttgccg  
1380  
gtggcggcgg tggtggtgtc ggggaagacc ccggcggcgc tggaggcga gatcggcccg  
1440  
ttcggcgaac tggccgcgaa cgggaacgcg ctcgacgtgg cgtactcggc cgcgacaggg  
1500  
cgcgctgttc tggagcaccg ggcggtgctg gtgggccccg agaccgtgac cggctcggtc  
1560  
gccgagggca aggtggcgtt cctgttcacc gggcagggga gtcaacgcct gggatggga  
1620  
cgggagttgt acgagacgtt ccccgccctc gcgtcggcgt tcgacgaggt gtgtgcggtg  
1680



ctcgaccccg ctgtgcgtga ggtgatgtgg ggcgacgaag aagtactgag ccgtacggag  
1740  
ttcaccacagc ctgcatctt cgctcttgag gtggcgttgt tccggttggg ggagtcgtgg  
1800  
ggggtcaagc cggacttcct ggtggggcat tcgatcggtg agctggcggc ggcgcatgtg  
1860  
gcggggggtgt tcggtctgga ggatgcgggc cggttgatct cggcgcgtgg gcggttgatg  
1920  
caggcgttgc cggcgggtgg ggcgatggtc gcatccagg ccaccgaaga agaagtactc  
1980  
ccgtctca cggacgaggt gagtgtcgcg gcggtcaaca gcccgtctc ggtggtgatt  
2040  
tcgggcgagg agaaggcggg gacggcggtc gcgagcggg tcaccgaccg caagcgcaac  
2100  
cgtctgacgg tctcgcacgc ctccactca ccgtgatgg acccgatgct ggacgacttc  
2160  
cgcgagatcg ccgagagcgt cacgtacagc gaaccggcca tccggctgac caaggacgtc  
2220  
ggttcggcgg agtactgggt cgggcacgta cgcgacgcgg tgcggttcgc cgatgatgtc  
2280  
cggcatttgc aggacgaggg tgtgacgcgg ttcttgaga tcggtccgga tgggtgtgctg  
2340  
acggcgatgg cgggacagag tgccgacggc accctggcgc ccacgctgcg acgcgaccgc  
2400  
cccgaggtgg agagtgtgtt cgccgggggtg ggccggttgt tcgcgccggg tgttccggtc  
2460  
gactgggacg cgggtgttcga cgggcgtggt gcgcggcggg tggatctgcc cacgtacgcc  
2520  
ttccagcga agcgttactg gctgatcagc cagtccaagt ccacggccgg cgcggacgcc  
2580  
gtcgaccacc cctgctgac ctccggcatg gtgctcccgg acaccggcgg tgtggtgttc  
2640  
accggacgtc tctccctcga caccacccc tggctcggcg accacgacgc ctccggcggc  
2700  
gtggtcgtgc ccggggaagt gttcctggaa ctcgccctga cggcagcggg gcagacgggt  
2760  
cgcgggcgcg tcgaagcgtc cgacctcac gcgccgctgg tcatggccgg caccgaggac  
2820  
gacacgacac tgcgcgcgat ggtcgacgag gccggggcgt tcagcgtgta cgcgcgccgt  
2880  
ggcgaccgga cggcggcctg ggtcaagcac gccaccggcg tcctgacctc cgaggccacc  
2940  
gagtctcgt ggacacgtga cgacgagacg tacgccacgg tgctcgaagc cgccgcgcac  
3000

accgccggac tccaccccga cggcgagccg acgctcgccc atgtgtggcg cggcgcctcg  
 3060  
 gtgagcgtcg gccaggacgg acagcccgtt ctgtcgggtg cctcggtcga gacccgctcg  
 3120  
 atcaccacgg acgaggtacg ggcacacacc ggcggccgtg aatcgctgtt ccacgtcgac  
 3180  
 tgggtcgcag cccccgcacc cgtcggctcc gagaccggtg ctcatgtcgt ctacgagtgc  
 3240  
 ccgccactgg acgagcccgc gcccgaaggc ttccggacga tcacccatca ggtgctcccc  
 3300  
 gtcattcagc ggtggctgga cgacgagcag gcggcatccg ccaagctcgt cgtcgtcacg  
 3360  
 cgcggcgcga ccgacggaag cgacctcggg cacgcggcgg tctggggcct ggtgcgttcg  
 3420  
 gcggagtccg agcaccggg ccggttcgtc ctggtggacc tggacgccga agccgaactc  
 3480  
 cccgacgagg tgctgcgcct cgccgagccc gagatcgccg tacgggacgg cgagatccgg  
 3540  
 gtggcgaggc tggcccgcgt ccccgcggtc gaaggccgcg tccagtcgtg gggcacgcac  
 3600  
 ggcacgatcc tgatcaccgg tggcctgaac gggctcggcg ccctggccgc ccggcacctg  
 3660  
 gccgcggagc acggcgcgaa gagcctgctg ctgaccagcc gccggggcat ggacacgccg  
 3720  
 ggcgcggccg aactggtggc ggagctgacc ggatcgggcg ccgaggtgga ggtcgtcgtc  
 3780  
 tgcgacgtca ccgaccggga cgcgctggcc gccctgctgg ccgagcggcc cctgacgggc  
 3840  
 ctcgtgcact cggccggtgt gctggacgac ggactcatgg gagcgttgac gcccgagcgg  
 3900  
 ctggacacgg tcatgaggcc gaaggtcgac gccgcctggc atctgcacga actcacctg  
 3960  
 gaccacgatc tctcgtcgtt cgtgatcttc tcctccgctg ccggcacgct cggcggcgcg  
 4020  
 ggacagggca actacgccgc cgcgaacgcc tggctggacg cgctcgccag gcatcgcgcg  
 4080  
 accaaggggc tgcccgcgct gaccctcggc tggggaccct ggaccgaggt cgggggcatg  
 4140  
 gccgaccgtc tggacgacgc cgagctggcg cggctcaggc gctcgggaat gccgcccgtg  
 4200  
 tccccggacg aggggctggc gctgatggat gccgcgaccc tgcgcggccg gccgcccgatg  
 4260  
 acgctgcccg tccggttcga cctggcgggtg atgcggtccc tcgccgagac ggggacactg  
 4320

cccgcgatat tcggcggact ggtacgcggc cagaaccgtc gcgcgccggc aggccgggcg  
 4380  
 tcgtgggagc ggctgtccgg actgaccgtg ccggaacgcg agaagttcct cctggagttg  
 4440  
 gtacggggcc acgtcgccca ggtgctcggg cacggcggcg ccgacgaggt gcccccgac  
 4500  
 cgcgctttca acgaactcgg cctcacctcg ctgggcgcgg tggaactgcg caacgccttc  
 4560  
 aacaccgaga cagggctgtc actgcccccg acctggcct tcgactaccg gacccccctg  
 4620  
 gccatcgccg agctcgtgca cgacgggctg cgccccggagg aggccgacgg ggcggccgcg  
 4680  
 ctgctggccg aactgaaccg gctggaagcc gtgctgaccg agctggcacc gggcgaccgc  
 4740  
 gaagcccacg ccaaggtcac ggccccctc gaaaccatgt cgcgacgggc gcgagacgcc  
 4800  
 gggagcggcg tgacggacga gcccgtcggc ggccctcgtga ccgaatccga tgacgagttg  
 4860  
 ttcgcggtgc tgaacaagga actcggactg tcctga  
 4896

<210> 37  
 <211> 1631  
 <212> PRT  
 <213> *Streptomyces* sp. MP28-13  
 <400> 37

5

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

10

ES 2 543 203 T3

Ser Pro Arg Glu Ala Leu Glu Thr Asp Pro Gln Gln Arg Leu Leu Leu  
 115 120 125  
 Glu Ala Ser Trp Glu Val Leu Glu Arg Ala Gly Ile Asp Pro Ala Ser  
 130 135 140  
 Leu Lys Gly Ser Pro Thr Gly Val Phe Ala Gly Val Met Tyr His Asp  
 145 150 155 160  
 Tyr Ala Gly Gly Ser Ser Gly Gly Ser Ile Val Ser Gly Arg Val Ala  
 165 170 175  
 Tyr Thr Leu Gly Leu Val Gly Pro Ala Val Thr Met Asp Thr Ala Cys  
 180 185 190  
 Ser Ser Ser Leu Val Ala Leu His Thr Ala Ile Gln Ser Leu Arg Ser  
 195 200 205  
 Gly Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr  
 210 215 220  
 Pro Glu Met Phe Leu Tyr Phe Ser Glu Gln Arg Gly Leu Ser Ile Asp  
 225 230 235 240  
 Gly Arg Cys Lys Ser Phe Ser Asp Ala Ala Asn Gly Met Ser Cys Ser  
 245 250 255  
 Glu Gly Val Gly Val Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg  
 260 265 270  
 Leu Gly His Pro Val Leu Ala Val Val Arg Gly Ser Ala Leu Asn Gln  
 275 280 285  
 Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala Gln Arg  
 290 295 300  
 Arg Val Ile Lys Gln Ala Leu Ala Lys Ala Gly Leu Ser Thr Ala Asp  
 305 310 315 320  
 Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro  
 325 330 335  
 Ile Glu Ala Gln Ala Leu Leu Glu Thr Tyr Gly Gln Gly Arg Pro Glu  
 340 345 350  
 Gly Arg Pro Leu Trp Leu Gly Ser Ile Lys Ser Asn Ile Gly His Thr  
 355 360 365  
 Gln Ala Ala Ala Gly Val Ala Gly Ile Ile Lys Met Val Glu Ala Met

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

Gln Ala Leu Pro Ala Gly Gly Ala Met Val Ala Ile Gln Ala Thr Glu  
 645 650 655  
 Glu Glu Val Leu Pro Leu Leu Thr Asp Glu Val Ser Val Ala Ala Val  
 660 665 670  
 Asn Ser Pro Ser Ser Val Val Ile Ser Gly Glu Glu Lys Ala Val Thr  
 675 680 685  
 Ala Val Ala Glu Arg Phe Thr Asp Arg Lys Arg Asn Arg Leu Thr Val  
 690 700  
 Ser His Ala Phe His Ser Pro Leu Met Asp Pro Met Leu Asp Asp Phe  
 705 710 715 720  
 Arg Glu Ile Ala Glu Ser Val Thr Tyr Ser Glu Pro Ala Ile Arg Leu  
 725 730 735  
 Thr Lys Asp Val Gly Ser Ala Glu Tyr Trp Val Gly His Val Arg Asp  
 740 745 750  
 Ala Val Arg Phe Ala Asp Asp Val Arg His Leu Gln Asp Glu Gly Val  
 755 760 765  
 Thr Arg Phe Leu Glu Ile Gly Pro Asp Gly Val Leu Thr Ala Met Ala  
 770 775 780  
 Gly Gln Ser Ala Asp Gly Thr Leu Ala Pro Thr Leu Arg Arg Asp Arg  
 785 790 795 800  
 Pro Glu Val Glu Ser Val Phe Ala Gly Val Gly Arg Leu Phe Ala Ala  
 805 810  
 Gly Val Pro Val Asp Trp Asp Ala Val Phe Asp Gly Arg Gly Ala Arg  
 820 825 830  
 Arg Val Asp Leu Pro Thr Tyr Ala Phe Gln Arg Lys Arg Tyr Trp Leu  
 835 840 845  
 Ile Glu Gln Ser Lys Ser Thr Ala Gly Ala Asp Ala Val Asp His Pro  
 850 855 860  
 Leu Leu Thr Ser Gly Met Val Leu Pro Asp Thr Gly Gly Val Val Phe  
 865 870 875 880  
 Thr Gly Arg Leu Ser Leu Asp Thr His Pro Trp Leu Ala Asp His Asp  
 885 890 895  
 Ala Phe Gly Ala Val Val Val Pro Gly Glu Val Phe Leu Glu Leu Ala

900					905					910					
Leu	Thr	Ala	Ala	Glu	Gln	Thr	Gly	Arg	Gly	Arg	Val	Glu	Ala	Leu	Asp
		915					920					925			
Leu	His	Ala	Pro	Leu	Val	Met	Ala	Gly	Thr	Glu	Asp	Asp	Thr	Thr	Leu
	930					935					940				
Arg	Ala	Met	Val	Asp	Glu	Ala	Gly	Ala	Phe	Ser	Val	Tyr	Ala	Arg	Arg
945					950					955					960
Gly	Asp	Arg	Thr	Ala	Ala	Trp	Val	Lys	His	Ala	Thr	Gly	Val	Leu	Thr
				965					970					975	
Ser	Glu	Ala	Thr	Glu	Ser	Ser	Trp	Thr	Arg	Asp	Asp	Glu	Thr	Tyr	Ala
			980					985					990		
Thr	Val	Leu	Glu	Ala	Ala	Ala	His	Thr	Ala	Gly	Leu	His	Pro	Asp	Gly
		995					1000					1005			
Glu	Pro	Thr	Leu	Ala	His	Val	Trp	Arg	Gly	Ala	Ser	Val	Ser	Val	
	1010					1015					1020				
Gly	Gln	Asp	Gly	Gln	Pro	Val	Leu	Ser	Val	Ala	Ser	Val	Glu	Thr	
	1025					1030					1035				
Arg	Ser	Ile	Thr	Thr	Asp	Glu	Val	Arg	Ala	His	Thr	Gly	Gly	Arg	
	1040				1045						1050				
Glu	Ser	Leu	Phe	His	Val	Asp	Trp	Val	Ala	Ala	Pro	Ala	Pro	Val	
	1055				1060						1065				
Gly	Ser	Glu	Thr	Gly	Ala	His	Val	Val	Tyr	Glu	Cys	Pro	Pro	Leu	
	1070				1075						1080				
Asp	Glu	Pro	Ala	Pro	Glu	Gly	Phe	Arg	Thr	Ile	Thr	His	Gln	Val	
	1085				1090						1095				
Leu	Pro	Val	Ile	Gln	Arg	Trp	Leu	Asp	Asp	Glu	Gln	Ala	Ala	Ser	
	1100				1105						1110				
Ala	Lys	Leu	Val	Val	Val	Thr	Arg	Gly	Ala	Thr	Asp	Gly	Ser	Asp	
	1115				1120						1125				
Leu	Gly	His	Ala	Ala	Val	Trp	Gly	Leu	Val	Arg	Ser	Ala	Glu	Ser	
	1130				1135						1140				
Glu	His	Pro	Gly	Arg	Phe	Val	Leu	Val	Asp	Leu	Asp	Ala	Glu	Ala	
	1145				1150						1155				

Glu Leu Pro Asp Glu Val Leu Arg Leu Ala Glu Pro Glu Ile Ala  
 1160 1165 1170  
 Val Arg Asp Gly Glu Ile Arg Val Ala Arg Leu Ala Arg Val Pro  
 1175 1180 1185  
 Ala Val Glu Gly Arg Val Gln Ser Trp Gly Thr His Gly Thr Ile  
 1190 1195 1200  
 Leu Ile Thr Gly Gly Leu Asn Gly Leu Gly Ala Leu Ala Ala Arg  
 1205 1210 1215  
 His Leu Ala Ala Glu His Gly Ala Lys Ser Leu Leu Leu Thr Ser  
 1220 1225 1230  
 Arg Arg Gly Met Asp Thr Pro Gly Ala Ala Glu Leu Val Ala Glu  
 1235 1240 1245  
 Leu Thr Gly Ser Gly Ala Glu Val Glu Val Val Val Cys Asp Val  
 1250 1255 1260  
 Thr Asp Arg Asp Ala Leu Ala Ala Leu Leu Ala Glu Arg Pro Leu  
 1265 1270 1275  
 Thr Gly Leu Val His Ser Ala Gly Val Leu Asp Asp Gly Leu Met  
 1280 1285 1290  
 Gly Ala Leu Thr Pro Glu Arg Leu Asp Thr Val Met Arg Pro Lys  
 1295 1300 1305  
 Val Asp Ala Ala Trp His Leu His Glu Leu Thr Leu Asp His Asp  
 1310 1315 1320  
 Leu Ser Ser Phe Val Ile Phe Ser Ser Val Ala Gly Thr Leu Gly  
 1325 1330 1335  
 Gly Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Trp Leu Asp  
 1340 1345 1350  
 Ala Leu Ala Arg His Arg Ala Thr Lys Gly Leu Pro Ala Leu Thr  
 1355 1360 1365  
 Leu Gly Trp Gly Pro Trp Thr Glu Val Gly Gly Met Ala Asp Arg  
 1370 1375 1380  
 Leu Asp Asp Ala Glu Leu Ala Arg Leu Arg Arg Ser Gly Met Pro  
 1385 1390 1395  
 Pro Leu Ser Pro Asp Glu Gly Leu Ala Leu Met Asp Ala Ala Thr



1400						1405						1410		
Leu	Arg	Gly	Arg	Pro	Pro	Met	Thr	Leu	Pro	Val	Arg	Phe	Asp	Leu
	1415					1420					1425			
Ala	Val	Met	Arg	Ser	Leu	Ala	Glu	Thr	Gly	Thr	Leu	Pro	Ala	Ile
	1430					1435					1440			
Phe	Gly	Gly	Leu	Val	Arg	Gly	Gln	Asn	Arg	Arg	Ala	Pro	Ala	Gly
	1445					1450					1455			
Arg	Ala	Ser	Trp	Glu	Arg	Leu	Ser	Gly	Leu	Thr	Val	Pro	Glu	Arg
	1460					1465					1470			
Glu	Lys	Phe	Leu	Leu	Glu	Leu	Val	Arg	Gly	His	Val	Ala	Gln	Val
	1475					1480					1485			
Leu	Gly	His	Gly	Gly	Ala	Asp	Glu	Val	Pro	Pro	Asp	Arg	Ala	Phe
	1490					1495					1500			
Asn	Glu	Leu	Gly	Leu	Thr	Ser	Leu	Gly	Ala	Val	Glu	Leu	Arg	Asn
	1505					1510					1515			
Ala	Leu	Asn	Thr	Glu	Thr	Gly	Leu	Ser	Leu	Pro	Pro	Thr	Leu	Ala
	1520					1525					1530			
Phe	Asp	Tyr	Pro	Thr	Pro	Leu	Ala	Ile	Ala	Glu	Leu	Val	His	Asp
	1535					1540					1545			
Gly	Leu	Arg	Pro	Glu	Glu	Ala	Asp	Gly	Ala	Ala	Ala	Leu	Leu	Ala
	1550					1555					1560			
Glu	Leu	Asn	Arg	Leu	Glu	Ala	Val	Leu	Thr	Glu	Leu	Ala	Ser	Gly
	1565					1570					1575			
Asp	Arg	Glu	Ala	His	Ala	Lys	Val	Thr	Ala	Arg	Leu	Glu	Thr	Met
	1580					1585					1590			
Ser	Arg	Arg	Ala	Arg	Asp	Ala	Gly	Ser	Gly	Val	Thr	Asp	Glu	Pro
	1595					1600					1605			
Val	Gly	Gly	Leu	Val	Thr	Glu	Ser	Asp	Asp	Glu	Leu	Phe	Ala	Val
	1610					1615					1620			
Leu	Asn	Lys	Glu	Leu	Gly	Leu	Ser							
	1625					1630								

<210> 38  
 <211> 288  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

5

<400> 38

atgcaccctg actacgtggc cggcccgcga acacgggaac ccgtcgccgt cacacctgcc  
60  
gagcacctcg tgcgcgtcgt caggggagcgt ccgggagaga tggaaactggc ggcgctgctc  
120  
gcggtgttga gcgccctgac cggatcccgg gagcaggccg tcgcgtcctc cgcggcccga  
180  
cagcggcgca actccgcctg gggcgcggtc cgttacgcct caccgggtgc ctggtacaca  
240  
ccgcccggcg cctggacggg cggccaggac agagagtggc ggcactga  
288

5  
<210> 39  
<211> 95  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

10  
<400> 39  
Met His Pro Asp Tyr Val Ala Gly Pro Arg Thr Arg Glu Pro Val Ala  
1 5 10 15  
Val Thr Pro Ala Glu His Leu Val Arg Val Val Arg Gly Ala Pro Gly  
20 25 30  
Glu Met Glu Leu Ala Ala Leu Leu Ala Val Leu Ser Ala Leu Thr Gly  
35 40 45  
Thr Arg Glu Gln Ala Val Ala Ser Ser Ala Ala Arg Gln Arg Arg Asn  
50 55 60  
Ser Ala Trp Gly Ala Val Arg Tyr Ala Ser Pro Gly Ala Trp Tyr Thr  
65 70 75 80  
Pro Pro Gly Ala Trp Thr Gly Gly Gln Asp Arg Glu Trp Arg His  
85 90 95

15  
<210> 40  
<211> 774  
<212> ADN  
<213> *Streptomyces* sp. MP28-13

20  
<400> 40  
atgaccgcca tatcgagcga caacgcgtcc aattggattc gagaatttca tccggcggac  
60  
cggacatccc cgaggatgat ctgcttcccc cacgcgggag gtgcggcgag ctactacttc  
120  
cccgtctccc gggcgctggc cgggaagatc gaagtcctcg ccatccagta cccggggcgc  
180

caggaccggtt acacggaacc ggccatcggc aacgtcgagg ccctcgccgc cgcggtcttc  
240  
cgtgagcttc cgacggagga cctggaccgg acctggctct tcgggcacag catgggggccc  
300  
gccgtcgctt tcgaggtggc ccggctgatg gaacgggagt tgaaccagtc gcctgtcggg  
360  
atcatcctct ccggccggcg cgcaccgtcc cggttccgtc ccgagaccct ccacctgcag  
420  
ggcgacgcgg cgatcatcgc caacgtgcag tcgctcagcg gtaccgacgc gatcctcttc  
480  
gaggaccccg acacccagcg gctgatcatg ccggcgctgc gagccgacta ccggggccatc  
540  
gagacctacc ggccgcccgg cactccacgc gtcgctgtcc cgatccacac cttcgtgggc  
600  
gacgccgacc cgttggccac gctggacgag gtcggcagct ggcgcgacca cacctcggcc  
660  
gagtacaccc tgcgcgtttt cccgggtgac cacttctatc tgacggcgcg tgccgtcgag  
720  
gtcatctccg cgatctccca gctgatcgtg gagcccaccc agaccgcggg ctga  
774

<210> 41  
<211> 257  
<212> PRT  
<213> *Streptomyces* sp. MP28-13

5

<400> 41

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

10

Glu Leu Asn Gln Ser Pro Val Gly Ile Ile Leu Ser Gly Arg Arg Ala  
 115 120 125  
 Pro Ser Arg Phe Arg Pro Glu Thr Leu His Leu Gln Gly Asp Ala Ala  
 130 135 140  
 Ile Ile Ala Asn Val Gln Ser Leu Ser Gly Thr Asp Ala Ile Leu Phe  
 145 150 155 160  
 Glu Asp Pro Asp Thr Gln Arg Leu Ile Met Pro Ala Leu Arg Ala Asp  
 165 170 175  
 Tyr Arg Ala Ile Glu Thr Tyr Arg Pro Pro Gly Thr Pro Arg Val Ala  
 180 185 190  
 Cys Pro Ile His Thr Phe Val Gly Asp Ala Asp Pro Leu Ala Thr Leu  
 195 200 205  
 Asp Glu Val Gly Ser Trp Arg Asp His Thr Ser Ala Glu Tyr Thr Leu  
 210 215 220  
 Arg Val Phe Pro Gly Asp His Phe Tyr Leu Thr Ala Arg Ala Val Glu  
 225 230 235 240  
 Val Ile Ser Ala Ile Ser Gln Leu Ile Val Glu Pro Thr Gln Thr Arg  
 245 250 255

Gly

<210> 42  
 <211> 714  
 <212> ADN  
 <213> *Streptomyces* sp. MP28-13

5

<400> 42

gtgttctact acgtactgaa gtacgtgctg ttggggcccg tgctgcggtt gctcttccgg  
 60  
 ccccgatcg aggggctcga acacatcccg gcggacggcg ccgcatcgt cgcgggcaat  
 120  
 catctctctt tctccacca cttctgatg cccgcatca tccggcggcg gatcacgttt  
 180  
 ctgcgaagg ccgagtactt caccgggtccc ggcgtcaagg gacgcctcac cgcctccttc  
 240  
 ttccgcgggc tcggccagat cccggtcgac cgggccggca aggaggccgg gaaggccgcg  
 300  
 atccgggaag ggctcggggg gctcggcaag ggtgagttgc tggggatcta cccggagggc  
 360  
 acgcgctcgc acgacggacg gctctacaag ggcaaggctc ggggtggcggg gatggccatc  
 420

10

agggcgcagg tcccgggtggt gccgtgcgcg atggtgggta cgttcgagat ccagccgccc  
480  
ggtcagaaga tcccgaacat ccggcggggtc acgatccggt tcggtgagcc gctggacttc  
540  
tcgcgctacg cgggtctgga gaaccagaag gcggcgggtcc gcgcggtcac cgacgagatc  
600  
atgtacgcga tcctcgggtct gtccggggcag gagtacgtgg accggtacgc cgccgaggtg  
660  
aaggccgagg aggcgcagca ggcgccgaag aagttcccgc gcctgcgacg ctga  
714

<210> 43  
<211> 237  
<212> PRT  
<213> *Streptomyces* sp. MP28-13  
  
<400> 43

5

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

10

Pro Leu Asp Phe Ser Arg Tyr Ala Gly Leu Glu Asn Gln Lys Ala Ala  
 180 185 190  
 Val Arg Ala Val Thr Asp Glu Ile Met Tyr Ala Ile Leu Gly Leu Ser  
 195 200 205  
 Gly Gln Glu Tyr Val Asp Arg Tyr Ala Ala Glu Val Lys Ala Glu Glu  
 210 215 220  
 Ala Gln Gln Ala Pro Lys Lys Phe Pro Arg Leu Arg Arg  
 225 230 235

<210> 44  
 <211> 1611  
 <212> ADN  
 <213> *Streptomyces sp.* MP28-13

5

<400> 44

ttgagacgca cagcagtgct tggctcggcc ggcactctgc tcgcgggcac gctcatagcg  
 60  
 accacgctcg ccgcctcgtc ggcgagcgcc gacggccgcg ccgacagtgg cccggccgcg  
 120  
 tcgggcaagg aacagaagac gctgaacgcc caggcggccg gcgccgagat cgccgccgag  
 180  
 cgggccgca agaagggcat cgactggcg gactgcccgg ccgactgggc cctggagaag  
 240  
 ccgatccagt gcggctacat cagtgtcccg ctcgactacg cccgtccgaa cggcaagcag  
 300  
 atcaagatag ccgtcgaccg gatcggcaac accgggacgg cgggtgagcg tcagggctcc  
 360  
 ctcgtctaca accccggtgg ccccggcgcg tccggcatgg cgttcccgcg ccgcgtcgtg  
 420  
 accaagaacg ccatctgggc ggacgccgcc aaggcgtacg acttcgtcgg cttcgacccg  
 480  
 cgcggcgtcg ggcgctcgac gcccatctcc tgcgtcgacc cgcaggagtt cgtcaaggct  
 540  
 cccaaggccg accccgtccc cgacaccgag gccgacaagc gcgccccagcg caagctcgcg  
 600  
 gccgagtacg cggacggctg caaggagcgc agcggctgga tgctgccgca catgacgacg  
 660  
 cccaacagcg cgcgcgacct ggacgtcctg cgggccgcg tcggcgacaa gaagctcaac  
 720  
 tacgtgggtg tctcctacgg cacctacctg ggcgccgtct acggcacgct cttcccgtcc  
 780  
 catgtacggc gcatggtgct ggacagcgtg gtcaaccctg cgaaggagaa gatctggtac  
 840

10

caggccaacc tggaccagga cgtcgccttc gagacacgct tcgacgactg gaagaagtgg  
 900  
 gtcgccgaga acgacgcggc gttccacatc ggcgacaccg tcgccaaggt cgagaagcag  
 960  
 tgggacaagc tccgcgccac cgccaagaag aaccgatcg gcggcgtcgt gggaccggcc  
 1020  
 gaactcatcg ggctgttcca gagcgcgccc tactacgact ccagctgggt gccggtcgcc  
 1080  
 gacacctgga gcaagtacct ggccggagac acccaggcgc tcgtcgacgc cgccgcgccg  
 1140  
 gacctgtccg acacggtggg caacacccgc gccgagaaca gcaacgcggt ctacaccgcc  
 1200  
 gtcgagtgcg ccgacgcca gtggcccacg agctggcgca cctgggaccg ggacaacacc  
 1260  
 cggtccacc gcgaccacc gttctcacc tgggccaacg cctggatgaa cctgccctgt  
 1320  
 tcgacctggg gcgtagagca gcagacaccg atcgaggctg gtacgggccg cggcctgccg  
 1380  
 cccgtactga tcgcgcagtc cacgcgtgac gccgccacc cgtacggggg cgccgtcgag  
 1440  
 ctgcacaagc gcttcaaggg ctcacgcctg atcaccgagc gggacgcccg ttcgcacggc  
 1500  
 atcaccaatg tcgcgaacgc gtgcatcaac gaccgggtcg agtcgtatct gctcagcggc  
 1560  
 gagctggacc gccgcgacgt gacgtgcgca ccgcacgcca cacccaagcc g  
 1611

<210> 45  
 <211> 537  
 <212> PRT  
 <213> *Streptomyces* sp. MP28-13  
 <400> 45

5

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

10

Pro Ile Gln Cys Gly Tyr Ile Ser Val Pro Leu Asp Tyr Ala Arg Pro  
 85 90 95  
 Asn Gly Lys Gln Ile Lys Ile Ala Val Asp Arg Ile Gly Asn Thr Gly  
 100 105 110  
 Thr Ala Gly Glu Arg Gln Gly Ser Leu Val Tyr Asn Pro Gly Gly Pro  
 115 120 125  
 Gly Ala Ser Gly Met Ala Phe Pro Arg Arg Val Val Thr Lys Asn Ala  
 130 135 140  
 Ile Trp Ala Asp Ala Ala Lys Ala Tyr Asp Phe Val Gly Phe Asp Pro  
 145 150 155 160  
 Arg Gly Val Gly Arg Ser Thr Pro Ile Ser Cys Val Asp Pro Gln Glu  
 165 170 175  
 Phe Val Lys Ala Pro Lys Ala Asp Pro Val Pro Asp Thr Glu Ala Asp  
 180 185 190  
 Lys Arg Ala Gln Arg Lys Leu Ala Ala Glu Tyr Ala Asp Gly Cys Lys  
 195 200 205  
 Glu Arg Ser Gly Trp Met Leu Pro His Met Thr Thr Pro Asn Ser Ala  
 210 215 220  
 Arg Asp Leu Asp Val Leu Arg Ala Ala Leu Gly Asp Lys Lys Leu Asn  
 225 230 235 240  
 Tyr Val Gly Val Ser Tyr Gly Thr Tyr Leu Gly Ala Val Tyr Gly Thr  
 245 250 255  
 Leu Phe Pro Ser His Val Arg Arg Met Val Leu Asp Ser Val Val Asn  
 260 265 270  
 Pro Ser Lys Glu Lys Ile Trp Tyr Gln Ala Asn Leu Asp Gln Asp Val  
 275 280 285  
 Ala Phe Glu Thr Arg Phe Asp Asp Trp Lys Lys Trp Val Ala Glu Asn  
 290 295 300  
 Asp Ala Ala Phe His Ile Gly Asp Thr Val Ala Lys Val Glu Lys Gln  
 305 310 315 320  
 Trp Asp Lys Leu Arg Ala Thr Ala Lys Lys Asn Pro Ile Gly Gly Val  
 325 330 335  
 Val Gly Pro Ala Glu Leu Ile Gly Leu Phe Gln Ser Ala Pro Tyr Tyr



		340						345						350		
Asp	Ser	Ser	Trp	Val	Pro	Val	Ala	Asp	Thr	Trp	Ser	Lys	Tyr	Leu	Ala	
		355					360					365				
Gly	Asp	Thr	Gln	Ala	Leu	Val	Asp	Ala	Ala	Ala	Pro	Asp	Leu	Ser	Asp	
	370					375					380					
Thr	Val	Gly	Asn	Thr	Arg	Ala	Glu	Asn	Ser	Asn	Ala	Val	Tyr	Thr	Ala	
385					390					395					400	
Val	Glu	Cys	Ala	Asp	Ala	Lys	Trp	Pro	Thr	Ser	Trp	Arg	Thr	Trp	Asp	
				405					410					415		
Arg	Asp	Asn	Thr	Arg	Leu	His	Arg	Asp	His	Pro	Phe	Leu	Thr	Trp	Ala	
			420					425					430			
Asn	Ala	Trp	Met	Asn	Leu	Pro	Cys	Ser	Thr	Trp	Gly	Val	Glu	Gln	Gln	
		435					440					445				
Thr	Pro	Ile	Glu	Val	Gly	Thr	Gly	Arg	Gly	Leu	Pro	Pro	Val	Leu	Ile	
	450					455					460					
Ala	Gln	Ser	Thr	Arg	Asp	Ala	Ala	Thr	Pro	Tyr	Gly	Gly	Ala	Val	Glu	
465					470					475					480	
Leu	His	Lys	Arg	Phe	Lys	Gly	Ser	Arg	Leu	Ile	Thr	Glu	Arg	Asp	Ala	
				485					490					495		
Gly	Ser	His	Gly	Ile	Thr	Asn	Val	Ala	Asn	Ala	Cys	Ile	Asn	Asp	Arg	
			500					505					510			
Val	Glu	Ser	Tyr	Leu	Leu	Ser	Gly	Glu	Leu	Asp	Arg	Arg	Asp	Val	Thr	
		515					520					525				
Cys	Ala	Pro	His	Ala	Thr	Pro	Lys	Pro								
	530					535										

&lt;210&gt; 46

&lt;211&gt; 1490

&lt;212&gt; ADN

<213> *Streptomyces* sp. MP28-13

&lt;400&gt; 46

tagagtttga tcatggctca ggacgaacgc tggcggcgtg cttaacacat gcaagtcgaa  
60cgatgaagcc ttcgggtgga ttagtggcga acgggtgagt aacacgtggg caatctgccc  
120tgcactctgg gacaagccct ggaaacgggg tctaataccg gataatactg tgcccctcat  
180

5

10

gggggacggt tgaaagctcc ggcggtgcag gatgagcccc cggcctatca gcttgttggt  
 240  
 ggggtaatgg cctaccaagg cgacgacggg tagccggcct gagagggcga ccggccacac  
 300  
 tgggactgag acacggccca gactcctacg ggaggcagca gtggggaata ttgcacaatg  
 360  
 ggcgaaagcc tgatgcagcg acgccgcgtg agggatgacg gccttcgggt tgtaaacctc  
 420  
 tttcagcagg gaagaagcga aagtgcgggt acctgcagaa gaagcgcggg ctaactacgt  
 480  
 gccagcagcc gcggtataac gtagggcgca agcgttgtcc ggaattattg ggcgtaaaga  
 540  
 gctcgtaggc ggtctgtcac gtcgggtgtg aaagcccggg gcttaacccc gggctctgat  
 600  
 tcgatacggg cagactagag tgtggtaggg gagatcggaa ttcttggtgt agcggtgaaa  
 660  
 tgcgcagata tcaggaggaa caccggtggc gaaggcggat ctctgggcca ttactgacgc  
 720  
 tgaggagcga aagcgtgggg agcgaacagg attagatacc ctggtagtcc acgccgtaaa  
 780  
 cgttgggaac taggtggttg cgacattcca cgtcgtcggg gccgcagcta acgcattaag  
 840  
 ttccccgcct ggggagtacg gccgcaaggc taaaactcaa aggaattgac gggggcccgc  
 900  
 acaagcagcg gagcatgtgg cttaatcga cgcaacgcga agaaccttac caaggcttga  
 960  
 catacaccgg aaagcatcag agatggtgcc ccccttggtg tcggtgtaca ggtggtgcat  
 1020  
 ggctgctgtc agctcgtgtc gtgagatggt gggttaagtc ccgcaacgag cgcaaccctt  
 1080  
 gttctgtggt gccagcatgc ctttcggggt gatggggact cacaggagac cgccggggtc  
 1140  
 aactcggagg aaggtgggga cgacgtcaag tcatcatgcc cttatgtct tgggctgcac  
 1200  
 acgtgctaca atggccggta caatgagctg cgataaccgca aggtggagcg aatctcaaaa  
 1260  
 agccggtctc agttcggatt ggggtctgca actcgacccc atgaagtcgg agttgctagt  
 1320  
 aatcgcagat cagcattgct gcggtgaata cgttccccgg ccttgtagac accgcccgtc  
 1380  
 acgtcacgaa agtcggtaac acccgaagcc ggtggcccaa ccccttggtg gagggagctg  
 1440  
 tcgaaggtgg gactggcgat tgggacgaag tcgtaacaag gtagccgtaa  
 1490

<210> 47  
 <211> 20  
 <212> ADN

ES 2 543 203 T3

<213> Secuencia artificial  
<220>  
<223> cebador degenerado KSMA-D

5 <400> 47

**tsgcsatgga cccscagcag**  
**20**

<210> 48  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> cebador degenerado KSMB-I

10  
15 <400> 48

**ccsgtscctgt gsgcctcsac**  
**20**

<210> 49  
<211> 17  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> cebador directo M13 (-20)

20  
25 <400> 49

**gtaaaacgac ggccagt**  
**17**

<210> 50  
<211> 16  
<212> ADN  
<213> Secuencia artificial

30  
35 <220>  
<223> cebador inverso M13

<400> 50

40

**aacagctatg accatg**  
**16**

<210> 51  
<211> 725  
<212> ADN  
<213> *Streptomyces sp.* MP28-13

<220>  
<221> misc\_feature  
<222> (633)..(633)  
<223> n es a, c, g o t

45  
50 <400> 51

tccggtcccg tgcgcctcga ccgcgtccac atccgccgtc gacagacccg ccttcgccag  
60  
cgcctgcttg atcacgcgac gctgggaggg gccgttcggc gcggtgatgc cgttgctggc  
120  
gccgtcctgg ttgagcgcgc tcccgcgcac caccgcgagc accggatgcc ccagccgacg  
180  
cgcctccgac agacgctcca ccaccaggac gcccgcgccc tcgccccacc cggtgccgtc  
240  
ggtggaggac gagaaggact tgcagcggcc gtccgcggcc aggccgcgct gctcgtgaa  
300  
gtcgatgaac gaccgcggtg ttcccatgac ggtcacgccg cccacgaggg cgagcgagca  
360  
ctcgcggaa cgcagtgcct gcgcccacca gtgcagggcc accagggacg acgagcatgc  
420  
cgtgtccacg ctcaccgcgg ggccttcgag accgagggtg taggccaccc ggcccgacac  
480  
gaggctgccg ccgccgggtc cgccggggta gtcgtggtac atcaccgccg cgaacacacc  
540  
ggtcgggctg cccttcagcg tggtggtgc gatccccgca cgctccagcg cctcccacga  
600  
ggtctccagc agcaaccgct gctgcgggtc canggccgaa atcacgaatt ctggatccga  
660  
tacgtaacgc gtctgcagca tgcgtggtac cgagctttcc ctatagtgag tcgtattaga  
720  
gcttg  
725

5 <210> 52  
<211> 19  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<223> cebador SuperCos directo  
<400> 52

ggccgcaatt aaccctcac  
19

15 <210> 53  
<211> 20  
<212> ADN  
<213> Secuencia artificial

20 <220>  
<223> Cebador SuperCos inverso  
<400> 53

ggccgcataa tacgactcac  
20

25

## REIVINDICACIONES

1. Una molécula de ácido nucleico que comprende:

- 5 (a) una secuencia de nucleótidos como se muestra en la SEC ID N° 1; o  
 (b) una secuencia de nucleótidos que es el complemento de la SEC ID N° 1; o  
 (c) una secuencia de nucleótidos que tiene una identidad de secuencia de al menos 85 % (preferentemente una  
 identidad de secuencia de al menos 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %) con  
 la SEC ID N° 1; o  
 10 (d) una parte de una cualquiera de (a) a (c), en la que dicha parte comprende:

(i) una secuencia de nucleótidos como se muestra en una cualquiera de las SEC ID N° 2, 6, 8, 10, 12, 14, 16,  
 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 o 44 o una secuencia de nucleótidos que es complementaria de  
 las mismas o que tiene una identidad de secuencia de al menos 85 % con las mismas; o

- 15 (ii) una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos seleccionadas de SEC  
 ID N° 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45 o que tiene una identidad de  
 secuencia de al menos 85 % (preferentemente al menos 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %  
 o 99 %) con las mismas,

20 en la que dicha molécula de ácido nucleico codifica o es complementaria a una molécula de ácido nucleico que  
 codifica uno o más polipéptidos que tienen actividad funcional en la síntesis de BE-14106.

2. La molécula de ácido nucleico de la reivindicación 1, en la que dicha molécula codifica un sistema biosintético  
 NRPS-PKS para la síntesis de BE-14106.

25 3. Un polipéptido codificado por una molécula de ácido nucleico como se define en la reivindicación 1 o 2, en el que  
 dicho polipéptido comprende:

- 30 (a) la totalidad de una secuencia de aminoácidos como se muestra en una cualquiera o más de las SEC ID N° 3,  
 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45; o  
 (b) la totalidad de una secuencia de aminoácidos que tiene una identidad de secuencia de al menos 85 % con  
 una cualquiera o más de las SEC ID N° 3, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 o 45,  
 preferentemente al menos una identidad de secuencia del 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97  
 %, 98 % o 99 %, en el que dicho polipéptido tiene una actividad funcional en la síntesis de BE-14106.

35 4. El uso de una molécula de ácido nucleico como se define en la reivindicación 1 o 2 en la preparación de un grupo  
 de genes biosintéticos de BE-14106 para la preparación de una molécula de BE-14106 modificada.

40 5. Un método de preparación de una molécula de ácido nucleico que codifica un sistema NRPS-PKS modificado de  
 BE-14106, comprendiendo dicho método modificar una molécula de ácido nucleico que codifica dicho sistema  
 NRPS-PKS de BE-14106 y que comprende:

- (a) una secuencia de nucleótidos como se muestra en la SEC ID N° 1; o  
 45 (b) una secuencia de nucleótidos que tiene una identidad de secuencia de al menos 85 % (preferentemente una  
 identidad de secuencia de al menos 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %) con  
 la SEC ID N° 1; o  
 (c) una parte de (a) o (b).

50 6. El método de la reivindicación 5 en el que dicha molécula de ácido nucleico se modifica introduciendo, mutando,  
 delecionando, reemplazando o inactivando una o más actividades o proteínas  
 codificadas por dicha molécula de ácido nucleico.

7. El método de la reivindicación 5 o 6 en el que se modifica una o más de las secuencias de nucleótidos  
 seleccionadas de:

- 55 (i) una secuencia de nucleótidos como se muestra en una cualquiera de las SEC ID N° 2, 4, 6, 8, 10, 12, 14, 16,  
 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 o 44; o  
 (ii) una secuencia de nucleótidos que tiene una identidad de secuencia de al menos 85 % con una secuencia de  
 nucleótidos como se muestra en una cualquiera de las SEC ID N° 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28,  
 30, 32, 34, 36, 38, 40, 42 o 44 y que codifica un polipéptido que tiene actividad funcional en la síntesis de BE-  
 60 14106; o  
 (iii) una secuencia de nucleótidos que es complementaria de (i) o (ii).

65 8. El método de una cualquiera de las reivindicaciones 5 a 7 en el que dicha molécula de ácido nucleico se modifica  
 de una o más de las siguientes maneras:

- (i) modificación de una secuencia que codifica una PKS para modificar un dominio de carga para alterar la naturaleza de la unidad iniciadora;
- (ii) modificación de una secuencia que codifica una PKS para modificar el número de módulos, preferentemente para disminuir el número de módulos;
- 5 (iii) modificación de una secuencia que codifica una PKS para modificar un dominio AT para alterar su especificidad por una unidad extendidora;
- (iv) modificación de una secuencia que codifica una PKS para alterar la actividad de un dominio de deshidratasa (DH) o de cetorreductasa (KR), preferentemente para inactivar o deletar un dominio DK o KR;
- 10 (v) modificación de una secuencia que codifica una hidroxilasa como se muestra en la SEC ID N° 26 para inactivar la enzima hidroxilasa o alterar su especificidad;
- (vi) delección de una secuencia que codifica una PKS o la modificación de una secuencia que codifica una PKS para inactivar la enzima PKS codificada;
- (vii) introducción de una secuencia de nucleótidos que codifica una enzima de glucosilación.
- 15 9. El método de la reivindicación 8 en el que la modificación de la molécula de ácido nucleico comprende uno o más de:
- (i) la delección o inactivación de una secuencia de nucleótidos que codifica el dominio DH como se expone en la Tabla 3;
- 20 (ii) la delección o inactivación de una secuencia de nucleótidos que codifica el dominio KR como se expone en la Tabla 4;
- (iii) la delección o inactivación de *becA* como se muestra en la SEC ID N° 5 o de uno de sus módulos;
- (iv) la delección o inactivación de una o más secuencias de nucleótidos que codifican un módulo de BecB, BecD, BecE, BecF o BecG como se define por las posiciones de nucleótidos indicadas en la Tabla 2;
- 25 (v) la delección o inactivación de *becO* como se muestra en la SEC ID N° 26.
10. El método de una cualquiera de las reivindicaciones 5 a 9, en el que la molécula de ácido nucleico está endógenamente presente en un microorganismo que produce BE-14106 o uno de sus derivados, y el método se realiza en dicho microorganismo, preferentemente en el que dicho microorganismo es *Streptomyces* sp como el depositado en la DSMZ el 25 de enero de 2008 con el número de depósito DSM21069 o uno de sus mutantes o una de sus cepas modificadas que produce BE-14106 o uno de sus derivados.
- 30 11. Un método para preparar un sistema NPRS-PKS modificado de BE-14106, comprendiendo dicho método preparar una molécula de ácido nucleico modificada de acuerdo con el método de una cualquiera de las reivindicaciones 5 a 10 y expresar dicha molécula de ácido nucleico en un microorganismo.
- 35 12. Un método para preparar una molécula de macrolactama; comprendiendo dicho método preparar una molécula de ácido nucleico modificada de acuerdo con el método de una cualquiera de las reivindicaciones 5 a 10 y expresar dicha molécula de ácido nucleico en un microorganismo, que además comprende preferentemente preparar la molécula de macrolactama.
- 40 13. Una cepa de *Streptomyces* depositada en la DSMZ con el número de depósito DSM21069 o uno de sus mutantes o una de sus cepas modificadas que produce BE-14106 o uno de sus derivados.
- 45 14. Un análogo de BE-14106 que comprende una modificación seleccionada de una cualquiera o más del grupo que comprende, 3-, 5-, 7-, 11-, 13-, 15-, 17- o 23-hidroxi BE-14106, 3-, 5-, 7-, 9-, 11-, 13-, 15-, 17-, o 23-oxo BE-14106 o una de sus combinaciones o un análogo que comprende una combinación de un grupo 8-desoxi con una o más modificaciones seleccionadas de la introducción de un grupo hidroxilo u oxo en cualquiera de las posiciones 3, 5, 7, 11, 13, 15, 17 o 23 o un grupo oxo en la posición 9.
- 50 15. Una célula hospedadora que contiene una molécula de ácido nucleico como se define en la reivindicación 1 o 2, en la que dicha molécula se ha introducido en una célula hospedadora heteróloga.

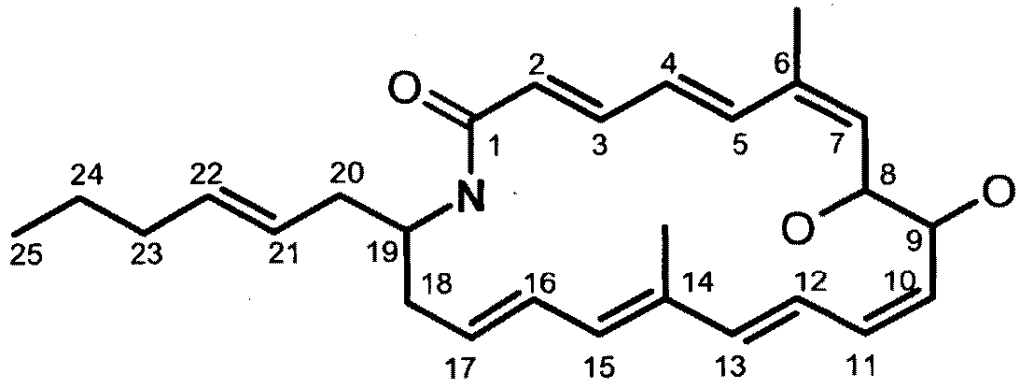


Figura 1

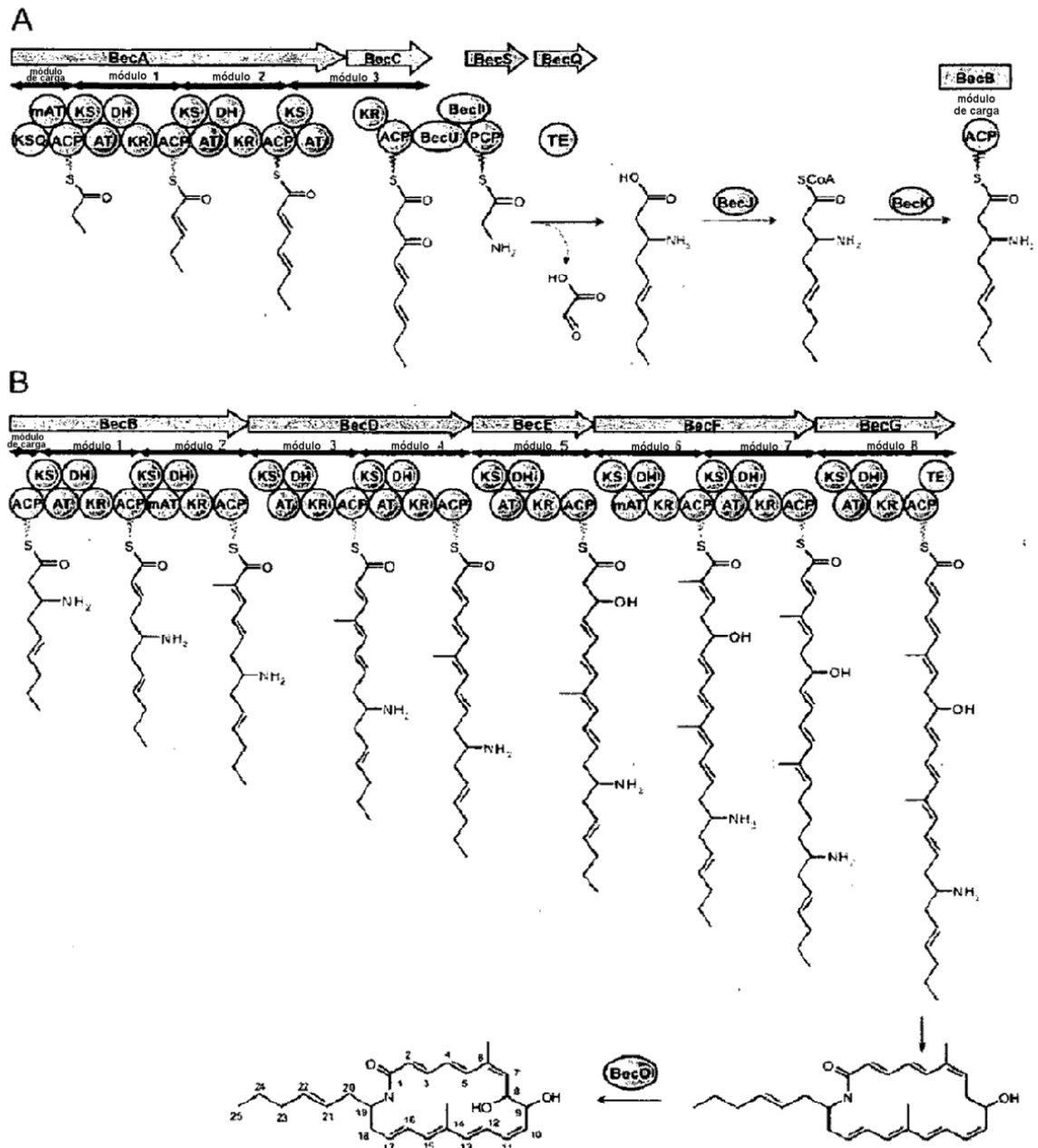


Figura 2