

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 397 885**

21 Número de solicitud: 201130010

51 Int. Cl.:

**C12N 15/00** (2006.01)

**A61K 31/44** (2006.01)

**A61P 31/00** (2006.01)

**A61P 35/00** (2006.01)

12

## PATENTE DE INVENCION

B1

22 Fecha de presentación:

**05.01.2011**

43 Fecha de publicación de la solicitud:

**12.03.2013**

Fecha de la concesión:

**29.01.2014**

45 Fecha de publicación de la concesión:

**05.02.2014**

73 Titular/es:

**UNIVERSIDAD DE OVIEDO (50.0%)  
Plaza del Riego 4, Edificio Histórico  
33003 Oviedo (Asturias) ES y  
ENTRECHEM, S.L. (50.0%)**

72 Inventor/es:

**GARCÍA LLORENTE, Ignacio;  
MIGUEL VIOR, Natalia;  
SIALER GUERRERO, Carlos Alberto;  
GONZÁLEZ SABIN, Javier;  
FERNÁNDEZ BRAÑA, Alfredo;  
MÉNDEZ FERNÁNDEZ, Carmen;  
SALAS FERNÁNDEZ, José Antonio y  
MORÍS VARAS, Francisco**

74 Agente/Representante:

**ILLESCAS TABOADA, Manuel**

54 Título: **DERIVADOS DE COLISMICINA.**

57 Resumen:

Derivados de colismicina. La presente invención describe el aislamiento, clonación y secuenciación de la agrupación de genes implicados en la biosíntesis de colismicina por *Streptomyces* spp. CS40, y el uso de dichos genes para incrementar la producción de colismicina y/o compuestos análogos o derivados relacionados por medio de cepas productoras. Dichos compuestos producidos en la presente invención son aplicables en el tratamiento de diversas enfermedades como, por ejemplo, el cáncer, enfermedades neurodegenerativas y enfermedades infecciosas.

ES 2 397 885 B1

## DESCRIPCIÓN

Derivados de colismicina.

5 **CAMPO DE LA INVENCION**

La presente invención está comprendida dentro del campo de la biología, la farmacia y la medicina.

10 **ESTADO DE LA TÉCNICA**

10

Dentro de la naturaleza las moléculas que presentan en su estructura un grupo 2,2'-bipiridil son compuestos con un gran número de actividades biológicas descritas (Cristalli *et al.*, 1986); (Gomi *et al.*, 1994); (Tsuge *et al.* 1999). Una de estas moléculas es la colismicina producida por *Streptomyces* spp. CS40 (FIG. 1A), que presenta actividad antibiótica frente a bacterias Gram-positivas y Gram-negativas, antifúngica frente a un amplio espectro de hongos y citotóxica frente a células de leucemia P388 (Gomi *et al.* 1994); (Tsuge *et al.*, 1999). Adicionalmente la colismicina está descrita como un inhibidor de la unión entre la dexametasona y los receptores de glucocorticoides (Shindo *et al.*, 1994) y en la patente WO/2007/017146 se describe la capacidad de la colismicina para inhibir el estrés oxidativo en células.

15

20

El desarrollo de la tecnología de ADN recombinante se está convirtiendo en una poderosa herramienta a la hora de incrementar nuestro conocimiento sobre los genes que participan en la biosíntesis de compuestos bioactivos. Esta tecnología puede hoy en día ser aplicada a la mejora en los niveles de producción de distintos compuestos bioactivos y a la obtención de nuevas moléculas derivadas con mejores propiedades clínicas a través de la combinación de genes de distintas rutas de biosíntesis de moléculas bioactivas y su expresión en microorganismos productores de estos compuestos, en lo que se ha denominado biosíntesis combinatoria.

25

La tecnología de ADN recombinante ha hecho posible el aislamiento de agrupaciones de genes completas para la biosíntesis de distintos compuestos bioactivos utilizando, entre otras estrategias, la clonación, la selección o el análisis de genotecas de los microorganismos productores de moléculas con interés farmacológico mediante sondas de ADN. Esta estrategia se basa en la existencia de información genética previa sobre la ruta de biosíntesis o rutas relacionadas biosintéticamente, lo que permite utilizar o diseñar sondas genéticas a partir de la secuencia total o parcial de un enzima de biosíntesis.

30

35

En la literatura científica no existe apenas información acerca de la maquinaria celular encargada de la biosíntesis de compuestos del tipo 2,2'-bipiridil. Tan sólo se han realizado estudios sobre la biosíntesis de una molécula de la familia 2,2'-bipiridil, la caerulomicina. La caerulomicina es una molécula con una estructura similar a la de la colismicina (FIG. 1B) producida por *Streptomyces caeruleus* (Funk y Divekar, 1959).

40

A través de experimentos de incorporación de precursores marcados radiactivamente se identificaron algunas de las moléculas intermediarias durante la biosíntesis de caerulomicina, una de estas moléculas es el ácido picolínico (Vining *et al.*, 1988). Este compuesto ha sido descrito, además, como un metabolito intermediario en la biosíntesis de otras moléculas bioactivas por parte de algunas especies de *Streptomyces* como es el caso de la nikomicina D (Bruntner y Bormann, 1998). En todos los casos descritos en la literatura el precursor para la biosíntesis del ácido picolínico ó 3-hidroxipicolínico (en el caso de la biosíntesis de virginiamicina) es el aminoácido lisina y las enzimas encargadas de la generación de este compuesto han sido identificadas, siendo la primera de ellas un enzima con actividad lisina 2-aminotransferasa (Bruntner y Bormann,. 1998; Namwat *et al.*, 2002).

45

50

La presente invención describe la clonación y secuenciación de la agrupación de genes de biosíntesis de colismicina en el microorganismo productor *Streptomyces* sp. CS40. La información genética disponible sobre enzimas del tipo lisina-2-aminotransferasa, ha sido usada para diseñar oligonucleótidos y construir una sonda genética que ha permitido el aislamiento y clonación de la agrupación de genes de biosíntesis de colismicina.

Desde el punto de vista genético, no hay descripciones previas en la literatura relativas al aislamiento y secuenciación de la agrupación de genes de biosíntesis de colismicina u otra molécula de la familia 2,2'-bipiridil producida por especies de *Streptomyces*.

55

La invención proporciona una importante herramienta para la manipulación genética de esta agrupación de genes en el sentido de aumentar la producción de colismicina y/o obtener nuevos derivados de esta molécula o moléculas estructuralmente relacionadas con propiedades mejoradas.

60 **DESCRIPCIÓN DE LA INVENCION**

60

La presente invención se relaciona con el aislamiento e identificación de una agrupación de genes que participan en la biosíntesis de colismicina por *Streptomyces* spp. CS40, y proporciona una herramienta para la manipulación genética de esta agrupación génica para aumentar la producción de colismicina y obtener nuevos derivados con propiedades mejoradas para su aislamiento y utilización, entre otros, en el sector farmacéutico o químico.

65

Esta invención describe una secuencia de ADN que contiene 24 genes implicados en la biosíntesis de colismicina y sus precursores, incluyendo aquellos implicados en la regulación del agrupamiento génico. El agrupamiento génico incluye genes que codifican para un sistema híbrido policétido sintasa-péptido sintetasa no ribosomal (PCS-NRPS) encargado de sintetizar la estructura central de la colismicina, genes que codifican enzimas implicados en la biosíntesis del precursor ácido picolínico de la colismicina, genes que codifican enzimas implicados en modificaciones de la molécula tales como deshidrogenasas, aminotransferasas y metiltransferasas. La invención por tanto se refiere a nuevos genes y moléculas de ácidos nucleicos que codifican proteínas/polipéptidos que muestran actividades funcionales implicadas en la biosíntesis de colismicina, y su potencial aplicación en el incremento de los niveles de producción de colismicina en *Streptomyces spp.* CS40 y en la producción de nuevos derivados de colismicina.

Los procedimientos experimentales aplicados a la presente invención incluyen métodos de biología molecular convencionales. Una descripción detallada de los métodos utilizados y no detallados aquí se puede obtener de Hopwood *et al.* (1985); Sambrook *et al.* (1989) y Kieser *et al.*, (2000).

Con el fin de clonar la agrupación de genes de biosíntesis de colismicina, se construyó una genoteca de ADN cromosómico de *Streptomyces spp.* CS40 en *Escherichia coli*, utilizando el cósmido pWE15 (ver ejemplo 2).

Para el aislamiento de la agrupación de genes de biosíntesis de colismicina se utilizó una sonda genética consistente en un fragmento de PCR de 412 bp procedente de la amplificación parcial del gen que codifica para una lisina 2-aminotransferasa implicada en la biosíntesis de colismicina. Para su amplificación se utilizó como molde el ADN cromosómico de la cepa CS40 y dos oligonucleótidos: L2ATson1 (SEQ ID NO: 31) y L2ATson2 (SEQ. ID NO: 32) diseñados en base a la secuencia nucleotídica de una lisina 2-aminotransferasa cuya secuencia, de aproximadamente 500 bp, se identificó previamente usando los oligonucleótidos degenerados L2ATFW2 (SEQ. ID NO: 29) y L2ATRV2 (SEQ ID NO: 30).

La utilización de esta sonda en la hibridación de la genoteca de ADN cromosómico de *Streptomyces spp.* CS40 permitió aislar un cósmido (cos1C3) que define una región de aproximadamente 41 kb en el cromosoma. La participación del ADN clonado en el cósmido cos1C3 en la biosíntesis de colismicina se determinó utilizando un fragmento *Bam*HI de 3128 bp, que contenía un fragmento interno a un gen que codifica para una NRPS. Este fragmento fue clonado en el plásmido pOJ260 (Bierman *et al.*, 1992). La construcción resultante se usó para la disrupción génica en *Streptomyces spp.* CS40 generando un mutante no productor de colismicina (ver ejemplo 4) como prueba de la implicación del ADN clonado en la biosíntesis de colismicina.

No obstante, el análisis de la secuencia del cósmido cos1c3 mostró la necesidad de extender la región hacia la izquierda para buscar otros genes no presentes en el cósmido cos1C3, posiblemente implicados en la biosíntesis de colismicina. Un fragmento amplificado por PCR usando los oligonucleótidos 1c3-5FW (SEQ ID NO: 33) y 1c3-5RV (SEQ ID NO: 42) de 1,9 kb del extremo del cósmido cos1C3 se usó como sonda para analizar de nuevo la genoteca de *Streptomyces spp.* CS40 aislándose un nuevo cósmido solapante con el cósmido cos1C3, cos3B11, usado para completar la secuencia de 46672 bp de la región que contiene el agrupamiento génico implicado en la biosíntesis de colismicina (FIG. 2).

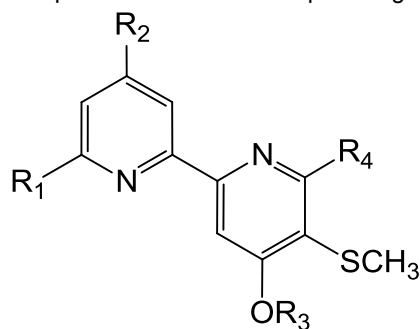
La asignación de funciones en la biosíntesis a los distintos productos génicos se realizó mediante comparación de sus secuencias aminoacídicas con las de otras proteínas presentes en bases de datos (ver ejemplo 3). Algunos genes codifican enzimas biosintéticos estructurales (incluyendo PCSs, NRPSs, etc.), activadores transcripcionales, proteínas implicadas en la exportación al exterior de la célula, y proteínas implicadas en el aporte de precursores para la biosíntesis de colismicina. La participación de esta agrupación génica en la biosíntesis de colismicina se demostró mediante inactivaciones de diferentes genes tanto por disrupción como por reemplazamiento génico (ver ejemplos 4, 6 y 8), generándose en algunos de ellos nuevos derivados de colismicina (ver ejemplos 7 y 8). La actividad antitumoral de los nuevos compuestos generados por manipulación de la ruta de biosíntesis de colismicina se ha valorado mediante ensayos de citotoxicidad in vitro frente a distintas líneas celulares tumorales (ver ejemplo 9). La actividad neuroprotectora de los nuevos compuestos se ha valorado en ensayos utilizando el modelo del pez cebra (ver ejemplo 10). No todos los genes identificados en la región de ADN presentada en esta invención forman parte del agrupamiento génico implicado en la biosíntesis de colismicina. El establecimiento de los límites del agrupamiento génico se realizó mediante disrupción de los genes *orf1* y *orf27* y reemplazamiento del gen *orf25* (ver ejemplo 5).

La presente invención describe procedimientos para la manipulación de los genes de biosíntesis, encaminados a obtener nuevos derivados de la colismicina por técnicas de disrupción y/o reemplazamiento génico generando mutantes en la biosíntesis de colismicina que acumulen distintos intermediarios. Las nuevas cepas generadas, productoras de análogos de colismicina por manipulación genética son las denominadas *Streptomyces spp.* CLM-A, *Streptomyces spp.* CLM-L, *Streptomyces spp.* CLM-M2 y han sido depositadas en la Colección Española de Cultivos Tipo (CECT), con números de depósito 7755, 7754 y 7756 respectivamente.

La presente invención proporciona asimismo compuestos análogos de colismicina A y colismicina C obtenidos mediante acilación enzimática catalizada por una lipasa del grupo oxima o hidroxilo, respectivamente.

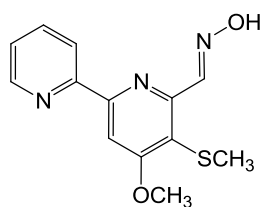
En el sentido de la presente invención se entiende por acilación enzimática la transformación de un sustrato en un derivado acilado a partir de su reacción con un agente acilante catalizada por lipasa. Lipasas útiles para la acilación pueden encontrarse en *Tetrahedron* 2004, 60, 501-519; *Chem. Soc. Rev.* 2004, 33, 201-209; o *Adv. Synth. Catal.* 2006, 348, 797-812. Más concretamente, en la presente invención se utiliza la lipasa de *Burkholderia cepacia* (PS-C), para obtener derivados de colismicina A y colismicina C acilados. Estas lipasas se presentan en diferentes formas de inmovilización sobre soportes hidrófobos y mecánicamente resistentes o sobre resinas acrílicas, como puede ser una resina epoxiacrílica activada con grupos deca-octilo. Agentes acilantes útiles para la presente invención son aquellos que pueden actuar como sustratos de la lipasa utilizada dando lugar a la acilación de la colismicina A y colismicina C, y pueden ser ésteres, carbonatos y anhídridos. La reacción puede llevarse a cabo en una gran variedad de disolventes orgánicos. Más concretamente, en la presente invención se utiliza el éter metil tertbutilico (MTBE). En general la temperatura debe mantener la estructura de la enzima intacta sin que se produzcan fenómenos de desnaturalización. La reacción puede llevarse a cabo entre 5 y 60°C, preferiblemente entre 10 y 60°C, más preferiblemente entre 20 y 50°C.

Asimismo, la presente invención proporciona compuestos caracterizados por la siguiente fórmula (I):

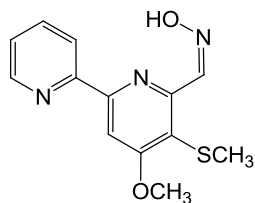


(I)

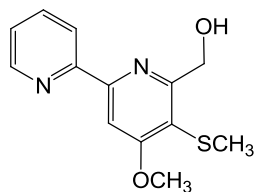
donde R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> y R<sub>4</sub> son, cada uno e independientemente, hidrógeno o un grupo protector. El grupo protector puede consistir en un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo alquilo halogenado, un grupo alcoxialquílico, un grupo alqueno, un grupo alquínico, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos; exceptuando los compuestos con las siguientes fórmulas:



, también conocido como Colismicina A y como (E)-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído oxima,



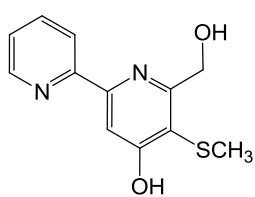
, también conocido como Colismicina B, y como (Z)-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído oxima, y



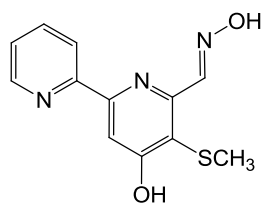
, también conocido como Colismicina C y como 5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-metanol.

En particular, la presente invención proporciona, entre otros, los compuestos de Fórmula II a XVI comprendidos en la Fórmula I:

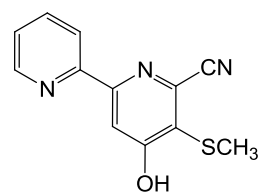
- Fórmula II: 6-(hidroximetil)-5-(metiltio)-[2,2'-bipiridin-4-ol].
- Fórmula III: (*E*)-4-hidroxi-5-(metiltio)-[2,2'-bipiridin-6-il]-carbaldehído oxima.
- Fórmula IV: 4-hidroxi-5-(metiltio)-[2,2'-bipiridinil-6-il]-carbonitrilo.
- Fórmula V: ácido 4-hidroxi-5-(metiltio)-[2,2'-bipiridin-6-il]-carboxílico.
- Fórmula VI: (*E*)-6'-metil-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído oxima.
- Fórmula VII: (*E*)-4'-metil-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído oxima.
- Fórmula VIII: 4'-metil-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-metanol.
- Fórmula IX: (*E*)-4-hidroxi-4'-metil-5-(metiltio)-[2,2'-bipiridin-6-il]-carbaldehído oxima.
- Fórmula X: 4-hidroxi-4'-metil-5-(metiltio)-[2,2'-bipiridin-6-il]-carbonitrilo.
- Fórmula XI: acetato de 5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-metilo.
- Fórmula XII: butanoato de 5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-metilo.
- Fórmula XIII: benzoato de 5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-metilo.
- Fórmula XIV: (*E*)-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído *O*-acetil oxima.
- Fórmula XV: (*E*)-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído *O*-cloroacetil oxima.
- Fórmula XVI: (*E*)-5-(metiltio)-4-metoxi-[2,2'-bipiridin-6-il]-carbaldehído *O*-butanoil oxima.



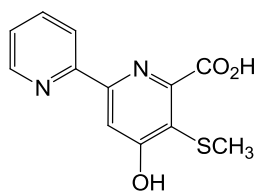
(II)



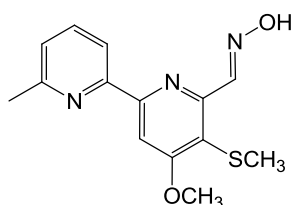
(III)



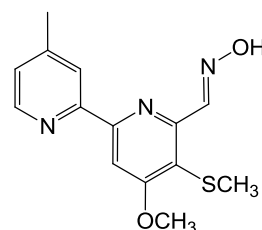
(IV)



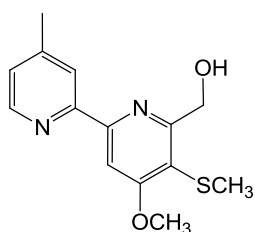
(V)



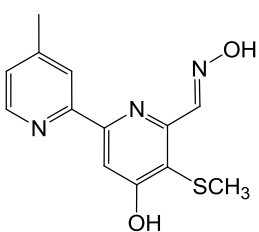
(VI)



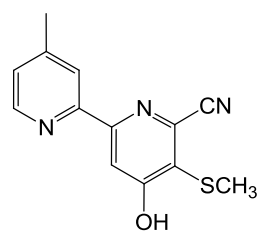
(VII)



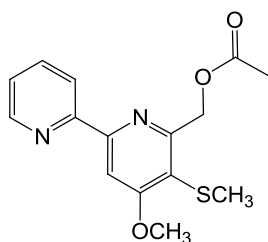
(VIII)



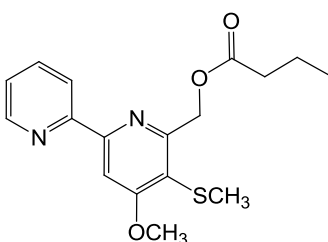
(IX)



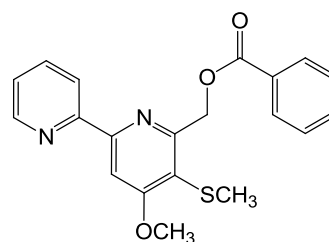
(X)



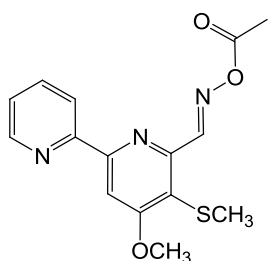
(XI)



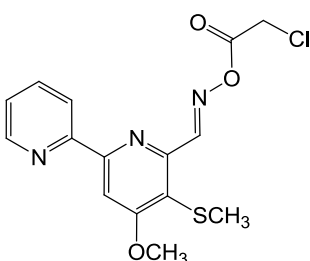
(XII)



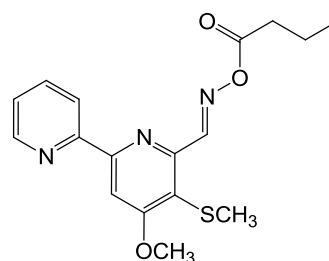
(XIII)



(XIV)



(XV)



(XVI)

Estos compuestos son derivados o análogos de colismicina A con un grupo metilo en posiciones 4' o 6' (compuestos de fórmula VI y VII, respectivamente), colismicina A desmetilada en 4 (compuesto de fórmula III), colismicina A desmetilada en 4 y con un grupo metilo en 4' (compuesto de fórmula IX), colismicina C desmetilada en 4 (compuesto de fórmula II), colismicina C metilada en 4' (compuesto de fórmula VIII), o bien compuestos análogos con la unidad 2,2'-bipiridil característica y un grupo funcional nitrilo o ácido carboxílico en posición 6 (compuestos de fórmula IV, V y X). Asimismo, los compuestos de fórmula XI, XII y XIII son ésteres derivados de colismicina C y los compuestos de fórmula XIV, XV y XVI son ésteres de oxima derivados de colismicina A.

La presente invención describe asimismo el uso de los nuevos compuestos como antibióticos, antitumorales o neuroprotectores.

Los compuestos de la invención son inhibidores del crecimiento de tumores y son por tanto útiles en el tratamiento del cáncer. Así mismo los compuestos de la invención son inhibidores de la lisis o la apoptosis neuronal inducida por estrés oxidativo y por tanto útiles en el tratamiento de aquellas enfermedades neurodegenerativas causadas por agentes oxidantes, tanto exógenos como endógenos, al paciente.

5 De esta forma, son objeto de la presente invención las composiciones farmacéuticas que comprenden una cantidad terapéuticamente efectiva de un compuesto de Fórmula (I) o una sal o solvato farmacéuticamente aceptable del mismo en la manufacturación de un medicamento.

10 Es también objeto de la presente invención el uso de un compuesto de Fórmula (I) o una sal o solvato farmacéuticamente aceptable del mismo para inhibir el crecimiento de un tumor o la lisis o apoptosis neuronal causada por estrés oxidativo.

15 Tal como es usado aquí, "inhibir" significa disminuir, hacer más lento, o detener. Por tanto, un compuesto de esta invención puede disminuir, hacer más lento, o detener el crecimiento de una célula tumoral o bien disminuir hacer más lenta, o detener la lisis o la apoptosis neuronal en condiciones de estrés oxidativo.

20 Tal como es usado aquí, "crecimiento" significa aumento en tamaño, o proliferación, o ambos. Por tanto, un compuesto de esta invención puede inhibir el aumento de tamaño de una célula tumoral y/o puede impedir que la célula tumoral se divida y aumente el número de células tumorales. Una "célula tumoral" es una célula que constituye un neoplasma (crecimiento nuevo), el cual puede ser canceroso (maligno) o no canceroso (benigno). Una célula tumoral cancerosa puede invadir los tejidos normales a su alrededor y los vasos sanguíneos/linfáticos y formar metástasis en tejidos alejados del tumor original. Por el contrario, una célula tumoral no cancerosa puede crecer y comprimir los tejidos normales adyacentes pero no puede invadir tejidos normales y vasos sanguíneos/linfáticos, y tampoco puede formar metástasis en tejidos alejados del tumor original.

25 Tal como es usado aquí, enfermedad neurodegenerativa significa aquel tipo de enfermedad cuyos síntomas son el resultado de la muerte por lisis o apoptosis de neuronas. El estrés oxidativo en el contexto de enfermedad neurodegenerativa se refiere a aquella situación anómala en la que el citoplasma de las neuronas presenta un aumento de la cantidad de especies reactivas de oxígeno ( $H_2O_2$ ,  $HO_2$ ,  $O^{2-}$  y  $OH^{\cdot}$ ). Estas especies reactivas de oxígeno pueden ser generadas por el propio metabolismo de la neurona enferma, por otros órganos del individuo o ser exógenas al individuo.

30 Tal como es usado aquí, un compuesto neuroprotector frente al estrés oxidativo es aquel compuesto capaz de detener, minimizar o ralentizar la lisis o la apoptosis neuronal causada por agentes oxidantes.

35 Es también objeto de la presente invención el uso de un compuesto de Fórmula (I) o una sal o solvato farmacéuticamente aceptable del mismo para tratar el cáncer o enfermedades neurodegenerativas causadas o agravadas por la lisis o apoptosis neuronal inducida por estrés oxidativo.

40 Es también objeto de la presente invención el uso de un compuesto de Fórmula (I) o una sal o solvato farmacéuticamente aceptable del mismo en la manufacturación de un medicamento con actividad antitumoral o neuroprotectora frente a estrés oxidativo.

45 Es también objeto de la presente invención el uso de un compuesto de Fórmula (I) o una sal o solvato farmacéuticamente aceptable del mismo en la manufacturación de un medicamento para el tratamiento del cáncer o de ciertas enfermedades neurodegenerativas.

50 Es también objeto de la presente invención un método de tratamiento de un sujeto, incluyendo un ser humano, diagnosticado con cáncer o con una enfermedad neurodegenerativa, que consiste en tratar a dicho sujeto con una cantidad terapéuticamente efectiva de un compuesto de fórmula (I) o una sal o solvato farmacéuticamente aceptable del mismo.

55 Tal como es usado aquí, un "sujeto" puede incluir animales domesticados (por ejemplo, gatos, perros, etc.), ganado (por ejemplo, vacas, caballos, cerdos, ovejas, cabras, etc.), animales de laboratorio (por ejemplo, ratones, conejos, cobayas, etc.) y pájaros. De manera preferente, el sujeto es un mamífero tal como un primate y, con mayor preferencia, un ser humano.

60 En general, una "cantidad terapéuticamente eficaz" de un compuesto es aquella cantidad necesaria para conseguir el resultado deseado. Por ejemplo, la cantidad eficaz de un compuesto de la presente invención trata el cáncer mediante la inhibición del crecimiento de las células que constituyen el tumor, con lo que previene la invasión de tejidos normales y vasos sanguíneos/linfáticos por parte de las células tumorales y, por tanto, previene metástasis, o bien cantidad efectiva de un compuesto de la presente invención es aquella capaz de ralentizar o inhibir la apoptosis neuronal en condiciones de estrés oxidativo. La expresión "composición farmacéutica aceptable" se refiere a un material adecuado biológicamente, es decir, que el material puede ser administrado al sujeto sin causarle efectos biológicos sustancialmente dañinos.

65

Las dosis o cantidades de los compuestos de la invención deben ser suficientemente grandes para producir el efecto deseado. Sin embargo, la dosis no debe ser tan alta que cause efectos secundarios adversos, por ejemplo reacciones cruzadas indeseadas, reacciones anafilácticas, y similares. Generalmente, la dosis variará con la edad, condición, sexo y el grado de la enfermedad del sujeto, y puede ser determinada por cualquier experto en la materia. La dosis puede ser ajustada por cada médico, en base a la condición clínica del sujeto implicado. La dosis, régimen de dosificación y ruta de la administración pueden variarse.

Los compuestos de la invención pueden ser útiles para la investigación en bioquímica o biología celular.

Cualquiera de los compuestos de la invención puede ser utilizado terapéuticamente formando parte de una composición farmacéutica aceptable. Las composiciones farmacéuticas aceptables pueden consistir en soluciones estériles en agua, soluciones salinas, o soluciones tamponadas a pH fisiológico. Cualquiera de los compuestos de la invención puede ser preparado en forma de composición farmacéutica. Las composiciones farmacéuticas pueden incluir diversos agentes transportadores, espesantes, diluentes, tamponantes, conservantes, tensoactivos, y otros, además del compuesto de la invención. Las composiciones farmacéuticas pueden incluir también ingredientes activos tales como agentes antimicrobianos, antiinflamatorios, anestésicos, etc.

Los compuestos de la invención pueden ser administrados al sujeto de varias maneras distintas, dependiendo de si se desea que el tratamiento sea local o sistémico, y dependiendo del área a ser tratada. Así por ejemplo, un compuesto de la presente invención puede ser administrado en forma de solución oftálmica, de aplicación en la superficie del ojo. Además un compuesto puede ser administrado a un sujeto por vía vaginal, rectal, intranasal, oral, por inhalación, o por vía parenteral, ya sea por ruta intradérmica, subcutánea, intramuscular, intraperitoneal, intrarrectal, intraarterial, intralinfática, intravenosa, intratecal e intratraqueal. La administración parenteral, si se emplea, se realiza generalmente mediante inyección. Los inyectables pueden ser preparados de diversas formas, tales como soluciones o suspensiones líquidas, formas sólidas adecuadas para ser disueltas o puestas en suspensión antes de la inyección, o como emulsiones. Otras formas de administración parenteral emplean sistemas de liberación lenta o sostenida, de tal forma que se consigue mantener una dosis constante (ver, por ejemplo, patente US 3,710,795). Las preparaciones para la administración parenteral incluyen soluciones estériles acuosas o no acuosas, suspensiones, y emulsiones, que además pueden contener tampones y aditivos diluentes y otros. Ejemplos de solventes no acuosos son: propilenglicol, polietilenglicol, aceites vegetales tales como el aceite de oliva, y ésteres orgánicos inyectables tales como etiloleato. Ejemplos de solventes acuosos son: agua, soluciones alcohólico-acuosas, emulsiones o suspensiones, incluyendo soluciones salinas y tamponadas. Ejemplos de vehículos parenterales son: solución de cloruro sódico, dextrosa de Ringer, cloruro de sodio y dextrosa, etc. También pueden estar presentes conservantes y otros aditivos, tales como, por ejemplo, agentes antimicrobianos, antioxidantes, quelantes, gases inertes, etc. Las formulaciones para administración tópica pueden incluir cremas, lociones, geles, gotas, supositorios, sprays, líquidos y polvos. También pueden ser necesarios o deseables ciertos transportadores farmacéuticos convencionales, bases acuosas, oleosas, o en polvo, espesantes, etc. Las composiciones para administración oral pueden incluir polvos o gránulos, suspensiones o soluciones en agua o medio no acuoso, cápsulas, o tabletas. Puede ser deseable la inclusión de agentes espesantes, saborizantes, diluentes, emulsionantes, dispersantes, etc.

A los efectos de la presente invención y su descripción, el término "derivado" o "análogo" de colismicina debe interpretarse como un compuesto cubierto por la Fórmula general (I) que no necesariamente debe derivar de la colismicina sino que puede ser sintetizado de novo.

La presente invención se describe en detalle a continuación con una serie de ejemplos sin que en ningún caso sean ejemplos que limiten su utilización a los que se mencionan.

## BREVE DESCRIPCIÓN DE LAS FIGURAS

### Figura 1.

A. Estructura de la colismicina A producida por *Streptomyces* spp CS40.  
B. Estructura de la caerulomicina producida por *Streptomyces caeruleus*.

**Figura 2.** Diagrama en el que se muestra una representación esquemática del mapa de restricción, utilizando el enzima de restricción BamHI (posiciones numeradas en el esquema), de la agrupación génica para la biosíntesis de colismicina en *Streptomyces* spp. CS40 contenida en la secuencia de nucleótidos descrita como SEQ IN NO: 1. La escala se muestra en kilobases (kb). *cos1C3* y *cos3B11* representan los cósmidos en los cuales se ha aislado la secuencia de nucleótidos descrita como SEQ IN NO: 1. Los genes presentes en la agrupación génica se representan por números bajo las flechas. Los genes que se han inactivado dentro del agrupamiento génico se muestran como flechas grises y el resto como flechas negras. Los genes no implicados en la biosíntesis de colismicina se muestran como flechas blancas.

**Figura 3.** Análisis por cromatografía líquida (UPLC) y cromatografía líquida acoplada a espectrometría de masas (HPLC/MS). La FIG. 3A muestra la producción de colismicina A (pico a 3.259 min) en la cepa silvestre *Streptomyces* spp. CS40 analizado por UPLC. En las ordenadas se representa el tiempo en minutos (min.) y en las abscisas como unidades arbitrarias (AU). La FIG. 3B muestra el espectro de absorción de colismicina A. En las abscisas se representa la longitud de onda en nanómetros (nm) y en las ordenadas como unidades arbitrarias. La FIG. 3C muestra colismicina



A purificada producida en la cepa silvestre *Streptomyces spp.* CS40 y analizada por HPLC/MS. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 3D muestra el espectro de masas de la colismicina A marcándose en cada pico el valor de masas (sin unidades).

5 **Figura 4.** El análisis por UPLC de la cepa silvestre *Streptomyces spp.* CS40 productora de colismicina A se muestra en la FIG. 4A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). En la FIG. 4B se muestra el análisis por UPLC del mutante CLM-L obtenido por reemplazamiento génico, no productor de colismicina A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). En la FIG. 4C se muestra el análisis por UPLC del mutante CLM-12, obtenido por disrupción génica, no productor de colismicina A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU).

10 **Figura 5.** El análisis por UPLC de la cepa mutante CLM-22D productora de colismicina A se muestra en la FIG. 5A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). En la FIG. 5B se muestra el análisis por UPLC de la cepa mutante CLM-24, productora de colismicina A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU).

15 **Figura 6.** Esquema de los pasos iniciales en la ruta de biosíntesis de colismicina A. El aminoácido lisina, en una reacción catalizada por el producto génico de los genes orf19 y orf20, se transforma en ácido picolínico. Posteriormente el resto de genes implicados en la biosíntesis continuaran modificando este compuesto hasta producir colismicina A.

20 **Figura 7.** La FIG. 7A muestra un esquema de la organización genética del mutante CLM-L (cepa depositada como CECT 7754) que presenta interrumpido el gen orf19 que codifica para una lisina 2-aminotransferasa. En la FIG. 7B. se muestra el análisis por cromatografía líquida (UPLC) del extracto de la cepa mutante no productora de colismicina A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 7C muestra el análisis por cromatografía líquida de muy alta resolución (UPLC) del extracto de la cepa mutante no productora de colismicina A suplementada con ácido picolínico, compuesto capaz de rescatar la producción de colismicina A. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU).

25 **Figura 8.** Análisis de la producción de colismicina C y del compuesto de fórmula II por la cepa mutante CLM-A (depositada como CECT 7755). En la FIG. 8A se muestra el esquema de la organización genética del mutante CLM-A, que presenta interrumpido el gen orf11, que codifica para una aminotransferasa. La FIG. 8B muestra el análisis por UPLC del extracto del mutante CLM-A, el pico a 2.86 min corresponde al compuesto de fórmula II y el pico a 2.987 a la Colismicina C. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 8C muestra el espectro de absorción del compuesto de fórmula II. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG 8D muestra el espectro de absorción de Colismicina C. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 8E muestra el perfil de MS correspondiente al compuesto de fórmula II marcándose en cada pico el valor de masas (sin unidades). La FIG. 8F muestra el perfil de MS correspondiente a la Colismicina C marcándose en cada pico el valor de masas (sin unidades).

30 **Figura 9.** Análisis de la producción de colismicinas modificadas, compuestos de fórmula III y de fórmula IV por la cepa mutante CLM-M2 (depositada como CECT 7756). En la FIG. 9A se muestra el esquema de la organización genético del mutante CLM-M2, que presenta interrumpido el gen orf9, que codifica para una metiltransferasa. La FIG. 9B muestra el análisis por UPLC del extracto del mutante CLM-M2, señalando los compuestos de fórmula III y de fórmula IV. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 9C muestra el espectro de absorción del compuesto de fórmula III. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG 9D muestra el espectro de absorción del compuesto de fórmula IV. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 9E muestra el perfil de MS del compuesto de fórmula III marcándose en cada pico el valor de masas (sin unidades). La FIG. 9F muestra el perfil de MS del compuesto de fórmula IV marcándose en cada pico el valor de masas (sin unidades).

35 **Figura 10.** Análisis de la producción de colismicina modificada, compuesto de fórmula V por la cepa mutante CLM-G. En la FIG. 10A se muestra el esquema de la organización genético del mutante CLM-G, que presenta interrumpidos los genes orf13 y orf14, que codifican para proteínas similares GriD (YP\_001825756) y GriC (YP\_001825755). La FIG. 10B muestra el análisis por HPCL/MS del extracto de la cepa silvestre *Streptomyces spp.* CS40. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 10C muestra el análisis por HPCL/MS del extracto de la cepa mutante CLM-G. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 10D muestra el espectro de absorción del compuesto de fórmula V. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 10E muestra el perfil de MS del compuesto de fórmula V marcándose en cada pico el valor de masas (sin unidades).

65

**Figura 11.** Análisis de la producción de colismicinas modificadas, compuestos de fórmula VI y de fórmula VII, producidos por la cepa mutante CLM-L (depositada como CECT 7754) suplementada con ácido 6-metil picolínico y ácido 4-metilpiridina-2-carboxílico respectivamente. En la FIG. 11A se muestra el análisis por UPLC del extracto del mutante CLM-L suplementado con ácido 6-metil picolínico, señalando el compuesto de fórmula VI. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 11B muestra el espectro de absorción del compuesto de fórmula VI. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 11C muestra el perfil de MS del compuesto de fórmula VI marcándose en cada pico el valor de masas (sin unidades). En la FIG. 11D se muestra el análisis por UPLC del extracto del mutante CLM-L suplementado con ácido 4-metilpiridina-2-carboxílico, señalando el compuesto de fórmula VII. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 11E muestra el espectro de absorción del compuesto VII. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 11F muestra el perfil de MS del compuesto de fórmula VII marcándose en cada pico el valor de masas (sin unidades).

**Figura 12.** Análisis de la producción de colismicina modificada, compuesto de fórmula VIII, producido por la cepa mutante CLM-A (depositada como CECT 7755) suplementada con ácido 4-metilpiridina-2-carboxílico. En la FIG. 12A se muestra el análisis por UPLC del extracto del mutante CLM-A suplementado con ácido 4-metilpiridina-2-carboxílico, señalando el compuesto de fórmula VIII. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 12B muestra el espectro de absorción del compuesto de fórmula VIII. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 12C muestra el perfil de MS del compuesto de fórmula VIII marcándose en cada pico el valor de masas (sin unidades).

**Figura 13.** Análisis de la producción de colismicinas modificadas, compuestos de fórmula IX y X, producidos por la cepa mutante CLM-M2 (depositada como CECT 7756) suplementada con ácido 4-metilpiridina-2-carboxílico. En la FIG. 13A se muestra el análisis por UPLC del extracto del mutante CLM-M2 suplementado con ácido 4-metilpiridina-2-carboxílico señalando los compuestos de fórmula IX y de fórmula X. En las abscisas se representa el tiempo en minutos (min.) y en las ordenadas como unidades arbitrarias (AU). La FIG. 13B muestra el espectro de absorción del compuesto de fórmula IX. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 13C muestra el espectro de absorción del compuesto de fórmula X. En las abscisas se representa la longitud de onda en nanómetro (nm) y en las ordenadas como unidades arbitrarias. La FIG. 13D muestra el perfil de MS del compuesto de fórmula IX marcándose en cada pico el valor de masas (sin unidades). La FIG. 13E muestra el perfil de MS del compuesto de fórmula X marcándose en cada pico el valor de masas (sin unidades).

**Figura 14.** Capacidad neuroprotectora de la colismicina A y de los compuestos de fórmula VI y VII. En todos los casos de la FIG 14 la presencia de células apoptóticas se muestra por los puntos teñidos intensamente en el cerebro de embriones de pez cebra. En la FIG 14A se muestra el cerebro de embriones de pez cebra sin tratar. En la FIG 14B se muestra el cerebro de embriones de pez cebra tratados con ácido retinoico 10  $\mu$ M. En la FIG 14C se muestra el cerebro de embriones de pez cebra tratados con ácido retinoico 10  $\mu$ M y con colismicina A 1  $\mu$ M. En la FIG 14D se muestra el cerebro de embriones de pez cebra tratados con ácido retinoico 10  $\mu$ M y el compuesto de fórmula VI 1  $\mu$ M. En la FIG 14E se muestra el cerebro de embriones de pez cebra tratados con ácido retinoico 10  $\mu$ M y el compuesto de fórmula VII 1  $\mu$ M. En la FIG 14F se muestra el cerebro de embriones de pez cebra tratados con ácido retinoico 10  $\mu$ M y ácido lipoico 1  $\mu$ M. En la FIG 14G se muestra una representación gráfica de la capacidad neuroprotectora de la colismicina A y los compuestos de fórmula II, XI, VI y VII respectivamente. El eje de ordenadas muestra el porcentaje de apoptosis en el cerebro de los embriones de pez cebra. Se consideró el 0% de apoptosis como el nivel de apoptosis basal de los cerebros de embriones no tratados (EW) y el 100% de apoptosis como el nivel de apoptosis presente en los cerebros de embriones tratados con ácido retinoico 10  $\mu$ M (AR). Los ensayos de neuroprotección se realizaron tratando los embriones de pez cebra con ácido retinoico en presencia de los distintos compuestos a ensayar. Como control de neuroprotección se usó ácido lipoico (AR+LP). Los compuestos ensayados como neuroprotectores fueron la colismicina A (AR+Col.A), el compuesto de fórmula II (AR + II), la colismicina C (AR+Col.C), el compuesto de fórmula VI (AR + VI) y el compuesto de fórmula VII (AR + VII). Como puede comprobarse en la gráfica, el compuesto de fórmula VII presenta una actividad neuroprotectora mejorada respecto a la colismicina A y a la colismicina C.

Por lo tanto, un primer aspecto de la presente invención hace referencia a un procedimiento de aislamiento y purificación de un fragmento de ADN, que contiene la agrupación de genes de la ruta de biosíntesis de colismicina de la bacteria *Streptomyces spp.* CS40, que se incluye en un fragmento de 46672 bp del genoma de *Streptomyces spp.* CS40. El proceso comprende las siguientes etapas: (a) obtención de una genoteca de ADN genómico de un microorganismo productor de colismicina; (b) transfección de clones de dicha genoteca en células hospedadoras; (c) diseño de oligonucleótidos para el aislamiento de la agrupación de genes de biosíntesis de colismicina; (d) construcción de una sonda que comprende una secuencia nucleotídica de una agrupación de genes de biosíntesis de colismicina; (e) utilización de sondas heterólogas para el aislamiento de la agrupación de genes de biosíntesis de colismicina; (f) hibridación de dichas sonda con una genoteca de ADN genómico obtenida de dicho microorganismo y (g) aislamiento de dicha agrupación génica a partir de los clones con hibridación positiva.

El segundo aspecto de la presente invención hace referencia a una molécula de ácido nucleico que consiste en una secuencia de nucleótidos descrita como SEQ ID NO: 1; o una secuencia de nucleótidos complementaria a SEQ ID NO: 1; o una secuencia de nucleótidos degenerada respecto a SEQ ID NO: 1; o una secuencia de nucleótidos capaz de

5 hibridar bajo condiciones restrictivas con SEQ ID No. 1, con la hebra complementaria de SEQ ID NO: 1, o con una sonda de hibridación derivada de SEQ ID NO: 1 o de su hebra complementaria; o una secuencia de nucleótidos que posee al menos un 80 % de identidad de secuencia con SED ID NO:1; o una secuencia de nucleótidos que posee al menos 65 % de identidad de secuencia con SEQ ID NO:1 y que preferiblemente codifica o es complementaria a una  
 10 secuencia que codifica al menos un enzima biosintético de colismicina o una parte de él. En una realización preferida de la invención la molécula de ácido nucleico se caracteriza porque tiene al menos 15 nucleótidos de longitud. En otra realización preferida de la invención la molécula de ácido nucleico codifica uno o más polipéptidos, o que incluye uno o más elementos genéticos, que poseen una actividad funcional en la síntesis de un antibiótico 2,2'-bipiridílico o un precursor de un 2,2'-bipiridilo. En otra realización preferida de la invención dicho antibiótico 2,2'-bipiridílico es colismicina o un precursor de colismicina. En otra realización preferida de la invención la molécula de ácido nucleico codifica uno o más polipéptidos, o incluye uno o más genes y/o una o más secuencias reguladoras, y/o uno o más elementos genéticos codificadores o no codificadores, que tienen actividad funcional en la síntesis de un 2,2'-bipiridilo o un precursor de un 2,2'-bipiridilo. En otra realización preferida de la invención la molécula de ácido nucleico se caracterizada porque dicho antibiótico 2,2'-bipiridílico o precursor de un 2,2'-bipiridilo es colismicina o un precursor de colismicina. En otra realización preferida de la invención la molécula de ácido nucleico se caracteriza por incluir una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos de las descritas en SEQ ID NOs: 2 a 28, o una secuencia de nucleótidos que es complementaria o degenerada con respecto a una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos que poseen al menos un 60% de identidad de secuencia con cualquiera de SEQ ID NOs: 2 a 28. En otra realización preferida de la invención la molécula de ácido nucleico codifica una o más secuencias aminoacídicas que poseen al menos un 85% de identidad de secuencia con cualquiera de SEQ ID NOs: 2 a 28.

25 El tercer aspecto de la presente invención hace referencia a un polipéptido codificado por la molécula de ácido nucleico arriba descrita. En una realización preferida de la invención el polipéptido incluye: una o más secuencias aminoacídicas completas, o partes de las mismas, descritas en SEQ ID NOs: 2 a 28; o una o más secuencias aminoacídicas completas, o partes de las mismas, que poseen al menos un 60% de identidad de secuencia con cualquiera de SEQ ID NOs: 2 a 28. En otra realización preferida de la invención el polipéptido está caracterizado porque las secuencias aminoacídicas arriba mencionadas poseen al menos un 85% de identidad de secuencia con cualquiera de SEQ ID NOs: 2 a 28. En otra realización preferida de la invención el polipéptido posee una actividad funcional en la síntesis de un antibiótico 2,2'-bipiridílico.

35 El cuarto aspecto de la presente invención hace referencia a una molécula de ADN recombinante, que incluye el fragmento de ADN arriba mencionado, o una parte con similares características, clonada en un vector que se replica en *Streptomyces* o en *E. coli*. En una realización preferida de la invención el ADN recombinante es el cósmido cos1c3 o el cósmido cos3b11.

El quinto aspecto de la presente invención hace referencia a una célula hospedadora o un organismo transgénico no humano que contiene una molécula de ácido nucleico arriba mencionada.

40 Así, los genes codificados por el fragmento de ADN arriba citado, y las células hospedadoras u organismos transgénicos que comprenden dicho fragmento de ADN, pueden usarse en la producción de metabolitos 2,2'-bipiridílicos, particularmente en la producción de colismicina, derivados de colismicina o precursores de colismicina; para incrementar la producción de metabolitos 2,2'-bipiridílicos; para incrementar la producción de colismicina, derivados de colismicina o precursores de colismicina; en la inactivación de genes implicados en la biosíntesis de colismicina; en técnicas de amplificación por PCR encaminadas al aislamiento y/o utilización de genes implicados en la biosíntesis de colismicina.

50 El sexto aspecto de la presente invención hace referencia a un proceso para incrementar la producción de 2,2'-bipiridilos en un hospedador bacteriano, que comprende la transferencia del fragmento de ADN arriba mencionado a un hospedador de *Streptomyces*, cultivo de la cepa recombinante obtenida, y aislamiento del 2,2'-bipiridilo producido. En una realización preferida de la invención el hospedador es del género *Streptomyces* es *Streptomyces spp.* CS40. En otra realización preferida de la invención el hospedador *Streptomyces spp.* CS40 es un mutante derivado de *Streptomyces spp.* CS40. En otra realización preferida de la invención el compuesto 2,2'-bipiridílico es colismicina, un derivado de colismicina o un precursor de colismicina.

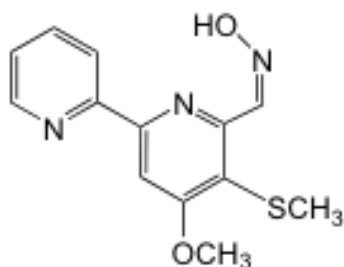
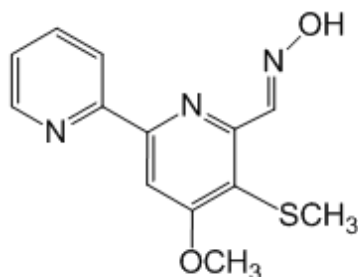
55 El séptimo aspecto de la presente invención hace referencia a un proceso para generar derivados de colismicina o precursores de colismicina por inactivación de genes codificados por el fragmento de ADN arriba mencionado. En una realización preferida de la invención el proceso se focaliza en la utilización de intermediarios de colismicina o derivados de colismicina como compuestos de partida en la síntesis química de productos 2,2'-bipiridílicos.

60 El octavo aspecto de la presente invención hace referencia a un proceso para generar derivados o análogos de colismicina A y colismicina C mediante acilación enzimática catalizada por una lipasa.

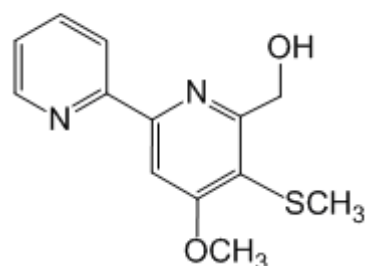
65 El noveno aspecto de la presente invención hace referencia a la cepa mutante *Streptomyces spp.* CLM-A depositada en la Colección Española de Cultivos Tipo con el número de identificación 7755 y a los productos acumulados por la misma.

El décimo aspecto de la presente invención hace referencia a la cepa mutante *Streptomyces spp.* CLM-M depositada en la Colección Española de Cultivos Tipo con el número de identificación 7756 y a los productos acumulados por la misma.

5 El undécimo aspecto de la presente invención hace referencia a un compuesto de Fórmula (I), exceptuando a los compuestos de fórmula:



y



40

45 donde  $R_1$ ,  $R_2$ ,  $R_3$  y  $R_4$  son, cada uno e independientemente, hidrógeno o un grupo protector. El grupo protector puede consistir en un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo alquilo halogenado, un grupo alcoxilquilo, un grupo alqueno, un grupo alquino, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo cetona, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos. En una realización preferida de la presente invención el compuesto de Fórmula (I) se selecciona entre los compuestos de Fórmula II a X.

50 El duodécimo aspecto de la presente invención hace referencia al uso de los compuestos de Fórmula I para la elaboración de una composición farmacéutica destinada al tratamiento del cáncer, al tratamiento enfermedades neurodegenerativas, o su uso como neuroprotector, o al tratamiento enfermedades infecciosas, o a su uso como antibiótico. Además, este aspecto también hace referencia a los compuestos de Fórmula I para ser usados en el tratamiento del cáncer, en el tratamiento enfermedades neurodegenerativas, o como neuroprotector, o en el tratamiento enfermedades infecciosas, o como antibiótico.

55 El décimo tercer aspecto de la presente invención hace referencia a una composición farmacéutica que comprende al menos uno de los compuestos de Fórmula I y al menos un excipiente farmacéuticamente aceptable.

60 El último aspecto de la presente invención hace referencia a un método para el tratamiento del cáncer, para el tratamiento de enfermedades neurodegenerativas o para el tratamiento de enfermedades infecciosas que comprende la administración al paciente de una cantidad terapéuticamente eficaz de un compuesto de Fórmula I o de una composición farmacéutica que comprenda al menos un compuesto de Fórmula I.

65

## EJEMPLOS

**Ejemplo 1. Aislamiento de la cepa productora de colismicina *Streptomyces* spp. CS40**

5 La cepa *Streptomyces* spp. CS40 fue aislada a partir de unas hormigas cortadoras de hojas de especie *Acromyrmex octospinosus* recolectadas en el departamento Lambayeque en Perú. La secuenciación y análisis de su 16SrDNA  
10 mostró su ubicación dentro del género *Streptomyces* sin niveles de identidad concluyentes a nivel de especie. La cepa *Streptomyces* spp. CS40 se encuentra depositada en la Colección Española de Cultivos Tipo con número de depósito 7757.

**Ejemplo 2. Clonación de la agrupación génica implicada en la biosíntesis de colismicina.****2.1. Microorganismos, plásmidos y condiciones de cultivo.**

15 Los microorganismos y plásmidos utilizados se describen en la Tabla 1. *Streptomyces* spp. CS40 y los mutantes generados a partir de él se cultivaron para su esporulación en medio A (Fernández *et al.*, J. Bacteriol., 180, 4929-4937, 1998); para la producción de antibiótico se cultivó en medio líquido R5A (Fernández *et al.*, J. Bacteriol., 180, 4929-4937, 1998); utilizando un inóculo previamente cultivado en medio líquido TSB (Tryptone Soya Broth, Merck). La conjugación intergenérica desde *Escherichia coli* ET12567 (pUB307) a *Streptomyces* spp. CS40 se realizó según Sambrook *et al.*,  
20 Molecular cloning: a laboratory manual. Cold Spring Harbour, NY: Cold Spring Harbour Laboratory press (1989). Las cepas de *E. coli* se cultivaron y transformaron como se describe por Sambrook *et al.*, (1989). Los medios de cultivo fueron suplementados con los antibióticos apropiados a cada marcador de resistencia en las concentraciones siguientes: 100 µg/ml ampicilina, 20 µg/ml tobramicina, 25 µg/ml apramicina, 50 µg/ml tiostreptona, 25 µg/ml kanamicina  
25 10 µg/ml tetraciclina, 25 µg/ml cloramfenicol y 50 µg/ml ácido nalidíxico.

**Tabla 1.** Cepas bacterianas y plásmidos usados en este estudio.

Cepa, plásmido	Propiedades	Fuente o referencia
<i>E. coli</i> DH10B	Huésped general de clonación	Invitrogen
<i>E. coli</i> XLI Blue MR	Huésped para la construcción de la genoteca	Stratagene
<i>E. coli</i> ET12567 (pUB307)	Cepa para la conjugación intergenérica	Kieser <i>et al.</i> , Practical <i>Streptomyces</i> Genetics. The John Innes Foundation. Norwich, 2000
<i>Streptomyces</i> spp. CS40	Productor de colismicina A	Aislado en nuestro laboratorio
pWE15	Cósmido para la construcción de la genoteca	Stratagene
pOJ260	Plásmido para la disrupción génica	Bierman <i>et al.</i> , Gene 116, 43-49, 1992
pEM4T	Plásmido para la expresión de genes en <i>Streptomyces</i>	Menéndez <i>et al.</i> , Appl. Environ. Microbiol. 72, 167-177, 2006
pCR-BLUNT pU09090	Plásmido para la clonación de productos de PCR Plásmido fuente del gen de resistencia a apramicina	Invitrogen Prado <i>et al.</i> , Mol Gen Genet. 201, 216-225, 1999
pBluescript SK+ pHZ1358	Vector de clonación en <i>E. coli</i> Plásmido para el reemplazamiento génico	Stratagene Sun <i>et al.</i> , Appl. Microbiol. Biotechnol 82, 303-310, 2009
pGemT	Plásmido para la clonación de productos de PCR	Promega

**2.2. Análisis de la producción de colismicina A.**

30 La producción de colismicina A se realizó de forma rutinaria en 1,5 mL de medio R5A sólido (Fernández *et al.*, J. Bacteriol., 180, 4929-4937, 1998) en placas de 25 pocillos. Para su inóculo se utilizaron esporas de *Streptomyces* spp. CS40 y los cultivos se mantuvieron durante 7 días a 30°C, extrayéndose tras ese tiempo con 1 ml de acetato de etilo. La producción de colismicina A se realizó también en cultivos líquidos de 5 a 7 días crecidos en un agitador orbital a 30 °C  
35 y 250 rpm. Para ello se utilizaron matraces Erlenmeyer de 250 ml conteniendo 50 ml de medio R5A líquido (Fernández *et al.*, J. Bacteriol, 180: 4929-4937, 1998). Para su inóculo se utilizó un volumen del 2% de un preinóculo de *Streptomyces* spp. CS40 realizado en medio TSB (50 ml en matraces de 250 ml) que se recogió después de dos días de incubación en un agitador orbital a 30°C y 250 rpm. La colismicina A presente en los cultivos líquidos se extrajo con

volúmenes variables de acetato de etilo. Los extractos de acetato de etilo se evaporaron utilizando una centrifuga acoplada a vacío (Speed-Vac) y una vez evaporadas las muestras se resuspendieron en metanol para su análisis.

5 La identificación y análisis cuantitativo de colismicina A se llevó a cabo mediante cromatografía en fase reversa en un equipo Acquity UPLC utilizando una columna BEH C18 (2,1 x 100 mm Waters) y utilizando acetonitrilo y 0,1% TFA en agua como solventes. Las muestras fueron eluidas con acetonitrilo al 10% durante 1 min seguido por un gradiente lineal de acetonitrilo desde el 10% al 80% durante 7 min. El flujo utilizado fue de 0,5 ml/min y la temperatura de la columna de 35°C. Para el análisis de masas acoplado a HPLC (HPLC/MS) se usó un sistema cromatográfico Alliance acoplado a un espectrómetro de masas ZQ4000 y a una columna Symmetry C18 (2.1 x 150 mm, Waters). Los solventes  
10 utilizados fueron los mismos que los descritos anteriormente y la elución se realizó manteniendo inicialmente un 10% de acetonitrilo durante 4 min, seguido de un gradiente lineal desde el 10% al 88% de acetonitrilo durante 26 min, utilizando para ello un flujo de 0,25 ml/min. El análisis de masas se realizó por ionización electrospray en modo positivo, con un voltaje de capilar de 3 kV y un voltaje de cono de 20 V. La detección de los picos y la caracterización de su espectro de absorción se realizó en ambos casos con un sistema de fotodiodos en línea y utilizando el software Empower de  
15 Waters, extrayéndose cromatogramas bidimensionales a una longitud de onda de 332 nm.

Para la caracterización estructural de los derivados de colismicina A mencionados en esta invención se realizaron cultivos de las cepas de *Streptomyces spp.* CS40 correspondientes. Tras centrifugación para eliminar las células, los caldos resultantes se sometieron a extracción en fase sólida en cartuchos de extracción con relleno C18 (Sep-Pak Vac, Waters). Los compuestos buscados se purificaron a partir de los extractos resultantes utilizando HPLC preparativa en fase reversa. Todas las purificaciones se hicieron en condiciones isocráticas con mezclas de acetonitrilo o metanol y 0.05% TFA en agua, en proporciones optimizadas para cada compuesto. Las columnas empleadas fueron una XTerra PrepRP18 (19 x 300 mm, Waters) y una Symmetry C18 (7,8 x 300 mm, Waters). Las soluciones con los picos purificados se evaporaron parcialmente en rotavapor para reducir su contenido en solvente orgánico y posteriormente se cargaron en un cartucho de extracción en fase sólida (Sep-Pak C18, Waters). Los compuestos retenidos se lavaron sucesivamente con agua, 0.1% amoníaco en agua y de nuevo agua, para eliminar totalmente el TFA. Finalmente se eluyeron con metanol, se secaron en vacío y por último se liofilizaron. La identificación y análisis cuantitativo de colismicina A se llevó a cabo mediante cromatografía en fase reversa en un equipo Acquity UPLC utilizando una columna BEH C18 (2.1 x 100 mm, Waters) y utilizando acetonitrilo y 0.05% TFA como solventes. Las muestras fueron eluidas con acetonitrilo al 10% durante 1 min seguido por un gradiente lineal de acetonitrilo desde el 10% al 80% durante 7 min. El flujo utilizado fue de 0,5 ml/min y la temperatura de la columna de 30°C. Para el análisis de masas acoplado a HPLC (HPLC/MS) se usó un sistema cromatográfico Alliance acoplado a un espectrómetro de masas ZQ4000 y a una columna Symmetry C18 (2.1 x 150 mm, Waters). Los solventes utilizados fueron los mismos que los descritos anteriormente y la elución se realizó con un gradiente isocrático inicial de acetonitrilo al 10% mantenido durante 4 min y seguido por un gradiente lineal de acetonitrilo desde el 10% al 88% durante 26 min, utilizando para ello un flujo de 0,25 ml/min. El análisis de masas se realizó por ionización electrospray en modo positivo, con un voltaje de capilar de 3 kV y un voltaje de cono de 20 V. La detección de los picos y la caracterización de su espectro de absorción se realizaron en ambos casos con un sistema de fotodiodos en línea y utilizando el software Empower de Waters,  
40 extrayéndose cromatogramas bidimensionales a una longitud de onda de 360 nm.

La colismicina A analizada en UPLC presenta una retención de 3.30 min y un espectro de absorción con máximos a 250 y 332 nm. La colismicina A analizada en HPLC/MS presenta una retención de 15.13 min y muestra un ión molecular con una masa de 276 m/z [M+H]<sup>+</sup>. (FIG. 3)

45 Para la caracterización estructural de los derivados de colismicina mencionados en esta invención se realizaron cultivos de las cepas de *Streptomyces spp.* CS40 correspondientes y los extractos fueron disueltos en 5 ml de una mezcla de DMSO y metanol a partes iguales. Posteriormente se centrifugaron y se eliminó la capa superior correspondiente a la fracción lipídica. El primer paso de purificación se realizó por cromatografía en una columna XTerra PrepRP18 (19 x 300 mm, Waters) usando como solventes acetonitrilo y ácido trifluoroacético (TFA) al 0,05% disuelto en agua. Se utilizó un gradiente lineal desde el 30% al 100% de acetonitrilo durante 7 min seguido por 3 min de acetonitrilo al 100%. El flujo utilizado fue de 15 ml/min. Los picos de interés fueron recolectados sobre tampón fosfato 0,1 M, pH 7,0. Las soluciones obtenidas fueron parcialmente evaporadas en un rotavapor para reducir la concentración de acetonitrilo y posteriormente se aplicaron a un cartucho de extracción en fase sólida (Sep-Pak C18, Waters), se lavaron posteriormente con agua para eliminar las sales y se eluyeron con metanol. Las purificaciones posteriores se realizaron en condiciones isocráticas, optimizadas para cada pico, utilizando una columna Symmetry C18 (7,8 x 300 mm, Waters) y mezclas de acetonitrilo y TFA al 0,05% disuelto en agua, usando un flujo de 7 ml/min. Tal como se mencionó anteriormente, los picos se recogieron siempre sobre tampón fosfato 0,1 M, pH 7,0, se desalaron utilizando extracción en fase sólida y finalmente se liofilizaron.  
55

### 2.3. Manipulación de ADN.

La preparación de plásmidos, ADN total, digestiones con enzimas de restricción, ligaciones de ADN, etc., se llevó a cabo siguiendo métodos estandarizados previamente descritos (Sambrook *et al.*, Molecular cloning: a laboratory manual. Cold Spring Harbour, NY: Cold Spring Harbour Laboratory press, 1989; Kieser *et al.*, Practical *Streptomyces* Genetics. The John Innes Foundation. Norwich, 2000). Los fragmentos de ADN fueron aislados de geles de agarosa utilizando el kit de extracción QUIAquick Gel Extraction Kit de QIAGEN (Hilden, Germany) y marcados usando el kit de detección y marcaje DIG DNA Labelling and Detection Kit de Roche Diagnostics (Manheim, Germany) utilizado para el análisis por Southern de acuerdo a las instrucciones del fabricante. La secuenciación fue realizada sobre ADN de doble cadena utilizando el método de descrito por Sanger *et al.*, Proc. Natl. Acad. Sci. USA. 74, 5463-5467 (1977) y el kit de secuenciación Cy5 Autocycle Sequencing Kit (Pharmacia Biotech). La electroforesis de las muestras fue realizada en un secuenciador automático Alf-express (Pharmacia Biotech). Las secuencias obtenidas se analizaron usando el paquete informático de programas del GCG, del Genetics Computer Group de la Universidad de Wisconsin (Devereux *et al.*, Nucleic Acids Res. 12, 387-395, 1984) y el programa BLAST (Altschul *et al.*, J. Mol. Biol. 215, 403-410, 1990). El análisis de las regiones transmembrana de posibles proteínas transmembranales se realizó utilizando el programa TMHMM v. 2.0 (Krogh *et al.*, J. Mol. Biol. 305, 567-580, 2001). El análisis de PCS y NRPS se realizó utilizando los programas ASMPKS (Tae *et al.*, BMC Bioinformatics. 8, 327-335, 2007) y NRPSpredictor (Rausch *et al.*, Nucleic Acids Res. 33, 5799-5808, 2005).

### 2.4. Amplificación de fragmentos de ADN por PCR y clonación de un fragmento de ADN que codifica parte de una lisina 2-aminotransferasa del genoma de *Streptomyces* spp. CS40.

La estrategia empleada para la clonación de la agrupación de genes de biosíntesis de colismicina fue la utilización de la homología genética con algunas proteínas codificadas por genes previamente caracterizados y para los cuales debería existir un homólogo en la ruta de biosíntesis de colismicina.

La información disponible sobre la biosíntesis de ácido picolínico (Bruntner y Bormann, 1998; Namwat *et al.*, 2002), un intermediario en la ruta de biosíntesis de caerulomicina (Vining *et al.*, 1988) y posiblemente también en la de colismicina fue usada para diseñar los oligonucleótidos degenerados L2ATFW2 (SEQ ID NO: 29) y L2ATRV2 (SEQ ID NO: 30) y tratar de amplificar una secuencia perteneciente a un gen homólogo a lisina 2-aminotransferasas.

Para la obtención del ADN total de *Streptomyces* spp. CS40 el microorganismo se cultivó en medio líquido TSB y el ADN total se aisló como ha sido descrito por Kieser *et al.*, (2000). El ADN total de *Streptomyces* spp. CS40 fue utilizado como molde para la reacción en cadena de la polimerasa utilizando los oligonucleótidos L2ATFW2 (SEQ ID NO: 29) y L2ATRV2 (SEQ ID NO: 30). Se asumió que, como consecuencia de la amplificación se obtendría un fragmento de ADN de aproximadamente 0.6 kb que contendría la parte interna del gen que codifica para una lisina 2-aminotransferasa. La reacción en cadena de la polimerasa se llevó a cabo en un volumen total de 50 µl y la mezcla de reacción contenía 0,1 µg de ADN total de *Streptomyces* spp. CS40, 2.5% de dimetilsulfóxido (DMSO), 200 pmoles de cada cebador, dNTPs (concentración final 200 µM), 1xPCR del enzima Taq DNA polimerasa (Invitrogen). La reacción en cadena se realizó en un termociclador GeneAmp® PCR System9700 de Applied Biosystems con el siguiente programa: 1 ciclo de desnaturalización a 98°C (5 min), 30 ciclos de desnaturalización/ anillamiento/ síntesis a 94°C (1 min) / 55°C (1 min) / 72°C (1 min) y 1 ciclo de extensión final a 72°C (10 min). El fragmento de ADN obtenido con este procedimiento fue clonado en el vector de *E. coli* pGemT usando el procedimiento recomendado por el fabricante y fue sometido a secuenciación utilizando técnicas estandarizadas. El análisis de la secuencia del fragmento amplificado por PCR reveló que contenía parte de una proteína que presentaba claras similitudes a nivel de aminoácidos con varias lisina 2-aminotransferasas previamente caracterizadas: NikC de *Streptomyces tendae*, número de acceso CAA75797 (Bruntner y Bormann, 1998) y VisA de *Streptomyces virginiae*, número de acceso BAB83671 (Namwat *et al.*, 2002).

A partir de esta secuencia homóloga a lisina 2-amino transferasas amplificada fueron diseñados dos oligonucleótidos para construir una sonda genética. Los oligonucleótidos sintéticos L2ATson1 (SEQ ID NO: 31) y L2ATson2 (SEQ ID NO: 32) fueron utilizados como cebadores para la amplificación de la sonda genética por medio de la reacción en cadena de la polimerasa (PCR) y usando como ADN molde el ADN cromosómico de *Streptomyces* spp. CS40. La reacción en cadena de la polimerasa se llevó a cabo en un volumen total de 50 µl y la mezcla de reacción contenía 0,1 µg de ADN total de *Streptomyces* spp. CS40, 2.5% de dimetilsulfóxido (DMSO), 200 pmoles de cada cebador, dNTPs (concentración final 200 µM), 1xPCR del enzima pfx DNA polimerasa (Invitrogen). La reacción en cadena se realizó en un termociclador GeneAmp® PCR System9700 de Applied Biosystems con el siguiente programa: 1 ciclo de desnaturalización a 98°C (5 min), 30 ciclos de desnaturalización/ anillamiento/ síntesis a 94°C (1 min) / 60°C (1 min) / 68°C (1 min) y 1 ciclo de extensión final a 68°C (10 min). El fragmento de ADN obtenido con este procedimiento fue clonado en el vector de *Escherichia coli* pCR-Blunt y fue sometido a secuenciación utilizando técnicas estandarizadas.

Una vez confirmado que el fragmento amplificado formaba parte de la región codificadora de la lisina 2-aminotransferasa este fragmento se utilizó como sonda genética para el análisis de una genoteca de ADN cromosómico de *Streptomyces* spp. CS40.

### 2.5. Construcción y análisis de la genoteca de ADN cromosómico de *Streptomyces* spp. CS40.

La genoteca de ADN cromosómico de *Streptomyces* spp. CS40 fue construida en el cósmido pWE15 que es capaz de replicarse en *E. coli*. El ADN genómico de *Streptomyces* spp. CS40 fue aislado como se ha descrito anteriormente, fue digerido parcialmente con Sau3AI y los fragmentos obtenidos, de un tamaño aproximado de 35-40 kb, fueron desfosforilados por tratamiento con fosfatasa alcalina (Roche Diagnostics, Mannheim). El cósmido pWE15, utilizado como vector, fue linearizado con BamHI. Los fragmentos de ADN y el vector fueron ligados y empaquetados *in vitro* usando un kit comercial de empaquetamiento Gigapack III Gold packaging Extract kit siguiendo las instrucciones de la casa comercial (Stratagene). Las partículas de ADN recombinante fueron utilizadas para infectar células de *E. coli* XLI Blue MR y los transductantes fueron seleccionados en placas con medio LA (Luria-Bertani agar) conteniendo como antibiótico de selección ampicilina. Aproximadamente 1000 colonias transductantes fueron cultivadas en placas de microtitulación conteniendo medio LB (Luria-Bertani broth) y el antibiótico de selección. Después de su incubación a 37°C durante 24 h, fueron mantenidas en presencia de glicerol al 25% a -70°C para su preservación.

Con objeto de clonar la agrupación de genes de biosíntesis de colismicina se llevó a cabo el análisis de la genoteca de *Streptomyces* spp. CS40 mediante hibridación *in situ* de colonias con la sonda mencionada anteriormente. Los transductantes fueron transferidos de las placas de microtitulación a placas de medio LA conteniendo como antibiótico de selección ampicilina y tras una noche de crecimiento a 37°C las colonias fueron transferidas a filtros de nylon para su hibridación *in situ* siguiendo los protocolos descritos por Sambrook *et al.*, (1989). Los filtros fueron analizados con las sondas marcadas utilizando el kit comercial DIG DNA labeling and detection kit de Roche. De este modo (y como se detalló anteriormente) se aisló el cósmido cos1C3.

#### Ejemplo 3. Obtención y análisis de la secuencia nucleotídica del agrupamiento génico responsable de la biosíntesis de colismicina y deducción de las funciones de los genes.

El cósmido cos1C3 y el fragmento EcoRV/HindIII de 7,7 kb procedente del cósmido cos3B11 subclonado en los mismos sitios en el vector pBluescriptSK+ fueron secuenciados en su totalidad. La secuenciación fue realizada sobre ADN de doble cadena utilizando el método de descrito por Sanger *et al.*, (1977) y utilizando el kit de secuenciación Cy5 Autocycle Sequencing Kit (Pharmacia Biotech). La electroforesis de las muestras fue realizada en un secuenciador automático Alf-express (Pharmacia Biotech) y los datos obtenidos se analizaron usando el paquete informático de programas del GCG, del Genetics Computer Group de la Universidad de Wisconsin (Devereux *et al.*, 1984) y el programa BLAST (Altschul *et al.*, 1990). El análisis de las regiones transmembrana de posibles proteínas transmembranales se realizó utilizando el programa TMHMM v. 2.0 (Krogh *et al.*, J. Mol. Biol. 305, 567-580, 2001). El análisis de PCS y NRPS se realizó utilizando los programas ASMPKS (Tae *et al.*, 2007) y NRPSpredictor (Rausch *et al.*, 2005).

El análisis informático de la secuencia de ADN de 46668 bp (SEQ ID NO: 1) mostró la presencia de 27 pautas de lectura abierta (ORFs) (FIG. 2 y Tabla 2) con un alto contenido en G+C característico del ADN de *Streptomyces*. Además se muestra un contenido especialmente alto en G+C alto en la tercera posición de los codones característico de genes de *Streptomyces*. Las funciones de los genes fueron deducidas por comparación de las secuencias aminoacídicas, traducidas conceptualmente a partir de la secuencia nucleotídica, con secuencias conocidas disponibles en bases de datos. Los resultados obtenidos se muestran en la Tabla 2 referidos a los datos de secuencia que se incluyen en la solicitud. 23 de las 27 ORFs encontradas, que ocupan una región de 37,5 kb, están probablemente implicadas en la biosíntesis de colismicina. Esta región esta flanqueada por 4 ORFs probablemente no implicadas en la biosíntesis de colismicina (en blanco en la FIG. 2). De estas ORFs, *orf1*, mostró gran similitud con una posible helicasa de *S. viridochromogenes* DSM 40736 (ZP\_05534927) y una posible kinasa de *S. scabiei* 87.22, (YP\_003487313).

De las 23 ORFs que se proponen implicadas en la biosíntesis de colismicina A, 16 ORFs están implicadas en la biosíntesis de la parte estructural de la colismicina, 5 de ellas (*orf3-7*) están implicadas en el transporte de colismicina, 2 ORFs (*orf2* y *orf8*) están probablemente implicadas en procesos de regulación.



**Tabla 2.** Genes identificados en la región del cromosoma de *Streptomyces spp.* CS40 implicada en la biosíntesis de colismicina.

Gen	Posición	Aminoácidos	Función deducida	Notas
<i>orf1</i>	549-2237 compl.	563	Desconocida	SeqID.NO:2
<i>orf2</i>	2857-3414	186	Regulador transcripcional homólogo a TetR	SeqID.NO:3
<i>orf3</i>	3514-5247 compl.	578	Transportador	SeqID.NO:4
<i>orf4</i>	5256-7016 compl.	587	Transportador	SeqID.NO:5
<i>orf5</i>	7225-8193	323	Transportador	SeqID.NO:6
<i>orf6</i>	8322-9305	328	Transportador	SeqID.NO:7
<i>orf7</i>	9314-10099	262	Transportador	SeqID.NO:8
<i>orf8</i>	10363-10869	169	Regulador transcripcional homólogo a LuxR	SeqID.NO:9
<i>orf9</i>	11042-12070 compl.	343	Metil transferasa	SeqID.NO:10
<i>orf10</i>	12105-13703 compl.	533	Deshidrogenasa	SeqID.NO:11
<i>orf11</i>	13861-15495 compl.	545	Aminotransferasa	SeqID.NO:12
<i>orf12</i>	15620-16606	329	Metiltransferasa	SeqID.NO:13
<i>orf13</i>	16688-18076 compl.	463	Proteína similar a GriD (YP_001825756) Arylcarboxilato reductasa componente	SeqID.NO:14
<i>orf14</i>	18078-19109 compl.	344	Proteína similar a GriC (YP_001825755) Arylcarboxilato reductasa componente	SeqID.NO:15
<i>orf15</i>	19405-20613	403	Monoxigenasa	SeqID.NO:16
<i>orf16</i>	20680-21921 compl.	414	Amidohidrolasa	SeqID.NO:17
<i>orf17</i>	21926-23614 compl.	563	Enzima formadora de adenilato	SeqID.NO:18

<i>orf18</i>	23611-23871 compl.	87	Proteína de unión a fosfopanteteína	SeqID.NO:19
<i>orf19</i>	24108-25484	459	Lisina 2-aminotransferasa	SeqID.NO:20
<i>orf20</i>	25481-26662	394	Sarcosina oxidasa monomérica	SeqID.NO:21
<i>orf21</i>	26548-34206	2553	Proteína híbrida PCS/NRPS	SeqID.NO:22
<i>orf22</i>	34203-37478	1092	NRPS	SeqID.NO:23
<i>orf23</i>	37475-38659	395	Deshidrogenasa	SeqID.NO:24
<i>orf24</i>	38652-39380	243	Tioesterasa	SeqID.NO:25
<i>orf25</i>	39442-40665 compl.	408	Desconocida	SeqID.NO:26
<i>orf26</i>	41103-41708	202	Desconocida	SeqID.NO:27
<i>orf27</i>	42017-46105	1363	Desconocida	SeqID.NO:28

#### Ejemplo 4. Inactivación de la agrupación de genes de biosíntesis de colismicina mediante disrupción génica.

5 Con objeto de demostrar la implicación de la agrupación de genes clonados en la biosíntesis de colismicina se llevó a cabo la inactivación del gen *orf21*.

10 Para la inactivación de la *orf21* se aisló un fragmento BamHI del cósmido cos1C3. Este fragmento de 3128 bp es interno a *orf21* (sitios BamHI 13 y 14, FIG. 2) y fue clonado en el plásmido pOJ260 digerido con BamHI. La construcción resultante, pOJB12 fue introducida en *Streptomyces spp.* CS40 mediante conjugación intergenérica desde *E. coli* ET12567 (pUB307) para generar la cepa mutante CLM-12 que fue seleccionada por su resistencia a apramicina.

15 La mutación en el gen *orf21* fue comprobada mediante análisis por Southern. El mutante se demostró no productor de colismicina A mediante análisis por UPLC de muestras de cultivos de *Streptomyces spp.* CS40 y CLM-12 extraídas con acetato de etilo (FIG. 4), confirmando de este modo la implicación de la agrupación de genes en la biosíntesis de colismicina.

#### Ejemplo 5. Establecimiento de los límites de la agrupación de genes de biosíntesis de colismicina mediante reemplazamiento génico.

20 Para establecer los límites del agrupamiento génico implicado en la biosíntesis de colismicina se realizaron dos mutantes por disrupción génica y un mutante por reemplazamiento génico en los genes *orf1* (mutante CLM-H) y *orf27* (mutante CLM-24) y *orf25* (mutante CLM-22D) respectivamente. En el caso de los mutantes CLM-H y CLM-24 las ORFs se interrumpen por introducción de un gen de resistencia a apramicina y en el caso del mutante CLM-22D la ORF se sustituye por el gen de resistencia a apramicina.

25 Para la obtención del mutante en la *orf1* se amplificó una región de 910 pb correspondiente a un fragmento interno a la *orf1* con los oligonucleótidos HelDisr1 (SEQ ID NO: 34) y HelDisr2 (SEQ ID NO: 35). Dicho fragmento se clonó en el vector pCR-Blunt para generar el plásmido pCRBH y se sometió a secuenciación utilizando técnicas estandarizadas. Posteriormente este fragmento se obtuvo del plásmido pCRBH digiriendo el mismo con EcoRI y se clonó en el mismo sitio de restricción del vector pOJ260, generando así el plásmido pOJBH que se usó para la disrupción génica en *Streptomyces spp.* CS40.

30 En ningún caso fuimos capaces de obtener mutantes que llevaran interrumpida la *orf1* sugiriendo que este gen es esencial para la viabilidad celular. Este dato descarta la participación de la *orf1* en la biosíntesis de colismicina, puesto que los distintos mutantes (CLM-L, CLM-12) que presentan afectada la biosíntesis de colismicina son perfectamente viables, excluyendo así este gen del cluster de biosíntesis de colismicina.

Para la obtención del mutante CLM-24 se obtuvo un fragmento ClaI-BamHI de 2271 bp, procedente del cósmido cos1C3 que corresponde a un fragmento interno a la *orf27*. Dicho fragmento se hizo romo y se clonó en el vector pOJ260 en el sitio EcoRV generando así el plásmido pOJB23P que se usó para la disrupción génica en *Streptomyces spp.* CS40 generando el mutante CLM-24.

Para la obtención del mutante CLM-22D se empleó la técnica de reemplazamiento génico. Se amplificó mediante PCR un fragmento de 1526 bp que se encuentra flanqueando corriente arriba al gen *orf25*. Para ello se usaron los oligonucleótidos ORF22del1 (SEQ ID NO: 36) y ORF22del2 (SEQ ID NO: 37) a los que se les introdujeron las dianas de restricción EcoRI y HindIII respectivamente, dicho fragmento se clonó en el vector pCR-Blunt y fue sometido a secuenciación utilizando técnicas estandarizadas. Posteriormente dicho fragmento se clonó en el vector pUO9090 en los sitios de restricción EcoRI y HindIII generando el plásmido pUO909022A. Adicionalmente se amplificó mediante PCR un segundo fragmento de 1496 bp que se encuentra flanqueando corriente abajo al gen *orf25*. Para ello se usaron los oligonucleótidos ORF22del3 (SEQ ID NO: 38) y ORF22del4 (SEQ ID NO: 39) a los que se les introdujo la diana EcoRV, este fragmento se clonó en el vector pCR-Blunt y fue sometido a secuenciación utilizando técnicas estandarizadas. Posteriormente dicho fragmento se clonó en el vector pUO909022A usando el sitio de restricción EcoRV para generar el vector pUO909022AB.

Para el reemplazamiento génico el cassette de reemplazamiento se obtuvo del vector pUO909022AB digerido con SpeI y se clonó en el vector pHZ1358. La construcción resultante, pHZ22AB, fue usada para el reemplazamiento génico en *Streptomyces spp.* CS40 generando el mutante CLM-22D.

La integración del gen de resistencia a apramicina en el caso de los mutantes generados por disrupción así como el reemplazamiento en el cromosoma de la copia silvestre del gen por la mutada fue confirmado en los transconjugantes mediante análisis por Southern.

Cada uno de los mutantes fue analizado para conocer la producción de colismicina A, en paralelo con la cepa parental *Streptomyces spp.* CS40 mediante UPLC. Los mutantes CLM-22D y CLM-24, se mostraron como productores de colismicina A (FIG. 5), indicando que los genes mutados no participan en la biosíntesis de colismicina y confirmando de este modo los límites del agrupamiento génico implicado en la biosíntesis de colismicina.

#### **Ejemplo 6. Generación de un mutante en los pasos iniciales de la biosíntesis de colismicina.**

El primero de los pasos propuestos para la biosíntesis de colismicina (FIG 6) consiste en la formación de ácido picolínico a partir de lisina y el primer gen implicado en este proceso sería una lisina 2-aminotransferasa codificada por la *orf19*. Para confirmar este dato se inactivó el gen *orf19*.

Para la inactivación de la *orf19* se empleó la técnica de reemplazamiento génico. Se amplificó mediante PCR un fragmento de 1589 bp que se encuentra flanqueando corriente arriba al gen *orf19*. Para ello se usaron los oligonucleótidos ORF16del1 (SEQ ID NO: 40) y ORF16del2 (SEQ ID NO: 41) a los que se les introdujeron las dianas de restricción EcoRI y HindIII respectivamente, dicho fragmento se clonó en el vector pCR-Blunt y fue sometido a secuenciación utilizando técnicas estandarizadas. Posteriormente dicho fragmento se clonó en el vector pUO9090 en los sitios de restricción EcoRI y HindIII generando el plásmido pUO909016A. Adicionalmente se amplificó mediante PCR un segundo fragmento de 1559 bp que se encuentra flanqueando corriente abajo al gen *orf19*. Para ello se usaron los oligonucleótidos ORF16del3 (SEQ ID NO: 43) y ORF16del4 (SEQ ID NO: 44) a los que se les introdujeron las dianas de restricción BamHI y EcoRV respectivamente, este fragmento se clonó en el vector pCR-Blunt y fue sometido a secuenciación utilizando técnicas estandarizadas. Posteriormente dicho fragmento se clonó en el vector pUO909016A usando los sitios de restricción EcoRV y BamHI para generar el vector pUO909016AB.

Para el reemplazamiento génico el cassette de reemplazamiento se obtuvo del vector pUO909016AB digerido con SpeI y se clonó en el vector pHZ1358. La construcción resultante, pHZ16AB, fue introducida en *Streptomyces spp.* CS40 mediante conjugación intergenérica desde *E. coli* ET12567 (pUB307) para generar la cepa mutante CLM-L. La cepa CLM-L se encuentra depositada en la Colección Española de Cultivos Tipo con número de depósito 7754.

El mutante se demostró no productor de colismicina mediante análisis por UPLC de muestras de cultivos de *Streptomyces spp.* CS40 y CLM-L extraídas con acetato de etilo (FIG. 7B), confirmando de este modo la implicación de este gen en la biosíntesis de colismicina.

Con el fin de corroborar la participación del gen *orf19* en los pasos iniciales de biosíntesis se analizó por UPLC la producción de colismicina A por parte del mutante CLM-L en presencia de ácido picolínico (concentración final 1 mM), un posible metabolito intermediario, en el medio de cultivo. Las muestras extraídas con acetato de etilo mostraron que el mutante CLM-L en presencia de ácido picolínico recupera la capacidad para producir colismicina A (FIG 7C). Este dato confirma la participación de *orf19* en los pasos iniciales de biosíntesis y la existencia de ácido picolínico como metabolito intermediario en la ruta.

**Ejemplo 7. Generación de mutantes productores de nuevos derivados de colismicina mediante reemplazamiento génico.**

- 5 Para generar mutantes de *Streptomyces spp.* CS40 capaces de producir colismicinas químicamente modificadas se seleccionaron de manera individual los genes *orf11* y *orf9* y de manera conjunta *orf13* y *orf14* (FIG. 2), todos ellos posiblemente implicados en pasos finales de la biosíntesis de colismicina. Para ello se construyeron los plásmidos plásmidos pHZ6AB, pHZ8AB y pHZ101AB.
- 10 En primer lugar se amplificó mediante PCR, usando los oligonucleótidos sintéticos orf8del3 (SEQ ID NO: 45) y orf8del4 (SEQ ID NO: 46) un fragmento de 1499 bp correspondiente a la región corriente abajo del gen *orf11*, que se clonó en el vector pCR-Blunt para dar el plásmido pCRB8B que fue sometido a secuenciación utilizando técnicas estandarizadas. Este fragmento se extrajo del plásmido pCRB8B usando los sitios EcoRI existentes en el polilinker del vector pCRB-Blunt, se hizo romo y subclonó en el plásmido pU09090 para dar el plásmido pUO8B.
- 15 Posteriormente se amplificó por PCR usando los oligonucleótidos sintéticos orf8del1 (SEQ ID NO: 47) orf8del2 (SEQ ID NO: 48) que incluían sitios de restricción EcoRI y HindIII respectivamente para facilitar la subclonación, un fragmento de 1516 bp correspondiente a la región corriente arriba del gen *orf11*, que se clonó en el vector pCR-Blunt para dar el plásmido pCRB8A que fue sometido a secuenciación utilizando técnicas de secuenciación estandarizadas. Este fragmento se extrajo del plásmido pCRB8A usando los sitios de restricción EcoRI y HindIII y se subclonó en los mismos sitios del plásmido pUO8B para generar el plásmido PUO8AB.
- 20 El cassette de reemplazamiento se extrajo del plásmido pUO8AB usando los sitios de restricción SpeI y se clonó en el vector pHZ1358 para generar el plásmido pHZ8AB que se introdujo en *Streptomyces spp.* CS40 por conjugación intergenérica desde *E. coli* ET12567 (pUB307) para generar la cepa mutante CLM-A. La cepa CLM-A se encuentra depositada en la Colección Española de Cultivos Tipo con número de depósito 7755.
- 25 Para la generación de la cepa mutante CLM-M2 se amplificó mediante PCR un fragmento de 1504 bp perteneciente a la región corriente arriba del gen *orf9*. Para ello se utilizaron los oligonucleótidos orf6del1 (SEQ ID NO: 49) y orf6del2b (SEQ ID NO: 50) que llevaban los sitios de restricción EcoRI y HindIII respectivamente. Este fragmento se clonó en el vector pCR-Blunt generando así el plásmido pCRB6A que se secuenció usando técnicas estandarizadas. Posteriormente se obtuvo el fragmento digiriendo el plásmido pCRB6A con EcoRI y HindIII y se clonó en los mismos sitios del vector pU09090 generando el vector pUO6A.
- 30 Después se amplificó mediante PCR un fragmento de 1504 bp perteneciente a la región corriente abajo del gen *orf9*. Para ello se utilizaron los oligonucleótidos orf6del3 (SEQ ID NO: 51) y orf6del4 (SEQ ID NO: 52). Este fragmento se clonó en el vector pCR-Blunt generando así el plásmido pCRB6B que se secuenció usando técnicas estandarizadas. Posteriormente se obtuvo el fragmento digiriendo el plásmido pCRB6A con EcoRI, se hizo romo y se clonó en el sitio EcoRV del plásmido pUO6A generando así el plásmido pUO6AB.
- 35 El cassette de reemplazamiento se extrajo del plásmido pUO6AB usando los sitios de restricción SpeI y se clonó en el vector pHZ1358 para generar el plásmido pHZ6AB que se introdujo en *Streptomyces spp.* CS40 por conjugación intergenérica desde *E. coli* ET12567 (pUB307) para generar la cepa mutante CLM-M2. La cepa CLM-M2 se encuentra depositada en la Colección Española de Cultivos Tipo con número de depósito 7756.
- 40 Para la generación del doble mutante en los genes *orf13* y *orf14* se amplificó mediante PCR un fragmento de 1496 bp perteneciente a la región corriente arriba del gen *orf14*. Para ello se utilizaron los oligonucleótidos orf101del1 (SEQ ID NO: 53) y orf101del2 (SEQ ID NO: 54) que llevaban los sitios de restricción EcoRI y HindIII respectivamente (secuencia subrayada). Este fragmento se clonó en el vector pCR-Blunt generando así el plásmido pCRB101A que se secuenció usando técnicas estandarizadas. Posteriormente se obtuvo el fragmento digiriendo el plásmido pCRB101A con EcoRI y HindIII y se clonó en los mismos sitios del vector pU09090 generando el vector pUO101A.
- 45 Después se amplificó mediante PCR un fragmento de 1522 bp perteneciente a la región corriente abajo del gen *orf13*. Para ello se utilizaron los oligonucleótidos orf101del3 (SEQ ID NO: 55) y orf101del4 (SEQ ID NO: 56) que llevaban el sitio de restricción EcoRV (secuencia subrayada). Este fragmento se clonó en el vector pCR-Blunt generando así el plásmido pCRB101B que se secuenció usando técnicas estandarizadas. Posteriormente se obtuvo el fragmento digiriendo el plásmido pCRB101B con EcoRV y se clonó en el mismo sitio del plásmido pUO101A generando así el plásmido pUO101AB.
- 50 El cassette de reemplazamiento se extrajo del plásmido pUO101AB usando los sitios de restricción SpeI y se clonó en el vector pHZ1358 para generar el plásmido pHZ101AB que se introdujo en *Streptomyces spp.* CS40 por conjugación intergenérica desde *E. coli* ET12567 (pUB307) para generar la cepa mutante CLM-G. La cepa CLM-G se encuentra depositada en la Colección Española de Cultivos Tipo con número de depósito CECT7861.
- 55 En todos los casos los transconjugantes en los que había ocurrido un acontecimiento de doble sobrecruzamiento fueron seleccionados por su resistencia a apramicina y su sensibilidad a tiestreptona. El reemplazamiento en el cromosoma fue confirmado en los transconjugantes mediante análisis por Southern.
- 60
- 65

El análisis de cultivos de la cepa CLM-A por UPLC mostró dos picos con absorbancia característica de colismicina A pero con movilidades de 2,86 min y 2,99 min que corresponden a compuestos de fórmula II y a colismicina C (FIG. 8). El análisis posterior por HPLC/MS determinó que el compuesto de fórmula II presenta una movilidad de 13,42 min y un ión de 249 m/z [M+H]<sup>+</sup>. La colismicina C presenta una movilidad de 13,56 min y un ión de 263 m/z [M+H]<sup>+</sup>.

El análisis de cultivos de la cepa CLM-M2 por UPLC mostró dos picos con absorbancia característica de colismicina A pero con movilidades de 3,05 min y 4,1 min correspondientes a los compuestos de fórmula III y de fórmula IV (FIG. 9). El análisis posterior por HPLC/MS determinó que el compuesto de fórmula III presenta una movilidad de 15,16 min y un ión de 262 m/z [M+H]<sup>+</sup>. El compuesto de fórmula IV presenta una movilidad de 19,12 min y un ión de 244 m/z [M+H]<sup>+</sup>.

El análisis de cultivos de la cepa CLM-G por HPLC/MS mostró un pico con absorbancia característica de colismicina A ausente en la cepa silvestre (FIG 10A y B) correspondiente al compuesto de fórmula V y que presenta una movilidad de 6,15 min (FIG 10C) y un ión de 263 m/z [M+H]<sup>+</sup>.

#### 5 **Ejemplo 8. Generación de nuevos derivados de colismicina mediante la adición de distintos precursores a la cepas mutantes CLM-L, CLM-A y CLM-M2.**

Con el fin de generar nuevas moléculas derivadas de colismicina se suplementaron los medios de cultivo en los que se inocularon los mutantes CLM-L, CLM-A y CLM-M2 con dos análogos estructurales del ácido picolínico esperando que dichos compuestos fueran incorporados a la ruta de biosíntesis de colismicina.

Para ello se añadieron ácido 6-metil picolínico y ácido 4-metilpiridina-2-carboxílico a concentración final 0,7 mM al medio R5A en el que se inocularon los mutantes CLM-L, CLM-A y CLM-M2. Tras 6 días de crecimiento a 30°C se extrajeron los nuevos compuestos generados con acetato de etilo.

El análisis del cultivo de la cepa CLM-L al que se le añadió ácido 6-metil picolínico, por UPLC mostró un pico con absorbancia característica de colismicina pero con movilidad 3,06 min que corresponde con el compuesto de fórmula VI. El análisis posterior por HPLC/MS determinó que el compuesto VI presenta un ión de 290 m/z [M+H]<sup>+</sup> (FIG. 11A, B y C).

El análisis del cultivo de la cepa CLM-L al que se le añadió ácido 4-metilpiridina-2-carboxílico, por UPLC mostró un pico con absorbancia característica de colismicina pero con movilidad 3,14 min que corresponde con el compuesto de fórmula VII. El análisis posterior por HPLC/MS determinó que el compuesto de fórmula VII presenta un ión de 290 m/z [M+H]<sup>+</sup> (FIG 11D, E y F).

El análisis del cultivo por UPLC de la cepa CLM-A al que se le añadió ácido 4-metilpiridina-2-carboxílico, mostró un pico con absorbancia característica de colismicina pero con movilidad 3,14 min que corresponde con el compuesto de fórmula VIII (FIG. 12). El análisis posterior por HPLC/MS determinó que el compuesto de fórmula VIII presenta un ión de 277 m/z [M+H]<sup>+</sup>.

El análisis del cultivo de la cepa CLM-M2 al que se le añadió ácido 4-metilpiridina-2-carboxílico, por UPLC mostró dos picos con absorbancia característica de colismicina pero con movilidades 3,16 min y 3,51 min que corresponde con el compuesto de fórmula IX y fórmula X. El análisis posterior por HPLC/MS determinó que los compuestos de fórmula IX y fórmula X presentan un ión de 276 m/z [M+H]<sup>+</sup> y 258 m/z [M+H]<sup>+</sup> respectivamente (FIG. 13).

#### 45 **Ejemplo 9. Generación de nuevos derivados de colismicina mediante acilación enzimática catalizada por una lipasa.**

Las acilaciones enzimáticas de colismicina A y colismicina C catalizadas por PS-C se llevaron a cabo incubando a 45 °C y 250 rpm una suspensión de 20 mg de lipasa en una disolución de 5 mg de colismicina A o colismicina C en 2 ml de éter metil tertbutílico y 1 ml del agente acilante correspondiente. La conversión de la biotransformación fue monitoreada mediante HPLC, empleando un equipo cromatográfico Agilent Technologies 1200 Series, usando como solventes acetonitrilo y agua (sin TFA) y una columna de fase reversa (Zorbax Eclipse XDB-C18, RR, 1.8 μm, 4.6 x 50 mm, Agilent). Las muestras fueron eluidas empleando un método de tres gradientes lineales, el primero del 10% al 60% de acetonitrilo a lo largo de 5.7 minutos, a continuación otro gradiente del 60% al 100% de acetonitrilo durante 0.4 minutos, a continuación 0.55 minutos en isocrático a 100% de acetonitrilo y un gradiente final del 100% al 10% de acetonitrilo durante 0.35 minutos, a un flujo de 2 ml/min. La longitud de onda a la que se obtuvieron los cromatogramas fue de 332 nm. En este método la colismicina A y la colismicina C presentan una movilidad de 2.29 y 1.90 min respectivamente. Cuando la conversión alcanzó un valor próximo al 90%, el enzima se filtró a vacío en placa filtrante y se lavó con abundante éter metil tertbutílico y metanol. Para el caso de la Colismicina C, el filtrado se concentró a vacío y el residuo resultante, previamente disuelto en 1 ml de metanol, fue cromatografiado en una columna XBridge Prep C18 (30x150 mm, Waters), utilizando como fase móvil mezclas de acetonitrilo y agua a un flujo de 20 ml/min. Los picos de interés se diluyeron cuatro veces con agua y posteriormente se concentraron mediante extracción en fase sólida. Para el caso de la Colismicina A, el filtrado se concentró a vacío y el residuo resultante se purificó mediante cromatografía de columna empleando mezclas de hexano:acetato de etilo. Por último, los compuestos obtenidos fueron liofilizados para su conservación. Los nuevos derivados acilados se identificaron inicialmente mediante análisis por HPLC, comparando el

espectro de absorción y el tiempo de retención. A continuación, se caracterizaron mediante el análisis de MS y RMN de acuerdo con los ejemplos anteriores.

5 - Para el caso particular de emplear colismicina C y acetato de vinilo como agente acilante el tiempo de reacción fue de 6 horas y se alcanzó una conversión de 95%. Por HPLC se identificó un nuevo compuesto con tiempo de retención de 4.51 minutos el cual se purificó empleando como fase móvil una mezcla de acetonitrilo y agua (50:50) a un flujo de 20 ml/min. Una vez aislado, el análisis posterior por RMN confirmó la estructura del compuesto de acuerdo a la fórmula XI.

10 - Para el caso particular de emplear colismicina C y butanoato de vinilo como agente acilante el tiempo de reacción fue de 10 horas y se alcanzó una conversión de 95%. Por HPLC se identificó un nuevo compuesto con tiempo de retención de 5.56 minutos el cual se purificó empleando como fase móvil una mezcla de acetonitrilo y agua (55:45) a un flujo de 20 ml/min. Una vez aislado, el análisis posterior por RMN confirmó la estructura del compuesto de acuerdo a la fórmula XII.

15 - Para el caso particular de emplear colismicina C y benzoato de vinilo como agente acilante el tiempo de reacción fue de 14 horas y se alcanzó una conversión de 90%. Por HPLC se identificó un nuevo compuesto con tiempo de retención de 6.22 minutos el cual se purificó empleando como fase móvil una mezcla de acetonitrilo y agua (60:40) a un flujo de 20 ml/min. Una vez aislado, el análisis posterior por RMN confirmó la estructura del compuesto de acuerdo a la fórmula XIII.

20 - Para el caso particular de emplear colismicina A y acetato de vinilo como agente acilante el tiempo de reacción fue de 5 horas y se alcanzó una conversión de 95%. Por HPLC se identificó un nuevo pico con absorbancia característica de colismicina y con movilidad de 2.72 minutos que corresponde con el compuesto de fórmula XIV. El análisis posterior por HPLC/MS determinó que el compuesto de fórmula XIV presenta un ión de 318 m/z [M+H]<sup>+</sup>.

25 - Para el caso particular de emplear colismicina A y cloroacetato de vinilo como agente acilante el tiempo de reacción fue de 12 horas y se alcanzó una conversión de 90%. Por HPLC se identificó un nuevo pico con absorbancia característica de colismicina y con movilidad de 2.60 minutos que corresponde con el compuesto de fórmula XV. El análisis posterior por HPLC/MS determinó que el compuesto de fórmula XV presenta un ión de 352 m/z [M+H]<sup>+</sup>.

30 - Para el caso particular de emplear colismicina A y butanoato de vinilo como agente acilante el tiempo de reacción fue de 9 horas y se alcanzó una conversión de 95%. Por HPLC se identificó un nuevo pico con absorbancia característica de colismicina y con movilidad de 3.66 minutos que corresponde con el compuesto de fórmula XVI. El análisis posterior por HPLC/MS determinó que el compuesto de fórmula XVI presenta un ión de 346 m/z [M+H]<sup>+</sup>.

35 Datos de espectros de RMN del compuesto de fórmula II:  
<sup>1</sup>H RMN (DMSO-*d*<sub>6</sub>, 600 MHz): 2.35 (s), 4.75 (s), 6.30 (sa), 7.10 (s), 7.55 (sa), 8.00 (sa), 8.25 (sa), 8.80 (sa), 11.10 (sa)  
<sup>13</sup>C RMN (DMSO-*d*<sub>6</sub>, 150 MHz): 15.9, 59.1, 112.1, 117.9, 121.2, 125.8, 138.7, 142.2, 148.5, 149.7, 152.0, 177.2

40 Datos de espectros de RMN del compuesto de fórmula III:  
<sup>1</sup>H RMN (acetona-*d*<sub>6</sub>, 600 MHz): 2.48 (s), 4.00 (s), 7.68 (dd, *J* = 7.6, 4.4 Hz), 7.89 (s), 8.16 (dt, *J* = 7.6, 1.6 Hz), 8.29 (d, *J* = 7.6 Hz), 8.82 (d, *J* = 4.4 Hz), 8.86 (s), 10.80 (s)  
<sup>13</sup>C RMN (acetona-*d*<sub>6</sub>, 150 MHz): 16.0, 110.0, 121.3, 122.9, 126.1, 138.6, 143.7, 145.9, 147.3, 149.5, 150.2, 174.1

45 Datos de espectros de RMN del compuesto de fórmula IV:  
<sup>1</sup>H RMN (acetona-*d*<sub>6</sub>, 600 MHz): 2.57 (s), 3.50 (s), 7.50 (dd, *J* = 7.6, 4.4 Hz), 7.99 (dt, *J* = 7.6, 1.6 Hz), 8.26 (s), 8.41 (d, *J* = 7.6 Hz), 8.70 (d, *J* = 4.4 Hz)  
<sup>13</sup>C RMN (acetona-*d*<sub>6</sub>, 150 MHz): 16.9, 110.2, 116.3, 121.0, 124.9, 125.3, 137.5, 137.8, 149.3, 153.6, 157.7, 166.3

50 Datos de espectros de RMN del compuesto de fórmula V:  
<sup>1</sup>H RMN (DMSO-*d*<sub>6</sub>, 600 MHz): 2.36 (s), 7.47 (dd, *J* = 7.8, 4.2 Hz), 7.94 (t, *J* = 7.8 Hz), 7.96 (s), 8.29 (d, *J* = 7.8 Hz), 8.68 (d, *J* = 4.2 Hz), 11.7 (sa)  
<sup>13</sup>C RMN (DMSO-*d*<sub>6</sub>, 150 MHz): 17.7, 108.4, 116.1, 121.2, 125.1, 137.9, 149.8, 154.5, 156.0, 157.6, 166.7, 168.4

55 Datos de espectros de RMN del compuesto de fórmula VI:  
<sup>1</sup>H RMN (acetona-*d*<sub>6</sub>, 600 MHz): 2.43 (s), 2.67 (s), 4.19 (s), 7.45 (d, *J* = 7.8 Hz), 7.96 (t, *J* = 7.8 Hz), 8.41 (d, *J* = 7.8 Hz), 8.18 (s), 8.88 (s), 10.70 (s)  
<sup>13</sup>C RMN (acetona-*d*<sub>6</sub>, 150 MHz): 17.2, 23.0, 56.2, 103.3, 118.8, 122.5, 124.7, 138.5, 146.7, 152.3, 153.0, 155.1, 157.8, 168.1

60 Datos de espectros de RMN del compuesto de fórmula VII:  
<sup>1</sup>H RMN (DMSO-*d*<sub>6</sub>, 600 MHz): 2.36 (s), 2.47 (s), 4.08 (s), 7.45 (d, *J* = 4.8 Hz), 8.03 (s), 8.32 (s), 8.61 (d, *J* = 4.8 Hz), 8.73 (s), 10.70 (s)  
<sup>13</sup>C RMN (DMSO-*d*<sub>6</sub>, 150 MHz): 17.3, 20.4, 56.1, 103.9, 121.9, 122.4, 126.2, 147.2, 148.6, 150.0, 153.0, 153.8, 155.4, 167.4

Datos de espectros de RMN del compuesto de fórmula VIII:

$^1\text{H}$  RMN (acetona- $d_6$ , 600 MHz): 2.39 (s), 2.49 (s), 4.59 (s), 4.86 (s), 7.30 (d,  $J = 4.8$  Hz), 8.12 (s), 8.48 (s), 8.56 (d,  $J = 4.8$  Hz)

$^{13}\text{C}$  RMN (acetona- $d_6$ , 150 MHz): 16.5, 20.2, 62.3, 102.6, 117.9, 121.6, 125.1, 148.2, 149.0, 154.8, 155.7, 160.8, 167.2

Datos de espectros de RMN del compuesto de fórmula XI:

$^1\text{H}$  RMN (DMSO- $d_6$ , 600 MHz): 2.14 (s), 2.35 (s), 4.06 (s), 5.40 (s), 7.49 (dd,  $J = 7.3$ , 4.6 Hz), 7.98 (dt,  $J = 7.3$ , 1.8 Hz), 8.02 (s), 8.35 (d,  $J = 7.3$  Hz), 8.71 (dd,  $J = 4.6$ , 1.8 Hz)

$^{13}\text{C}$  RMN (DMSO- $d_6$ , 150 MHz): 17.7, 21.1, 56.8, 65.8, 103.3, 120.3, 121.1, 125.1, 137.9, 149.7, 154.9, 156.1, 157.5, 167.3, 170.6

Datos de espectros de RMN del compuesto de fórmula XII:

$^1\text{H}$  RMN (DMSO- $d_6$ , 600 MHz): 0.91 (t,  $J = 7.2$  Hz), 1.60 (sext,  $J = 7.2$  Hz), 2.35 (s), 2.40 (t,  $J = 7.2$  Hz), 4.06 (s), 5.41 (s), 7.48 (dd,  $J = 7.8$ , 4.5 Hz), 7.97 (dt,  $J = 7.8$ , 1.8 Hz), 8.01 (s), 8.35 (d,  $J = 7.8$  Hz), 8.71 (dd,  $J = 4.5$ , 1.8 Hz)

$^{13}\text{C}$  RMN (DMSO- $d_6$ , 150 MHz): 14.0, 17.7, 18.5, 35.8, 56.8, 65.8, 103.3, 120.4, 121.1, 125.1, 137.8, 149.7, 154.9, 156.1, 157.6, 167.2, 172.9

Datos de espectros de RMN del compuesto de fórmula XIII:

$^1\text{H}$  RMN (DMSO- $d_6$ , 600 MHz): 2.38 (s), 4.08 (s), 5.68 (s), 7.44 (dd,  $J = 7.8$ , 4.4 Hz), 7.57 (t,  $J = 7.2$  Hz), 7.70 (t,  $J = 7.2$  Hz), 7.80 (dt,  $J = 7.8$ , 1.2 Hz), 8.03 (d solapado,  $J = 7.2$  Hz), 8.02 (s), 8.18 (d,  $J = 7.8$  Hz), 8.69 (dd,  $J = 4.4$ , 1.2 Hz)

$^{13}\text{C}$  RMN (DMSO- $d_6$ , 150 MHz): 17.6, 56.8, 66.6, 103.4, 120.4, 121.0, 125.1, 129.3, 129.7, 130.2, 133.9, 137.7, 149.7, 154.9, 156.0, 157.3, 166.1, 167.3

#### Ejemplo 10. Actividad antitumoral de los nuevos compuestos generados.

Para los ensayos de actividad antitumoral se utilizaron varias líneas celulares tumorales: A549 (pulmón), HT29 (colon), MDA-MB-231 (mama) y HCT116 (colon), así como la línea control no tumoral de fibroblastos NIH373.

Las células se repartieron en placas de 96 pocillos a razón de 5.000 células/pocillo en 100 mL por pocillo de medio DMEM suplementado con 10% FBS y 2 mM glutamina. A continuación se preincubaron durante 24 h sin fármaco ensayo para que las células se adhieran a la placa. Al día siguiente se añadieron las concentraciones adecuadas de los distintos compuestos diluidos en medio DMEM y siempre en un volumen de 10 mL. Como controles se añadió medio sin compuesto para cuantificar la viabilidad de la línea celular y SDS al 0.1% como control de 100% de muerte. Transcurridas 24 h de incubación en presencia de cada compuesto se añadieron a cada pocillo 10 mL del reactivo WST-8 (Cell Counting Kit-8, Dojindo, Probiol, Alemania) y se incubaron durante 2 h a 37°C tras lo que se determinó la absorbancia a 450 nm en un lector ELISA (ELx800, BioTek). La actividad de los distintos compuestos se muestra en la tabla 3.

**Tabla 9:** Actividad antitumoral (IC50) de los distintos compuestos frente a distintas líneas celulares.

Línea celular	➤	NIH 373	A549	HT29	MDA-MB-231	HCT116
Compuesto	▼	Fibroblastos	Pulmón	Colon	Pecho	Colon
Colismicina A		>100 $\mu\text{M}$	5 $\mu\text{M}$	>100 $\mu\text{M}$	>100 $\mu\text{M}$	0,6 $\mu\text{M}$
fórmula VI		>100 $\mu\text{M}$	100 $\mu\text{M}$	>100 $\mu\text{M}$	>100 $\mu\text{M}$	1 $\mu\text{M}$
fórmula VII		>100 $\mu\text{M}$	2,5 $\mu\text{M}$	>100 $\mu\text{M}$	>100 $\mu\text{M}$	0.7 $\mu\text{M}$
fórmula II		N/D	90 $\mu\text{M}$	100 $\mu\text{M}$	100 $\mu\text{M}$	>100 $\mu\text{M}$
fórmula XI		N/D	>100 $\mu\text{M}$	100 $\mu\text{M}$	>100 $\mu\text{M}$	>100 $\mu\text{M}$

Algunos compuestos ensayados frente a determinadas líneas celulares mostraron valores de IC50 de orden micromolar, como es el caso de los compuestos de fórmula VI y de fórmula VII frente a HCT116. Sin embargo, algunos compuestos presentaron valores menos elevados de citotoxicidad, de un orden superior a 100 micromolar frente a determinadas líneas celulares.

#### Ejemplo 11. Actividad neuroprotectora de los nuevos compuestos generados.

Los ensayos de actividad neuroprotectora se llevaron a cabo utilizando el modelo del pez cebra. El manejo de los embriones, su mantenimiento y cría se realizó siguiendo procedimientos estandarizados (Westerfield, 1993; Kimmel *et al.*, 1995). Una vez recolectados fueron mantenidos en agua de embriones (135  $\mu\text{M}$   $\text{CaCl}_2$ , 623  $\mu\text{M}$   $\text{MgSO}_4$ , 1.14mM  $\text{NaHCO}_3$ , 402.41  $\mu\text{M}$  KCl, 10.7 mM NaCl, 348.5  $\mu\text{M}$   $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ , Scharlau Chemie, SA, 08016 Barcelona, Spain).

Los embriones fueron tratados tras 3 días postfertilización durante 24 horas con 10  $\mu\text{M}$  ácido retinoico (CAS #302-79-4, Sigma-Aldrich) (Selderslaghs *et al.*, 2009) en presencia de cada uno de los compuestos neuroprotectores a ensayar y utilizando como control positivo ácido  $\alpha$ -lipoico (CAS 1077-28-7, Sigma-Aldrich) a 1  $\mu\text{M}$ . Dimetil sulfóxido (DMSO; 1%) fue incorporado a los ensayos al ser el solvente en el que iban disueltos los compuestos a ensayar. El ácido retinoico y los compuestos potencialmente neuroprotectores fueron añadidos al agua de embriones descrito anteriormente, utilizando 15 embriones de pez zebra. Al final del tratamiento embriones de 4 días postfertilización fueron lavados tres veces con agua y sumergidos en una solución de 2  $\mu\text{g/ml}$  de cloruro de acridinio (Sigma-Aldrich) durante 60 min. A continuación se lavaron abundantemente tres veces (5 min por lavado) y se anestesiaron con anestésico MS222 (metanosulfonato de tricaina). Finalmente se posicionaron adecuadamente para su observación por microscopía de fluorescencia. Esta se llevó a cabo utilizando un microscopio de fluorescencia Leica DMIL LED (Leica Microsistemas, S.A., Barcelona), equipado con un filtro verde FITC (excitación: 488 nm, emisión: 515 nm), y una cámara digital EC3 (Leica Microsistemas, S.A., Barcelona). Las imágenes fueron procesadas con LAS EZ V1.6.0 (Leica Microsistemas, S.A., Barcelona) y Adobe Photoshop 7.0 software (Adobe, San Jose, CA). El análisis de partículas (Wasabi, Hamamatsu Photonics Germany GmbH) fue usado para cuantificar la fluorescencia. Para comparar la eficacia entre los distintos compuestos se utilizó el análisis de la varianza (ANOVA) para detectar si había diferencias significativas entre compuestos en los experimentos. El t-test de Dunnett fue usado para identificar compuestos que exhibían diferencias significativas en comparación al control (n=3). Todos los cálculos fueron hechos usando el software estadístico GraphPad Prism (GraphPad Software, Inc., San Diego, CA). Un P-value menor de 0,05 fue considerado estadísticamente significativo.

Los neuroprotectores fueron ensayados a la concentración NOEC (Non-observed effect concentration). Esta NOEC, tras un estudio previo, fue determinada como 1  $\mu\text{M}$  para todos los compuestos ensayados. El control, sin tratamiento con ácido retinoico mostró pocas o ninguna célula apoptótica (FIG. 14A). Por el contrario, los embriones tratados con ácido retinoico mostraron un aumento significativo en las apoptosis cerebrales (FIG. 14B). Las imágenes que se muestran en las FIG 14C, D y E son ejemplos del co-tratamiento con ácido retinoico y cada uno de los compuestos ensayados, incluyendo ácido lipoico como control positivo (Packer *et al.*, 1997. *Free Radic Biol Med.*;22(1-2):359-78.) (FIG. 20F) y usando la concentración NOEC. Los compuestos de fórmula VI y VII mostraron protección frente al ácido retinoico, siendo claramente significativo el efecto del compuesto de fórmula VII (FIG. 14E).

### 30 Microorganismos depositados.

Los siguientes microorganismos fueron depositados en las fechas abajo indicadas en la Colección Española de Cultivos Tipo (CECT), cumpliendo con el Tratado de Budapest:

Microorganismo	Número de Identificación	Fecha de depósito
<i>Streptomyces spp.</i> CS40	7757	18.06.2010
<i>Streptomyces spp.</i> CLM-L	7754	18.06.2010
<i>Streptomyces spp.</i> CLM-A	7755	18.06.2010
<i>Streptomyces spp.</i> CLM-M2	7756	18.06.2010
<i>Streptomyces spp.</i> CLM-G	7861	16.12.2010

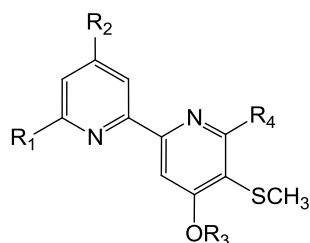


## BIBLIOGRAFÍA

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. y Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215, 403-10.
- 5 2. Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N. y Schoner, B. E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* 116, 43-9.
3. Bruntner, C. y Bormann, C. (1998). The *Streptomyces tendae* Tu901 L-lysine 2-aminotransferase catalyzes the initial reaction in nikkomycin D biosynthesis. *Eur J Biochem* 254, 347-55.
- 10 4. Cristalli, G., Franchetti, P., Grifantini, M. y Ripa, S. (1986). 2,2'-Bipyridyl-6-carboxamidoximes with potential antitumor and antimicrobial properties. *Farmaco [Sci]* 41, 499-507.
- 5 5. Devereux, J., Haeblerli, P. y Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12, 387-95.
- 6 6. Fernandez, E., Weissbach, U., Sanchez Reillo, C., Brana, A. F., Mendez, C., Rohr, J. y Salas, J. A. (1998). Identification of two genes from *Streptomyces argillaceus* encoding glycosyltransferases involved in transfer of a disaccharide during biosynthesis of the antitumor drug mithramycin. *J Bacteriol* 180, 4929-37.
- 15 7. Funk, A. y Divekar, P. V. (1959). Caerulomycin, a new antibiotic from *Streptomyces caeruleus* Baldacci. I. Production, isolation, assay, and biological properties. *Can J Microbiol* 5, 317-21.
- 8 8. Gomi, S., Amano, S., Sato, E., Miyadoh, S. y Kodama, Y. (1994). Novel antibiotics SF2738A, B and C, and their analogs produced by *Streptomyces* sp. *J Antibiot (Tokyo)* 47, 1385-94.
- 20 9. Hopwood, D., Bibb, M. y Chater, K. (1985). *Genetic Manipulation of Streptomyces: a Laboratory Manual*. The John Innes Foundation, Norwich, UK.
- 10 10. Kieser, T., Bibb, M., Chater, K., Butter, M. y Hopwood, D. (2000). *Practical Streptomyces genetics: a laboratory manual*. The John Innes Foundation, Norwich, UK.
- 25 11. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. y Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev Dyn* 203, 253-310.
- 3 12. Krogh, A., Larsson, B., von Heijne, G. y Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305, 567-80.
- 30 13. Martinez Gil, A., Usan Egea, P., Medina Padilla, M., GARCÍA PALOMERO, E., Pérez Baz, J., Fernández Chimeró, R. I., Fernández Medarde, A., Cañedo Hernández, L. M., Romero Millán, F., Castro Morena, A. et al. (2007). Use of collismycin and derivatives as oxidative stress inhibitors. Spain.
- 35 14. Menendez, N., Nur-e-Alam, M., Fischer, C., Brana, A. F., Salas, J. A., Rohr, J. y Mendez, C. (2006). Deoxysugar transfer during chromomycin A3 biosynthesis in *Streptomyces griseus* subsp. *griseus*: new derivatives with antitumor activity. *Appl Environ Microbiol* 72, 167-77.
- 40 15. Namwat, W., Kinoshita, H. y Nihira, T. (2002). Identification by heterologous expression and gene disruption of VisA as L-lysine 2-aminotransferase essential for virginiamycin S biosynthesis in *Streptomyces virginiae*. *J Bacteriol* 184, 4811-8.
- 45 16. Prado, L., Lombo, F., Brana, A. F., Mendez, C., Rohr, J. y Salas, J. A. (1999). Analysis of two chromosomal regions adjacent to genes for a type II polyketide synthase involved in the biosynthesis of the antitumor polyketide mithramycin in *Streptomyces argillaceus*. *Mol Gen Genet* 261, 216-25.
- 50 17. Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W. y Huson, D. H. (2005). Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* 33, 5799-808.
- 55 18. Sambrook, J., Fritsch, E. y Maniatis, T. (1989). *Molecular cloning: A laboratory manual 1-3: cold spring harbor laboratory press*.
- 60 19. Sanger, F., Nicklen, S. y Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74, 5463-7.
- 20 20. Selderslaghs, I. W., Van Rompay, A. R., De Coen, W. y Witters, H. E. (2009). Development of a screening assay to identify teratogenic and embryotoxic chemicals using the zebrafish embryo. *Reprod Toxicol* 28, 308-20.
- 25 21. Shindo, K., Yamagishi, Y., Okada, Y. y Kawai, H. (1994). Collismycins A and B, novel non-steroidal inhibitors of dexamethasone-glucocorticoid receptor binding. *J Antibiot (Tokyo)* 47, 1072-4.
- 30 22. Sun, Y., He, X., Liang, J., Zhou, X. y Deng, Z. (2009). Analysis of functions in plasmid pHZ1358 influencing its genetic and structural stability in *Streptomyces lividans* 1326. *Appl Microbiol Biotechnol* 82, 303-10.
- 35 23. Tae, H., Kong, E. B. y Park, K. (2007). ASMPKS: an analysis system for modular polyketide synthases. *BMC Bioinformatics* 8, 327.
- 40 24. Tsuge, N., Furihata, K., Shin-Ya, K., Hayakawa, Y. y Seto, H. (1999). Novel antibiotics pyrisulfoxin A and B produced by *Streptomyces californicus*. *J Antibiot (Tokyo)* 52, 505-7.
- 45 25. Vining, L. C., McInnes, A., McCulloch, A. W., Smith, D. G. y Walter, J. A. (1988). The biosynthesis of caerulomycins in *Streptomyces caeruleus*. Isolation of a new caerulomycin and incorporation of picolinic acid and glycerol into caerulomycin A. *Canadian Journal of Chemistry* 66(1), 191-194.
- 50 26. Westerfield, M. (1993). *The zebrafish book: a guide for the laboratory use of zebrafish (Brachydanio rerio)*: University of Oregon Press Eugene, OR.

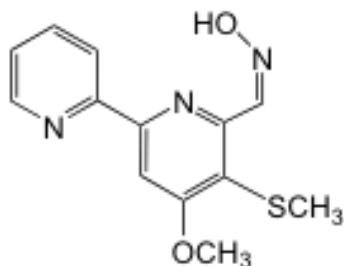
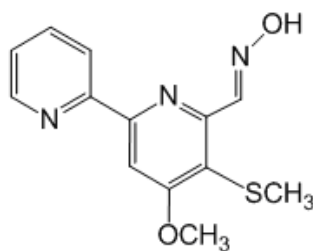
## REIVINDICACIONES

1. Molécula de ácido nucleico que consiste en al menos un fragmento de la SEQ ID NO: 1 capaz de codificar para al menos una de las secuencias SEQ ID NO: 2 a SEQ ID NO: 28.
2. Molécula de ácido nucleico, según la reivindicación 1, que consiste en la SEQ ID NO: 1.
3. Vector de expresión que comprende la molécula de ácido nucleico de las reivindicaciones 1 ó 2.
4. Vector de expresión, según la reivindicación 3, caracterizado por ser el cósmido cos1c3 o cos3b11.
5. Célula u organismo transgénico no humanos que comprende el vector de expresión de las reivindicaciones 3 ó 4.
6. Cepa de *Streptomyces spp.* con número de identificación CECT 7754, CECT 7755, CECT 7756, CECT 7757 o CECT 7861.
7. Compuesto de Fórmula (I):



(I)

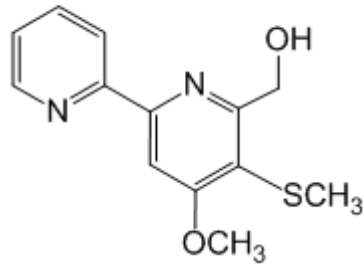
donde R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> y R<sub>4</sub> son, cada uno e independientemente, hidrógeno o un grupo protector; donde el grupo protector comprende un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo alquilo halogenado, un grupo alcoxialquilo, un grupo alquenilo, un grupo alquinilo, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo cetona, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos; exceptuando los compuestos con las siguientes fórmulas:



y

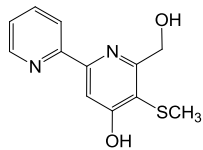
5

10

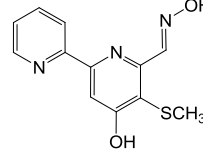


15

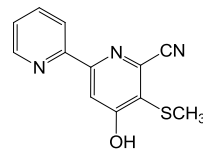
8. Compuesto, según la reivindicación 7, seleccionado entre los siguientes compuestos de Fórmula II a XVI:



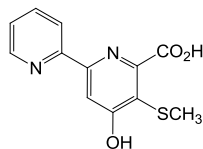
(II)



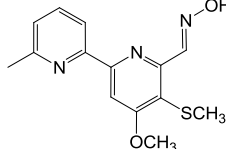
(III)



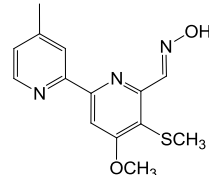
(IV)



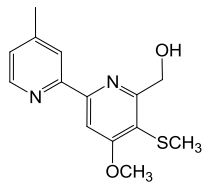
(V)



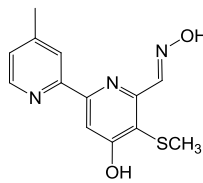
(VI)



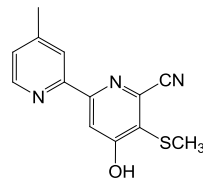
(VII)



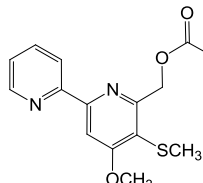
(VIII)



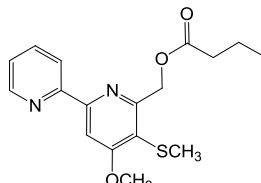
(IX)



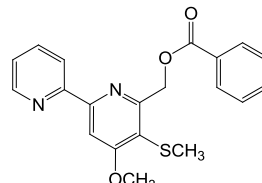
(X)



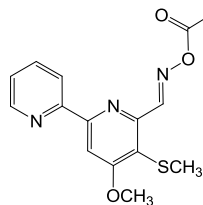
(XI)



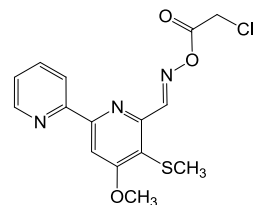
(XII)



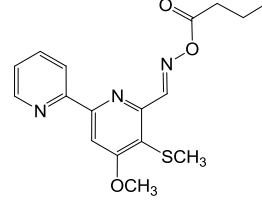
(XIII)



(XIV)

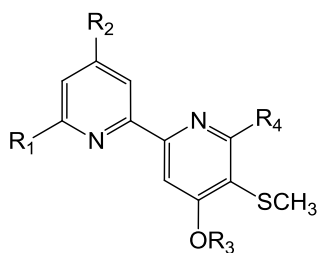


(XV)



(XVI)

9. Uso de un compuesto de Fórmula (I):



(I)

donde  $R_1$ ,  $R_2$ ,  $R_3$  y  $R_4$  son, cada uno e independientemente, hidrógeno o un grupo protector; donde el grupo protector comprende un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo alquilo halogenado, un grupo alcoxialquilo, un grupo alquenilo, un grupo alquinilo, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo cetona, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos; para la elaboración de una composición farmacéutica destinada al tratamiento del cáncer.

5

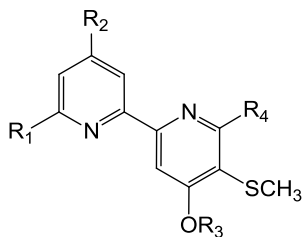
10

10. Uso, según la reivindicación 9, donde el tipo de cáncer se selecciona entre: cáncer de mama, cáncer de pulmón y cáncer de colon.

11. Uso de un compuesto según las reivindicación 7 u 8 para la elaboración de una composición farmacéutica destinada al tratamiento enfermedades neurodegenerativas.

15

12. Uso de un compuesto de Fórmula (I):



(I)

donde  $R_1$ ,  $R_2$ ,  $R_3$  y  $R_4$  son, cada uno e independientemente, hidrógeno o un grupo protector; donde el grupo protector comprende un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo alquilo halogenado, un grupo alcoxialquilo, un grupo alquenilo, un grupo alquinilo, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo cetona, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos; para la elaboración de una composición farmacéutica destinada al tratamiento de enfermedades infecciosas.

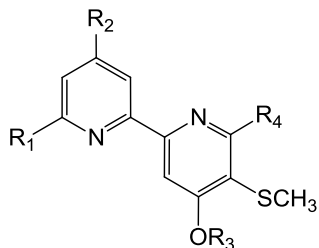
20

25

13. Composición farmacéutica que comprende al menos uno de los compuestos de las reivindicaciones 7 u 8 y al menos un excipiente farmacéuticamente aceptable.

30

14. Compuesto de Fórmula (I):



(I)

35

donde  $R_1$ ,  $R_2$ ,  $R_3$  y  $R_4$  son, cada uno e independientemente, hidrógeno o un grupo protector; donde el grupo protector comprende un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo

alquilo halogenado, un grupo alcoxilquilo, un grupo alquenoilo, un grupo alquinilo, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo cetona, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos; para ser usado en el tratamiento del cáncer.

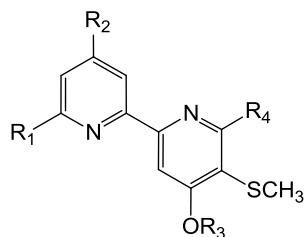
5

15. Compuesto, según la reivindicación 14, para ser usado en el tratamiento de un tipo de cáncer seleccionado entre: cáncer de mama, cáncer de pulmón y cáncer de colon.

10

16. Compuesto según las reivindicación 7 u 8 para ser usado en el tratamiento de enfermedades neurodegenerativas.

17. Compuesto de Fórmula (I):



(I)

15

donde  $R_1$ ,  $R_2$ ,  $R_3$  y  $R_4$  son, cada uno e independientemente, hidrógeno o un grupo protector; donde el grupo protector puede consistir en un grupo alquilo, un grupo cicloalquilo, un grupo cicloalquilo heterocíclico, un grupo hidroxialquílico, un grupo alquilo halogenado, un grupo alcoxilquilo, un grupo alquenoilo, un grupo alquinilo, un grupo arilo, un grupo arilo heterocíclico, un grupo alquilarilo, un grupo éster, un grupo cetona, un grupo carbonato, un grupo ácido carboxílico, un grupo aldehído, un grupo cetona, un grupo oxima, un grupo nitrilo, un grupo uretano, un grupo sililo, un grupo sulfoxi o una combinación de ellos; para ser usado en el tratamiento de enfermedades infecciosas.

20

18. Procedimiento para la obtención de derivados acilados de colismicina A y/o colismicina C, que comprende reaccionar colismicina A y/o colismicina C con un agente acilante en presencia de una enzima hidrolasa.

25

FIG. 1

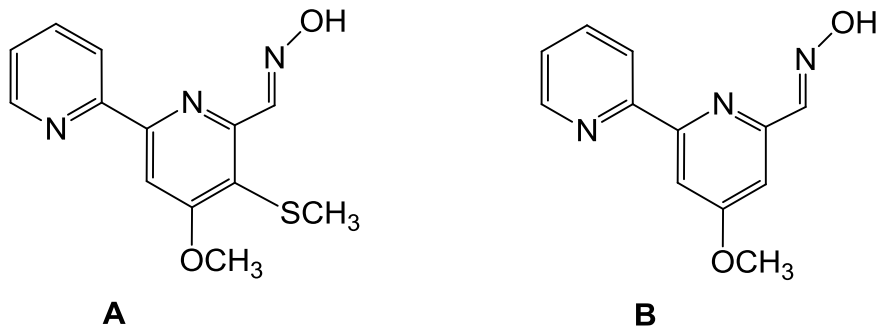


FIG. 2

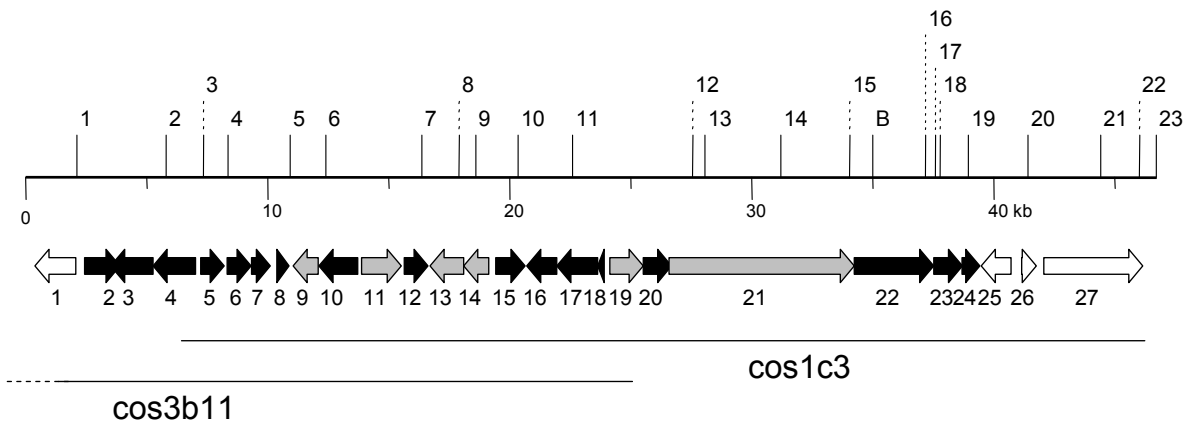


FIG. 3

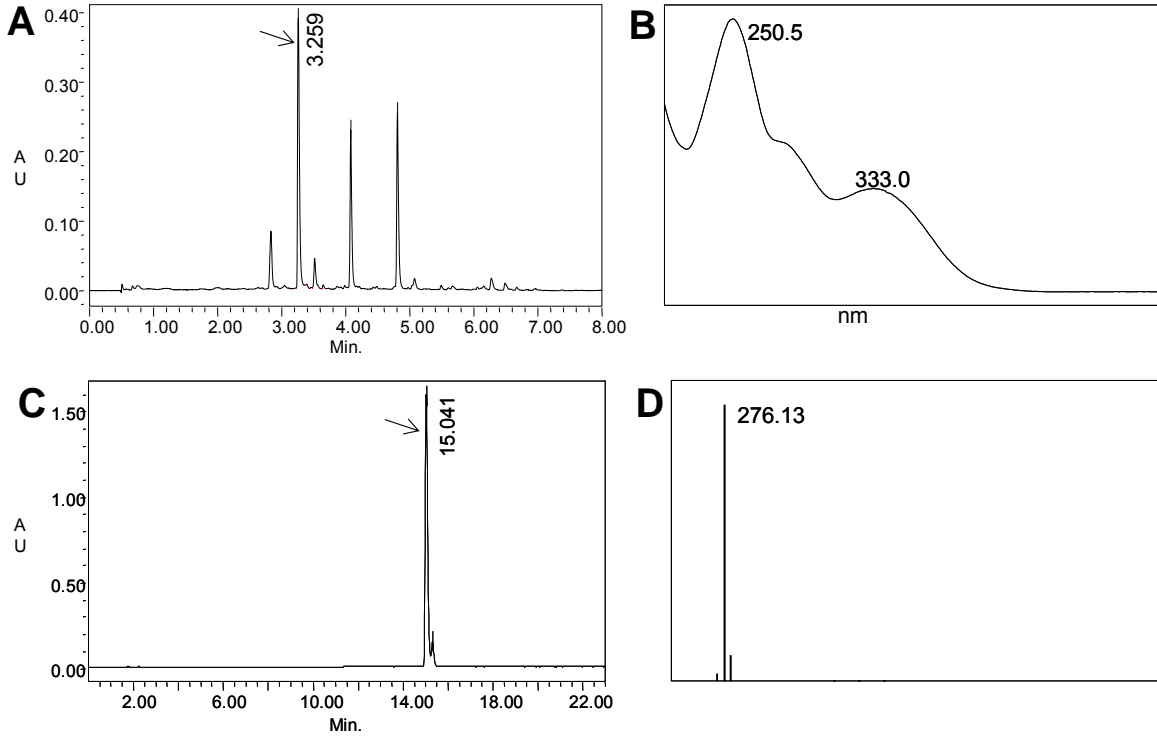




FIG.4

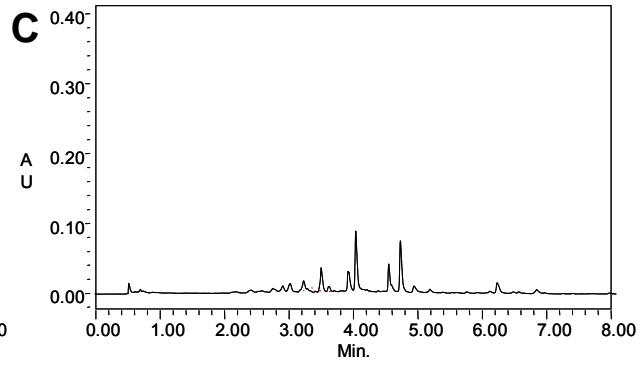
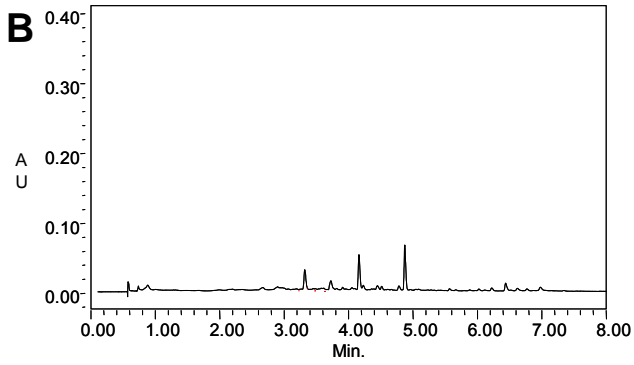
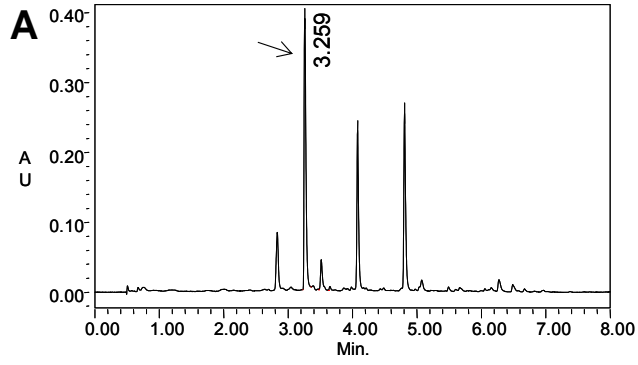


FIG. 5

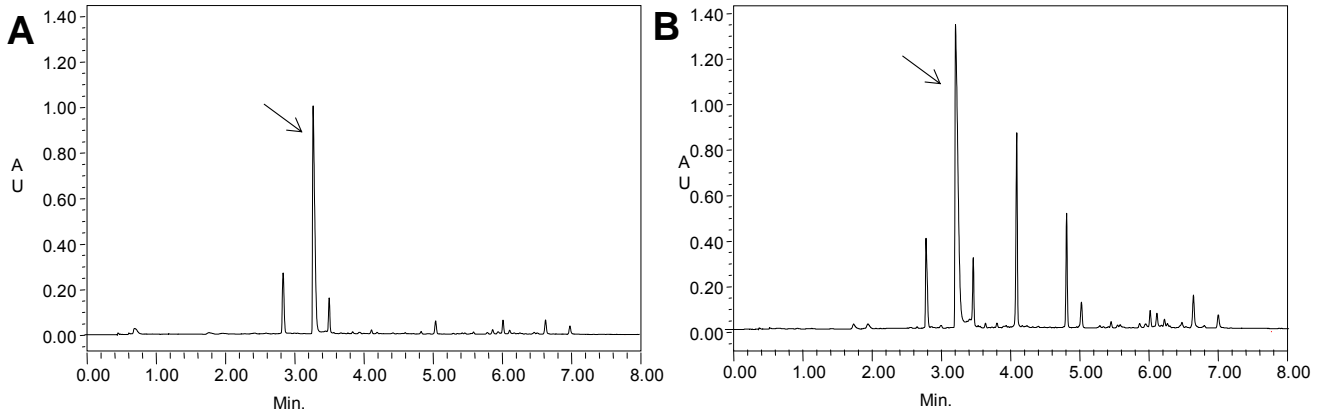


FIG. 6

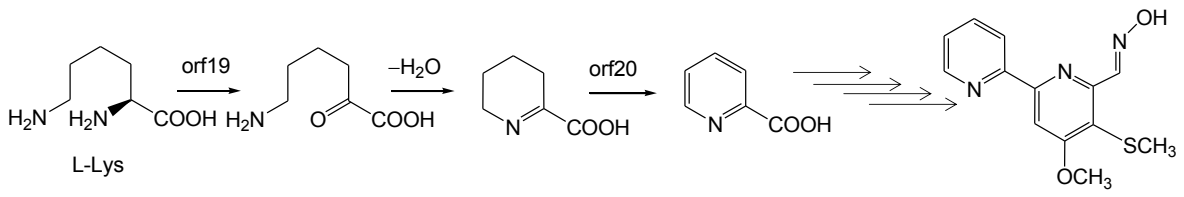


FIG. 7

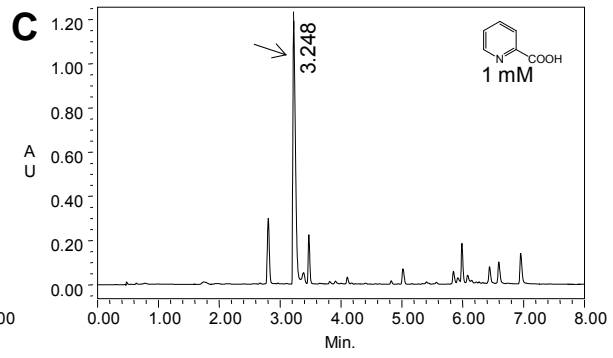
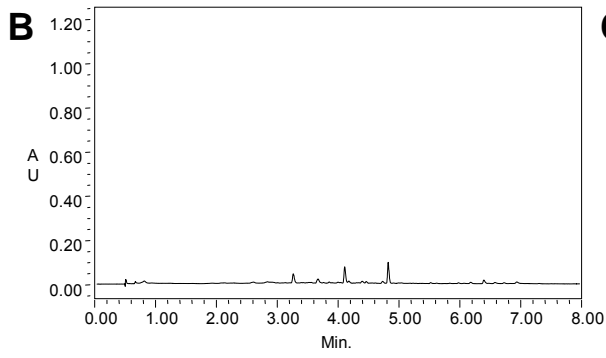
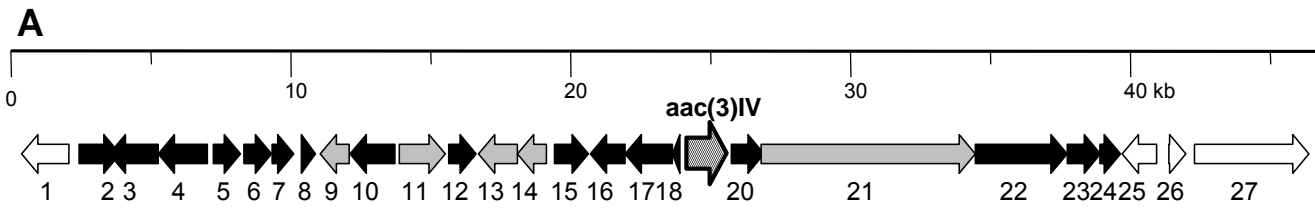


FIG. 8

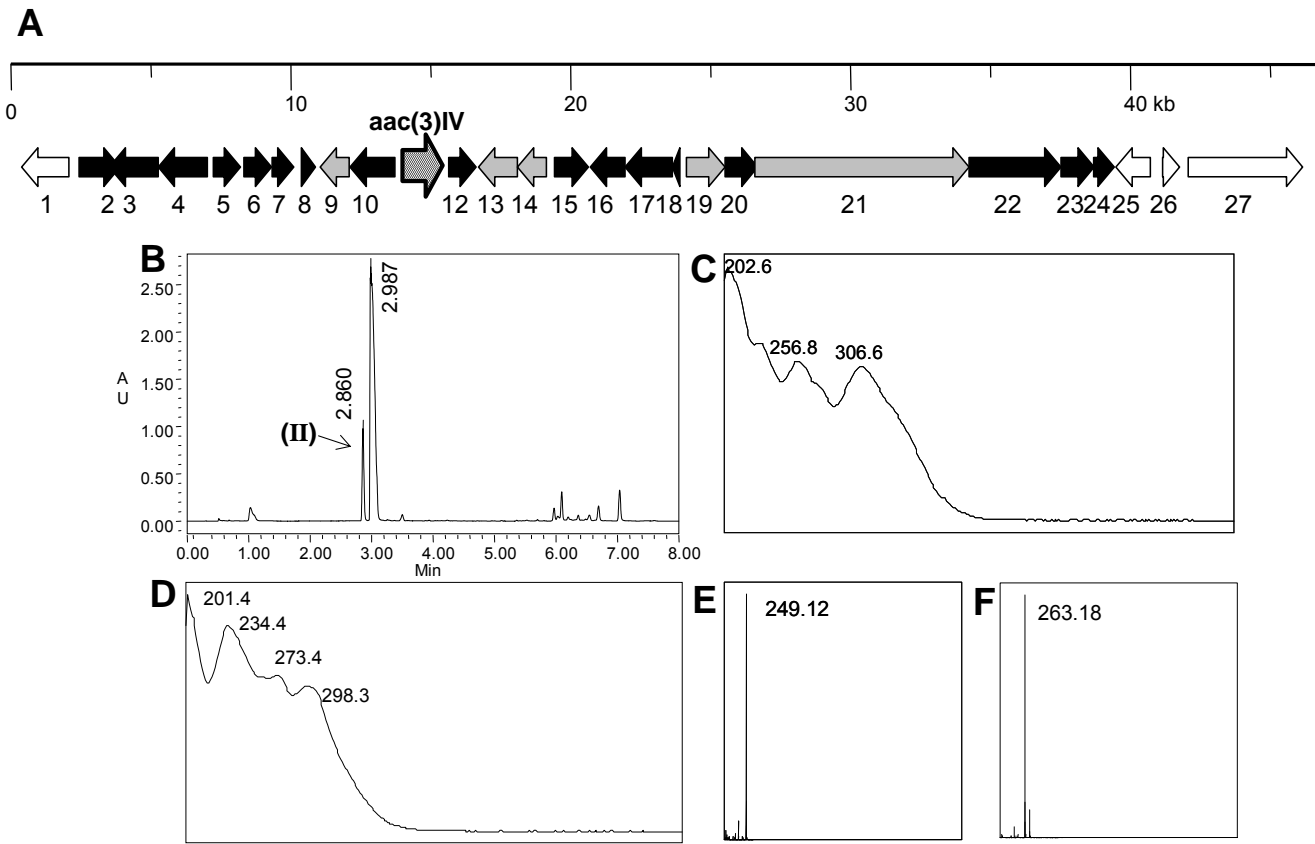
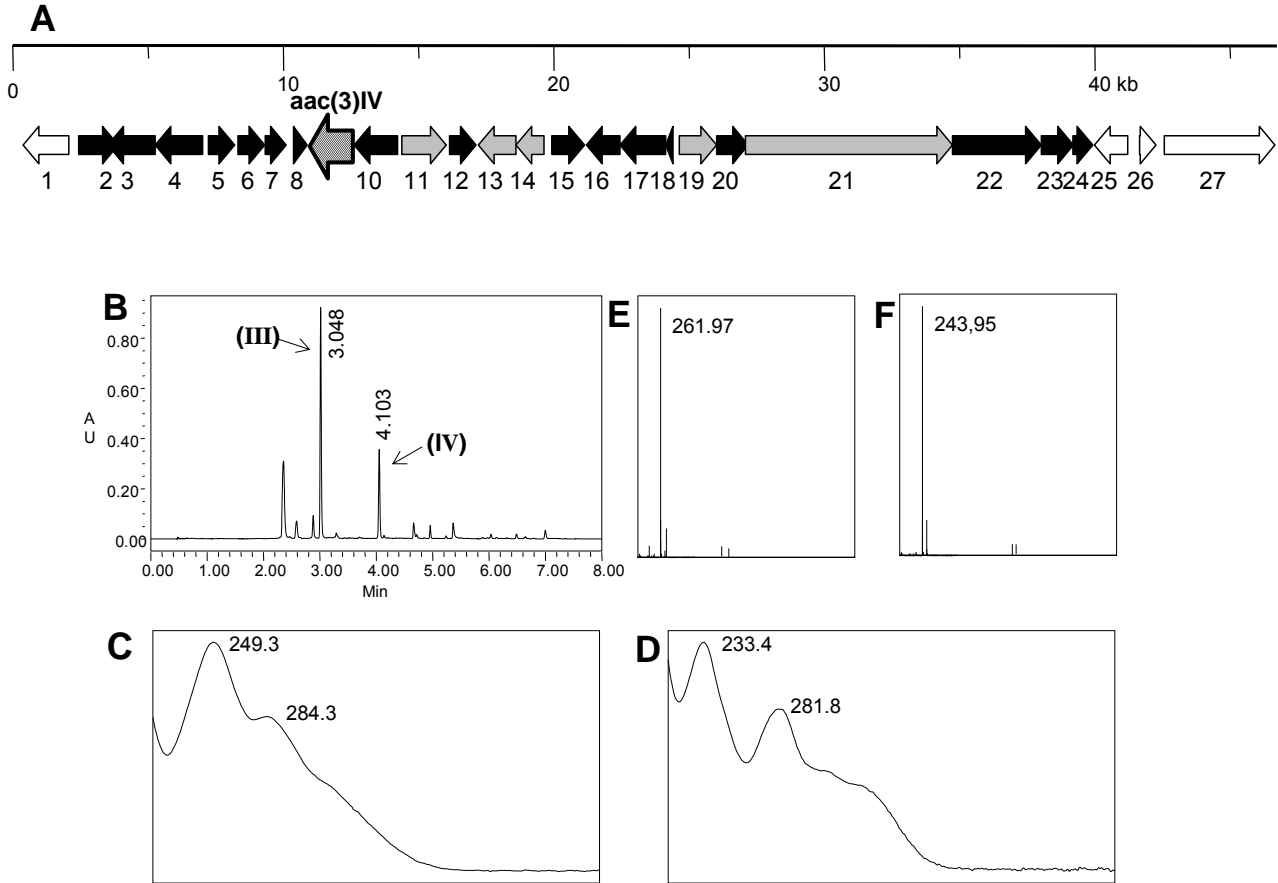


FIG. 9



**A**

**FIG. 10**

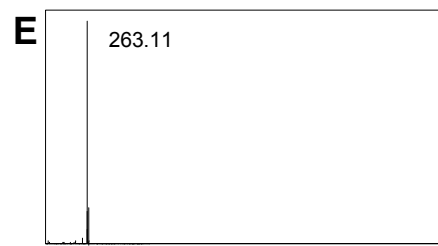
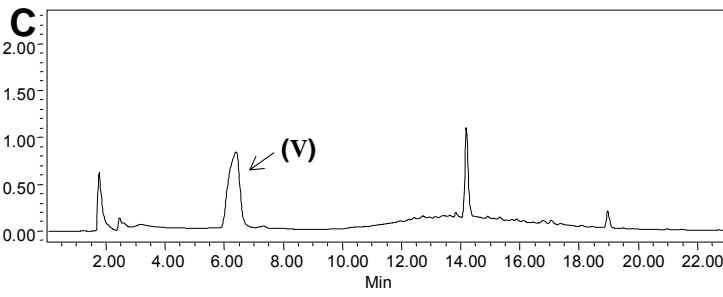
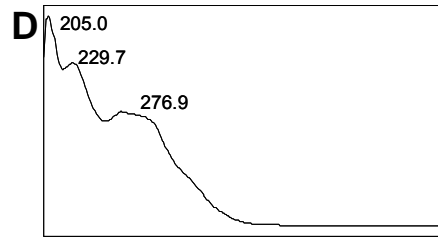
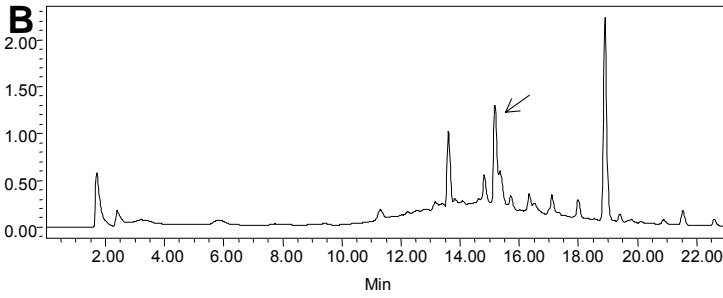
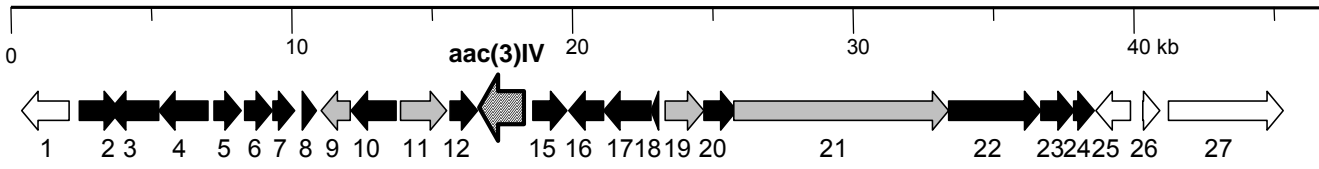


FIG. 11

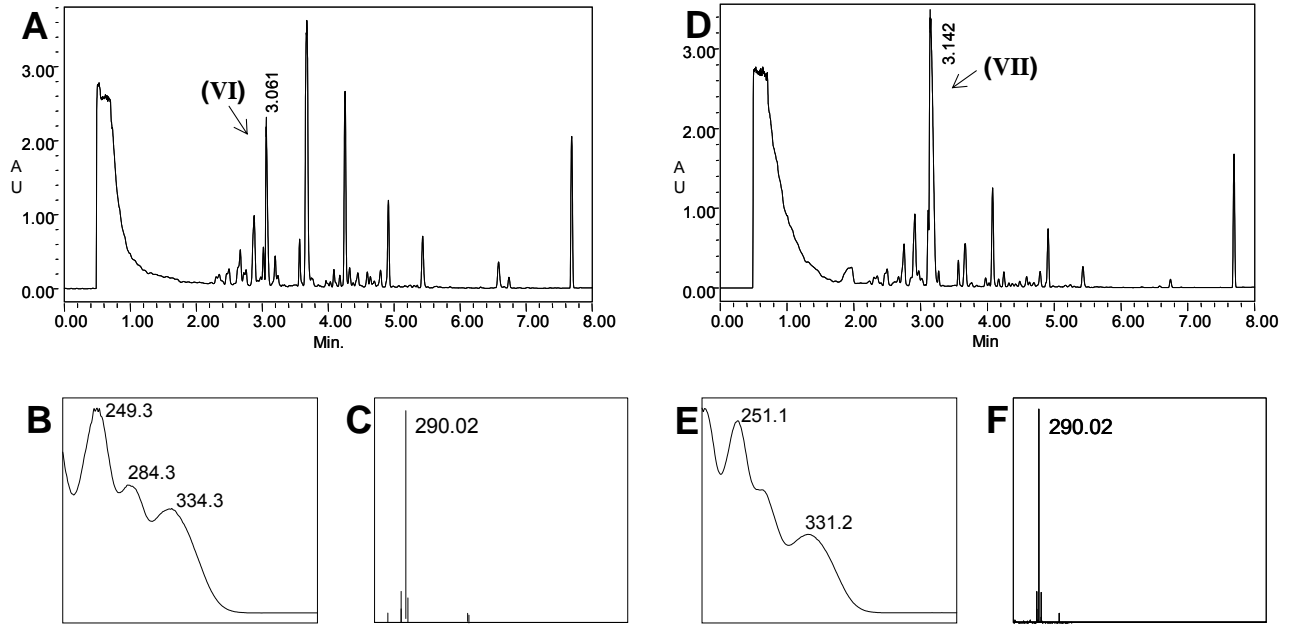




FIG. 12

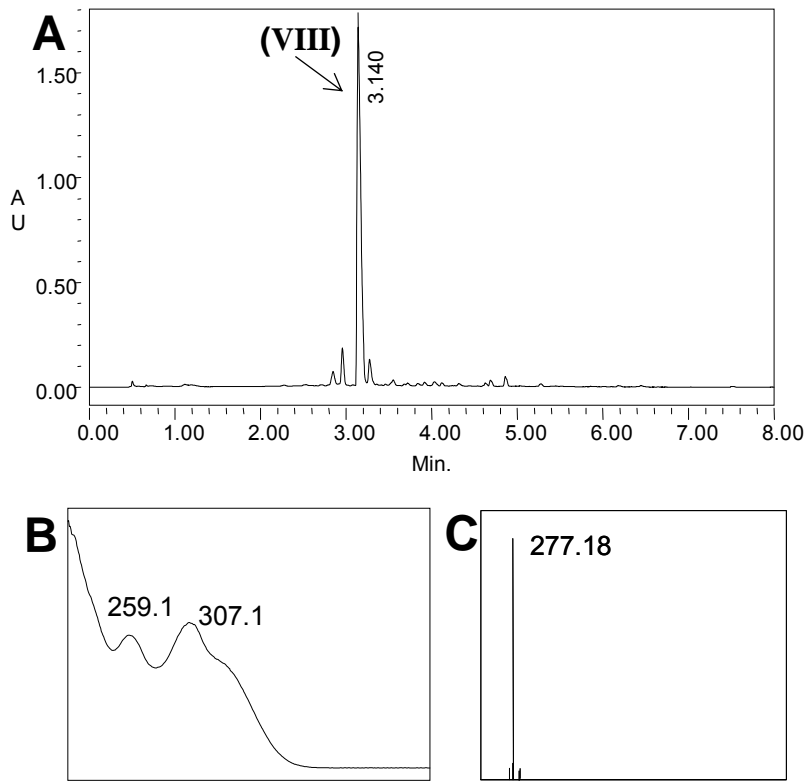


FIG. 13

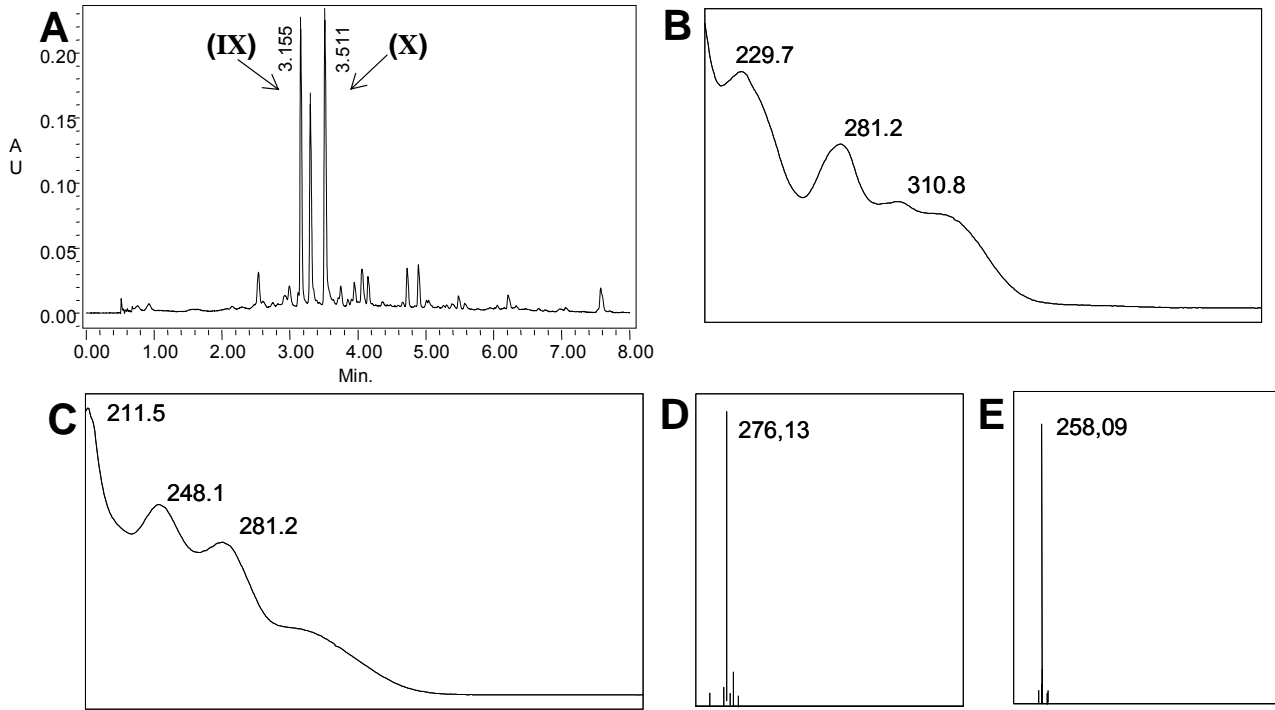
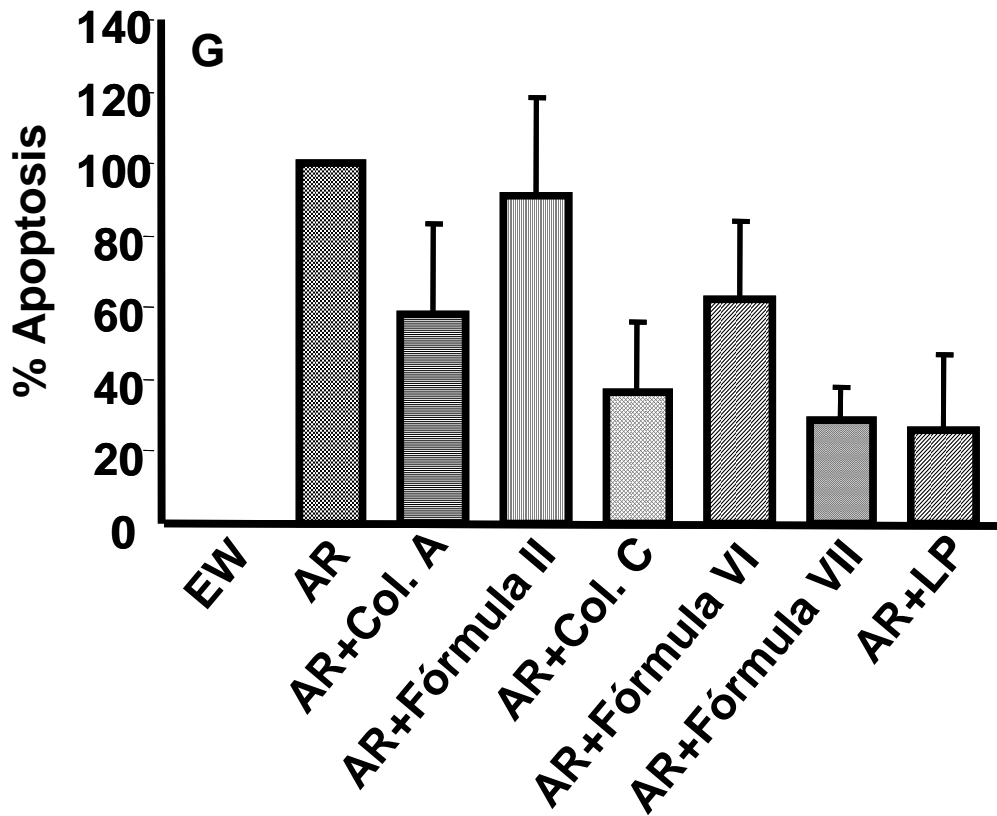
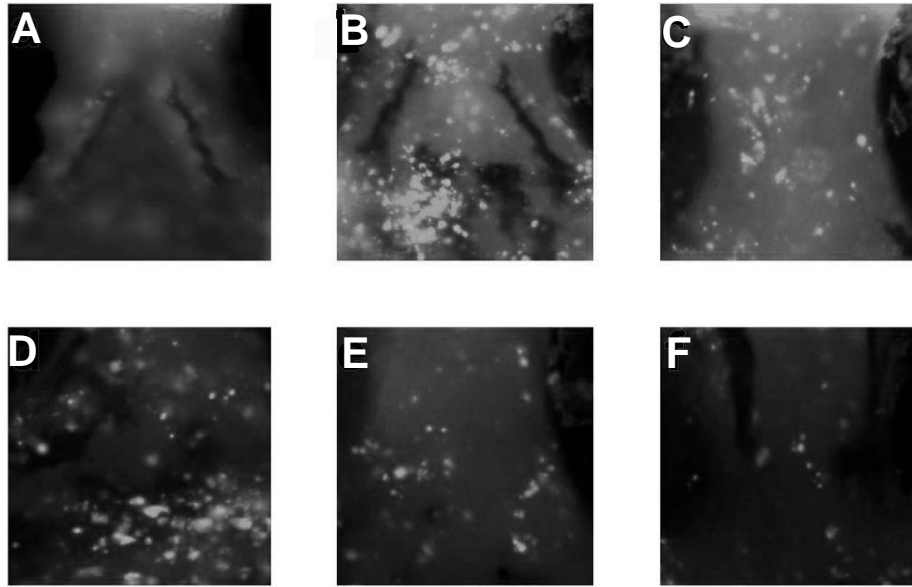


FIG. 14



LISTADO DE SECUENCIAS

<110> UNIVERSIDAD DE OVIEDO  
 <110> ENTRECHEM S.L

5 <120> DERIVADOS DE COLISMICINA

<130> P-03987

<160> 56

10 <170> PatentIn version 3.5

<210> 1

<211> 46672

15 <212> DNA

<213> Streptomyces ssp.

<220>

<223> Cepa CS40

20 <400> 1

gatatcgagc gcgccgaggc cgctcgacggc ggcgaggggtg ttctccggcg cgtagcccga 60  
 cgccccccgg tgggccacga cgaccgggtt cgcccgtgc gcggagacct ggtccggggc 120  
 25 ctggggccgg tcggcgggca gcagcagggc gccggcgccg agcagggcgg cggccgtggt 180  
 ggaggcgggt gcggtgcgga cgaacacggg gactcctcac gtcgggagtg tgcggacacg 240  
 30 cagagagtga cagcggcccc ccaacgcccg gcggggcggc gatggcccgc aatgaacgg 300  
 aacctcggt ccagaggacc accaccgagg cacggcctgc gaccgtgccg ataaaatgcg 360  
 35 ccttctttg tctgtctgcc gggactcgcc ccttggggac cctcctcgac cccgggcatt 420  
 ctccaagggc gcacgcgcat gcaaggaaca gtcgacggat tcacgtacgg cctggtgact 480  
 ccggtcgcgg cgttctcat ggccgcgctg ggagccgccc tcgggctccg ctgcaccacg 540  
 40 cggtccttc acacgacacg ctctgtcgg atcaggctcg accggtcgtc cggccgggtc 600  
 agcatgagct ccagcatcac gtagttcgtc agcaggatgt ccggagggtt cttgcgcagc 660  
 45 tccttgccc gctcctcgga ctctggccg gtgtaacggg cgaaggtgac cggctcctgc 720  
 cccgcgccga aaccgtggcg caggtactc tccagctcca tcagctcga gttggccagg 780  
 gcgttcatcg ggtagacgac gatcgcgagg acccggcccc cctcgttacg cctcccgtg 840  
 50 ccgttcctc cggcgcgag gtcgcccgc tgacgctcct tcaggacgtg gttcacgac 900  
 ggcacgatgt acgccaggga cttgccgaa cccgtgccg tggcagcac gtaggaatcc 960  
 ccggcctgag cggcttcgat gcctgcccg tgatggagat ggaaggtcag cggtcgccc 1020  
 55 tcggggcgtc gcgaactctc cttctgccc gcctggaaga tctccgcga cttcgggtgc 1080  
 agcagtcct cctggacgag gtcggtgacc ttgcctccgt cggcgaagaa ggggttcagg 1140  
 60 gagagccagg ggtcgggcca ctgcgactc gccccaggc cgtcctgac gaagccctc 1200  
 atccgggct cgcggaacac tgctgcgctc ttggtgaagc tctctactc gttgatcagg 1260  
 65 tcggcatgga taccgaagac gtccatggtc tcgggcaggc gcggttcct gcgcggcacc 1320

# ES 2 397 885 B1

ggcacggggg cggccggggg gcgctccggg agcaggggtcc gcagccgcgc caccgggtcc 1380  
atgccaccc agtgcggagc cagcagggcg gccgtgtcgt ccgcgccctg cgcgagcgg 1440  
5 accagcacgt catggagcag ggccgcgttc tggggccgaa gctcggcctt cttctccagc 1500  
agacggtcca ccaggcggac caactcgacg gggaggtccg gacgtacgag cgtcaggagg 1560  
ggaggtggg agtgcgtggg ctgttcggag agttcgtacg gggtcctggc ggagaacggc 1620  
10 ggctgtccga ccagcagttc gtacagaatg cagccgagcg catagaggtc cgcgaggccg 1680  
gagacgtgtt cggcacggaa ctgctcgggt gccatgtagc gagccgtgcc gacgctgacg 1740  
15 cccgtgctcg tcagcccgct ctgatcgggg tcgtcgacga tcgagcccat gccgaagtgc 1800  
aggatcttga cggtgccgtt gctggtcagc atgacattgg cgggtttcag gtcacggtgg 1860  
acgaccccgg ccgtgtgccc ggctgtcaga ccggcggcga tctgcgcgcc gatggcggcg 1920  
20 acccaggaga cggggagctg gggttcctcg tcgatgaggt ccgcgagagg gtgaccgtcc 1980  
agcagctcca tggccagata gggcagaccg ctgccgctg cgtcccgtc gacgccccg 2040  
25 tcgatcagcc gggtagggtt cgggtggcct tgggagagca tgcgcatgat ccgacactca 2100  
cggcggaaac ggtcgaactc cctggtgat cccgaggtgt cgacggcgat cccggtccgg 2160  
ccgcgagca ccgtctgac cgcgacctcc cgtgcgggg agtcgggtgc tgcccgaagg 2220  
30 tctcctgctc tgtcacctc gcccatgttc ccgcgcccga cgatccccgt gatccggaac 2280  
cgtgcgtcga ccgtctccag gctcactgcg cctccccgtc ttctccacct gttgcacaca 2340  
35 aggcagacac atcctcgggg agcgtatacc ttgagcagat caccaggatg acatggttct 2400  
gttgacatcg tcaacaagtg cttcgggtcg tctgcgccga gcgggaggag cgggttcggg 2460  
tatgcgccga tgggagcccc gtaatccca cgagtgggtg gcgtccatga ccctccgga 2520  
40 gttccgtagc ctcgggtccg ccctctgggg cggcgcgaag tcgtcgttcg agacgatgac 2580  
tgaactcgt gctcacgtct gtgccgtgc cgcactccg cgcaccgttc tcgtccgact 2640  
45 catggaccac gggcggtcc tgaagtctt cgctgctcg tcctgcgcg atctggacgg 2700  
gaactccgca tgtcggcgt gacattgcc aggcgggagg gccacccat cgccgggccc 2760  
50 cccgcccgg cacgtagact ccgtgaaatg agtgtgcgcc gaatcaccac tggccgccc 2820  
cgaaccaaac ccgccgagga gcggcgcgcc gacctcctgg acgcggccga gggagtgttc 2880  
gccgagcggg gaatcgtatc tgctcggatc gacgagatca cggagcgcgc ggggtcgcg 2940  
55 atcggcacct tctatctgca cttcagctcc aagcgcgaca tcatggccgc cgtacaggcc 3000  
cgtttcgtcg accggctggt cgagcggcag cagcggcgg cgcagaacct gcccgccgac 3060  
gactggatcg gacgcgtgga gcctggctc agcgacgcgg tccgcatcta tgtggaacat 3120  
60 gcccaactgc acgacgtgct ctacggccac acgccgatca acgccacgac ggagatccag 3180  
gtcagccccg agaacgcgca tctcagggcg ttccggcggc tcatcgcga gcgtcccgac 3240  
65 ccgccggcgg acgccccgaa ccccgcatg gcggtcttc tgatctacag cgctgttac 3300

## ES 2 397 885 B1

ggcggcacc atgcgctgct gcaccacgag gccgacgaca tcgaccggt gccgacgag 3360  
ctcatcgcc acataccga cctcgcccac cgctacctgc gctccgaac ctgaaccggg 3420  
5 ctgccgacgc cccgagttcg ggccgtcggc agcggaaagg gttcggtagc gcgcaccgcc 3480  
gtgaggcgca ccctgggac cggttcccc gggtcagtga ctgccgtgg cgatgcgcca 3540  
gtgccgccc cgttcgtact gggaactcag ccgcgcaaac gtcccacccg tgcgcagcag 3600  
10 ttctcgcgc gtgccgact ccacgacccg gccgcccgtc agcgagacca ccaggtccgc 3660  
ggcctccagg gtggcggggc ggtgggcgat cagatcacc gtacggtcgg ggtcggcgc 3720  
15 caggttcgcg atggcccggg tgatcgccc ttcgtctcc gggtcgagag cggagctggc 3780  
ctcgtcgacg aggacgatgc gcgcccgtt gagcatgcc cggcgatcg cgacgcgctg 3840  
gcgttcgcc cccgacagct gtgcccgcc ctcgccacc cgggtctgcc agccggcggg 3900  
20 cagccgctcg atcacctcgt cgagccgtc cgcggtggcg gcctcccga actcgtcca 3960  
ggtggcctcc ggacgggcca gccgagatt gtctcgtatg gtgtcgtcga acagatagac 4020  
25 gtctggaag acgatggcga tctcgtcag caggggtggtc gggtcgagct tgcgcacatc 4080  
gacgccgccc atgcgcagct cccccgagtc gatgtcgaag aagcgggcca tgagccgggt 4140  
cgccgtcgtc tcccggacc ccgacggtcc gaccagagcg gtggtcgtc cgcgccgaca 4200  
30 gcgcagggag acgttcgca gcgcgggggt gtcaccgcc ggataggtga acgtgacgtc 4260  
cgcgaactcg acctccgct cgtcgtaccg ccgcaccggc tcggcccga cccgagtcg 4320  
35 ctcagtgccg agcagctct cgatgcgccc gatctggtc cccaacgccc gcagggcgc 4380  
gatgagttcg atgagattgc cgagcggctc caggaaacgg gcggccagca cgagcagcac 4440  
gatcggcgtg gcggcgtcga tgcgtccgt gagcaccagc tcggcggtca gggcgagcac 4500  
40 gcgcagcaaa ccggccatga tcacaccgt gtagacgaag aagggtgcc gcgcgccc 4560  
catgccctcc cgttacgtct cgcggtggtc ctccagcgc gcgcgatcc ggggagtgcc 4620  
45 ctcggcgtc ccggcggccc gcaggaccgg ctggcctgg ccagctcga tcgccggct 4680  
ggcgacccc gccgcccgc gctccagttc gatctcggc atctcggaga tccggcggc 4740  
acggcgcagc gccagatagg cgaccagggc gatggcgcac agcagcagcg ccatccgcca 4800  
50 gtcgaccgtg aacgtcacga cgacgaccgt cgcgggcagc agcgtcaggg tgatcacggg 4860  
cccggcgtc gtcacggcca ggtgcgccc gctgccgacg tccggtcga ccgccgccc 4920  
55 gagccggcc ttgtgctcgg cggtgaacca gccagcgtc agcgtgctca catggtgcat 4980  
cagcccccgg cgcagctcgc ccgcgagcgc gcccccgcg gcgaagccca ccggggtggc 5040  
gacgatgctc agtaccgct acagcggcc gccaccgcg ccgagacca gccaggtcgt 5100  
60 ggccgcgccc atgtccggtt cggggcgcac gacggcgtcc aggatcggga tgagcagggc 5160  
gagcaacaac ccctggagca ccgctcgc gcggtgagc agccacagcc gggccagcag 5220  
65 ttccgattc ggccagatc ggtacagcg gcggtcacg cgagcaccct ttctgtgcg 5280

# ES 2 397 885 B1

gcccacatgc gggcgtacag cccgtcgcgc gcgaggagt cttcgtggcg gccctgttcg 5340  
accagccgcc cgtttccag tacgaggatc tggcgggcgc ccgcatggtg gtgcagccgg 5400  
5 tggcgatca cgatcaccgt cttgccacc gccaggtgg ccagggttc ctggaccgcc 5460  
gattcgtgt cggcgtcag ggagccgtc gcctctcca ggacgacgat cggcgtccg 5520  
gacaggatcg cccgcgcgat cgtcagccgc tggcgtccc cgccggacag actgcccgcg 5580  
10 cccgcgtcga gcaggtgtc gtaccctcg gccatgctct cgatcacgtc gtggatgtg 5640  
gcggcggcgc ccgctcgcg gacctcggcg tcgctcgcgc cgggcccgcg gatgcggatg 5700  
15 ttctcccca cgctgtcgcg cagcaggcgc acgtcctgga agaccaggga catcgaccgc 5760  
agcagcttcg ccgacgggat cgagcgcacg tccaccccgc cgatctgat ggatccccg 5820  
gacacttcgt agaagcgcgg cagcaggcgc gccagggtcg tcttgccgc gccggacgga 5880  
20 ccgacgatcg cgtcaccgc gcccggtgga cagacggcgc tgatgtctc gaccgcgtg 5940  
gtcaccctcg ctgaggagaa ggagaccgc tcgaactcga cgccgtggcc gacgggctcc 6000  
25 tgcggtccc gcggtcggg cagcggcga cggcgagca gctcctgat gttggccgcg 6060  
cccatccggc cctgtcgcac tccctggcgc gcggtcacgg cggggaggat cgagttgggc 6120  
agcccgatgg cgacgaccag gaaccacc aggtcggcga cgccgagcga gccgcggtcg 6180  
30 acgaggacca ggccgacggc cgccacggcg gtgaggacgg ccattcctga gccgaacagc 6240  
cgggagcccg cggagctgt gcggacctc gcgaccagg ccgcaacgc ctgggtgtgc 6300  
35 tcgtccacgg cgtcgtcga gcggcgcagg acacggccgc ccgtgccga cgtttgacc 6360  
acctgatgc cgtcggcga ctcgacgctg gcggcgtga tccggcgtc gccgaggacc 6420  
agccggtta tatgggtcgt catggaacgc atcgagacgc ggaagaagat gccatcagg 6480  
40 agcaggacgc ccagtaccac gaaggccatg cggacatcga cgaggagcag atagcccgtc 6540  
gcgaccgtca tgacggtgac ggccccgacc aactggcca gggcgtcgc gatgacctc 6600  
45 tgcattctc ccagatgcc cgtcatgcc cgttcacct tgcccagcc cacggcccgg 6660  
aaccagcca gcggcagct cgcgaggtg gtcacgatc gcacgcgcac gtcgtcagg 6720  
50 atgcccgct cggcgtagt cccacgcgg gaggactgca cgaacaggac gagccagacg 6780  
cccgcgccg cgccccgac accaccag gtccagatgg ctgtgtgta gcccggggg 6840  
tcgtcagca tgagccgcg gatctcggc acggcgtat agggcacgat gcccgccc 6900  
55 gcgctcagg acgaaaggac gcacagatg acgagatgg tccgatcgg gcgagcagg 6960  
cgcattaac gcggggccc tccagcccc ggagtctgc atgtgatct atcattagg 7020  
gtaccctaac atggcgactc aagctcattc aataacgtg gctcggccg gagatgcctc 7080  
60 cgcggcgcga tgtccgatg ccctgagagc cttaccgaa ggaactcct cttcatgtc 7140  
cagccgcacc acgaccgcc gtaacaagc tctcgtcgc cacgcgccg tccgaagcgg 7200  
65 cccggggcgc cctcgtcag tgcctgatc gcgctcggc ccgtcgggc gaccgggtgc 7260

# ES 2 397 885 B1

ggaggcggcg acgccgacgc acgccagaag ggcgagaaga agacgatcgc cttcgaatcc 7320  
tgtagtcga cggctgagct cgaccggatc ccgaagcgcg tggcgatcac cacggacgcc 7380  
5 atcgccgaca cttgtttga gctcggcgtc ggcgaccgga tcgtcgcaa gacccgcggc 7440  
gagtccgcc cgcacccga gctgaaggag cgcctcggcg ccctgcccag cctcgggaca 7500  
10 cgcaatccga gcgtcgaggc gctggtcggc gccaaagccg atctcctgat caccgatcag 7560  
gtcgagaagg tctccggcaa gctgggcagc ccgagcatcg ccgagctgga gcggctgggc 7620  
atgccacct atgtgtagg cggcggctgt ggcggcgacc tgagcgagga cacctccggc 7680  
15 ctgaaagcgc tcgacggaga catccggcag ctcggcacgg tcttcggcgt ggaggcccgc 7740  
gcccgaagc ttgccgaaa gctgaacggc agcctcgacg acgtacgacg gcagaccgcc 7800  
20 caggagcccc ggaccaaggt cgccaagctc tcgaggctc cggggcagct ctacgtgacc 7860  
tccggggcc tctccgacga cgtgatcag cggggggcg ggacgaacgt ctcgcccag 7920  
ctgcccggcc agttcggccc ggtcagcccc gaacagatcg tcgcgcgca tccgagctc 7980  
25 atcatcgtc acaactcac ggcgacggcg gccggtgaga acgaggcagat ggcgtacctc 8040  
aagcgacgt tccccaccgt cgaggcggc aagaagcaac ggtcctcgt catcgacgcg 8100  
30 gccaaatccg gggcgcgcg cagtacgagg ccggtggagg gcgtcgtcga gatagcgcgc 8160  
ttctgcacc cgagcagtc ccgtcccag tagggctcgc gatgccgaag caggaaacgc 8220  
ctgctaccg caccggcag gaggaacgcc cgcgaggcgc ccgccggccg actcccgtac 8280  
35 gcggaagccg acatgccgcc ctcccgtcgt tctgcccgt gctgtcgtc gtcctcctc 8340  
gcgcggtctg cctcggcgtg tcggctggat ccgtcggcgt gcggccggcg accgtcgtca 8400  
40 aggtcatcgc cgatcacctg atcggcgtcg ggacggccgc cgccgtcatg gacgatcaga 8460  
tcgtctgaa tctcggctg ccgcgtgtc tgctggccgc cgccgtcggc ggggggctc 8520  
ccgtcgtcg cgtcacctg caggcgaccg tccgcaacc gtcgcccag ccctacgtc 8580  
45 tcggtgtc cgccggccg gggctgatg cctccatgt gatcacctc ggatcggctg 8640  
cggtgggccg cctctgaca tcggccgcg cctcgtcgg ggcgctcgt gcgatggtg 8700  
50 cggctcctg cctctcgc cgggcggcc gtgtcatccc gagccggctg ctgctcggc 8760  
gggtgacgt ctctatctg ttcagcggg ccaccagctt cgtcatctc cgcagcggca 8820  
acgcgacgc ggcgactcgt gtgctgtt ggtcgtgg ctcctcgtc gaggcctcgt 8880  
55 ggtcgaacct gtccctgcc gcggcggccg tctggtcgt cggcgtgtac ctgatctcc 8940  
aggcccggc cctcaacgc ctcgcccgc gtgacgacgc ggcgctcctc ctcggcgtg 9000  
60 ccgtccaccg cgtgcgcat cgctcctc tcgtgcctc cctcctgacc ggcgtgctc 9060  
tcgctgtc cggcgggatc ggctcgtg gactcatct tcccctc gtcggctg 9120  
tcatggccc ggaccaccgt cggctgctc ccgtagccat gctcgtcggc gcggtctacc 9180  
65 ttgtcgtc cgtatcgtc tcgctgtc tcgtcgtc cgaggaact cccatcgga 9240



## ES 2 397 885 B1

tcgtaccgc cgtgctcggc gcgccggtct tctgtggct gctgcgacgc tcggaggcgt 9300  
catgagactc gccctgtccg gcgtcagtgt ccggatcgat tccagcccga tcgtcgcgga 9360  
5 gtgcgatctc accgtcgagg acggcgaacg cgtcggcctc gtcggcccca acggcagtgg 9420  
caagacctcg ctgttgcgta cggctaccg cgcgctcgac cccttcgcg gcaagtgcgc 9480  
10 gctgtccggc gacgacgtga cgacgctcag ccagcgggag gtcgcacgga gggccgcgct 9540  
cgtcgccag gacagcacc cggactcga ctcaccgtg gaggaggtcg tgccatggg 9600  
gcgcgggccg tggtccggg ccttcgagtc gaccaccggc ggcggggacg cgacgatcac 9660  
15 cgccgactg gagcgggtcc gtctggcggg ccaccgcgcc cgcaggctct ccacctgtc 9720  
ggcggtgag cggcagcggg ccctgggtgc ccgcgcgctc gccaggaga gccctctgt 9780  
20 cctcctcgac gagccgacca accacctga tgtccacgcg gcgctggaac tctggagct 9840  
cgtgaaggac ttgagcggg cgacgctgtg tgtgctgac gacctcaatc tggcggccgc 9900  
ttattgcgac cgtatctacg tctgcacgg cgggaggatc gtcgcggacg gccgcccgc 9960  
25 cgaggtgctc gaccccgaac tcgtccggac cgtcttcgga gtcacctgta cgcactgac 10020  
ccatccggtc accggcggac tgttctcgc gttctgccc cagcagcccg tccggtcacc 10080  
gcactcggg aaggggtgag aactcccga cgtcggatc gcgtcaaata aattccctc 10140  
30 gcacgctgc tgcgtccctc ataatgctt ggtgcgtcat aagggcggc gggcggccg 10200  
gtatcggcc gtctggtg ccccggtcga ccgtagtcgc aaaagtgtt tgctgttga 10260  
35 cggcggcaga atcatgtgc gcccttgga ccgtgaagt aatctctga agtccgaagt 10320  
gctcgtcat atgcgtatc caaacgaaag gcgagtcacc cgggatcga tcgtgtgaat 10380  
gtcataccgt taacggagc ccagaatgaa gacgaagccc aggcggccat catccagatc 10440  
40 gacgaagtcc tgacgctct ccaggagtg ggcgacgtc cgcgccgcg acggatgaa 10500  
ctggtgccc cccggcaccg ggtactcgt gtcaccgcg ctcaggtacc ctctccgga 10560  
45 acctccgca cccggcaccg gcccgagggc gaaccgccc agcagcccgc tgcccgggc 10620  
ggcgggccc gagccgacca gaagcccct cgcctcacc ggcgccagaa ggaagtgctc 10680  
aatctgctg cccagggcct gtccaaccg cgcactggc gcgccctgca catcaccgag 10740  
50 cagacgtga aggccatct ccacatggtg taccacaagc tcggcgtcgc ggaccggaca 10800  
gaggccgtg tcactgccct gcgccagggg ctggttcagc atcagcgcag ggaggaccg 10860  
55 ccctcctgac cggcccgcgc gccgacgga ggcggcagtg gcaggagccc gcgccaccg 10920  
cggaaacgga tccgctcgt gcggcgtgaa cgtgcaggcc cggccctggg acggctcggc 10980  
cggggcttg gcgcatcgg ccgaccctc atgcgaacgg cccgggctc cccgcgaccg 11040  
60 gtcaggccc gatacgggc acgacctga tgagcgcgc gttgagctc tggccagcg 11100  
gcacgacct ggtgacggc agcccggccc gcgcgagat cgcggtgtac tccgccggg 11160  
65 tgcgctggc gccgccgatg gtgacgagca tatgaagatc ggtcagatag ggttctcca 11220

## ES 2 397 885 B1

gcgccgctg gttgaggag tccgatcgg cgggtcggg cagcaccggc tcggcgatga 11280  
gcagctccc gtgtcgggc atggcgtccc ggaccggcg cagaattgtc caggcgcgt 11340  
5 cgctgtcca gttgagcacc acgttctga tgagatagag gtccgcccc ttggcactt 11400  
cggcgaagaa gtctccgcg accgcgggc atctgtccgt cagccccgac cggcgatgg 11460  
10 tctcccgcg ccgtccgcc ccggcctcg tctccagcac cgtcccgtc aggtcgggt 11520  
accggtcgag gaacgcgggc agcagcgtgc cgtgccacc gccacatcg acgaccgtgc 11580  
tgaaacgcc caggtcgtag acggcggaga tggcgcggc catctccagc gggcgatgac 11640  
15 gcctgtcat ggccgctgg aacagggcg ccagctccg ccggcccgc agatagtcgt 11700  
acaccggcag tccgaacgc gtgtcaagg cggcagtc ggtgtccagc gactcctga 11760  
gattcagcca gttgcctgg aagacgtcat gggtaggaa cgcgcgaac gccagcagc 11820  
20 acgccgggc gtccggcg agcagcggc ccgctcggg gagcgcgaac cggccgggt 11880  
cggactccac gcacaggccg agcagtgcca gggcgcgag cagtctccc agtgttcg 11940  
25 cgggcacacc gcaggcagg gccagtcgt cgtatcggc cccactctc ccgatccgt 12000  
ccgccagtc cagccgcgt gccgcgcca cggcctggg gcccagggt ccgaagatc 12060  
ctgatatcac ggtgcgccc gcatcggct cggctgtcat gaactacca ttctctgt 12120  
30 gccggtgat ggggtccga ttccggctc cggctcgg gtgccgatat cgacaaggc 12180  
tggtgaaaa tgccaacgg gtccactcc gccttgacct cttgcagcct gggatagtt 12240  
35 tcgccaat acagcgtgt ccaggaacg ccggaggtt tccattccg gtcggccaga 12300  
tcgacatcc cgtaattgat gtaggcggc gtgtgacct cgtccggcac cggaacccc 12360  
cccgtagccc tgtcatgtc ccgtacagc tcgcgatcc accgaggtg ccggtcgtcc 12420  
40 tcggcgggt cctccaggt gaccatgtg aacgttca gcagggcgc ccgcggaatc 12480  
gccgtgcct cggcgggac ggcgttact ctgcccggt acgcatgta ctccaccgc 12540  
45 gcgtaggtg cccgtagtc gtccgggtg agatggcgt ggagggcggc gagctgatc 12600  
ggcgggtgc gggcgcgag atcgccgat tgacctgc accgagtc gatggccgc 12660  
ggcgggtct cgggtacggc cagatactg ctgcggcca gccacggcag ggaggtg 12720  
50 cggaccgtg ggcgatgcc cagccggcg cccagcggc ccaggaacgc gtccatg 12780  
gcgtggccc gcggcctc ctacggagg tggcgtgca ggtgatgtc gcccgccgac 12840  
55 cggtcgggc agtcaggag tgcgtacaga tcggcgtcg ccgagtcac cgcgctgtc 12900  
gcgatctcc agtccatgaa gttgccgacc aggggacga agtccgccc ggtgagccg 12960  
tccacggcc agctcacct ggcatgtg agggcggc gcggcgcg cagcaggccc 13020  
60 gccggttcg ccgggacgc ctccgggaa cgcattcag accgggtgat cagccgaag 13080  
ttcctccg cgcggcgt atggcccac cagaggtctc tgtcgggtc ggtttctc 13140  
65 cgtgtgcca ccaccgtca cagttccg tctgtcca cgaccacc ctcgaccgc 13200

# ES 2 397 885 B1

tacaggtggt cgacgctgag ccccagcagc cgggacagcg gtccgtagcc gccgcccgtg 13260  
atgtgtccac cggcgcccc gcccatacag aaccggcggg ggacgggtgac gccccagtgt 13320  
5 tggaaagagg tggtcagcat cttcccacc ggagcgcccc cctcgtatgac gaaggcgttg 13380  
cgctcccgtt cgtaatcgac gcgggacatc ggggagaggt ccagaatcac ttccacatcg 13440  
10 ggtgcggcga cgaagtctc accgcagtgc ccgcgcgac gcacgggtgat ccgcttccc 13500  
gctgtcaccg ctcccggac gcctcttcg acctgccgcg gcgaatgac gagacggaaa 13560  
tagtccggct tcgcatgaa ccggtagtgt taggaacggc gcaggttctc gtagcggatg 13620  
15 tcgtcggggc tgacgcggga ggaaggcggc gtcggactgc ccggcccgtt ctctcttcg 13680  
gcgtctggt cgtcacgtt catgggggca agctaccgg gaaataaggc aggccacgcg 13740  
20 gaacataaac gaaggctcga gccggtaaa tccggagatt tagaccgtt gccggtggtg 13800  
gacgcgccc tgtaaggctc ggaatcaca aagtgacag acggatcgtg cacctcgca 13860  
ctgcattcat cgacggctgt acgaaactc acggctatac gaaagaggca ccgaatgagc 13920  
25 agtccggccg gcggaccctt cttgcggac gcggtgctca atgattggct gacgagtgtg 13980  
gggctcggca tcgagtatt ccgagcggaa gccaacacgc tgtattacgt ggacgaggac 14040  
30 ggtcaggagg ttccggtact cgaccatgtc tgcggattc gctcgtgat cttcggtcat 14100  
aacaaccgg acatcgtgc gcatcgaag acggtgctgg accggcaggt tccggtgtac 14160  
gtccagctgt cccatcagtc gcacgccaac gatatgccg ccgtactcaa tgccatattg 14220  
35 cggcgggaga tcccggcgg cgacgggtgac tacgccgca tattcgcaa cagtggagcc 14280  
gaggccgtc agatatgtt gaagcacgcg gagctggagc gccgcgcacg cgtcgcgaaa 14340  
ctcaccgac agaccagcc gaacgcggag gaggcgcgc aggcggtcgg cgcgggggcg 14400  
40 gccacggtc cggagaacc gtagctccg ttcgacggg cgagcgggtg gtccgacagg 14460  
ctggagcagc tgctgccga gggcgggcg aggaacgcgg agctgatcgc gcgggggccc 14520  
45 gtccatctg cctggagaa cgcgttccac ggcaagctg tcgccagcat ccagctcacc 14580  
cagaaccgc actggcggct gccgttacc tcgctggcct cctcaccgg cttcctcgc 14640  
50 gccgaccggc cggacgagat gaaggcggc gtcgaggaac tccggacgag cctgctggac 14700  
gtcctggtg acgacggcgt ggtgacggtg gtggagcggg acttcccct cgtcggcgcc 14760  
ttctcgtg aaccgtcca gggcgcgaac ggcacgctc cctgaccga ggcggcggcc 14820  
55 cgcgagatc gtgcggtgt cgacgcggtc ggctgccgt tgatcgtcga cgagatccac 14880  
agcggtatg gacggacgg cgcctctc gccagttgc acatggcct gcggggcgac 14940  
60 tactacacc tggcaagat catcggcga gggatcgca agaacgcgtt gccctctt 15000  
caccgggacc ggttccccc ggagttcag gtcacaca gctcaccct cgcgaaggac 15060  
gggttctcc cgccatcgc cctcaaggtc ctggagatg tggaggcca cgacggggcg 15120  
65 gcgtacagga tcgcggcca gcggggggac cggctgaagg gggcgtcac cgcggtccc 15180

# ES 2 397 885 B1

gccgacttcc cggacgtcgt ggacgcggtc cacggtatcg gcctcatgct cgcggtggag 15240  
ttcaaggacc agaagggggc gtcgtccgag ccgctgcggg agaaggcggc gtcgggcatg 15300  
5 ctcgggtact tcatcgcggg cctgatcctg cgcgagcacc gcatacgggt gctcccggtc 15360  
ggacctgccg gaaactccgt ccggttcgaa ccctcgatct acctaccga cgccgacatc 15420  
gcccgtacgg agaacgcgct gcgcatgctc tgcacgatcc tccgcgacca ggacggagac 15480  
10 cgtctcacc cctgacagcc ttcacgaggg agaggagcag cgcgctgaac gacgagacga 15540  
ggcagatgag ttgagcgggc agaagggcag ttcgacatg agcggcgcgga tacctaccga 15600  
15 gtcctcgtg tcgggactga tggagcatgc cctcagcgcg gcctgcgcg cctcgggtcg 15660  
cgcgccgggtg acctggggc tgcccggagc cctcggggaa gcgccggcca cggccgacga 15720  
gctggcgaca gcggtcggcg ccgatccggg ggccctgcgg cggctcctgc ggtcgtgac 15780  
20 ctctacggc gtgtcggcg aggaggggac tggttcgttc gtccacaccg agaagtccc 15840  
tgctctcgg gaggactccc cggacagcat caagtacctg gtgctgtgtt gcaccgaacc 15900  
25 atggctctg tcgctgtgg gagacctga cgaatcgggt cggacgggcg gcgagatctt 15960  
caccaggacg cacggggcga ggttctacga gcacctgcac acctgatggc ccgagtccgc 16020  
gcggatcttc aaccgggcca tgaccagca gaccgggtg tccgcgaccg tcatcgcgga 16080  
30 catgctccc atgtcggcg ccggcaccgt cgcgacgctc ggcgaggac aggtctcgt 16140  
tctcggcacg ttgtggaac gccacccga cgtacgggga ttctcctgg acctccccga 16200  
35 ggtcgtggcc aacgtggacg cgcgactgca cccggcggt gaactggccg accgggtgcg 16260  
gctggtgcc ggtgactgct tcgaggggat ctccgtcgag gccgacgtct accttcaa 16320  
gaacatcctg gggccgacg acgacacatc ggtccggtc ctgcgcaacg ccatgaaggc 16380  
40 ggcccggcg ggtgccgga tggatcatgt ggagaacttc gtcgacgacg gcccggcgga 16440  
gaggctgcc tccgcgctgg atctcggtat gctgctcgtc atcggcgggc agaagcacac 16500  
45 ccgcccggg ctgctgggga tcgaggagcg ggccggcctg acctgcggg acgtacgcc 16560  
ggtgactcc tcgctcaca tgatcgagac gttgtaccg ggtagaccgt cgcaccacga 16620  
acagcaccg tcacctgtg ggccccggc atccgatgcc gggccccac aggtgccggg 16680  
50 ggagacgtca gccgacgacg accgccttg accgcatgag gaaagagccg agatacct 16740  
cgtcgggaag gccgggagg gtccacgtg tcggacggtc cccgatgtg atctgtcga 16800  
55 tggacggctc ggcgaccagc gattcgatga gcccgctgct gtcggtgaag gcggtgagcg 16860  
acaaggtgct ccgaggggc gccagccagt cgccctccc cgccacggc agcaccaca 16920  
cgcaggggaa gccatctc ataccggcct cggggcgctc cggctctgag agcaggaaca 16980  
60 ccgcccggac caacaccgc gagccgtccc cgagttcggc gacgaccgtg tccgaccca 17040  
gcaacggccg cgcgtcggc gccccggccc gcagatacgc gtccatctc tgccgacgg 17100  
65 cggcccggcg gaccggcagc cggggccgct cgtcaccgg cggcagcgag ggaagctcg 17160

## ES 2 397 885 B1

ccagccgctg cgcgatggcc tcggccacgg gctccggatc gccctccacg aacaccgccc 17220  
tcgcgttac acaccgggtg ccggcgtgac cggcgatcga ctccgtgatg gtggccagat 17280  
5 gccgcccgac atccgtgccg gaggtgagga ggatcttcca ccggcccggc ccgaacggca 17340  
tcaccagggt gcttgaccgg tacttccggg ccacctcgtc gccgccgaac gcgatggcga 17400  
cgtccgcctc ggcgaccagc cggtcggccg tcgcgtggtc cgtcggcaac aggacgacct 17460  
10 ggtcatggtc gtaaccacc tggcgcaggg ccgagaccag ccggtgggcg gtgaacggct 17520  
cccgtgcga cggcccggacc gccacgcgat accccagagc cagagcgtcc agccacgagg 17580  
15 tgtcacggc aggggagtg gcgggagcgt gcacatcag cacctcgcca cgacgggtcc 17640  
acaccccgga gccgtgccg gcccgctgat cgcgccagct gtcggccgcg cccaccggcc 17700  
gggctgacg cacattctcg tacgcggtcg cggcgtagtc gccgacctgc tggccgact 17760  
20 gccgtacggt ggcgatcggg atgccggaca cctcggcgac ggtccggtgg taggccgcg 17820  
cggtcacgcc gtcgaccgag ccgctcgcga agagccagcc cgcgccgctc aggagttct 17880  
25 tgcgccgctc cggatccgtt ggtgcggccc ggcgcagcgc ccgatgctg cgctgggcca 17940  
acagccaggg aacgaggctc agttcggcca ccggttcgcc ccggacatcg tgcacgactt 18000  
cgaggttccg gctccgttac ggacccccg tgcgagcgc gtcgagggcg ggcggcgcg 18060  
30 aggacgatgc ccgacgtca gtagaccctc tcatcacgg gggccgacc gaactcccgt 18120  
accggcttga actcggacac ggcgtcacc gggagccca gccggtgctg gtggcgatg 18180  
35 ccggtgtccc ggtccaggtt gttggcagg aacaggttc gggtaggtg atgggtgagg 18240  
acctggccgc gctccccgta cggcacgttc ttcccgtgt cgggatcga caccgagaac 18300  
atggagaacg gcgacggcgg gtcgaagacc ggggactcgt cggccggtga gtccggccgc 18360  
40 tccggcatcg cgcagaagat catcgtgctg ccgaagatgc tgacgaagtt gatctcggg 18420  
aagagctcgg tccggaacag gtggcgggtg tcctcgtcca tcgacgtacc gccgtagatg 18480  
45 accgtgttga ccttctggt gatgagatcc accagatgat cgcgcccgca gaccgcctcc 18540  
aggagcggcg gggtagtac catgaccgg atgtctcggc tccgaggat ccattcgtc 18600  
tggttacca ggtgggtgat gtaggcgttc gcctcgtcga cggcggcgg gccgatgacg 18660  
50 agtttcgcc accgcggtc gaagtccacg gtgaagaaga ttcccggaa gcgctgggcg 18720  
gtgtcctggg ccaggatgcc ggcatgtgc ggtcctcgg gcatgatccc gagccagttg 18780  
55 tcgcccggc cgatgccgtg ctccaccagc ctgtttagt accaggccca ggactgctcc 18840  
cggacgtcgg gcagcacgaa caccgcctc gccgccccg tcgtcccgc gctctgccc 18900  
accaccggcg gcctggacag cctgtccgg ccgcccgtac cccgggggac caggtcctc 18960  
60 atgcgctcgt cgcgcagttc gtccacgac ttgggaaca gcgtcagatc ctgacggtc 19020  
cgcacgtcct tgcgcgatc gaattccagc tttctgctt gctcctgcca gaacggcgat 19080  
65 ccggtcttcg ggtcgaatg ccatgccatg gcttccgga cgagatcctg gcctccggc 19140

# ES 2 397 885 B1

gggctgtccc acggaacgtc gagtgcggat gcatcgcgtg aggtcattgt ctgtctctc 19200  
ggcattctct ttcacgggc agctgtattt tgaagtaca aggattcctg gtcctcggc 19260  
5 catacccgct gtctcgtccg atgggtcaac tggttttaga accgtggatc tagactcatt 19320  
gccgcccgcg cggagttgtg cgtagcgtgc tgcgtgctgt tctgttttc gtagccatcg 19380  
10 catacgtgcg ggtgggagtg gaagatgaga gagagccgct acgacgtgat cgtcgtgggt 19440  
gcgcggtgtg ccgcatcgcc gaccgcgatg ctgtggcga ggaaggggta ccgggtcctg 19500  
gtcgtggacc gggcggctt tcccagcgc accctgtcga cccatctggt ccaccgccc 19560  
15 ggcgtcgcgg cgtcgcgcgg ctgggggctg ctcgaccggc tggggccac gggctgtccg 19620  
cccatccaca cctacgagt cgaactcggc tccctcgtcc tgccgggccc cccggggagc 19680  
20 gaggccgaac cgtacgcgta cgcgccccgt cgcactgttc tggacaagct gctggtggac 19740  
gcggcgaggg aggcgggggc cgaggtccgc gaggggttca cggtcaccgg gctcgtctc 19800  
ggcgacccg gcgaggtggt gggcgtccgg gggcgcgggc cccgcggccc ggaggtgacc 19860  
25 gagcgggccc gcgtggtgct cggtgcggac ggtctccatt cgtcgtcgc gcgtgccgtg 19920  
gacgcgccc ggtacaacga gcatccgaaa ctgatggtcg gttactacag ctattcagc 19980  
30 gggctggaca tggacggtgt ttcaaggcc cattcccgc cgtaccgcag ttcggtgcc 20040  
tggcccacgc atgacggact gacgctggtc ggtggtgct ggccgttcgc ggagttcaac 20100  
gacatacga aagacatcga ggggaattac ctaaagaact tcgcgctggc gccgcctgg 20160  
35 gaagaacgca tacgtgatgc ccggcgcgag gaccggatcg tcggagcggc cctgccgaac 20220  
ttctccgca aaccttcgg tcccggctgg gcgctggtcg gtgacgcccg ttattgcaag 20280  
40 gacttctca ccgccaggg aatcagcgc gcgttcatct ccgccgagat gtgcgccgga 20340  
tccctggacg acgccctgtc gggcgcgcgg ccgttcgaca cggcgtggc gcctatcag 20400  
gcggcccgcg accggcatgc gcgaccgctc tacgacttca cgtccaggt ctccacgctg 20460  
45 gaaccctgt ccccgaggt cgagaaggtc ctggagggca tcgacgggaa tcagcagggc 20520  
atggacgct tcgccaggt gaacgcgggt gtgacgtcga tggagcgggt ctccgccgac 20580  
50 tggggtggcg ccgtccgacc cgtaccccgg tgacgtgacg aaaccgtccg cggctcggcg 20640  
gacggttcc ccgctcacc gggcgggagg ccggacggct caccgggagg tgacgggctc 20700  
cccggactc cacaccggtg tgatcggtc cgggtcggcc agcacggtga tctcggccag 20760  
55 cggattccct tccaccgcca ggagatcggc gtcgtagccg gctccagcc gcccgaccg 20820  
gggcgcccgg ggcccagcg tgccggggcc gtgggccgtg gccgcccgtg tggcctccag 20880  
60 cgggctcagc cccgcggaga ccagatgggc gaactcgtc gcgtgcccgc cccaggacag 20940  
cgggccgcca cgtcgtctg tgcccagatc ggtccgagc gcgatggtga ccccgccc 21000  
gtggcgatg ccgatcgcag tgaggtgccc ctcggccatc acctcgaagc ggtcccgcca 21060  
65 gcccggcggc agcgcggcca cgtgtgccc gaacgcctcg tagatcgtcc gcgtgggac 21120

## ES 2 397 885 B1

cagcgtcatg ccccgtcgg ccatcaggtc tgcggtctcc tcgtcgatct ccgtgccgtg 21180  
ctcgaccgtg cggcatcccg cgtgatcgc cgccaggatg ccggcccgcc cgtggcagtg 21240  
5 ggccgccacg atccggtcgg cccgcgggc ctcggtgacg atcgcgttca actcctcgt 21300  
gcggtactgc tggtgaccg ggtgtccac ctcgctcagc acaccgcccg aggtgcagat 21360  
cttgatcagt tcggccccg cccgcagttg gagggggacc gcccgcacgc actcgtccac 21420  
10 cccgtcggcg atcgcaggg tcccggcct ccggcagga tcggtcacc agcggtagg 21480  
caggcgggtc gcgtcgtgt gcccggcgt ctgcccgatg acctggttcg cactgtagat 21540  
15 gttcggcccg gtgaaggctc cttcccggat ggcttcagcg agtacgcacc cgtggccgcc 21600  
catgtcgcgc acgtggtga agccggcccg cagcggcgtc tccgctctc tcacgtgctg 21660  
ggcgacggcc agcgtctcgg ggtgagcat cagttcctcg gtgctgacct ggccgcgtat 21720  
20 cccggcgaag tgcacatggc agtcccacag gccgggcagc agagtggcga cccgtgctgg 21780  
ctccaacgcc cgggtctcgc ccgacagttc ggccggtcgt cccgcgaagc ggatacggcc 21840  
25 gtccgatcg acgacgtcg cgttcacgtt gggctcggc gcgccgggga tcagcaggtc 21900  
cgctcgtat cgatgttcca tgggtctacg cctccacggc tgtccggcgc gcgatgtcgg 21960  
cgcgagcac ttctgtcg atctgcccga tggcggctgt cggaactgg gccaccgtct 22020  
30 ccatcagttc gggcagctc cagacggcga gaccgagtc gcgcaggaac cggcgcagct 22080  
ccagcagggt gggctcccg cccgcctccg gcacgacgta caggcagacg acctccccgt 22140  
35 acagctcgtg cggggcgccg acggcggcgg acgcgcgcac gtccggatgc cgtgccacca 22200  
gcaactccag gtctcggcc gggatctct cccgcccgcg gttgatgatg tcgcaacc 22260  
ggccgttgc gacgagatcg ccccggacg tccgggtcgc caggtcggcc gtacggtaga 22320  
40 agccgtccgg ggtgaacgcc ttcggtcgc cggccgggtc ccgtagtac cccgcgacgg 22380  
ttagggccc cgggtgagg agtccccg tcgtccgtc gccaccgga cggcctgt 22440  
45 cgtccacgat gacacctcg tccatcggg aggcggggcg gccctgggta ccgacgatca 22500  
ccgactcggg gtctcagc cccgtgtagc tgagcagccc ctcgctcatg ccgtagcact 22560  
gctggagcgt cccccagc tcggggcga tccgggccc cagctccgga gcggggcggg 22620  
50 cgccaccgtt ctggagcacc ctacagctgc tgagatcggc ggtcgtcccc tcggcggccc 22680  
ccagccactg catggccacc gtcgggacca gcgcgcagtg tctacgcgc tcgctcga 22740  
55 tcagcccag cgctcggg ggtcgggtg tctcgggag cagcaccgtg ccgcccggg 22800  
ccaactgcc gagaacccc gggcagttca ggacgaatcc gtggcgggc gccatcaccg 22860  
ccaggtacac cgcgtctgc gagacgtccg cgattgcga ggccgtcgg atcatgtacg 22920  
60 cgtatcccgc gttgccccg gggatcacct tcggcgggcc ggtggtgcc cccgacagca 22980  
ggaacacggc ggcgtcgtg agggcggagc cgcgcccgg gcgcgggtgc gcggaggagg 23040  
65 cggggccac cggcgtcct cacggcgcac cggggtcga cagtccgcc aggtcagct 23100

## ES 2 397 885 B1

caccggccgc gacatcgccg agcgcgagga catggtcaag ggttctggtg tctctgcgca 23160  
gtgccgggc cgattccacg gggtcggacg acctgctccc ggcgctcacc gcgagggcga 23220  
5 ccggccgggc ggaccgcacc acatggccca gttcgtagtc gccgaacccg ggcagcacca 23280  
ggaccggtg ggcgccgatg tcgagcagcg ccaggaccag gaccacgaac tcggcacagt 23340  
10 tgggcagttg gaccaccacc cggccccgg cacggacccc gagcccctcc agccgcccgc 23400  
cggccacggt gacggccgtg tgcagctcgc cgtgcgtcag ccgctcgtgta ccgctcacca 23460  
cggcgaggcg gtcgggggtg cgtagccgcg cggcccgcag cagctcggcg acgcccctct 23520  
15 cgcgccacag cccggcccgc agataggcct cggcccggcg ccggtcgtac accggtacgt 23580  
ccggtatcag cggccggctc tgatcgggc tcacggggc atctcccgt tccctgccgc 23640  
caactgctc atggtgtagt ccgcctgtc ccacagattg atccggacgc cgaatcgttc 23700  
20 gcgcaactga atctccaggt cggtggttc gaccgaactg agaccgagat cggcacggag 23760  
ccgggtaccg gttcgtatc cggtcagcgt ccgctccgcg ataccgatat cgagcaacaa 23820  
25 ctgctgagt acatccacg acacggcgc tcctgattt cccgacggca aggcaggccg 23880  
aggctattcg acgtaccccg gcctgacaat ggactagtac ctcggtacta gccggcacc 23940  
cggaaacccg gtgcgggtgg gcgaggctg agaacggcac tagaccacc ggtgtattc 24000  
30 cgatgggagc gggatttgg agtgttcgc cgatgccgg cgaatatcg acgtcagcg 24060  
tgcggtggtg accgcggaat actcgcgccg ttctcggat cagcgtagtg atcggcgag 24120  
35 tgctcagcg agcctcggac cagcacatat cgatgccga actgtcggat gacacgtgta 24180  
aaggcggctg ggagccttg ggttgacgg tggaacggtc gtacgagatt ccctacccgg 24240  
aactcggctc cgtgctcga gagcaggaaa tcggggtgct caccgggtg gtcacatccg 24300  
40 gtgagaatct ttcgggagg cgatgctgg aggagtcca gcgctgttc cggaatacc 24360  
tcgactgcc gcacgccct tcggtgacca gtggaaccgt ggcgctggag atcgcgatac 24420  
45 ggtgctcga cctcaatgag ggcgacgagg tcatagcag cccgcagacc tacaaggcca 24480  
gtgtcaacc gctgtgaac tacccgtga aggtgcgct ctgcatgctc gggccgaaca 24540  
cgctcaacat cgacccgga cacttcaat cgctgatcac cagccgcacc aaggccgtga 24600  
50 tcctcgtgca ctacggcgt ctgccctgc acatggacgc catcatggcg atcgcgcc 24660  
ggcacggaat caccgtcatc gaggactcg cccacgcgt cggcgcgag tacggggac 24720  
55 gcaagccggg cgcgctggc gacatcgct gctcagct cactccagc aagaacatca 24780  
ccaccctcg cgagggcgg atgatcacg tttcgacc ggcgctggc gaacgggcg 24840  
accgatccg ctccaacgac gccacgcg tctaccgcg acaggccagg gccatcggga 24900  
60 acaccaccg cgcgacccg tggatgctc acccggcgc gccttcacc cagactgct 24960  
cgacgatccg ttacggcggc acgaacgcca cctcgcga gccgaacgcc gccgtgggca 25020  
65 ccgtacaact ggcgaagctg gaccggctg tgcgcccag cgcggagatc gcggcgct 25080



## ES 2 397 885 B1

acacggacgt cctcaaacag caccgccgcg tgcggatgca cgagggaccc gatccgtcc 25140  
ggcacgcca ccacctgtc acgttctcg ccgaccggc cgacggcatc ctgaggacc 25200  
5 ggctggtctc acggctcgac gcgctggcg tgcagatgca gttgcgtac ttcccatgc 25260  
acctctcgc cgaatggcg gccagaggcc acaccggcg tgagtcccg gtcgccgaac 25320  
10 gcctgtggtt cgagcaacag gtcaacctgc cctgccacc ggcatgacc gaccggcagg 25380  
tcggccagg ctgtcccgg ctggacacc tctcggcca ggaggcccg gaccaaccga 25440  
tcgctccgt gacgacccc acagcggagg gctggaacc atgagcgaac ctgagcacta 25500  
15 cgacatcgc gtgatcggc gcggccccgt cggactcgc tccgctggc acgccccc 25560  
ccgcgcgaa cgggtcggc tctcagca gttctctc ggcaacgaac agtgcggcac 25620  
gagcgggct gagcggcact ggcggctga gtacaccga ccgatctgt gccggctcac 25680  
20 cggggaagcc ctgccctct ggcgggaact cgaacggcc accgggacc agctctcca 25740  
cgcttcgac agcctctgt tcggcgacat cgatcgc accaacgagg gcaggatctc 25800  
25 ggcgaccgc cggaccatg acgacctgc catccgtac gagggtc cgcggccga 25860  
catcgagcgc cgctacgat tcaccggct gccggccac ttcgaagggt tcgtccagcc 25920  
cgacggcga gccgtcgac tacgcgccac cgtggaagg ctgctcggc tgaccagga 25980  
30 agcgggctgt gccctcggg cgcacgagc ggtgctgga ctgatcccg accggcggc 26040  
cgtgaccctg cgcaccgac gcggccgct cggggcggc aaggctcgc tggccaacg 26100  
35 cgctacgcc aaaaactgc tcgaaccgt gggctcccg ctcgacctc acgtgtcga 26160  
gatggcgtg gtgacctgc gccagcga cccaaggc cgctaccgt tctggtcgt 26220  
cttcaggag ccgaccgag aggacacaa tctgtctac gggttcccg ccaacgctg 26280  
40 gcaggacacc gacacgtcc ggtcggccc ggtgtcag gtcaacgctc tggccgacc 26340  
cgcgggcg accggcacc ccgatccg ccatgtggc cggatgtcg agtgggtcga 26400  
45 acggcatctg ccgtgtgg acccagggc gctggcccg gacacctgc tcgctctc 26460  
gcccggcgc cctgagcgc agtctctc cggcagcg aaggccggt tcgacggcg 26520  
cgagaactg gtgatcga ccggcgggt ggggtcaag ttcgtcccg tgctgggca 26580  
50 ggtctcgc gatctgtcg tggacggcg gaccggctc cacgtggacc gcctgatct 26640  
gcccagcgc gccaccctg gagcgttc gcgccgac gtgccgacc ctcccagc 26700  
55 aaaggagatt cccggtgt tccgaagc ccgcccag gtgcccaca cgacgtacg 26760  
tccctgaca tcgagctg cgggatggc gcccgttc ccggcggc cgacctgac 26820  
gcctactgc acaacctgc ctccggcgt ggtcgtac agcggctc cgaggacgac 26880  
60 ctgctggcg agggcgtc cccgagct atcggcgac cgggtatgt gccctcgc 26940  
ccgtgctg agggcatga cctctgac gccggtct tcggctcacc ccccgggag 27000  
65 gccgcgtc tcgaccaca gcagcggct tcttgaga gcgctggc ccgatggag 27060

## ES 2 397 885 B1

cacgccggga tcgaccccg cccgtgctgg accgccgctg tggctcggg cgcaacatg 27120  
cccgcctatc tgatgtcaa cctgctggc ggggcccggg tcgtcctga ctggccatg 27180  
5 ttcgagctcc agatacaca cgacaaggac ttctggcca gccgcaccg ctacaagctc 27240  
ggactgaccg gacccgctg gaacgtccag acggcgtgct ccacctcgt ggtggccgtg 27300  
10 caccaggccg ccgcccctc gcgtccggc gactgcgaga tcgccctggc cggcggggg 27360  
tgctccggg tgccgaccg ggtcggctac cgctacgaac agggcctgat ctacgcgcc 27420  
gacgggctc gcccccgtt cgacgccgac ggcgcccga cggcttctg caacggagcc 27480  
15 ggtgcgctc tcctcaagc gctggccgac gccggcgcg acggcgaccg gatcctcgc 27540  
gtcctcaag gctccgctc caacaacgac ggggcccga aggtcggta cacctcccc 27600  
agcgtgagc gacaggaagc cgtgctgcc gccgcgatc ccgacagcg ggtgcccgc 27660  
20 cggctgatc ccgcatga ggcgcaccg acggcacc acgtcgtga cccatcag 27720  
atcacggc tcagcaggc gtcggacc cacacgacc acacggggt ctgcgccgc 27780  
25 ggtcctca agtgaacat cggcctc gaatcggcc cggcctcgc cagttcctc 27840  
aaggccgtc tgcaactca ccaccgacc ctggtccca gcctgactt cgagcggcc 27900  
30 aaccgctg tcgactga gccacccc ttctctca acacggaact gcgtgctg 27960  
cccaggggg agcaccacg cgggatagg gtcagctct tcggcatgg cggcaccac 28020  
gccatgtg tgctggagc ggcccggat ccggtcccc ccgagccct cggcggccg 28080  
35 gagctggtg tggctcgc caagtcccc gcggcctg acggccac cgaggcctc 28140  
gcggagaag tcgccccc cgacgcgag ccgctcgg acatcggca cacgtccag 28200  
accggccg gggcatgct ctaccgag gccgtctg ccgccggc cgccgaggc 28260  
40 gcggcctc tctccggg cgaccggg cgggtgcga gcgcggacc gggcactgc 28320  
ccgggaagg tgggttct cttcccgc caggcgcgc agtaccgg gatgagccg 28380  
45 ggctgtac cgagcaacc ggtgtcgc gaggcctg acgctgcgc ggactgctg 28440  
gccgaggc tcggatga cctgcggacc gtctctcc ccgacgacc cgccgaggc 28500  
gggtcacc acacgagct cgcacagcc gccctgtc ccaccgagta cgcgatggc 28560  
50 acgtgctg ggtcctggg cgtcaacc gacgtatg tcgggcag catcggcag 28620  
ttaccggc cggctctc cgggtgct tcctgaag acggcgcg cctgctgc 28680  
55 ctgcggca ggtcatga ggaccggc accgggcca tggctccat cgccgaccg 28740  
gccccgaca tgaaccct cctgccgc ggggtctca tcgcccat caaccccc 28800  
gtgctctg tcgctcag gccgcagc gcggtggc aactcggc gatcctggc 28860  
60 gccaggaga tcaccgct cccgctcc acctccac cttccact cgcgatgat 28920  
gaccggctg tggagcgt caccgaggc gtcgccgaa cccgctgc cgccccggc 28980  
65 ctgctgtc tctctcgt caccggct ccatcacc ccgagctg cacgaccg 29040

## ES 2 397 885 B1

cagfactggg gcacgcacct gcgccgcccc gtccgcttcg ccgacgccgt gcgcaccgcg 29100  
atcggcgacg gccccgccgt cctcgtcgag gtgggccccg gcaacacct gagcacctg 29160  
5 gcccgcgccg gggccggcac cggcgggccc cgctgcgcg ccgtaaccac cctgcgcagg 29220  
cccgcagagg ccgccgacga cggccaggtg ctccgcaccg ccgtcgggga catctggctc 29280  
10 ttcggcggcg cggcggactg gcccgcctg caccagggcc gccgcaaccg ggtcgaactc 29340  
cccggctacc cctccagcg cgaccgctac tggatcgaac cgcggggctc cgcgaccggc 29400  
acgcccctgg tcgccgactt cgccgagcac gaggaggccg agacggagcc ggccggcagg 29460  
15 gcgaccggcg cgacacctt ggtcaccgcc tacgtggcac cggcggacga gctggagacc 29520  
acgatcggcg ggatctggga ggagatgttc ggcacgcgc cgatcggcac gcgggacgac 29580  
20 ttctcgagc tcggcgcca ctgcctctc gccatccagg tcctcaaccg cctccaggcc 29640  
acctccggcg tcacctgga gctgggccgg ctgctggcca cccccacat cgggggcatg 29700  
gccgaggagc tcgcgcggcg cggcggccc gccaccgacg accggctccc caccgtcgtc 29760  
25 ccgcgcccgg acctcggta cgagccgttc ccgctaccg agatgcagca ggcccagtgg 29820  
atcggccggc tgtccagctt cgacatggcg ggagtcgccc ctcatctgta cttcagttc 29880  
30 gacagccgca ccatcgagac cgcccggctg gagcgggctt gccagcgagt ggtcagcgc 29940  
cacgacatgc tcgggatgtt ggtcctccc gagggccggc agcagatct cgacaacacc 30000  
gagccgtacc gttcaggtt gctcgacctg cgcaccaccg acccggagga ggcagagcgc 30060  
35 caactggccc agatccgca ccggatggcc accgaggtgc gccccgccga cgtctggccg 30120  
ctgtgggagg tacgggtcgg cctgctgcc gaccaccggg tcgggtgca catcagcttc 30180  
40 gatctgttg tcgccagct ctctcgttc ttctaccaac tgctccgca atggcgggag 30240  
ttctaccacc atcccgaaca cgaccccgag cccctcggcc tctcctccg ggactacgtg 30300  
ctgccgagg aggagttcg gcggacgccc cgttacgaac ggtcgtgga gtactggcgc 30360  
45 aagcgggtcc gcgagctgcc ccgccaccc gaactgcca cggcgcaggg tcggggggc 30420  
ggcgaacgcg tggggtcgt ccgacggcac gcccgctgg acgccagct tggggcccg 30480  
50 atcaaggcga agcgggggga gttcggcgtc acccgtcca gcgcatgct cgccgcttc 30540  
gccgtacca tcggcacctg gagcaagtcc cagcggttca cctcaactt cacggccgtc 30600  
aaccggctcc ccgtcatga ggaagtgcac gatgtgtgg gcgagttcg ctcctcgac 30660  
55 ctactggagg tggacgcggt ctggcaccg gactcggccg gtctcgtcg ggaactgca 30720  
cggcagagct gggcgactt cgaccaccg tacgtcagcg gcgtacgcat cctgcgcgag 30780  
60 cgggcccggg cacgtggcgg cgcgggcat gtcacggcg tcgtcttac cagcgcgctc 30840  
ggctcggacg tggacggca gcccgcgccc tcgccgtgg actggctcgg cgagcagagc 30900  
tactcatct cgcagacccc gcagtgacc atcgaccact tctcgtgga gttcggggc 30960  
65 aacctggagc tggcctggca ccgctcgac gggctctcc ccgacgggt gatggaggag 31020

## ES 2 397 885 B1

atgttcagg cctaccagga cttcgtggtg ggcttgccg agaccgacgg ctggcaccgc 31080  
ccccgggtgc tcgacctgcc cgccggtcag ctgcacccc gcgcccgcg caacgacacg 31140  
5 gcgggcgagc tgcccgacgg cgtgctgccc gcccgatcc tggcccgggc cggctccgcc 31200  
gaaccggcgg tgataccga ggaccgtacg ctgactacg ccgaactcac cggccgggcc 31260  
10 gtggcgttg cccgcgaact gaccgagcgg ggctacggcc gaggggctgt cgtcggatc 31320  
ggactcgcca agggctggcg ccagaccgtg gcggcgctcg cggcctccgc gcccggtgc 31380  
acgtacgtac cactggaccc cggccttccc gagggcccgc gccgctggct ggtcgaacag 31440  
15 gccggtatcg gctgcgtact ggccgaaccg gacacggcgg cactctggcc gaacgcgcc 31500  
cgggtgctgc cggctgcgga ggacgcgagg tgggacccc cggacacggc ggctggagc 31560  
20 tgtcccgcc gcccgagga taccgctac gtcatctaca cctcgggctc caccggaacc 31620  
cccaagggcg tggcggtag ccaccgggcc gcgctcaaca cctggtgga catcgaggag 31680  
cggttcgga tccggccgg agaccgggtg ctgggtctct ccgccctcaa ctcgacctc 31740  
25 tccgtcttc acgtcttcg gatgctcgc gcgggagggt cgcgctgct gccggaggcg 31800  
gcgaccggc gcaaccgga ccgctggacc gagctgtgcc ggccaccagg ggtcacggtg 31860  
30 tggaactcgg tgcccact gatgcagatg ctggtcgaac acctggagag ccggggccc 31920  
gccgacgac cgggccacct ccccggtcgg cggctcggc tgcacagcg ggactggata 31980  
ccgctgtccc tgccggaccg gatacggcgg gtggccccg caccgacgt gatcagcctc 32040  
35 ggccggcga ccgaggccgc cgtctgttcc atgccacc cgatcggcga ggtcgtatcc 32100  
gactggccgt ccgtcccga cggacgccc ctgcgcaacc agcgttcca cgtctcaac 32160  
gaccggctgc ggcacgcgcc ggtgtgggtg ccggccagc tgcacatgc cggggccgga 32220  
40 ctgccgagg ggtactggcg ggacgagcgg cgcaccgccc agtcgtcat cacgacccc 32280  
gagacgggg agcggctgta ccgaccggc gacctggac gctatctcc ggacggcacc 32340  
45 atcgagtcc tcggccgga cgaactccag gtgaagatc gcggccacc gatcgagctg 32400  
ggcgagatc agcacgcgt cggctcccac ccgagctgc tcaacgccgt ggtgtccgc 32460  
50 cccggggaac gcaaccgca acgctggtc gccatgtc tcccggcga cccgggacg 32520  
cggaacgat cggacttgc cgaaccgctg cgggaccacc tcaccaccac cctgccctc 32580  
tacetatcc cgtccgacat cgtgtcacc gacgcatc cgtgagcgc caacggcaag 32640  
55 gtcgaccgt ccgactgcc cgaccctcag cggacggcgg acgcccaggc cgcggcgagc 32700  
gcccggagg acgacgggga ggaggccacc gggcgctgc gcaccctgt ggtcctggca 32760  
60 gccacctcc tcggcgtgaa cgggccccg ccgcccggca acttctcga gctcggcggc 32820  
gactccatca tgggggtgca gctcgtcggc agggccaacg ccgaaggcat cccatcacc 32880  
ccgagaacc tgtcagag caccacctc ctgaactgg ccgcccgtt acccgtcag 32940  
65 ccgggaacc acgacaggg caggcgggtg gactcacc cgcaccagac cctgccccac 33000

## ES 2 397 885 B1

gcccaggtgg gctcggfact cctcgacgta cgggacgct tcgacccggc gtccgccgc 33060  
cgggcactga acgccctggc cgaccgccac cccgccctgc gaaccgcgt gcgtaccgag 33120  
5 gacggtcagc ggttcgccgt acgccccggc cccggcgagg acttcgacgt accggagatc 33180  
gacctgccg ccctgccga cgacgtacgg gccgaagcgg tggcggagat gatcggcgag 33240  
10 atggccgggg aggtggacat cgagaccggt ccggcgggtca agttcgggt gttccggctg 33300  
ggcgaacgcg gcagcgtact ggctgcacg gcggcccagg gcctgatga cgacgcctcc 33360  
gtcctcctgc tctgccgga actgatccag gcgtacgacc ggtggccgc cggccggccg 33420  
15 gtggtgtgga gcgacggcg gggctcaccg caggcatga accggggcct gcgccggaag 33480  
cccgcgacc cggccgggt cgccgaaat ccggaacag caggggagtt gccccggcag 33540  
cggaacatgg agctggacgc cgcgctacg gcaggcctgt tcaccgcagc cgcgggcagc 33600  
20 caccactgg accccaccga ggtgctctg gccgcggcgt ccgccccct cggcagggcg 33660  
ctgcccgaac cgccgcaact cctcgtgga cgttcgtcc gcgacgacct ggccgccggc 33720  
25 gacgaaccg ccggacgct ggtgggccg acgaccgagc tgcgcacggt gcagccggtg 33780  
gcggcgggta cccccctga cacggcctg acctcgtga agggccggt ccggacggcg 33840  
gaccggacc cgtccgggg gacgacggtc gcggtcagg aagtcgtgac ctgggaccgc 33900  
30 gtcgagggcg cgtcagagt gcccgggac ttcgccggag tgaccggcct cgcgggctg 33960  
cacgaggaga cgtggggca gctgagcgc gccgtggtg acggcgcct gcggatccgc 34020  
35 tggcaactgg ccgcatccgt gccggaggac gcggcgacc gcctggccga cgcgttcgc 34080  
acggtgctcg gggagatgc cgagcactgc ccggggtg cggagggctc ctacgaaccg 34140  
tccgactcc cctggcgga cctcgggg gacgaactg ccgagttct cgatgaactg 34200  
40 cggtgacggc gaccggtgac cgacgaatg cgacaacggc ggacgggact gtgatgactg 34260  
acaacggtgc gacgaagca ccggtcccg acatcgcgga catctacga ctctgccga 34320  
45 tccaacagg cctgctctac gaacagctg cgacgggg tctcggcatc tacgtggagc 34380  
agctggggct ggagttctc gggacgatg accccgagca ctcgagcgg gcgtggcagc 34440  
50 tggctgca ccggcaccg atctcgca cctcctcca ctggcgaag gacggaagcg 34500  
ccgtcaggt cgtgcacggc tccgccggc tccgctgga gacactggac tggcgcgacc 34560  
tcgacgaacg taccaggag gagcggctgc gggcgtcct ggacgccgag agggccgagg 34620  
55 ggttcgatc caccgatgt ccattgatg gcagcagct gatccggcg gcgacgagc 34680  
ggtggactt ctctggcg ttctgcacc tgctgatga cggctgtcc ttcaccctg 34740  
ccatccagga ctcatcgac cactaccgg tgctgtgcc gggcggggc cgcagctct 34800  
60 caccggggcg ctctaccg gactacctt cctgtggcg cgaccggac cccgaggagg 34860  
cacgggagtt ctggcgcgag gagctggcg actaccggc ggtggagcag gtgcactgg 34920  
65 gcgggaccg gattccgaa ggggagcca cccacgcca ctcgagcgg atctggcg 34980

## ES 2 397 885 B1

atctggcccc ccggtgacc gccctggccc gcgcggaaca gctgacactc gccacgctgg 35040  
cccaaggcgc ctggttcacg gtgctggccc gcttctcgg ccgtacggat ctgcctgcg 35100  
5 ggatcacat ggcgaccgg cccccgacc tcgtcggctc ccaggacatc ctggggccga 35160  
tgatgccac cctgccgctg cgcaggaggc tggacccgc gatgatctg cggctgtggc 35220  
10 tgcgggagtt ggcaagcac ggcacgagg cgagcgggca cagcggctg cccctgaccg 35280  
agatgcaggc cctgctcga acggactccg cgattccgat cctgcagagc agcgtgtcct 35340  
acgagaacgt gccgatgcc gacttcgacc tggcggacgt cggcgggag atgaccgaac 35400  
15 tggctatga cggcaggccc cactcccga tcaccatgt gatcatgcc ggcccgata 35460  
tgccctgcg ggtcctcac gaccggcga aggtctcga cgaggtggc gagcggctc 35520  
20 ccggggaagt cgtctcgtg ctgacgaga tgatgaacg cccggacgtg acgtgggag 35580  
agctgacct cctcagcacc ccgagccc ggaacacccc cctgaccgag ccggacgcc 35640  
aatccctga cgagacgtc cggcggcac gcgcgctcgc cccggatgc accgccgtc 35700  
25 gctcggagg ccgcgccctc acctaccgg agctggacgc gtactccgac cggatgcc 35760  
ccaccctcg cgcgcggtg ccggcgtga cccgcgtgg gctgtcctg ccccgctcga 35820  
30 tcgagctgt gcccgcatg atcggggtg tcaaggccg gcgcgctat gtcccctgg 35880  
accggagta cccggcggc cggctggccg acatgctgc cgactcggc gcggagctg 35940  
tgctaccga cggcgcacc gccgacgdc tcaccgccc caaggccgg ctggtaccc 36000  
35 tcccagat ggacggagag ccggaccaca acgctccacc gccggtgcc gccgacccc 36060  
acgccccgc ctacctctc tacacctcg gctcaccg gcggccaag ggcgtcccga 36120  
40 tcaccaccg caactccag agcctgctg ccgcccggc cgaggtctc ggctcacgg 36180  
ccgaggacgt gtggacctc gcgactctc tcgctcga ctactcgtg tggagatct 36240  
ggggcgccct gggcaacgc gcgtcgtg tcgtgtgga ccacgagacc ggacgcgac 36300  
45 cgcgcgccct gggccggtg atcggcagg agcgggtgac ggtgctcagc gagacccc 36360  
cgctctcga gcactggtg cccgaactg ccgacgacac gtcactgdc aggtgttc 36420  
50 tggcgggga cggctggac ccggcatcc tgcggcctg gttcggcgc ttcggtgacc 36480  
gggacggag cggccgggag accaccggc gggccccgc cccggcatc gagctgaca 36540  
acctctacg agtcaccgag gccaccgtg tctcacgta ccaccgggtg cgcgaggag 36600  
55 acgtaagggc cggccggccg gtcccgatc gcgcgcgct gcccaaccag cgggtgtacc 36660  
tgctcgaga ggacgaccg ccggtcccg tggcggccac cggccagtg tgcgtggcc 36720  
60 ggcacgggt agcctccga taccagacc gggacggact caccgggaa cggctcggct 36780  
ccgaccgct gggggtcct tctcggcg cctccccct gtaccgacc ggcgacctg 36840  
ccacggccac cccggacgg gaggtgact tctcggccg cggcagacc caggtaagg 36900  
65 tgccggctt ccgctcga ccggcgaga tcaggcccg gctcgcgag acgggggg 36960

## ES 2 397 885 B1

tgcggtcggc gacggtgacc gtccacggct ccggaacggc gcggcgcctg gtcggctatg 37020  
ccgtgcccga ggaccggac gcggtcctca cagcggggcc ggtfccacc gaaccgctcc 37080  
5 ggaacacct gcgcacccgg ctgcccgagc acatggtgcc ggccgcggtg tactggatcg 37140  
accgatccc gaccacgccc ggcggcaagg tggacgtcgc cgccctccc gtaccggatg 37200  
10 ccggagggac cgaccggaac accgcgcca tgaccgagc cgagcgcctg ctgcccggc 37260  
tgctgaccga ggtcctccag gtcccagc tgggcgcca cgacacgctg ggccgctgg 37320  
gcctgactc gctcggcgc atcggctgg cggcccggct gcgcccgcg tacgccctc 37380  
15 atctggcgt gagcgtctg ccggccacgc gtacggtggc cgagctggca cgccggtgg 37440  
aagcccccg cccggtcgc ggggagggg cacggtgaac gcgaacccc gacagacccc 37500  
20 gacggacccc acggatcacc ccgacagccc ggagaaccag cgctggcggg agcgatccg 37560  
gcactcgcg gagaaggaga tcgcccgc ctccaccacg atggacagga cggcactct 37620  
ggacgcccgc ctgcccggc gcctcttcg cgaggggctg atgtcgtgg agatcccgc 37680  
25 cggctacggc ggcacggcg gcaccctc ccaactgat ctcaccatc aggaggtgg 37740  
gcgggtgat cccgctcgc ccgtcgggt gcacgtcac aacgtcctc tcgcccggc 37800  
30 gctcctcgc cacgctcgc gcgatcagc ccggcagtc ctccgcagc tgccaccgg 37860  
gaagatcggc gcgttcgcc tctccagga acagggggc agcagcgcg tcgctgac 37920  
caccctcgc cggcaggac aggcgggcta tctcgtacc ggccgcaagc ggtggaccag 37980  
35 caacgcggc aacccgatc tgctcctc cttcgcctg gccgacgcg gcggcccac 38040  
ggcgtcgtg gtcccgcgc acgcccgg ggtgtccctg gacgaccggg tgcagcagat 38100  
40 ggggtcgcg gcggggcca cctcgacgt gatctcgc gggaccccg tccgtaccg 38160  
ccagcgcgc gcccccgc gcggggcca gacggtggc ctgtccggc tcggcctcg 38220  
gcggctcgc atcggcgc agatgaccg cctcggcag ggagccctg acgcccgg 38280  
45 cggctacagc aggtccgg agcagttcg cggccgata gccgaccac agggggtcgc 38340  
gtcccctg gccgatgg cgagcccgt cgcggcgcc cggcactgc tgtaccggc 38400  
50 ggtcaccct cacggccc gcacggacc ggtcagctg atgaggtga ccgcatggc 38460  
gaagtacgc gcgtccagg tggccaacg gcggcctc gtcgctgg aaaccctcg 38520  
cgggaacgg tacaccgac cctaccccgt ggagcggtc taccggacg ccaaggccg 38580  
55 gaagatcac gagggcacg ccaacgtct gctccgacg atcgcctca tctgatcg 38640  
aggaagccc ggtgactg ccctcgcgc tctctgtt cccgtacgc gcggaaacg 38700  
60 ccagaccta cgtccggtg gcgggcacc tggcaccga catcaggtc tgcccatgc 38760  
aactcccgg ccacggcag cggatcggc agcccggc ccaccctg gacgacctc 38820  
tggccgacat ccgaccaga ctgaccgac tgacgactc tgaggaccg ccgatcccc 38880  
65 tttcggcca cagcctggc gcctcctc cctcagtg ccccggatc ctggttccg 38940

## ES 2 397 885 B1

agcacggcat ccggccggcc cgtctgctgg tctcgggcca ccgggccc cacctcccc 39000  
tgccgggagga gacactccac cacctgccg acacggagtt cctcagagg ctcaagtgagc 39060  
5 ggtcgcgcac gctcgcgcg ttgaccgacc ccgagttccg caaactcctc cttcccatgc 39120  
tgccggccga cttaccgcg tcggagacct acacgttccg ggagggaccg acgctgacct 39180  
gcccgatcac gcccctgggc ggcgagaggg acgaggacgc gaccctgggc gaggtcgccg 39240  
10 cctggcaacg gcaactacg gcccggttcg agctgacggc ctccccggc gatcacttct 39300  
tcatcgacga cgctgggag gcggtgtga ccgcggtcgg cgaccggctg cggtcacgag 39360  
15 aggggagcac accgggctga ggggcgcgaa ggcccgcacg gaaccgcag gtggagagac 39420  
ggctccggct ccgtggacgg gctacgccag gaccggaagg gacacgagtc cgcggacgag 39480  
acgggttctg cgccactcca gcagctctc cggtagcggca aggtgattc ctgaaagcg 39540  
20 ggccaggaca gcccgaagag cgatctcgag ctccgccttg gccaggggtg caccaacgca 39600  
gcggtggatg ccgtggccga aggcaagggt agcagtcgcg ttccggttca gatcgagctg 39660  
25 atccggcgc ccggaaccgc tgggttccc gttggtgcc ccgagggcga ccaggaccgg 39720  
gggtcgccg gggacatcgg tgccgccgag ggtgacggct tcccggtga accggaaggt 39780  
ggctgtactc accgcggcgt cgaagcggag cagttctcc agcgcggcgg gcatctctc 39840  
30 ggggttcca cgaaggcgt ccagctcggc ggggtgccg agcagcgcga ggagggcatt 39900  
gccaagggcg ttggtgtgg tctcgttcc ggctacgagc agcagcacgg ccatggagat 39960  
35 cagttctcc tcggtcaggt ggccgtctc gtcgcgcgc gagatgagcc ggtcgaggaa 40020  
ggacctgcc gggttttcc gttgtcggc gaccaggtct gcatgtagt cggccatcgc 40080  
gtcgaggca gcgtgatga tgccgggctc tcccgcccg aacagtccg cggaccagcg 40140  
40 ctggacatcg ggccggtcgg ccggcggcac tccgagcagt tcgagatca cgatgaccgg 40200  
cagtgacc cccagggcgg ccacgacgac gaacggccc tggcgggcc actggtccag 40260  
45 aaggtgtca gtgatccgg cgatgaacgg gcgagctgt gcgagggccc ctctggtgaa 40320  
cgcttggtc accagttgc gcagccgggt gtgccgcggc ggatcgctg ccagcatggt 40380  
tcgtcgacc gcaggtgca gacggcgtc cgatttctg cctgcgaaga agacggcgg 40440  
50 gtcctcgag agccgggct cgccgagggc tcccgggct tccgcgtagc cggtagaccag 40500  
gtagccggg cgctcggcgg agccggcctg cagcggctgt accgggcatg cggagcgcag 40560  
55 ggcgtacac gtcgtagg ggttccggag gaagtgggg tcctctctc agtcagtc 40620  
ggctccaggc tccaacgggt tccggcggg catcgtcggc ggcacgcctc gcgtggcgg 40680  
cgtaggcgga ggccaccagg aggagccact ccgttccc cttcaggtt tacggtagg 40740  
60 cggcctccg cgatcgcgga gcaacgtgt cggagaactc ggcccaccg tcgcccgtt 40800  
ccgagtaacc ggttcgccc gacatgcgca tcaccagga gcggggtcgg gcccttact 40860  
65 ctgcccgcct tcagcttct ggacgtcgc ccagcccaa gagcccgtc tcgctccgt 40920



## ES 2 397 885 B1

acccccgtcg catggttcc atcgtgctgc tcatcgtca taccgtgccc ggcacaagct 40980  
cagaccaagg gtccgcaccg gcagcgcacc ctgcgcgatg gccacctggc gcaacctgc 41040  
5 catcagagcg ctccgctcgg ccggagtgag gaacatcgcg gccagtctgc ggcacaaggc 41100  
ccgtgaccg cgacggctcc tcgccgtccc cggcctcaac tgatcacaaa ccaggtgtca 41160  
10 cgcgctacgc cgaggccctg acaactacgc tgagtttcat gactgacgtt tctctcggcg 41220  
acgaactgac tcgcccggtt gaagacctg aagaacgagt ctctctgctc ttccaggaaac 41280  
accagccgc tgagacgaac aacggggtcg gaacagctga tcatgtagct gaagccgagg 41340  
15 acccgctca gctcgcacc aggtcctcag cacgttcgga tccggacgtg ttggcggctc 41400  
gctgaacgc gctcgggac ccggtgcgtc tgcgtatcct gctcgcctgt cttgacggcc 41460  
cgcgcagggc tgcggaactc gcagcgcaga ccgatatggg cagtacaggc cagatctacc 41520  
20 atcactgcg ccaactggtg aaccagggat ggcttcagc ctacgcaga ggccactatg 41580  
aagtaccccg cgaagccctt gaggttgtg ctgcggttct tgcggcgacg ttctgggacg 41640  
25 gcaacgccga ccccatcgc tcgaagctg cttcagggc tcgcacgtcc gctgcatcct 41700  
cgacgtgaac cccatgagt tcggaccgat acgcggtcgt ggaggccat cgggagacac 41760  
30 atggtgaggc catacatgtg ccgccccgc ctgcgggagc ggctgtccgg tccggggcag 41820  
cgcgaactc tacgcacctg cccgggcccgc gcagggcgac cacacttca gtgccgatcg 41880  
ggcatgccc atcgggtata ggcgttcggc catgcccta cgggcgaccg ggtgggtgc 41940  
35 ccttctgcgc gcgtaagga gcgtctgagt agggctttc ggtctgac ggacctggtc 42000  
atgatgtgcc cgcaaggta gctgttctg tcgaagtcgc gccgtgagcc gcgaccgaac 42060  
40 cgtctccgc gctcccggc tcgagcgca cttcagccgc gccatcgata ggggtaggaa 42120  
attgtcgag ggatctgcgc ctccacggt ccggacagcg gaccggctgc ccgacgat 42180  
tcaccggctg gtggagtccc agccgaacg tgcgccgac gccgtcggg tggagaccga 42240  
45 ggacgggaag ctacactacc gcgaactgga cgcgcgcgcc aaccagttcg cccgccacct 42300  
ccggtcggc ggcgtcggg gggagtccct cgtcgcctc cacatggagc ggggctcct 42360  
50 gacaccggtc gtctgctcg gcatcctcaa aacgggcgcg gcctacctgc cgctcgacac 42420  
cgaatcggc gcggagcgc tcgcgcggt gctcggcgc gcggcggcc cgcgggtgt 42480  
cacggccgc ccctgccgc cggtcgctg cccgtgatc gacctgaca ccgacctgcc 42540  
55 cgcatcacc gcactaccg ccgagcccct cacggacgtc gaggagccgg gcccgaccg 42600  
cctggcctac gtcatgtca cctccggctc caccggctc cccaagggcg tactggtgga 42660  
gcaccgcgc gtatccggc tgatccgga gcagagctac gcgaggctcg gccccgatc 42720  
60 caccatctg ctgctcgc cgctgcctt cgacgcctc acctggaga tctggggcg 42780  
cctggcgac ggccggcggc tcgtcgtc ggccccgga gcgcgcaccg tcgaccagct 42840  
65 cggccgacc ctgcggacc ggcgggtgac caccctgtg ctgaccgct cgcttcaa 42900

## ES 2 397 885 B1

cctcgtggtc gacgaggacc cgtccatcct cgccggggtc ggcgacctc tcacggcgg 42960  
cgaggcgctg tcggtcaacc atgtccggac ggcacggaag gcgctgccg acacggctgt 43020  
5 caccaacggc tacggcccga ccgagacgac cacgttcgcc tgcaccacg ccatccggcc 43080  
gcaggacctc gacggcgcgt cgatccgat cggcggggcc atgcccaca ccgaagtcca 43140  
10 tgtcctcgac gaggactcg accggctgc gcccgcgag gcgggtgaac tgtcatcgg 43200  
cggccccgc ctggcgcgc gctacctaa ccggccggc ctgaccgcg agcgcttct 43260  
cgcgcaccg gccgccacc agcccgggtc ccggctctac cgcaccggc accgggtcag 43320  
15 ggtgcggccc gacggcacc tggagtacct cgggcgcctg gacgaccagg tcaactgcg 43380  
cggcttcgg atcgaacccg gcgaggtcc cgccggactg accgggtgc cgcaggtcag 43440  
20 ggacgcggtc gtcgtcggc ggggcgggccc ctcggaccgg cgcctggtc cctacgtgt 43500  
gcccgaggcc gacgccacc ccggcatgga caacgagcgg gagcaggtc cgcactggga 43560  
ggcgggttc gacgagacct accgacgag tgtgggagcc gccgaaggac gctgggagct 43620  
25 gagcgggtg gtcggcagc gtgacggcct gccctcccc gccgaccaga tgcgggaatg 43680  
gacggacgcc accgtcgaac gcatccgac cctcggcgc cggcgcgtc tcgagatcg 43740  
30 ctgcggcacc ggctgctc cgatcggct cccccggac gcggaacgct atgtcggcag 43800  
cgacctgac gccgtggca tccgaggct gcgcgcacag atggacgccc cgggactcga 43860  
ccacaccgag ctggtgcac gcgccggcga cgacctgac gcggttccc gcggcacctt 43920  
35 cgacgtcgt gactgaact ccatcgtgca gtacctgcc tcggcgcagt atctgcgca 43980  
ggtcatcgag cgcgcggccc cacggctggc acccggcggg cacctgttc tcggagact 44040  
40 gcgcagctc ccgctcctc acgcctcca cctcagcgc gaactgaag gcggccacga 44100  
ggacgcgta cccctggcga cgctggcga ggcggtacgc gaacgggccc cggccgagaa 44160  
ggaactcgt gtcgtccct cctcttcac cgacctgag gcccgggcag gcatcgacca 44220  
45 cgtacgggtg acccgcggc gcggcaggca ccgcaacgag atgacgcagt tccgctacga 44280  
cgcggtgctc cgggtgcgc gagccgaacc gcgcgcgct cccgaccgt gctcactg 44340  
gcgggacgaa ggactcactc tggaggactc gcacggatc ctccagacc agcggccgca 44400  
50 gcacctggc ctgcgcggtg tcaccgacgc gcgcgtgcc gacgaggtg cccgcctcgt 44460  
ccggctgac gaggacgccc aagggaccgt ggccgcgctc cgggagacc ggcacgacg 44520  
55 cccggccgct gagatcgac acgtgtacga cctggcggcc cgcgcctct acaccgtgga 44580  
cgtcagcgt gcgggcagt gcggggcga tccttcgac gtcactctgt ggaccgacgc 44640  
cgaaccgga ccggtcgcgt tcgccccgg cccggcggaa gcacggggcc cccgaccag 44700  
60 catgccgct gcgaccgca cgtcgggca cctcagacg ctggtgcgc attcctcgc 44760  
cgagctcct cccgctaca tgatccggc ggtgtcgtc tcatggacg cgtgccgct 44820  
65 cacctccacc ggcaagatc accgctcgc cctgcccag ccgccgccc ggacctcgc 44880

## ES 2 397 885 B1

cggcggcgcg ggacggcggg ccgccaccgc gaccgaacgc gcactcgagc ccctgtggcg 44940  
cgatctgctc gccttgaaa ccgtccactg cgacgacgac ttcttcgac tcggcggcca 45000  
5 ctccctgctc ggcaccggcg tgctgtcccg cgtccggggc ctgtggggtg tcgaactctc 45060  
gctcggcgcc ctgttctccg ctcccacct cggcgcctc gccgcccga tcgactccgc 45120  
10 gcgccaggac acgcccggcc tccccggcac cctcgcggac aaggcggacc ccggatcggc 45180  
gccaccgctg tccccggcac agcaccggct ctggctcgtc gaacagctca ccccgggcaa 45240  
ccccgctac accgtgcccg tcgctacag gatgcgcgga ccgatcgaca cggcggcctc 45300  
15 ccaggccgcc ctgacaccc tcgtcggccg acacgaggtg ctgcgacca ccttcccctc 45360  
ccacgacggc accccccgcc aggtcgtcgc cccctccgga cgcattccca tcgaacgggc 45420  
20 cgacgtcggc ggcgagggtg ccgacgcccc cggcggccgc cacaacatcc tgaccggca 45480  
ggcgagccgc tggtcgcagc tgcagtccgg ccccctggcc gccgccacc tcgtccggct 45540  
cgccgaggac gaccacgtcc tgtcctgac tctccatcac atgatctcgc acggctggtc 45600  
25 cctggacctc ctgcggccg aactgagcga gggctacaac gcccgctcg cccgcccggc 45660  
accgcagctg ccgggatcc accaccacac caggaccgga taccggaccg gacaccccct 45720  
30 ccaacagcac cacctcacgg aaccggcctc catacgcagc ctgagtcacc accagccca 45780  
caccggcatc cgcaaatg aactgacac ggtccgccc atagccagga tottgccgag 45840  
ccccagcagc accacgtgcc ccccgaggcc tcggggcggg cggcgcagcg aggaggcgg 45900  
35 gcgcagggtg ccgaacgcct ccagaacggc ggcgatcagc agcggcagca ggagcagacc 45960  
ggcgatgccg gagaggatcc gctccacca gctcgtccc gtgcaggcct cgctcgaagc 46020  
40 ggcgttcagc gaactacca ggacgagcgt cgaataccag gcgcgcacct ccaccggggc 46080  
accggacgcg gaggctccct catgaagccc gctgcttcc cgcggcgcg tccggagca 46140  
cgcgccgcg cctgacct cggcggcatg ctccgctcg agtgagcga gatctacgag 46200  
45 ggcgaggtcc gctacgccc cgacgcctc gaggcctgc gcctcatgga cgcgctcgc 46260  
ggcgtcaagc gcggcgtgcc gggcgcggcc ctgcccgagc tgaagcagc cgggtccc 46320  
50 aagaaggatg tgccgtcct ggaggtcagc gagcccagg gctcgggtcg ctggacgtc 46380  
tccatcaca accgatcgc cgcgctcgg ctggcgaca tccggggcgc cgcccagcag 46440  
tcgaaccgga tgctggaccg taccgtccg tcgctggcct ccggcgagg cgacgcccc 46500  
55 gcccggtg gctcctcgt ggtcagctg cgcgcacgg tggaggacct cgaccgccc 46560  
gacaccccgg gcgcggggc gcggcgctg ctgtcggc tgccgggccc gcggcggtg 46620  
60 cgcgacctg tcgcaagta cgcctccc aaggccacc tcgaccgat cc

<210> 2

<211> 562

<212> Proteína

65 <213> Artificial

&lt;220&gt;

&lt;223&gt; Orf1

&lt;400&gt; 2

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

ES 2 397 885 B1

485            490            495  
 Lys Tyr Leu Arg His Gly Phe Gly Ala Gly Gln Glu Pro Val Thr Phe  
           500            505            510  
 5 Ala Arg Tyr Thr Gly Gln Glu Ser Glu Glu Arg Arg Lys Glu Leu Arg  
           515            520            525  
 Lys Asn Pro Pro Asp Ile Leu Leu Thr Asn Tyr Val Met Leu Glu Leu  
           530            535            540  
 10 Met Leu Thr Arg Pro Asp Asp Arg Ser Ser Leu Ile Arg Gln Glu Arg  
       545            550            555            560  
 Val Val

15 <210> 3  
     <211> 185  
     <212> Proteína  
     <213> Artificial

20 <220>  
     <223> Orf2  
     <400> 3  
 Leu Asp Ala Ala Glu Gly Val Phe Ala Glu Arg Gly Ile Asp Ala Ala  
   1            5            10            15  
 25 Arg Ile Asp Glu Ile Thr Glu Arg Ala Gly Val Ala Ile Gly Thr Phe  
           20            25            30  
 Tyr Leu His Phe Ser Ser Lys Arg Asp Ile Met Ala Ala Val Gln Ala  
           35            40            45  
 Arg Phe Val Asp Arg Leu Val Glu Arg Gln His Ala Ala Ala Gln Asn  
           50            55            60  
 30 Leu Pro Ala Asp Asp Trp Ile Gly Arg Val Asp Ala Trp Leu Ser Asp  
       65            70            75            80  
 Ala Val Arg Ile Tyr Val Glu His Ala Gln Leu His Asp Val Leu Tyr  
           85            90            95  
 35 Gly His Thr Pro Ile Asn Ala Thr Thr Glu Ile Gln Val Ser Pro Glu  
           100            105            110  
 Asn Ala His Val Glu Ala Phe Arg Arg Leu Ile Ala Glu Arg Pro Asp  
           115            120            125  
 Pro Pro Ala Asp Ala Pro Asn Pro Ala Met Ala Ala Leu Leu Ile Tyr  
           130            135            140  
 40 Ser Ala Trp Tyr Gly Gly Thr His Ala Leu Leu His His Glu Ala Asp  
       145            150            155            160  
 Asp Ile Asp Arg Leu Ala Asp Glu Leu Ile Ala Asp Ile Thr Asp Leu  
           165            170            175  
 45 Ala His Arg Tyr Leu Arg Ser Glu Arg  
           180            185

50 <210> 4  
     <211> 577  
     <212> Proteína  
     <213> Artificial

55 <220>  
     <223> Orf3  
     <400> 4

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

65            70            75            80  
 Leu Ala Ala Gln Leu Arg Arg Arg Leu Met His His Val Ser Thr Leu  
                  85            90            95  
 5    Thr Leu Gly Trp Phe Thr Ala Glu His Lys Ala Arg Leu Ala Arg Ala  
                  100           105           110  
 Val Thr Ala Asp Val Gly Ser Ala Ala His Leu Ala Val Thr Ile Gly  
                  115           120           125  
 Gly Pro Val Ile Thr Ser Thr Leu Leu Pro Ala Thr Val Val Val Val  
                  130           135           140  
 10    Thr Phe Thr Val Asp Trp Arg Met Ala Leu Leu Leu Cys Ala Ile Ala  
                  145           150           155           160  
 Leu Val Ala Tyr Leu Ala Leu Arg Arg Ala Ala Arg Ile Ser Glu Ile  
                  165           170           175  
 15    Ala Glu Ile Glu Leu Glu Arg Ala Ala Ala Gly Val Ala Ser Arg Ala  
                  180           185           190  
 Ile Glu Leu Gly Gln Ala Gln Pro Val Leu Arg Ala Ala Gly His Ala  
                  195           200           205  
 Glu Gly Thr Pro Arg Met Arg Ala Ala Leu Glu Asp His Arg Glu Thr  
                  210           215           220  
 20    Tyr Arg Glu Gly Met Arg Arg Ala Arg Gln Pro Phe Phe Val Tyr Thr  
                  225           230           235           240  
 Gly Val Ile Met Ala Gly Phe Val Ala Val Leu Ala Leu Thr Ala Glu  
                  245           250           255  
 25    Leu Val Leu Ser Gly Arg Ile Asp Ala Ala Thr Ala Ile Val Leu Leu  
                  260           265           270  
 Val Leu Ala Ala Arg Phe Leu Glu Pro Leu Gly Asn Leu Ile Glu Leu  
                  275           280           285  
 30    Ile Gly Ala Leu Arg Ala Leu Gly Asn Gln Ile Ala Arg Ile Glu Glu  
                  290           295           300  
 Leu Leu Ala Thr Glu Ala Leu Pro Val Pro Ala Glu Pro Val Arg Arg  
                  305           310           315           320  
 Ile Asp Glu Ala Glu Val Glu Phe Ala Asp Val Thr Phe Thr Tyr Pro  
                  325           330           335  
 35    Gly Gly Asp Thr Pro Ala Leu Arg Asn Val Ser Leu Arg Cys Pro Ala  
                  340           345           350  
 Gly Ser Thr Thr Ala Leu Val Gly Pro Ser Gly Ser Gly Lys Thr Thr  
                  355           360           365  
 Ala Thr Arg Leu Ile Ala Arg Phe Phe Asp Ile Asp Ser Gly Glu Leu  
                  370           375           380  
 40    Arg Ile Gly Gly Val Asp Val Arg Lys Leu Asp Pro Thr Thr Leu Leu  
                  385           390           395           400  
 Asp Glu Ile Ala Ile Val Phe Gln Asp Val Tyr Leu Phe Asp Asp Thr  
                  405           410           415  
 45    Ile Glu Asp Asn Leu Arg Leu Ala Arg Pro Glu Ala Thr Trp Asp Glu  
                  420           425           430  
 Leu Arg Glu Ala Ala Thr Ala Ala Arg Leu Asp Glu Val Ile Glu Arg  
                  435           440           445  
 Leu Pro Ala Gly Trp Gln Thr Arg Val Gly Glu Gly Gly Ala Gln Leu  
                  450           455           460  
 50    Ser Gly Gly Glu Arg Gln Arg Val Ala Ile Ala Arg Ala Met Leu Lys  
                  465           470           475           480  
 Arg Ala Arg Ile Val Leu Val Asp Glu Ala Ser Ser Ala Leu Asp Pro  
                  485           490           495  
 55    Glu Asn Glu Ala Ala Ile Thr Arg Ala Ile Ala Asn Leu Gly Ala Asp  
                  500           505           510  
 Pro Asp Arg Thr Val Ile Val Ile Ala His Arg Pro Ala Thr Leu Glu  
                  515           520           525  
 Ala Ala Asp Leu Val Val Ser Leu Asp Gly Gly Arg Val Val Glu Ser  
                  530           535           540  
 60    Gly Thr Arg Glu Glu Leu Leu Arg Thr Gly Gly Thr Phe Ala Arg Leu  
                  545           550           555           560  
 Ser Ser Gln Tyr Glu Arg Ala Arg His Trp Arg Ile Ala Ser Gly Ser  
                  565           570           575  
 65    His

<210> 5  
 <211> 586  
 <212> Proteína  
 <213> Artificial  
 5  
 <220>  
 <223> Orf4  
 <400> 5  
 10  
 Met Asn Arg Ser Thr Ser Gln Thr Pro Gly Leu Ala Gly Pro Gly Ala  
 1 5 10 15  
 Leu Met Arg Leu Leu Ala Pro Ile Arg Thr His Leu Val Ile Cys Ala  
 20 25 30  
 15  
 Val Leu Ser Cys Leu Ser Ala Ala Ala Gly Ile Val Pro Tyr Ile Ala  
 35 40 45  
 Val Ala Glu Ile Ala Arg Leu Met Leu Asp Asp Pro Ala Gly Ser His  
 50 55 60  
 20  
 Thr Ala Ile Trp Thr Trp Val Gly Val Gly Ala Ala Gly Ala Gly Val  
 65 70 75 80  
 Trp Leu Val Leu Phe Val Gln Ser Ser Arg Val Gly His Tyr Ala Asp  
 85 90 95  
 Ala Ala Ile Leu His Asp Val Arg Val Arg Ile Val Thr His Leu Gly  
 100 105 110  
 25  
 Lys Leu Pro Leu Gly Trp Phe Arg Ala Val Gly Ser Gly Lys Val Lys  
 115 120 125  
 Arg Ala Met Thr Gly Asp Leu Glu Glu Met His Glu Val Ile Ala His  
 130 135 140  
 30  
 Ala Leu Gly Gln Leu Val Gly Ala Val Thr Val Met Thr Val Ala Thr  
 145 150 155 160  
 Gly Tyr Leu Leu Leu Val Asp Val Arg Met Ala Phe Val Val Leu Gly  
 165 170 175  
 Val Leu Leu Leu Met Gly Ile Phe Phe Arg Val Ser Met Arg Ser Met  
 180 185 190  
 35  
 Thr Thr His Met Asn Arg Leu Val Leu Ala Asp Ala Arg Ile Ser Ala  
 195 200 205  
 Ala Ser Val Glu Tyr Ala Asp Gly Ile Gln Val Val Lys Thr Phe Gly  
 210 215 220  
 40  
 Thr Gly Gly Arg Val Leu Arg Arg Phe Asp Asp Ala Val Asp Glu His  
 225 230 235 240  
 Thr Gln Ala Phe Ala Ala Trp Val Ala Glu Val Arg His Ser Ser Ala  
 245 250 255  
 Ala Ser Arg Leu Phe Gly Ser Glu Met Ala Val Leu Thr Ala Val Ala  
 260 265 270  
 45  
 Ala Val Gly Leu Val Leu Val Asp Arg Gly Ser Leu Gly Val Ala Asp  
 275 280 285  
 Leu Val Ala Phe Leu Val Val Ala Ile Gly Leu Pro Asn Ser Ile Leu  
 290 295 300  
 50  
 Pro Ala Val Thr Ala Ala Gln Gly Val Arg Lys Gly Arg Met Gly Ala  
 305 310 315 320  
 Ala Asn Ile Glu Gln Leu Leu Ala Arg Thr Pro Leu Pro Glu Pro Arg  
 325 330 335  
 Glu Pro Gln Glu Pro Val Gly His Gly Val Glu Phe Asp Arg Val Ser  
 340 345 350  
 55  
 Phe Ser Tyr Asp Gly Val Thr Asn Ala Val Glu Asp Ile Ser Ala Val  
 355 360 365  
 Cys Pro Pro Gly Arg Val Thr Ala Ile Val Gly Pro Ser Gly Ala Gly  
 370 375 380  
 60  
 Lys Thr Thr Leu Ala Gly Leu Leu Pro Arg Phe Tyr Glu Val Ser Arg  
 385 390 395 400  
 Gly Ser Ile Arg Ile Gly Gly Val Asp Val Arg Ser Ile Pro Ser Ala  
 405 410 415  
 Lys Leu Leu Ala Ser Met Ser Leu Val Phe Gln Asp Val Ala Leu Leu  
 420 425 430  
 65  
 Arg Asp Ser Val Ala Glu Asn Ile Arg Ile Gly Arg Pro Gly Ala Ser  
 435 440 445

Asp Ala Glu Val Arg Glu Ala Ala Ala Ala Ala His Ile His Asp Val  
 450 455 460  
 Ile Glu Ser Met Pro Asp Gly Tyr Asp Thr Leu Leu Asp Ala Gly Gly  
 465 470 475 480  
 5  
 Gly Ser Leu Ser Gly Gly Glu Arg Gln Arg Leu Thr Ile Ala Arg Ala  
 485 490 495  
 Ile Leu Ser Gly Ala Pro Ile Val Val Leu Asp Glu Ala Thr Ala Ser  
 500 505 510  
 10  
 Leu Asp Ala Asp Ser Glu Ser Ala Val Gln Glu Ala Leu Ala Asn Leu  
 515 520 525  
 Ala Val Gly Lys Thr Val Ile Val Ile Ala His Arg Leu His Thr Ile  
 530 535 540  
 Ala Gly Ala Ala Gln Ile Leu Val Leu Glu Asn Gly Arg Leu Val Glu  
 15  
 545 550 555 560  
 Gln Gly Arg His Glu Glu Leu Leu Ala Arg Asp Gly Leu Tyr Ala Arg  
 565 570 575  
 Met Trp Ala Ala Gln Glu Gly Val Leu Ala  
 580 585  
 20  
 <210> 6  
 <211> 322  
 <212> Proteína  
 <213> Artificial  
 25  
 <220>  
 <223> Orf5  
 <400> 6  
 30  
 Leu Ile Ala Leu Ala Ala Val Gly Ala Thr Gly Cys Gly Gly Gly Asp  
 1 5 10 15  
 Ala Asp Ala Arg Gln Lys Gly Glu Lys Lys Thr Ile Ala Phe Glu Ser  
 20 25 30  
 35  
 Cys Ser Arg Thr Val Glu Leu Asp Arg Ile Pro Lys Arg Val Ala Ile  
 35 40 45  
 Thr Thr Asp Ala Ile Ala Asp Thr Leu Phe Glu Leu Gly Val Gly Asp  
 50 55 60  
 Arg Ile Val Ala Lys Thr Arg Gly Glu Ser Ala Pro Ala Pro Glu Leu  
 40  
 65 70 75 80  
 Lys Glu Arg Leu Ala Ala Leu Pro Ser Leu Gly Thr Arg Asn Pro Ser  
 85 90 95  
 Val Glu Ala Leu Val Gly Ala Lys Pro Asp Leu Leu Ile Thr Asp Gln  
 100 105 110  
 45  
 Val Glu Lys Val Ser Gly Lys Leu Gly Ser Pro Ser Ile Ala Glu Leu  
 115 120 125  
 Glu Arg Leu Gly Ile Ala Thr Tyr Val Val Gly Gly Gly Cys Ala Ala  
 130 135 140  
 Asp Leu Ser Glu Asp Thr Ser Gly Leu Glu Ala Leu Asp Gly Asp Ile  
 50  
 145 150 155 160  
 Arg Gln Leu Gly Thr Val Phe Gly Val Glu Ala Arg Ala Arg Lys Leu  
 165 170 175  
 Ala Asp Lys Leu Asn Gly Ser Leu Asp Asp Val Arg Arg Gln Thr Ala  
 180 185 190  
 55  
 Gln Glu Pro Arg Thr Lys Val Ala Lys Leu Ser Gln Val Ala Gly Gln  
 195 200 205  
 Leu Tyr Val Thr Ser Gly Gly Leu Ser Asp Asp Val Ile Glu Arg Ala  
 210 215 220  
 Gly Gly Thr Asn Val Phe Ala Asp Leu Pro Gly Gln Phe Ala Pro Val  
 60  
 225 230 235 240  
 Ser Pro Glu Gln Ile Val Ala Arg Asp Pro Gln Ser Ile Ile Val Asp  
 245 250 255  
 Asn Phe Thr Ala Thr Ala Ala Gly Glu Asn Glu Ala Met Ala Tyr Leu  
 260 265 270  
 65  
 Lys Arg Thr Phe Pro Thr Val Glu Ala Val Lys Lys Gln Arg Val Leu  
 275 280 285



Val Ile Asp Ala Ala Lys Ser Gly Ala Arg Gly Ser Thr Arg Pro Val  
 290 295 300  
 Glu Gly Val Val Glu Ile Ala Arg Phe Leu His Pro Ser Thr Ser Arg  
 305 310 315 320  
 5 Ser Gln  
  
 <210> 7  
 <211> 327  
 10 <212> Proteína  
 <213> Artificial  
  
 <220>  
 <223> Orf6  
 15  
 <400> 7  
 Leu Phe Val Val Leu Leu Gly Ala Val Cys Leu Ala Val Ser Val Gly  
 1 5 10 15  
 Ser Val Gly Val Arg Pro Ala Thr Val Val Lys Val Ile Ala Asp His  
 20 20 25 30  
 20 Leu Ile Gly Val Gly Thr Pro Ala Ala Val Met Asp Asp Gln Ile Val  
 35 40 45  
 Trp Asn Leu Arg Leu Pro Arg Val Leu Leu Ala Ala Ala Val Gly Gly  
 50 55 60  
 25 Gly Leu Ala Val Val Gly Val Thr Leu Gln Ala Thr Val Arg Asn Pro  
 65 70 75 80  
 Leu Ala Asp Pro Tyr Val Leu Gly Val Ser Ala Gly Ala Gly Leu Met  
 85 90 95  
 Ala Ser Ile Val Ile Thr Leu Gly Ser Val Ala Val Ala Gly Leu Ser  
 30 100 105 110  
 Thr Ser Ala Ala Ala Phe Val Gly Ala Leu Val Ala Met Val Ala Val  
 115 120 125  
 Leu Gly Leu Ser Arg Arg Ala Gly Arg Val Ile Pro Ser Arg Leu Leu  
 130 135 140  
 35 Leu Ala Gly Val Thr Leu Ser Tyr Leu Phe Ser Gly Ala Thr Ser Phe  
 145 150 155 160  
 Val Ile Phe Arg Ser Gly Asn Ala Asp Ala Ala His Ser Val Leu Phe  
 165 170 175  
 Trp Leu Leu Gly Ser Leu Ser Glu Ala Ser Trp Ser Asn Leu Ser Leu  
 40 180 185 190  
 Pro Ala Ala Ala Val Leu Val Val Gly Val Tyr Leu Met Leu Gln Ala  
 195 200 205  
 Arg Thr Leu Asn Ala Leu Ala Ala Gly Asp Asp Ala Ala Leu Ser Leu  
 210 215 220  
 45 Gly Val Ala Val His Arg Val Arg Ile Arg Leu Leu Val Val Ala Ser  
 225 230 235 240  
 Leu Leu Thr Gly Val Leu Val Ala Val Ser Gly Gly Ile Gly Phe Val  
 245 250 255  
 Gly Leu Ile Val Pro His Leu Val Arg Leu Val Met Gly Pro Asp His  
 50 260 265 270  
 Arg Arg Leu Leu Pro Val Ala Met Leu Val Gly Ala Val Tyr Leu Val  
 275 280 285  
 Val Val Asp Leu Leu Cys Arg Val Leu Val Arg Pro Glu Glu Leu Pro  
 290 295 300  
 55 Ile Gly Ile Val Thr Ala Val Leu Gly Ala Pro Val Phe Leu Trp Leu  
 305 310 315 320  
 Leu Arg Arg Ser Glu Ala Ser  
 325  
 60  
  
 <210> 8  
 <211> 261  
 <212> Proteína  
 65 <213> Artificial

<220>

<223> Orf7

<400> 8

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

40

<210> 9

<211> 168

<212> Proteína

<213> Artificial

45

<220>

<223> Orf8

<400> 9

50

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

Val Lys Ala His Leu His Met Val Tyr His Lys Leu Gly Val Ala Asp  
 130 135 140  
 Arg Thr Glu Ala Val Val Ile Ala Leu Arg Gln Gly Leu Val Gln His  
 145 150 155 160  
 5 Gln Arg Arg Glu Asp Pro Pro Ser  
 165  
  
 <210> 10  
 10 <211> 342  
 <212> Proteína  
 <213> Artificial  
  
 <220>  
 15 <223> Orf9  
  
 <400> 10  
 Val Ile Ser Ala Ile Phe Gly Thr Leu Ala Thr Gln Ala Val Gly Ala  
 1 5 10 15  
 20 Ala Ala Arg Leu Glu Leu Ala Asp Arg Ile Gly Glu Ser Gly Ala Asp  
 20 25 30  
 Thr Asp Glu Leu Ala Leu Ala Cys Gly Val Pro Ala Glu Gln Leu Gly  
 35 40 45  
 25 Arg Leu Leu Arg Ala Leu Ala Ser Leu Gly Leu Cys Val Glu Ser Glu  
 50 55 60  
 Pro Gly Arg Phe Ala Leu Thr Glu Ala Gly Ala Leu Leu Arg Arg Asp  
 65 70 75 80  
 Asp Pro Ala Ser Leu Leu Ala Phe Ala Ala Phe Leu Thr His Asp Val  
 85 90 95  
 30 Phe Gln Arg Asn Trp Leu Asn Leu Gln Glu Ser Leu Asp Thr Gly Leu  
 100 105 110  
 Pro Ala Phe Asp Thr Ala Phe Gly Leu Pro Val Tyr Asp Tyr Leu Ser  
 115 120 125  
 35 Gly Arg Pro Glu Leu Ala Ala Leu Phe His Ala Ala Met Ser Arg Arg  
 130 135 140  
 His Arg Pro Leu Glu Met Ala Ala Ala Ile Ser Ala Val Tyr Asp Leu  
 145 150 155 160  
 Gly Arg Phe Ser Thr Val Val Asp Val Gly Gly Gly Asp Gly Thr Leu  
 165 170 175  
 40 Leu Ala Ala Phe Leu Asp Arg Tyr Pro His Leu Thr Gly Thr Val Leu  
 180 185 190  
 Glu Thr Glu Ala Gly Ala Ala Arg Ala Arg Glu Thr Ile Ala Gly Ser  
 195 200 205  
 45 Gly Leu Thr Asp Arg Cys Arg Ala Val Ala Gly Asp Phe Phe Ala Glu  
 210 215 220  
 Val Pro Lys Gly Ala Asp Leu Tyr Leu Ile Lys Asn Val Val Leu Asn  
 225 230 235 240  
 Trp Asp Asp Glu Arg Ala Trp Thr Ile Leu Arg Arg Val Arg Asp Ala  
 245 250 255  
 50 Met Pro Asp His Gly Lys Leu Leu Ile Ala Glu Pro Val Leu Pro Asp  
 260 265 270  
 Thr Ala Asp Ala Asp Ser Leu Asn His Ala Ala Leu Glu Asn Pro Tyr  
 275 280 285  
 55 Leu Thr Asp Leu His Met Leu Val Thr Ile Gly Gly Arg Gln Arg Thr  
 290 295 300  
 Arg Ala Glu Tyr Thr Ala Ile Cys Ala Arg Ala Gly Leu Arg Val Thr  
 305 310 315 320  
 Glu Val Val Pro Leu Ala Gln Glu Leu Asn Ala Ser Leu Ile Glu Val  
 325 330 335  
 60 Val Pro Asp Ile Ala Ala  
  
 340  
  
 65 <210> 11  
 <211> 532

<212> Proteína  
 <213> Artificial

<220>  
 <223> Orf10

<400> 11

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

Ala Asp Val Asp Leu Ala Asp Pro Glu Trp Asn Thr Ser Gly Val Pro  
 465 470 475 480  
 Trp His Thr Leu Tyr Tyr Gly Asp Asn Tyr Pro Arg Leu Gln Glu Val  
 485 490 495  
 5 Lys Ala Glu Trp Asp Pro Leu Asp Ile Phe His His Ala Leu Ser Ile  
 500 505 510  
 Ser Ala Pro Glu Ala Gly Arg Pro Glu Ser Gly Pro Arg Thr Pro Ala  
 515 520 525  
 10 Gln Arg Glu Trp  
 530

<210> 12  
 <211> 544  
 15 <212> Proteína  
 <213> Artificial

<220>  
 <223> Orf11  
 20

<400> 12  
 Leu His Ser Ser Thr Ala Val Arg Asn Ser Thr Ala Ile Arg Lys Arg  
 1 5 10 15  
 25 His Arg Met Ser Ser Pro Ala Gly Gly Pro Val Phe Ala Asp Arg Val  
 20 25 30  
 Leu Asn Asp Trp Leu Ser Ser Val Gly Leu Gly Ile Glu Tyr Ser Arg  
 35 40 45  
 Ala Glu Ala Asn Thr Leu Tyr Tyr Val Asp Glu Asp Gly Gln Glu Val  
 50 55 60  
 30 Pro Val Leu Asp His Val Cys Gly Phe Gly Ser Leu Ile Phe Gly His  
 65 70 75 80  
 Asn Asn Pro Asp Ile Val Ala His Ala Lys Thr Val Leu Asp Arg Gln  
 85 90 95  
 35 Val Pro Val Tyr Val Gln Leu Ser His Gln Ser His Ala Asn Asp Ile  
 100 105 110  
 Ala Ala Val Leu Asn Ala Ile Leu Arg Arg Glu Ile Pro Gly Gly Asp  
 115 120 125  
 Gly Asp Tyr Ala Ala Ile Phe Ala Asn Ser Gly Ala Glu Ala Val Glu  
 130 135 140  
 40 Ile Cys Val Lys His Ala Glu Leu Glu Arg Arg Ala Arg Val Ala Lys  
 145 150 155 160  
 Leu Thr Asp Glu Thr Ser Arg Asn Ala Glu Glu Ala Arg Glu Ala Val  
 165 170 175  
 45 Gly Ala Gly Arg Ala Thr Val Ala Glu Asn Pro Tyr Val Arg Phe Asp  
 180 185 190  
 Gly Ala Ser Gly Glu Ser Asp Arg Leu Glu Gln Leu Leu Ala Glu Ala  
 195 200 205  
 Gly Arg Arg Asn Ala Glu Leu Ile Ala Arg Gly Pro Val His Leu Ala  
 210 215 220  
 50 Leu Glu Asn Ala Phe His Gly Lys Leu Val Ala Ser Ile Gln Leu Thr  
 225 230 235 240  
 Gln Asn Pro His Trp Arg Leu Pro Phe Thr Ser Leu Ala Ser Ser Thr  
 245 250 255  
 55 Arg Phe Leu Arg Ala Asp Arg Pro Asp Glu Met Lys Ala Ala Val Glu  
 260 265 270  
 Glu Leu Arg Thr Ser Leu Leu Asp Val Leu Val Asp Asp Gly Val Val  
 275 280 285  
 Thr Val Val Glu Arg Asp Phe Pro Leu Val Gly Ala Phe Phe Val Glu  
 290 295 300  
 60 Pro Val Gln Gly Ala Asn Gly Met Arg Pro Leu Thr Glu Ala Ala Ala  
 305 310 315 320  
 Arg Glu Ile Arg Ala Val Cys Asp Ala Val Gly Cys Pro Leu Ile Val  
 325 330 335  
 65 Asp Glu Ile His Ser Gly Met Gly Arg Thr Gly Ala Phe Leu Ala Ser  
 340 345 350  
 Ser His Met Gly Leu Arg Gly Asp Tyr Tyr Thr Leu Ala Lys Ser Ile

355            360            365  
 Gly Gly Gly Ile Ala Lys Asn Ala Val Ala Leu Phe His Arg Asp Arg  
 370            375            380  
 5 Phe Arg Pro Glu Phe Glu Val Met His Ser Ser Thr Phe Ala Lys Asp  
 385            390            395            400  
 Gly Phe Ser Ala Ala Ile Ala Leu Lys Val Leu Glu Met Leu Glu Ala  
 405            410            415  
 Asp Asp Gly Arg Ala Tyr Arg Ile Ala Ala Glu Arg Gly Asp Arg Leu  
 420            425            430  
 10 Lys Gly Ala Leu Thr Ala Val Ala Ala Asp Phe Pro Asp Val Val Asp  
 435            440            445  
 Ala Val His Gly Ile Gly Leu Met Leu Ala Val Glu Phe Lys Asp Gln  
 450            455            460  
 15 Lys Gly Ala Ser Ser Glu Pro Leu Arg Glu Lys Ala Ala Ser Gly Met  
 465            470            475            480  
 Leu Gly Tyr Phe Ile Ala Gly Leu Ile Leu Arg Glu His Arg Ile Arg  
 485            490            495  
 Val Leu Pro Val Gly Pro Ala Gly Asn Ser Val Arg Phe Glu Pro Ser  
 500            505            510  
 20 Ile Tyr Leu Thr Asp Ala Asp Ile Ala Arg Thr Glu Asn Ala Leu Arg  
 515            520            525  
 Asp Val Cys Thr Ile Leu Arg Asp Gln Asp Gly Asp Arg Leu Thr Pro  
 530            535            540  
  
 25 <210> 13  
 <211> 328  
 <212> Proteína  
 <213> Artificial  
  
 30 <220>  
 <223> Orf12  
  
 35 <400> 13  
  
 Met Glu His Ala Leu Ser Ala Ala Cys Ala Ala Ser Val Arg Ala Ala  
 1            5            10            15  
 Val Thr Leu Gly Leu Pro Glu Ala Leu Gly Glu Ala Pro Ala Thr Ala  
 20            25            30  
 40 Asp Glu Leu Ala Thr Ala Val Gly Ala Asp Pro Gly Ala Leu Arg Arg  
 35            40            45  
 Leu Leu Arg Ser Leu Thr Ser Tyr Gly Val Phe Ala Glu Glu Gly Thr  
 50            55            60  
 45 Gly Ser Phe Val His Thr Glu Lys Ser Arg Ala Leu Arg Glu Asp Ser  
 65            70            75            80  
 Pro Asp Ser Ile Lys Tyr Leu Val Leu Trp Cys Thr Glu Pro Trp Leu  
 85            90            95  
 50 Trp Ser Leu Trp Gly Asp Leu Asp Glu Ser Val Arg Thr Gly Gly Glu  
 100            105            110  
 Ile Phe Thr Arg Thr His Gly Arg Arg Phe Tyr Glu His Leu His Thr  
 115            120            125  
 Arg Trp Pro Glu Ser Ala Arg Ile Phe Asn Arg Ala Met Thr Gln Gln  
 130            135            140  
 55 Thr Arg Leu Ser Ala Thr Val Ile Ala Asp Met Leu Pro Met Ser Gly  
 145            150            155            160  
 Ala Gly Thr Val Ala Asp Val Gly Gly Gly Gln Gly Leu Val Leu Gly  
 165            170            175  
 60 Thr Leu Leu Glu Arg His Pro His Val Arg Gly Phe Leu Leu Asp Leu  
 180            185            190  
 Pro Glu Val Val Ala Asn Val Asp Ala Arg Leu His Pro Gly Gly Glu  
 195            200            205  
 Leu Ala Asp Arg Val Arg Leu Val Pro Gly Asp Cys Leu Glu Gly Ile  
 210            215            220  
 65 Ser Val Glu Ala Asp Val Tyr Leu Phe Lys Asn Ile Leu Gly Ala Asp  
 225            230            235            240

Asp Asp Thr Ser Val Arg Ile Leu Arg Asn Ala Met Lys Ala Ala Arg  
                   245                  250                  255  
 Pro Gly Ala Arg Met Val Ile Val Glu Asn Phe Val Asp Asp Gly Pro  
                   260                  265                  270  
 5 Gly Glu Arg Leu Ala Ser Ala Leu Asp Leu Arg Met Leu Leu Val Ile  
                   275                  280                  285  
 Gly Gly Gln Lys His Thr Arg Ala Gly Leu Leu Gly Ile Ala Glu Arg  
                   290                  295                  300  
 10 Ala Gly Leu Thr Val Arg Asp Val Arg Pro Val Asp Ser Ser Leu His  
                   305                  310                  315                  320  
 Met Ile Glu Thr Val Val Pro Gly  
                   325  
  
 15 <210> 14  
     <211> 462  
     <212> Proteína  
     <213> Artificial  
  
 20 <220>  
     <223> Orf13  
  
     <400> 14  
 25 Val Arg Ala Ser Ser Ser Ala Pro Pro Ala Leu Asp Ala Leu Gly Thr  
     1          5                  10                  15  
 Gly Gly Pro Tyr Arg Ser Arg Asn Leu Glu Val Val His Asp Val Arg  
           20                  25                  30  
 Gly Glu Pro Val Ala Glu Leu Ser Leu Val Pro Trp Leu Phe Ala Gln  
           35                  40                  45  
 30 Arg Ser Ile Arg Ala Leu Arg Arg Ala Ala Pro Pro Asp Pro Glu Arg  
     50                  55                  60  
 Arg Lys Lys Leu Leu Thr Arg Ala Gly Trp Leu Phe Ala Ser Gly Ser  
     65                  70                  75                  80  
 Val Asp Gly Val Thr Ala Ala Ala Tyr His Arg Thr Val Ala Glu Val  
           85                  90                  95  
 35 Ser Gly Ile Pro Ile Ala Thr Val Arg Gln Ser Ala Gln Gln Val Gly  
           100                  105                  110  
 Asp Tyr Ala Ala Thr Ala Tyr Glu Asn Val Arg Gln Ala Arg Pro Val  
           115                  120                  125  
 40 Gly Ala Ala Asp Ser Trp Arg Asp Gln Arg Ala Arg His Gly Ser Gly  
     130                  135                  140  
 Val Trp Thr Arg Arg Gly Glu Val Leu Met Val His Ala Pro Ala Asn  
     145                  150                  155                  160  
 Ser Pro Ala Val His Thr Ser Trp Leu Asp Ala Leu Ala Leu Gly Tyr  
     165                  170                  175  
 45 Arg Val Ala Val Arg Pro Ser Gln Arg Glu Pro Phe Thr Ala His Arg  
           180                  185                  190  
 Leu Val Ser Ala Leu Arg Gln Val Gly Tyr Asp His Asp Gln Val Val  
           195                  200                  205  
 50 Leu Leu Pro Thr Asp His Ala Thr Ala Asp Arg Leu Val Ala Glu Ala  
     210                  215                  220  
 Asp Val Ala Ile Ala Phe Gly Gly Asp Glu Val Ala Arg Lys Tyr Gly  
     225                  230                  235                  240  
 Ser Ser Thr Leu Val Met Pro Phe Gly Pro Gly Arg Ser Lys Ile Leu  
           245                  250                  255  
 55 Leu Thr Ser Gly Thr Asp Val Gly Arg His Leu Ala Thr Ile Thr Glu  
           260                  265                  270  
 Ser Ile Ala Gly His Ala Gly Thr Gly Cys Val Asn Ala Thr Ala Val  
           275                  280                  285  
 60 Phe Val Glu Gly Asp Pro Glu Pro Val Ala Glu Ala Ile Ala Gln Arg  
     290                  295                  300  
 Leu Gly Glu Leu Pro Ser Leu Pro Pro Gly Asp Glu Arg Ala Arg Leu  
     305                  310                  315                  320  
 Pro Val Arg Arg Ala Ala Val Ala His Glu Met Asp Ala Tyr Leu Arg  
           325                  330                  335  
 65 Ala Gly Ala Gly Asp Ala Arg Pro Leu Leu Gly Ala Asp Thr Val Val

ES 2 397 885 B1

340 345 350  
 Ala Glu Leu Gly Asp Gly Ser Ala Val Leu Arg Pro Ala Val Phe Leu  
 355 360 365  
 5 Leu Asp Arg Pro Asp Ala Pro Gln Ala Arg Ile Glu Met Gly Phe Pro  
 370 375 380  
 Cys Val Trp Val Leu Pro Trp Arg Arg Glu Gly Asp Trp Leu Ala Pro  
 385 390 395 400  
 Leu Arg Asp Thr Leu Ser Leu Thr Ala Phe Thr Asp Asp Asp Gly Leu  
 405 410 415  
 10 Ile Glu Ser Leu Val Ala Glu Pro Ser Ile Asp Lys Ile His Ile Gly  
 420 425 430  
 Asp Arg Pro Thr Thr Trp Thr Leu Pro Gly Leu Pro His Glu Gly Tyr  
 435 440 445  
 15 Leu Gly Ser Phe Leu Met Arg Ser Lys Ala Val Val Val Gly  
 450 455 460

<210> 15  
 <211> 343  
 20 <212> Proteína  
 <213> Artificial

<220>  
 <223> Orf14

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



Gly Asp Ala Val Ser Glu Phe Lys Pro Val Arg Glu Phe Gly Ala Ala  
 325 330 335  
 Pro Val Ile Glu Gly Val Tyr  
 340

5  
 <210> 16  
 <211> 402  
 <212> Proteína  
 <213> Artificial

10  
 <220>  
 <223> Orf15

15  
 <400> 16

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

30  
 Ala Ala Arg Glu Ala Gly Ala Glu Val Arg Glu Gly Phe Thr Val Thr  
 115 120 125  
 Gly Leu Val Leu Gly Asp Ala Gly Glu Val Val Gly Val Arg Gly Arg  
 130 135 140  
 Gly Pro Arg Gly Pro Glu Val Thr Glu Arg Ala Arg Val Val Leu Gly  
 145 150 155 160  
 Ala Asp Gly Leu His Ser Leu Val Ala Arg Ala Val Asp Ala Pro Arg  
 165 170 175  
 Tyr Asn Glu His Pro Lys Leu Met Val Gly Tyr Tyr Ser Tyr Phe Ser  
 180 185 190

40  
 Gly Leu Asp Met Asp Gly Val Phe Lys Ala His Ser Arg Pro Tyr Arg  
 195 200 205  
 Ser Phe Gly Ala Trp Pro Thr His Asp Gly Leu Thr Leu Val Gly Gly  
 210 215 220  
 Cys Trp Pro Phe Ala Glu Phe Asn Asp Ile Arg Lys Asp Ile Glu Gly  
 225 230 235 240  
 Asn Tyr Leu Lys Asn Phe Ala Leu Ala Pro Ala Trp Glu Glu Arg Ile  
 245 250 255  
 Arg Asp Ala Arg Arg Glu Asp Arg Ile Val Gly Ala Ala Leu Pro Asn  
 260 265 270

50  
 Phe Phe Arg Lys Pro Phe Gly Pro Gly Trp Ala Leu Val Gly Asp Ala  
 275 280 285  
 Gly Tyr Cys Lys Asp Phe Phe Thr Ala Gln Gly Ile Ser Asp Ala Phe  
 290 295 300  
 Ile Ser Ala Glu Met Cys Ala Gly Ser Leu Asp Asp Ala Leu Ser Gly  
 305 310 315 320  
 Arg Ala Pro Phe Asp Thr Ala Met Ala Ala Tyr Gln Ala Ala Arg Asp  
 325 330 335  
 Arg His Ala Arg Pro Val Tyr Asp Phe Thr Leu Gln Val Ser Thr Leu  
 340 345 350

60  
 Glu Pro Leu Ser Pro Glu Phe Glu Lys Val Leu Glu Gly Ile Asp Gly  
 355 360 365  
 Asn Gln Gln Gly Met Asp Ala Phe Ala Gln Val Asn Ala Gly Val Thr  
 370 375 380  
 Ser Met Glu Arg Phe Ser Ala Asp Trp Gly Gly Ala Val Arg Pro Val  
 385 390 395 400  
 Pro Arg

<210> 17  
 <211> 413  
 5 <212> Proteína  
 <213> Artificial  
  
 <220>  
 <223> Orf16  
 10 <400> 17  
  
 Met Glu His Arg Ile Asp Ala Asp Leu Leu Ile Pro Gly Ala Gly Glu  
 1 5 10 15  
 15 Pro Thr Val Asn Gly Ser Val Val His Ala Asp Gly Arg Ile Arg Phe  
 20 25 30  
 Ala Gly Pro Thr Ala Glu Leu Ser Gly Glu His Arg Ala Leu Glu Pro  
 35 40 45  
 20 Thr Arg Val Ala Thr Leu Leu Pro Gly Leu Trp Asp Cys His Val His  
 50 55 60  
 Phe Ala Gly Ile Arg Gly Arg Val Ser Thr Glu Glu Leu Met Leu Thr  
 65 70 75 80  
 Pro Glu Thr Leu Ala Val Ala Arg Ser Val Lys Asp Ala Glu Thr Ala  
 85 90 95  
 25 Leu Arg Ala Gly Phe Thr Ser Val Arg Asp Met Gly Gly His Gly Cys  
 100 105 110  
 Val Leu Ala Glu Ala Ile Arg Glu Gly Thr Phe Thr Gly Pro Asn Ile  
 115 120 125  
 30 Tyr Ser Ala Asn Gln Val Ile Gly Gln Thr Gly Gly His Ser Asp Ala  
 130 135 140  
 His Arg Leu Pro Tyr Arg Trp Val Thr Asp Pro Cys Arg Ser Gly Gly  
 145 150 155 160  
 Thr Leu Arg Ile Ala Asp Gly Val Asp Glu Cys Val Arg Ala Val Arg  
 165 170 175  
 35 Leu Gln Leu Arg Ala Gly Ala Glu Leu Ile Lys Ile Cys Thr Ser Gly  
 180 185 190  
 Gly Val Leu Ser Glu Val Asp Asn Pro Val His Gln Gln Tyr Arg Ser  
 195 200 205  
 40 Glu Glu Leu Asn Ala Ile Val Thr Glu Ala Ala Arg Ala Asp Arg Ile  
 210 215 220  
 Val Ala Ala His Cys His Gly Arg Ala Gly Ile Leu Ala Ala Ile Asn  
 225 230 235 240  
 Ala Gly Cys Arg Thr Val Glu His Gly Thr Glu Ile Asp Glu Glu Thr  
 245 250 255  
 45 Ala Asp Leu Met Ala Glu Arg Gly Met Thr Leu Val Pro Thr Arg Thr  
 260 265 270  
 Ile Tyr Glu Ala Phe Arg His Asn Val Ala Ala Leu Pro Pro Ala Trp  
 275 280 285  
 Arg Asp Arg Phe Glu Val Met Ala Glu Arg His Leu Thr Ala Ile Gly  
 290 295 300  
 50 Ile Ala His Arg Ala Gly Val Thr Ile Ala Leu Gly Thr Asp Leu Gly  
 305 310 315 320  
 Thr Ser Asp Arg Gly Gly Pro Leu Ser Trp Gly Gly His Ala Ser Glu  
 325 330 335  
 55 Phe Ala His Leu Val Ser Ala Gly Leu Ser Pro Leu Glu Ala Ile Thr  
 340 345 350  
 Ala Ala Thr Ala His Gly Pro Gly Thr Leu Gly Pro Arg Ala Pro Arg  
 355 360 365  
 Ser Gly Arg Leu Glu Ala Gly Tyr Asp Ala Asp Leu Leu Ala Val Glu  
 370 375 380  
 60 Gly Asn Pro Leu Ala Asp Ile Thr Val Leu Ala Asp Pro Asp Arg Ile  
 385 390 395 400  
 Thr Arg Val Trp Lys Ser Gly Glu Pro Val Thr Ser Gly  
 405 410  
 65

<210> 18  
 <211> 562  
 <212> Proteína  
 <213> Artificial  
 5  
 <220>  
 <223> Orf17  
 <400> 18  
 10 Met Ser Pro Tyr Gln Ser Arg Pro Leu Asp Pro Asp Val Pro Val Ile  
 1 5 10 15  
 Asp Arg Arg Arg Ala Glu Ala Tyr Leu Ala Ala Gly Leu Trp Arg Glu  
 20 25 30  
 15 Glu Gly Val Ala Glu Leu Leu Arg Ala Ala Arg Leu Arg His Pro Asp  
 35 40 45  
 Arg Leu Ala Val Val Ser Gly Thr Thr Arg Leu Thr His Gly Glu Leu  
 50 55 60  
 His Thr Ala Val Thr Val Ala Gly Arg Arg Leu Glu Gly Leu Gly Val  
 65 70 75 80  
 20 Arg Ala Gly Asp Arg Val Val Val Gln Leu Pro Asn Cys Ala Glu Phe  
 85 90 95  
 Val Val Leu Val Leu Ala Leu Leu Asp Ile Gly Ala Pro Pro Val Leu  
 100 105 110  
 Val Leu Pro Gly Phe Gly Asp Tyr Glu Leu Gly His Val Val Arg Ser  
 25 115 120 125  
 Ala Arg Pro Val Ala Leu Ala Val Ser Ala Gly Ser Arg Ser Ser Asp  
 130 135 140  
 Pro Val Glu Ser Ala Arg Arg Leu Arg Glu Asp His Glu Thr Leu Asp  
 145 150 155 160  
 30 His Val Leu Ala Leu Gly Asp Val Ala Ala Gly Asp Val Asp Leu Ala  
 165 170 175  
 Ala Leu Cys Asp Pro Gly Ala Pro Trp Asp Ala Pro Val Gly Pro Ala  
 180 185 190  
 Ser Ser Ala His Pro Arg Pro Gly Arg Gly Ser Ala Leu Thr Asp Ala  
 35 195 200 205  
 Ala Val Phe Leu Leu Ser Gly Gly Thr Thr Gly Pro Pro Lys Val Ile  
 210 215 220  
 Pro Arg Gly Asn Ala Gly Tyr Ala Tyr Met Ile Arg Thr Ala Cys Ala  
 225 230 235 240  
 40 Ile Ala Asp Val Ser Gln Asp Ala Val Tyr Leu Ala Val Met Pro Ala  
 245 250 255  
 Ala His Gly Phe Val Leu Asn Cys Pro Gly Val Leu Gly Thr Leu Ala  
 260 265 270  
 Gln Gly Gly Thr Val Val Leu Gly Asp Thr Thr Asp Pro Arg Gln Ala  
 45 275 280 285  
 Leu Gly Leu Ile Glu Arg Glu Arg Val Thr His Cys Ala Leu Val Pro  
 290 295 300  
 Thr Val Ala Met Gln Trp Leu Ala Ala Ala Glu Gly Thr Thr Ala Asp  
 305 310 315 320  
 50 Leu Ser Ser Leu Arg Val Leu Gln Thr Gly Gly Ala Arg Pro Ala Pro  
 325 330 335  
 Glu Leu Ala Ala Arg Ile Arg Pro Glu Leu Gly Ala Thr Leu Gln Gln  
 340 345 350  
 Cys Tyr Gly Met Ser Glu Gly Leu Leu Ser Tyr Thr Gly Leu Asp Asp  
 55 355 360 365  
 Pro Glu Ser Val Ile Val Gly Thr Gln Gly Arg Pro Ala Ser Pro Met  
 370 375 380  
 Asp Glu Val Leu Ile Val Asp Glu Gln Gly Arg Pro Val Ala Asp Gly  
 385 390 395 400  
 60 Thr Thr Gly Glu Leu Leu Thr Arg Gly Pro Tyr Thr Val Ala Gly Tyr  
 405 410 415  
 Tyr Arg Asp Pro Ala Ala Thr Ala Lys Ala Phe Thr Pro Asp Gly Phe  
 420 425 430  
 Tyr Arg Thr Gly Asp Leu Ala Thr Arg Thr Ser Gly Gly Asp Leu Val  
 65 435 440 445  
 Val Asn Gly Arg Val Arg Asp Ile Ile Asn Arg Gly Gly Glu Lys Ile

450            455            460  
 Pro Ala Glu Asp Leu Glu Leu Val Ala Arg His Pro Asp Val Arg  
 465            470            475            480  
 Ala Ser Ala Ala Val Gly Ala Pro His Glu Leu Tyr Gly Glu Val Val  
 5            485            490            495  
  
 Cys Leu Tyr Val Val Pro Glu Ala Gly Arg Glu Pro Thr Leu Leu Glu  
           500            505            510  
 Leu Arg Arg Phe Leu Arg Gly Leu Gly Leu Ala Val Trp Lys Leu Pro  
 10            515            520            525  
 Glu Leu Met Glu Thr Val Ala Gln Leu Pro Thr Thr Ala Ile Gly Lys  
           530            535            540  
 Ile Asp Lys Lys Val Leu Arg Ala Asp Ile Ala Arg Arg Thr Ala Val  
 15            545            550            555            560  
 Glu Ala  
  
 <210> 19  
 <211> 86  
 20 <212> Proteína  
     <213> Artificial  
  
 <220>  
 <223> Orf18  
 25  
 <400> 19  
 Leu Pro Ser Gly Asn Gln Gly Ala Ala Val Ser Val Asp Val Leu Lys  
   1            5            10            15  
 Gln Leu Leu Leu Asp Ile Gly Ile Ala Glu Arg Thr Leu Thr Glu Ile  
 30            20            25            30  
 Glu Pro Gly Thr Arg Leu Arg Ala Asp Leu Gly Leu Ser Ser Val Glu  
           35            40            45  
 Thr Thr Asp Leu Glu Ile Gln Leu Arg Glu Arg Phe Gly Val Arg Ile  
           50            55            60  
 35 Asn Leu Trp Asp Lys Ala Asp Tyr Thr Met Glu Gln Leu Ala Ala Gly  
       65            70            75            80  
 Ile Arg Glu Met Pro Arg  
           85  
  
 40  
 <210> 20  
 <211> 458  
 <212> Proteína  
 <213> Artificial  
 45  
 <220>  
 <223> Orf19  
  
 <400> 20  
 50 Val Ile Gly Ala Val Leu Ser Ala Ala Ser Asp Gln His Ile Ser Met  
   1            5            10            15  
 Pro Glu Leu Ser Asp Asp Thr Cys Lys Gly Gly Arg Glu Pro Trp Gly  
           20            25            30  
 Trp Thr Val Glu Arg Ser Tyr Glu Ile Pro Tyr Pro Glu Leu Gly Ser  
 55            35            40            45  
 Val Leu Gly Glu Gln Glu Ile Gly Val Leu Thr Arg Leu Val Thr Ser  
           50            55            60  
 Gly Glu Asn Leu Ser Gly Gly Arg Cys Arg Glu Glu Phe Glu Arg Cys  
       65            70            75            80  
 60 Phe Arg Glu Tyr Leu Asp Val Pro His Ala Leu Ser Val Thr Ser Gly  
           85            90            95  
 Thr Val Ala Leu Glu Ile Ala Ile Arg Leu Leu Asp Leu Asn Glu Gly  
           100            105            110  
 Asp Glu Val Ile Ala Thr Pro Gln Thr Tyr Lys Ala Ser Val Gln Pro  
 65            115            120            125  
 Leu Leu Asn Tyr Pro Val Lys Val Arg Phe Cys Asp Val Gly Pro Asn

130            135            140  
 Thr Leu Asn Ile Asp Pro Gly His Phe Glu Ser Leu Ile Thr Ser Arg  
 145            150            155            160  
 5 Thr Lys Ala Val Ile Leu Val His Tyr Gly Gly Leu Pro Cys Asp Met  
           165            170            175  
 Asp Ala Ile Met Ala Ile Ala Arg Arg His Gly Ile Thr Val Ile Glu  
           180            185            190  
 Asp Cys Ala His Ala Leu Gly Ala Glu Tyr Arg Gly Arg Lys Pro Gly  
 10 Ala Leu Ala Asp Ile Gly Cys Phe Ser Phe His Ser Ser Lys Asn Ile  
           195            200            205  
           210            215            220  
 Thr Thr Leu Gly Glu Gly Gly Met Ile Thr Leu Phe Asp Pro Ala Leu  
 15 Ala Glu Arg Ala Asp Arg Ile Arg Ser Asn Asp Ala Asp Ala Val Tyr  
           225            230            235            240  
           245            250            255  
 Arg Ala Gln Ala Arg Ala Ile Gly Asn Thr Thr Ser Ala His Pro Trp  
           260            265            270  
 Met Leu His Pro Gly Ala Ala Phe Thr His Asp Cys Ser Thr Ile Arg  
 20 Tyr Gly Gly Thr Asn Ala Thr Leu Ala Glu Pro Asn Ala Ala Val Gly  
           275            280            285  
           290            295            300  
 Thr Val Gln Leu Ala Lys Leu Asp Arg Leu Val Arg Arg Ala Glu  
 25 Ile Ala Ala Ala Tyr Thr Asp Val Leu Lys Gln His Pro Gly Val Arg  
           305            310            315            320  
           325            330            335  
 Met His Glu Gly Pro Asp Pro Val Arg His Ala His His Leu Phe Thr  
           340            345            350  
 Phe Phe Ala Asp Pro Ala Asp Gly Ile Leu Arg Asp Arg Leu Val Ser  
 30 Arg Leu Asp Ala Leu Gly Val Gln Met Gln Leu Arg Tyr Phe Pro Met  
           355            360            365  
           370            375            380  
 His Leu Leu Ala Glu Trp Arg Ala Arg Gly His Thr Ala Gly Glu Cys  
 35 Pro Val Ala Glu Arg Leu Trp Phe Glu Gln Gln Val Asn Leu Pro Cys  
           385            390            395            400  
           405            410            415  
 His Pro Ala Met Thr Asp Arg Gln Val Gly Gln Val Val Ser Arg Leu  
           420            425            430  
 Asp Thr Val Leu Gly Gln Glu Ala Arg Asp Gln Pro Ile Ala Ser Val  
           435            440            445  
 40 Thr Thr Pro Thr Ala Glu Gly Trp Asn Arg  
           450            455

45 <210> 21  
     <211> 393  
     <212> Proteína  
     <213> Artificial

50 <220>  
     <223> Orf20

<400> 21  
 Met Ser Glu Thr Glu His Tyr Asp Ile Ala Val Ile Gly Gly Gly Pro  
 1            5            10            15  
 55 Val Gly Leu Ala Ser Ala Trp His Ala Ala Arg Arg Gly Glu Arg Val  
           20            25            30  
 Ala Val Phe Glu Gln Phe Ser Phe Gly Asn Glu Gln Cys Gly Thr Ser  
           35            40            45  
 Gly Ala Glu Arg His Trp Arg Leu Gln Tyr Thr Glu Pro Asp Leu Cys  
 60            50            55            60  
 Arg Leu Thr Gly Glu Ala Leu Pro Leu Trp Arg Glu Leu Glu Arg Ala  
 65            65            70            75            80  
 Thr Gly His Gln Leu Leu His Ala Phe Gly Ser Leu Trp Phe Gly Asp  
           85            90            95  
 65 Ile Asp Val Ala Thr Asn Glu Gly Arg Ile Ser Ala Thr Ala Arg Thr  
           100            105            110

Met Asp Asp Leu Ser Ile Pro Tyr Glu Trp Leu Thr Ala Ala Asp Ile  
 115 120 125  
 Glu Arg Arg Tyr Gly Phe Thr Gly Leu Pro Gly His Phe Glu Gly Phe  
 130 135 140  
 5 Val Gln Pro Asp Gly Gly Ala Val Asp Val Arg Ala Thr Val Glu Gly  
 145 150 155 160  
 Leu Leu Arg Leu Thr Glu Glu Ala Gly Cys Ala Leu Arg Ala His Glu  
 165 170 175  
 10 Pro Val Leu Glu Leu Ile Pro Asp Gly Gly Gly Val Thr Leu Arg Thr  
 180 185 190  
 Ala Arg Gly Arg Cys Arg Ala Gly Lys Val Val Val Ala Asn Gly Ala  
 195 200 205  
 Tyr Ala Asn Lys Leu Leu Glu Pro Leu Gly Ser Arg Leu Asp Leu His  
 210 215 220  
 15 Val Phe Glu Met Ala Leu Val Thr Leu Arg Gln Arg Asp Pro Lys Val  
 225 230 235 240  
 Arg Tyr Pro Phe Trp Phe Val Phe Gln Glu Pro Thr Glu Glu Asp Thr  
 245 250 255  
 20 Asn Leu Phe Tyr Gly Phe Pro Pro Asn Ala Trp Gln Asp Thr Asp Thr  
 260 265 270  
 Val Arg Val Gly Pro Val Phe Glu Val Asn Ala Leu Ala Asp Pro Ala  
 275 280 285  
 Arg Ala Thr Gly Thr Pro Asp Pro Arg His Val Ala Arg Met Cys Glu  
 290 295 300  
 25 Trp Val Glu Arg His Leu Pro Val Val Asp Pro Arg Pro Leu Ala Arg  
 305 310 315 320  
 Asp Thr Cys Leu Ala Val Leu Pro Ala Asp Pro Glu Arg Gln Phe Phe  
 325 330 335  
 30 Leu Gly Thr Ala Lys Gly Arg Phe Asp Gly Gly Glu Asn Val Val Ile  
 340 345 350  
 Ala Thr Gly Gly Trp Gly Phe Lys Phe Val Pro Leu Leu Gly Lys Val  
 355 360 365  
 Cys Ala Asp Leu Cys Val Asp Gly Ala Thr Gly Tyr His Val Asp Arg  
 370 375 380  
 35 Leu Met Leu Pro Asp Ala Ala His Pro  
 385 390

40 <210> 22  
 <211> 2552  
 <212> Proteína  
 <213> Artificial

45 <220>  
 <223> Orf21

<400> 22  
 Val Gly Val Gln Val Arg Pro Ala Ala Gly Gln Gly Leu Arg Arg Ser  
 1 5 10 15  
 50 Val Arg Gly Arg Arg Asp Arg Leu Pro Arg Gly Pro Pro Asp Ala Ala  
 20 25 30  
 Arg Arg Gly Pro Pro Val Ser Arg Ser Ala Pro Thr Arg Gly Arg Pro  
 35 40 45  
 55 Phe Pro Ser Lys Gly Asp Ser Pro Val Val Ser Glu Ala Pro Ala Thr  
 50 55 60  
 Gly Ala His His Asp Gly Thr Ser Val Asp Ile Ala Val Val Gly Met  
 65 70 75 80  
 Ala Gly Arg Phe Pro Gly Ala Pro Asp Leu Asp Ala Tyr Trp His Asn  
 85 90 95  
 60 Leu Arg Ser Gly Val Glu Ser Ile Glu Arg Leu Thr Glu Asp Asp Leu  
 100 105 110  
 Leu Ala Glu Gly Val Asp Pro Glu Leu Ile Gly Ala Pro Gly Tyr Val  
 115 120 125  
 65 Pro Val Ala Pro Val Leu Glu Gly Ile Asp Leu Phe Asp Ala Arg Phe  
 130 135 140  
 Phe Gly Phe Thr Ala Arg Glu Ala Ala Leu Leu Asp Pro Gln Gln Arg

145            150            155            160  
 Leu Phe Leu Glu Ser Ala Trp His Ala Met Glu His Ala Gly Ile Asp  
                  165            170            175  
 5 Pro Ala Arg Cys Gly Thr Ala Ala Val Phe Ala Gly Gly Asn Met Pro  
                  180            185            190  
 Ala Tyr Leu Met Ser Asn Leu Leu Gly Gly Ala Arg Val Val Leu Asp  
                  195            200            205  
 Ser Ala Met Phe Glu Leu Gln Ile His Asn Asp Lys Asp Phe Leu Ala  
                  210            215            220  
 10 Ser Arg Thr Ala Tyr Lys Leu Gly Leu Thr Gly Pro Ala Val Asn Val  
                  225            230            235            240  
 Gln Thr Ala Cys Ser Thr Ser Leu Val Ala Val His Gln Ala Ala Ala  
                  245            250            255  
 15 Ala Leu Arg Ser Gly Asp Cys Glu Ile Ala Leu Ala Gly Gly Val Cys  
                  260            265            270  
 Val Arg Val Pro His Arg Val Gly Tyr Arg Tyr Glu Gln Gly Leu Ile  
                  275            280            285  
 Tyr Ala Pro Asp Gly Arg Cys Arg Pro Phe Asp Ala Asp Gly Ala Gly  
                  290            295            300  
 20 Thr Val Phe Gly Asn Gly Ala Gly Ala Val Val Leu Lys Arg Leu Ala  
                  305            310            315            320  
 Asp Ala Arg Arg Asp Gly Asp Arg Ile Leu Ala Val Leu Lys Gly Ser  
                  325            330            335  
 25 Ala Val Asn Asn Asp Gly Ala Glu Lys Val Gly Tyr Thr Ser Pro Ser  
                  340            345            350  
 Val Ser Gly Gln Glu Ala Val Val Ala Ala Ala Ile Ala Asp Ser Gly  
                  355            360            365  
 Val Pro Ala Arg Ser Ile Thr Ala Ile Glu Ala His Gly Thr Gly Thr  
                  370            375            380  
 30 His Val Gly Asp Pro Ile Glu Ile Thr Ala Leu Ser Arg Ala Phe Gly  
                  385            390            395            400  
 Arg His Thr Thr Asp Thr Gly Phe Cys Ala Val Gly Ser Val Lys Ser  
                  405            410            415  
 35 Asn Ile Gly His Leu Glu Ser Ala Ala Gly Ile Ala Ser Phe Ile Lys  
                  420            425            430  
 Ala Val Leu Gln Leu His His Arg Thr Leu Val Pro Ser Leu His Phe  
                  435            440            445  
 Glu Arg Pro Asn Pro Arg Ile Asp Phe Asp Ala Thr Pro Phe Phe Val  
                  450            455            460  
 40 Asn Thr Glu Leu Arg Ala Trp Pro Glu Gly Glu His Pro Arg Arg Ile  
                  465            470            475            480  
 Gly Val Ser Ser Phe Gly Ile Gly Gly Thr Asn Ala His Val Val Leu  
                  485            490            495  
 45 Glu Gln Ala Pro Asp Pro Val Pro Ala Glu Pro Ser Gly Arg Pro Glu  
                  500            505            510  
 Leu Val Val Val Ser Ala Lys Ser Pro Ala Ala Leu Asp Ala Ala Thr  
                  515            520            525  
 Glu Ala Leu Ala Glu Lys Leu Ala Ala Pro Asp Ala Gln Pro Leu Ala  
                  530            535            540  
 50 Asp Ile Ala His Thr Leu Gln Thr Gly Arg Gly Ala Met Arg Tyr Arg  
                  545            550            555            560  
  
 Arg Ala Val Val Ala Ala Gly Thr Ala Glu Ala Ala Ala Leu Leu Ser  
                  565            570            575  
 55 Gly Ala Asp Pro Gly Arg Val Arg Ser Ala Asp Ala Gly Thr Ala Pro  
                  580            585            590  
 Ala Lys Val Val Phe Leu Phe Pro Gly Gln Gly Ala Gln Tyr Pro Gly  
                  595            600            605  
 Met Ser Arg Gly Leu Tyr Ala Ser Glu Pro Val Phe Ala Glu Ala Leu  
                  610            615            620  
 60 Asp Ala Cys Ala Asp Leu Leu Ala Glu Glu Leu Gly Ile Asp Leu Arg  
                  625            630            635            640  
 Thr Val Leu Phe Pro Asp Ala Pro Ala Glu Asp Gly Leu Thr His Thr  
                  645            650            655  
 65 Thr Leu Ala Gln Pro Ala Leu Phe Ala Thr Glu Tyr Ala Met Ala Thr  
                  660            665            670

Leu Leu Arg Ser Trp Gly Val Glu Pro Asp Val Met Val Gly His Ser  
 675 680 685  
 Ile Gly Glu Phe Thr Ala Ala Val Leu Ser Gly Val Leu Ser Leu Lys  
 690 695 700  
 5 Asp Ala Ala Arg Leu Val Ala Leu Arg Gly Arg Leu Met Gln Asp Arg  
 705 710 715 720  
 Pro Thr Gly Ala Met Val Ser Ile Ala Ala Pro Ala Ala Asp Ile Glu  
 725 730 735  
 10 Pro Leu Leu Pro Ala Gly Val Ser Ile Ala Ala Ile Asn Ala Pro Val  
 740 745 750  
 Leu Cys Val Ala Ser Gly Pro His Glu Ala Val Ala Glu Leu Gly Glu  
 755 760 765  
 Ile Leu Ala Ala Lys Glu Ile Thr Val Arg Pro Leu His Thr Ser His  
 770 775 780  
 15 Ala Phe His Ser Ala Met Met Asp Pro Val Val Glu Pro Phe Thr Glu  
 785 790 795 800  
 Ala Val Ala Gly Thr Pro Leu Ala Ala Pro Gly Leu Pro Phe Val Ser  
 805 810 815  
 20 Cys Val Thr Gly Leu Pro Ile Thr Ala Glu Leu Ala Thr Asp Pro Gln  
 820 825 830  
 Tyr Trp Gly Thr His Leu Arg Arg Pro Val Arg Phe Ala Asp Ala Val  
 835 840 845  
 Arg Thr Ala Ile Gly Asp Gly Pro Ala Val Leu Val Glu Val Gly Pro  
 850 855 860  
 25 Gly Asn Thr Leu Ser Thr Leu Ala Arg Ala Gly Ala Gly Thr Gly Gly  
 865 870 875 880  
 Pro Arg Cys Ala Ala Val Thr Thr Leu Arg Arg Pro Asp Glu Ala Ala  
 885 890 895  
 30 Asp Asp Gly Gln Val Leu Arg Thr Ala Val Gly Asp Ile Trp Leu Phe  
 900 905 910  
 Gly Gly Ala Val Asp Trp Pro Ala Leu His Gln Gly Arg Arg Asn Arg  
 915 920 925  
 Val Glu Leu Pro Gly Tyr Pro Phe Gln Arg Asp Arg Tyr Trp Ile Glu  
 930 935 940  
 35 Pro Arg Gly Ser Ala Thr Gly Thr Pro Leu Val Ala Asp Phe Ala Glu  
 945 950 955 960  
 His Glu Glu Ala Glu Thr Glu Pro Ala Gly Arg Ala Thr Arg Pro Ser  
 965 970 975  
 40 Thr Leu Val Thr Ala Tyr Val Ala Pro Ala Asp Glu Leu Glu Thr Thr  
 980 985 990  
 Ile Ala Gly Ile Trp Glu Glu Met Phe Gly Ile Ala Pro Ile Gly Thr  
 995 1000 1005  
 Arg Asp Asp Phe Glu Leu Gly Gly His Ser Leu Leu Ala Ile  
 1010 1015 1020  
 45  
  
 50 Gln Val Leu Asn Arg Leu Gln Ala Thr Ser Gly Val Thr Val Glu  
 1025 1030 1035  
 Leu Gly Arg Leu Leu Ala Thr Pro Thr Ile Gly Gly Met Ala Glu  
 1040 1045 1050  
 Glu Leu Arg Ala Ala Gly Ala Ala Gly Thr Asp Asp Arg Leu Pro  
 1055 1060 1065  
 55 Thr Val Val Pro Arg Pro Asp Leu Arg Tyr Glu Pro Phe Pro Leu  
 1070 1075 1080  
 Thr Glu Met Gln Gln Ala Gln Trp Ile Gly Arg Leu Ser Ser Phe  
 1085 1090 1095  
 Asp Met Gly Gly Val Ala Pro His Leu Tyr Phe Glu Phe Asp Ser  
 1100 1105 1110  
 60 Arg Thr Ile Glu Thr Ala Arg Leu Glu Arg Ala Trp Gln Arg Val  
 1115 1120 1125  
 Val Gln Arg His Asp Met Leu Arg Met Val Val Leu Pro Asp Gly  
 1130 1135 1140  
 Arg Gln Gln Ile Leu Asp Asn Thr Glu Pro Tyr Arg Phe Glu Val  
 1145 1150 1155  
 65 Leu Asp Leu Arg Thr Thr Asp Pro Glu Glu Ala Glu Arg Gln Leu



	1160	1165	1170
	Ala Gln	Ile Arg Asp Arg Met	Ala Thr Glu Val Arg Pro Ala Asp
	1175	1180	1185
5	Val Trp	Pro Leu Trp Glu Val	Arg Val Gly Leu Leu Pro Asp His
	1190	1195	1200
	Arg Val	Arg Val His Ile Ser	Phe Asp Leu Leu Val Ala Asp Val
	1205	1210	1215
	Ser Ser	Phe Phe Tyr Gln Leu	Leu Pro Gln Trp Arg Glu Phe Tyr
	1220	1225	1230
10	His His	Pro Glu His Asp Pro	Glu Pro Leu Ala Leu Ser Phe Arg
	1235	1240	1245
	Asp Tyr	Val Leu Ala Glu Glu	Glu Leu Arg Arg Thr Pro Arg Tyr
	1250	1255	1260
	Glu Arg	Ser Leu Glu Tyr Trp	Arg Lys Arg Val Arg Glu Leu Pro
15	1265	1270	1275
	Ala Ala	Pro Glu Leu Pro Thr	Val Gln Gly Ala Gly Gly Gly Glu
	1280	1285	1290
	Arg Leu	Gly Phe Val Arg Arg	His Ala Arg Leu Asp Ala Glu Leu
	1295	1300	1305
20	Trp Gly	Arg Ile Lys Ala Lys	Ala Gly Glu Phe Gly Val Thr Pro
	1310	1315	1320
	Ser Ser	Ala Met Leu Ala Ala	Phe Ala Val Thr Ile Gly Thr Trp
	1325	1330	1335
	Ser Lys	Ser Gln Arg Phe Thr	Leu Asn Phe Thr Ala Val Asn Arg
25	1340	1345	1350
	Leu Pro	Val His Glu Glu Val	Asp Asp Val Val Gly Glu Phe Ala
	1355	1360	1365
	Ser Phe	Asp Leu Leu Glu Val	Asp Ala Val Ser Ala Pro Asp Phe
	1370	1375	1380
30	Ala Gly	Leu Val Arg Glu Leu	Gln Arg Gln Ser Trp Ala Asp Phe
	1385	1390	1395
	Asp His	Arg Tyr Val Ser Gly	Val Arg Ile Leu Arg Glu Arg Ala
	1400	1405	1410
	Arg Ala	Arg Gly Gly Ala Gly	Asp Val Met Pro Val Val Phe Thr
35	1415	1420	1425
	Ser Ala	Leu Gly Ser Asp Val	Asp Gly Lys Pro Ala Pro Ser Pro
	1430	1435	1440
	Val Asp	Trp Leu Gly Glu Gln	Ser Tyr Phe Ile Ser Gln Thr Pro
	1445	1450	1455
40	Gln Val	Thr Ile Asp His Phe	Leu Leu Glu Phe Gly Gly Asn Leu
	1460	1465	1470
	Glu Leu	Ala Trp His Ala Val	Asp Gly Leu Phe Pro Asp Gly Leu
	1475	1480	1485
45	Met Glu	Glu Met Phe Gln Ala	Tyr Gln Asp Phe Val Val Gly Leu
	1490	1495	1500
	Ala Glu	Thr Asp Gly Trp His	Arg Pro Pro Val Leu Asp Leu Pro
	1505	1510	1515
	Ala Gly	Gln Leu Ala Pro Arg	Ala Ala Ala Asn Asp Thr Ala Gly
50	1520	1525	1530
	Glu Leu	Pro Asp Gly Val Leu	Pro Ala Arg Ile Leu Ala Arg Ala
	1535	1540	1545
	Gly Ser	Ala Glu Pro Ala Val	Ile Thr Glu Asp Arg Thr Leu Asp
	1550	1555	1560
55	Tyr Ala	Glu Leu Thr Gly Arg	Ala Val Ala Leu Ala Arg Glu Leu
	1565	1570	1575
	Thr Glu	Ala Gly Tyr Gly Arg	Gly Ala Val Val Gly Ile Gly Leu
	1580	1585	1590
	Ala Lys	Gly Trp Arg Gln Thr	Val Ala Ala Leu Ala Ala Ser Ala
60	1595	1600	1605
	Ala Gly	Cys Thr Tyr Val Pro	Leu Asp Pro Gly Leu Pro Glu Ala
	1610	1615	1620
	Arg Arg	Arg Trp Leu Val Glu	Gln Ala Gly Ile Gly Cys Val Leu
	1625	1630	1635
65	Ala Glu	Pro Asp Thr Ala Ala	Leu Trp Pro Asn Ala Pro Arg Val
	1640	1645	1650

Leu Pro Val Ala Glu Asp Ala Arg Trp Asp Pro Ala Asp Thr Ala  
 1655 1660 1665  
 Ala Trp Ser Cys Pro Ala Arg Pro Glu Asp Thr Ala Tyr Val Ile  
 1670 1675 1680  
 5 Tyr Thr Ser Gly Ser Thr Gly Thr Pro Lys Gly Val Ala Val Ser  
 1685 1690 1695  
 His Arg Ala Ala Leu Asn Thr Leu Val Asp Ile Glu Glu Arg Phe  
 1700 1705 1710  
 10 Gly Ile Arg Pro Gly Asp Arg Val Leu Gly Leu Ser Ala Leu Asn  
 1715 1720 1725  
 Phe Asp Leu Ser Val Phe Asp Val Phe Gly Met Leu Ala Ala Gly  
 1730 1735 1740  
 Gly Ala Val Val Leu Pro Glu Ala Ala Asp Arg Arg Asn Pro Asp  
 1745 1750 1755  
 15 Arg Trp Thr Glu Leu Cys Arg Arg His Gly Val Thr Val Trp Asn  
 1760 1765 1770  
 Ser Val Pro Ala Leu Met Gln Met Leu Val Glu His Leu Glu Ser  
 1775 1780 1785  
 20 Arg Gly Pro Ala Asp Asp Ala Gly His Leu Pro Gly Leu Arg Leu  
 1790 1795 1800  
 Ala Leu Leu Ser Gly Asp Trp Ile Pro Leu Ser Leu Pro Asp Arg  
 1805 1810 1815  
 Ile Arg Ala Val Ala Pro Ala Thr Asp Val Ile Ser Leu Gly Gly  
 1820 1825 1830  
 25 Ala Thr Glu Ala Ala Val Trp Ser Ile Ala His Pro Ile Gly Glu  
 1835 1840 1845  
 Val Asp Pro Asp Trp Pro Ser Val Pro Tyr Gly Arg Pro Leu Arg  
 1850 1855 1860  
 30 Asn Gln Arg Phe His Val Leu Asn Asp Arg Leu Arg His Ala Pro  
 1865 1870 1875  
 Val Trp Val Pro Gly Gln Leu His Ile Ala Gly Ala Gly Leu Ala  
 1880 1885 1890  
 Glu Gly Tyr Trp Arg Asp Glu Arg Arg Thr Ala Glu Ser Phe Ile  
 1895 1900 1905  
 35 Thr His Pro Glu Thr Gly Glu Arg Leu Tyr Arg Thr Gly Asp Leu  
 1910 1915 1920  
  
 Gly Arg Tyr Leu Pro Asp Gly Thr Ile Glu Phe Leu Gly Arg Asp  
 1925 1930 1935  
 40 Asp Phe Gln Val Lys Ile Gly Gly His Arg Ile Glu Leu Gly Glu  
 1940 1945 1950  
 Ile Glu His Ala Leu Gly Ser His Pro Glu Leu Leu Asn Ala Val  
 1955 1960 1965  
 45 Val Ser Ala Pro Gly Glu Arg Asn Arg Gln Arg Leu Val Ala His  
 1970 1975 1980  
 Val Val Pro Ala Asp Pro Gly Thr Arg Asn Asp Ala Asp Phe Ala  
 1985 1990 1995  
 Asp Arg Leu Arg Asp His Leu Thr Thr Thr Leu Pro Ser Tyr Met  
 2000 2005 2010  
 50 Ile Pro Ser Asp Ile Val Leu Ile Asp Ala Met Pro Leu Ser Ala  
 2015 2020 2025  
 Asn Gly Lys Val Asp Arg Ser Ala Leu Pro Asp Pro Gln Arg Thr  
 2030 2035 2040  
 55 Gly Asp Ala Glu Ala Ala Ala Ser Ala Ala Glu Asp Asp Gly Glu  
 2045 2050 2055  
 Glu Ala Thr Gly Ala Leu Arg Thr Leu Leu Val Leu Ala Ala Asp  
 2060 2065 2070  
 Leu Leu Gly Val Asn Gly Pro Arg Pro Arg Asp Asn Phe Phe Glu  
 2075 2080 2085  
 60 Leu Gly Gly Asp Ser Ile Met Gly Val Gln Leu Val Gly Arg Ala  
 2090 2095 2100  
 Asn Ala Glu Gly Ile Pro Ile Thr Pro Gln Asn Leu Phe Glu Ser  
 2105 2110 2115  
 Thr Thr Phe Leu Glu Leu Ala Ala Ala Val Pro Val Glu Pro Gly  
 2120 2125 2130  
 65 Thr Asp Asp Thr Gly Glu Ala Val Ala Leu Thr Pro His Gln Thr

2135            2140            2145  
 Leu Ala His Ala Gln Val Gly Ser Val Leu Leu Asp Val Pro Asp  
 2150            2155            2160  
 5 Ala Phe Asp Pro Ala Ser Ala Ala Arg Ala Leu Asn Ala Leu Ala  
 2165            2170            2175  
 Asp Arg His Pro Ala Leu Arg Thr Arg Val Arg Thr Glu Asp Gly  
 2180            2185            2190  
 Gln Arg Phe Ala Val Arg Pro Gly Pro Gly Glu Asp Phe Asp Val  
 2195            2200            2205  
 10 Pro Glu Ile Asp Leu Ala Ala Leu Pro Asp Asp Val Arg Ala Glu  
 2210            2215            2220  
 Ala Val Ala Glu Met Ile Gly Glu Met Ala Gly Glu Val Asp Ile  
 2225            2230            2235  
 15 Glu Thr Gly Pro Ala Val Lys Phe Ala Val Phe Arg Leu Gly Glu  
 2240            2245            2250  
 Arg Gly Ser Val Leu Ala Cys Thr Ala Ala Gln Gly Leu Met Asp  
 2255            2260            2265  
 Asp Ala Ser Val Leu Leu Leu Cys Arg Glu Leu Ile Gln Ala Tyr  
 2270            2275            2280  
 20 Asp Arg Leu Ala Ala Gly Arg Pro Val Val Trp Ser Asp Gly Ala  
 2285            2290            2295  
 Gly Ser Pro Gln Ala Trp Asn Arg Gly Leu Arg Arg Lys Pro Ala  
 2300            2305            2310  
 25 His Pro Ala Gly Leu Ala Glu Asn Pro Gly Thr Ala Gly Glu Leu  
 2315            2320            2325  
 Pro Arg Gln Arg Asn Met Glu Leu Asp Ala Ala Arg Thr Ala Gly  
 2330            2335            2340  
 Leu Phe Thr Ala Ala Ala Gly Ser His His Leu Asp Pro Thr Glu  
 2345            2350            2355  
 30 Val Leu Val Ala Ala Ala Ser Ala Ala Leu Gly Arg Ala Leu Pro  
 2360            2365            2370  
  
 Glu Pro Pro Gln Leu Leu Val Glu Arg Ser Leu Arg Asp Asp Leu  
 2375            2380            2385  
 35 Ala Ala Gly Asp Glu Pro Ala Gly Arg Leu Val Gly Arg Thr Thr  
 2390            2395            2400  
 Glu Leu Arg Thr Val Gln Pro Val Ala Ala Gly Thr Pro Leu Asp  
 2405            2410            2415  
 40 Thr Ala Leu Thr Ser Val Lys Gly Arg Leu Arg Thr Ala Asp Pro  
 2420            2425            2430  
 Asp Pro Val Arg Gly Thr Thr Val Ala Val Arg Glu Val Val Thr  
 2435            2440            2445  
 Trp Asp Arg Val Glu Gly Ala Val Glu Val Pro Ala Asp Phe Ala  
 2450            2455            2460  
 45 Gly Val Thr Gly Leu Ala Gly Trp His Glu Glu Thr Val Gly Gln  
 2465            2470            2475  
 Leu Ser Ala Ala Val Val Asp Gly Ala Leu Arg Ile Arg Trp Gln  
 2480            2485            2490  
 50 Leu Ala Ala Ser Val Pro Glu Asp Ala Ala Thr Arg Leu Ala Asp  
 2495            2500            2505  
 Ala Phe Gly Thr Val Leu Gly Glu Ile Ala Glu His Cys Arg Arg  
 2510            2515            2520  
 Val Ala Glu Gly Ser Tyr Glu Pro Ser Asp Phe Pro Leu Ala Asp  
 2525            2530            2535  
 55 Leu Ser Gly Asp Glu Leu Ala Glu Phe Leu Asp Glu Leu Arg  
 2540            2545            2550  
  
 <210> 23  
 <211> 1091  
 60 <212> Proteína  
 <213> Artificial  
  
 <220>  
 <223> Orf22  
 65

<400> 23  
 Val Thr Ala Thr Gly Asp Arg Arg Met Ala Thr Thr Ala Asp Gly Thr  
 1 5 10 15  
 Val Met Thr Asp Asn Gly Ala Thr Lys Arg Pro Val Arg Asp Ile Ala  
 5 20 25 30  
 Asp Ile Tyr Glu Leu Ser Pro Ile Gln Gln Gly Leu Leu Tyr Glu Gln  
 35 40 45  
 Leu Ala Gln Pro Gly Leu Gly Ile Tyr Val Glu Gln Leu Gly Leu Glu  
 50 55 60  
 10 Phe Ser Gly Thr Met His Pro Glu His Phe Glu Arg Ala Trp Gln Leu  
 65 70 75 80  
 Val Val Asp Arg His Pro Ile Leu Arg Thr Ser Phe His Trp Arg Lys  
 85 90 95  
 15 Asp Gly Ser Ala Val Gln Val Val His Gly Ser Ala Arg Leu Pro Leu  
 100 105 110  
 Glu Thr Leu Asp Trp Arg Asp Leu Asp Glu Arg Thr Gln Glu Glu Arg  
 115 120 125  
 Leu Arg Ala Ser Leu Asp Ala Glu Arg Ala Glu Gly Phe Asp Leu Thr  
 130 135 140  
 20 Asp Val Pro Leu Met Arg Ser Thr Leu Ile Arg Arg Gly Asp Glu Arg  
 145 150 155 160  
 Trp Thr Phe Ser Trp Arg Phe Ser His Leu Leu Met Asp Gly Trp Ser  
 165 170 175  
 25 Phe Thr Leu Ala Ile Gln Asp Phe Ile Asp His Tyr Arg Val Leu Cys  
 180 185 190  
 Arg Gly Gly Arg Pro Thr Leu Ser Pro Gly Arg Ser Tyr Arg Asp Tyr  
 195 200 205  
 Leu Ser Trp Trp Arg Asp Arg Asp Pro Glu Glu Ala Arg Glu Phe Trp  
 210 215 220  
 30 Arg Glu Glu Leu Ala Asp Tyr Arg Pro Val Glu Gln Val His Leu Gly  
 225 230 235 240  
 Gly Thr Gly Ile Pro Glu Gly Glu Pro Thr His Ala His Phe Glu Arg  
 245 250 255  
 35 Ile Leu Gly Asp Leu Ala Pro Arg Leu Thr Ala Leu Ala Arg Ala Glu  
 260 265 270  
 Gln Leu Thr Leu Ala Thr Leu Ala Gln Gly Ala Trp Phe Ile Val Leu  
 275 280 285  
 Gly Arg Phe Leu Gly Arg Thr Asp Leu Ala Cys Gly Ile Thr Met Ala  
 290 295 300  
 40 His Arg Pro Pro Asp Leu Val Gly Ser Gln Asp Ile Leu Gly Pro Met  
 305 310 315 320  
 Ile Ala Thr Leu Pro Leu Arg Arg Arg Leu Asp Pro Ala Met His Leu  
 325 330 335  
 45 Arg Ser Trp Leu Arg Glu Phe Gly Lys His Gly Ile Glu Ala Ser Gly  
 340 345 350  
 His Ser Ala Val Pro Leu Thr Glu Met Gln Ala Leu Leu Gly Thr Asp  
 355 360 365  
 Ser Ala Ile Pro Ile Leu Gln Ser Ser Val Ser Tyr Glu Asn Val Pro  
 370 375 380  
 50 Met Pro Asp Phe Asp Leu Ala Asp Val Gly Ala Glu Met Thr Glu Leu  
 385 390 395 400  
 Val Tyr Asp Gly Arg Pro His Phe Pro Ile Thr Met Val Ile Met Pro  
 405 410 415  
 55 Gly Ala Asp Met Pro Leu Arg Val Val His Asp Arg Arg Lys Val Ser  
 420 425 430  
 Asp Glu Val Ala Glu Arg Phe Ala Gly Glu Val Val Ser Val Leu Thr  
 435 440 445  
 Gln Met Ile Glu Arg Pro Asp Val Thr Leu Gly Glu Leu Thr Phe Leu  
 450 455 460  
 60 Ser Thr Pro Gln Pro Gly Asn Thr Pro Leu Thr Glu Pro Asp Ala Glu  
 465 470 475 480  
 Ser Leu His Glu Thr Phe Arg Arg His Ala Arg Leu Arg Pro Asp Ala  
 485 490 495  
 65 Thr Ala Val Arg Cys Gly Gly Arg Ala Leu Thr Tyr Arg Glu Leu Asp  
 500 505 510  
 Ala Tyr Ser Asp Arg Ile Ala Ala Thr Leu Arg Ala Arg Cys Pro Gly

515            520            525  
 Val Thr Arg Val Gly Leu Cys Leu Pro Arg Ser Ile Glu Leu Val Ala  
 530            535            540  
 Ala Met Ile Gly Val Phe Lys Ala Gly Ala Ala Tyr Val Pro Leu Asp  
 5    545            550            555            560  
 Pro Glu Tyr Pro Ala Asp Arg Leu Ala Asp Met Leu Ala Asp Ser Ala  
 565            570            575  
 Ala Glu Leu Val Leu Thr Asp Gly Ala Pro Ala Asp Ala Leu Thr Ala  
 580            585            590  
 10 Gly Lys Ala Gly Leu Val Thr Leu Pro Glu Met Asp Gly Glu Pro Asp  
 595            600            605  
 His Asn Ala Pro Pro Pro Val Pro Ala Asp Pro Asp Ala Pro Ala Tyr  
 610            615            620  
 Leu Leu Tyr Thr Ser Gly Ser Thr Gly Arg Pro Lys Gly Val Pro Ile  
 15 625            630            635            640  
 Thr His Arg Asn Val Gln Ser Leu Leu Ala Ala Gly Arg Glu Val Phe  
 645            650            655  
 Gly Phe Thr Ala Glu Asp Val Trp Thr Phe Ala His Ser Phe Ala Phe  
 660            665            670  
 20 Asp Tyr Ser Val Trp Glu Ile Trp Gly Ala Leu Gly Asn Gly Ala Ser  
 675            680            685  
  
 Leu Val Val Val Asp His Glu Thr Gly Arg Asp Pro Arg Ala Leu Ala  
 690            695            700  
 25 Arg Leu Ile Ala Glu Glu Arg Val Thr Val Leu Ser Glu Thr Pro Ala  
 705            710            715            720  
 Leu Phe Glu His Leu Val Pro Glu Leu Ala Asp Asp Thr Ser Leu Arg  
 725            730            735  
 Arg Val Phe Leu Gly Gly Asp Arg Leu Asp Pro Ala Ile Leu Arg Pro  
 30 740            745            750  
 Trp Phe Ala Arg Phe Gly Asp Arg Asp Gly Ala Gly Arg Glu Thr Thr  
 755            760            765  
 Gly Arg Ala Pro Ala Pro Gly Ile Glu Leu Tyr Asn Leu Tyr Gly Val  
 770            775            780  
 35 Thr Glu Ala Thr Val Val Ser Thr Tyr His Arg Val Arg Glu Glu Asp  
 785            790            795            800  
 Val Arg Ala Gly Arg Pro Val Pro Ile Gly Arg Ala Leu Pro Asn Gln  
 805            810            815  
 Arg Val Tyr Leu Leu Gly Glu Asp Asp Arg Pro Val Pro Val Gly Ala  
 40 820            825            830  
 Thr Gly Gln Leu Cys Val Ala Gly His Ala Val Ala Ser Gly Tyr His  
 835            840            845  
 Asp Arg Asp Gly Leu Thr Ala Glu Arg Phe Gly Ser Asp Pro Ser Ala  
 850            855            860  
 45 Gly Pro Ser Ser Ala Ala Phe Pro Leu Tyr Arg Thr Gly Asp Leu Ala  
 865            870            875            880  
 Thr Ala Thr Pro Asp Gly Glu Val His Phe Leu Gly Arg Ala Asp Thr  
 885            890            895  
 Gln Val Lys Val Arg Gly Phe Arg Val Glu Pro Gly Glu Ile Glu Ala  
 50 900            905            910  
 Ala Leu Arg Glu Thr Pro Gly Val Arg Ser Ala Thr Val Thr Val His  
 915            920            925  
 Gly Ser Gly Thr Ala Arg Arg Leu Val Gly Tyr Ala Val Pro Glu Asp  
 930            935            940  
 55 Pro Asp Ala Val Leu Thr Ala Gly Pro Val Pro Thr Glu Pro Leu Arg  
 945            950            955            960  
 Glu His Leu Arg Thr Arg Leu Pro Glu His Met Val Pro Ala Ala Val  
 965            970            975  
 Tyr Trp Ile Asp Arg Ile Pro Thr Thr Pro Gly Gly Lys Val Asp Val  
 60 980            985            990  
 Ala Ala Leu Pro Val Pro Asp Ala Gly Gly Thr Asp Arg Asn Thr Ala  
 995            1000            1005  
 Pro Met Thr Glu Ala Glu Arg Leu Leu Ala Gly Leu Leu Thr Glu  
 1010            1015            1020  
 65 Val Leu Gln Val Pro Asp Val Gly Ala Asp Asp Thr Leu Gly Ala  
 1025            1030            1035

Leu Gly Leu Asp Ser Leu Gly Ala Met Arg Leu Ala Ala Arg Leu  
 1040 1045 1050  
 Arg Gly Ala Tyr Ala Leu Asp Leu Ala Val Ser Asp Leu Pro Ala  
 1055 1060 1065  
 5 Thr Arg Thr Val Ala Glu Leu Ala Arg Ala Val Glu Ala Ala Arg  
 1070 1075 1080  
 Pro Val Ala Gly Glu Gly Ala Arg  
 1085 1090  
  
 10  
  
 <210> 24  
 <211> 394  
 <212> Proteína  
 15 <213> Artificial  
  
 <220>  
 <223> cIX5  
  
 20 <400> 24  
 Val Asn Ala Asn Pro Gly Gln Thr Pro Thr Asp Ala Thr Asp His Pro  
 1 5 10 15  
 Asp Ser Pro Glu Asn Gln Arg Trp Arg Glu Arg Ile Arg His Phe Ala  
 20 25 30  
 25 Glu Lys Glu Ile Ala Pro Leu Ser Thr Thr Met Asp Arg Thr Ala Thr  
 35 40 45  
 Leu Asp Ala Gly Leu Arg Glu Arg Leu Phe Ala Glu Gly Leu Met Ser  
 50 55 60  
 Val Glu Ile Pro Arg Gly Tyr Gly Thr Gly Gly Thr Leu Cys Gln  
 30 65 70 75 80  
 Leu Ile Leu Thr Ile Glu Glu Val Ala Arg Val Asp Pro Gly Val Ala  
 85 90 95  
 Val Gly Val His Val His Asn Val Leu Val Ala Gly Thr Leu Leu Arg  
 100 105 110  
 35 His Ala Ser Gly Asp Gln Arg Arg Gln Tyr Leu Pro Gln Leu Ala Thr  
 115 120 125  
 Gly Lys Ile Gly Ala Phe Ala Leu Ser Glu Glu Gln Ala Gly Ser Asp  
 130 135 140  
 Ala Phe Ala Leu Thr Thr Val Ala Arg Gln Asp Glu Ala Gly Tyr Leu  
 40 145 150 155 160  
 Leu Thr Gly Arg Lys Arg Trp Thr Ser Asn Ala Arg Asn Ala Asp Leu  
 165 170 175  
 Leu Leu Val Phe Ala Leu Ala Asp Ala Gly Gly Pro Thr Ala Phe Val  
 180 185 190  
 45 Val Pro Ala Asp Ala Pro Gly Val Ser Leu Asp Asp Arg Val Gln Gln  
 195 200 205  
 Met Gly Val Arg Ala Ala Ala Thr Ser Asp Val Ile Phe Asp Gly Thr  
 210 215 220  
 Pro Val Arg Thr Ala Gln Arg Val Gly Pro Pro Gly Gly Gln Thr  
 50 225 230 235 240  
 Val Ala Leu Ser Gly Leu Gly Leu Gly Arg Leu Gly Ile Ala Ala Gln  
 245 250 255  
 Met Thr Gly Leu Ala Gln Gly Ala Leu Asp Ala Ala Val Gly Tyr Ser  
 260 265 270  
 55 Arg Val Arg Glu Gln Phe Gly Gly Arg Ile Ala Asp His Gln Gly Val  
 275 280 285  
 Ala Phe Pro Leu Ala Asp Val Ala Ser Arg Leu Ala Ala Ala Arg Ala  
 290 295 300  
 Leu Leu Tyr Arg Ala Val Asp Leu His Gly Arg Gly Thr Asp Pro Val  
 60 305 310 315 320  
 Glu Leu Met Arg Leu Thr Ala Met Ala Lys Tyr Val Ala Ser Glu Val  
 325 330 335  
 Ala Glu Arg Ala Ala Ser Val Ala Val Glu Thr Leu Gly Gly Asn Gly  
 340 345 350  
 65 Tyr Thr Asp Ala Tyr Pro Val Glu Arg Phe Tyr Arg Asp Ala Lys Ala  
 355 360 365

Gly Lys Ile Tyr Glu Gly Thr Ser Asn Val Leu Leu Arg Thr Ile Ala  
 370 375 380  
 Ser Ile Leu Ile Gly Gly Ser Pro Gly Asp  
 385 390  
 5  
 <210> 25  
 <211> 242  
 <212> Proteína  
 <213> Artificial  
 10  
 <220>  
 <223> orf23  
 15  
 <400> 25  
 Val Thr Glu Pro Leu Arg Leu Leu Cys Phe Pro Tyr Ala Gly Gly Asn  
 1 5 10 15  
 Ala Gln Thr Tyr Val Arg Trp Arg Arg His Leu Ala Pro Asp Ile Glu  
 20 25 30  
 20 Val Cys Pro Met Gln Leu Pro Gly His Gly Glu Arg Ile Gly Glu Pro  
 35 40 45  
 Pro Arg His Arg Trp Asp Asp Leu Leu Ala Asp Ile Arg Thr Arg Leu  
 50 55 60  
 25 Thr Asp Leu Thr Thr Pro Glu Asp Arg Pro Ile Ala Leu Phe Gly His  
 65 70 75 80  
 Ser Leu Gly Ala Leu Leu Ala Phe Glu Cys Ala Arg Ile Leu Val Ser  
 85 90 95  
 Glu His Gly Ile Arg Pro Ala Arg Leu Leu Val Ser Gly His Arg Ala  
 100 105 110  
 30 Pro His Leu Pro Leu Arg Glu Glu Thr Leu His His Leu Pro Asp Thr  
 115 120 125  
 Glu Phe Leu Thr Arg Leu Ser Glu Arg Ser Arg Thr Leu Arg Ala Leu  
 130 135 140  
 35 Thr Asp Pro Glu Phe Arg Lys Leu Leu Leu Pro Met Leu Arg Ala Asp  
 145 150 155 160  
 Phe Thr Ala Ser Glu Thr Tyr Thr Phe Arg Glu Gly Pro Thr Leu Thr  
 165 170 175  
 Cys Pro Ile Thr Ala Leu Gly Gly Glu Arg Asp Glu Asp Ala Thr Leu  
 180 185 190  
 40 Gly Glu Val Ala Ala Trp Gln Arg His Thr Thr Gly Arg Phe Glu Leu  
 195 200 205  
 Thr Ala Phe Pro Gly Asp His Phe Phe Ile Asp Asp Ala Trp Glu Ala  
 210 215 220  
 45 Val Val Thr Ala Val Gly Asp Arg Leu Arg Ser Arg Glu Gly Ser Thr  
 225 230 235 240  
 Pro Gly  
 50  
 <210> 26  
 <211> 407  
 <212> Proteína  
 <213> Artificial  
 55  
 <220>  
 <223> orf25  
 <400> 26  
 Val Pro Pro Thr Met Pro Arg Arg Asn Pro Leu Glu Pro Gly Thr Met  
 1 5 10 15  
 60 Thr Asp Ser Thr Lys Asp Pro His Phe Leu Arg Asn Pro Tyr Pro Thr  
 20 25 30  
 Tyr Asp Ala Leu Arg Ser Ala Cys Pro Val Gln Pro Leu Gln Ala Gly  
 35 40 45  
 Ser Gly Glu Arg Pro Gly Tyr Leu Val Thr Gly Tyr Ala Glu Ala Arg  
 50 55 60  
 65 Glu Ala Leu Gly Asp Ala Arg Leu Ser Lys Asp Thr Ala Val Phe Phe

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

45  
<210> 27  
<211> 201  
50 <212> Proteína  
<213> Artificial

<220>  
<223> orf26

55 <400> 27

Val Thr Arg Asp Gly Ser Ser Pro Ser Pro Ala Ser Thr Asp His Lys  
1            5            10            15  
60 Pro Gly Val Thr Arg Tyr Ala Glu Ala Leu Thr Thr Thr Leu Ser Phe  
                  20            25            30  
Met Thr Asp Val Ser Leu Gly Asp Glu Leu Thr Arg Arg Val Glu Asp  
                  35            40            45  
Leu Glu Glu Arg Val Ser Leu Leu Phe Gln Glu His Pro Ala Ala Glu  
                  50            55            60  
65 Thr Asn Asn Gly Val Gly Thr Ala Asp His Val Ala Glu Ala Glu Asp  
65            70            75            80



Pro Arg Gln Leu Ala Thr Arg Leu Leu Ala Arg Ser Asp Pro Asp Val  
 85 90 95  
 Leu Ala Ala Arg Leu Asn Ala Leu Gly His Pro Val Arg Leu Arg Ile  
 100 105 110  
 5 Leu Leu Ala Cys Leu Asp Gly Pro Arg Arg Ala Ala Glu Leu Ala Ala  
 115 120 125  
 Gln Thr Asp Met Gly Ser Thr Gly Gln Ile Tyr His His Leu Arg Gln  
 130 135 140  
 10 Leu Val Asn Gln Gly Trp Leu Ser Ala Ser Arg Arg Gly His Tyr Glu  
 145 150 155 160  
 Val Pro Arg Glu Ala Leu Glu Val Val Ala Ala Val Leu Ala Ala Thr  
 165 170 175  
 Phe Trp Asp Gly Asn Ala Asp Pro Thr Ser Ser Glu Ala Ala Ser Gly  
 180 185 190  
 15 Ala Arg Thr Ser Ala Ala Ser Ser Thr  
 195 200

20 <210> 28  
 <211> 1362  
 <212> Proteína  
 <213> Artificial

25 <220>  
 <223> orf27

<400> 28  
 Val Ser Cys Cys Cys Arg Ser Arg Ala Val Ser Arg Asp Arg Thr Val  
 1 5 10 15  
 30 Ser Ala Ala Pro Gly Val Arg Ala His Phe Ser Arg Ala Ile Asp Arg  
 20 25 30  
 Gly Arg Lys Leu Ser Gln Gly Ser Ala Pro Ser Thr Val Arg Thr Ala  
 35 40 45  
 35 Asp Arg Leu Pro Ala Thr Ile His Arg Leu Val Glu Ser Gln Ala Glu  
 50 55 60  
 Arg Ala Pro Asp Ala Val Ala Val Glu Thr Glu Asp Gly Lys Leu Thr  
 65 70 75 80  
 Tyr Arg Glu Leu Asp Ala Arg Ala Asn Gln Phe Ala Arg His Leu Arg  
 85 90 95  
 40 Ser Ala Gly Val Arg Gly Glu Ser Leu Val Ala Val His Met Glu Arg  
 100 105 110  
 Gly Leu Leu Thr Pro Val Val Leu Leu Gly Ile Leu Lys Thr Gly Ala  
 115 120 125  
 45 Ala Tyr Leu Pro Leu Asp Thr Glu Ser Pro Ala Glu Arg Leu Ala Ala  
 130 135 140  
 Val Leu Ala Asp Ala Ala Pro Ala Ala Val Val Thr Ala Gly Pro Leu  
 145 150 155 160  
 Pro Pro Val Ala Val Pro Leu Ile Asp Leu Asp Thr Asp Leu Pro Ala  
 165 170 175  
 50 Ile Thr Ala Leu Pro Ala Glu Pro Leu Thr Asp Val Glu Glu Pro Gly  
 180 185 190  
 Pro Asp Arg Leu Ala Tyr Val Met Phe Thr Ser Gly Ser Thr Gly Val  
 195 200 205  
 55 Pro Lys Gly Val Leu Val Glu His Arg Ala Val Ile Arg Leu Ile Arg  
 210 215 220  
 Glu Gln Ser Tyr Ala Arg Leu Gly Pro Asp Ala Thr His Leu Leu Leu  
 225 230 235 240  
 Ala Pro Leu Ala Phe Asp Ala Ser Thr Leu Glu Ile Trp Gly Ala Leu  
 245 250 255  
 60 Ala His Gly Gly Arg Leu Val Val Ala Ala Pro Gly Ala Arg Thr Val  
 260 265 270  
 Asp Gln Leu Gly Arg Thr Leu Ala Asp Arg Arg Val Thr Thr Leu Trp  
 275 280 285  
 65 Leu Thr Ala Ser Leu Phe Asn Leu Val Val Asp Glu Asp Pro Ser Ile  
 290 295 300  
 Leu Ala Gly Val Gly Asp Leu Leu Ile Gly Gly Glu Ala Leu Ser Val

305            310            315            320  
 Asn His Val Arg Thr Ala Arg Lys Ala Leu Pro Asp Thr Val Val Thr  
                  325            330            335  
 5 Asn Gly Tyr Gly Pro Thr Glu Thr Thr Thr Phe Ala Cys Thr His Ala  
                  340            345            350  
 Ile Arg Pro Gln Asp Leu Asp Gly Ala Ser Ile Pro Ile Gly Gly Pro  
                  355            360            365  
 10 Ile Ala His Thr Glu Val His Val Leu Asp Glu Asp Phe Asp Pro Val  
                  370            375            380  
 Ala Pro Gly Glu Ala Gly Glu Leu Phe Ile Gly Gly Pro Arg Leu Ala  
 385            390            395            400  
 Arg Gly Tyr Leu Asn Arg Pro Gly Leu Thr Ala Glu Arg Phe Val Ala  
                  405            410            415  
 15 His Pro Ala Ala Thr Glu Pro Gly Ser Arg Leu Tyr Arg Thr Gly Asp  
                  420            425            430  
 Arg Val Arg Val Arg Pro Asp Gly Thr Leu Glu Tyr Leu Gly Arg Leu  
                  435            440            445  
 20 Asp Asp Gln Val Lys Leu Arg Gly Phe Arg Ile Glu Pro Gly Glu Val  
                  450            455            460  
 Arg Ala Gly Leu Thr Gly Leu Pro Gln Val Arg Asp Ala Val Val Val  
 465            470            475            480  
 Ala Arg Gly Gly Pro Ser Asp Arg Arg Leu Val Ala Tyr Val Val Pro  
                  485            490            495  
 25 Glu Ala Asp Ala Thr Ala Gly Met Asp Asn Glu Arg Glu Gln Val Ala  
                  500            505            510  
 Asp Trp Glu Ala Val Phe Asp Glu Thr Tyr Arg Asp Gly Val Gly Ala  
                  515            520            525  
 Ala Glu Gly Arg Trp Glu Leu Ser Gly Trp Val Gly Ser Gly Asp Gly  
                  530            535            540  
 30 Leu Pro Val Pro Ala Asp Gln Met Arg Glu Trp Thr Asp Ala Thr Val  
                  545            550            555            560  
 Glu Arg Ile Arg Ala Leu Gly Ala Arg Arg Val Leu Glu Ile Gly Cys  
                  565            570            575  
 35 Gly Thr Gly Leu Leu Ala Met Arg Leu Ala Pro Asp Ala Glu Arg Tyr  
                  580            585            590  
 Val Gly Ser Asp Leu Ser Ala Val Ala Ile Arg Arg Leu Arg Ala Gln  
                  595            600            605  
 Met Asp Ala Ala Gly Leu Asp His Thr Glu Leu Val His Ala Pro Ala  
                  610            615            620  
 40 Asp Asp Leu Asp Ala Val Pro Gly Gly Thr Phe Asp Val Val Val Leu  
                  625            630            635            640  
 Asn Ser Ile Val Gln Tyr Leu Pro Ser Ala Gln Tyr Leu Arg Glu Val  
                  645            650            655  
 45 Ile Glu Arg Ala Ala Ala Arg Leu Ala Pro Gly Gly His Leu Phe Val  
                  660            665            670  
 Gly Asp Val Arg Ser Leu Pro Leu Leu Asp Ala Phe His Leu Ser Ala  
                  675            680            685  
 Glu Leu Lys Arg Gly His Glu Asp Ala Val Pro Leu Ala Thr Leu Ala  
                  690            695            700  
 50 Glu Ala Val Arg Glu Arg Ala Ala Ala Glu Lys Glu Leu Val Val Ala  
                  705            710            715            720  
 Pro Ser Phe Phe Thr Asp Leu Ser Gly Arg Ala Gly Ile Asp His Val  
                  725            730            735  
 55 Arg Val Thr Pro Arg Arg Gly Arg His Arg Asn Glu Met Thr Gln Phe  
                  740            745            750  
 Arg Tyr Asp Ala Val Leu Arg Val Arg Gly Ala Glu Pro Ala Arg Val  
                  755            760            765  
 60 Pro Asp Arg Trp Leu Asp Trp Arg Asp Glu Gly Leu Thr Leu Glu Asp  
                  770            775            780  
 Val Ala Arg Ile Leu His Asp Gln Arg Pro Gln His Leu Ala Leu Arg  
                  785            790            795            800  
 Gly Val Thr Asp Ala Arg Val Ala Asp Glu Val Ala Arg Leu Val Arg  
                  805            810            815  
 65 Leu Arg Glu Asp Ala Glu Gly Thr Val Ala Ala Leu Arg Glu Thr Gly  
                  820            825            830

His Asp Gly Pro Ala Val Glu Ile Asp Asp Val Tyr Asp Leu Ala Ala  
 835 840 845  
 Arg Ala Ser Tyr Thr Val Asp Val Ser Val Ala Gly Ser Ala Ala Gly  
 850 855 860  
 5 Asp Ala Phe Asp Val Ile Leu Trp Thr Asp Ala Glu Pro Gly Pro Val  
 865 870 875 880  
 Ala Phe Ala Pro Gly Pro Ala Glu Ala Arg Gly Pro Arg Thr Ser Met  
 885 890 895  
 10 Pro Leu Ala Thr Ala Thr Ser Arg His Leu Thr Thr Leu Val Arg Asp  
 900 905 910  
 Ser Leu Arg Glu Leu Leu Pro Ala Tyr Met Ile Pro Ala Val Phe Val  
 915 920 925  
 Phe Met Asp Ala Leu Pro Leu Thr Ser Thr Gly Lys Ile Asp Arg Ser  
 930 935 940  
 15 Ala Leu Pro Glu Pro Pro Arg Arg Thr Ser Ala Gly Gly Ala Gly Arg  
 945 950 955 960  
 Arg Ala Ala Thr Ala Thr Glu Arg Ala Leu Glu Pro Leu Trp Arg Asp  
 965 970 975  
 20 Leu Leu Ala Leu Glu Thr Val His Val Asp Asp Asp Phe Phe Ala Leu  
 980 985 990  
 Gly Gly His Ser Leu Leu Gly Thr Arg Leu Leu Ser Arg Val Arg Gly  
 995 1000 1005  
 Leu Trp Gly Val Glu Leu Ser Leu Ala Ala Leu Phe Ser Ala Pro  
 1010 1015 1020  
 25 Thr Leu Gly Ala Leu Ala Ala Arg Ile Asp Ser Ala Arg Gln Asp  
 1025 1030 1035  
 Thr Pro Ala Leu Pro Gly Thr Leu Ala Asp Lys Ala Asp Pro Gly  
 1040 1045 1050  
 30 Ser Ala Pro Pro Leu Ser Pro Ala Gln His Arg Leu Trp Leu Val  
 1055 1060 1065  
 Glu Gln Leu Thr Pro Gly Asn Pro Arg Tyr Thr Val Pro Val Ala  
 1070 1075 1080  
 Tyr Arg Met Arg Gly Pro Ile Asp Thr Ala Ala Leu Gln Ala Ala  
 1085 1090 1095  
 35 Leu Asp Thr Leu Val Ala Arg His Glu Val Leu Arg Thr Thr Phe  
 1100 1105 1110  
 Pro Ser His Asp Gly Thr Pro Arg Gln Val Val Ala Pro Ser Gly  
 1115 1120 1125  
 40 Arg Ile Pro Ile Glu Arg Ala Asp Val Gly Gly Glu Gly Ala Asp  
 1130 1135 1140  
 Ala Pro Ala Ala Ala His Asn Ile Leu Thr Arg Gln Ala Ser Arg  
 1145 1150 1155  
 Trp Leu Asp Val Gln Ser Gly Pro Leu Ala Ala Ala Thr Leu Val  
 1160 1165 1170  
 45 Arg Leu Ala Glu Asp Asp His Val Leu Cys Leu Thr Leu His His  
 1175 1180 1185  
 Met Ile Cys Asp Gly Trp Ser Leu Asp Leu Leu Ala Ala Glu Leu  
 1190 1195 1200  
 50 Ser Glu Gly Tyr Asn Ala Arg Val Ala Arg Arg Thr Pro Gln Leu  
 1205 1210 1215  
 Pro Glu Ile His His His Thr Arg Thr Arg Tyr Arg Thr Gly His  
 1220 1225 1230  
  
 55 Pro Leu Gln Gln His His Leu Thr Glu Pro Ala Leu Ile Arg Ser  
 1235 1240 1245  
 Leu Ser His His Gln Pro His Thr Gly Ile Arg Gln His Glu Leu  
 1250 1255 1260  
 Asp Thr Val Arg Arg Ile Ala Arg Ile Leu Pro Ser Pro Ser Ser  
 1265 1270 1275  
 60 Thr Thr Cys Pro Pro Arg Pro Arg Gly Gly Arg Arg Ser Glu Glu  
 1280 1285 1290  
 Ala Val Arg Arg Val Pro Asn Ala Ser Arg Thr Ala Ala Ile Ser  
 1295 1300 1305  
 Ser Gly Ser Arg Ser Arg Pro Ala Met Pro Glu Arg Ile Arg Val  
 1310 1315 1320  
 65 His Gln Leu Val Pro Val Gln Ala Ser Leu Glu Ala Ala Phe Met

ES 2 397 885 B1

1325            1330            1335  
 Glu Leu Thr Arg Thr Ser Val Glu Tyr Gln Ala Arg Thr Ser Thr  
 1340            1345            1350  
 Gly Ala Pro Asp Ala Glu Val Arg Ser  
 5            1355            1360

10 <210> 29  
 <211> 17  
 <212> ADN  
 <213> Artificial

15 <220>  
 <223> L2ATFW2

<400> 29  
 5'-CTCGGCGACATCGGGTG-3'

20 <210> 30  
 <211> 18  
 <212> ADN  
 <213> Artificial

25 <220>  
 <223> L2ATFV2

<400> 30  
 5'-CAGGTGCARNGGGAAGTA-3'

30 <210> 31  
 <211> 17  
 <212> ADN  
 <213> Artificial

<220>  
 <223> L2ATson1

40 <400> 31  
 5'-ATCACCACCCTCGGCGA-3'

45 <210> 32  
 <211> 16  
 <212> ADN  
 <213> Artificial

<220>  
 <223> L2ATson2

50 <400> 32  
 5'-ATGCCGTCGGCCGGGT-3'

55 <210> 33  
 <211> 18  
 <212> ADN  
 <213> Artificial

60 <220>  
 <223> 1c35FW

65 <400> 33  
 5'-TCCCGCCGGACAGACTG-3'

5 <210> 34  
 <211> 20  
 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> Heldisr1  
  
 10 <400> 34  
 5'-GGTCAGCACGTAGGAATCCC-3'  
  
 15 <210> 35  
 <211> 21  
 <212> ADN  
 <213> Artificial  
  
 20 <220>  
 <223> Heldisr2  
  
 25 <400> 35  
 5'-ACCTGAAACCCGCCAATGTCA-3'  
  
 30 <210> 36  
 <211> 22  
 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> orf22del1  
  
 35 <400> 36  
 5'-GAATTCCGGTGAATCGTCGCGG-3'  
  
 40 <210> 37  
 <211> 21  
 <212> ADN  
 <213> Artificial  
  
 45 <220>  
 <223> orf22del2  
  
 50 <400> 37  
 5'-AAGCTTCGGCACGCCTCGCGT-3'  
  
 55 <210> 38  
 <211> 23  
 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> orf22del3  
  
 60 <400> 38  
 5'-GATATCGCGTAGCCCGTCCACGG-3'  
  
 65 <210> 39  
 <211> 24  
 <212> ADN  
 <213> Artificial

<220>  
 <223> orf22del4  
 5 <400> 39  
 5'-GATATCGCTATCTGCTGACCGGCC-3'  
  
 10 <210> 40  
 <211> 26  
 <212> ADN  
 <213> Artificial  
  
 15 <220>  
 <223> orf16del1  
  
 <400> 40  
 5'-GAATTCGCCCTCGCTCATGCCGTAG-3'  
  
 20 <210> 41  
 <211> 28  
 <212> ADN  
 <213> Artificial  
  
 25 <220>  
 <223> orf16del2  
  
 <400> 41  
 30 5'-AAGCTTTGAGCACTGCGCCGATCACTAC-3'  
  
 <210> 42  
 <211> 18  
 35 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> 1c35RV  
 40 <400> 42  
 5'-AGGCGCTCCTTCAGCTCG-3'  
  
 45 <210> 43  
 <211> 29  
 <212> ADN  
 <213> Artificial  
  
 50 <220>  
 <223> orf16del3  
  
 <400> 43  
 55 5'-GGATCCACCGATGAGCGAACTGAGCACT-3'  
  
 <210> 44  
 <211> 27  
 60 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> orf16del4  
 65 <400> 44

5'-GATATCCAGGAAGAGCCGCTGCTGTGG-3'

5 <210> 45  
 <211> 24  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> orf8del3

<400> 45  
 5'-GGATCCCACCCCCTGACAGCCTTC-3'

15 <210> 46  
 <211> 24  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> orf8del4

<400> 46  
 25 5'-GATATCCGCGGTGTTCTGCTCGA-3'

<210> 47  
 <211> 24  
 <212> ADN  
 <213> Artificial

<220>  
 <223> orf8del1

35 <400> 47  
 5'-GAATTCGTACTIONCCACCGCCGCGTA-3'

40 <210> 48  
 <211> 24  
 <212> ADN  
 <213> Artificial

45 <220>  
 <223> orf8del2

<400> 48  
 50 5'-AAGCTTGACCGTCCTCGTCCACGT-3'

<210> 49  
 <211> 24  
 <212> ADN  
 <213> Artificial

<220>  
 <223> orf6del1

60 <400> 49  
 5'-GAATTCTCTCGTGCATTCGCCGCG-3'

65 <210> 50  
 <211> 30  
 <212> ADN

<213> Artificial  
 <220>  
 <223> orf6del2b  
 5  
 <400> 50  
 5'-AAGCTTTCCGAAGATCGCTGATATCACGGT-3'  
 10  
 <210> 51  
 <211> 18  
 <212> ADN  
 <213> Artificial  
 15  
 <220>  
 <223> orf6del3  
 <400> 51  
 5'-GCCCGATATCGCGGCCTG-3'  
 20  
 <210> 52  
 <211> 25  
 <212> ADN  
 <213> Artificial  
 25  
 <220>  
 <223> orf6del4  
 <400> 52  
 30  
 5'-GATATCCGCTCGTCGCCCAGGACAG-3'  
 <210> 53  
 <211> 24  
 <212> ADN  
 <213> Artificial  
 35  
 <220>  
 <223> orf101del1  
 <400> 53  
 5'-GAATTCAAGGGAATCCGCTGGCCG-3'  
 40  
 <210> 54  
 <211> 25  
 <212> ADN  
 <213> Artificial  
 45  
 <220>  
 <223> orf101del2  
 <400> 54  
 5'-AAGCTTCGTCCGATGGGTCAACTGG-3'  
 55  
 <210> 55  
 <211> 24  
 <212> ADN  
 <213> Artificial  
 60  
 <220>  
 <223> orf101del3  
 <400> 55  
 5'-GATATCAAGGCGGTCGTCGTCGGC-3'  
 65



ES 2 397 885 B1

5 <210> 56  
<211> 25  
<212> ADN  
<213> Artificial

10 <220>  
<223> orf101del4

<400> 56  
5'-GATATCCTTCCCGGACGTCGTGGAC-3'



- ②① N.º solicitud: 201130010  
 ②② Fecha de presentación de la solicitud: 05.01.2011  
 ③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
A	EP 1749552 A1 (NEUROPHARMA S.A.) 07.02.2007, párrafos [0012],[0013],[0017],[0018].	1-18
A	JP 9012550 A (HOUSE FOODS CORP) 14.01.1997, (resumen), [en línea], [recuperado el 08.05.2012]. Recuperado de EPODOC Database.	1-18
A	JP 5078322 A (MEIJI SEIKA KAISHA) 30.03.1993, (resumen), [en línea], [recuperado el 08.05.2012]. Recuperado de EPODOC Database.	1-18

Categoría de los documentos citados

X: de particular relevancia  
 Y: de particular relevancia combinado con otro/s de la misma categoría  
 A: refleja el estado de la técnica

O: referido a divulgación no escrita  
 P: publicado entre la fecha de prioridad y la de presentación de la solicitud  
 E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

**El presente informe ha sido realizado**

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe  
24.01.2013

Examinador  
S. González Peñalba

Página  
1/4

## CLASIFICACIÓN OBJETO DE LA SOLICITUD

**C12N15/00** (2006.01)

**A61K31/44** (2006.01)

**A61P31/00** (2006.01)

**A61P35/00** (2006.01)

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12N, A61K, A61P

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, NPL, BIOSIS, MEDLINE, EMBASE, XPESP, EBI, STN

Fecha de Realización de la Opinión Escrita: 24.01.2013

**Declaración**

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones 1-18	<b>SI</b>
	Reivindicaciones	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP11/1986)</b>	Reivindicaciones 1-18	<b>SI</b>
	Reivindicaciones	<b>NO</b>

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

**Base de la Opinión.-**

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

**1. Documentos considerados.-**

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	EP 1749552 A1 (NEUROPHARMA S.A.)	07.02.2007
D02	JP 9012550 A (HOUSE FOODS CORP)	14.01.1997
D03	JP 5078322 A (MEIJI SEIKA KAISHA)	30.03.1993

**2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**

La presente solicitud de patente tal y como ha sido redactada, hace referencia a una molécula de ácido nucleico, que consiste en al menos un fragmento de la SEQ ID NO:1 capaz de codificar para al menos una de las secuencias SEQ ID NO:2 a SEQ ID NO: 28 (reivindicaciones 1-2); al vector de expresión que comprende dicha molécula de ácido nucleico (reivindicaciones 3 y 4); a la célula u organismo transgénico no humano que comprende dicho vector de expresión (reivindicación 5); a las cepas de *Streptomyces spp* CECT (7754, 7755, 7756, 7757, 7861, 8069, 8070 y 8071 (reivindicación 6); a los compuestos de fórmula (I) a (XX) (reivindicaciones 7-9, 19-22); al uso de dichos compuestos para el tratamiento del cáncer, enfermedades neurodegenerativas y enfermedades infecciosas (reivindicaciones 10-13); a composiciones farmacéuticas que comprenden al menos uno de los compuestos (reivindicaciones 14); a un método de tratamiento de dichas enfermedades (reivindicaciones 15-18) y al procedimiento de obtención de dichos compuestos (reivindicación 23).

**NOVEDAD Y ACTIVIDAD INVENTIVA ARTS . 6 Y 8 DE LA LP**

El documento D01 hace referencia al uso de compuestos de colismicina para la preparación de un medicamento para tratar enfermedades o condiciones inducidas por estrés oxidativo. Las colismicinas son moléculas de 2,2'-bipiridina que se han aislado de especies de *Streptomyces*. Se ha encontrado que la colismicina A y sus derivados sintéticos cercanos muestran una fuerte inhibición de estrés oxidativo (véase página 4, párrafo [0012]). Dichos compuestos presentan la fórmula general de la página 9, líneas 1-35. Preferiblemente, unas de las enfermedades o condiciones inducidas por estrés oxidativo son las neurodegenerativas (véase página 5, párrafo [0017]), concretamente, en este documento se cita, la enfermedad neurodegenerativa de Alzheimer (véase página 5, párrafo [0018]).

El documento D02 se refiere a derivados de 2,2'-bipiridina que tienen actividad antineoplásica obtenidos a partir de *Streptomyces*. Son obtenidos dichos derivados de la cepa *Sptreptomyces californicus* BS-75.

El documento D03 describe compuestos con actividad antitumoral y antibiótica obtenidos a partir del cultivo de bacterias de *Streptomyces sp*.

Por lo tanto, ninguno de los documentos citados, considerados solos o en combinación, revelan la molécula de ácido nucleico que consiste en al menos un fragmento de la SEQ ID NO: 1, ni las cepas, ni los compuestos comprendidos en la presente invención. Ni tampoco, en dichos documentos citados existen sugerencias que dirijan al experto en la materia hacia la invención definida en las reivindicaciones 1-18, por lo que el objeto de las reivindicaciones 1-18 cumple el requisito de novedad y actividad inventiva de acuerdo con los artículos 6 y 8 de la LP.