



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 287 552**

(51) Int. Cl.:

C07D 313/00 (2006.01)

A61K 31/335 (2006.01)

A61P 33/06 (2006.01)

A61P 17/06 (2006.01)

A61P 35/00 (2006.01)

A61P 9/10 (2006.01)

A61P 19/02 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **03786170 .5**

(86) Fecha de presentación : **24.12.2003**

(87) Número de publicación de la solicitud: **1576160**

(87) Fecha de publicación de la solicitud: **21.09.2005**

(54) Título: **Derivados de borrelidina y su utilizacion en medicina.**

(30) Prioridad: **27.12.2002 GB 0230217**

(45) Fecha de publicación de la mención BOPI:
16.12.2007

(45) Fecha de la publicación del folleto de la patente:
16.12.2007

(73) Titular/es: **Biotica Technology Limited**
Chesterford Research Park, Little Chesterford
Nr Saffron Walden, Essex CB10 1XL, GB
The University of Oviedo

(72) Inventor/es: **Salas, José A.;**
Méndez, Carmen;
Olano, Carlos;
Sánchez, César;
Brana, Alfredo F.;
Wilkinson, Barrie;
Martin, Christine J.;
Moss, Steven;
Leadlay, Peter F. y
Oliynyk, Marko

(74) Agente: **Lehmann Novo, María Isabel**

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

DESCRIPCIÓN

Derivados de borrelidina y su utilización en medicina.

5 **Campo de la invención**

La presente invención se refiere a nuevos poliquétidos y usos de los mismos.

Antecedentes de la invención

10

Los poliquétidos son productos naturales producidos por una extensa gama de organismos, y particularmente por microorganismos. Los poliquétidos tienen muchos usos farmacéuticos, veterinarios y agrícolas importantes. Los poliquétidos abarcan una gama enorme de espacio químico estructural, y tienen una amplia gama de actividades biológicas asociadas. Poliquétidos con utilización en tratamientos médicos incluyen antibióticos, inmunosupresores, agentes antitumorales, otros agentes quimioterapéuticos, y otros compuestos que poseen una extensa gama de propiedades terapéuticas y biológicas. La bacteria Gram-positiva *Streptomyces* y sus géneros afines son productores prodigiosos de poliquétidos, y la genética y bioquímica de la biosíntesis de los poliquétidos en estos organismos están caracterizadas relativamente bien (Hopwood, 1997). Los genes para la biosíntesis de poliquétidos en *Streptomyces* están agrupados y el aprovechamiento de la tecnología del DNA ha hecho posible aislar agrupaciones de genes biosintéticos completos por exploración de genotecas de genes con sondas de DNA que codifican los genes responsables para su biosíntesis. Así, números crecientes de agrupaciones de genes para biosíntesis de poliquétidos en *Streptomyces* y otros microorganismos han sido aislados y secuenciados, incluyendo, por ejemplo, los correspondientes al poliéter monensina (WO 01/68867), el polieno nistatina (WO 01/59126) y para rapamicina (Schwecke *et al.*, 1995).

Los poliquétidos se sintetizan por la condensación repetida de bloques de construcción que contienen una función ácido carboxílico. En cada etapa del proceso esto da como resultado la formación de una nueva función β -ceto y una ramificación de cadena lateral α en la cadena creciente. La diversidad estructural de los poliquétidos se deriva de varios aspectos de su camino biosintético que incluye: la gran diversidad de unidades iniciadoras que pueden utilizarse en su biosíntesis; las diferentes longitudes de cadenas de poliquétidos que son posibles; las diversas cadenas laterales α que se introducen sea durante o después del ensamblaje de la cadena del poliquétido; las diversas sustituciones β que pueden introducirse durante o después del ensamblaje de la cadena del poliquétido; los diversos grados de procesamiento que pueden sufrir los grupos β -ceto (ceto, hidroxilo, enoílo, y metileno); y las diversas estereoquímicas que son posibles en los centros α y β .

La síntesis de poliquétidos es catalizada por una enzima, o por un complejo de enzimas, denominado(a) la poliquétido-sintasa (PKS) de una manera similar a la de la biosíntesis de los ácidos grasos. Las PKSs de *Streptomyces* y géneros afines caen dentro de tres categorías principales: tipo I, tipo II y tipo III. Las PKSs de Tipo III son pequeñas proteínas afines a las calcona-sintasas vegetales que han sido descubiertas sólo recientemente (Moore & Hopke, 2000). Sistemas de tipo III han sido implicados en la biosíntesis de un pequeño número de metabolitos secundarios, pero pueden estar implicados más generalmente en la biosíntesis de pigmentos solubles (Cortes *et al.*, 2002). Las PKSs de tipo II están constituidas por varias proteínas monofuncionales que actúan como un complejo multi-polipeptídico. Poliquétidos aromáticos simples tales como actinorrodina se forman por varias tandas de ensamblaje de cadenas, que se realizan iterativamente sobre una serie de enzimas PKS de tipo II que son codificadas por una serie de genes PKS (Hopwood, 1997). Las PKSs de tipo I son proteínas multifuncionales que se requieren para la síntesis de poliquétidos más complejos tales como eritromicina y rapamicina. Como el foco de esta patente, se describen a continuación en detalle la organización y función de las PKS de tipo I:

Las PKSs de tipo I están organizadas en módulos, en donde cada módulo está constituido por varios “dominios” catalíticos que se requieren para llevar a cabo una tanda de ensamblaje de cadenas (Staunton & Wilkinson, 1997). En general, una PKS modular contiene el número correcto de módulos (módulos de carga más extensión) para seleccionar y condensar el número correcto de unidades de carga y extensión. Por ejemplo, la PKS de eritromicina consta de 7 módulos (un módulo de carga y 6 módulos de extensión) para seleccionar y condensar la unidad iniciadora y 6 unidades de tensión requeridas para la biosíntesis del precursor de eritromicina, 6-desoxieritronolida B. Así pues, existe una relación de uno a uno entre el número de módulos presentes en la PKS y el número de unidades incorporadas. Esta relación de uno a uno se describe como “co-linealidad”.

El término “módulo de extensión”, tal como se utiliza en esta memoria, hace referencia a la serie de dominios contiguos, desde el dominio de proteína-sintasa portadora de β -cetoacil-acilo (KS) al dominio de la proteína portadora de acilo (ACP) siguiente, que realiza un ciclo de extensión de la cadena de poliquétido. El término “módulo de carga” como se utiliza en esta memoria, hace referencia a cualquier grupo de dominios contiguos que realiza la carga de la unidad iniciadora en la PKS y la pone así disponible para el dominio KS del primer módulo de extensión. Aparte de la condensación de la unidad ácido carboxílico extendidora siguiente (o quétido) sobre la cadena de poliquétido creciente, que es realizada por la actividad catalítica del dominio esencial KS, los módulos de PKSs de tipo I pueden contener dominios con actividades de β -cetorreductasa (KR), deshidratasa (DH), y enoil-reductasa (ER), que son responsables del procesamiento ulterior de los grupos β -ceto recién formados durante la extensión de la cadena. La acil-transferasa (AT) y los dominios ACP presentes en cada módulo son responsables de la elección de la unidad extendidora, y la ligadura de la cadena en crecimiento durante su paso sobre la PKS, respectivamente. Los dominios AT de una PKS modular pueden encontrarse también como proteínas discretas (Cheng *et al.*, 2003). La cadena de

poliquétido completada se libera generalmente de las PKSs por la acción de un dominio de tioesterasa terminal (TE) que está implicado también generalmente en la ciclación (lactonización) del producto final. Se emplean también otras estrategias de terminación/ciclación de cadenas, tales como la que consiste en la adición de un residuo de aminoácido y formación de macrolactama como se ha observado para la rapamicina (Schwecke *et al.*, 1995), para la formación de macrolactamas y para rifamicina (August *et al.*, 1998), y para incorporación de aminoácidos seguida por eliminación reductora como en el caso de la biosíntesis de la mixalamida (Silakowski *et al.*, 2001). En resumen, existe un solo dominio enzimático presente para cada paso catalítico sucesivo que tiene lugar durante la biosíntesis en la PKS, y los mismos se utilizan en una secuencia definida que depende de su localización en la proteína y de la función particular que realicen. Este mecanismo se denomina “procesivo”.

La configuración modular de las PKSs de tipo I fue confirmada por primera vez por mutación de la PKS de eritromicina (conocida también como 6-desoxieritronolida B-sintasa, DEBS) por una delección en marco de una región del dominio KR del módulo 5 (Donadio *et al.*, 1991). Esto condujo a la producción de los análogos de eritromicina, 5,6-didesoxi-3- α -micarosil-5-oxoeritronolida B y 5,6-didesoxi-5-oxoeritronolida B, debido a la incapacidad del dominio KR mutado para reducir el grupo β -ceto 5 en esta etapa de la biosíntesis procesiva. Análogamente, la alteración de los residuos de sitio activo en el dominio ER del módulo 4 de DEBS2, por modificación mediante ingeniería genética del DNA correspondiente codificante de PKS y su introducción en *Saccharopolyspora erythraea*, condujo a la producción de 6,7-anhidroeritromicina C (Donadio *et al.*, 1993). Adicionalmente, la longitud de la cadena de poliquétido formada por DEBS ha sido alterada por la relocalización específica del dominio TE de DEBS3 al final de DEBS1. El producto triquetido-lactona esperado se produjo con rendimiento satisfactorio (Cortés *et al.*, 1995). Debe indicarse que los cambios descritos implicaban modificación por delección de secuencia, o por desactivación específica de secuencia, o por la yuxtaposición alternativa de la secuencia de DNA desde el interior de la misma agrupación PKS (es decir, los mismos se consideran “cambios homólogos”). Otros cambios “homólogos” de este tipo a la PKS de eritromicina se describen en WO 93/13663.

La organización modular de los genes PKS de tipo I se presta por sí misma a la manipulación de estos genes a fin de producir estructuras de poliquétido alteradas. Las PKSs de tipo I representan una línea de ensamblaje para biosíntesis de poliquétidos que puede manipularse por cambio del número de módulos; por cambio de sus especificidades hacia diferentes unidades iniciadoras de ácido carboxílico y unidades extendedoras; por desactivación, mutación, eliminación, intercambio o inserción de dominios con actividades y especificidades diferentes; y por alteración del tamaño de la cadena o del anillo mediante el reposicionamiento de dominios de terminación o ciclación (Stauton & Wilkinson, 1997).

WO 98/01546 describe la producción de ensamblajes de genes PKS híbridos que comprenden la incorporación de DNA *heterólogo*. WO 98/01546 describe métodos para generación de PKSs híbridas en los cuales la sustitución de genes que codifican módulos heterólogos, sub-módulos o dominios para los genes nativos genera nuevos poliquétidos con estructuras alteradas. Específicamente, por ejemplo los dominios AT de DNA heterólogo de las PKSs de rapamicina o monensina pueden intercambiarse por el nativo para la PKS de eritromicina a fin de generar nuevos poliquétidos con ramificación alquilo alterada. Un intercambio de dominio AT de este tipo representó el primer ejemplo de la producción de una PKS verdaderamente híbrida (Oliynyk *et al.*, 1996). WO 98/01546 describe también en términos generales la producción de ensamblajes de PKS híbridas que comprenden un modo de carga y al menos un módulo de extensión. Dicho documento describe específicamente la construcción de un ensamblaje de gen PKS híbrido por injerto del módulo de carga de especificidad amplia para la PKS productora de avermectina sobre la primera proteína de la PKS de eritromicina (DEBS1) en lugar del módulo de carga normal (véase también Marsden *et al.*, 1998). Ejemplos adicionales que comprenden intercambios de módulos de carga que son específicos de sustrato han sido descritos también (WO 00/00618; US 5.876.991; Kuhstoss *et al.*, 1996). WO 00/01827 describe métodos para variar la capacidad de procesamiento de β -ceto de un módulo PKS por la capacidad de intercambiar “bucles reductores”, es decir la capacidad para, rápidamente y de manera combinatoria, alterar el número y tipo de dominios de cetorreductasa, deshidratasa y enoil-reductasa dentro de un módulo. Además de cambiar el nivel de procesamiento del grupo β -ceto, tales cambios pueden conducir también a cambios en la estereoquímica de los grupos α -alquilo y β -hidroxilo así formados por los módulos alterados.

Aunque las PKSs modulares operan “normalmente” de una manera co-lineal y procesiva como se ha descrito arriba, ejemplos de una desviación de este modo de operación han sido descritos y se exponen más adelante.

La agrupación de genes PKS de picromicina en *Streptomyces venezuelae* es responsable de la biosíntesis tanto de picromicina (un macrólido heptaquétido de 14 miembros) y metimicina (un macrólido hexaquétido de 12 miembros) (Xue *et al.*, 1998). La capacidad de una PKS simple para producir dos macrólidos afines, de tamaños de anillo diferentes, deriva de la expresión alternativa del gen PKS final *pikA4* (Xue & Sherman, 2000). Cuando ocurre la expresión “normal” y se forma *PikA4* de longitud total, se incorpora una sexta unidad de extensión y se produce la picromicin-aglicona; cuando ocurre expresión alternativa y se produce una forma truncada en el terminal N de *PikA4*, no se incorpora unidad sexta de extensión alguna y la cadena de poliquétido creciente se pasa directamente al dominio TE, lo que conduce a la formación de la metimicin-aglicona. Así pues, ocurre una ruptura de la co-linealidad y se forma un producto “de anillo contraído”. La base bioquímica para este fenómeno ha sido investigada y se ha demostrado que es una transferencia de ACP5 a ACP6, con omisión de la unión covalente al dominio KS6 interpuesto; dicha ruptura de la co-linealidad ha sido denominada “salto” (Beck *et al.*, 2002).

Se ha observado que también ocurre salto cuando se interpoló un módulo de extensión adicional de la PKS de rapamicina en la PKS de eritromicina a fin de convertir la PKS natural productora de heptaquétido en una productora

de octaquétido (Rowe *et al.*, 2001). Se produjo el octaquétido esperado macrólido de 16 miembros, pero el producto principal era el producto 6-desoxieritronolida heptaquétido normal. Se cree que este “salto” del módulo interpolado ocurre debido a que el módulo interpolado actúa en algunas ocasiones como una “lanzadera”, pasando la cadena creciente del módulo precedente al módulo siguiente aguas abajo sin realizar una tanda de extensión de cadena. Posteriormente se demostró que el dominio ACP del módulo interpolado es esencial en el paso de la cadena de poliquétido creciente desde el dominio ACP precedente y el paso de la misma al dominio KS del módulo siguiente durante el salto (Thomas *et al.*, 2002), un mecanismo similar al descrito anteriormente para la biosíntesis de metimicina. Se demuestra que puede producirse salto sin el nucleófilo del sitio activo del dominio KS. Se ha informado de una nemadectina (un macrólido antiparasitario) de anillo contraído (por salto), procedente de un mutante de un aislado de *Streptomyces* del suelo que se había modificado por mutación química (Rudd *et al.*, 1990); se anuló la biosíntesis del producto PKS natural.

Una manera alternativa en la cual las PKSs modulares se desvían de la operación co-lineal implica la operación iterativa de módulos. Por ejemplo, el módulo 4 de la PKS de eritromicina parece operar iterativamente, a nivel bajo, para producir un macrólido octaquétido de anillo expandido de 16 miembros afín a la 6-desoxieritronolida B (Wilkinson *et al.*, 2000). La capacidad de la PKS de eritromicina para realizar esta operación ha sido denominada “tartamudeo”. El “tartamudeo” de la PKS de eritromicina está considerado como un proceso aberrante, dado que los productos de este tartamudeo se forman con rendimiento bajo y el producto principal de la PKS de eritromicina es el heptaquétido normal 6-desoxieritronolida B formado por operación co-lineal. Productos que parecen formarse tanto por tartamudeo como por salto han sido también comunicados como componentes menores del productor de epotilona *Sorangium cellulosum* (Hardt *et al.*, 2001). La agrupación biosintética de estigmatelina de *Stigmatella aurantiaca* codifica una PKS que comprende diez módulos (uno de carga y nueve de extensión) (Gaitatzis *et al.*, 2002); sin embargo, sobre la base de los resultados de la elucidación estructural y la alimentación de sustratos marcados con isótopos estables, la estigmatelina está formada por 11 unidades derivadas de módulos. Así, parecería que uno de los módulos de PKS de estigmatelina opera iterativamente (dos veces).

Desde la presentación de prioridad de la presente solicitud, la secuencia de la PKS responsable para la biosíntesis del macrólido lankacidina por *Streptomyces rochei* ha sido descrita (Mochizuki *et al.*, 2003). Esta PKS parece contener también un número muy pequeño de módulos en comparación con el número de ciclos de extensión requeridos para la biosíntesis de la lankacidina, aunque el mecanismo por el cual podría ocurrir esto no está claro.

Puede generarse una diversidad estructural adicional mediante la modificación de poliquétidos por la acción de enzimas distintas de las PKS, sea durante el proceso de ensamblaje de cadenas como se ve durante la biosíntesis de algunas ansamicinas (Floss, 2001), o después del proceso de ensamblaje de cadenas siguiente a la liberación de la PKS. Tales reacciones no mediadas por PKS pueden incluir, pero sin carácter limitante, las siguientes: reducción, oxidación, hidroxilación, acilación, alquilación aminación, descarboxilación, deshidratación, isomerización/migración de enlaces dobles, ciclación, escisión de anillos, conjugación, glicosilación, eliminación reductora y cualquier combinación de éstas. Cuando estas reacciones ocurren después del ensamblaje de cadenas, las mismas se conocen como pasos post-PKS o pasos de adaptación. Tales pasos de adaptación son por regla general, pero no siempre, esenciales para dotar de actividad biológica al producto poliquétido natural.

Adicionalmente, la diversidad estructural de poliquétidos que pueden obtenerse por biosíntesis puede mejorarse adicionalmente por el uso de enzimas adaptadoras heterólogas post-PKS definidas así como por el uso de aquéllas que modifican naturalmente el poliquétido natural (Gaisser *et al.*, 2000). El documento WO 01/79520 describe la modificación heteróloga de estructuras de poliquétidos macrólidos por glicosilación, epoxidación, hidroxilación, y metilación. La capacidad para generar análogos del compuesto agrícola espinosina por glicosilación con sustituyentes desoxihexosa alternativos ha sido comunicada (Gaisser *et al.*, 2002).

La borrelidina 1 (Figura 1) es un macrólido de 18 miembros producido por varias cepas bacterianas que incluyen, pero sin carácter limitante, *Streptomyces rochei* ATCC 23956, *Streptomyces parvulus* Tü113 y *Streptomyces parvulus* Tü4055. La borrelidina se demuestra en esta memoria que se deriva de un ácido iniciador, ácido *trans*-ciclopentano-1,2-dicarboxílico, 3 unidades extendedoras malonil-CoA y 5 unidades extendedoras metilmalonil-CoA (véase la Figura 2). A partir de la estereoquímica absoluta de la borrelidina, basada en la estructura cristalina y configurada recientemente por síntesis total, se predice que el ácido iniciador real es ácido *trans*-ciclopentano-(1*R*,2*R*)-dicarboxílico. La borrelidina aislada después de la alimentación de sustratos acetato y propionato marcados con isótopos estables indicaba claramente la incorporación esperada de estos bloques de construcción; adicionalmente, se ha demostrado en la presente solicitud que la alimentación de ácido *trans*-ciclopentano-1,2-dicarboxílico era suficiente para restablecer la biosíntesis de borrelidina en mutantes en los cuales los genes específicos que se creía están involucrados en la formación de la unidad iniciadora habían sido desorganizados. La borrelidina contiene un grupo nitrilo unido a la posición C12, que se demuestra en esta memoria surge por la acción de enzimas adaptadoras que actúan sobre una ramificación metilo derivada de metilmalonil-CoA presente en esta posición. La estructura global de la borrelidina fue elucidada por primera vez en 1967 (Keller-Scheirlein, 1967), y fue refinada subsiguientemente por análisis NMR detallado (Kuo *et al.*, 1989). La configuración absoluta de la borrelidina fue confirmada por cristalografía de rayos X (Anderson *et al.*, 1989). Su co-identidad como el antibiótico treponemicina ha sido comprobada (Maehr & Evans, 1987).

Varios grupos han comunicado la síntesis de fragmentos de la estructura de borrelidina, y desde la presentación de prioridad de la presente solicitud, han sido consignadas dos síntesis totales independientes de borrelidina (Hanessian *et al.*, 2003; Duffey *et al.*, 2003).

La borrelidina fue descubierta por primera vez debido a su actividad antibacteriana (Berger *et al.*, 1949), aunque esta actividad antibacteriana se extiende solamente a un número limitado de micrococcos, y no se encuentra contra todas las bacterias de test comunes. El modo de acción en los microorganismos sensibles implica la inhibición selectiva de la treonil-tRNA-sintetasa (Paetz & Nass, 1973). Otras actividades contra espiroquetas del género *Treponema* (Singh *et al.*, 1985; US 4.759.928), contra virus (Dickinson *et al.*, 1965, usos para el control de plagas animales y malezas (DE 3.607.287) y uso como fungicida agrícola (DE 19835669; US 6.193.964) han sido consignadas. Adicionalmente, desde la presentación de prioridad de la presente solicitud, se ha informado que la borrelidina posee actividad antimarial contra cepas fármaco-resistentes de *Plasmodium falciparum* (Otoguro *et al.*, 2003). Entre todos estos informes, únicamente dos han comunicado algunos derivados modificados sintéticamente. El primero de éstos describe el éster bencílico y su derivado bis-O-(4-nitrobenzoílo) (Berger *et al.*, 1949). El segundo de ellos describe el éster metílico de borrelidina, el derivado de éster metílico bis-O-acetilo, y los derivados del éster metílico Δ_{14-15} -dihidro-, $\Delta_{14-15,12-13}$ -tetrahidro-, y $\Delta_{14-15,12-13}$ -tetrahidro-C12-amino (Anderton & Rickards, 1965). No se ha consignado actividad biológica alguna para ninguno de estos compuestos.

Una descripción reciente particularmente interesante es el descubrimiento de que la borrelidina exhibe actividad anti-angiogénesis (Wakabayashi *et al.*, 1997). La angiogénesis es el proceso de la formación de nuevos vasos sanguíneos. La angiogénesis ocurre sólo local y transitoriamente en adultos, estando implicada, por ejemplo, en la reparación subsiguiente a traumatismos locales y el ciclo reproductivo de las hembras. Se ha establecido como componente fundamental en varios procesos patogénicos que incluyen cáncer, artritis reumatoide y retinopatía diabética. Su importancia en la habilitación del crecimiento de los tumores más allá de un diámetro de 1-2 cm fue establecida por Folkman (Folkman, 1986), y está provocada por la respuesta de los tumores a la hipoxia. En sus consecuencias aguas abajo, la angiogénesis es fundamentalmente un proceso derivado del hospedador, por lo que la inhibición de la angiogénesis ofrece un potencial importante en el tratamiento de los cánceres, evitando las dificultades de otras modalidades terapéuticas anticancerígenas tales como la diversidad de tipos de cáncer y la resistencia a los fármacos (Matter, 2001). Presenta interés adicional el hecho de que publicaciones recientes han descrito la implicación funcional de tirosinil- y triptofanil-tRNA-sintetasas en la regulación de la angiogénesis (Wakasugi *et al.*, 2002; Otani *et al.*, 2002).

En el cultivo matriz de la aorta de la rata utilizado como modelo de angiogénesis, la borrelidina exhibe un efecto potente inhibidor de la angiogénesis y causa también la desorganización de los tubos capilares formados de una manera dependiente de la dosis por inducción de la apoptosis de las células formadoras de capilares (Wakabayashi *et al.*, 1997). La borrelidina inhibía la formación de tubos capilares con un valor CI_{50} de 0,4 ng/ml (0,8 nM). En el mismo estudio, se demostró que la borrelidina posee actividad anti-proliferativa frente a las células endoteliales de la vena umbilical humana (HUVEC) en un ensayo de crecimiento celular; el valor CI_{50} se midió como 6 ng/ml, que es 15 veces menor que la actividad anti-angiogénesis medida en el mismo medio. Se demostró que esta actividad anti-proliferativa de la borrelidina es general frente a diversas líneas de células. Además de estos datos, los autores consignan que la borrelidina inhibe la tRNA-sintetasa y la síntesis de proteínas en las células de rata cultivadas; sin embargo, el valor CI_{50} para la actividad anti-angiogénesis (0,4 ng/ml) era 50 veces menor que el consignado para la inhibición de la síntesis de proteínas (20 ng/ml), lo que indica actividades diferentes del compuesto.

La borrelidina exhibe también una inhibición potente de la angiogénesis *in vivo* utilizando el modelo del saco de aire dorsal del ratón (Funahashi *et al.*, 1999), que examina la angiogénesis inducida por VEGF y es un modelo excelente para estudiar la angiogénesis de los tumores. La borrelidina se administró a una dosis de 1,8 mg/kg por inyección intraperitoneal y se demostró que reduce significativamente el incremento de volumen vascular inducido por las células WiDr, y en mayor grado que lo hace TNP-470, que es un inhibidor de angiogénesis sintético en pruebas clínicas. Controles detallados verificaron que estos datos se refieren a la inhibición de la angiogénesis y no a inhibición del crecimiento de las células tumorales. Los autores demostraron también que la borrelidina es eficaz para la inhibición de la formación de metástasis espontáneas en el pulmón de las células del melanoma B16-BL6 para la misma dosificación por inhibición de los procesos angiogénicos implicados en su formulación.

Los documentos JP 9-227.549 y JP 8-173.167 confirman que la borrelidina es eficaz contra la línea de células WiDr del cáncer de colon humano, y también contra las líneas de células PC-3 del cáncer de próstata humano. JP 9-227.549 describe la producción de borrelidina por *Streptomyces rochei* Mer-N7167 (Ferm P-14670) y su aislamiento del cultivo de fermentación resultante. Además de borrelidina 1, han sido descritos 12-desnitrilo-12-carboxil-borrelidina 2 (supuestamente un compuesto intermedio biosintético o metabolito de derivación), 10-desmetil-borrelidina 3 (supuestamente un análogo biosintético que surge por la incorporación errónea de una unidad extendidora malonil-CoA alternativa en el módulo 4 de la PKS de borrelidina), 11-epiborrelidina 4 y el análogo de C14,C15-cis-borrelidina 5 (véase la Figura 1). Así, JP 9-227.549 especifica borrelidina y análogos de borrelidina en los cuales un grupo nitrilo o carboxilo está unido al esqueleto de carbonos en C12, y un átomo de hidrógeno o grupo alquilo inferior está unido al esqueleto de carbonos en C10.

WO 01/09113 describe la preparación de análogos de borrelidina que han sufrido modificación por síntesis en el resto ácido carboxílico del anillo de ciclopentano. La actividad de estos compuestos se examinó utilizando la proliferación de las células endoteliales y ensayos de formación de capilares endoteliales de una manera similar a la arriba descrita. En general, la modificación del resto carboxilo mejoraba la selectividad para inhibir la formación de capilares: la razón principal para esta mejora en la selectividad se debe a una disminución en la actividad de inhibición de la proliferación celular, mientras que la actividad inhibidora de la formación de capilares se alteraba en un grado mucho menor. Específicamente, el éster borrelidin-morfolinoetílico exhibía un índice de selectividad de 60 veces, la borrelidin-amida exhibía un índice de selectividad de 37 veces, el éster borrelidin-(2-piridil)-etílico exhibía un índice

de selectividad de 7,5 veces y la borrelidin-morfolinoetil-amida exhibía un índice de selectividad de 6 veces, para la actividad inhibidora de la formación de capilares en comparación con la proliferación celular con respecto a la borrelidina. La actividad inhibidora de la formación de capilares de estos y otros derivados de borrelidina se comprobó utilizando un ensayo de formación de micro-vasos. Adicionalmente, los autores demostraron que la borrelidina inhibía débilmente la propagación de nódulos metastásicos, después de la eliminación del tumor primario, cuando se utilizó un modelo de adenocarcinoma de pulmón de Lewis. No obstante, se informó que el derivado de borrelidin-(3-picolilamida) inhibe muy considerablemente el aumento de micrometástasis en las ratas después de administración intraperitoneal y también con la administración *per os* a dosis subtóxicas. Análogamente, utilizando el modelo bazo-hígado-colon 38 la capacidad formadora de metástasis de las células del adenocarcinoma de colon del ratón trasplantadas en bazo de ratón se reducía considerablemente después de tratamiento con una dosis subtóxica de este derivado de borrelidina. Estos datos confirman la capacidad comunicada previamente de la borrelidina y sus derivados para inhibir la formación de metástasis.

La borrelidina ha sido identificada también como un inhibidor de la quinasa Cdc28/Cln2 dependiente de ciclina de *Saccharomyces cerevisiae* con un valor CI_{50} de 12 $\mu\text{g/ml}$ (24 μM) (Tsuchiya *et al.*, 2001). Se demostró que la borrelidina detiene tanto las células haploides como las diploides en la fase tardía G_1 (en un momento indiferenciable de la feromona de apareamiento α), y a concentraciones que no afectan a la biosíntesis global de las proteínas. Estos datos se tomaron para indicar que la borrelidina tiene potencial como compuesto conductor para desarrollar agentes antitumorales.

Desde la presentación de prioridad de la presente solicitud, se han publicado dos informes adicionales concernientes a la actividad biológica de la borrelidina. El primero de éstos indica que los efectos anti-angiogénicos de la borrelidina están mediados por caminos distintos (Kawamura *et al.*, 2003). Se encontró que concentraciones altas de treonina atenúan la capacidad de la borrelidina para inhibir tanto la formación de tubos capilares en el modelo de cultivo de la aorta de la rata como la proliferación de las células HUVEC; sin embargo, la misma no afectaba a la capacidad de la borrelidina para colapsar los tubos capilares formados o para inducir la apoptosis en las células HUVEC. Se encontró también que la borrelidina activa la caspasa-3 y la caspasa-8, y los inhibidores de estas dos suprimían la apoptosis en HUVEC inducida por borrelidina. El segundo de estos documentos utilizó el método de la presentación de perfiles de mRNA celulares globales para proporcionar comprensión en cuanto a los efectos de borrelidina sobre *Saccharomyces cerevisiae* (Eastwood y Schaus, 2003). Este análisis demostró la inducción de las enzimas de la biosíntesis de aminoácidos de una manera dependiente del tiempo por tratamiento con borrelidina, y se comprobó que la inducción de este camino implica el factor de transcripción *GCN4*.

En suma, el efecto inhibidor de la angiogénesis de la borrelidina está dirigido hacia los efectos duales tumoral-biológicos de proliferación y formación de capilares. Adicionalmente, se ha demostrado que la borrelidina y sus derivados, inhiben la propagación de las metástasis. La borrelidina tiene también indicaciones para uso en la modulación del ciclo celular. Así pues, la borrelidina y los compuestos afines son dianas particularmente atractivas para investigación como agentes terapéuticos para el tratamiento de tejidos tumorales, sea como agentes individuales o para uso como producto adjunto a otras terapias. Adicionalmente, aquéllos pueden utilizarse para tratamiento de otras enfermedades en las cuales está implicada la angiogénesis en el proceso patogénico, con inclusión, pero sin carácter limitante, de la lista siguiente: artritis reumatoide, psoriasis, aterosclerosis, retinopatía diabética y diversos trastornos oftálmicos.

Sumario de la invención

La presente invención proporciona la secuencia de ácido nucleico entera de la agrupación de genes biosintéticos responsables del control de la síntesis del poliquétido macrólido borrelidina en *Streptomyces parvulus* Tü4055. Se proporciona también el uso de la totalidad o parte del DNA clonado y las secuencias de ácido nucleico del mismo en la detección específica de otros agrupaciones de genes poliquétidos biosintéticos, en la modificación por ingeniería genética de cepas mutantes de *Streptomyces parvulus* y otras cepas hospedadoras adecuadas para la producción de niveles incrementados de borrelidina, o para la producción de poliquétidos modificados o nuevos, y de genes recombinantes que codifican sistemas de PKS para la biosíntesis de poliquétidos modificados o nuevos.

La secuencia de nucleótidos completa de la agrupación de genes biosintéticos de borrelidina de *Streptomyces parvulus* Tü4055 se muestra en SEQ ID NO: 1. Su organización se presenta en la Figura 3 y comprende genes y marcos de lectura abierta designados en lo sucesivo en esta memoria como: *borA1*, *borA2*, *borA3*, *borA4*, *borA5*, *borA6*, *borB*, *borC*, *borD*, *borE*, *borF*, *borG*, *borH*, *borI*, *borJ*, *borK*, *borL*, *borM*, *borN*, *borO*, *orfB1*, *orfB2*, *orfB3*, *orfB4*, *orfB5*, *orfB6*, *orfB7*, *orfB8*, *orfB9*, *orfB10*, *orfB11*, *orfB12*, *orfB13*, *orfB14*, *orfB15*, *orfB16*, *orfB17*, *orfB18*, *orfB19*, *orfB20*, *orfB21* y *orfB22*.

Las funciones propuestas de los genes clonados se describen en las Figuras 4 (biosíntesis propuesta de la unidad iniciadora), 5 (organización de la PKS de borrelidina y biosíntesis de la pre-borrelidina) y 6 (introducción del resto nitrilo en C12) y se describen a continuación.

La presente descripción proporciona por tanto una molécula de ácido nucleico aislada que comprende:

- (a) una secuencia de nucleótidos como se muestra en SEQ ID NO: 1, o una porción o fragmento de la misma;

- (b) una secuencia de nucleótidos que es el complemento de SEQ ID NO: 1, o una porción o fragmento de la misma; o
- (c) una secuencia de nucleótidos que está degenerada con una secuencia codificante de SEQ ID NO: 1, o una porción o fragmento de la misma.

Como se utiliza en esta memoria, el término “fragmento” con respecto a una secuencia de nucleótidos hace referencia a un tramo de residuos de ácido nucleico que tienen una longitud de al menos 10, preferiblemente al menos 20, al menos 30, al menos 50, al menos 75, al menos 100, al menos 150 o al menos 200 nucleótidos. Una porción o fragmento preferida(o) de SEQ ID NO: 1 es la secuencia que se extiende entre las posiciones de los nucleótidos 7603 y 59966 de SEQ ID NO: 1.

La secuencia puede codificar o ser complementaria de una secuencia que codifica un polipéptido de una agrupación de genes poliquétidos biosintéticos, o una porción de los mismos. Por “un polipéptido de una agrupación de genes biosintéticos de un poliquétido” se entiende un polipéptido codificado por uno o más marcos de lectura abierta de una agrupación de genes biosintéticos de un poliquétido, y particularmente la agrupación del gen biosintético de borrelidina.

Una agrupación de genes biosintéticos de poliquétidos es un segmento de DNA que comprende una pluralidad de genes que codifican polipéptidos que tienen actividad en la biosíntesis de un resto poliquétido o macrólido. Esto no está restringido a componentes de la poliquétido-sintasa (PKS) que funcionan *inter alia* en la síntesis de la cadena principal del poliquétido y al procesamiento reductor de los grupos laterales, sino que abarca también polipéptidos que tienen funcionales auxiliares en la síntesis del poliquétido. Así, los polipéptidos de la agrupación de genes biosintéticos pueden actuar también en la modificación del anillo macrólido o la cadena del poliquétido (v.g. catalizando una reacción en la formación del resto nitrilo en C12 de la borrelidina), en la síntesis de una unidad precursora o iniciadora para un poliquétido o resto macrólido (v.g. catalizando una reacción en la síntesis de la unidad iniciadora de ácido *trans*-ciclopentano-1,2-dicarboxílico para la PKS de la borrelidina, o responsable de la activación de moléculas tales como los tioésteres de la coenzima-A de las unidades iniciadoras y extendedoras de la cadena), actividad reguladora (v.g. regulación de la expresión de los genes o proteínas implicados en la síntesis del poliquétido o el macrólido), actividad transportadora (v.g. en el transporte de sustratos para el resto poliquétido o macrólido en la célula, o de productos de síntesis tales como la molécula de poliquétido o macrólido fuera de la célula), y en conferir resistencia de la célula productora a los productos sintetizados (v.g. por fijación específica a la molécula sintetizada, o como un reemplazamiento para otras proteínas endógenas a las cuales puede fijarse la molécula sintetizada dentro o fuera de la célula).

La agrupación de genes incluye también regiones no codificantes, tales como promotores y otras secuencias reguladoras de la transcripción que están enlazadas operativamente a las regiones codificantes de la agrupación de genes. Las personas expertas están perfectamente capacitadas para identificar tales elementos sobre la base de la fracción proporcionada en esta memoria, y éstos están dentro del alcance de la presente exposición.

Los genes y marcos de lectura abierta codificados dentro de SEQ ID NO: 1 representan partes o fragmentos preferidos de SEQ ID NO: 1. Así, una molécula de ácido nucleico aislada puede comprender una secuencia que codifica un polipéptido de una agrupación de genes biosintéticos de borrelidina, en donde dicho polipéptido tiene una secuencia de aminoácidos seleccionada del grupo constituido por SEQ ID Nos. 2 a 43 y 113.

En realizaciones preferidas, la secuencia de ácido nucleico comprende un marco de lectura abierta seleccionado del grupo de marcos de lectura abierta de SEQ ID NO: 1 constituidos por *borA1*, *borA2*, *borA3*, *borA4*, *borA5*, *borA6*, *borB*, *borC*, *borD*, *borE*, *borF*, *borG*, *borH*, *borI*, *borJ*, *borK*, *borL*, *borM*, *borN*, *borO*, *orfB1*, *orfB2*, *orfB3*, *orfB4*, *orfB5*, *orfB6*, *orfB7*, *orfB8*, *orfB9*, *orfB10*, *orfB11*, *orfB12*, *orfB13a*, *orfB13b*, *orfB14*, *orfB15*, *orfB16*, *orfB17*, *orfB18*, *orfB19*, *orfB20*, *orfB21* y *orfB22*, estando descritos respectivamente dichos marcos de lectura abiertos por las bases 16184*-18814, 18875-23590, 23686-34188, 34185*-39047, 39122*-45514, 45514-50742, 7603-8397c, 8397-9194c, 9244-9996c, 9993-11165c, 11162-11980c, 11992-13611c, 13608-15659*c, 50739*-52019, 52113-53477, 53486-54466, 54506-56176, 56181*-57098, 57112-57858, 57939-59966, 2-313 (incompleta), 501*-3107, 3172-3810c, 3935-4924c, 5123-5953, 5961-6518*c, 6564*-7538, 60153-60533*c, 60620-61003, 61188*-61436, 61526-61738, 61767-62285c, 62750-63067c, 62586-62858c, 63155-65071c, 65374-65871, 65942-68305*c, 68290-68910*c, 69681-70436, 70445-71848, 71851-72957, 73037-73942 y 73995-74534c de SEQ ID No.1.

En la lista anterior, “c” indica que el gen está codificado por la cadena complementaria a la que se muestra en SEQ ID NO: 1. Cada marco de lectura abierta anterior representa el marco de lectura abierta probable más largo presente. A veces ocurre que puede identificarse más de un codón potencial inicial. Un experto en la técnica reconocerá esto y será capaz de identificar codones iniciales posibles alternativos. Aquellos genes que tienen más de un codón de iniciación posible se indican con un símbolo “*”. A lo largo de todo el documento, los autores han indicado lo que se cree que es el codón de iniciación; sin embargo, una persona con experiencia en la técnica apreciará que puede ser posible generar proteína activa utilizando un codón de iniciación alternativo, y las proteínas generadas utilizando estos codones de iniciación alternativos se consideran también dentro del alcance de la presente exposición.

Debe indicarse que cierto número de estos marcos de lectura abierta comienzan con un codón (GTG, CTG o TTG) distinto del codón de iniciación ATG más normal. Es bien sabido que en algunos sistemas bacterianos tales

codones, que denotan normalmente valina (GTG) o leucina (CTG, TTG) pueden leerse como codones de iniciación que codifican metionina en el término N de la cadena del polipéptido. En las secuencias de aminoácidos (SEQ ID Nos. 2 a 43 y 113) proporcionados en esta memoria, tales codones se traducen por tanto como metionina.

5 Se proporcionan también moléculas de ácido nucleico que comprenden porciones de los marcos de lectura abierta identificados en esta memoria. Por ejemplo, una secuencia de ácido nucleico de este tipo puede comprender uno o más dominios aislados derivados de los marcos de lectura abierta identificados en esta memoria. Los polipéptidos codificados por estas porciones aisladas de los marcos de lectura abierta pueden tener actividad independiente, v.g. actividad catalítica. En particular, los polipéptidos que constituyen la PKS de borrelidina tienen estructuras modulares en las cuales dominios individuales tienen actividades catalíticas particulares como se ha expuesto anteriormente. Así, cualquiera de estos dominios puede expresarse solo o en combinación, con otros polipéptidos de la PKS de borrelidina descritos en esta memoria o dominios de los mismos, o con polipéptidos de la PKS de otros poliquétidos. En particular, cualquiera de estos dominios puede emplearse en sustitución de los dominios equivalentes sea dentro de la PKS de borrelidina o en otras poliquétido-sintasas y dominios adicionalmente equivalentes de otras PKSs pueden emplearse en sustitución de los dominios dentro de la PKS de borrelidina. En este contexto, un dominio equivalente incluye dominios que tienen el mismo tipo de función pero difieren, por ejemplo, en su especificidad; un ejemplo de sustituciones contempladas por la presente exposición incluye: el empleo de un dominio AT específico de malonil-CoA en lugar de un dominio AT específico de metilmalonil-CoA, o el empleo de un bucle reductor que contiene un dominio KR únicamente en lugar de uno que contiene KR, DH y ER. En realizaciones preferidas, los dominios expresados representan al menos un módulo de PKS como se describe más adelante.

El término “dominio de PKS” tal como se utiliza en esta memoria, hace referencia a una secuencia de polipéptido, capaz de plegarse independientemente del resto de la PKS, y que tiene una actividad enzimática individual diferenciada u otra función en la síntesis del poliquétido o macrólido con inclusión, pero sin carácter limitante, de la proteína-sintasa portadora de β -cetoacil-acilo (KS), proteína portadora de acilo (ACP), acil-transferasa (AT), β -ceto-reductasa (KR), deshidratasa (DH), enoil-reductasa (ER) o tioesterasa terminal (TE).

De acuerdo con ello, la exposición proporciona adicionalmente:

- 30 (a) una molécula de ácido nucleico aislada que comprende una secuencia que codifica un dominio de PKS seleccionado de AT0 y ACP0, estando descritos dichos dominios por, respectivamente, los aminoácidos 322-664 y 694-763 de SEQ ID NO: 2. En una realización preferida, el dominio de PKS comprende una secuencia seleccionada del grupo constituido por las bases 17147-18175 y 18263-18472 de SEQ ID NO: 1;
- 35 (b) una molécula de ácido nucleico aislada que comprende una secuencia que codifica un dominio de PKS seleccionado de KS1, AT1, KR1 y ACP1, estando descritos dichos dominios, respectivamente, por los aminoácidos 34-459, 557-885, 1136-1379 y 1419-1486 de SEQ ID NO: 3. En una realización preferida, el dominio de PKS comprende una secuencia seleccionada del grupo constituido por las bases 18974-20251, 20543-21529, 22280-23011 y 23129-23332 de SEQ ID NO: 1;
- 40 (c) una molécula de ácido nucleico aislada que comprende una secuencia que codifica un dominio de PKS seleccionado de KS2, AT2, DH2, KR2, ACP2, KS3, AT3, DH3, KR3 y ACP3, estando descritos dichos dominios, respectivamente, por los aminoácidos 34-459, 559-887, 903-1050, 1354-1597, 1628-1694, 1724-2149, 2245-2576, 2593-2734, 3060-3307 y 3340-3406 de SEQ ID NO: 4. En una realización preferida, el dominio de PKS comprende una secuencia seleccionada del grupo constituido por las bases 23785-25062, 25360-26346, 26392-26835, 27745-28476, 28567-28767, 28855-30132, 30418-31413, 31462-31887, 32863-33606 y 33703-33903 de SEQ ID NO: 1;
- 45 (d) una molécula de ácido nucleico aislada que comprende una secuencia que codifica un dominio de PKS seleccionado de KS4, AT4, KR4 y ACP4, estando descritos dichos dominios por, respectivamente, los aminoácidos 34-459, 555-886, 1179-1423 y 1459-1525 de SEQ ID NO: 5. En una realización preferida, el dominio de PKS comprende una secuencia seleccionada del grupo constituido por las bases 34284-35561, 35847-36842, 37719-38453 y 38559-38759 de SEQ ID NO: 1;
- 50 (e) una molécula de ácido nucleico aislada que comprende una secuencia que codifica un dominio de PKS seleccionado de KS5, AT5, DH5, ER5, KR5 y ACP5, estando descritos dichos dominios, respectivamente, por los aminoácidos 34-457, 553-888, 905-1046, 1401-1690, 1696-1942 y 1975-2041 de SEQ ID NO: 6. En una realización preferida, el dominio de PKS comprende una secuencia seleccionada del grupo constituido por las bases 39221-40492, 40778-41785, 41834-42259, 43322-44191, 44207-44947 y 45044-45244 de SEQ ID NO: 1;
- 55 (f) una molécula de ácido nucleico aislada que comprende una secuencia que codifica un dominio de PKS seleccionado de KS6, AT6, KR6, ACP6 y TE, estando descritos dichos dominios, respectivamente, por los aminoácidos 37-457, 555-883, 1101-1335, 1371-1437 y 1461-1708 de SEQ ID NO: 7. En una realización preferida, el dominio de PKS comprende una secuencia seleccionada del grupo constituido por las bases 45622-46884, 47176-48162, 48814-49518, 49624-49824 y 49894-50637 de SEQ ID NO: 1.
- 60
- 65

En otro de sus aspectos, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia que codifica un módulo PKS, seleccionándose dicho módulo del grupo constituido por los aminoácidos 322-763 de SEQ ID NO: 2, 34-1486 de SEQ ID NO: 3, 34-1694 de SEQ ID NO: 4, 1724-3406 de SEQ ID NO: 4, 34-1525 de SEQ ID NO: 5, 34-2041 de SEQ ID NO: 6 y 37-1437 o 1708 de SEQ ID NO: 7. En una realización preferida, el módulo comprende una secuencia seleccionada del grupo constituido por las bases 17147-18472, 18974-23332, 23785-28767, 28855-33903, 34284-38759, 39221-45244, 45622-49824 o 50637 de SEQ ID NO: 1.

El término “módulo”, como se utiliza en esta memoria, hace referencia a un polipéptido simple que comprende una pluralidad de dominios PKS, cada uno de los cuales tiene una actividad enzimática individual distinta en la síntesis de poliquétidos o macrólidos que incluye, pero sin carácter limitante, la proteína-sintasa portadora de β -cetoacil-acilo (KS), aciltransferasa (AT), proteína portadora de acilo (ACP), β -ceto-reductasa (KR), deshidratasa (DH), o enoil-reductasa (ER) o tioesterasa terminal (TE). Un módulo de extensión comprende típicamente un dominio KS, AT y ACP (aunque algunas PKSs modulares pueden codificar sus dominios AT como proteínas independientes). Un módulo de extensión puede comprender adicionalmente uno o más dominios capaces de reducir un grupo β -ceto a un grupo hidroxilo, enófilo o metileno (haciéndose referencia en esta memoria a dicho grupo de dominios como un “bucle reductor”). Así pues, un módulo que comprende un bucle reductor contiene típicamente un dominio KR, dominios KR y DH, o dominios KR, DH y ER.

Una PKS puede comprender además un dominio TE para realizar la terminación de la cadena y/o la ciclación del producto final, o alternatively puede contener otra funcionalidad que se sepa realiza una función similar tal como la adición de un residuo de aminoácido y formación de macrolactama como se observa para la rapamicina (Schwecke *et al.*, 1995), para formación de macrolactama como en el caso de la rifamicina (August *et al.*, 1998), y para incorporación de aminoácidos seguida por eliminación reductora como en el caso de la biosíntesis de la mixalamida (Silakowski *et al.*, 2001).

Las secuencias proporcionadas en esta memoria proporcionan medios con los cuales es posible manipular y/o mejorar la síntesis de los poliquétidos. Así, se proporciona un método de modificación de una poliquétido-sintasa parental, que comprende expresar un dominio de una poliquétido-sintasa de borrelidina o un derivado del mismo como se describe en esta memoria en una célula hospedadora que expresa dicha poliquétido-sintasa parental, de tal modo que el dominio se incorpora en dicha poliquétido-sintasa parental. Se proporciona adicionalmente un método de modificación de una poliquétido-sintasa parental, que comprende introducir en una célula hospedadora un ácido nucleico que codifica un dominio de una poliquétido-sintasa de borrelidina, un derivado del mismo, en donde la célula hospedadora contiene ácido nucleico que codifica dicha poliquétido-sintasa parental, de tal manera que, cuando se expresa, el dominio se incorpora en dicha poliquétido-sintasa parental. El dominio de PKS de borrelidina puede insertarse además de los dominios nativos de la PKS parental, o puede reemplazar un dominio parental nativo. Típicamente, la PKS parental será una PKS tipo I.

La presente descripción proporciona adicionalmente métodos de modificación de una PKS de borrelidina parental. Un dominio donante (v.g. de una PKS tipo I) puede expresarse en una célula hospedadora que expresa dicha PKS de borrelidina parental. Se proporciona adicionalmente un método de modificación de una poliquétido-sintasa de borrelidina parental que comprende introducir en una célula hospedadora un ácido nucleico que codifica un dominio de una poliquétido-sintasa donante, en donde la célula hospedadora contiene ácido nucleico que codifica dicha poliquétido-sintasa de borrelidina parental, de tal manera que, cuando se expresa, el dominio se incorpora en dicha poliquétido-sintasa de borrelidina parental.

Adicional o alternatively, un dominio de la PKS parental puede deleccionarse o desactivarse de otro modo; v.g. un dominio parental puede deleccionarse simplemente, o reemplazarse por un dominio de una PKS donante, o puede añadirse un dominio de una PKS donante al parental. En el caso en que se añade o reemplaza un dominio, el dominio donante puede derivarse de la sintasa parental, o de una sintasa diferente.

Estos métodos pueden utilizarse para mejorar la biosíntesis de la borrelidina, para producir nuevos derivados o análogos de borrelidina, u otras estructuras de poliquétidos o macrólidos. El número y la naturaleza de módulos en el sistema puede alterarse para cambiar el número y tipo de unidades extendedoras reclutadas, y para cambiar las diversas actividades de sintasa, reductasa y deshidratasa que determinan la estructura de la cadena del poliquétido. Tales cambios pueden hacerse por alteración del orden de los módulos que comprenden la PKS, por la duplicación o eliminación de módulos que comprenden la PKS, por la introducción de módulos de fuentes heterólogas, o por cualquier combinación de estos diversos enfoques.

Así, los dominios o módulos de la PKS de borrelidina pueden deleccionarse, duplicarse, o intercambiarse con otros dominios o módulos de la PKS de borrelidina, o de sistemas PKS responsables de la síntesis de otros poliquétidos (sistemas PKS heterólogos, particularmente sistemas PKS tipo I), que pueden ser de diferentes cepas o especies bacterianas. Alternatively, dominios o módulos de la PKS de borrelidina pueden introducirse en sistemas PKS heterólogos a fin de producir nuevos poliquétidos o macrólidos. Módulos combinatorios pueden intercambiarse también entre la poliquétido-sintasa de borrelidina y otras poliquétido-sintasas, extendiéndose estos módulos combinatorios entre dominios correspondientes de dos módulos de tipo natural, v.g. desde el AT de un módulo al AT del siguiente.

Por ejemplo, un módulo extendedor particular puede intercambiarse por uno que tenga especificidad para una unidad extendedora diferente (como se describe v.g. en WO 98/01571 y WO 98/01546), o mutarse para exhibir es-

pecificidad o selectividad para una unidad extendidora diferente, v.g. como se describe más adelante. Adicional o alternativamente, introducción, delección, intercambio o mutación de dominios o módulos, tales como los dominios KR, DH y ER responsables del procesamiento de un resto β -ceto dado, pueden utilizarse para alterar el nivel de procesamiento reductor de una unidad extendidora durante la síntesis de los poliquétidos. Tales cambios pueden conducir también a cambios en la estereoquímica de los grupos alfa-alquilo y β -hidroxilo así formados por módulos alterados. En una realización preferida, el módulo borA5 puede introducirse en una PKS parental para proporcionar adición iterativa de unidades extendedoras a una cadena principal de poliquétido, v.g. por expansión del tamaño del anillo de un poliquétido macrólido con relación al producido naturalmente por la PKS parental.

El módulo de carga de borrelidina es el primer módulo de carga de PKS identificado que tiene especificidad para una unidad iniciadora de ácido di-carboxílico alicíclico. Así pues, este módulo o un derivado del mismo pueden utilizarse para producir unidades iniciadoras alicíclicas en las poliquétido-sintasas heterólogas. Esto no requiere estar limitado al uso del ácido *trans*-ciclopentano-1,2-dicarboxílico utilizado normalmente como la unidad iniciadora de borrelidina. Se demuestra también en esta memoria que la molécula de carga de borrelidina es capaz de dirigir la incorporación de otras unidades iniciadoras que incluyen ácido *trans*-ciclobutano-1,2-dicarboxílico, ácido 2,3-metilsuccínico y ácido 2-metilsuccínico. La unidad iniciadora de borrelidina puede modificarse también en una célula productora de borrelidina, o reemplazarse por un módulo de carga heterólogo, para introducir unidades iniciadoras alternativas en el camino de síntesis de la borrelidina.

La posición del módulo de carga de la PKS puede seleccionarse (v.g. por fusión del mismo a una localización particular dentro de la PKS) a fin de controlar el tamaño del anillo de las moléculas poliquétido/macrólido resultantes.

Los dominios AT que determinan las unidades extendedoras ácido carboxílico-CoA-tioéster pueden deleccionarse, modificarse o reemplazarse. Los dominios ACP pueden también deleccionarse, modificarse o reemplazarse. Adicionalmente, pueden insertarse dominios que no están presentes normalmente en la PKS de borrelidina pero que se encuentran en otros sistemas PKS modulares y/o PKS/NRPS mixtos. Ejemplos incluyen, pero sin carácter limitante: dominios O-metil-transferasa, dominios C-metil-transferasa, dominios de epimerización, dominios de monooxigenasa y dominios de deshidrogenasa, dominios aminotransferasa y dominios de péptido-sintetasas no ribosómicos.

Adicionalmente, el dominio de tioesterasa de la PKS de borrelidina puede alterarse o reposicionarse (v.g. fusionarse a una localización seleccionada dentro de la PKS) a fin de cambiar su especificidad y/o con objeto de liberar moléculas poliquétido/macrólido con un tamaño de anillo seleccionado. Alternativamente, pueden insertarse dominios heterólogos de tioesterasa en la PKS de borrelidina para producir moléculas con tamaño de anillo alterado con relación a la molécula producida normalmente por la PKS parental, o para producir un ácido libre.

En otra alternativa adicional, los dominios de incorporación de aminoácidos y formación de macrolactama procedentes de sistemas NRPS/PKS mixtos tales como el de la rapamicina, o para sistemas afines tales como para la biosíntesis de rifamicina y biosíntesis de mixalamida, o módulos de sistemas NRPS (tal como para la biosíntesis de bleomicina) pueden insertarse en la PKS para producir nuevas moléculas afines a poliquétidos de origen mixto.

Los marcos de lectura abierta que codifican la PKS descrita en esta memoria pueden comprender también porciones codificantes de porciones no activas enzimáticamente que sin embargo tienen un papel funcional como regiones de entramado que separan y estabilizan los dominios enzimáticamente activos y/o módulos de la PKS a distancias y orientaciones apropiadas, y que pueden tener funciones de reconocimiento y enlace que ordenan los dominios y módulos de la PKS en la configuración espacial correcta. Así, las secuencias de ácido nucleico de la presente descripción comprenden secuencias que codifican tales regiones de entramado, sea solas o en combinación con secuencias que codifican dominios o módulos como se han descrito arriba. Se apreciará que las diversas manipulaciones de las secuencias codificantes de PKS arriba descritas pueden dar lugar a genes o sistemas PKS híbridos. Así pues, la presente descripción proporciona también ácidos nucleicos que codifican tales sistemas PKS híbridos. La descripción proporciona por tanto un constructo de ácido nucleico que comprende al menos una primera porción de ácido nucleico que codifica al menos un dominio de una PKS de borrelidina y una segunda porción o porciones de ácido nucleico que codifican al menos un dominio de PKS tipo I que es heterólogo a dicha PKS de borrelidina. En realizaciones preferidas, el constructo comprende un gen de poliquétido-sintasa híbrido, codificando dicho gen al menos un dominio de una PKS de borrelidina y al menos un dominio de PKS tipo I que es heterólogo a dicha PKS de borrelidina. Realizaciones preferidas adicionales son como se ha descrito arriba.

En un aspecto adicional, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia que codifica un polipéptido que cataliza un paso en la síntesis de una unidad iniciadora o sustrato para síntesis de poliquétidos, preferiblemente en la síntesis del resto ácido *trans*-ciclopentano-1,2-dicarboxílico utilizado como unidad iniciadora por la PKS de borrelidina. El polipéptido puede tener actividad como una deshidrogenasa, 3-oxoacil-ACP-reductasa, ciclasa, deshidrogenasa dependiente de F420, o 2-hidroxihepta-2,4-dieno-1,7-dioato-isomerasa. Preferiblemente, el polipéptido comprende la secuencia codificada por uno del grupo de genes constituidos por *borC*, *borD*, *borE*, *borF*, *borG*, *borH*, *borK*, *borL*, *borM* y *borN*, como se muestra en SEQ ID NO: 8, 9, 10, 12, 13, 14, 17, 18, 19 ó 20.

Estos genes pueden estar deleccionados, desorganizados, o desactivados de cualquier otro modo en una célula productora de borrelidina a fin de anular la producción de borrelidina. Líneas de células resultantes de tales cambios pueden complementarse químicamente por la adición de ácidos carboxílicos exógenos que pueden incorporarse en

lugar de la unidad iniciadora natural. Así, pueden sintetizarse nuevas moléculas afines a borrelidina, que están iniciadas por el ácido carboxílico alimentado exógenamente. Un enfoque de este tipo se denomina mutasíntesis. Los genes responsables de la síntesis del ácido *trans*-ciclopentano-1,2-dicarboxílico pueden introducirse en una célula productora de poliquétidos heterólogos para hacer posible que dicha célula sintetice el ácido dicarboxílico alicíclico como unidad iniciadora para su propia PKS.

Así, la presente descripción proporciona adicionalmente un método para la producción de borrelidina y análogos de borrelidina con títulos mejorados, comprendiendo dicho método desorganizar *borG* en la célula hospedadora, fermentar la línea de células resultante y alimentar un ácido carboxílico exógeno. En diversas realizaciones preferidas, el ácido carboxílico exógeno es ácido *trans*-ciclopentano-1,2-dicarboxílico o el ácido carboxílico exógeno se selecciona del grupo constituido por ácido *trans*-ciclobutano-1,2-dicarboxílico, ácido 2,3-dimetil-succínico y ácido 2-metilsuccínico y/o el método adicional comprende delecionar, modificar o reemplazar uno o más genes biosintéticos de borrelidina, o dominios o módulos de poliquétido-sintasa de borrelidina. Una persona con experiencia en la técnica es sabedora de que las poliquétido-sintasas pueden expresarse también en hospedadores heterólogos, por lo que la presente descripción contempla también un método para la producción de títulos mayores de borrelidina y análogos de borrelidina en un hospedador heterólogo, comprendiendo dicho método transformar una célula hospedadora con la agrupación de genes de borrelidina entero con la excepción de *borG* o desorganizar el gen *borG* *in situ* una vez que se ha transferido la agrupación de genes.

Alternativamente, genes responsables para la síntesis de la unidad iniciadora pueden sobre-expresarse a fin de mejorar los títulos de fermentación de borrelidina o moléculas afines a borrelidina. Así pues, la presente descripción proporciona adicionalmente un método para aumentar el título de borrelidina y derivados de borrelidina o moléculas afines a borrelidina y sus derivados, comprendiendo dicho método regular en sentido creciente un gen biosintético de borrelidina implicado en la producción de la unidad iniciadora, seleccionándose dicho gen del grupo constituido por *borC*, *borD*, *borE*, *borF*, *borH*, *borK*, *borL*, *borM* y *borN*, siendo en una realización preferida el gen regulado en sentido creciente *borE* o *borL*.

En otro enfoque, los genes responsables de la síntesis de la unidad iniciadora pueden modificarse, o reemplazarse por otros genes sintéticos que dirigen la producción de ácidos carboxílicos alterados, conduciendo a la producción de moléculas afines a borrelidina. Estas técnicas pueden complementarse por la modificación del módulo de carga de la PKS como se ha descrito arriba.

En un aspecto adicional, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia que codifica polipéptido que cataliza un paso en la modificación de una cadena lateral de un resto de poliquétido, por ejemplo en la conversión de un grupo metilo en un resto nitrilo, v.g. en C12 de pre-borrelidina (14). El polipéptido puede tener actividad como una oxidasa del citocromo P450, amino-transferasa, o NAD/quinona-oxidoreductasa. Preferiblemente, el polipéptido comprende la secuencia codificada por uno del grupo de genes constituido por *borI*, *borJ*, *borK* como se muestra en SEQ ID NO: 15, 16 ó 17.

Varios de estos genes pueden estar delecionados/desactivados de tal manera que se acumulen moléculas afines a borrelidina, o metabolitos de derivación de las mismas, que representan etapas intermedias del proceso que introduce el resto nitrilo. La adición de genes heterólogos a tales sistemas puede permitir la elaboración alternativa de cualesquiera compuestos intermedios biosintéticos acumulados o metabolitos de derivación de los mismos. Alternativamente, los genes pueden mutarse a fin de alterar su especificidad de sustrato de tal modo que funcionen en posiciones alternativas de moléculas pre-borrelidina a fin de proporcionar moléculas afines a borrelidina. Adicionalmente, los genes responsables de la formación del grupo nitrilo pueden sobre-expresarse a fin de mejorar los títulos de fermentación de borrelidina o moléculas afines a borrelidina.

Alternativamente, uno, algunos o la totalidad de estos genes pueden introducirse en células capaces de producir otros poliquétidos para proporcionar el procesamiento deseado de la cadena lateral de dicho poliquétido, v.g. la introducción de un resto nitrilo. Esto abre la posibilidad de introducción específica biosintética de restos nitrilo en poliquétidos, particularmente en cadenas laterales derivadas de unidades extendedoras metilmalonil-CoA o etilmalonil-CoA. También pueden utilizarse enzimas purificadas (véase más adelante) para efectuar la conversión de las cadenas laterales de poliquétido en restos nitrilo *in vitro*.

Las secuencias de nucleótidos de la descripción pueden ser porciones de la secuencia que se muestra en SEQ ID NO: 1, o el complemento de la misma, o mutantes, variantes, derivados o alelos de estas secuencias. Las secuencias pueden diferir de la representada por un cambio que es uno o más de adición, inserción, delección y sustitución de uno o más nucleótidos de la secuencia representada. Cambios a una secuencia de nucleótidos codificante pueden dar como resultado un cambio de aminoácido al nivel de proteína, o no, tal como se determina por la redundancia del código genético. Así, el ácido nucleico de acuerdo con la presente descripción puede incluir una secuencia diferente de la secuencia representada en SEQ ID NO: 1 pero codificar todavía un polipéptido con la misma secuencia de aminoácidos. Preferiblemente mutantes, variantes, derivados o alelos de las secuencias proporcionadas codifican polipéptidos que tienen la misma actividad enzimática que los descritos en esta memoria.

En el caso en que la secuencia es una secuencia codificante, el polipéptido codificado puede comprender una secuencia de aminoácidos que difiere en uno o más residuos de aminoácido de las secuencias de aminoácidos representadas en SEQ ID Nos. 2 a 43 y 113. Se proporciona adicionalmente por la presente descripción un ácido nucleico

que codifica un polipéptido que es una secuencia de aminoácidos mutante, variante, derivado o alelo de cualquiera de las secuencias mostradas. Tales polipéptidos se exponen más adelante. Un ácido nucleico que codifica un polipéptido de este tipo puede exhibir una identidad mayor que aproximadamente 60% con la secuencia codificante de SEQ ID NO: 1, identidad mayor que aproximadamente 70%, identidad mayor que aproximadamente 80%, o identidad mayor que aproximadamente 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% con el mismo. El porcentaje de identidad puede calcularse utilizando uno de los programas tales como BLAST o BestFit comprendidos dentro del paquete de soporte lógico Genetics Computer Group (GCG) Version 10, disponible de la Universidad de Wisconsin, utilizando parámetros por defecto.

En realizaciones preferidas, sean codificantes o no codificantes, las secuencias de nucleótidos de la descripción son capaces de hibridación específica con al menos una porción de la secuencia de SEQ ID NO: 1 o el complemento de la misma.

Por ejemplo, pueden realizarse hibridaciones de acuerdo con el método de Sambrook *et al.* (Sambrook *et al.*, 1989), utilizando una solución de hibridación que comprende: 5X SSC, 5X Reactivo Denhardt, 0,5-1,0% SDS, 100 µg/ml de DNA de esperma de salmón desnaturalizado y fragmentado, 0,05% de pirofosfato de sodio y hasta 50% de formamida. La hibridación se lleva a cabo a 37-42°C durante al menos 6 horas. Después de la hibridación, los filtros se lavan como sigue: (1) 5 minutos a la temperatura ambiente en 2X SSC y 1% SDS; (2) 15 minutos a la temperatura ambiente en 2X SSC y 0,1% SDS; (3) 30 minutos-1 hora a 37°C en 1X SSC y 1% SDS; (4) 2 horas a 42-65°C en 1X SSC y 1% SDS, cambiando la solución cada 30 minutos.

Una fórmula común para calcular las condiciones de severidad requeridas para conseguir hibridación entre moléculas de ácido nucleico de una homología de secuencia especificada es (Sambrook *et al.*, 1989):

$$T_m = 81,5^{\circ}\text{C} + 16,6\text{Log}[\text{Na}^+] + 0,41 (\% \text{ G} + \text{C}) - 0,63 (\% \text{ formamida}) - 600/\text{\#pb en dúplex}$$

Como ilustración de la fórmula anterior, utilizando $[\text{Na}^+] = [0,368]$ y 50% formamida, con contenido de GC de 42% y un tamaño medio de sonda de 200 bases, la T_m es 57°C. La T_m de un dúplex de DNA disminuye en 1-1,5°C con cada 1% de disminución de homología. Así pues, dianas con identidad de secuencia mayor que aproximadamente 75% podrían observarse utilizando una temperatura de hibridación de 42°C. Dicha hibridación se consideraría sustancialmente específica para la secuencia de ácido nucleico de la presente descripción.

Los ácidos nucleicos de la presente descripción comprenden preferiblemente al menos 15 nucleótidos contiguos de SEQ ID NO: 1. Los mismos pueden comprender 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 150, 200, 300, 500 ó más nucleótidos contiguos de SEQ ID NO: 1.

Los ácidos nucleicos pueden utilizarse como p.ej. cebadores o sondas para la identificación de nuevos genes u otros elementos genéticos, tales como secuencias reguladoras de la transcripción, de agrupaciones de genes biosintéticos de poliquétidos o macrólidos, v.g. secuencias codificantes de enzimas de la PKS o dominios o módulos de las mismas enzimas implicadas en la biosíntesis de una unidad iniciadora, enzimas modificadoras de cadenas laterales de restos de poliquétidos, transportadores, genes de resistencia y moléculas reguladoras como se describe.

Así pues, la presente descripción proporciona un método de identificación de una nueva agrupación de genes biosintéticos de poliquétidos, o una porción de la misma, que comprende hibridar una muestra de ácido nucleico diana con un ácido nucleico de la presente descripción capaz de hibridarse específicamente a un ácido nucleico que tenga la secuencia de SEQ ID NO: 1 o una porción del mismo. El ácido nucleico diana puede ser cualquier ácido nucleico adecuado, y preferiblemente es DNA genómico bacteriano.

Típicamente, el método comprende además el paso de detectar hibridación entre la muestra de ácido nucleico y el ácido nucleico de la descripción. La hibridación puede medirse utilizando cualquiera de una diversidad de métodos a disposición de los expertos en la técnica. Por ejemplo, las sondas pueden marcarse por radiactividad, por fluorescencia o por adición de enzimas. Otros métodos que no emplean marcación de las sondas incluyen amplificación utilizando PCR, escisión con RNAsas, y utilización de sondas de oligonucleótidos alelo-específicas.

Un método puede incluir la hibridación de una o más (v.g. dos) sondas o iniciadores al ácido nucleico diana. Donde el ácido nucleico es DNA bicatenario, la hibridación irá precedida generalmente por desnaturalización para producir DNA monocatenario. La hibridación puede ser como parte de un procedimiento PCR, o como parte de un procedimiento de sondeo que no implique PCR. Un procedimiento ilustrativo podría ser una combinación de PCR e hibridación de baja severidad. Un procedimiento de exploración, seleccionado de los muchos disponibles para los expertos en la técnica, se utiliza para identificar con éxito sucesos de hibridación y el ácido nucleico hibridado aislado.

Los expertos en la técnica están perfectamente capacitados para emplear condiciones adecuadas de la severidad deseada para hibridación selectiva, teniendo en cuenta factores tales como la longitud de oligonucleótidos y composición de las bases, temperatura, etcétera, como se ha descrito arriba.

Una molécula de ácido nucleico aislada de la descripción puede ser una molécula de ácido nucleico aislada existente naturalmente (es decir aislada o separada de los componentes con los cuales se encuentra normalmente en la

naturaleza) tal como libre o esencialmente libre del ácido nucleico flanqueante del gen en el genoma bacteriano, excepto posiblemente una o más secuencias reguladoras para la expresión. El ácido nucleico puede ser total o parcialmente sintético y puede incluir DNA genómico, cDNA o RNA. En los casos en que el ácido nucleico de acuerdo con la descripción incluye RNA, la referencia a la secuencia representada debería interpretarse como referencia al RNA equivalente, con U en lugar de T.

La presente descripción proporciona adicionalmente un vector que comprende un ácido nucleico de acuerdo con la presente descripción. El vector es preferiblemente un vector de expresión que comprende un ácido nucleico que codifica un polipéptido de una agrupación de genes biosintéticos de poliquétido (preferiblemente una agrupación de genes biosintéticos de borrelidina), o una porción de la misma, como se describe. Pueden seleccionarse o construirse vectores adecuados que comprenden ácido nucleico para introducción en bacterias u hospedadores eucariotas, que contienen secuencias reguladoras apropiadas, que incluyen secuencias promotoras, fragmentos terminadores, secuencias intensificadoras, genes marcadores y otras secuencias en caso apropiado. Los vectores pueden ser plásmidos, virales, v.g. “fago”, o “fagémido”, según sea apropiado. Para detalles adicionales véase, por ejemplo, Sambrook *et al.*, 1989. Muchas técnicas conocidas y protocolos para manipulación de ácido nucleico, por ejemplo en la preparación de constructos de ácido nucleico, mutagénesis, secuenciación, introducción de DNA en células y expresión de genes, y análisis de proteínas, se describen en detalle en Short Protocols in Molecular Biology, Segunda Edición, Ausubel *et al.* Eds, John Wiley & Sons 1992. Las descripciones de Sambrook *et al.* y Ausubel *et al.* se incorporan en esta memoria por referencia.

En otro de sus aspectos, la presente descripción proporciona un polipéptido aislado codificado por una molécula de ácido nucleico de la descripción como se describe en esta memoria. Más particularmente, se proporciona un polipéptido aislado que comprende una secuencia de aminoácidos como se muestra en una cualquiera o más de SEQ ID Nos. 2 a 43 y 113 o una porción de las mismas. Como se ha expuesto anteriormente, estas secuencias de aminoácidos representan traducciones de los marcos de lectura abierta más largos posibles presentes en la secuencia de SEQ ID NO: 1 y el complemento de la misma. El primer aminoácido se representa siempre como Met, con indiferencia de si el codón de iniciación es ATG, GTG, CGT o TTG.

Tal como se utiliza en esta memoria, el término “polipéptido(s)” incluye péptidos, polipéptidos y proteínas, utilizándose estos términos de modo intercambiable a no ser que se especifique otra cosa.

Un polipéptido que es una secuencia de aminoácidos variante, alelo, derivado o mutante de una cualquiera de las secuencias de aminoácidos representadas puede exhibir al menos 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad con el polipéptido de una cualquiera de las SEQ ID Nos. 2 a 43 y 113, o con una porción de la misma. Variantes de secuencias de aminoácidos particulares pueden diferir de las representadas por inserción, adición, sustitución o delección de 1 aminoácido, 2, 3, 4, 5-10, 10-20, 20-30, 30-50, 50-100, 100-150, o más de 150 aminoácidos. El porcentaje de identidad puede calcularse utilizando uno de los programas tales como FASTA o BestFit del paquete de soporte lógico de Genetics Computer Group (GCG) Version 10, disponible de la Universidad de Wisconsin, utilizando parámetros por defecto.

La presente descripción incluye también porciones activas, fragmentos, y derivados de los polipéptidos de la descripción.

Una “porción activa” significa un péptido que es menor que el polipéptido de longitud total, pero que requiere al menos algo de su actividad biológica esencial. Por ejemplo, dominios o módulos aislados de la PKS como se ha descrito arriba pueden considerarse como porciones activas de la PKS.

Un “fragmento” significa un tramo de residuos de aminoácidos de al menos cinco, al menos seis, o al menos siete aminoácidos contiguos, a menudo al menos ocho o al menos nueve aminoácidos contiguos, típicamente al menos 10, al menos 13 aminoácidos contiguos y, muy preferiblemente, al menos 20, al menos 25, al menos 30, al menos 50, al menos 75, al menos 100 o más aminoácidos contiguos. Fragmentos de la secuencia pueden comprender determinantes antigénicos o epítopes útiles para generar anticuerpos para una porción del polipéptido relevante. Así, el polipéptido no necesita comprender una secuencia completa proporcionada en cualquiera de SEQ ID Nos. 2 a 43 y 113, sino que puede comprender una porción de las mismas que tenga la actividad deseada, v.g. un dominio o módulo aislado, tal como los de la PKS arriba descrita. Debe indicarse que los términos parte, porción y fragmento se utilizan intercambiablemente en esta memoria descriptiva; no debe adscribirse significado particular alguno al uso específico de uno de estos términos en cualquier contexto particular.

Un “derivado” de un polipéptido de la descripción o un fragmento del mismo significa un polipéptido modificado por variación de la secuencia de aminoácidos de la proteína, v.g. por manipulación del ácido nucleico codificante de la proteína o por alteración de la proteína propiamente dicha. Tales derivados de la secuencia de aminoácidos natural pueden implicar inserción, adición, delección o sustitución de uno, dos, tres, cinco o más aminoácidos, sin alterar fundamentalmente la actividad esencial del polipéptido de tipo salvaje.

Los polipéptidos de la descripción se proporcionan en forma aislada, v.g. aislados de uno o más componentes con los cuales se encuentran normalmente asociados en la naturaleza. Aquéllos pueden aislarse de un hospedador en el cual se expresan los mismos naturalmente, o pueden ser sintéticos o recombinantes.

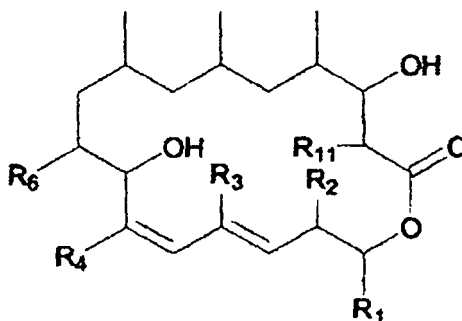
La presente descripción abarca también un método de producción de un polipéptido (como se describe), incluyendo el método la expresión de ácido nucleico codificante del polipéptido (generalmente ácido nucleico de acuerdo con la descripción). Esto puede conseguirse convenientemente dejando crecer una célula hospedadora en cultivo, que contenga un vector de expresión como se ha descrito arriba, en condiciones apropiadas que causen o permitan la expresión del polipéptido. Los polipéptidos pueden expresarse también en sistemas *in vitro*, tales como sistemas de lisado de reticulocitos.

El método puede incluir el paso de introducir el ácido nucleico en una célula hospedadora. La introducción, a la cual puede hacerse referencia generalmente sin limitación como "transformación" (particularmente para la introducción *in vitro*), puede emplear cualquier técnica disponible. Para células eucariotas, técnicas adecuadas pueden incluir transfección con fosfato de calcio, DEAE-Dextrano, electroporación, transfección mediada por liposomas y transducción utilizando retrovirus u otros virus, v.g. vaccinia o, en el caso de células de insecto, baculovirus. Para células bacterianas, técnicas adecuadas pueden incluir transformación con cloruro de calcio, conjugación, electroporación y transfección utilizando bacteriófago. Como alternativa, podría emplearse la inyección directa del ácido nucleico. Genes marcadores tales como genes de resistencia o sensibilidad a antibióticos pueden utilizarse en la identificación de clones que contienen el ácido nucleico de interés, como es bien conocido en la técnica.

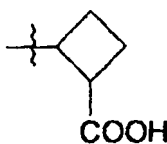
Células hospedadoras preferidas incluyen Actinomicetos, preferiblemente Streptomicetos, y en particular los seleccionados del grupo constituido por *Saccharopolyspora erythraea*, *Streptomyces coelicolor*, *Streptomyces avermitilis*, *Streptomyces griseofuscus*, *Streptomyces cinnamonensis*, *Micromonospora griseorubida*, *Streptomyces hygrosopicus*, *Streptomyces fradiae*, *Streptomyces longisporoflavus*, *Streptomyces lasaliensis*, *Streptomyces tsukubaensis*, *Streptomyces griseus*, *Streptomyces venezuelae*, *Streptomyces antibioticus*, *Streptomyces lividans*, *Streptomyces rimosus* y *Streptomyces albus*, *Streptomyces rochei* ATCC23956, *Streptomyces parvulus* Tü113 y *Streptomyces parvulus* Tü4055, más preferiblemente seleccionados del grupo constituido por *Streptomyces rochei* ATCC23956, *Streptomyces parvulus* Tü113 y *Streptomyces parvulus* Tü4055.

En su aspecto principal, la invención proporciona moléculas que pueden derivarse de los objetos de la descripción y compuestos modificados formados a partir de ellas y métodos para su producción. Las moléculas derivadas de los objetos de la descripción se representan por la Fórmula 1 y se extienden a sales farmacéuticamente aceptables de la misma, en donde:

Fórmula 1



R₁ es un grupo cicloalquilo sustituido como se muestra a continuación;



en donde R₁ puede estar sustituido también opcionalmente con uno o más átomos de halógeno o uno o más grupos alquilo C₁ a C₃; R₂, R₃, R₆, o R₁₁ son cada uno independientemente H, OCH₃, CH₃ o CH₂CH₃; y R₄ es CN, CO₂H o CH₃.

En realizaciones preferidas,

(a) R₄ es CH₃ o COOH

(b) R₁ es ciclobutano-1'-carboxilato

(c) R₆ es CH₃, R₂ y R₁₁ son H, y R₁ es ciclobutano-1'-carboxilato

(d) R₁ es ciclobutano-1'-carboxilato, y R₄ es CH₃ o COOH.

Los compuestos de la presente invención pueden tener actividad inhibidora de la tRNA-sintetasa (v.g. los mismos pueden inhibir treonil-, tirosinil-, o triptofanil-tRNA-sintetasa). Pueden exhibir actividad anti-microbiana, con inclusión de actividad contra parásitos intra- o extracelulares y organismos tales como bacterias, espiroquetas (v.g. *Treponema*), malaria, virus y hongos. Adicional o alternativamente, aquéllos pueden tener actividad anti-proliferativa contra células de mamífero, y/o actividad anti-angiogénica, sea como resultado de la inhibición de la tRNA-sintetasa, o por cualquier otro modo de acción. Esto puede hacer que los compuestos de la presente invención sean particularmente adecuados como agentes anti-cáncer (v.g. agentes para el tratamiento de cáncer intestinal, cáncer de próstata u otros), y pueden proporcionar también aplicación en el tratamiento de otros trastornos proliferativos, tales como psoriasis, o condiciones en las cuales ocurre vascularización inadecuada, tales como psoriasis, artritis reumatoide, aterosclerosis y retinopatía diabética.

Los compuestos de la presente invención pueden formularse en composiciones farmacéuticamente aceptables, v.g. por mezcla con un excipiente, vehículo, tampón, estabilizador u otros materiales farmacéuticamente aceptables bien conocidos por los expertos en la técnica. Tales composiciones caen también dentro del alcance de la presente invención.

Dichos materiales farmacéuticamente aceptables deben ser no tóxicos y no deben interferir con la eficacia del ingrediente activo. La naturaleza precisa del vehículo u otro material puede depender de la ruta de administración, v.g. las rutas oral, intravenosa, cutánea o subcutánea, nasal, intramuscular o intraperitoneal.

Las composiciones farmacéuticas para administración oral pueden encontrarse en forma de tableta, cápsula, polvo o líquido. Una tableta puede incluir un vehículo sólido tal como gelatina o un adyuvante. Las composiciones farmacéuticas líquidas incluyen generalmente un vehículo líquido tal como agua, éter de petróleo, aceites animales o vegetales, aceite mineral o aceite sintético. Puede incluirse solución salina fisiológica, dextrosa u otra solución de sacárido o glicoles tales como etilenglicol, propilenglicol o polietilenglicol.

Para inyección intravenosa, cutánea o subcutánea, o inyección en el sitio de aflicción, el ingrediente activo se encontrará en la forma de una solución acuosa parenteralmente aceptable que está exenta de pirógenos y tiene pH, isotonicidad y estabilidad adecuados. Las personas con experiencia relevante en la técnica están perfectamente capacitadas para preparar soluciones adecuadas utilizando, por ejemplo, vehículos isotónicos tales como Inyección de Cloruro de Sodio, Inyección de Ringer o Inyección de Ringer con Lactato. En caso requerido, pueden incluirse conservantes, estabilizadores, tampones, antioxidantes y/u otros aditivos. Ejemplos de las técnicas y protocolos arriba mencionados pueden encontrarse en Remington's Pharmaceutical Sciences, Edición 20^a, 2000, pub. Lippincott, Williams & Wilkins.

La invención proporciona adicionalmente los compuestos y composiciones arriba descritos para uso en un método de tratamiento médico. Se proporciona también el uso de los compuestos de la invención en la preparación de un medicamento para el tratamiento de afecciones microbianas (con inclusión de la malaria), para la inhibición de la angiogénesis, para el tratamiento de trastornos proliferativos, o para el tratamiento de afecciones caracterizadas por vascularización inadecuada, como se ha descrito arriba.

Breve descripción de los dibujos

La Figura 1 ilustra la estructura de borrelidina y algunos metabolitos afines aislados de organismos productores de borrelidina.

La Figura 2 ilustra los patrones de incorporación para sustratos de extensión marcados con el isótopo estable ¹³C y la posición de los carbonos derivados de la unidad iniciadora ácido *trans*-ciclopentano-1,2-dicarboxílico.

La Figura 3 ilustra la organización de la agrupación de genes biosintéticos de borrelidina. Sitios de restricción: B, *Bam*HI; Bc, *Bc*I; E, *Eco*RI; X, *Xho*I.

La Figura 4 ilustra un esquema que muestra el camino de biosíntesis propuesto para la unidad iniciadora ácido *trans*-ciclopentano-1,2-dicarboxílico.

La Figura 5 ilustra la organización de la PKS de borrelidina y la biosíntesis de la molécula de pre-borrelidina.

La Figura 6 ilustra la ruta biosintética propuesta para la introducción del resto nitrilo en la posición C12 de borrelidina.

La Figura 7 ilustra la estructura propuesta de la molécula 6.

La Figura 8 ilustra la estructura propuesta de las moléculas 7 & 8.

La Figura 9 ilustra la caracterización molecular de la del camino catabólico del ácido 4-hidroxifenilacético en *E. coli* W.

La Figura 10 ilustra las estructuras de las moléculas 18-20.

La Figura 11 ilustra las estructuras de las moléculas 21-26.

Descripción detallada de la invención

Una genoteca de cósmidos de DNA genómico de *S. parvulus* Tü4055 se construyó utilizando fragmentos obtenidos de una digestión parcial con *Sau3AI* que se clonaron en pWE15 y se introdujeron en células *E. coli* utilizando el kit Gigapack® III Gold Packaging Extract (Stratagene). Una genoteca de 3000 transformantes de *E. coli* se exploró con respecto a homología utilizando una sonda marcada que se generó utilizando el DIG DNA Labelling and Detection Kit (Roche). La sonda utilizada era un fragmento *BglIII-BamHI* de 1,7 kpb obtenido del gen que codifica el módulo 6 de la tercera subunidad de la PKS de oleandomicina de *Streptomyces antibioticus* (Swan *et al.*, 1994).

Se seleccionaron clones que daban una respuesta positiva y se aisló el DNA de cósmido. El DNA de cósmido se digirió con *BamHI* y los fragmentos de tamaño menor que 3 kpb se sub-clonaron en pOJ260 (Bierman *et al.*, 1992). Los plásmidos se utilizaron luego para transformar protoplastos de *S. parvulus* Tü4055 y los mutantes resultantes se exploraron respecto a la capacidad para producir borrelidina. Se identificaron dos mutantes como no productores de borrelidina, los dos cuales se derivaban de plásmidos que contenían fragmentos de cosBor32A2. Estos dos fragmentos tenían un tamaño de 1,97 y 2,80 kpb, y se identificaron más tarde como fragmentos adyacentes que codificaban partes de la PKS de borrelidina (*borA2* & *borA3*). Utilizando cosBor32A2 como sonda, se identificó un segundo cósmido solapante, cosBor19B9, de la genoteca original. Estos dos cósmidos son suficientes para cubrir la agrupación de genes biosintéticos de borrelidina entero (véase la Figura 3).

La secuencia de nucleótidos completa de cosBor32A2 y cosBor19B9 se determinó por secuenciación con perdigonada de una genoteca de subclones derivados de *Sau3AI* para cada cósmido, constituida por fragmentos de 1,5-2,0 kpb en pHSG397 (Takeshita *et al.*, 1987). Detalles específicos se proporcionan en el Ejemplo 3. La secuencia codificante de nucleótidos completa solapante para cosBor32A2 y cosBor19B9 se presenta como SEQ ID NO: 1. La región codificada por el cósmido cosBor32A2 representa la secuencia de las posiciones de nucleótidos 0-40217 pb de SEQ ID NO: 1. La región codificada por el cósmido cosBor19B9 está superpuesta a esta región por 4452 nucleótidos, y corresponde a las posiciones de nucleótidos 35766-74787 pb de SEQ ID NO: 1. Como se describe con mayor detalle en el texto que sigue, los autores de la presente invención han realizado experimentos de desactivación de genes utilizando muchos de los marcos de lectura abierta identificados codificados dentro de SEQ ID NO: 1, y esto ha conducido a los autores de la invención a identificar los límites de la agrupación. La agrupación de genes biosintéticos de borrelidina está contenida entre las posiciones de nucleótidos 7603 y 59966 de SEQ ID NO: 1 (*borB* a *borO*, lo que incluye la región *borA*). Así, estos esfuerzos combinados han conducido a los autores de la invención a la identificación y secuenciación de la región de DNA que abarca la agrupación de genes biosintéticos de borrelidina entero, y a la identificación y descripción de las secuencias funcionales codificadas dentro de esta región.

Genes PKS

Codificados entre las posiciones 16184-50742 de SEQ ID NO: 1 se encuentran 6 marcos de lectura abierta que exhiben una homología muy alta con los genes que codifican las PKSs de organismos productores de macrólidos conocidos. Estos genes se designan *borA1*, *borA2*, *borA3*, *borA4*, *borA5* y *borA6*, y codifican la PKS de borrelidina como se demostró anteriormente por desorganización de una región de 1,97 kpb dentro de *borA2*. Los 6 marcos de lectura abierta están dispuestos de manera cabeza-a-cola y cada uno de ellos está terminado por un codón de parada en marco. La secuencia de nucleótidos y los detalles de la secuencia de polipéptidos correspondiente se muestran a continuación en la Tabla 1:

TABLA 1

Gen codificante de PKS	Posición de nucleótidos en SEQ ID NO: 1	Número de la secuencia de polipéptidos correspondiente
<i>borA1</i>	16184-18814	SEQ ID No.2
<i>borA2</i>	18875-23590	SEQ ID No.3
<i>borA3</i>	23686-34188	SEQ ID No.4
<i>borA4</i>	34185-39047	SEQ ID No.5
<i>borA5</i>	39122-45514	SEQ ID No.6
<i>borA6</i>	45514-50742	SEQ ID No.7

El gen *borA1* codifica el módulo de comienzo o de carga (SEQ ID NO: 1, posición 16184-18814). La asignación del codón inicial no es obvia para este marco de lectura abierta. El codón inicial dado aquí es el que los autores de la invención creen es el verdadero codón inicial, pero existen al menos otros 3 codones iniciales posibles entre el primero y el comienzo de la secuencia de dominio AT0, y una persona con experiencia en la técnica apreciará que puede ser posible generar proteína activa utilizando uno de estos codones iniciales alternativos. El codón inicial dado aquí deja una cola importante N-terminal de 321 aminoácidos que preceden al dominio AT0. Para comparación, la cola N-terminal que precede al AT0 del módulo de carga de eritromicina tiene 108 aminoácidos, y la del módulo de carga

ES 2 287 552 T3

de avermectina tiene 28 aminoácidos. Por consiguiente, es posible que uno de los otros codones iniciales candidato pudiera ser correcto; los más probables de éstos se encuentran en las posiciones 16298, 16607 y 16901 de SEQ ID NO: 1. La longitud de la cola N-terminal sugiere que la misma podría representar posiblemente una actividad catalítica, aunque no tiene homología significativa alguna con otras secuencias en las bases de datos. La posición de la secuencia de nucleótidos y la secuencia de aminoácidos correspondiente para cada uno de los dominios funcionales dentro del módulo inicial se identifican a continuación en la Tabla 2:

TABLA 2

Dominio en <i>borA1</i>	Bases en SEQ ID NO: 1	Aminoácidos en SEQ ID NO: 2
AT0	17147-18175	322-664
ACP0	18263-18472	694-763

El gen *borA2* codifica el primer módulo de extensión (SEQ ID NO: 1, Posición 18875-23590). La posición de la secuencia de nucleótidos y la secuencia de aminoácidos correspondiente para cada uno de los dominios funcionales dentro del primer módulo de extensión se identifican a continuación en la Tabla 3:

TABLA 3

Dominio en <i>borA2</i>	Bases en SEQ ID NO: 1	Aminoácidos en SEQ ID NO: 3
KS1	18974-20251	34-459
AT1	20543-21529	557-885
KR1	22280-23011	1136-1379
ACP1	23129-23332	1419-1486

El gen *borA3* codifica los módulos de extensión segundo y tercero (SEQ ID NO: 1, posición 23686-34188). La posición de la secuencia de nucleótidos y la secuencia de aminoácidos correspondiente para cada uno de los dominios funcionales dentro de los módulos de extensión segundo y tercero se identifican a continuación en la Tabla 4:

TABLA 4

Dominio en <i>borA3</i>	Bases en SEQ ID NO: 1	Aminoácidos en SEQ ID NO: 4
KS2	23785-25062	34-459
AT2	25360-26346	559-887
DH2	26392-26835	903-1050
KR2	27745-28476	1354-1597
ACP2	28567-28767	1628-1694
KS3	28855-30132	1724-2149
AT3	30418-31413	2245-2576
DH3	31462-31887	2593-2734
KR3	32863-33606	3060-3307
ACP3	33703-33903	3340-3406

ES 2 287 552 T3

El gen *borA4* codifica el cuarto módulo de extensión (SEQ ID NO: 1, posición 34185-39047). La posición de la secuencia de nucleótidos y la secuencia de aminoácidos correspondiente para cada uno de los dominios funcionales dentro del cuarto módulo de extensión se identifican a continuación en la Tabla 5:

TABLA 5

Dominio en <i>borA4</i>	Bases en SEQ ID NO: 1	Aminoácidos en SEQ ID NO: 5
KS4	34284-35561	34-459
AT4	35847-36842	555-886
KR4	37719-38453	1179-1423
ACP4	38559-38759	1459-1525

El gen *borA5* codifica el quinto módulo de extensión (SEQ ID NO: 1, posición 39122-45514). La posición de la secuencia de nucleótidos y la secuencia de aminoácidos correspondiente para cada uno de los dominios funcionales dentro del quinto módulo de extensión se identifican a continuación en la Tabla 6:

TABLA 6

Dominio en <i>borA5</i>	Bases en SEQ ID NO: 1	Aminoácidos en SEQ ID NO: 6
KS5	39221-40492	34-457
AT5	40778-41785	553-888
DH5	41834-42259	905-1046
ER5	43322-44191	1401-1690
KR5	44207-44947	1696-1942
ACP5	45044-45244	1975-2041

El gen *borA6* codifica el sexto módulo de extensión y la tioesterasa de terminación de cadena (SEQ ID NO: 1, posición 45514-50742). La posición de la secuencia de nucleótidos y la secuencia de aminoácidos correspondiente para cada uno de los dominios funcionales dentro del sexto módulo de extensión se identifican a continuación en la Tabla 7:

TABLA 7

Dominio en <i>borA6</i>	Bases en SEQ ID NO: 1	Aminoácidos en SEQ ID NO: 7
KS6	45622-46884	37-457
AT6	47176-48162	555-883
KR6	48814-49518	1101-1335
ACP6	49624-49824	1371-1437
TE	49894-50637	1461-1708

La identificación de dominios funcionales y sus límites como se describen en lo mencionado anteriormente se determinan sobre la base de las similitudes con las secuencias de aminoácidos conservadas de otras PKSs modulares tales como las correspondientes a las biosíntesis de rapamicina (Schwecke *et al.*, 1995; Aparicio *et al.*, 1996) y eri-

tromicina (Cortés *et al.*, 1990). Los límites de los dominios catalíticos se establecen sobre la base de la homología con otras agrupaciones de PKS y el punto seleccionado en el cual comienza o termina un dominio no está definido absolutamente, sino que se selecciona basándose en las consideraciones mencionadas anteriormente. En el caso de los dominios de procesamiento β -ceto es menos obvio, dado que existe típicamente una región de gran tamaño no asimilada a un dominio funcional que precede al dominio KR. Esta región puede ser estructuralmente importante, o puede ser requerida para estabilidad del dímero de PKS. Una característica extraña de la PKS de borrelidina es que todos los dominios enzimáticos individuales parecen ser catalíticamente competentes sobre la base de su secuencia de oligonucleótidos/aminoácidos, y todos ellos son necesarios a fin de proporcionar el procesamiento β -ceto requerido para producir los grupos funcionales observados en la borrelidina. Esto es bastante extraño, dado que la mayoría de las secuencias PKS modulares comunicadas hasta ahora contienen uno o más dominios inactivos, siendo una excepción por ejemplo la PKS de espinosina (Waldron *et al.*, 2001; US 6.274.50).

Un experto en la técnica está familiarizado con la degeneración del código genético; por tanto, el técnico experto puede modificar las secuencias de DNA específicas proporcionadas por esta descripción para producir proteínas que tengan las mismas características alteradas o mejoradas comparadas con los polipéptidos que se proporcionan específicamente en esta memoria. Un experto en la técnica puede modificar también las secuencias de DNA para expresar un polipéptido idéntico a los proporcionados, aunque expresado a niveles más altos. Adicionalmente, un experto en la técnica está familiarizado con medios para preparar sintéticamente, sea parcialmente o en su totalidad, secuencias de DNA que podrían ser útiles en la preparación de vectores de DNA recombinante o secuencias codificantes que están abarcados por la presente descripción. Adicionalmente, medios recombinantes para modificar las secuencias de DNA proporcionadas pueden incluir por ejemplo delección orientada o mutagénesis orientada. Estos métodos son bien conocidos por los expertos en la técnica y no precisan ser explicados aquí adicionalmente. Por consiguiente, como se utiliza en esta memoria, el DNA aislado de fuentes naturales, preparado sintéticamente o semi-sintéticamente, o que está modificado por métodos de DNA recombinantes, está dentro del alcance de la presente descripción.

Análogamente, los expertos en la técnica reconocerán que los polipéptidos de la descripción pueden expresarse recombinantemente. De modo alternativo, dichos polipéptidos pueden sintetizarse sea totalmente o en parte, por técnicas convencionales conocidas no recombinantes; por ejemplo, síntesis en fase sólida. Así, la presente descripción no debería interpretarse como limitada necesariamente a cualquier construcción de vector específica o medios para producción de las moléculas de agrupaciones biosintéticas específicas con inclusión de las moléculas de poliquétido-sintasa citadas como ejemplos.

El módulo de carga de la PKS de borrelidina existe como una proteína discreta. Esto es más bien extraño, dado que la mayoría de los módulos de carga se encuentran en la misma proteína como el primer módulo de extensión. Excepciones a ello incluyen, por ejemplo, las PKSs de nistatina (Brautaset *et al.*, 2000) y de anfotericina (Caffrey *et al.*, 2001). El módulo de carga, que está constituido por un didominio AT-ACP, es similar al módulo de carga de especificidad amplia de la PKS de avermectina, que acepta cierto número de adiciones iniciadoras alternativas, y son de uso en la generación de genotecas de nuevos poliquétidos (Marsden *et al.*, 1998; Pacey *et al.*, 1998). El dominio AT del módulo de carga de PKS de borrelidina diverge de la gran mayoría de los dominios AT dado que el residuo serina del sitio activo está reemplazado con una cisteína de tal manera que el motivo del sitio activo es GXCXG (específicamente GH CYG). En las secuencias de dominios AT de PKS tipo I disponibles, el motivo del sitio activo conservado es GX SXG; el mismo motivo se observa en lipasas, sintasas de ácidos grasos y la mayoría de las tioesterasas. La serina nucleofílica está sustituida por cisteína en dos dominios RNPS de tioesterasa, específicamente las sintetasas responsables de la producción de micobactina y piochelina (Shaw-Reid *et al.*, 1999). Un motivo GX CXG se observa también en un dominio semejante a tioesterasa de orf1 en la agrupación de bialafós (Raibaud *et al.*, 1991). Se ha sugerido que, dado que no es posible moverse entre los dos tipos de codones serina por un solo cambio de base, los sitios activos que contienen un residuo serina esencial pueden encontrarse en dos líneas de descenso desde una enzima antigua ancestral que tenía una cisteína en lugar de una serina en su sitio activo (Brenner, 1988). La presencia de enzimas que contienen cisteína en el sitio activo puede soportar este punto de vista. Alternativamente, puede ocurrir que la cisteína se presente en estos sitios activos debido a que es posible moverse desde un tipo de codón serina al otro por la vía de una cisteína que podría mantenerse activa. Los dominios AT de las PKSs seleccionan una unidad de ácido carboxílico particular como sustrato. Se ha demostrado que esta selectividad está en correlación con ciertas configuraciones de motivos dentro del dominio AT (Reeves *et al.*, 2001; WO 02/14482). El motivo del dominio AT del módulo de carga de borrelidina difiere de cualesquiera descritos hasta ahora, lo cual no es sorprendente, dado que este dominio AT es el primero en ser secuenciado que selecciona un ácido dicarboxílico alicíclico. Los dominios AT para los módulos de extensión de PKS de borrelidina exhiben el motivo del sitio activo esperado GX SXG, y contienen también cada uno los motivos esperados para la selección de malonil-CoA o metilmalonil-CoA (Reeves *et al.*, 2001; WO 02/14482). Los dominios AT selectivos de malonil-CoA (AT1, AT2 y AT6) exhiben semejanza muy alta unos respecto a otros, tanto a nivel de proteína como a nivel de DNA. Lo mismo sucede para los dominios AT selectivos de metilmalonil-CoA (AT3, AT4 y AT5); dos de estos dominios AT (AT3 y AT4) tienen secuencias de aminoácidos idénticas a todo lo largo de la región conservada. La alta semejanza de AT5 con AT3 y AT4 es evidencia de que la unidad extendedora seleccionada en el módulo 5 es metilmalonil-CoA, y que el grupo metilo C12 de la borrelidina así incorporado se modifica subsiguientemente para dar una función nitrilo después de la incorporación en la PKS.

Para demostrar que es posible alterar la estructura de borrelidina derivada de PKS, el dominio AT del módulo 4 (el dominio AT codificado por *borA4*) se reemplaza por el dominio AT del módulo 2 de la PKS de rapamicina (rapAT2) utilizando una estrategia de reemplazamiento (véase el Ejemplo 6). Esto da la cepa *S. parvulus* Tü4055/467. Después de fermentación y análisis por LCMS de los extractos de cultivo de este mutante, puede determinarse que se

produce algo de borrelidina y se observa también un compuesto nuevo más polar con un valor m/z 14 unidades menor que la borrelidina. Esto es consistente con la incorporación de un malonato más bien que una unidad extendidora metilmalonato por el módulo 4 de la PKS para producir 10-desmetil-borrelidina 3. Además de la producción por métodos de intercambio de dominios, 3 se genera también por introducción de mutaciones específicas en el motivo de selectividad del dominio AT del módulo 4 (Reeves *et al.*, 2001; WO 02/14482) (véase el Ejemplo 7). Un cambio de este tipo afecta la selectividad del dominio AT de tal modo que selecciona una molécula sustrato de malonil-CoA con preferencia a metilmalonil-CoA. Así pues, el motivo de aminoácidos YASH en las posiciones 739 a 742 de SEQ ID NO: 5 está mutado a HAFH para dar la cepa *S. parvulus* Tü4055/472. Después de fermentación y análisis por LCMS de los extractos de cultivo de este mutante, se determina que se produce borrelidina además de un nuevo compuesto más polar con un valor m/z 14 unidades menor que la borrelidina. Este nuevo compuesto es idéntico al descrito arriba y por tanto es consistente con la incorporación de un malonato más bien que una unidad extendidora metilmalonato por el módulo 4 de la PKS para producir 3.

Estos resultados indican claramente que la PKS de borrelidina es susceptible de manipulación genética y del intercambio de la secuencia nativa por la de una cepa heteróloga. Está claro para un experto en la técnica que la modificación por ingeniería genética de biosíntesis por los métodos arriba descritos, de la PKS de borrelidina conducirá a la producción de nuevas moléculas afines a borrelidina.

Genes distintos de PKS

Tanto aguas arriba como aguas abajo de los genes codificantes de PKS se encuentran otros marcos de lectura abiertos implicados en la biosíntesis de borrelidina. Un marco de lectura abierto (orf) está diseñado como constituido por al menos 100 nucleótidos contiguos, que comienzan con un codón de iniciación apropiado y termina con un codón de parada apropiado, y que tiene una preferencia de codones apropiada para regiones codificantes de proteínas de un organismo cuyo DNA es rico en los nucleótidos guanina y citosina. En la secuencia de DNA tanto aguas arriba como aguas abajo de los genes PKS de borrelidina (*borA1-borA6*) existen cierto número de orfs que podrían identificarse por comparación con otras secuencias en la base de datos NCBI (véase la Figura 3). Los detalles de secuencias de nucleótidos de estos orfs se dan a continuación en la Tabla 8:

(Tabla pasa a página siguiente)

ES 2 287 552 T3

TABLA 8

Gen	Bases en SEQ ID NO.1	Número de la secuencia de polipéptidos correspondiente
<i>borB</i>	7603-8397c	SEQ ID No.8
<i>borC</i>	8397-9194c	SEQ ID No.9
<i>borD</i>	9244-9996c	SEQ ID No.10
<i>borE</i>	9993-11165c	SEQ ID No.11
<i>borF</i>	11162-11980c	SEQ ID No.12
<i>borG</i>	11992-13611c	SEQ ID No.13
<i>borH</i>	13608-15599c*	SEQ ID No.14
<i>borI</i>	50739*-52019	SEQ ID No.15
<i>borJ</i>	52113-53477	SEQ ID No.16
<i>borK</i>	53486-54466	SEQ ID No.17
<i>borL</i>	54506-56176	SEQ ID No.18
<i>borM</i>	56181*-57098	SEQ ID No.19
<i>borN</i>	57112-57858	SEQ ID No.20
<i>borO</i>	57939-59966	SEQ ID No.21
<i>orfB1</i>	2-313	SEQ ID No.22
<i>orfB2</i>	501*-3107	SEQ ID No.23
<i>orfB3</i>	3172-3810c	SEQ ID No.24
<i>orfB4</i>	3935-4924c	SEQ ID No.25
<i>orfB5</i>	5123-5953	SEQ ID No.26
<i>orfB6</i>	5961-6518*c	SEQ ID No.27
<i>orfB7</i>	6564*-7538	SEQ ID No.28
<i>orfB8</i>	60153-60533c	SEQ ID No.29
<i>orfB9</i>	60620-61003	SEQ ID No.30
<i>orfB10</i>	61188*-61436	SEQ ID No.31
<i>orfB11</i>	61526-61738	SEQ ID No.32
<i>orfB12</i>	61767-62285c	SEQ ID No.33
<i>oriB13a</i>	62750-63067c	SEQ ID No.34
<i>orfB13b</i>	62586-62858c	SEQ ID No.113
<i>orfB14</i>	63155-65071c	SEQ ID No.35
<i>orfB15</i>	65374-65871	SEQ ID No.36
<i>orfB16</i>	65942-68305c*	SEQ ID No.37
<i>orfB17</i>	68290-68910c*	SEQ ID No.38
<i>orfB18</i>	69681-70436	SEQ ID No.39
<i>orfB19</i>	70445-71848	SEQ ID No.40

TABLA 8 (continuación)

Gen	Bases en SEQ ID NO.1	Número de la secuencia de polipéptidos correspondiente
<i>orfB20</i>	71851-72957	SEQ ID No.41
<i>orfB21</i>	73037-73942	SEQ ID No.42
<i>orfB22</i>	73995-74534c	SEQ ID No.43
[Nota 1: c indica que el gen está codificado por la cadena de DNA del complemento; Nota 2: para cada marco de lectura abierto dado arriba, se describe el marco de lectura abierto probable más largo. A veces ocurre que puede identificarse más de un codón de iniciación candidato potencial. Un experto en la técnica reconocerá esto y será capaz de identificar codones de iniciación posibles alternativos. Se han indicado aquellos genes que tienen más de un codón de iniciación posible con un símbolo "c". A todo lo largo de este documento se ha indicado el que los autores creen es el codón de iniciación; sin embargo, una persona con experiencia en la técnica apreciará que puede ser posible generar proteína activa utilizando un codón de iniciación alternativo, considerándose también dentro del alcance de la presente descripción las proteínas generadas con utilización de estos codones de iniciación alternativos. Nota 3: la SEQ ID NO: para <i>orfB13b</i> se designó originalmente como SEQ ID NO: 34, pero para claridad se ha asignado una secuencia y un SEQ ID NO separado.]		

Funciones potenciales de los polipéptidos predichos (SEQ ID NO: 7 a 43) se obtuvieron de la base de datos NCBI utilizando una búsqueda BLAST. Los apareamientos óptimos obtenidos de estas búsquedas se describen a continuación en la Tabla 9:

(Tabla pasa a página siguiente)

TABLA 9

Gen	Apareamientos de proteína significativos	Registro	Acceso a GenBank	Función propuesta
<i>orfB1</i>	Proteína hipotética, ausencia de aciertos de longitud total, alta preferencia de codones GC			Desconocida
<i>orfB2</i>	SCM2.07, proteína hipotética (<i>S. coelicolor</i>)	998	NP_625154	Desconocida
<i>orfB3</i>	SCF76.07, proteína hipotética, (<i>S. coelicolor</i>)	359	NP_62786	Desconocida
<i>orfB4</i>	SCF76.06, regulador de transcripción de la familia <i>araC</i> (<i>S. coelicolor</i>)	412	NP_624785	Desconocida
<i>orfB5</i>	SCF76.05c, cloroperoxidasa distinta de hemo (<i>S. coelicolor</i>)	495	NP_624784	Cloroperoxidasa distinta de hemo
<i>orfB6</i>	SCF76.09, proteína hipotética (<i>S. coelicolor</i>)	159	NP_624788	Desconocida
<i>orfB7</i>	SCF76.08c, proteína hipotética (<i>S. coelicolor</i>)	473	NP_624787	Desconocida
<i>borB</i>	PteH, tioesterasa de polieno macrólido (<i>S. avermitilis</i>)	244	BAB69315	Tioesterasa tipo II
<i>borC</i>	XF1726, 2,5-dicloro-2,5-ciclohexadieno-1,4-diol-deshidrogenasa (<i>Xylella fastidiosa</i> , cepa 9a5c)e	160	NP_299015	Deshidrogenasa
<i>borD</i>	FabG, precursor de 3-oxoacil-ACP-reductasa (<i>Plasmodium falciparum</i>)	124	AAK83686	3-oxoacil-ACP-reductasa

ES 2 287 552 T3

TABLA 9 (continuación)

Gen	Apareamientos de proteína significativos	Registro	Acceso a GenBank	Función propuesta
<i>borE</i>	FN1586, O-succinilbenzoil-CoA-sintasa (<i>Fusobacterium nucleatum</i> subesp. <i>nucleatum</i> ATCC 25586)	88	NP_602402	Ciclasa (miembro de la superfamilia enolasas)
<i>borF</i>	Homólogo supuesto de lisofosfolipasa (<i>Arabidopsis thaliana</i>)	57	NP_565066	Desconocida
<i>borG</i>	MTH1444, acetolactato-sintasa, subunidad grande (<i>Methanothermobacter thermautotrophicus</i>)	120	NP_276558	Desconocida
<i>borH</i>	PA3592, proteína hipotética conservada (<i>Pseudomonas aeruginosa</i>)	116	NP_252282	Desconocida
<i>borI</i>	TyH1, citocromo P450 (<i>Streptomyces fradiae</i>)	285	AAD12167	Citocromo P450-oxidasa
<i>borJ</i>	BioA, DAPA-aminotransferasa (<i>Kurthia</i> sp. 538-KA26)	346	BAB39453	Amino-transferasa
<i>borK</i>	Adh1, alcohol-deshidrogenasa (<i>Aquifex aeolicus</i>)	191	NP_213938	NAD/quinona-oxidoreductasa
<i>borL</i>	Proteína GH3 supuesta regulada por auxina (<i>Arabidopsis thaliana</i>)	92	NP_176159	Desconocida
<i>borM</i>	SCL6.10, proteína hipotética similar a la deshidrogenasa dependiente de F420 supuesta (<i>S. coelicolor</i>)	108	CAB76875	Deshidrogenasa dependiente de F420
<i>borN</i>	SC1C2.27, proteína hipotética, superfamilia 2-hidroxihepta-2,4-dieno-1,7-dioato-isomerasa (<i>S. coelicolor</i>)	215	NP_629680	2-hidroxihepta-2,4-dieno-1,7-dioato-isomerasa
<i>borO</i>	ThrS, treonil-tRNA-sintetasa	627	NP_301410	Treonil-tRNA-

TABLA 9 (continuación)

	<i>(Mycobacterium leprae)</i>			sintetasa, gen de au- torresistencia
<i>orfB8</i>	Proteína hipotética conser- vada (<i>Methanosarcina ace- tivorans</i> str. C2A). (Pfam expulsa la familia MarR débil)	37	NP_617908	Posible regulador
<i>orfB9</i>	Antagonista supuesto anti- factor sigma (<i>Streptomyces coelicolor</i>)	113	NP_631789	Antagonista anti-factor sigma
<i>orfB10</i>	Proteína hipotética conser- vada (<i>S. coelicolor</i>)	95	NP_631790	Desconocida
<i>orfB11</i>	Proteína hipotética, sin aciertos de longitud total, alta preferencia de codones GC			Desconocida

TABLA 9 (continuación)

Gen	Apareamientos de proteína significativos	Registro	Acceso a GenBank	Función propuesta
<i>orfB12</i>	Regulador supuesto (<i>S. coelicolor</i>)	92	NP_631494	Regulador (de un sistema de dos componentes, puede ser sensor de membrana)
<i>orfB13a</i>	acetiltransferasa supuesta (<i>S. coelicolor</i>);	58	NP_625155	asignación provisional de acetiltransferasa en dos marcos, o error de secuenciación y debería encontrarse en un solo marco
<i>orfB13b</i>	acetiltransferasa supuesta (<i>S. coelicolor</i>)	100	NP_625155	
<i>orfB14</i>	Lipoproteína supuesta (<i>S. coelicolor</i>)	386	NP_631245	Desconocida
<i>orfB15</i>	Proteína hipotética (<i>S. coelicolor</i>)	41	NP_631424	Desconocida
<i>orfB16</i>	Formiato-deshidrogenasa supuesta (<i>S. coelicolor</i>) (Pfam se aparea con molibdopterina-oxidoreductasa-/formiato-deshidrogenasa, subunidad alfa)	915	NP_626265	Oxidoreductasa
<i>orfB17</i>	Proteína hipotética conservada, <i>S. coelicolor</i> SCBAC25F8.16	175	NP_631569	desconocida
<i>orfB18</i>	Producto desconocido (<i>Streptomyces aureofaciens</i>)	396	AAD23399	Desconocida
<i>orfB19</i>	Aldehído-deshidrogenasa supuesta (<i>S. aureofaciens</i>)	635	AAD23400	Aldehído-deshidrogenasa
<i>orfB20</i>	Alcohol-deshidrogenasa supuesta (<i>S. coelicolor</i>)	450	NP_630527	Alcohol-deshidrogenasa
<i>orfB21</i>	Proteína hipotética (<i>S. coelicolor</i>)	395	NP_630528	Desconocida
<i>orfB22</i>	Proteína de fijación de calcio supuesta (<i>S. Coelicolor</i>)	160	NP_631687	Proteína de fijación de calcio

El análisis de las funciones de los productos génicos supuestos indica que los genes *borB* a *borO* forman muy probablemente los límites de la agrupación biosintética de borrelidina. La evidencia para soportar esto llegó de la desorganización de *borB2* que produjo borrelidinaa niveles indistinguibles de la cepa parental de tipo salvaje. Adicionalmente, *borB3* a *borB7* tienen homólogos en el genoma de *Streptomyces coelicolor* A3(2) codificados en el cósmido SCF76; están presentes los mismos orfs, pero en un orden diferente. Los orfs *borB8* a *borB10* están dispuestos idénticamente a los homólogos en el cósmido SC5E3 de *S. coelicolor* A3(2). Los orfs *borB18* a *borB21* tienen homólogos

que están dispuestos análogamente en el cósmido SC1A2 de *S. coelicolor* A3(2). El orf *borB13* contiene un desplazamiento de marco y por tanto cualquier producto génico sería muy probablemente inactivo. Además, no puede deducirse fácilmente función alguna para los productos de estos orfs durante la biosíntesis de la borrelidina.

5 Genes de biosíntesis de la unidad iniciadora

Con objeto de identificar los genes que están implicados en la biosíntesis de la unidad iniciadora *trans*-ciclopentano-1,2-dicarboxílico, cada uno de los genes *borB* a *borN* se desorganizó (v.g. véanse los Ejemplos 8-11). Esto se hizo de una manera diseñada para minimizar la posibilidad de efectos polares, lo que se verificó por complementación satisfactoria en *trans* con una copia de longitud total del gen desorganizado bajo el control del promotor ermE*, lo que

proporcionó aproximadamente niveles de tipo salvaje (WT) de producción de borrelidina en cada caso. Cada uno de los mutantes desorganizados se dejó crecer por triplicado como se describe en el Ejemplo 1, y se evaluó la producción de borrelidina. Junto con estos, cada mutante se dejó crecer por triplicado y se complementó, después de 24 horas, con el ácido iniciador exógeno hasta una concentración final de 1 mM, y se evaluó la producción de borrelidina. La extracción y el análisis para borrelidina proporcionaron los datos que se describen a continuación en la Tabla 10:

TABLA 10

Gen biosintético de borrelidina desorganizado	Producción de borrelidina sin alimentación (% con relación a WT)	Producción de borrelidina con alimentación (% con relación a WT no alimentado)
Tipo salvaje (control)	100 ± 16, (100 ± 2)	363 ± 65, (269 ± 49)
<i>borB</i>	75 ± 11, (43 ± 20)	172 ± 51
<i>borC</i>	0, (10 ± 3)	933 ± 42
<i>borD</i>	7 ± 1, (0)	75 ± 15
<i>borE</i>	2 ± 1	122 ± 23
<i>borF</i>	3 ± 2	201 ± 52
<i>borG</i>	11 ± 1, (32 ± 3)	1532 ± 142
<i>borH</i>	17 ± 2, (23 ± 13)	203 ± 40
<i>borI</i>	0, (0)	0, (0)
<i>borJ</i>	0, (0)	0, (0)
<i>borK</i>	0, (6 ± 1)	319 ± 54, (464 ± 18)
<i>borL</i>	0, (0)	408 ± 70, (399 ± 69)
<i>borM</i>	0, (6 ± 3)	461 ± 29, (553 ± 66)
<i>borN</i>	25 ± 9, (34 ± 3)	68 ± 12, (46 ± 9)
<i>borO</i>	N/A	N/A

[Nota 1 : Los valores dados entre paréntesis indican en qué casos se realizaron ejecuciones repetidas de algunos experimentos; Nota 2: N/A = no aplicable.]

Basándose en los datos de la Tabla 10, está claro para un experto en la técnica que los productos génicos BorC-F y K-M son esenciales o muy importantes para la biosíntesis del ácido *trans*-ciclopentano-1,2-dicarboxílico, dado que estos mutantes no producían cantidad alguna o niveles muy bajos de borrelidina sin la adición del ácido iniciador exógeno, por lo cual aquéllos producían borrelidina a niveles que se aproximaban a, o eran mejores que, el del organismo WT. Adicionalmente, los productos génicos BorG, H y N parecen estar implicados en, pero sin ser esenciales para, la biosíntesis de la unidad iniciadora, dado que los mismos producían niveles significativamente menores de borrelidina a no ser que se añadiera ácido iniciador exógeno, por lo que aquellos producían borrelidina a niveles que se aproximaron o eran mejores que el del organismo WT; esto era particularmente notable en el caso del mutante *borG*.

La función metabólica normal de los homólogos de *BorN* es la producción de 2-oxo-hepta-3-eno-1,7-dioato 10, un paso fundamental en el catabolismo de la tirosina por la vía del ácido 4-hidroxifenil-acético 9 (Figura 9) (Prieto *et al.*, 1996). Por consiguiente, 10 puede ser un compuesto intermedio en el camino biosintético para *trans*-ciclopentano-1,2-dicarboxílico. La capacidad del mutante desorganizado en *borN* para producir borrelidina, aunque a un nivel reducido,

está basada muy probablemente en la presencia de un homólogo en cualquier otra parte del genoma utilizado en el catabolismo de la tirosina durante el metabolismo primario.

El compuesto intermedio 10 contiene toda la funcionalidad requerida para la formación eventual de ácido *trans*-ciclopentano-1,2-dicarboxílico. El paso más probable siguiente de la biosíntesis es la reducción de la posición 3-eno en una reacción similar a la catalizada por una enoil-reductasa. Enzimas potenciales responsables de este paso son BorC, BorD, BorK o BorM; estas enzimas están involucradas todas ellas en la biosíntesis de unidades iniciadoras de borrelidina como se ve por los datos de la Tabla 10. El 2-oxohepta-1,7-dioato resultante 11 es un posible sustrato para ciclación por formación de un enlace nuevo C-C entre C6 y C2. Otro posible sustrato para esta ciclación podría ser 2-hidroxihepta-1,7-dioato 12 o alguna forma activada del mismo. Ésta podría formarse presumiblemente a partir de 11 por la acción de una oxidorreductasa tal como BorC, BorD o BorM.

El paso fundamental de la ciclación está catalizado muy probablemente por BorE, que exhibe semejanza con O-succinilbenzoil-CoA-sintasa y la cloromuconato-cicloiso-merasa. Estas enzimas pertenecen a la super-familia de las enolasas, cuyos miembros comparten la capacidad común para estabilizar la formación de un anión en el átomo de carbono adyacente a un grupo carboxilato (Schmidt *et al.*, 2001). Adicionalmente, es notable que el sustrato para la muconato-cicloisomerasa es un hexa-1,6-dioato, que es similar en estructura global a 11 y 12. La abstracción de un protón y formación de un anión equivalente en C6 de 11 ó 12 (o una forma activada del mismo, v.g. 13) con ciclación subsiguiente a C2 proporciona la estructura del anillo ciclopentano correctamente sustituido, aunque la intermediación de 11 como sustrato requeriría algún procesamiento adicional del ciclopentano sustituido, muy probablemente por eliminación de agua para dar el ácido ciclopent-1-eno-1,2-dicarboxílico simétrico, o posiblemente el compuesto Δ^1 -insaturado, ácido ciclopent-1-eno-1,2-dicarboxílico. Sin embargo, la alimentación de ácido ciclopent-1-eno-1,2-dicarboxílico, o ésteres etílicos del mismo, a cepas de *S. parvulus* Tü4055 desorganizadas en cualquiera de *borC-E*, o a cepas WT, no producía cantidad alguna de borrelidina, o no producía borrelidina en cantidad incrementada en comparación con los controles no alimentados. Estos datos indican que este compuesto no es probablemente un compuesto intermedio en la biosíntesis de la unidad iniciadora, y que el sustrato de BorE es posiblemente el 2-hidroxihepta-1,7-dioato 12, o una forma activada del mismo (v.g. 13). Un camino supuesto para el camino biosintético al ácido *trans*-ciclopentano-1,2-dicarboxílico se muestra en la Figura 4.

Los genes combinados específicos requeridos para los pasos biosintéticos hacia el ácido *trans*-ciclopentano-1,2-dicarboxílico no están claros, pero probablemente están codificados por alguna combinación de *borC-H*, *borK*, *borM* y *borN*. La carencia de ciertos homólogos de genes que estén involucrados en el catabolismo del ácido 4-hidroxifenil-acético 9, y que podrían actuar antes de *BorN* en el camino, es muy probablemente una indicación de que genes metabólicos primarios realizan estas tareas. La adición de ácido *trans*-ciclopentano-1,2-dicarboxílico exógeno a *S. parvulus* Tü4055 y cepas afines aumenta el título de borrelidina del orden de 2 a 3 veces en las condiciones empleadas por los autores de la invención, lo que indica que la biosíntesis del ácido iniciador es un factor limitante en la biosíntesis de la borrelidina. Estos datos son consistentes con la degradación metabólica primaria de la tirosina, que es la fuente del ácido *trans*-ciclopentano-1,2-dicarboxílico.

En un intento para esclarecer adicionalmente cuáles son los genes que pueden ser específicamente responsables de la biosíntesis de la unidad iniciadora, se realizaron varios experimentos de co-cultivo con combinaciones de los mutantes diferentes - éstos requieren el conocimiento de que los productos génicos de *borI* y *borJ* están involucrados específicamente en la formación del resto nitrilo en C12, lo cual está clarificado por los datos proporcionados más adelante en la sección siguiente en combinación con los datos de la Tabla 10. En suma, el co-cultivo de mutantes *borE⁻* & *borD⁻*, y de *borE⁻* & *borM⁻* no logró producir cantidad alguna de borrelidina, mientras que el co-cultivo de mutantes *borM⁻* & *borI⁻*, y *borM⁻* & *borK⁻* produjo borrelidina a niveles aproximadamente iguales a los de tipo salvaje. Estos datos, en combinación con el de la Tabla 10, y siguientes, indican claramente que *borD*, *borE* y *borM* están involucrados en la biosíntesis de la unidad iniciadora, mientras que *borI* y posiblemente *borK*, están implicados en la formación del resto nitrilo en C12 de la borrelidina.

Está claro por los datos de la Tabla 10 que la adición exógena de ácido *trans*-ciclopentano-1,2-dicarboxílico es suficiente para restablecer aproximadamente los niveles WT, o mejores, de producción de borrelidina en mutantes donde los genes que están involucrados en la biosíntesis de la unidad iniciadora han sido desorganizados. Estos datos indican que no existe problema alguno en la absorción activa de ácido carboxílico añadido por *S. parvulus* Tü4055, y que está presente una actividad que es capaz de convertir el ácido carboxílico en un CoA-tioéster equivalente. Así pues, dadas las tecnologías de mutagénesis conocidas, es obvio para un experto en la técnica que la adición de ácidos carboxílicos exógenos a uno de los mutantes arriba mencionados, por ejemplo la cepa *borE⁻* de *S. parvulus* Tü4055/*borE:aac3(IV)* descrita en el Ejemplo 8, puede conducir a la producción de análogos de borrelidina en los cuales el resto ácido carboxílico de la unidad iniciadora está reemplazado con un resto derivado del ácido carboxílico añadido exógenamente.

Para examinar esta posibilidad, se alimentó la cepa *S. parvulus* Tü4055/*borE:aac3(IV)* con un ácido *trans*-ciclobutano-1,2-dicarboxílico de acuerdo con el protocolo descrito en el Ejemplo 1 y se analizó luego como se describe en el Ejemplo 4. La estructura 18, descrita en la Figura 10, muestra la nueva estructura de borrelidina obtenida por alimentación de este ácido carboxílico; este compuesto 18 exhibía el cromóforo UV anticipado para borrelidina pero se eluía con un tiempo de retención más breve y exhibía la masa esperada por LCMS ($m/z = 474,3$ [M-H]⁻). La verificación de esta metodología fue proporcionada por la producción, aislamiento y caracterización de 18 (Ejemplo 16). Está claro para un experto en la técnica que podrían utilizarse también otros ácidos carboxílicos en experimentos

de alimentación similares para proporcionar adicionalmente nuevos análogos de borrelidina. Aunque es posible que no todos los ácidos carboxílicos pudieran incorporarse utilizando la metodología exacta descrita en esta memoria, una persona experta en la técnica es conocedora de numerosos métodos disponibles para mejorar la incorporación de las unidades iniciadoras alimentadas.

Además del uso de la cepa delecionada en *borE*, se observó (véase la Tabla 10) que la cepa *S. parvulus* Tü4055/*borG::aac3(IV)*, en la cual *borG* ha sido desorganizado, cuando se alimentó con la unidad iniciadora natural de la PKS de bor, ácido *trans*-ciclopentano-1,2-dicarboxílico, producía borrelidina con títulos significativamente mayores que los observados cuando se alimentó (aumento de 4 veces) o no se alimentó (aumento de 15 veces) el organismo de tipo salvaje. Para examinar esto ulteriormente, se repitió dicho experimento utilizando a la vez el ácido iniciador natural y un ácido iniciador no natural como sustratos exógenos, alimentados, en paralelo, al tipo salvaje, el mutante *borE* y el mutante *borG*. Los datos resultantes se describen en la Tabla 11.

TABLA 11

	No alimen- tado	Alimentado con ácido ciclopentano <i>trans</i> -1,2-dicarboxílico 1 mM	Alimentado con ácido ciclobutano <i>trans</i> -1,2-dicarboxílico 1 mM
<i>S. parvulus</i> Tü4055	2.3 mg/l	6.6 mg/l	-
<i>S. parvulus</i> Tü4055/ <i>borE::aac3(IV)</i>	0	4.7 mg/l	2.2 mg/l
<i>S. parvulus</i> Tü4055/ <i>borG::aac3(IV)</i>	0	88.9 mg/l	43.0 mg/l

Como puede verse por la Tabla 11, la utilización de *S. parvulus* Tü4055/*borG::aac3(IV)* en lugar de *S. parvulus* Tü4055/*borE::aac3(IV)* para la mutasíntesis aumenta el título aproximadamente 19 veces, y que *S. parvulus* Tü4055/*borG::aac3(IV)* alimentada con el ácido iniciador natural produce 38 veces más borrelidina A que el tipo salvaje por sí solo, o 13 veces más borrelidina A que la cepa de tipo salvaje alimentada con la misma cantidad de ácido ciclopentano-*trans*-1,2-dicarboxílico. Estos datos indican claramente que el uso de la cepa *S. parvulus* Tü4055/*borG::aac3(IV)* para experimentos de mutasíntesis es beneficioso para la producción de títulos mejorados de análogos de borrelidina. Este método tiene aplicabilidad general para la producción tanto de borrelidina como de análogos de borrelidina.

Sobre la base de este descubrimiento, se repitieron los experimentos de alimentación con ácidos carboxílicos alternativos en *S. parvulus* Tü4055/*borG::aac3(IV)*, y se extendieron para incluir ácido 2,3-dimetil-succínico y ácido 2-metilsuccínico; los nuevos compuestos derivados de la incorporación de estas unidades iniciadoras alternativas, 19 y 20 respectivamente, se describen en la Figura 10.

Formación del resto nitrilo en C12

El análisis de la secuencia del dominio AT del módulo 3 de la PKS de borrelidina indica que el sustrato utilizado para la tercera tanda de extensión de la cadena es metilmalonil-CoA. Así pues, el átomo de carbono del resto nitrilo procede muy probablemente del grupo metilo de metilmalonil-CoA. Esto se comprobó por experimentos de alimentación con isótopos estables. La alimentación de [2,3-¹³C₂]-propionato de sodio a *S. parvulus* Tü113 dio borrelidina que exhibía una marcación intacta de los carbonos en C4-C24, C6-C25, C8-C26, C10-C27 y C12-C28, e incorporaciones específicas idénticas (como se determina dentro de los límites de los métodos experimentales empleados por los inventores), como era de esperar (Figura 2). Estos datos indican que la conversión del grupo metilo C12 ocurre sea durante el ensamblaje de la cadena en el momento de o después de la incorporación de la tercera unidad de extensión, o que la misma ocurre después del ensamblaje de la cadena del poliquétido y liberación de la PKS. Basándose en asignaciones funcionales dadas a los genes biosintéticos de borrelidina, en asociación con los datos de desorganización de genes expuestos en la Tabla 10, tanto *borI* como *borJ* están involucradas claramente en la formación del resto nitrilo en C12, si bien pueden estarlo también otros tales como *borK*.

La hidroxilasa BorI del citocromo P450 comparte la máxima semejanza con TylHI, que cataliza la hidroxilación de un grupo metilo exocíclico de la tilosin-macrolactona antes de la adición de un resto desoxihexosa (Fouces *et al.*, 1999). Por esta razón, se cree que BorI cataliza la oxidación del grupo metilo C12 durante la biosíntesis de la borrelidina. De acuerdo con esto, el mutante *borI⁻* de *S. parvulus* Tü4055/*borI::aac3(IV)* no produce cantidad alguna de borrelidina pero acumula un nuevo producto 14 (Figura 6), que es menos polar que la borrelidina. 14 se transforma fácilmente en borrelidina cuando se alimenta al mutante *borE⁻* *S. parvulus* Tü4055/*borE::aac3(IV)* que carece de la capacidad para sintetizar la unidad iniciadora de la PKS pero mantiene el resto de los genes biosintéticos de borrelidina intactos. La fermentación de *S. parvulus* Tü4055/*borI::aac3(IV)* seguida por extracción y aislamiento proporcionó ~30 mg de 14 (Ejemplo 14). El análisis estructural completo de 14 identificó el mismo como 12-desnitrilo-12-metilborrelidina (pre-borrelidina). Esto es consistente con la función propuesta de BorI en la biosíntesis de la borrelidina y proporciona una ruta para nuevos análogos de borrelidina con un grupo metilo unido a C12 del anillo de macrolactona.

Se cree que la aminotransferasa BorJ supuesta dependiente de PLP cataliza la introducción de un átomo de nitrógeno en la borrelidina en la posición activada C28, probablemente por la vía de un resto formilo en C12. De acuerdo con esto, el mutante *borJ*⁻ de *S. parvulus* Tü4055/*borJ::aac3(IV)* no produce borrelidina y acumula un nuevo compuesto que es más polar que la borrelidina. Este nuevo compuesto no se transforma en borrelidina cuando se alimenta al mutante *S. parvulus* Tü4055/*borE::aac3(IV)*, lo que indica que es probablemente un metabolito de derivación más bien que un compuesto intermedio en la biosíntesis de la borrelidina. La fermentación de *S. parvulus* Tü4055/*borJ::aac3(IV)* permitió el aislamiento de 17 mg del compuesto acumulado (Ejemplo 15). El análisis estructural detallado identificó el acumulante como 12-desnitrilo-12-carboxil-borrelidina 2.

Además de los compuestos aislados por la mutación de los genes biosintéticos de borrelidina, 12-desnitrilo-12-formil-borrelidina 15 se aísla a partir del sobrenadante de fermentación de *S. parvulus* Tü113. Los medios de fermentación y las condiciones utilizadas para estos experimentos difieren de los que han sido descritos hasta ahora por los autores de la invención en esta memoria, pero están diseñados para maximizar la producción de borrelidina. Se propone que este medio alterado, en combinación con una disminución en la concentración de oxígeno disuelto que se observa ocurre durante esta fermentación específica, promovía la acumulación de 15. 15 se transforma fácilmente en borrelidina cuando se alimenta al mutante *S. parvulus* Tü4055/*borE::aac3(IV)* que carece de la capacidad para sintetizar la unidad iniciadora de PKS pero mantiene el resto de los genes biosintéticos de borrelidina intactos.

Los datos anteriores han conducido a los autores de la invención a proponer una ruta de biosíntesis para el resto nitrilo de borrelidina como se presente en la Figura 6. El carbono del metilo C12 de la pre-borrelidina 14 se oxida primeramente a BorI para introducir un grupo hidroxilo alílico en C28 (16). Este grupo hidroxilo se convierte luego en el resto formilo unido a C12 (15) utilizando un método seleccionado del grupo que comprende: oxidación espontánea (con inclusión de oxidación derivada por alguna enzima de fondo) la acción de un gen específico de la agrupación de genes biosintéticos de borrelidina; productos génicos candidato son por tanto BorI propiamente dicho, que actúa de una manera multifuncional y opera por la formación de una estructura de gem-diol en C12 seguido por deshidratación; o alternativamente, por la vía de uno de los genes codificantes de oxidorreductasa tales como *borC* o *borK*. Se anticipa que el paso siguiente es la transaminación catalizada por BorJ de 15 a fin de introducir un átomo de nitrógeno en C28, en forma de una amina, por un proceso mediado por piridoxamina-fosfato. El producto amínico 17 supuesto sufre luego oxidación, posiblemente de modo espontáneo, pero muy probablemente por una actividad enzimática tal como BorI (pueden extraerse ciertos paralelismos con la biosíntesis de nitrilos en las plantas (Celenza, 2001; Hahn *et al.*, 1999; Nielson y Møller, 1999)) o por los productos de uno de los genes codificantes de oxidorreductasas, v.g. *borC* o *borK*, o por una oxidorreductasa general dentro del proteoma.

A fin de examinar el camino propuesto con mayor detalle, se realizaron cierto número de experimentos de biotransformación utilizando pre-borrelidina 14 como sustrato para investigar la acción de *borI-K* individualmente y en combinación, utilizando pEM4 como vector y *S. albus* J1074 (Chater & Wilde, 1980) como cepa de expresión. La expresión de *borI* o *borJ* individualmente no aportaba producción de borrelidina por adición de 14. La 14 añadida se consumía únicamente durante la biotransformación con *borI* (y no en ninguno de los experimentos de control); el compuesto 14 añadido se identificó al convertirse en el metabolito de derivación 2. Sin embargo, la co-expresión de *borI* & *borJ* convertía de hecho el compuesto 14 añadido en borrelidina. Así pues, parece ser que o bien BorI o las actividades generales del proteoma en *S. albus* son capaces de oxidar el compuesto intermedio amínico propuesto 17 en el camino biosintético de la borrelidina. Además de la alimentación de pre-borrelidina 14, se alimentó también 12-desnitrilo-12-carboxil-borrelidina 2 a las tres cepas arriba descritas. No se observó conversión alguna de 2 en borrelidina en ninguno de estos experimentos, lo que refuerza la idea de que 2 es un metabolito de derivación.

La investigación detallada del DNA genómico de 3 cepas productoras de borrelidina, *S. rochei* ATCC23956, *S. parvulus* Tü113 y *S. parvulus* Tü4055, utilizando numerosos materiales digeridos por restricción y análisis subsiguiente por Transferencia Southern, indica que los agrupaciones de genes biosintéticos de borrelidina de estos tres organismos se conservan muy estrechamente. Por consiguiente, parece ser que los caminos biosintéticos de borrelidina de estas cepas son muy similares. Esta suposición permite a los autores de la invención considerar los datos anteriores, que se obtienen de cepas diferentes, como aplicables a un solo camino de biosíntesis.

Está claro para un experto en la técnica que la manipulación de los genes implicados en la formación del resto nitrilo en C12 de borrelidina, por ejemplo *borI*, o *borJ*, es un método de utilidad general para la producción de nuevas moléculas afines a borrelidina y derivados de borrelidina con funcionalidad alterada en C12. Adicionalmente, la transferencia de estos genes a otros organismos que producen otros productos poliquétidos naturales o modificados por ingeniería genética puede permitir la incorporación de restos nitrilo en tales compuestos.

En una extensión de este trabajo, se realizan por separado desorganizaciones en *borI* y *borJ* en la cepa *S. parvulus* Tü4055/*borG::aac3(IV)* para dar las cepas doblemente mutadas *S. parvulus* Tü4055/*borG::aac3(IV)/borI::hyg* y *S. parvulus* Tü4055/*borG::aac3(IV)/borJ::hyg* (Ejemplos 12 & 13 respectivamente). Estas cepas se alimentan con ácidos carboxílicos alternativos, ácido *trans*-ciclobutano-1,2-dicarboxílico, ácido 2,3-dimetilsuccínico y ácido 2-metilsuccínico (como se ha descrito arriba), y se encuentra que producen los análogos mutasintéticos de borrelidina que llevan, o bien una función metilo (21, 22 y 23 respectivamente) o una función carboxilo en C12 (24, 25 y 26 respectivamente) en lugar del grupo nitrilo, y que se derivan también de unidades iniciadoras alternativas correspondientes a los ácidos carboxílicos suministrados exógenamente. Esta genoteca ortogonal de nuevos compuestos se describe en la Figura 11 y se muestran los cromóforos UV y datos espectrales de masas observados para cada compuesto.

Ejemplos

Métodos generales

- 5 Las enzimas de restricción, otros reactivos de biología molecular, antibióticos y productos químicos se adquirieron de fuentes comerciales estándar. La digestión con endonucleasas de restricción y ligación siguieron métodos estándar (Sambrook, J. *et al.*, 1989).

Ejemplo 1

Fermentación de cepas de *S. parvulus*

- 15 El método que sigue es generalmente útil para cultivar *S. parvulus* para la producción de borrelidina y/o análogos de borrelidina:

- Se inoculó un matraz de siembra que contenía medio NYG (30 ml en un matraz Erlenmeyer de 250 ml) con un stock de trabajo (0,5 ml). El medio NYG contiene, en agua desionizada: extracto de carne de buey (0,3%), peptona Bacto (0,5%), glucosa (1%) y extracto de levadura (0,5%). Después de dos días de agitación mediante sacudidas en una incubadora rotativa (carrera de 2 pulgadas (5 cm); 30°C; 250 rpm) el cultivo crema resultante se utilizó para inocular el medio de producción PYDG (30 ml en un matraz Erlenmeyer de 250 ml; 10% de inóculo). El medio PYDG contiene por litro de agua desionizada: nutriente de leche peptonizada (1,5%), autolizado de levadura (0,15%), dextrina (4,5%) y glucosa (0,5%) ajustados a pH 7,0. Después de 5 días de agitación mediante sacudidas en una incubadora rotativa (carrera de 2 pulgadas (5 cm); 30°C; 250 rpm) se recogió el cultivo para análisis como se describe en el Ejemplo 4, o para propósitos de aislamiento en caso requerido. Estos experimentos se realizaron por triplicado para análisis cuantitativo.

El método siguiente es útil para la alimentación de ácidos carboxílicos exógenos a cepas de *S. parvulus*:

- 30 La cepa *S. parvulus* se desarrolló como se ha descrito arriba. Después de 24 horas de crecimiento en medio de producción PYDG, se añadió el ácido carboxílico de elección como una sola parte alícuota de 50 μ l (solución 0,6 M en metanol al 70% después de neutralización con NaOH 5 N). El cultivo resultante se recogió después de 5 días de fermentación total y se analizó como se describe en el Ejemplo 4. Para estudios cuantitativos, estos experimentos se realizaron por triplicado, y las cepas WT equivalentes alimentadas y sin alimentar sirvieron como controles.

Ejemplo 2

Crioconservación de cepas de *S. parvulus*

Stocks de trabajo

- 45 Se prepararon stocks de trabajo de micelios vegetativos por mezcla de un cultivo de siembra de 2 días desarrollado en medio NGY (0,5 ml) con crioconservante (0,5 ml). El crioconservante está constituido por 20% de glicerol y 10% de lactosa en agua desionizada.

Stocks de esporas

- 50 Se incubaron cepas de *S. parvulus* en placas de agar HA a 30°C. Después de 14 días, las esporas resultantes de una sola placa se cosecharon y se suspendieron en crioconservantes (1 ml). El agar HA contiene, en agua desionizada: 0,4% de extracto de levadura, 1% de extracto de malta, 0,4% de dextrosa y 1,5% de agar ajustado a pH 7,3.

Ejemplo 3

Clonación de la agrupación de genes biosintéticos de borrelidina y desorganización de *borA2* & *borA3*

Generación de una genoteca de cósmidos

- 60 Se construyó una genoteca de cósmidos en el vector cósmido pWE15 utilizando el kit Gigapack® III Gold Packaging Extract de acuerdo con el manual del fabricante (Stratagene). Se extrajo DNA cromosómico de *S. parvulus* Tü4055 de acuerdo con protocolos estándar (Kieser *et al.*, 2000) y se trató con *Sau3AI* antes de clonación en pWE15. Cierta número de los transformantes de *E. coli* resultantes (3300) se seleccionaron y se transfirieron a placas de microtitulación de 96 pocillos que contenían medio de Caldo Luria (LB) (0,1 ml por pocillo) con ampicilina (100 μ g/ml). Los clones resultantes se replicaron en placas de agar Luria (LA) que contenían ampicilina (100 μ g/ml). Después de incubación durante una noche a 37°C, se transfirieron las colonias a filtros de membrana de nailon para análisis de hibridación de colonias *in situ* de acuerdo con protocolos publicados (Sambrook *et al.*, 1989).

Escrutinio de la genoteca

La genoteca de cósmidos se exploró utilizando una sonda que se generó utilizando el kit DIG de marcación y detección de DNA (Roche) de acuerdo con las instrucciones de los fabricantes. La sonda utilizada era un fragmento 5 *BglII-BamHI* (1,7 kpb) obtenido a partir del gen que codifica el módulo 6 de la tercera subunidad de la PKS de la oleandomicina de *Streptomyces antibioticus* (Swan *et al.*, 1994).

Disgregación de la agrupación de genes biosintéticos de borrelidina

10 Los cósmidos que dieron una respuesta positiva cuando se exploraron como se ha descrito arriba se digirieron con *BamHI* y fragmentos de menos de 3 kpb se subclonaron en pOJ260 (Bierman *et al.*, 1992). Se utilizaron luego éstos para transformar protoplastos de *S. parvulus* Tü4055 como se describe en el Ejemplo 5. Los transformantes resultantes se evaluaron luego en cuanto a la capacidad para producir borrelidina. Dos clones no eran productores de borrelidina; se obtuvieron ambos de cosBor32A2 y contenían una secuencia típica de una PKS modular. Los cósmidos restantes se 15 exploraron luego utilizando sondas obtenidas a partir de los dos fragmentos *BamHI*, lo que condujo a la identificación del cósmido solapante cosBor19B9 que contenía el resto de la agrupación biosintético de borrelidina.

Secuenciación de cosBor32A2 y cosBor19B9

20 Los cósmidos cosBor32A2 y cosBor19B9 se transformaron en *E. coli* DH10B y los clones resultantes se dejaron crecer a 37°C en medios 2xTY (30 ml) que contenían ampicilina. Después de 15 horas, las células se recogieron y se utilizaron kits Qiagen Tip 100 para preparar el DNA de cósmido. Aproximadamente 5 µg del DNA de cósmido se digirieron con *Sau3AI* (1 U). Se tomaron muestras con intervalos de 2, 4, 6, 8 & 10 minutos después de la adición de la enzima y se apagaron en un volumen igual de EDTA 0,5 M enfriado con hielo. Las muestras se mezclaron y se 25 analizaron luego por electroforesis en gel, y se recuperaron del gel aquellos fragmentos comprendidos entre 1,5 y 2,0 kpb. Los fragmentos se clonaron en pHSG397 linealizado y desfosforilado (Takeshita *et al.*, 1987), y se transformaron en *E. coli* DH10B. Los clones resultantes que contenían la inserción se dejaron crecer en medio 2xTY (2 ml) que contenía cloranfenicol (30 µg/ml) y se purificaron utilizando kits Wizard (Promega).

30 Se llevó a cabo la secuenciación del DNA utilizando un robot CATALYST Molecular Biology 800 de Applied Biosystems para realizar las reacciones del terminador didesoxi, que se cargaron luego en un secuenciador automático ABI Prism 3700 (Applied Biosystems). Los datos de secuencia brutos se procesaron utilizando el paquete de soporte lógico Staden. El ensamblaje y la revisión de los cóntigos se realizó utilizando GAP (Genome Assembly Program), versión 4.2 (Bonfield *et al.*, 1995). Se utilizó el paquete GCG (Devereux *et al.*, 1984) versión 10.0 para el análisis de 35 secuencia.

Ejemplo 4

40 *Análisis químico de cepas de S. parvulus*

El método siguiente es útil para análisis de las fermentaciones (véase Ejemplo 1) para la producción de borrelidinas naturales y de análogos de borrelidina modificados por ingeniería genética:

45 En un tubo Eppendorf de 2 ml se ajustó una parte alícuota de caldo de fermentación de 5 días (1 ml) a pH-3 por adición de ácido fórmico al 90% (aprox. 20 µl). Se añadió acetato de etilo (1 ml) a la muestra y se mezcló enérgicamente durante 10 min utilizando una bandeja de agitación enérgica. La mezcla se separó por centrifugación en una microcentrífuga y la parte superior se llevó a un tubo Eppendorf limpio de 2 ml. Se eliminó el acetato de etilo por evaporación utilizando un Speed-Vac. Los residuos se disolvieron en metanol (950 µl) y se clarificaron 50 utilizando una microcentrífuga. El análisis se realizó en un sistema de HPLC Agilent HP1100 como se describe a continuación:

Volumen de inyección:	50 µl
55 Fase estacionaria de la columna:	Columna de 150 x 4,6 mm, gel de sílice de fase inversa desactivado con base, tamaño de partícula 3 µm (Hypersil C ₁₈ -BDS).
Fase móvil A:	10% acetonitrilo:90% agua, que contiene acetato de amonio 10 mM y 0,1% de TFA.
60 Fase móvil B:	90% acetonitrilo:10% agua, que contiene acetato de amonio 10 mM y 0,1% de TFA.
Gradiente de la fase móvil:	T=0 min, 25%B; T=15, 100%B; T=19, 100%B; T=19,5, 25%B; T=25, 25%B.
65 Caudal:	1 ml/min.

ES 2 287 552 T3

Detección: UV a 258 nm (adquisición DAD en 190-600 nm); detección MS por ionización mediante pulverización electrónica en el intervalo m/z 100-1000 amu, con mutación en modo iónico +/-ve.

5 Ejemplo 5

Protocolo de transformación de protoplastos para S. parvulus Tü4055

10 Un matraz de siembra que contenía medio de caldo triptona-soja (TSP) (10 ml en un matraz Erlenmeyer de 100 ml) se inoculó con un stock de trabajo (0,15 ml). Después de 3 días de agitación mediante sacudidas en una incubadora rotativa (30°C, 250 rpm), se utilizaron 5 ml del cultivo para inocular medio R5 (Kieser *et al.*, 2000) (50 ml en un matraz Erlenmeyer de 250 ml) que se agitó luego mediante sacudidas en una incubadora rotativa durante 24 horas (30°C, 250 rpm). La transformación de los protoplastos mediada por PEG se realizó luego de acuerdo con protocolos estándar publicados (Kieser *et al.*, 2000).

15 Ejemplo 6

Reemplazamiento de borAT4 con rapAT2 - producción de C10-desmetil-borrelidina

20 El dominio AT4 de PKS de borrelidina se reemplaza con el dominio AT2 de la poliquétido-sintasa de rapamicina como sigue:

Se digiere CosBor32A2 con *EcoRI* y se aísla la banda de 5429 pb. Se utiliza ésta como molde para PCR utilizando los oligos CM410 (5'-AAAATGCATTCGGCCTGAACGG-CCCCGCTGCTA-3') (SEQ ID NO: 44) y CM411 (5'-AAATGGCCAGCGAACACCAACACCACACCACCA-3') (SEQ ID NO: 45). CM410 introduce un sitio de restricción *NsiI* para propósitos de clonación y CM411 introduce un sitio *MscI* para uso en la introducción de un AT heterólogo.

30 El producto de ~1,1 kpb se clona en pUC18 digerido con *SmaI* y desfosforilado. La inserción puede ligarse en dos orientaciones y la orientación inversa se explora por análisis con enzimas de restricción, y se secuencian la inserción. Un plásmido correcto se designa pCJM462. El DNA deficiente en metilación (específicamente dcm⁻) de pCJM462 y pCJR26 (Rowe *et al.*, 1998) se aísla por paso de los plásmidos a través de *E. coli* ET12567. Cada plásmido se digiere luego con *MscI* y *XbaI* y el fragmento de ~7,8 kpb de pCJR26, que contiene el AT2 de rapamicina y secuencias situadas aguas abajo en pCJR26, se liga a la cadena principal de ~3,8 kpb generada por digestión de pCJM462. El plásmido pCJM463 se identifica por análisis de restricción.

40 Se digiere CosBor32A2 con *EcoRI* y *EcoRV* se aísla la banda de 2871 pb. Se utiliza ésta como molde para PCR utilizando los oligos CM412 (5'-AAAGTCCTAGGCGGCGGCCGGC-GGGTCGACCT-3') (SEQ ID NO: 46) y CM413 (5'-TTTAGATCTCGCGACGTCGCACGCGCCGAACGTCA-3') (SEQ ID NO: 47). CM412 introduce un sitio de restricción *AvrII* que une, en marco, la homología de borrelidina aguas abajo al AT heterólogo, y CM413 introduce un sitio *BglII* para propósitos de clonación. El producto de ~1,1 kpb se clona a pUC18 digerido con *SmaI* y desfosforilado. La inserción puede ligar en dos orientaciones y la orientación inversa se explora por análisis con enzimas de restricción y se secuencian la inserción. Un plásmido correcto se designa pCJM464.

45 Los plásmidos pCJM463 y pCJM464 se digieren con *AvrII* y *XbaI* y el fragmento de ~1,1 kpb de pCJM464 se liga a la cadena principal de ~4,7 kpb de pCJM463 para dar pCJM465, que se identifica por análisis con enzimas de restricción. pCJM465 contiene el híbrido de rapamicina AT2 con genes flanqueantes de secuencia de borrelidina que proporcionan homología para integración y recombinación secundaria.

50 El plásmido pCJM465 se digiere con *NsiI* y *BglII* y el fragmento de ~3 kpb se clona en pSL1180 digerido previamente con *NsiI* y *BamHI* para dar pCJM466. El plásmido pCJM466 se digiere luego con *NsiI* y la casete de apramicina se incorpora en un fragmento *PstI* de pEFBA (Lozano *et al.*, 2000) para dar el vector de reemplazamiento pCJM467. Se introduce pCJM467 en *S. parvulus* Tü4055 por transformación de protoplastos como se describe en el Ejemplo 5. Las colonias resistentes a apramicina (25 µg/ml) se identifican inicialmente, y se someten luego a pasos varias veces a través de medios MA sin selección de antibióticos a fin de promover la segunda recombinación (Fernández *et al.*, 1998). Varias colonias sensibles a apramicina se aíslan y se analizan por PCR y transferencia Southern. El nuevo mutante se designa *S. parvulus* Tü4055/467.

60 *S. parvulus* Tü4055/467 se analiza como se describe en el Ejemplo 1 y se demuestra que produce una mezcla de compuesto con el espectro UV correcto. Uno de los nuevos componentes principales que es más polar que la borrelidina tiene el tiempo de retención correcto para 10-desmetil-borrelidina 3. El análisis LCMS indica una relación m/z para un compuesto que tiene 14 unidades de masa menos que la borrelidina como era de esperar, y con un patrón de fragmentación de masas apropiado. Se produce también borrelidina propiamente dicha, pero a niveles menores que el organismo WT.

65

Ejemplo 7

Mutación del motivo selectivo metilmalonil-CoA de *borAT4* para generar 10-desmetil-borrelidina

5 La mutagénesis orientada de los dominios acil-transferasa puede utilizarse también para alterar la especificidad de un AT. En este ejemplo, la especificidad de *borAT4* está dirigida desde metil-malonil-CoA hacia malonil-CoA. Ha sido identificado un motivo de aminoácido (Reeves *et al.*, 2001; WO 02/14482) que dirige la especificidad de un AT. El motivo YASH, como se observa en *bor-AT4*, se encuentra en ATs específicos de metilmalonil-CoA y, en este ejemplo, está alterado a HAFH que se encuentra en ATs específicos de malonil-CoA.

10 Se digiere CosBor32A2 con *NcoI* y se aísla la banda de 5167 pb. Se utiliza ésta como molde para PCR utilizando los iniciadores CM414 (5'-AAACTGCAGAGTCGAACATCGGTCACAC-GCAGGC-3') (SEQ ID NO: 48) y CM415 (5'-AAAATGCATGATCCACA-TCGATACGACGCGCCCA-3') (SEQ ID NO: 49). CM414 introduce un sitio de restricción *PstI* para propósitos de clonación, y CM415 es un iniciador mutagénico que cubre la región codificante de motivos del AT que efectuará los cambios de aminoácidos y contiene un sitio *NsiI* para propósitos de clonación. El fragmento de ~1,1 kpb se clona en pUC18 digerido con *SmaI* y desfosforilado. La inserción puede ligarse en cualquier orientación y la alimentación directa se explora por análisis con enzimas de restricción, y se secuencian la inserción. Un plásmido correcto se designa pCJM468.

20 Se realiza una segunda reacción PCR utilizando el fragmento de 5167 pb *NcoI* de CosBor32A2 y los iniciadores CM416 (5'-TAAATGCATTCCATTCGGTGCAGGTGGAGTTGATCC-3') (SEQ ID NO: 50) y CM417 (5'-ATAGGATCCCCTCCGGGTGCTCCAGACCG-GCCACCC-3') (SEQ ID NO: 51). CM416 introduce un sitio de restricción *NsiI* y es también un iniciador mutagénico que cubre la región codificante del motivo del AT, y CM417 introduce un sitio *BamHI* para propósitos de clonación. El fragmento de ~1,1 kpb se clona en pUC18 digerido previamente con *SmaI* y desfosforilado. La inserción puede ligarse en dos orientaciones, y la orientación directa se explora por análisis con enzimas de restricción, después de lo cual se secuencian la inserción. Un plásmido correcto se designa pCJM469.

30 Los plásmidos pCJM468 y pCJM469 se digieren con *NsiI* y *XbaI*, y el fragmento de ~1,1 kpb de pCJM468 se liga a la cadena principal de ~3,8 kpb de pCJM469 para dar pCJM470, que se identifica por análisis con enzimas de restricción. pCJM470 contiene el motivo mutado de *borAT4* con ~1,1 kpb de DNA homólogo a ambos lados, que proporcionan homología para integración y recombinación secundaria.

35 El plásmido pCJM470 se digiere con *PstI* y *BamHI* y el fragmento de ~2,2 kpb se clona en pSL1180 (Amersham Biosciences) digerido previamente con *PstI* y *BamHI* para dar pCJM471. El plásmido pCJM471 se digiere luego con *PstI* y la casete de apramicina se incorpora en un fragmento *PstI* de pEFBA (Lozano *et al.*, 2000) para proporcionar el vector de reemplazamiento pCJM472.

40 El vector de reemplazamiento pCJM472 se introduce en *S. parvulus* Tü4055 por transformación de protoplastos como se describe en el Ejemplo 5. Las colonias resistentes a apramicina se identifican inicialmente y se someten luego a pasos varias veces a través de medio MA sin selección de antibióticos a fin de promover la segunda recombinación (Fernández *et al.*, 1998). Se aíslan varias colonias sensibles a la apramicina y se analizan por PCR y transferencia Southern, seleccionándose una que contiene la nueva secuencia AT4 que contiene el motivo mutado y el sitio *NsiI*. El nuevo mutante se denomina *S. parvulus* Tü4055/472.

45 *S. parvulus* Tü4055/472 se cultiva y se analiza como se describe en el Ejemplo 1, demostrándose que produce una mezcla de compuestos con el perfil UV correcto para borrelidina. Uno de los nuevos componentes principales, que es más polar que la borrelidina, tiene el tiempo de retención correcto para 3 auténtica. El análisis LCMS indica una relación *m/z* para un compuesto que tiene 14 unidades másicas menos que la borrelidina como era de esperar, y con un patrón de fragmentación de masa apropiado. Se produce también borrelidina propiamente dicha, pero a niveles menores que el organismo de tipo salvaje (WT).

Ejemplo 8

Disgregación de *borE* (*S. parvulus* Tü4055/*borE::aac3(IV)*)

55 A fin de desorganizar *borE*, se amplificó un fragmento interno de 761 pb del gen por PCR utilizando los iniciadores B25A (5'-TTCTGCAGCCGCGGCCCTTCG-3') (SEQ ID NO: 81) y B25B (5'-AGAATTCGCCGCGCCGCTG-3') (SEQ ID NO: 82) utilizando como molde cosBor32A2. El producto se purificó, se digirió con *PstI*-*EcoRI* y se clonó en pOJ260ermE* que se había digerido análogamente, para proporcionar pOJEd1. Este enfoque se utilizó a fin de evitar posibles efectos polares. El vector pOJEd1 se introdujo en *S. parvulus* Tü4055 por transformación de protoplastos como se describe en el Ejemplo 5, y se seleccionaron colonias respecto a resistencia a apramicina en R5 y luego en agar MA. La desorganización se verificó por hibridación Southern y el nuevo mutante se designó *S. parvulus* Tü4055/*borE::aac3(IV)*. La cepa *S. parvulus* Tü4055/*borE::aac3(IV)* se cultivó, se extrajo y se analizó como se describe en el Ejemplo 1. No se observó producción alguna de borrelidina, en tanto que un control de tipo salvaje producía borrelidina como era de esperar.

Para comprobar que no se introducían efecto polar alguno, se introdujo una copia de longitud total de *borE* bajo el control del promotor ermE* en *trans* respecto al mutante desorganizado. Se amplificó *borE* de longitud total por

PCR utilizando los iniciadores B7T1 (5'-GGCTGCAGACGCGGCTGAAG-3') (SEQ ID NO: 83) y B7T2 (5'-CCG GATCCCAGAGCCACGTC-3') (SEQ ID NO: 84), utilizando como molde cosBor32A2. El producto de 1216 pb se purificó, se digirió con *Pst*I-*Bam*HI y se clonó en pIJ2925 digerido con *Pst*I-*Xba*I (Janssen & Bibb, 1993), junto con un fragmento de pLHyg digerido con *Bam*HI-*Spe*I que contenía la casete de resistencia a la higromicina, para generar pIJEH. Un fragmento *Bam*HI de 2,8 kpb se cortó de pIJEH y se clonó en pEM4 (Quiros *et al.*, 1998), que se había digerido análogamente, para dar pborEH (en el cual el gen *borE* estaba clonado en la orientación correcta para expresión del gen). Se introdujeron pborEH y el plásmido de control pEM4 en *S. parvulus* Tü4055/*borE::aac(3)IV* por transformación de protoplastos como se describe en el Ejemplo 5. La cepa resultante de *S. parvulus* Tü4055/*borE::aac(3)IV/pborEH* se analizó como se describe en el Ejemplo 1 y se demostró que producía borrelidina a un título similar a un control de tipo salvaje WT; la cepa de control *S. parvulus* Tü4055/*borE::aac(3)IV/pEM4* no producía borrelidina.

La complementación química de *S. parvulus* Tü4055/*borE::aac(3)IV* con ácido *trans*-1,2-diciclopentano-dicarboxílico, siguiendo el protocolo descrito en el Ejemplo 1, demostró que la cepa así desarrollada era capaz de producción de borrelidina en una proporción de $122 \pm 23\%$ del control parental de tipo salvaje WT. Así pues, *borE* es necesario para la biosíntesis del ácido *trans*-1,2-ciclopentano-dicarboxílico.

Ejemplo 9

Disgregación de *borG* (*S. parvulus* Tü4055/*borG::aac(3)IV*)

Con objeto de desorganizar *borG*, se amplificó una región interna de 885 pb por PCR utilizando los iniciadores B23A (5'-ATCTGCAGCGGCATCGGTGT-3') (SEQ ID NO: 105) y B23B (5'-AGAATTCTCCACTGCGGTCTG-3') (SEQ ID NO: 106) y el cósmido Bor32A2 como molde. El producto resultante se purificó y se digirió en los sitios flanqueantes *Pst*I-*Eco*RI y se subclonó luego en pOJ260P, aguas abajo del promotor *ermE**, para generar pOJGd1.

El vector pOJGd1 se introdujo en *S. parvulus* Tü4055 por transformación de protoplastos como se describe en el Ejemplo 5. Las colonias resistentes a apramicina se seleccionaron sobre agar MA. La desorganización se verificó por hibridación Southern y el nuevo mutante se designó *S. parvulus* Tü4055/*borG::aac(3)IV*. La cepa *S. parvulus* Tü4055/*borG::aac(3)IV* se cultivó, se extrajo y se analizó como se describe en el Ejemplo 1. La producción de borrelidina se comparó con un control de tipo salvaje. Adicionalmente, *S. parvulus* Tü4055/*borG::aac(3)IV* se complementó químicamente con ácido *trans*-1,2-diciclo-pentano-dicarboxílico, siguiendo el protocolo descrito en el Ejemplo 1.

Ejemplo 10

Disgregación de *borI* (*S. parvulus* Tü4055/*borI::aac(3)IV*)

El gen *borI* y el DNA circundante se amplificó a partir de cosBor19B9 utilizando los iniciadores PCR BP4501 (5'-CGTATGCATGGCGCCATGGA-3') (SEQ ID NO: 85) y BP4502 (5'-AGCCAATTGGTGCCTCCAG-3') (SEQ ID NO: 86). El producto de 2,32 kpb se purificó, se digirió con *Nsi*I-*Mfe*I y se clonó en pSL1180 digerido con *Nsi*I-*Eco*RI, para dar el plásmido pSLI. La casete de resistencia a apramicina se cortó de pEFBA como un fragmento *Eco*RI y se clonó en pSLI digerido con *Eco*RI para dar el plásmido pSLIA. Finalmente, la casete de resistencia a higromicina se cortó con *Spe*I-*Pst*I de pLHyg y se clonó en pSLIA que se había digerido con *Nsi*I-*Spe*I para dar el plásmido pSLIr1.

El vector de reemplazamiento pSLIr1 se introdujo en *S. parvulus* Tü4055 por transformación de protoplastos como se describe en el Ejemplo 5. Se seleccionaron las colonias resistentes a apramicina (25 µg/ml) y se sometieron luego a pasos varias veces a través de medio MA sin selección. El reemplazamiento se comprobó por hibridación Southern y el nuevo mutante se designó *S. parvulus* Tü4055/*borI::aac(3)IV*.

Se cultivó *S. parvulus* Tü4055/*borI::aac(3)IV* y se analizó como se describe en el Ejemplo 1. No se observó producción alguna de borrelidina, en tanto que se observaron varios compuestos nuevos a niveles significativamente menores. Uno de los compuestos menos polares exhibía un máximo de absorbancia UV de 240 nm, y el análisis LCMS indicó una relación *m/z* de que era 11 unidades másicas menor que en el caso de la borrelidina, lo cual es consistente con la presencia de un grupo metilo en lugar de un grupo nitrilo en C12.

Para comprobar que no se introducía efecto polar alguno, se introdujo una copia de longitud total de *borI* bajo el control del promotor *ermE** en *trans* respecto al mutante desorganizado. Se recuperó un fragmento *Nsi*I-*Avr*II de 2,1 kb que contenía *borI* a partir de pSLI y se subclonó en los sitios *Pst*I-*Xba*I de pEM4, junto con el fragmento *Nhe*I-*Spe*I de pLHyg que contenía el gen *hyg*. Ambos fragmentos se subclonaron en la misma orientación, generando pborIH. El plásmidopborIH y el plásmido de control pEM4 se introdujeron en *S. parvulus* Tü4055/*borI::aac(3)IV* por transformación de protoplastos como se describe en el Ejemplo 5. La cepa resultante *S. parvulus* Tü4055/*borI::aac(3)IV/pborIH* se analizó como se describe en los Ejemplos 1 & 4, demostrándose que produce borrelidina con un título similar a un control de tipo salvaje WT.

ES 2 287 552 T3

Ejemplo 11

Disgregación de borJ (S. parvulus Tü4055/borJ::aac3(IV))

5 El gen *borJ* y el DNA circundante se amplificaron a partir de cosBor19B9 utilizando los iniciadores PCR BNHT1 (5'-GTCATGCATCAGCGCACCCG-3') (SEQ ID NO: 87) y BNHT2 (5'-GTGCAATTGCCCTGGTAGTC-3') (SEQ ID NO: 88). El producto de 2,75 kpb se purificó, se digirió con *NsiI-MfeI* y se clonó en pSL1180 que se había digerido con *NsiI-EcoRI*, para dar el plásmido pSL. La casete de resistencia a higromicina se cortó de pLHyg como un fragmento *PstI-SpeI* y se clonó en pSL digerido con *NsiI-SpeI*, para dar pSLJH. Finalmente, se cortó la casete de resistencia a apramicina de pEFBA con *SpeI-BamHI* y se clonó en pSLJH que había sido pre-digerido con *AvrII-BglII* a fin de eliminar un fragmento de 453 pb de *borJ*, para dar el plásmido pSLJr1.

15 El vector de reemplazamiento pSLJr1 se introdujo en *S. parvulus* Tü4055 por transformación de protoplastos como se describe en el Ejemplo 5. Se seleccionaron las colonias resistentes a apramicina (25 g/ml), y se sometieron luego a pasos varias veces a través de medio MA sin selección. El reemplazamiento se comprobó por hibridación Southern. El nuevo mutante se designó *S. parvulus* Tü4055/*borJ::aac3(IV)*.

20 *S. parvulus* Tü4055/*borJ::aac3(IV)* se cultivó y se analizó como se describe en el Ejemplo 1. No se observó producción alguna de borrelidina, mientras que se observó un compuesto nuevo más polar que la borrelidina con un máximo UV a 262 nm. El análisis LCMS indicó un compuesto parental de 508 amu, lo cual es consistente con una función ácido carboxílico más bien que una función nitrilo en C12.

25 Para comprobar que no se introducía efecto polar alguno, se introdujo una copia de longitud total de *borJ* bajo el control del promotor ermE* en *trans* respecto al mutante desorganizado. Un fragmento *NsiI-SphI* de pSLJ de 2,4 kb, que contenía *borJ* se subclonó en los sitios *PstI-XbaI* de pEM4 junto con el gen *hyg* como un fragmento *SphI-SpeI* de pLHyg; ambos fragmentos se subclonaron en la misma orientación que la transcripción de los genes. El constructo final se designó pborJH. El plásmido pborJH y el plásmido de control pEM4 se introdujeron en *S. parvulus* Tü4055/*borJ::aac3(IV)* por transformación de protoplastos como se describe en el Ejemplo 5. La cepa resultante *S. parvulus* Tü4055/*borJ::aac3(IV)/pborJH* se analizó como se describe en los Ejemplos 1 & 4, y se demostró que produce borrelidina con un título similar a un control de tipo salvaje WT.

Ejemplo 12

Disgregación de borG y borI (S. parvulus Tü4055/borG::aac3(IV)/borI::hyg)

35 El gen *hyg* se aísla a partir de pLHyg como un fragmento *EcoRV* y se clona en pSLI (Ejemplo 10) digerido con *EcoRI* y tratado con fragmento Klenow para dar pSLIH; el gen *hyg* se clona en la misma orientación que *borI*. Se introduce pSLIH en *S. parvulus* Tü4055/*borG::aac3(IV)* por transformación de protoplastos, como se describe en el Ejemplo 5, y se selecciona para resistencia tanto a apramicina como a higromicina, después de lo cual se somete a pasos varias veces en medio MA sin selección a fin de promover la recombinación doble. Las colonias resistentes a apramicina e higromicina se analizan por hibridación Southern y PCR para verificar el reemplazamiento.

Ejemplo 13

Disgregación de borG y borJ (S. parvulus Tü4055/borG::aac3(IV)/borJ::hyg)

50 Se aísla el gen *hyg* a partir de pLHyg como un fragmento *EcoRV* y se clona en pSLJ (Ejemplo 11) digerido con *AvrII-BglIII* y tratado con Klenow, para dar pSLJH; el gen *hyg* se clona en la misma orientación que *borI*. Se introduce pSLJH en *S. parvulus* Tü4055/*borG::aac3(IV)* por transformación de protoplastos, como se describe en el Ejemplo 5, y se selecciona para resistencia tanto a apramicina como a higromicina, después de lo cual se somete a pasos varias veces en medio MA sin selección a fin de promover la recombinación doble. Las colonias resistentes a apramicina e higromicina se realizan por hibridación Southern y PCR para verificar el reemplazamiento.

Ejemplo 14

Producción de 12-desnitrilo-12-metil-borrelidina 14 (pre-borrelidina)

55 Se inocularon stocks de trabajo de *S. parvulus* Tü4055/*borI::aac3(IV)* (0,5 ml) en precultivos vegetativos primarios de NYG como se describe en el Ejemplo 1. Se prepararon pre-cultivos secundarios (como en el Ejemplo 1 pero con 250 ml de NYG en matraces Erlenmeyer de 2 l). Medio de producción PYDG (4 l), preparado con el Ejemplo 1 y con 0,01% de Plutronic L0101 añadido para reprimir la formación de espuma, se inoculó con pre-cultivo secundario (12,5% de inóculo). Un segundo fermentador que contenía medio de punto central (4 l) y 0,01% de Plutronic L0101 para reprimir la formación de espuma, se ajustó en paralelo y se inoculó también con pre-cultivo secundario (12,5% de inóculo). El medio de producción de punto central contiene, por litro de agua desionizada: polvo de leche desnatada Tesco (1,5%), Avidex W-80 (4,5%), glucosa (0,5%) y autolizado de levadura (0,15%) ajustado a pH 7,0 con NaOH 5 M.

ES 2 287 552 T3

Se dejaron fermentar cada uno de estos lotes en un fermentador Applikon de 7 l durante 6,5 días a 30°C. Se ajustó la corriente de aire a 0,75 vvm (volumen por volumen por minuto), con deflectores inclinados y se controló la velocidad del impulsor entre 400 y 800 rpm para mantener una tensión de oxígeno disuelto a o por encima de 30% de la saturación de aire. No se añadió cantidad alguna de antiespumante adicional. Al cabo de 22 horas se añadió a la fermentación el ácido de iniciación, ácido *trans*-ciclopentano-1,2-dicarboxílico, como una solución neutralizada de 1:1 MeOH/NaOH 5 M, a través de un filtro instalado en línea (0,22 μ m). La concentración final de ácido iniciador exógeno en el recipiente fermentador era 0,5 mM.

Después de 6,5 días de fermentación, se combinaron los caldos y se acidificaron a pH 3,5 con HCl concentrado (~6 ml), después de lo cual se clarificaron por centrifugación a 3500 rpm durante 10 minutos. Se extrajo el sobrenadante en acetato de etilo (3 x 1 volumen equivalente durante 4 horas cada vez) y el sedimento de células se dejó en remojo en metanol (2 x 1,5 litros durante 4 horas cada vez). Se reunieron las fases orgánicas y se eliminaron a presión reducida para dar una goma alquitranosa. Se resuspendió la goma en tampón de bórax 0,1 M (500 ml a pH 9,4) y se lavó con hexanos (500 ml) y acetato de etilo (500 ml). La capa acuosa se acidificó luego con HCl concentrado a pH 3,5 y se extrajo con acetato de etilo (3 x 500 ml), que se combinaron y se llevaron a sequedad. La goma resultante se disolvió en metanol (15 ml), se diluyó con agua (285 ml) y se cargó por gravedad en un cartucho C₁₈ de fase inversa (50 g, preparado en metanol acuoso al 5%). El cartucho se lavó con metanol acuoso al 20% y al 50% (300 ml de cada uno) y se eluyó con metanol al 100% (500 ml). Esta última fracción se llevó a sequedad a presión reducida para dar un aceite gomoso negro (600 mg) que se recogió en metanol. Este residuo se purificó finalmente por HPLC preparativa secuencial en fase inversa (eluida con las fases móviles utilizadas en el Ejemplo 4, sin TFA añadido, realizándose la ejecución isocráticamente para 40% B). Las fracciones activas se reunieron y se desalaron en un cartucho C₁₈ (1 g), para dar 28 mg de un aceite oscuro (3,5% mg/l de rendimiento aislado). La Tabla 12 resume los datos de desplazamiento químico NMR ¹H y ¹³C para la 12-desnitrilo-12-metil-borreledina 14 en CDCl₃.

(Tabla pasa a página siguiente)

ES 2 287 552 T3

TABLA 12

Posición	d _H (ppm)	Multiplicidad	Acoplamiento (Hz)	d _C (ppm)
1	-	-	-	174.5
2a	2.29	m	-	37.8
2b	2.26	m	-	-
3	3.85	dt	9.0, 3.0	71.9
4	1.83	m	-	35.1
5a	1.19	bt	13.5	43.6
5b	0.91	m	-	-
6	1.75	m	-	27.0
7a	1.08	m	-	49.2
7b	0.88	m	-	-
8	1.69	m	-	26.5
9a	0.97	m	-	38.3
9b	0.45	t	12.5	-
10	1.62	m	-	34.1
11	3.53	d	9.0	85.7
12	-	-	-	138.4
13	5.84	d	11.0	127.7
14	6.28	ddd	14.5, 11.0, 1.0	129.6
15	5.48	ddd	14.5, 10.5, 3.5	129.9
16a	2.53	m	-	39.1
16b	2.22	m	-	-
17	5.07	ddd	11.0, 8.0, 3.0	76.5
18	2.52	m	-	48.0
19a	1.92	m	-	30.4
19b	1.32	m	-	-
20a	1.74	m	-	26.2
20b	1.71	m	-	-

TABLA 12 (continuación)

Posición	d _H (ppm)	Multiplicidad	Acoplamiento (Hz)	d _C (ppm)
21a	1.96	m	-	32.0
21b	1.84	m	-	-
22	2.45	m	8.0	49.3
23	-	-	-	182.3
4-CH ₃	0.78	d	6.5	18.5
6-CH ₃	0.77	d	6.5	18.8
8-CH ₃	0.75	d	6.5	20.6
10-CH ₃	0.94	d	6.5	16.3
12-CH ₃	1.64	s	-	11.4
Se hace referencia a los desplazamientos químicos como CDCl ₃ (para ¹ H a 7,26 ppm y para ¹³ C a 77,0 ppm)				

Ejemplo 15

Producción de 12-desnitrilo-12-carboxi-borrelidina 2

Se inocularon stocks de trabajo de *S. parvulus* Tü4055/borJ::aac3(IV) (0,5 ml) en precultivos vegetativos primarios de NYG como se describe en el Ejemplo 1. Se prepararon pre-cultivos secundarios (como en el Ejemplo 1 pero con 250 ml de NYG en matraces Erlenmeyer de 2 l). El medio de producción PYDG (4 l), preparado como en el Ejemplo 1 y con 0,01% de Plutronic L0101 añadido para control de la espuma, se inoculó con el pre-cultivo secundario entero (10% inóculo). Se dejó fermentar éste en un fermentador Applikon de 7 l durante 6 días a 30°C. Se ajustó el caudal de aire a 0,75 vvm, con deflectores inclinados y con la velocidad del impulsor controlada entre 250 y 600 rpm para mantener una tensión de oxígeno disuelto a o por encima de 30% de la saturación de aire. No se añadió antiespumante adicional alguno. Se realizó una segunda fermentación exactamente como en el caso anterior, pero que se alimentó por lotes con 0,2 mol de glucosa como solución acuosa cada 12 horas a partir de las 60 horas después de la inoculación.

Después de 6 días se recogieron las fermentaciones y se combinaron. Se clarificó el caldo por centrifugación (3500 rpm, 10 minutos) y el sobrenadante resultante se acidificó con HCl 10 M (aq) hasta pH ~3,5. Esta solución se extrajo luego en acetato de etilo por agitación (3 x 1 equivalente volumétrico durante 4 horas cada uno). El sedimento de células se extrajo dos veces por remojo de las células en metanol/acetato de etilo 1:1 (500 ml). Todas las fases orgánicas se reunieron y se eliminaron a presión reducida para dar un lodo acuoso. El lodo se diluyó a 500 ml con agua, se acidificó a pH ~3,5 con HCl 10 M y se extrajo en acetato de etilo (3 x 300 ml). Las fases orgánicas se concentraron a presión reducida hasta ~300 ml y se extrajeron con bórax 0,1 M (3 x 150 ml, pH = 9,4). Las soluciones de bórax reunidas se acidificaron con HCl 10 M a pH ~3,5 y se extrajeron con 6 x 300 ml de acetato de etilo. La HPLC analítica demostró que algo del acumulante permanecía todavía en la solución de bórax y por tanto se cargó ésta, por gravedad, en un cartucho C₁₈ de fase inversa (50 g). El cartucho se lavó con agua y el acumulante se eluyó en 100% metanol. Las fases orgánicas que contenían el acumulante se reunieron y se redujeron a una solución metabólica de 40 ml. Se cargó ésta en una columna Sephadex LH-20 (70 g, hinchada durante una noche en metanol, columna de 60 cm x 2,5 cm), que se reveló con metanol al 100%; las fracciones activas se reunieron y se llevaron a sequedad. El material se procesó luego adicionalmente por HPLC preparativa en fase inversa (eluida con las fases móviles utilizadas en el Ejemplo 4, sin TFA añadido, ejecutándose la realización isocráticamente para 40% B). Las fracciones activas reunidas se llevaron a sequedad, se disolvieron en metanol (4 ml) y se diluyeron con agua (200 ml). Esta mezcla se diluyó en 2 fracciones iguales y se cargaron cada una, por gravedad, en un cartucho C₁₈ de fase inversa (20 g). Las columnas se eluyeron luego con 3 volúmenes de columna de 5%, 10%, 20%, 30%, 40%, 50%, 60%, 75%, 90%, y 100% de metanol acuoso. El acumulante se eluyó en todas las fracciones desde 60% a 100% de metanol, que se reunieron y se llevaron a sequedad. El acumulante (disuelto en DMSO) se purificó luego finalmente por HPLC preparativa secuencial en fase inversa (eluida con las fases móviles utilizadas en el Ejemplo 4, sin TFA añadido, realizándose la ejecución isocráticamente para 40% B). Las fracciones activas se reunieron y se desalaron en un cartucho C₁₈ (1 g), para dar 17 mg de un aceite pardo (2,1 ml de rendimiento aislado). La Tabla 13 resume los datos de desplazamiento químico NMR ¹H y ¹³C para 12-desnitrilo-12-carboxi-borrelidina 2 en δ₄-metanol.

ES 2 287 552 T3

TABLA 13

Posición	d _H (ppm)	Multiplicidad	Acoplamiento (Hz)	d _C (ppm)
1	-	-	-	173.27
2a	2.40	dd	15.8, 4.1	39.31
2b	2.29	dd	15.8, 8.2	
3	3.87	m		71.64
4	1.80	m		36.51
5a	1.29	m		44.24
5b	0.90	m		
6	1.59	m		27.48
7a	1.09	m		~49.0*
7b	1.03	m		
8	1.72	m		28.17
9a	1.12	m		38.42
9b	0.79	m		
10	2.03	m		36.43
11	3.90	m		81.95
12	-	-	-	132.35
13	6.43	d	11.0	140.83
14	6.96	dd	14.5, 11.5	130.91
15	5.91	ddd	15.0, 9.5, 5.0	138.93
16a	2.61	m	15.0	38.57
16b	2.36	m		
17	5.04	m		77.40
18	2.50	m		49.80
19a	1.90	m		30.59
19b	1.32	m		
20a	1.85	m		26.34
20b	1.41	m		
21a	1.97	m		32.40
21b	1.75	m		
22	2.52	m		~48.0*
23	-	-	-	180.27
4-CH ₃	0.83	d	7.0	18.76
6-CH ₃	0.80	d	6.0	17.06
8-CH ₃	0.81	d	6.5	20.60
10-CH ₃	0.93	d	6.5	16.61
12-CO ₂ H -		-	-	170.49

Los desplazamientos químicos están referenciados a metanol (para ¹H a 3.35 ppm (quintete) y para ¹³C a 49.0 ppm (septete)); * Oscurecido por la señal de disolvente, δ₄-metanol.

Ejemplo 16

Producción por mutasíntesis de 17-des-(ácido ciclopentano-2'-carboxílico)-17-(ácido ciclobutano-2'-carboxílico) borrelidina 18

Se inocularon stocks de trabajo de *S. parvulus* Tü4055/*borE::aac3(IV)* (0,5 ml) en precultivos vegetativos primarios de NYG como se describe en el Ejemplo 1. Se prepararon pre-cultivos secundarios (como en el Ejemplo 1 pero con 250 ml de NYG en matraces Erlenmeyer de 2 l). El medio de producción PYDG (4 l), preparado como en el Ejemplo 1 y con 0,01% de Plutronic L0101 añadido para control de la espuma, se inoculó con pre-cultivo secundario (12,5% de inóculo). Se dispusieron 2 biorreactores adicionales del mismo modo. Estos lotes se dejaron fermentar cada uno en un fermentador Applikon de 7 l durante 5 días a 30°C. Se ajustó el caudal de aire a 0,75 vvm (volumen por volumen por minuto), con deflectores inclinados y con la velocidad del impulsor controlada entre 400 y 700 rpm para mantener la tensión de oxígeno disuelto a o por encima del 30% de la saturación del aire. No se añadió antiespumante adicional alguno. Al cabo de 22 horas de fermentación, se añadió el ácido iniciador, ácido *trans*-ciclobutano-1,2-dicarboxílico como una solución neutralizada de MeOH/NaOH 5 M 1:1. La concentración final en el recipiente fermentador de ácido de iniciación exógeno era 0,5 mM.

Después de 5 días de fermentación, se reunieron los caldos y se acidificaron a pH 4,0 con HCl concentrado, después de lo cual se clarificaron por centrifugación a 3500 rpm durante 10 minutos. Se absorbió el sobrenadante en resina Diaion HP-20SS (1 l), que se había tratado previamente con metanol (2 l) y a continuación con metanol acuoso al 5% (2 l), por filtración a una tasa de aproximadamente 100 ml/min. La resina se eluyó luego con metanol acuoso al 20% (2,5 l) y luego con acetona acuosa al 80% (4,5 l). Se eliminó el disolvente orgánico de la acetona acuosa y el lodo acuoso resultante (1 litro) se extrajo en acetato de etilo (3 x 1 l). Se reunieron las fases orgánicas y se recogieron a vacío para dar un aceite amarillo/pardo (1,7 g). Entretanto, el sedimento de células se dejó en remojo en metanol-acetato de etilo, 1:1 (3 x 1 l durante 4 horas cada vez), y los sobrenadantes orgánicos resultantes se redujeron a vacío para dar un lodo acuoso (400 ml). La materia constituida por partículas se disolvió en metanol (50 ml), y se añadió de nuevo al lodo acuoso, que se completó hasta 500 ml con agua. Este lodo se absorbió sobre resina Diaion HP-20SS (300 ml), que se había tratado previamente con metanol (500 ml) y luego con metanol acuoso al 5% (500 ml). La resina se eluyó luego con metanol acuoso al 20% (1 l) y después con acetona acuosa al 80% (1,5 l). El disolvente orgánico se separó de la acetona acuosa y el lodo acuoso resultante (llevado a 750 ml) se extrajo en acetato de etilo (3 x 750 ml). Se reunieron las fases orgánicas y se redujeron a vacío para dar un aceite amarillo/pardo (1,7 g). Los extractos brutos se reunieron (3,4 g), se disolvieron en acetato de etilo (10 ml), se adsorbieron luego en una columna de sílice (5 cm de diámetro interior x 10 cm, tratada con EtOAc), y se eluyeron con EtOAc. Las fracciones activas se reunieron y el disolvente se eliminó a vacío para dar una goma parda (1,08 g). Este residuo se purificó finalmente por HPLC preparativa secuencial en fase inversa (eluida con las fases móviles utilizadas en el Ejemplo 4, sin TFA añadido, realizándose la ejecución desde 25% B a 75% B durante 25 minutos con un gradiente lineal). Las fracciones activas se reunieron y se desalaron en un cartucho C₁₈ (5 g), para dar 83,9 mg (o 7,0 mg/l de rendimiento aislado). El espectro ¹³C-NMR de 18 se muestra en la Tabla 14.

Tabla pasa a página siguiente)

TABLA 14

d _c (ppm)	Posición	d _c (ppm)	Posición
177.1	COOH (C22)	37.3	9
172.2	1	35.7	4
144.0	13	35.1	10
138.7	15	34.4	16
126.9	14	30.9	18
118.3	12	27.3	6
115.8	CN	26.2	8
75.5	17	21.7	20
73.1	11	21.0	21
69.7	3	20.1	8-Me
47.6	5	18.1	6-Me
43.1	7	16.9	4-Me
40.1	19	14.9	10-Me

40,0	2		
Asignación de ¹³ C-NMR para 18 , en CDCl ₃ , utilizando dicha señal de carbono como referencia para d _c = 77,7 ppm			

Referencias

Anderson, B.F., Herit, A.J., Rickards, R.W., and Robertson, G.B. (1989) Crystal and molecular structures of two isomorphous solvates of the macrolide antibiotic borrelidin: absolute configuration determination by incorporation of a chiral solvent in the in the crystal lattice. *Aust. J. Chem.* **42**:717-730.

Anderton, K., and Rickards, R.W. (1965) Some structural features of borrelidin, an anti-viral antibiotic. *Nature* **206**:269.

Aparicio, J.F., Molnár, I., König, A., Haydock, S.H., Khaw., L.E., Staunton, J., and Leadlay, P.F. (1996) Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. *Gene* **169**:9-16.

August, P.R., Tang, L., Yoon, Y.J., Ning, S., Müller, R., Yu, T.-W., Taylor, M., Hoffmann, D., Kim, C.G., Zhang, X.H., Hutchinson, C.R., and Floss, H.G. (1998)

Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **5**:69-79.

Beck, J.B., Yoon, Y.J., Reynolds, K.A., and Sherman, D.H. (2002) The hidden steps of domain skipping: ring size determination in the pikromycin modular polyketide synthase. *Chem. Biol.* **9**:575-583.

Berger, J., Jampolsky, L.M., and Goldberg, M.W. (1949) Borrelidin, a new antibiotic with anti-Borrelia activity and penicillin enhancement properties. *Arc. Biochem.* **22**:476-478.

Bierman, M., Logan, R., O'Brian, K., Seno, E.T., Rao, N., and Schoner, B.E. (1992) Plasmid vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**:43-49.

Bonfield, J.K., Smith, K.F., and Staden, R. (1995) A new DNA sequence assembly program. *Nucleic Acids Research* **23**:4992-4999.

Brautaset, T., Sekurova, O.N., Sletta, H., Ellingsen, T.E., Strøm, A.R., Valla, S., and Zotchev, S.B. (2000) Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. *Chem. Biol.* **7**:395-403.

Brenner, S. (1998) The molecular evolution of genes and proteins: a tale of two serines. *Nature* **334**:528-530.

Brosius, J. (1989) Super-polylinkers in cloning and expression vectors. *DNA* **8**:759-777.

Butler, A.R., Bate, N., and Cundliffe, E. (1999) Impact of thioesterase activity on tylosin biosynthesis in *Streptomyces fradiae*. *Chem. Biol.* **6**:287-292.

Caffrey, P., Lynch, S., Flood, E., Finnan, S., and Oliynyk, M. (2001) Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem. Biol.* **8**:713-723.

Celenza, J.L. (2001) Metabolism of tyrosine and tryptophan - new genes for old pathways. *Curr. Opin. Plant Biol.* **4**:234-240.

Chater, K.F. and Wilde, L.C. (1980) *Streptomyces albus* G mutants defective in the SalG1 restriction modification system. *J. Gen. Microbiol.* **116**:323-334.

Cheng, Y.Q., Tang, G.L., and Shen B. (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci. USA.* **100**:3149-3154.

Cortés J., Haydock, S.F., Roberts, G.A., Bevitt, D.J., and Leadlay, P.F. (1990) An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **348**:176-178.

Cortés, J., Weissman, K.E.H., Roberts, G.A., Brown, M.J.B., Staunton, J., and Leadlay, P.F. (1995) Repositioning of a domain in a modular polyketide synthase to promote specific chain elongation. *Science* **268**:1487-1489.

Cortés, J., Velasco, J., Foster, G., Blackaby, A.P., Rudd, B.A.M., and Wilkinson, B. (2002) Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in *Saccharopolyspora erythraea*. *Mol. Micro.* **44**:1213-1224.

Devereux, J., Heaberli, P., and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**:387-395.

Dickinson, L., Griffiths, A.J., Mason, C.G., and Mills, R.F. (1965) Anti-viral activity of two antibiotics isolated from a species of *Streptomyces*. *Nature* **206**:265-268.

Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J., and Katz, L. (1991) Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**:675-679.

Donadio, S., McAlpine, J.B., Sheldon, P.J., Jackson, M., and Katz, L. (1993) An erythromycin analog produced by reprogramming of polyketide synthesis *Proc. Nat. Acad. Sci. USA* **90**:7119-7123.

Duffey, M.O., LeTiran, A., and Morken, J.P. (2003) Enantioselective total synthesis of borrelidin. *J. Am. Chem. Soc.* **125**:1458-1459.

Eastwood, E.L., and Schaus, S.E. (2003) Borrelidin induces the transcription of amino acid biosynthetic enzymes via a GCN4-dependent pathway. *Bioorg. Med. Chem. Lett.* **13**:2235-2237.

Fernandez, E., Weissbach, U., Sanchez-Reillo, C., Brana, A.F., Mendez, C., Rohr, J., and Salas, J.A. (1998) Identification of two genes from *Streptomyces argillaceus* encoding glycosyltransferases involved in transfer of a disaccharide during the biosynthesis of the antitumor drug mithramycin. *J. Bacteriol* **180**:4929-4937.

Floss, H.G. (2001) Antibiotic biosynthesis: from natural to unnatural compounds. *J. Ind. Micro. Biotech.* **27**:183-194.

Fouces, R., Mellado, E., Diez, B., and Barrado, J.L. (1999) The tylosin biosynthetic cluster from *Streptomyces fradiae*: genetic organization of the left region. *Microbiology* **145**:855-868.

Folkman, J. (1986) How is blood vessel growth regulated in normal and neoplastic tissue? G.H.A. Cloves Memorial Lecture. *Cancer Res.* **51**:467-473.

- Funahashi, Y., Wakabayashi, T., Semba, T., Sonoda, J., Kitoh, K., and Yoshimatsu, K.** (1999) Establishment of a quantitative mouse dorsal air sac model and its application to evaluate a new angiogenesis inhibitor. *Oncol. Res.* **11**:319-329.
- 5 **Gaisser, S., Reather, J., Wirtz, G., Kellenberger, L., Staunton, J., and Leadlay, P.F.** (2000) A defined system for hybrid macrolide biosynthesis in *Saccharopolyspora erythraea*. *Mol. Microbiol.* **36**:391-401.
- Gaisser, S., Martin, C.J., Wilkinson, B., Sheridan, R.M., Lill, R.E., Weston, A.J., Ready, S.J., Waldron, C., Crouse, G.C., Leadlay, P.F., and Staunton, J.** (2002) Engineered biosynthesis of novel spinosyns bearing altered deoxyhexose substituents. *Chem. Commun.* 618-619.
- 10 **Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R.** (2002) The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* **277**:13082-13090.
- 15 **Hanessian, S., Yang, Y., Giroux, S., Mascitti, V., Ma, J., and Raeppe, F.** (2003) Application of conformation design in acyclic stereoselection: total synthesis of borrelidin as the crystalline benzene solvate. *J. Am. Chem. Soc.* **125**:13784-13792.
- 20 **Hardt, I.H., Steinmetz, H., Gerth, K., Sassa, F., Reichenbach, H., and Höfle, G.** (2001) New natural epothilones from *Sorangium cellulosum*, strains So ce90/B2 and So ce90/D13: isolation, structure elucidation, and SAR studies. *J. Nat. Prod.* **64**:847-856.
- Heathcote, M.L., Staunton, J., and Leadlay, P.F.** (2001) Role of type II thioesterases: evidence for the removal of short acyl chains produced by aberrant decarboxylation of chain extender units. *Chem. Biol.* **8**:207-220.
- 25 **Hopwood, D.** (1997) Genetic contributions to understanding polyketide biosynthesis. *Chem. Rev.* **97**:2465-2497.
- Hunziker, D., Yu, T.-W., Hutchinson, C.R., Floss, H.G., and Khosla, C.** (1998) Primer unit specificity in biosynthesis principally resides in the later stages of the biosynthetic pathways. *J. Am. Chem. Soc.* **120**:1092-1093.
- 30 **Janssen, G.R., Bibb, M.J.** (1993) Derivatives of pUC18 that have *Bgl*III sites flanking a modified cloning site and that retain the ability to identify recombinant clones by visual screening of *E. coli* colonies. *Gene* **124**:133-134.
- 35 **Kahn, R.A., Fahrendorf, T., Halkier, B.A., and Möller, B.L.** (1999) Substrate specificity of the cytochrome P450 enzymes CYP79A1 and CYP71 E1 involved in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor*(L.) Moench. *Arch. Biochem. Biophys.* **363**:9-18.
- Kawamura, T., Liu, D., Towle, M.J., Kageyama, R., Tsukahara, N., Wakabayashi, T., and Littlefield, B.A.** (2003) Anti-angiogenesis effects of borrelidin are mediated through distinct pathways: Threonyl-tRNA synthetase and caspases are independently involved in suppression of proliferation and induction of apoptosis in endothelial cells. *J. Antibiot.* **56**:709-715.
- 40 **Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A.** (2000) *Practical Streptomyces Genetics*. The John Innes Foundation, Norwich.
- 45 **Keller-Scheirlein, W.** (1967) Composition of the antibiotic borrelidin. *Helv. Chim. Acta.* **50**:731-753.
- Kuo, M.S., Yurek, D.A., and Kloosterman, D.A.** (1989) Assignment of ¹H and ¹³C NMR signals and the alkene geometry at C-7 in borrelidin. *J. Antibiot.* **42**:1006-1007.
- 50 **Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W., and Rao, R. N.** (1996) Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* **183**:231-236.
- 55 **Lozano, M.J., Remsing, L.L., Quiros, L.M., Brana, A.F., Fernandez, E., Sanchez, C., Mendez, C., Rohr, J., and Salas, J.A.** (2000) Characterization of two polyketide methyltransferases involved in the biosynthesis of the antitumor drug mithramycin by *Streptomyces argillaceus*. *J. Biol. Chem.* **275**:3065-3074.
- Maehr, H., and Evans, R.H.** (1987) Identity of borrelidin with treponemycin. *J. Antibiot.* **40**:1455-1456.
- 60 **Marsden, A.F., Wilkinson, B., Cortés, J., Dunster, N.J., Staunton, J., and Leadlay, P.F.** (1998) Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* **279**:199-202.
- Matter, A.** (2001) Tumor angiogenesis as a therapeutic target. *Drug Disc. Today* **6**:1005-1024.
- 65 **Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K., and Kinashi, H.** (2003) The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. *Mol Microbiol.* **48**:1501-1510.

Moore, B.S., and Hopke, J.N. (2000) Discovery of a new bacterial polyketide biosynthetic pathway. *Chembiochem* **2**:35-38.

Nielsen, J.S., and Møller, B.L. (1999) Biosynthesis of cyanogenic glucosides in *Triglochin maritime* and the involvement of cytochrome P450 enzymes. *Arch. Biochem. Biophys.* **368**:121-130.

Olano, C., Wilkinson, B., Moss, S.J., Brana, A.F., Mendaz, C., Leadlay, P.F., and Sala, J.A. (2003) Evidence from engineered gene fusions for the repeated use of a module in a modular polyketide synthase. *Chem. Commun.* 2780-2782.

Oliynyk, M., Brown, M.J.B., Cortés, J., Staunton, J., and Leadlay, P.F. (1996) *Chem. Biol.* **3**:833-839.

Otani, A., Slike, B.M., Dorrell, H.I., Hood, J., Kinder, K., Cheresh, D.A., Schimmel, P., and Friedlander, M. (2002) A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. *Proc. Nat. Acad. Sci. USA* **99**:178-183.

Otoguru, K., Ui, H., Ishiyama, A., Kobayashi, M., Togashi, H., Takahashi, Y., Masuma, R., Tanaka, H., Tornado, H., Yamada, H., and Ōmura, S. (2003) *In vitro* and *in vivo* antimalarial activities of a non-glycosidic 18-membered macrolide antibiotic, borrelidin, against drug-resistant strains of *Plasmodia*. *J. Antibiot.* **56**:727-729.

Pacey, M.S., Dirlam, J.P., Geldart, L.W., Leadlay, P.F., McArthur, H.A.I., McCormick, E.L., Monday, R.A., O'Connell, T.N., Staunton, J., and Winchester, T.J. (1998) Novel erythromycins from a recombinant *Saccharopolyspora erythraea* strain NRRL 2338 pIG1 I. Fermentation, isolation and biological activity. *J. Antibiot.* **51**:1029-1034.

Paetz, W., and Nass, G. (1973) Biochemical and immunological characterization of threonyl-tRNA synthetase of two borrelidin-resistant mutants of *Escherichia coli* K12. *Eur. J. Biochem.* **35**:331-337.

Prieto, M.A., Díaz, E., and García, J.L. (1996) Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli* W: engineering a mobile aromatic degradative cluster. *J. Bacteriol.* **178**:111-120.

Quiros, L.M., Aguirrezabalaga, I., Olano, C., Mandez, C., and Salas, J.A. (1998) Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*. *Mol. Microbial.* **28**:1177-1185.

Raibaud, A., Zalacain, M., Holt, T.G., Tizard, R., and Thompson, C.J. (1991) Nucleotide sequence analysis reveals linked N-acetyl hydrolase, thioesterase, transport, and regulatory genes encoded by the bialophos biosynthetic gene cluster of *Streptomyces hygroscopicus*. *J. Bacteriol.* **173**:4454-4463.

Reeves, C.D., Murli, S., Ashley, G.W., Piagentini, M., Hutchinson, C.R., and McDaniel, R. (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* **40**:15464-15470.

Rowe, C.J., Böhm, I.U., Thomas, I.P., Wilkinson, B., Rudd, B.A.M., Foster, G., Blackaby, A.P., Sidebottom, P.J., Roddis, Y., Buss, A.D., (2001) *Chem. Biol.* **8**:475-485.

Rowe, C.J., Cortés, J., Gaisser, S., Staunton, J., and Leadlay, P.F. (1998) Construction of new vectors for high-level expression in actinomycetes. *Gene* **216**:215-223.

Rudd, B.A.M., Noble, D., Foster, S.J., Webb, G., Haxell, M. (1990) The biosynthesis of a family of novel antiparasitic macrotides. Proceedings of the 6th International Symposium on the Genetics of Industrial Microorganisms. Straus-bourg, Franca. *Abstract* A70. p.96. ISBN 2-87805-004-5.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbour, Laboratory Press. New York.

Schmidt, D.M.Z., Hubbard, B.K., and Gerit, J.A. (2001) Evolution of enzymatic activities in the enolase superfamily: functional assignment of unknown proteins in *Bacillus subtilis* and *Escherichia coli* as L-Ala-DIL-Glu epimerases. *Biochemistry* **40**:15707-15715.

Schwecke, T., Aparicio, J.F., Molnár, I., König, A., Khaw, L.E., Haydock, S.F., Oliynyk, M., Caffrey, P., Cortés, J., Lester, J.B., Böhm, G.A., Staunton, J., and Leadlay, P.F. (1995) The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Nat. Acad. Sci. USA* **92**:7839-7843.

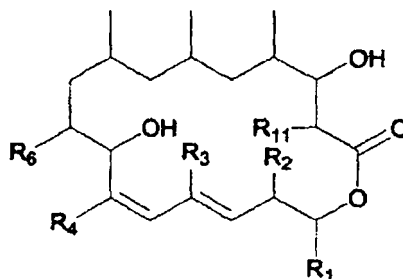
Shaw-Reid, C.A., Kelleher, N.L., Losey, H.C., Gehring, A.M., Berg, C., and Walsh, C.T. (1999) Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization. *Chem. Biol.* **6**:385-400.

- Silakowski, B., Nordsiek, G., Kunze, B., Blöcker, H., and Müller, R** (2001) Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantica* Sgal 5. *Chem. Biol.* **8**:59-69.
- 5 **Singh, S.K., Gurusiddaiah, S., and Whalen, J.W.** (1985) Treponemycin, a nitrile antibiotic active against *Treponema hyodysenteriae*. *Antimicrob. Agents Chemother.* **27**:239-245.
- Staunton, J., and Wilkinson, B.** (1997) Biosynthesis of erythromycin and rapamycin. *Chem. Rev.* **97**:2611-2629.
- 10 **Swan, D.G., Rodriguez, A.M., Vilches, C., Mendez, C., and Salas, J.A.** (1994) Characterization of a *Streptomyces antibioticus* gene encoding a type 1 polyketide synthase which has an unusual coding sequence. *Mol. Gen. Genet.* **242**:258-362.
- 15 **Takeshita, S., Sato, M., Toba, M., Masahashi, W., and Hashimoto-Gotoh, T.** (1987) High-copy number and lowcopy number plasmid vectors for *lacZ* alpha-complementation and chloroamphenicol- or kanamycin-resistance selection. *Gene* **61**:63-74.
- 20 **Thomas, I., Martin (née Rowe), C.J., Wilkinson, C.J., Staunton, J., and Leadlay, P.F.** (2002) Skipping in a hybrid polyketide synthase: evidence for ACP to ACP chain transfer. *Chem. Biol.* **9**:781-787.
- Tsuchiya, E., Yukawa, M., Miyakawa, T., Kimura, K.I., and Takahashi, H.** (2001) Borrelidin inhibits a cyclin-dependent kinase (CDK), Cdc28/Cln2, of *Saccharomyces cerevisiae*. *J. Antibiot.* **54**:84-90.
- 25 **Wakasugi, K., Slike, B.M., Hood, J., Otani, A., Ewalt, K.L., Friedlander, M., Cheresch, D.A., and Schimmel, P.** (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proc. Nat. Acad. Sci. USA* **99**:173-177.
- Wakabayashi, T., Kageyama, R., Naruse, N., Tsukahara, N., Funahashi, Y., Kitoh, K., and Watanabe, Y.** (1997) Borrelidin is an angiogenesis inhibitor, disruption of angiogenic capilla vessels in a rat aorta matrix cultura model. *J. Antibiot.* **50**:671-676.
- 30 **Waldron, C., Matsushima, P., Rosteck, P.R., Broughton, M.C., Tumer, J., Madduri, K., Crawford, K.P., Merla, D.J. and Baltz, R.H.** (2001) Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. *Chem. Biol.* **8**:487-499.
- 35 **Wilkinson, B., Foster, G., Rudd, B.A.M., Taylor, N.L., Blackaby, A.P., Sidebottom, P.J., Dawson, M.J., Buss, A.D., Gaisser, S., Böhm, I.U., Rowe, C.J., Cortés, J., Leadlay, P.F., and Staunton, J.** (2000) Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. *Chem. Biol.* **7**:111-117.
- 40 **Xue, Y.Q., Zhao, L.S., Lila, H.-W., and Sherman, D.H.** (1998) A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc. Nat Acad. Sci. USA* **95**:12111-12116.
- Xue, Y.Q., and Sherman, D.H.** (2000) Alternative modular polyketide synthase expression controls macrolactone structure. *Nature* **403**:571-575.
- 45
- 50
- 55
- 60
- 65

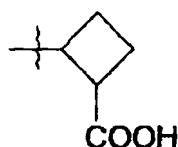
REIVINDICACIONES

1. Un compuesto de fórmula 1 o una sal farmacéuticamente aceptable del mismo, en donde:

Fórmula 1



R₁ es un grupo cicloalquilo sustituido como se muestra a continuación:



en donde R₁ puede estar sustituido también opcionalmente con uno o más átomos de halógeno o uno o más grupos alquilo C₁ a C₃;

R₂, R₃, R₆, o R₁₁ son cada uno independientemente H, OCH₃, CH₃ o CH₂CH₃; y R₄ es CN, CO₂H o CH₃.

2. Un compuesto o sal de acuerdo con la reivindicación 1, en donde R₄ es CH₃ o COOH.

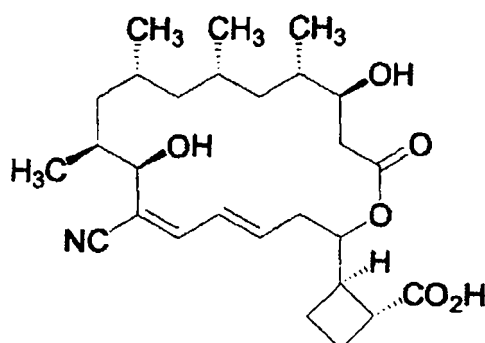
3. Un compuesto o sal de acuerdo con la reivindicación 1, en donde R₄ es CN.

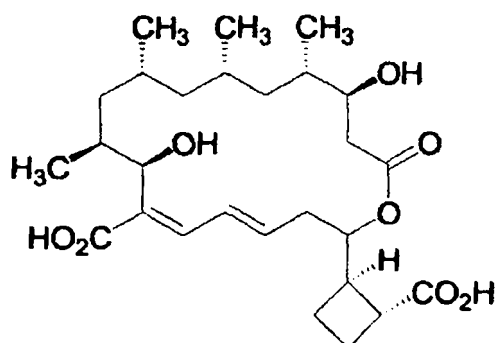
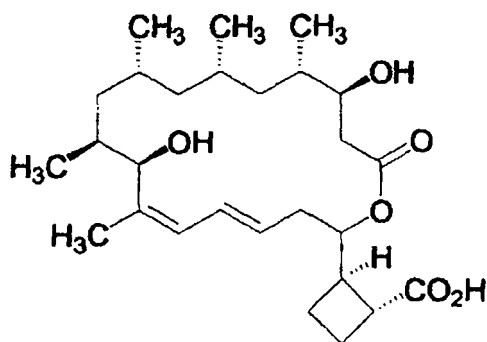
4. Un compuesto o sal de acuerdo con la reivindicación 1, en donde R₁ es ciclobutano-1'-carboxilato.

5. Un compuesto o sal de acuerdo con la reivindicación 4, en donde R₄ es CH₃ o COOH.

6. Un compuesto o sal de acuerdo con la reivindicación 1, en donde R₆ es CH₃, R₂ y R₁₁ son H y R₁ es ciclobutano-1'-carboxilato.

7. Un compuesto o sal del mismo de acuerdo con la reivindicación 1, en donde el compuesto se selecciona de la lista constituida por:





8. El uso de un compuesto de acuerdo con una cualquiera de las reivindicaciones 1 a 7 en la preparación de un medicamento para el tratamiento de afecciones microbianas que incluyen malaria, para la inhibición de la angiogénesis, para el tratamiento de trastornos proliferativos, o para el tratamiento de afecciones **caracterizadas** por vascularización inadecuada.

9. El uso de acuerdo con la reivindicación 8, en donde el medicamento es un agente anti-cáncer.

10. El uso de acuerdo con la reivindicación 8, en donde el trastorno proliferativo es psoriasis.

11. El uso de acuerdo con la reivindicación 8, en donde la afección **caracterizada** por vascularización inadecuada se selecciona del grupo constituido por psoriasis, artritis reumatoide, aterosclerosis y retinopatía diabética.

12. El uso de un compuesto de acuerdo con una cualquiera de las reivindicaciones 1 a 7 en la fabricación de un medicamento para la inhibición de la angiogénesis.

13. Un compuesto de acuerdo con una cualquiera de las reivindicaciones 1 a 7 para uso como medicamento.

14. Una composición farmacéutica que comprende un compuesto de acuerdo con una cualquiera de las reivindicaciones 1 a 7 en mezcla con un excipiente, vehículo, tampón o estabilizador farmacéuticamente aceptable.

15. Un compuesto de acuerdo con una cualquiera de las reivindicaciones 1 a 7 para uso como medicamento para el tratamiento de afecciones microbianas que incluyen malaria, para la inhibición de la angiogénesis, para el tratamiento de trastornos proliferativos, o para el tratamiento de afecciones **caracterizadas** por vascularización inadecuada.

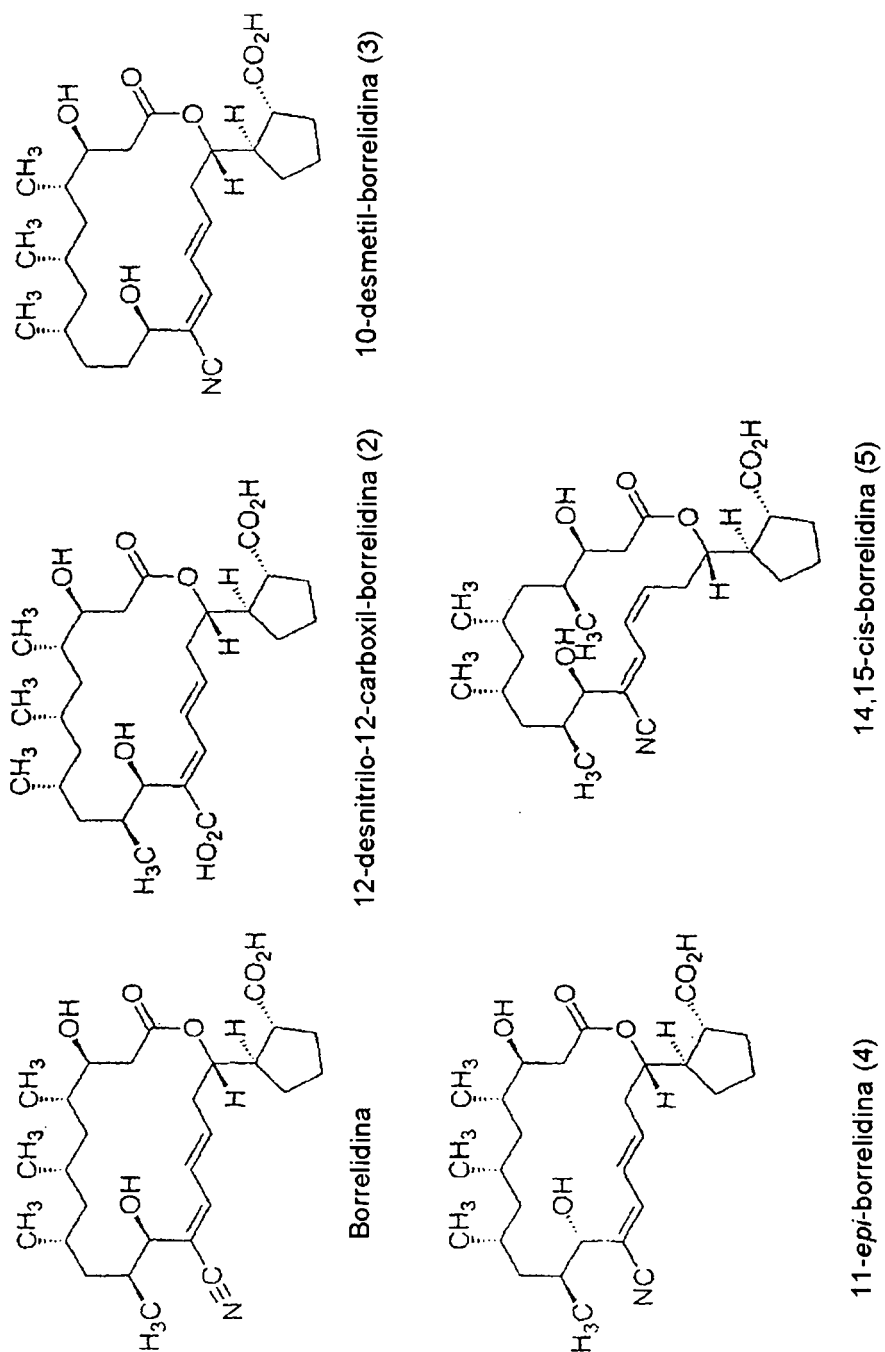
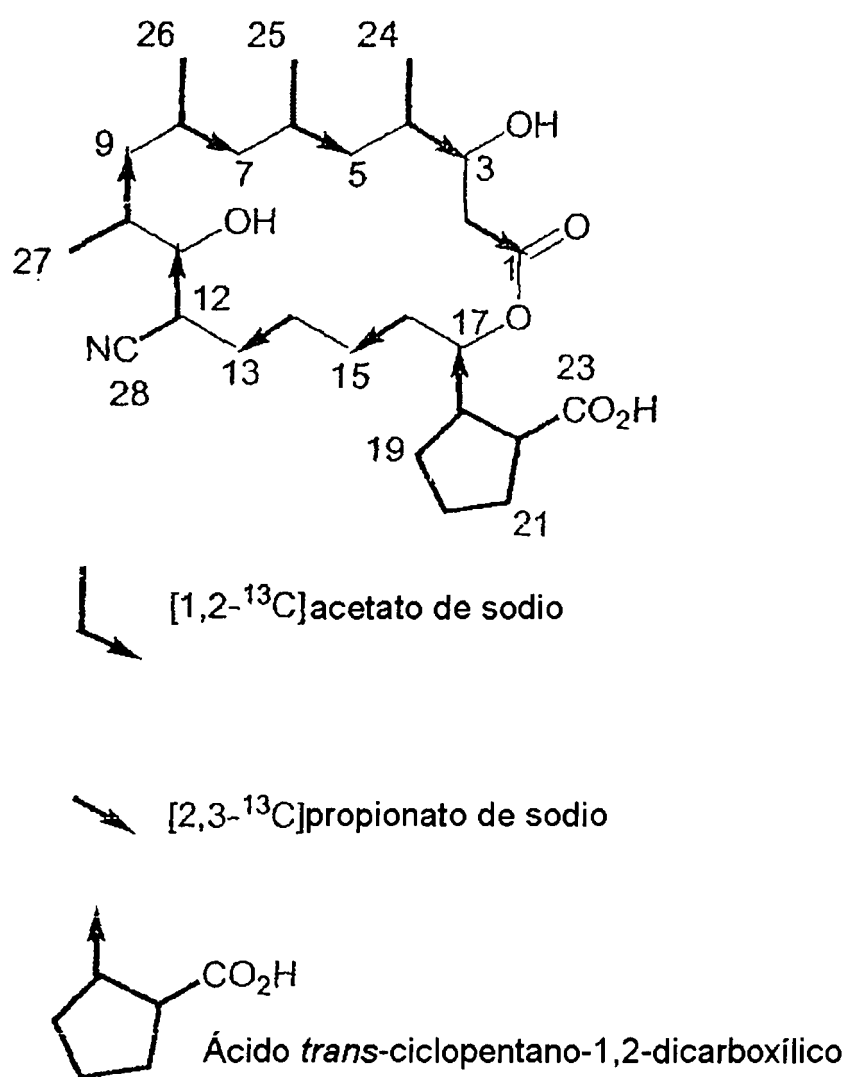


Figure 1

Figure 2



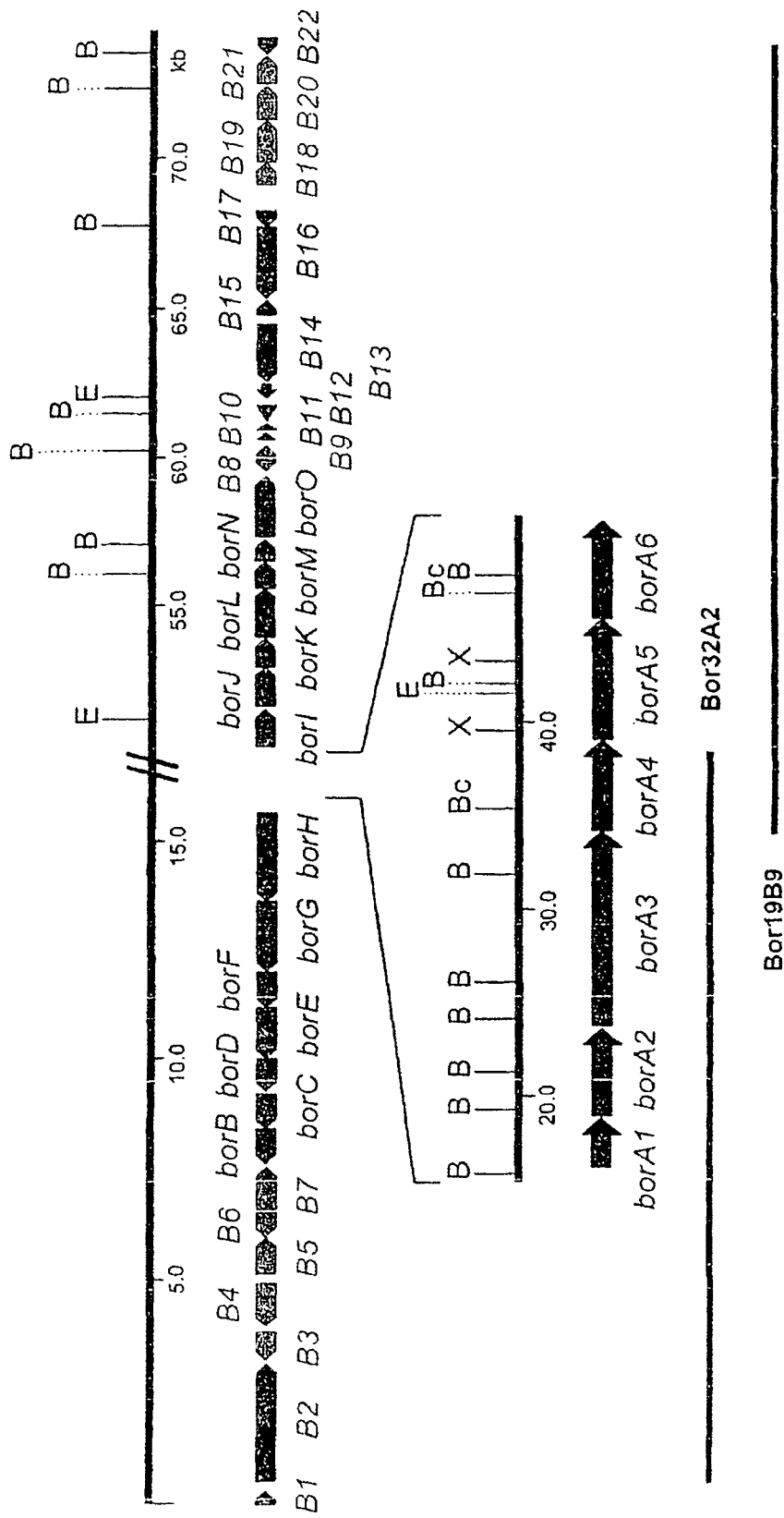


Figure 3

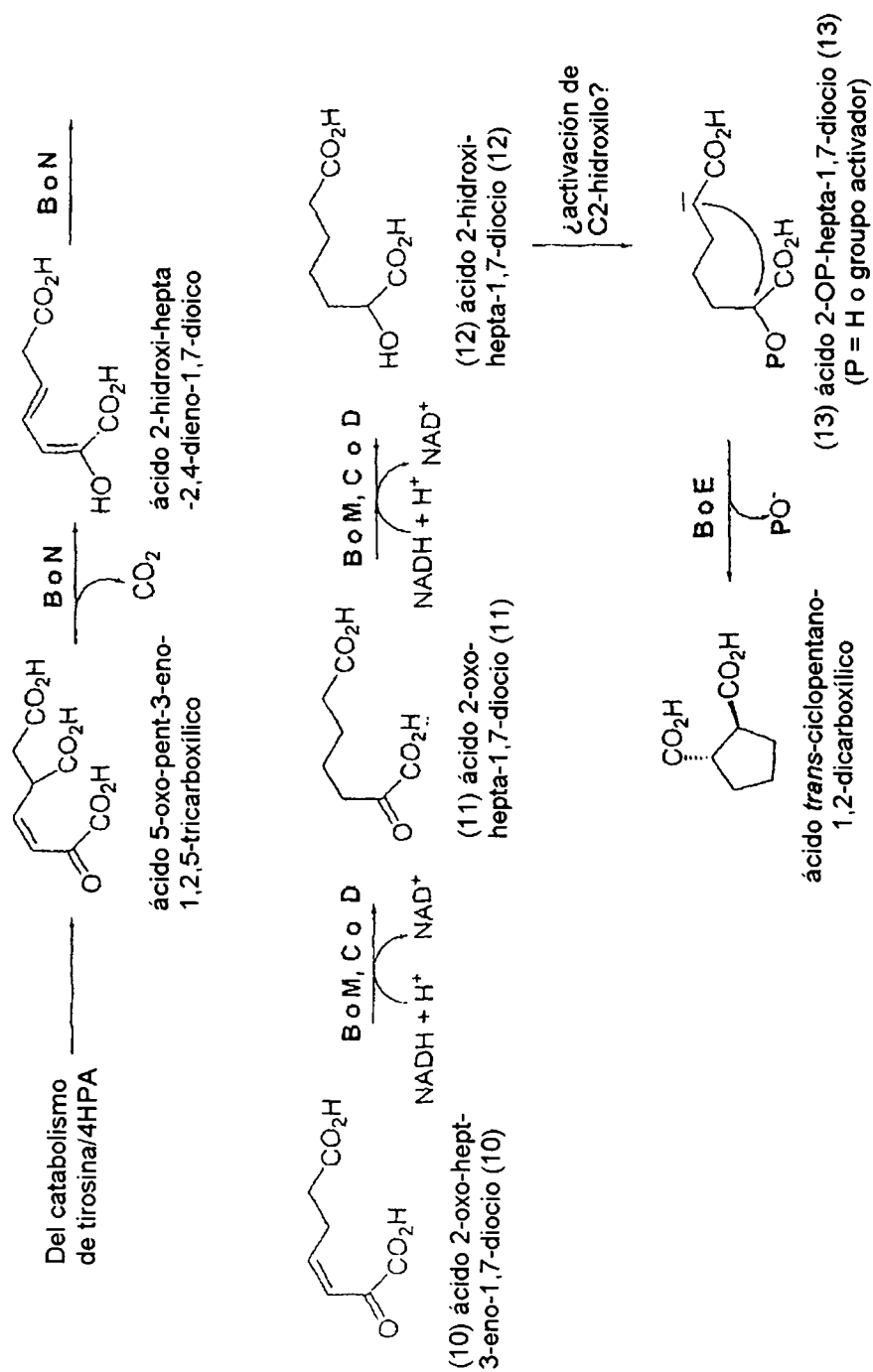


Figure 4

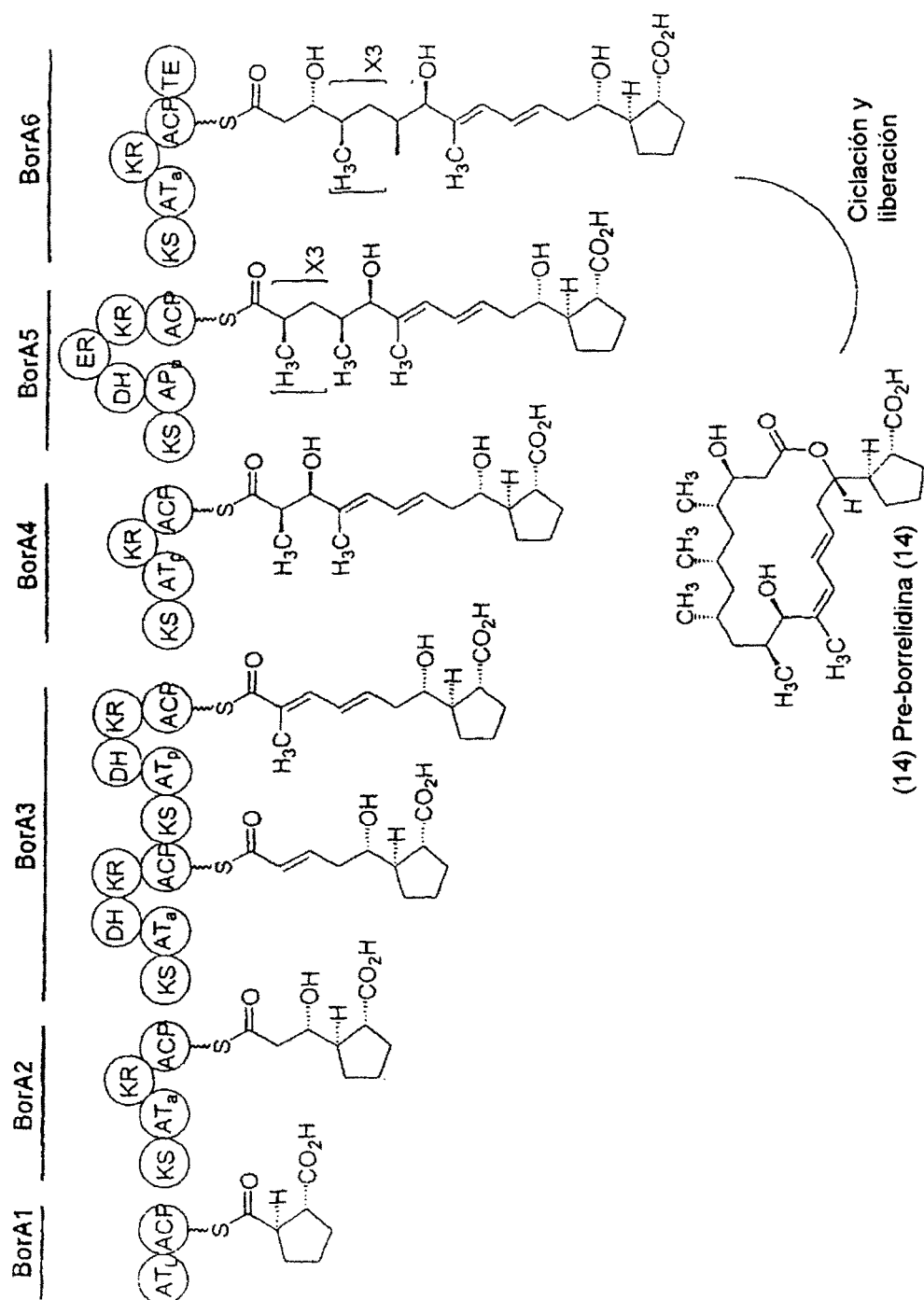


Figure 5

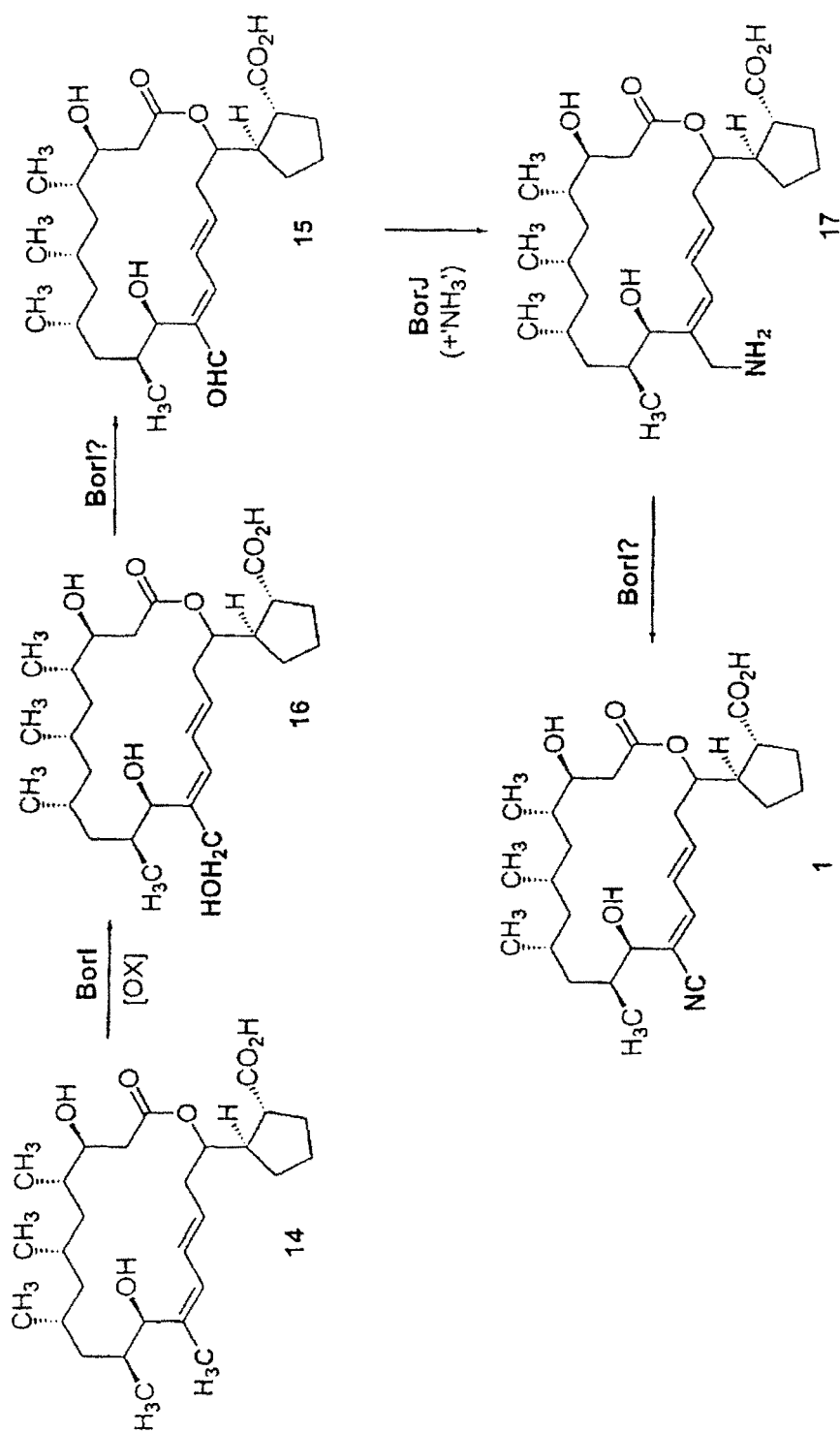
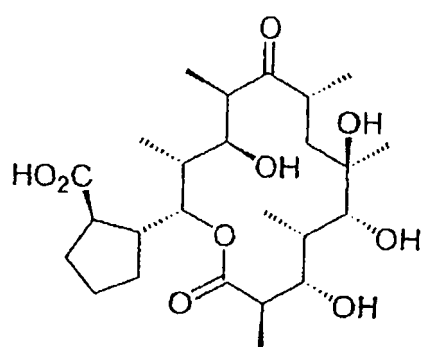


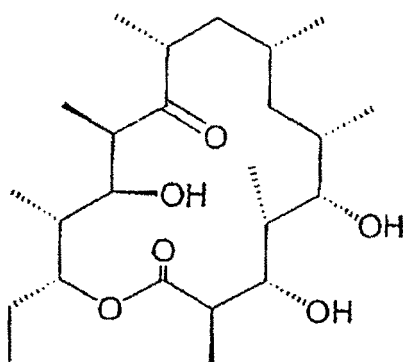
Figure 6

Figure 7

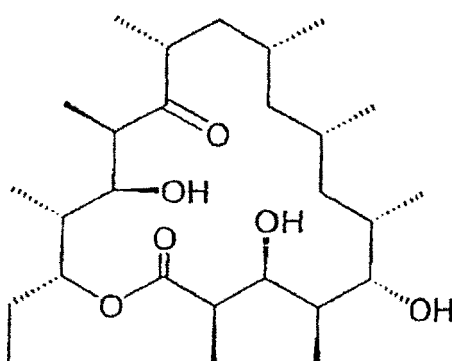


6

Figure 8



7



8

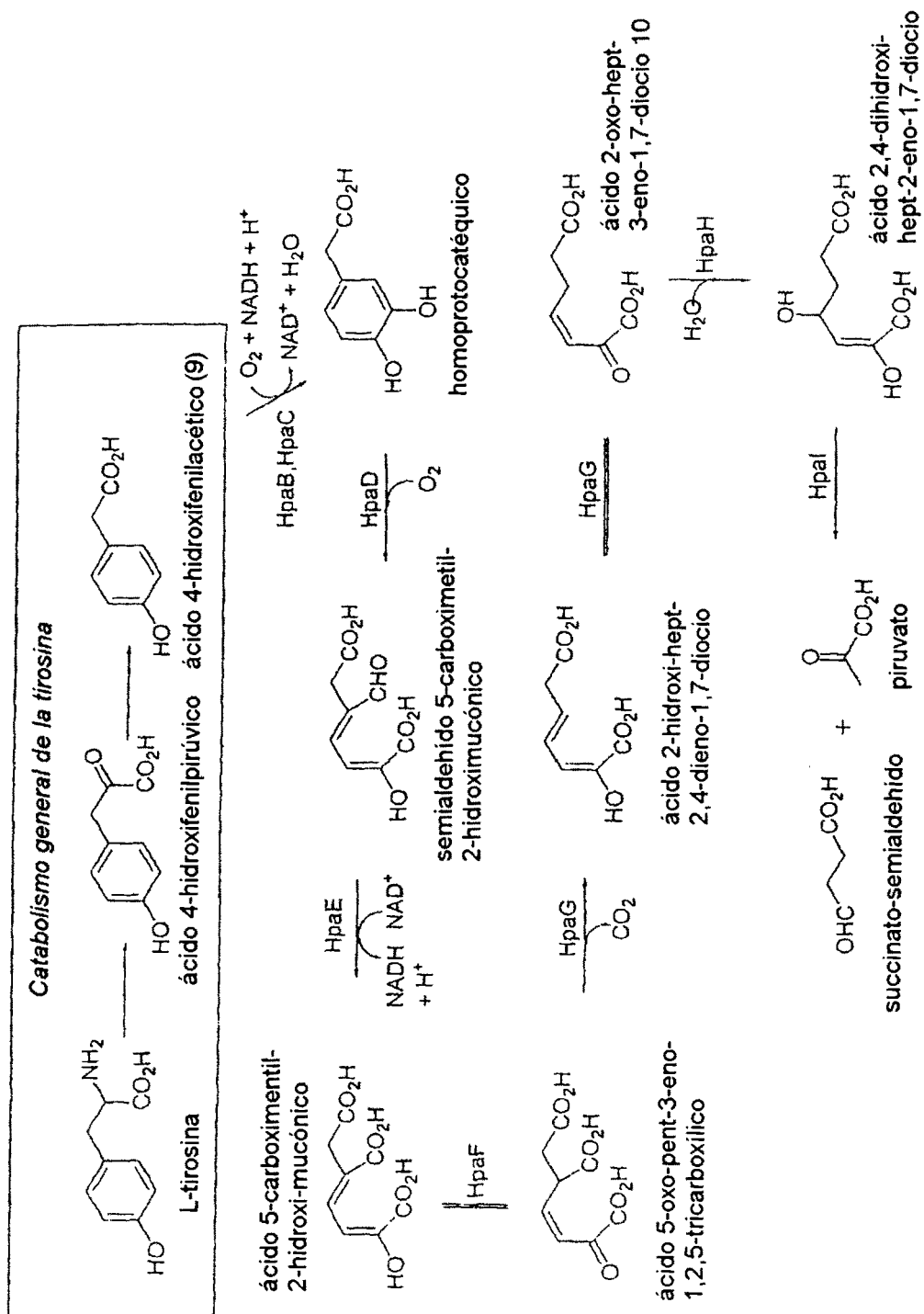
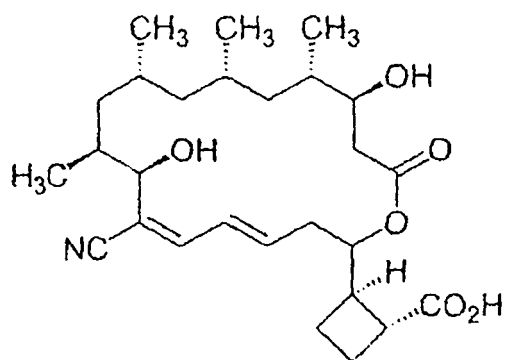
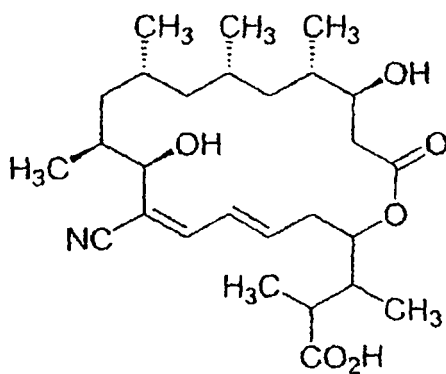


Figure 9

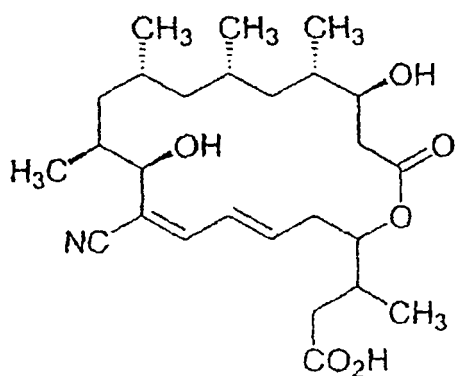
Figure 10



18: $m/z = 474.3$ ([M-H]⁻)



19: $m/z = 476.3$ ([M-H]⁻)



20: $m/z = 462.3$ ([M-H]⁻)

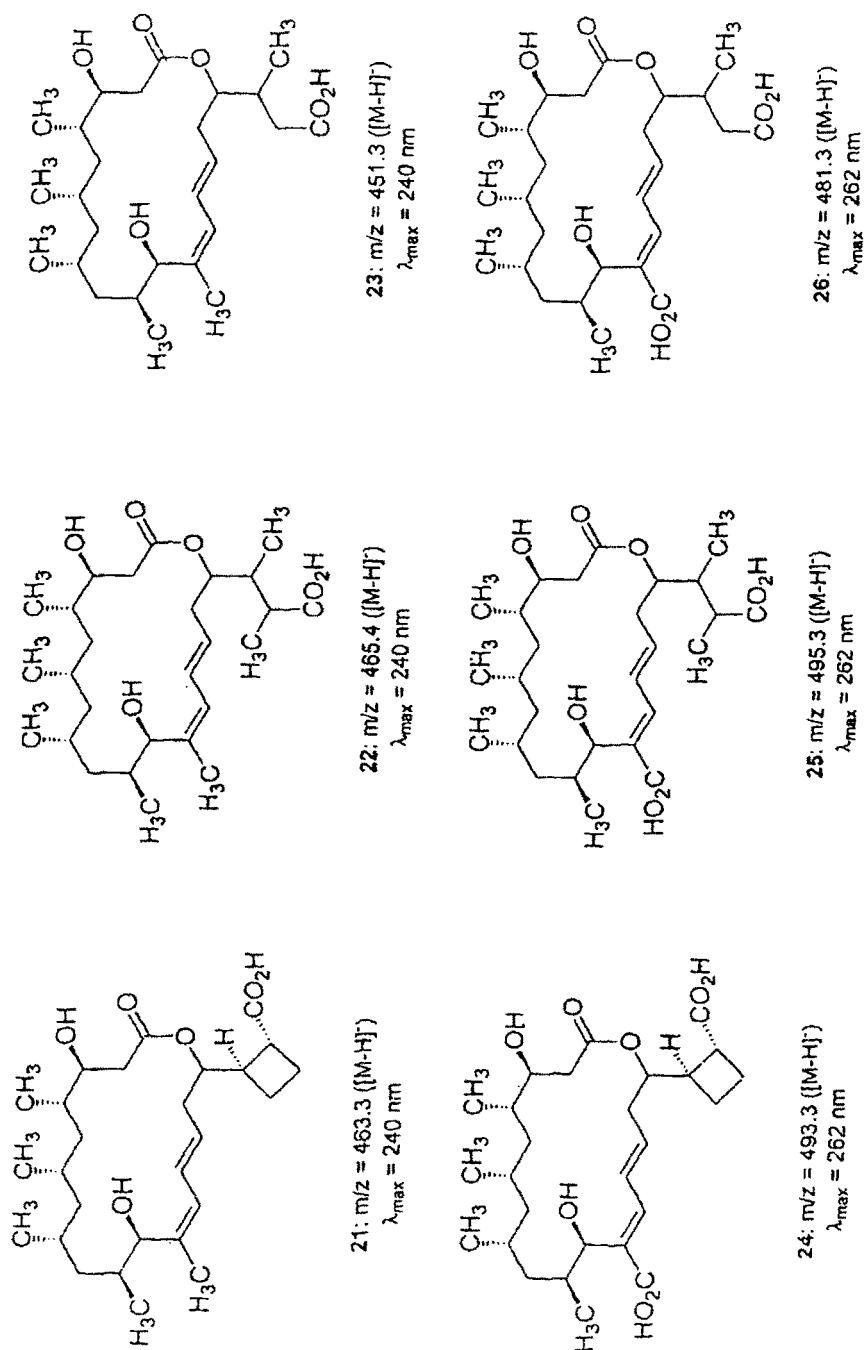


Figure 11

ES 2 287 552 T3

LISTA DE SECUENCIAS

<110> Brana, Alfredo F
 Biotica Technology Ltd
 5 Leadlay, Peter F
 Martin, Christine J
 Méndez, Carmen
 Moss, Steven
 Olano, Carlos
 10 Oliynyk, Marko
 Salas, Jose A.
 Sanchez, Cesar
 Wilkinson, Barrie
 University of Oviedo
 15

<120> BORRELIDIN PRODUCING POLYKETIDE SYNTHASE AND ITS USES

<130> POLEQUÉTIDO-SINTASA PRODUCTURA DE BORRELIDINA Y SUS USOS
 20

<150> GB0230217.2
 <151> 2002-12-27

25 <160> 113

<210> 1
 <211> 74787
 30 <212> DNA
 <213> *Streptomyces parvulus* Tü4055

35 <400> 1

gateccgcgc ggcacgcgc tcgacgtgct gcgggcgggc gacgcgtggc cccacagcgc	60
ggcaccgcgc caccggggac tcctcaacgc ctggtggggc gcctgggtct gggccacggt	120
40 cttcgaccgc tacgcgtcga ggacctacga cgacgccag gacgtcgacg cgatccacga	180
cgcggcggga ctggtcatgg ccggtgccgg attcgacatc ctgcgcgcgc tgcctcgcat	240
cctcttcgtg cgccggctga ccgcgcgaca gcacgcgaag gccctcgcgg ggcccacccc	300
45 gccgacgcac tgagcgccc gcacccgtga tcccgccccg cgatccccgg gcccgataaa	360
tgcgttgccc ccggcgcgcg cctgtggtgg gatgagcggc gacgggggcg gctccccggc	420
50 gtgcacccctt ctaccccttc tgcaaagatc ccgcgcgccc actctccgcc cccgttcttc	480
cgccccgagc cgtcgcgcgc gtggaggctt tcctgttgct cgccgcgcgc tccgtactgc	540
tgcgcgcgtga ccagagcgtc tacgtgacct cggggtccga gccggacggt ccgcccaggg	600
55 ccgcactgcg ccggctcgag gccgaactgc tcggccgcgg ccacgcgcgc tccgcgcgcg	660
tgcacgcggt cctcgccctc ttggactccg aggaactggc ggccgcccac gtacgcctcg	720
tcggactcgt cgacgacctg ctgggtcccg accgcaccca caccgcgcgc ttccgcgcgc	780
60 tccgcgcgcac cgtgccgcgc gacaccgagg cgctgtacgt ggaccgcgcgc ttgccttcc	840
tgcctgcagca gcccgagcag cctgcgtgc tgtgcgcgga ggccgcgcgc gtcctgcgcg	900
65 tgtcacccctg cgccgacctg gtctgcgcgc tgtgctggga ccgctccgac tacgcgggat	960

ES 2 287 552 T3

gcccgcctgtg ccaccgcagg atcgacgggg acgacccctt cctgcgtccg gtccgtgccg 1020
 tcggcgccgc cagggcgacc gtaccggggc cgctgcgact gctgcgcctg ggcaccgaca 1080
 5 tgaccgccga cggcaccacg gcggtggacg ccctgctggc ccgcgcgacc ccgctctccc 1140
 cgcaggaccg ggacgacctg ctcaccctgt tgcgctcac accggccggc cggggcgacc 1200
 10 tgccgcagga catcccggtc cgcgagacca aggcgctggt cctgggcgcg ctggtgcgcc 1260
 gggcaccgtc gcggccggcc ctgcggaggc tgcctgcga gcggctcacc accgccaccg 1320
 acgtgctgcg gctgctcgcc gtgctctcgg gcggcgacgc cgggctggtg acaccggcac 1380
 15 gggtcacgaa cgttccccgt tccctgcggc gtgacctget cgcgctctc gacggactgc 1440
 cggcgccgta cctggctgag gacatgctgc ggcacccac ggctgggaag cggggcgcgg 1500
 aggtgctgca ccccttcgag gggcacaccc ggcacccgcg cgcgcgctc gccaccgccc 1560
 20 tgctgcgcgc cacaccgttg gaccgggaca ccgccttcgg cgcgcctctg ctgaccacgg 1620
 ccgcgcgca ccgggacgcc gtgcgcccgg acggcacccg agtcgcgccg gccacctggg 1680
 25 cgggacggct ggagcaggcg atggccgagg gggacgcgc tcgggcccgc gccctcgccc 1740
 gggagcggcc cggcgaactg gtgcgcccgc tggacgtgtt gctgcgcctg cacaccgacg 1800
 aggcgctcgt gccggagctg gagaaggccc tgcggcacgg gctgcggaag gtgggcccgg 1860
 30 gcccgctget gtcggcgctc ggggcgctgc ggacacgcac cgaggaccgc accgggaccc 1920
 ggcgcgtgtt ctcccgcg ggcgacgtca cccgggccct gtccgtccc gagcggcgcc 1980
 35 ccgccttgc cgcggggccg gtgtccgagg tggctgcctt gctggagggg gaactgctgc 2040
 gccggttcgc cgcggggcg ccctacgagc tgtcgggtgtt ggaacccgga ctgaccgacc 2100
 tcaccgtgcc gttcaccgag cggaccgcgc ccaaggccct ggtgaccgtg ggccgcggca 2160
 40 gcgtccaggc actccccgag ggctccgtgc tccgactgtt cctgcactgg accgaacccc 2220
 ggggcaaccg cacogacctg gacctgtccg tgccttctt cgacgcgag tggacgttca 2280
 ccggcctgtg cgactacacg aacctggtgc acggtccgga cgcggcgatc cactccggcg 2340
 45 acctcacgtc gggcccgcg ccgcgcggcg ccaccgagta cgtggacctc gacctggagc 2400
 ggctggcgcg gcggggagac acctacgccg tcccgtggt gttcagctac aacaacgtcc 2460
 50 cgttcgagga actgccggac gccttcgccg ggttcattggc gctgcccgcg gaaggcccgc 2520
 gcgacgcgac ctacgacccg cgcacctgc ggcagcgctt cgacctcgcg ggcgactcca 2580
 aggtgtgctt gccgatgatc gtggacctgg ccgcgcggcg ggcgttgtgg accgacaccc 2640
 55 acctgccgtc cgcggggcg ttcagagca tcggttcgca cggcgggcgt gagctggccg 2700
 cgggtggccg tgacctctgg cagcagttca cctcgggcgg ccgggcgacc ctgtgggacc 2760
 60 tcgcgcctt gcgggcggcc gccctctcgc cggagggtgg ggtggtgtcc cgggagccgg 2820
 agcccgcggt gctgcgttac cggcgggcg cggccgagag cgaggcccg ttcgccgtcc 2880

65

ES 2 287 552 T3

gagtcgcgtc ccacaaggac gccgaggaac ggctggcgca caccgacccc gaetcggccg 2940

cgcccgggct cgccgccggc cggcgggtct tectcgcgac ggtccacggt gacgtccggc 3000

5 cgccgggggc gtcgggcacg tectaccggc tcttccccgg ggccggggac gcctcaccga 3060

ccctgacccg cgtgaccgcc ggggacctgc tcgccgagct gggctgagcc aggcgccggc 3120

10 ccgcgccggc ccgcgccggc ccgtccctgc ccgtgccgga gggctcgccg gtcactccgg 3180

ccaggcggag ttctcgatga cctcgacgaa gtccgtacgc cggaagccgg gcgcgaagtg 3240

ctccagcaca tccgcgttca ccgtgccgaa ggtggtcgcg ggccgggtgct cgaacccctc 3300

15 ggtgaacgcc cgcaggatct gcttcttgaa gtccggggcg ggatgcgcgg cggtagccgc 3360

gtcgatctgg gcccggtgga gattgccag ccgcaggccg agcacgtcg tctccacgcc 3420

ggcgggtggtc gccgcgatct cgggggcat ccggtacggc acctccggag tgggtgtcag 3480

20 ggcgacggcc gtccacacgg tgtccgcgtc ggccctgggg atgccgtgg cgagcaggaa 3540

ggcgtgggcc tggtcggcac cgtccatctc gaagcgtgg tcgtaccgc ggtagggcg 3600

25 caccaggccg gtgtcgtgga agagcgcggc gatgtacagc agctccgggt cggggcggtat 3660

gcccaggcg gcggcctgga ggctgccgaa gaggtacaca cggcgtgagt ggtggaagat 3720

cagcggcgga gtggtgtcgc ggatcaggtc ggtcgccctc cgcgccggcg cgtgtcggg 3780

30 aatctcgatg ccggcgatct gctcggccat ggctgccctc cggggaatcg gtgccgtcgt 3840

tgtgcctcc acctccgcc cggcgcgacc cggcgctccg ctaccgatg gccgacaacc 3900

35 ccttacaagc ggccatgtgc cccgcgcgc cgcctcagcc gccgtccggg cgcgggccgg 3960

cgtccggcac ggtggtggcg aagcgtgcc ggtagcgggt gggcgacagc cccagatgac 4020

ggggaaggc ccggcgagg ctctcgtagc tggggaaacc cgacagcgcg gcggcctcg 4080

40 tggcgttgtg cccggagtcg agcagcgct tggcgatgtc gaaacggatc agctccacgt 4140

acttcacggg cgtgacgtcc agctcgccc ggaacatccg ggtcagatgc cgggggtgga 4200

cccgcacgcg cgccgccaac gccgccagac tgtggtcggc ggccggatcg gcctgtacgg 4260

45 cgtcctggac ctgccgcagc acgggctcc gcggcgccgg gcccgcgaac gaggcggaga 4320

actgcgactg gccgccggcc cgtgcagggt acaccaccag cgagcgcgcg acctgcggg 4380

50 cgagatcggg cccgtggtcc tctccagca gcgcgaggc caggtcgatg cccgccgtca 4440

cgccggcgga cgtgtaggtc gcccgtct tgcgaagat cgcgtcggg tccacgcgtg 4500

tgcagggaca gcggcgggcc agcgcggtg tgtgtgccca gtgcgtcgtc gcccgctcgc 4560

55 cctccagcag acccgcgga cccagcacga aggcgcgggt gcacaccgag gcgacgcgtc 4620

cggcccgggc cgccagcgcc ttcgcgcggt cgatgagccg tgggtcgacg ggcgagccgg 4680

60 gcagcgcgtc accgcgacg acgacgagcg tgtccggcg gccggcgga cgcgcgtccg 4740

cctcgccggg gaccagcagg ccgatggacg aacgcaccgg cggcccgctc ggggagacga 4800

65

ES 2 287 552 T3

cgccgagccg gtaccggggc ccgaaccggt tggcctccgc gaagacctcc gccggccccg 4860
acaggtcgag catcttcatg ccgtcgaaga ccaggatgcc cacgctgtgc gctctcgccg 4920
5 tcatgtctcc ctctccggcg gccggggggc ccctgcgcgc cattgtcccg ccggccgtcc 4980
acgccggcgg ccggcggcgc gggcgggcgg cggtcggaat gaggcgcgcc ggacatcggc 5040
10 gtaggggtggc gagcgtgtgt tcggcccgcg tcccggagac cgcggaacgc aggacctttg 5100
gcaggcacgc ggaaggacag cgatgggtac ggtcaccacc tccgacggca cgagcatctt 5160
ctacaaggac tggggcccg cgcacgcccc gccgatcgtc ttccaccacg gctggccgct 5220
15 caccgcggac gactgggaca accagatgct gttcttctc tcgcacggct accgtgtgat 5280
cgcccaagac cggcgcgcc acggccgctc gggccagccc tcgacggggc acgagatgga 5340
cacctacgcc gccgacgtcg cggcgctgac cgaagcgctc gacctgcggg acgccgtcca 5400
20 catcgggcat tcgaccggcg gcggcgaggt cgcgcgctat gtggcgcgcg ccgaaccggg 5460
ccgggtcgcc aaggccgtgc tggtcggcgc cgtgccgccg gtgatggtca agtcgacgc 5520
25 caaccccgcc ggcaccccca tcgaggtctt cgacgggttc cgcacggccc tggccgcca 5580
ccggggcccag ttctacatcg acgtgccctc cggccccttc tacggattca accgggaggg 5640
cgcgaaggtc tcccagggcc tgatcgacaa ctggtggcgg cagggcatgt cgggcgcggc 5700
30 caacgcccac tacgagtga tcaaggcgtt ctccgagacc gacttcaccg aggacctcaa 5760
ggccatcgac gtgcgggtgc tggtcgcgca cggcaccgac gaccaggtcg tgcctacgc 5820
35 ggactcggcg ccgctgtcgg tgaagctcct gaagaacggc acctcaagt cgtacgaagg 5880
gctccgcac ggcattgctt ccacccaccc cgaggtggtc aaccccgacc tcttgactt 5940
cgtgaggtcc tagtcggcgc tcacgcggc gacacgggag cgggtgcggc gccgcgcacc 6000
40 ggggtgcttg tcaggacgga gaccgggtt aaggcggtga tgcgtatgc caccagatc 6060
acggcggaga cctcgtcgtc cgacaggacg ccccggtgct gcgcgtaggc ggcgctctgc 6120
45 gcggcgggcgt ccgcgggacg ggtggtcgcc tccgcgaggg cgagcgccgc ccgctccga 6180
gcggtgaaca gctcgggtgc ccgccaggcg ggcagcaccg ccaggcgctg ggtcgtctcg 6240
ccggcccga gcgcgcctt ggtgtgcaga ctgagacagt aggcacaggc attgagttg 6300
50 gagacgcgga tgttcaccag ttccacgagg aggcggtcca ggccggccgc cgcggcgggc 6360
tcccgacccg attccgcgc ggcacgaac gctttgtac cgcgggggt ctgcttgctg 6420
acgaagacc gccgctcgtc cgtcgccacc ggggectgct gtgtcacgtg gtctccttcg 6480
55 tcgcgtctc ttccggcggt tctatcatc accccatgg atgttgaaag tgaaacttcc 6540
aggtcggggc cgacggggg cgcgtggtga gcaacacgga gacacggccc gcggagatgc 6600
60 ggtgcggcgc cctcgaagac gaggtgccc cgcggggcgt cgaagtcctc accgcccgtg 6660
acgtccctt cggcgggccc cgcgccatga cgtgcggcg cacgctgcc cagcggggcc 6720

65

ES 2 287 552 T3

	ggacgctgat cggagcctgg tgettcgccc accactacgg tcccagacgac gtggccgctg	6780
	cgggcggcat ggacgtcgcc ccgcacccgc acatcgccct gcagacggtc agctggctgt	6840
5	tcagcgggga gatcgagcac cgggacagcc tcggcaccca cgccttcgtc aggcccgccg	6900
	aactcaacct gatgaccggc ggcttcggca tcgcccactc cgagggtctcg acccccgaca	6960
10	ccactgtcct gcacggcgtc cagctctggg tggcgctgcc ggaggagcac cgcgacaccg	7020
	gccgcgactt ccagcaccac gcaccccgcc cggtcgcctt cgacggcgcc acggcacgcg	7080
	tcttcctcgg ctcgctcgcc ggggacacct cgccttgag cacttcacg ccgctgctgg	7140
15	gcgccgagtt gacgctggtg ccggcgggca ccgccacct ggacgtcgac ccgggcttcg	7200
	agcacggcgt cctcgctgac agcggtgacg tacgctcgca gggcgccgtc gtgcgaccgg	7260
	ccgaactggg ctacgtcgcg ccgggtcgcg cgacgtgac cctgaccaac gagtggcccg	7320
20	caccgcgccg gctcactcct ctcgggggcc ccccgctccc cgaggagatc atcatgtggt	7380
	ggaacttcat cggccggtcg cagcagcaga tcgtgcgggc ccgcgaggac tggatgaagg	7440
25	gcgaccgctt cggcgagggtg cagggctacg acggggcacc cctgcccgcg ccggaactgc	7500
	cgaacgcacc cttgaagccg cgacgaaggg cgcgctgac tgccggggaca tgggttgga	7560
	ccaagggttt cggcgctgct cgatcaccga acccaccgcg agtcactctc gggtgagtcc	7620
30	cgaacggctg ccgggagcgc gtgagcacgt gcgcagatgc tcggcgatga tgccgagaat	7680
	cgcaccccg tgcctcagca ggtagaagtg accaccgcg aaggtgtcga gtgtgaacgg	7740
	gccgtccgtg tgttcggacc atgcccgggc ctcgaccggg gtgaccatcg ggtcatcctc	7800
35	cccgttcaag gcatggatgg ggacccgag ctccggggcc ggtcggtagc ggtaggtctc	7860
	ggcgccctg tagtcgccgc ggatggcggg gagggccata cgcaccagct cctcgctgtg	7920
40	gaagacctgc tcgcgggtgc cgtcgagggt cctcagctcg gccaccaact cctcgctccga	7980
	caggaggtgc accgtacccc ccgtcctctg ccgggacggg gcgggcccgg ccgagacgag	8040
	gagtgcctcc agggagatgc ccgcactctc gaaccgtcgg gccagttcga aggcgagggg	8100
45	ggcgcccatg ctgtgtccga acagcgcgac cggctggtga acacggggcc gcagcacggg	8160
	gaagagctgg ttcgcgagtt cgtcgatgtc ctccaggggc ttctccgcgc gccggtcctg	8220
	ccggccgggg tactggaccg cgagcacgtc gcaccggggg gccagcgcgg cagccacggg	8280
50	gtggtagaac gtcgcggagc cgcggcgctg cggcagacag atcaactggg gtgccgtggg	8340
	atgtgcgggg cgggtactgc tgateccacac gtcgctgtgg gtgttcgtac cggtcacgag	8400
55	cgggtgctgc ctccggcggt ggcggtgggt cgggggatgg ccgatccggc cgtgacgcct	8460
	ccgtcgaccc cgagggtctg cccggtgacg taggcggcga gggggctgag cagccagacg	8520
	accgcgttgg ccacctcctc gcaacttgcgc agggcgccca gcggagcccg gcgcgcctgt	8580
60	tgtgcgaggg cgctgggacg ggcgtaacag ctgcgcagca tgggggtgtc ggtcgaaccg	8640

ES 2 287 552 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

gggctgacca cgttgacgcg gatgccgtcg cccgcgtact gcagggccac cgacttgetc 8700
 aggccgatga ccgcgtgctt ggtggccgag tagagcgggc tctgggcgtg gccgatgtgc 8760
 ccggccactg atgcgcagtt cacgatcgcg ccgcgcgcg ccgtcagcat ggccctcgatc 8820
 tgtccgcgca tgcacgacca gacccacgc aggttggttg cgatcacgcg gtcgaagttg 8880
 tcggcggtgt cctggtgcag cggaccgaac gagccgaagg tcccggcgtt gttgaacgcc 8940
 ccgtccagcc gtccgaaccg gctcacccgc cgggccacgc agtccgccac ctgcttgctc 9000
 tcaccgacgt cgcagggcac caccaagtgg tgcgaggagg gtagtcgggc ggttgctctc 9060
 gtgagggccg actcgggtgcg gccaccagg acgacgcgtg ctccgtgccc caccgaggagc 9120
 cgggcgcgcg ccgcggccgat gccgtgccc gctccggtga ccatcatcac gcggtcggtg 9180
 agttccagac tcatcggtgt tccaacgctc cgtccctgct cgteggatgt gcgatccgct 9240
 gtgtcatatg tgcagtccgc cgttgacgtc gacgaccgtg ccggtggtgt atccggcgtc 9300
 ctgcgcgcac agatggcaga ccatgccgcg ggccctggcg acgtgccga agcggccggc 9360
 ggggatgttg ctgacgcggt cggcggtcca ctgcgggggc ttgtcctccc aagcccggcg 9420
 gatgcgctcg gtgccgatga cgcgtgggc gaccgcgtt accgtcacgc cgtgcggggc 9480
 cagttcgtag gcgcactgct tggtaaccc gatgacgcg gccttggcg cgacgtaggc 9540
 ggcatgtctg aaccgggtgt acgtgcgacc ggccacggac gccaggttga cgaccctgcc 9600
 ccaccccgcc gcgaccatcg ccgggacgca cagccgggtc atggtgaaca cgtcgcagc 9660
 gttgtgcgtg acggcctcct ggaggtcggc ctcggtcagt tcggtcaccg agcggggccc 9720
 ggtgtcgcca ccgacgcgct tgaccaggac gcccgggccg tctgcgggg cgagcgagtc 9780
 gacggcgga cccagggcgt gagggtcggc cagtcggcg accagcgggt cccggggcag 9840
 ccggtcgccg agcccgctcg cgaccgggtg caccgctcg gcgtccttgt cgagcaggac 9900
 caccgcagg ccccgggccg ccaggcctcg ggcgaacctc gcgcagatgc cgtgccccgc 9960
 tccggtcacc agagccacgt cgtgtcgtgc cgtcatgtgt tctccgcca gccgcgcgcg 10020
 gattccaggg tggccgcgca ccgggtgtgt cgcagtagtt cttcggtcgg ttcatgccc 10080
 gtgcccggac cggtaagggt ttcgaccggc tgcagtgacc ggtcgacggt gaacgcgggc 10140
 gtggtcagtg gcacggggaa ccactcgtcg gcccgggccc cctcgaccgt ctgccacagg 10200
 tcccatgcgg tggccagggt gcgcccggcc gccacagcg gcccacctc ggccacgtga 10260
 acgcccagct ggcaaccgac gccagctcg tcggcgcgca gtgccaggcg tgccgcggcg 10320
 aggaaccgc cgcacttcga cagccgtacg ttgatgtggc tggcgggccc gctggtggcg 10380
 gcggcggtga ggtcgggccg tccggtacag gactcgtcga gcatgacggg cagaccggtg 10440
 gccgcgcgca gccggcccaa ctcgggccag gaccgcggcg ggagcggctc ctccaccac 10500
 cccacgcgct ccagttcgcc cgcgaccttc tccgcttct cggccgtcca ggcgcggtg 10560

ES 2 287 552 T3

acgtccagtg agacacgggt gtcggcgggg aggcggtcct gggccgcgt cagccgggcc 10620

accgcccccg ccgggtccgc cacccttgatc ttcacgtgcc gcaacgccgc cagcgcccg 10680

ggcgtgagtg cgtccaggac ggtcgcgacg tcgcgcgaga ggtggatcac gaggtgacg 10740

gacgtcggtc cgtcccgccg tgatcgggca ggcggggcca ggaccgcag gacgtcggcg 10800

agcgcccggg cgaaatgccg gcacaccgcg tcgagcaggg cgatctccac ggcggccgc 10860

gccgacgagc cgtcgacgag ccgggtcagc ggcagctgtg cgatcgaggc gacggcgctc 10920

tcgaagtccc gccactcgat gcgctcggcc agtcccccg gatcgaggc ctggacggct 10980

cgcaccgcgc cgtccagggt ctacccgggt acgtagtcgc ggggcgctcc ctctcccat 11040

ccgcgggtgc ccgccagctc gatctcgacc agcagggacg ccgcgctgcg acgggagcgc 11100

gtggcggtgt cgaaggccgc ggccatgggc acgacggcgg tgtgcagccg tacgcgacgg 11160

atcacgcttc ctcttcagc cgcgtggcca gccagtcca gtacgccgtc cgggcccagc 11220

tgaactccac gtagtgcca tccgtggcga agacctctc gtgcacggct gacgtcagac 11280

gccgcagcat cgctcgcgcg gccgacaggt cgatgatcgg gtcgtgagtg gggagcgcca 11340

ggtcgacggg gagccgggtg cggggggcac cgcgggcata gtggtcctcc aggtgcacga 11400

gcgtcgctg cgtggccgag gtgacctgc gcagcatcag gtgatccccg gtgaggaaact 11460

cccggtagcg cggcaggtcg gtgtagtcgc cgctggcgag cccacgggc cgtagcccg 11520

tgccggtgag cgcgcggcg tcggcgagcg tgtcccggt gtggcgcgcc cgctgctgtc 11580

ccagcgcggg cgcgcacagg accaacctgc ggacgggcag atcgcggtg caccagagcg 11640

cggccagcac gctgccgccc aggtctctgc ccagggcgac cggcccggca ccgccgacct 11700

cggccgtcac ggcgtcgagg gcgcgggct agtcgtcgag gacgagatcg gccgacggca 11760

ggtggcccg agggccctcg ctgcggccc agcctctgcg gtccagggcg tagacgtcga 11820

tgccgctgc gttgagctcg ggcccgtct cgaacagcca gccgcgtgg ctctggatgc 11880

cgtggaggta gaagacggc gaggtggcg cggcggtgt ccagtgggtg agggtagacc 11940

cgggtccgct ggacgggtc agcatgctcg tgggtggcat gggctgcctc ctctagaccg 12000

gacgagattg acgtcgggg ccagccggac gtcgacgagg agcggtcctt cgagcgtcga 12060

caacaggctc ccgacggcgt cgagttcttc cgccttgcg aggtgaggg cacgtgcgcc 12120

catcgccgtg gccagcccgg cgaggtcggg ccaggcgaac gccgagtac cggggtcgta 12180

gccgtggttc ctgagtttgt agtgctcggc tccgtacgcc ccgtcgttga gtacgaccac 12240

gacgagcggc agccggtacc gtaccgccgt cgtgaactcc gacaggtgca tcatgaagcc 12300

gccgtcccc acggcggcga ccacgggccc gccgggtccc gccgtcgcc cgccgatcgc 12360

cccggcgacg ccgagcccga tcgagccgaa gccgcccag acggtgaagt gcagcgggtc 12420

cgccacgcgc agatacggcc agacaccac gtcgaagcgg ccgatatcgc tgacgacact 12480

ES 2 287 552 T3

gcgctcggcg ggcagtatcc ggtccagccg gatcatggcc gtccggatgt cgacgggtctc 12540

cgctccactg cggtcgtcga cgtcgtcctg cggcgagaaac ccggccagtt gcccggcgac 12600

5 gcgctccgcc caggcgccgt tggccgcggt gactccggcc tgatccagca ggacgttcat 12660

ggtctcggcc gtgcggcggg catccccggc cacgggctcg tcgacggggc tgtacgagcc 12720

10 gaaccgtgcc ggatcgggtg ccacgtgcac gactctcttg ccgaggagca gctcgcggtt 12780

gagcacggtc cacatgttca ggctcgcccc gaacgcgata acgcagtcg actcggcgat 12840

gaccgtgctc gccacgctgt gcgcgagcga gccgaagatg ccgacgtcgc ggggggtgac 12900

15 ggcgaacatc tccttgccga gcacgggtgt ggccagcgt gctccggtac ggtccgccag 12960

ctccaccagg gcctctcgcg caccggcgac ggccgcaccg tgcccggcga ggaccagcgg 13020

20 ccgcttgccc gagccgatca gcccagcgc gccgtccagc gcctcggcct ccggagcggc 13080

cagaggaccc ggcgccaccg ggagcgtgac cggcgcttgc tcgcccgcct ccgctgcat 13140

gaggtcgata ggcacattga gtacgacggg ccgccgctcg gccacgatcc gctggacggc 13200

25 ccggttcagg tcgcgacga gcgagggcgg tctgtggacg cgttcgtacc ccgcgccgc 13260

cgcggccgcg accgtcgcga tgtcgaagtg gtggaagtgc gtgggcaccg gtggcggatc 13320

acctgtgatc agcaggacct ggctgtggct acgagccgct tccacaagag gggtaagggc 13380

30 gttggtgaaa gccggcccggt gcgtcacgga cgcacaccg atgccgcgc acatacgtgc 13440

gcggccgctg gccatggcga cggcgccgc ctctgtggcg accgccacga accgtccgcc 13500

35 cgcgtcggcg aaggcgggca gatagagcag attggcgttg cccatgagac cgaagacggt 13560

atcgacgccg tgtgcgtca gagcgtcggc gagcgcgtgg aaaaccttca ttgctgtccc 13620

tcggtcgggg cgggctggag ccagacggga tcgttctggt cgaccggcgc gcagggtggg 13680

40 ggcgggtcct cgcggagcag ggaacgcagg tggttgccg tgatgagcac ggtggttccg 13740

gcgcagtcct gggcgacggc catccgccg gcgagaactt gggggtgccc caccgcgcgg 13800

45 gtcaggctgt gcagcggggc cgcgggca cgcagcgctg ccgccgcgtc cagtacggcg 13860

gcgacggggc cctcgccccg cggtggcggg ccttcgcgt cgacgaggac cgtgccgtcc 13920

gcccgccgga cgagggtggg ggcccgtgcg ggtggggcg tgaaccagcg gtcctgggtg 13980

50 agccagacag cgctgtcgaa cagcgcgacg tcgycggcgc atccgtgctc ggtgcgcaga 14040

cggacgtagg tggacacgac ggcggcggca gcggccaggt aggcaccgag gacgtcggcg 14100

ctggacactg gagtccctag cccggcgcgc gggccgcga cgaggcgcat gatgcccgac 14160

55 tcggcctgga tcacggtgtc gacgtgcgg tcggcggcg ccaggccatg gccggtgacc 14220

gtgcagtga cgacgccgtg ccgggacagg atctgatcgg gggcgaggcc ctgcgcggtg 14280

60 agtgtgtcgg ccgcgaggtt ggtgagcag atgtcgcac cggcgagcag ccgctcgaaac 14340

ccggcccggt cctcggcgtc ggcgaggtcg agccgacagg agcgtttgcc cgcgttgttg 14400

65

ES 2 287 552 T3

acgtagtaga ggtagccgac cccggccacc tgctgggcca gccgccggga cccttcgccg 14460
 tgccggcggt ccaccttcag tacgtcggcg ccgagttggg ccagcagccg gcccgcggtgc 14520
 5 ggteccggcg tgtagcgagc gacctccagc agggcgagcg cgcgcagcgg aggagtgccg 14580
 cgcgcgatcg gctcccagag gccgcccctgc cggggcatcg gaccgtcggg gacggcgggg 14640
 10 atgagggaac gcagtgggct gcccggtgta ccggacgggt cggtgaccag gccgcgacgc 14700
 cgggcggcgg ctccgtcgcg tacctcctcg gggggcgcga cctgggcgca cgggatgccg 14760
 gcggcccgca gggcggtcac cacgtccacg gcccgctgcc cggcggtoca cttgccgagg 14820
 15 atctcgtcga gctcgtcggc gttgcggacg cgggcggcgg tgcggcgaa gcgagggtcg 14880
 tcggggagggt ccgctcgtcc caggactgcg gtcagcctgt gccatatcgg ctgcgccatc 14940
 gtgcagatga cgacaggggc gtccctggcac gtgtagctgt tccaggggtgc cgccatgccg 15000
 20 tgtcggttgc cggtaocgac tggcggaaga ccggcgagcg cgacgctcgg aagcaagggtg 15060
 cccgtcagag tgaacaggct gtcgaactcg gcgatgtcca ggtagtcccc gcctcctccg 15120
 25 cgctcgcggc cgatgagccc ggccacgacg gcgatcagac cggacagggc cggcgtagcg 15180
 gacgccaggc ccaccacgga gaggaccgat ggttccccct cggtgccggt cggcgaggtc 15240
 agggccgcca gcgctgcaa ggtecgctcg gtggccgggg cgtcacgcag cgggcgggtg 15300
 30 agcccgaaac cgctcagccg caccgcgacc aactcggggc tgcgatcggg aagttccggc 15360
 gcccagagcc ccagagcggc gagccgctca tggccctcgg cgtcgcaac cagcacatcg 15420
 35 gccgtctgca gcagacgca tgccctggcc caaccggacg ccgactgcgc ggcggagtgc 15480
 agccaccgct cgaacggggc cccgtgatcg ggtgatcggc agagcgtcac gacacgtgcc 15540
 ccgaggtccg cgagcagctc cccagcagc ggggctgggt tactgcggcc ggccatgagt 15600
 40 acggcgatgc cctcagcgg ccctgccctt gtcattggaat tctccctcgc tccgcgcacc 15660
 gatgcggggc tcggtcgtca ccgctgattg gtcgtggacg tcggccgtga ggcgaccgcc 15720
 agggaaatca caccggcgcc gccgcgatcc gcgggggata tggccggcag tcccgatgcg 15780
 45 ccattaaagc gcgcatgatt cgttcctgtc cgaccgtacg accgagacgg cggaaaatca 15840
 tcgcacaccc ctgctccgga tccggaaacc ctgctcaggg ggcaaggggg aggggggtccg 15900
 50 taatggccaa aacgaaattt tacggagctt tacgtttgct ggacgatcta ttggtgagcg 15960
 cctcgacggg ctggacatgg cagtagtgaa tgtecgcat catggctatt agtaccgtga 16020
 ccctgatcac acgagccctg gttgacgggt gaaatttggg gctggcagag tgatgacgag 16080
 55 ctcccgctcg caaagtgggt gaataactgt tccgaaatct tcggcaattc aaaggagact 16140
 tacgggggat gcctctatta atgtattgct gtaggcgaa ataatgacag gcagtgtctgt 16200
 60 ttccggccca ttctgcagc ctccgaacc cgtctcaggg cactccgaac ggaaaagcga 16260
 tcccgctctt ctgctcggcg ccggacgcgg tgcccgcatg gcggatgccg tacgtgccgc 16320

65

ES 2 287 552 T3

cggegetcag gcgggcatcg acccgggcgt cctacggcgc acccgggcca ccttgatcac 16380
 cgcggggagc gcgggagccg caggccggt cgccgcgc ccgcgcctga ccggcgccac 16440
 5 gatctctctg gacaccgcg agacaccac actgctcgcc ctgcacctcg ccgccaagc 16500
 gctgcgggcg ggcgacacct cttacgccgt cgtcgggtgc gaacttccc acgggaactg 16560
 10 cggttgatc ctggccaggc agtcagcggc aaccgcgag ggggtgtgc ccaggcgat 16620
 cgtccgacc accacggcg accgcaccac cagcgcgat cagccctg cgccgacga 16680
 ccacggcagc ccggccctg aagccccga tgcaccgc acgttgtccc caggcatcac 16740
 15 ccaggccccc gccgagggt tcccgggcct gctggcgacc ctgcacgacg acacacctt 16800
 ggcggccacc gcggtcaccg agcagggcag cgacggcacc accgtcctcg tcctcctga 16860
 20 ccagcccgag gacggcgac ccggcgacc gctccctgg gtggtctcgg cccccacac 16920
 ccgcgcctc cgggccacgg ccgggacct ggccgtcac ctgcacacca caccggcgc 16980
 acccgccgac gtgcgcaca cctgctcac cgcgcgccc gaccggcacc gtgcccggt 17040
 25 cgtcggcgcg gaccgggcca cctcaccga cggactgcgc gactcgcca ccggaggcga 17100
 cgcgccccc ctcgtccacg gcacggcac cggatcgcc cgtccgtct tcgtctccc 17160
 cgccagggg tcgcagtggc ccggtatggc cgccgaact ctgaaacca gcgagccctt 17220
 30 tcacgacagc gtgcacgctt gcgcgacgc gctggccgag ttcgtcgact ggtcggtctt 17280
 cgacgtcctg cgccaggcac cggacgcgc accctgcgc cgggtggacg ttctccagc 17340
 35 caccctgtgg gcgacgatg tctccctggc cgaggtctgg cgtcgtacg gcgtggaacc 17400
 ggccgcgctc gtcggccact gctacggcga gatcgccgc gcgcaggtag ccggcgccct 17460
 cgacatgctg .gacggcgccc gactgctgc ccaccgcagc cgggcctggc tgcgactggt 17520
 40 gggcaagggc acggtcatct ccgtcgccac ctggggacag gacatcacc ggcgcatggc 17580
 ggccggccc gactcgtcg aactggcgc gctcaacggc ccgcgctccg tggcgctcgc 17640
 45 agggccgccc gacgtcctgg acggcatcgt caacgacctg accgaccagg gcatccagc 17700
 caaacgcac cccggcggtg acaccgtcg ccaactgtcc caggctcagg tcctccgcga 17760
 ccacctgctg gacgtcctgc gccgggtct gccccggccc gccgcgctgc cgttctactc 17820
 50 caccgtcgac ggaaccgaac gcgacaccac cagctggac accgactact ggtacctca 17880
 caccgcagc caggtcgct tccaccaggc cgtgcggaac ctgctcgccg ccggacaccg 17940
 55 ctggttcgtc gaggtgagc cgcaaccgct gctcggagc tccatcgagg acaccgcgc 18000
 cgagttcggc ctcgacgac tggccgctg cggaacctg cgtcgaggc agggcggcac 18060
 ccgcggggtc ctgaactcg tggcgaggc gtatgtccac ggcacgaca tcgactcac 18120
 60 gccgccttc accggcacga cccccaccg catcgacct ccgaccgtc aggaccacgg 18180
 catcgagggt cagcgcgac agggcggcga gacatggacc gaccgctca gaacctccc 18240

65

ES 2 287 552 T3

ggacgagcag cgcgaagagg ctttgcctgga cctcgtgtgc cgcaccgtcg ccgcggtgct 18300
 cgaagcggac ccggccggca cggcggacgc cgtcgcccc gacacggcgt tcaaggagat 18360
 5 gggcctcggc tcaactgagcg cggtcgggct gcgcaacggc ctccgcgagg ccaccggcgc 18420
 ccacctgccg gccaccatcg cctacgacca ccccaacccg gccgctctgg ccgccaacct 18480
 10 ggcgatgacc ctgttcgacg cgacggggcg cgcgccggcg gtcccgccac cgagcccgca 18540
 cgacgaaccg atcgacgccg agaccgctgt gctgaccgcg ctggaacggg ccgacgaggc 18600
 gctggaacgg ttgcggggcc cgcacgccc cgcgcccg caggagaccg gccggcggtat 18660
 15 cgacgagctg ctgcggtccc tgaccgacaa ggccaggcgg atgagacagg ccgacgccgt 18720
 cgatgatgtc gatgatccgg ccaccgaccg gttcgccgca gccaccgacg acgagatgtt 18780
 20 cgaactcctc gagaacagtt tcggcatctc ctgaggcgcg ccgacctccc gcaactgcag 18840
 tcgcttcccc caccgatccc gaaggcggca accgatggca catgaagaca aactgcgcca 18900
 cctcctcaag cgtgtcagtg ctgaactcga cgacaccag cgcgggtgc gtgagatgga 18960
 25 ggagagcgag cgcgagccga tcgcgatcgt ggggatgagc tgccgtctgc ccggcggggt 19020
 gaacagcccg ggggagttct ggtcgtgct ggaggccggg acggacgccg tctcggagtt 19080
 cccgcgggac cgtggctggg atgtggagaa cctctacgac ccggaccgg acgcccccg 19140
 30 gcggtcgtac gtccgcgagg gcggattcct ggacggggcc ggacagttcg acgcccctt 19200
 cttcggaaac tcgccccgtg aggcgtggtc gatggatccg cagcagcggc tgctgctgga 19260
 35 gtgctcgtgg gaggcgatcg agcggtcgcg gatcgaccgg aagacctgc acggcagccg 19320
 gaccggcgtc ttgcggggt ccaactggca ggactacaac accctgttgc tgaacgccga 19380
 ggagcgctcc cagagctacc tggccaccgg cgcctccgga agcgtgctgt ccgggcgcgt 19440
 40 ctctacacg ctgggcatgg aaggcccgcg gatcacgtg aacacggcgt gctcgtctc 19500
 tctggtcgcc gtccacctgg cggcccgttc cctgcgggcg ggggagtgcg acctcgccct 19560
 45 ggccggcgcc gtcacgggtc tgtccacacc gcagcttccg gtgccttct ccgggcagcg 19620
 cggactcgcc cctgacggtc gtcgaaagc cttcgcggtt tcggccgacg gcatgggctt 19680
 cggcgagggg gtggcggtgc ttgtgctgga gcggttgctg gtggcgcggc ggaacgggtc 19740
 50 tcgggtgttg gcgggtgtgc ggggttcggc ggtgaaccag gacggtgctg cgaacgggtc 19800
 gacggcgccc aacggtccgt cgcagcagcg ggtgatacgt gcggcggttg cgagtgcgg 19860
 55 gctgggtccg gccgatgtgg atgtggtgga ggcgcacgg acggggacgc ggttggtgga 19920
 tccgatcgag gcgcagggct tgctggcgac gtacggggcg gccggggacg cggagcgctc 19980
 gttgtggtg gggtcgggtga agtcgaacat cggatcatgc caggctgctg ccggtgtcgc 20040
 60 cgggtgcatc aagatggtgc tggccatgga gaaggccgt ctccctcgga cgctgcatgt 20100
 ggatgagccg tcgggtgagg tggactggga ctcgggtgcg gtgcggctgc tgaccgaggc 20160

65

ES 2 287 552 T3

gcgggactgg ccgtcggagg aaggtcgtct gcggcggggc ggtgtgtcgt cgttcgggat 20220
 ctcaggcacc aacgcgcacg tgatcatcga ggaagcaccg gaagaggggg aggaaccgga 20280
 5 gtcgcacgcg ggtggtgttg tgccgtgggt gctctccgcg cggacggaag gggcactgca 20340
 agcacaggcg gtgcaactga gcgagttcgt cggcgagtcg agtcgggtgg atgtgggttg 20400
 10 gtcgttggtt tcgacgcgtg cggcgttcga gcacggggcc gtggtggtgg ggcgcgggcg 20460
 ggacgagttg gtgcggggct tgtccgaggt cgcgcagggt cggggcgtga ggggtgtcgc 20520
 gtcttcggcg tcgggtggtc tcgcgtttgt ttttgcgtgt cagggcagtc agcggttggg 20580
 15 gatggggcgg gggttgtatg agcggttccc ggtgtttgcc gaggcgttcg acgaggtgtg 20640
 tgggcgggtc ggtccggggg tcggggaggt tgttttcggt tcggatgcgg gtgagttgga 20700
 20 ccggacggtg tgggcgcagg cggggttgtt cgcgttgag gtggcgtgtt ttcggttgtt 20760
 ggagtcctgg ggtgtgcggc cgggttgtct gatcgggcat tcggtcgtg agttgtcggc 20820
 ggcgtgtgtg gcggggttgt ggtcgttga ggatgcgtgt cgggtcgtgg ctgccgggc 20880
 25 gcggttgatg caggcgttgc cggcgggtgg ggtgatggtc gcggttcggg ccgagggcgg 20940
 ggagctggcc ggtttcctcg gtgaggacgt ggtgatcgcg tcggtgaacg cgcgggggca 21000
 ggtggtgac gctggtcctg aggggggtgt ggagcgtgtg gtggctgctt gtggggcgcg 21060
 30 gtcgcgtcgt ctggcgtct cgcgtctt tcattcgcct ttggtggagc cgatgcttg 21120
 ggagttcgt cgggttgtg agtcggtggc gttegggtgt cgcgtcttc ggggtgttct 21180
 35 caatgtcacg ggtgcgtggg tggatccgga ggagtgggg acgccggagt actgggtgcg 21240
 tcagggtcgt gagccggtgc gtttcgccga cggggtcgc accgttgcctg acgcgggtgt 21300
 gaggacgttc gtcgagctgg gtccgcgcg ggcgctcact tcgatggtca gccactgcgc 21360
 40 ggacgccacc gccacttcg tgacggtgt acctacctg cgcgccgatc acgatgagtc 21420
 gcggaccgtg ttgagtgcg cagcgtcctt gtacgtccag ggtcaccgg tcgactgggc 21480
 45 cccgtgttc ccgcggggcc gcacggtgga cctgccacc tacccttcc agcaccagca 21540
 ctactggctc gacgtacctc ctctgttcac cgcctcctcg gcggcccagg acggtggctg 21600
 gcgataccgc atccactggc ggcggtcgc cagagggac tccggggacc ggctctccgg 21660
 50 ccgctggttg ctgctggtgc ccgagtcgga cgggacggag ccctgggtgg agggggccga 21720
 gaagatgctg gccgagcgc ggtgcgaagt cgtccacgtg ccgatcgcg cgacggccga 21780
 55 ccgggacgcg atggtcggag ccgtgcgtga gagcgtcag gacggtcgg tcgacggtgt 21840
 gctcagcctg ctggcgtcg acggccgcc gcacccgat gcggctgcgg tgccgacagg 21900
 gttggtcgc accgcgcagg ttgtgcaggt cagtgcagag ctgggcatcg gccgctgtg 21960
 60 ggtcggcacc cgacaggcgg tctccgtcga cggggccgat gaggctgac gggccggtag 22020
 gaccaggaag gccgacgacc ccgccgatgt cgcgcaggcc gctgtgtgg ggtcggccg 22080

65

ES 2 287 552 T3

5 ggtagccgcg ctggagaagc ctcggttggtg gggcggcctc gtcgacctgc ccgcacgtgc 22140
 cgacgaacgg atgcgggacc tgggtggtca ggcctcacc gctcccgacg ccgaggacca 22200
 acttgccgtg cgggcccgcg gcatcgccgt tcgcccactg gtacgctccg ccgcgtcggc 22260
 cccggccgcg gactggcagc cgagcggcac cgtgctggtc accggcggca ccggaggcgt 22320
 10 cggagccaac gtggcgcggtt ggctggtcac ccaggacatc cagcaactgt tgetgggtcag 22380
 ccggcgcggc ccggacgccc ccggagccgc tgagctgctg gccgaactca gcgcctcagg 22440
 aacgtccgtg accatcgagc cctgcgacgt caccgacgcg gacgcggtag ggcgcctgat 22500
 15 cggcgccgta ccggccgaac ggccgctgag cacggtcgtc cacgcgcggg gcgtactgga 22560
 cgactgcttg atcgacgccc tgaccccgca ggcctcgcg gccgcactgg aggtcaaggc 22620
 caagggcgca ctgaacctcc acgagggcgc cggggaagcc cacttggtgc tcttcctc 22680
 20 gctggccgga acaaccggaa ccaagggaca gggcaactac gccgcgcaa acgcctatct 22740
 cgacgctctg gccgaacggc ggctggtga cggcctgcgc gccacttcgg tcgcctgggg 22800
 25 cgcctggcag ggcgcgggca tgggtggcga cgcgcgcgta gccaccgca cgcgcggtta 22860
 tggcctcccg ctcatgagcc ccgaccgcgc cgtcgccacc ctgcggcagg tcatggccga 22920
 gccggtggcc acgcaggtgg tggcggaagt cgactggcag cgattcgtcg ccgacttcac 22980
 30 cgcggtgcgc cccagccgcc tcctcgccga cctgccgaa gtgcgctccc tgggcgagca 23040
 gcgaaaggac ggcccgggcg gtcagggcga ggaggacggc ttggccagca agctggcagc 23100
 35 cctgcccgaa gccgaccgcc gacgagccgt gctggacctc gtggaggaac tcgtcctcgg 23160
 ggttctgggc cacgagacgc gcgcggcgat cggcccggac agttccttcc acgccatcgg 23220
 cttcgactcg ctaccgccg tcgaactgcg caacctgctg accgtacgcc tcgggatgaa 23280
 40 gctgcccgcg accctcgtct acgatcaccg gaccctgctg tcgctggccg accacctgca 23340
 cgagcaactg gttatcgacg gcacccccat gacggacacc gcggccgacc tgctcgccga 23400
 actcgacgca ctgcggcgga gactcgccgc cgtcgggctg gaaccggagg cgcgcgcccg 23460
 45 catcggacgc aggtcgaagg acatgcagac cgcctgcgaa ccaggtcgg agtcctcacg 23520
 cgacctgaag tccgcctcac gcaccgaagt gctcgacttc ctcaccaacg aactcggcat 23580
 50 ctcccgctga ccagttgacc gaccgcgacg aacggcgcac ctggctgcgg ctgctccacg 23640
 ccgaccttcg acctgcccg acgcccccg gacgggacta ccaccatgcc caacgacgaa 23700
 gaactcctcg actacctgaa gcggactgcc tcgaacctcc aggaggcgcg gcagcgggtg 23760
 55 cacgaactgg aggagagcga gcgcgagccg atcgcgatcg tgggatgag ctgccgtctg 23820
 cccggcgggg tgaacagccc ggaagagttc tggctcgtgc tggaggccgg gacggacgcc 23880
 60 gtctcgaggt tcccgcggga ccgtggctgg gacgtggagc ggctgtacga cccggacccg 23940
 gacgcccccg gcaagtctga cgtgcgggaa ggcggattcc tcgacggcgc gggccgggtc 24000

65

ES 2 287 552 T3

gaccccgcggt tcttcgggtat ctccccgcgg gaggccgtgg tcatggatcc gcagcagcgg 24060
 ctgctgctgg agtgctcgtg ggaggcgatc gagcggtcgc ggatcgaccc gaagaccctg 24120
 5 caccgcagcc gcgcggggtg gttcgtgggc tgaacggcc aggactacgg gacgcttctc 24180
 ctgctgccc acgaccgctc ccacgcctac ctgccacgg gcgcctccgc gagcgtgctc 24240
 10 tccggccgca tctctacac gctcggactg gagggccctg cggtcacgat cagtacggcc 24300
 tgctcgtcct cactgggtgc cctccacctg gcggcccgcg cctgcgggc gggggagtgc 24360
 gagctggcgc tcgccggcg tgtgacggtc atgccagca cccgcctgtt cgaggctctc 24420
 15 tcccggcagc gtggcctggc cggtgacggc cgctgcaagg ccttcgggc cggggccgac 24480
 ggcactggct ggggcgaggc cgtgggcgta ctgctcctgg agcgggtgtc ggtggcgcg 24540
 cggaacggtc atcgggtgtt ggcgggtgtg cgggggttcg cggtgacca ggacgggtgc 24600
 20 tcgaacggtc tgacggcgcc gaacgggtccg tcgcagcagc gggtgatccg cgcggccttg 24660
 gccagtgcac gcctggcccc cgaggacgtg gacgcgtag aggcacacgg caccggggacc 24720
 25 tccctggcg acccgatcga ggcgcaggcg ttgctggcga cgtacggcg gggccgggac 24780
 gcggagcgtc cgttgtggct ggggtcggtg aagtcgaaca tcggtcacgc gcaggccgt 24840
 gccgggtgtc ccgggtgtat caagatggc aaggcgatgc aggcgggcac gctgccccg 24900
 30 acgctgcctg tggatgagc gtcgggtgag gtggactggg actcgggtgc ggtgcggctg 24960
 ctgaccgagc cgcgggactg gccgtcggag gaaggctgtc tcggcgggc cgtgtgtctg 25020
 35 tcgttcggga tctccggcac caacgcgcac gtgattctcg aggagccgc ggcggaggac 25080
 gcggtaaccg agcctgaagc ggggtgatgtg gtgccgtggg ttctttcggc gcggtcggct 25140
 gaggcgttgc gggagcaggc tgcccggctg gcgtcggtg ctggtgggtt gaacgtgggtg 25200
 40 gatgtgggtt ggtcgttggc ttgcacgcgt gcggcgctcg agcaccgggc cgtagtgggtg 25260
 gggcgggagc ggggaagagc gctcgggggt ctgttcgtg tggctgggg acgccgggt 25320
 gcgaacgtgg tgacggggcc cgtcagctcc ggtcggcccg cctttgtttt tgctggtcag 25380
 45 ggcagtgcag ggttggggat gggcggggg ttgtatgagc ggttcccggt gtttgccgag 25440
 gcgttcgacg aggtgtgtgg gcgggtcggc ccgggggtgc gggaggttgt tttcgggttcg 25500
 50 gatgcgggtg agttggaccg gacgggtgtg gcgcaggcgg ggttggtcgc gttggagggtg 25560
 gcgtgttttc ggttgttga gtcctgggg gtgcggccgg gttgtctgat cgggcattcg 25620
 gtcggtgagt tgctgcggcc gtgtgtggcg gggttgtgtt cgttgaggga tgctgtcgg 25680
 55 gtcgtggctg cccgggcgcg gttgatgcag gcgttgccg cgggtgggg gatggtcgcg 25740
 gtccgggccc aggcggggga gctggccggt ttctcggtg aggcagtggt gatcgcgtcg 25800
 60 gtgaacgcgc cggggcagggt ggtgatcgtt ggtcctgagg ggggtgtgga gcgtgtgggtg 25860
 gctgcttgtg gggcgcggtc gcgtcgtctg gcgggtctgc atgcttttca ttcgcctttg 25920

ES 2 287 552 T3

gtggagccga tgcttgggga gttccgtcgg gttgtggagt cgggtggcgtt cgggtgtgccg 25980
 tcgttgccggg tggtttccaa tgtcacgggt gcgtgggtgg atccggagga gtgggggacg 26040
 5 ccggagtact ggggtgcgtca ggtccgtgag ccgggtgcgtt tcgccgacgg ggtcgccacg 26100
 ttgctcgacg cgggtgtgag gacgttcgtc gagctgggtc ccgctgggac gtcacttcg 26160
 10 atggtcagcc actgcgcgga cgcaccgcc acttcggtga cggctgtacc taccttgcg 26220
 cccgatcacg atgagtcgag gaccgtgttg agtgccgag cgtccttgta cgtccaggg 26280
 caccgggtcg actgggcccc gctgttcccg cgggcccga cggtggaact gccacctac 26340
 15 cccttcacg accagcacta ctggatgatg aacaccgaa gtgccgcca gccggcggag 26400
 ctggggctcg gcgatgccg tcatccgtg ctgggttcg tcgtcacgt cgcgggggac 26460
 gacaaggctg tcttcgccgg gcggtggtg ctgctcacac acccctggct ggcgaccac 26520
 20 accgtgctcg acgcggtctt gctgccgct acggccttc tcgaactggc cgtgcgcgc 26580
 ggtgaggagg tgagctgtcc ggtcgtacac gacctgacg tgcaccgacc gctggtcgta 26640
 25 cccgagcggg gcgccgtgca ggtacagatg gctgtggcg caccggaagc cgtgggcca 26700
 cgtgaggctc ggggtgtact ccgccccgac gacgacgagc agcacgagtg gacgtgcac 26760
 gccgctggac tgctggcgtc ggccgccacg gcggagccc cgtggcggc cggtgccctg 26820
 30 ccgccgccgg aggcgcaggc cgtggacctc gacggcttct accccggact cgcgagcac 26880
 ggctaccact acggcccgt gttccagggc gtccgggccc cgtggcggct gggcgacgac 26940
 gttctcgccg agatcgtgct gcccgaggcg gccggcgccc accccgccc gtacggcatg 27000
 35 catccggccc tgctcgacg cgtcctgcac gcggcacggc tgggcgcctt ccgtgagcgg 27060
 tcggaggaga agtacctgcc gttcgcctgg gaaggcgtga ccctgcgtac caggggagcg 27120
 40 accgccgtac gtgctcgaat ctcccgggcc ggtaccgacg ccatccggct ggacgtcacc 27180
 gacaccgcgg accggccgggt cctcacggcc gaatcgctca cgtgcgacc ggtctccgc 27240
 ggtcagctca tggccgtccc gcgcgactca ctgttccggg tcgactgggt ttccgcgcc 27300
 45 gccgcgaacg gtcccggcct gcggtggtc cgtgcccca ccgtggaggc ggccctcgcg 27360
 gcggacgcgg acatcgtggt cgtgccatgc ctgcagagt agggtcgca tcaggcgacg 27420
 50 taccaggcac tggagctgct acagcgtgg ctggccgcc acaccgtac caccacgctc 27480
 gccctgctca cccaccgtgc cgtggcggtc ggcgacgacg tccacgacct ccaccacgcg 27540
 cctctgtggg gcctggctcg caccgccag accgaacacc ccggtgctt ccggtcgtc 27600
 55 gactcgacg accccgaccc gacgacggac gtcttgccc cggcgctcg caccggsgaa 27660
 cccaggtcg cgtaccgtga cggcgccgtc ctggccccgc ggtgaccgc ggcctccgcg 27720
 60 ccgcgggagc cggccgagtg ggacgccgag ggaacagtcc tcatcaccg cggatcgggc 27780
 gccctcgag ggatcgtggc ccagcaacct gtcgcactc accgctacg ccgactcgtc 27840

65

ES 2 287 552 T3

5 ctcgcgagcc gcagcggcag gccgcacccg ggggccgacc tgctcgacgc cgacgtcacg 27900
 gccgtgtcct gcgacgtctc cgaccgcgac gccgtggccg cgctgctcgc ctccgtgccg 27960
 10 gacgaacacc cgctcaccgc cgtcgtgcac accgcaggcg tactggacga cggcgctcctg 28020
 cacgccctca cgaccgagcg catcgacacc tcgttcgcgg cgaaggctcga cggcgcccgt 28080
 catctccacg aactcacctc ccacctggat ctcaccgcgt tcgtgctgtt ctccctccgcg 28140
 tcggccgtgc tgggcgcgcg cggacagggc aactacgccg cggccaacgc ctacctcgac 28200
 gcgctcgccg cccaccgtcg cagcaacgac ctgccgcgcg tgtctctcgc gtgggggctg 28260
 15 tgggccgagc acgagggcat ggcccgcgga ctcggtgacg ccgagctgac gcgtatttcc 28320
 cggatcgcgg tcaccgcgct gagcgcggag gacggcatgc ggctgttcga cgcgggatgc 28380
 gccggcgatc agtcacagct cgtgccgatg cgggtggaca ccgcggcgct gcgcgcacgg 28440
 20 cgtgaccacc ttcccgaccc gatgtggagc ctggtccccg agcggacccg agcggcacgt 28500
 acacagcctg ccgcctcgct tcgggacagg ctgccgaac tgaccgcccc cgaacgcaag 28560
 cgcacgggtc tcaacctggt gcgcaacgcg gtcgccgaca cactcggcca caacgccgcc 28620
 gacggagtac cgcccgacca gagcctcgac gccgcgggt tcgactcgct caccgccgtc 28680
 gagttccgca accggctctc cgccgtcacc gacctgcgc tggccgccac cctcacctac 28740
 30 gatcacccca ccccgcggcg catcgccgag cacatcctga cccgcctcac cctgctgaag 28800
 gagaccgcg ccccgccgt cggcacccgc ccggttgccg cgcgcaccga agacgatgcg 28860
 atcgtcatcg tgggcatggc gggccgcttc cctggcgccg tcgcgcacacc cgaaggctctt 28920
 tgggacctcg tccactccgg cagggacgcc atctcggagt ggcccaccga ccgcggctgg 28980
 gacgtggaga acctctacga cccggacccc gacgcgctcg gcaagtccta cgtacggcac 29040
 40 ggcggattec tgcacgacgt cgccggcttc gacgcgggct tcttcgggat ctccgcgcgt 29100
 gaggcgctgg cgatggaccc gcagcagcgg ctctgctgg agtgctcgta cgaggccctg 29160
 gagcggggcg gcacgcaccc ggccacgctc agaggcagcc ggtcgggct gtacgccgga 29220
 45 gtgatgtacc acgagtacgc ctcccgctg ggccgcacgc ccgcaggctt cgaaggcaca 29280
 ctccggaccc gaagctcggg cagcatcgcc tcggggcgca tctctacac attcgacctc 29340
 accggggccg cggtcaccgt cgacaccgca tgttcacact cctcgtagg cctgcacctg 29400
 gccgtgcagg ctctgcgggc cggtgagtgc gaactggccc tcgccggcg cgtcaccgtc 29460
 atgcacacgc cgcgcccctt cgtcgagttc tccgcacagc gggcctggc cgcggacggc 29520
 55 cggagcaagg ctttcgcggc ctccgcgcac ggggtggcct gggccgaagg cgcgggaatc 29580
 ctgctcctgg agcggctgtc ggccgcgcgg cggaacggtc atcgggtgtt ggcggtggtg 29640
 60 cggggttcgg cggatgaacca ggacgggtcg tcgaacggtc tgacggcgcc gaacgggtccg 29700
 tcgcagcagc gggatgatac tgcggccttg gcgagtgcg ggctgggtcc ggccgatgtg 29760

65

ES 2 287 552 T3

gatgtcgtcg aggccacacg caccggcacg gccctcggcg atccgatcga ggccgagggc 29820
 ttgctggcga cgtacgggCG gggcggtgac gcggatcgtc cgttgtggct ggggtcgggtg 29880
 5 aagtcgaaca tcggtcacac gcaggcggcc gcgggtgttg caagcgtgat caagatgggtg 29940
 caggcgatgc aggcgggCGt gctgcccgcg acgctgcatg tggacgagcc gtcgggtgag 30000
 10 gtggattggg actcgggtgc ggtgcggctg ctgaccgagg cgcgcgagtg gccgtcgggg 30060
 gagggcgctg tgcggcgggc ggggtgtgtcg tcgttcggga tctccgggac gaacgcgcac 30120
 gtgatccttg aggagccgcc ggcgaggagc gcgctgccgg agcctgaagc ggggtgatgtg 30180
 15 gtgccgtggg ttctttcggc gcggtcggca gaggcgttgc gggagcaggc tgcccggctg 30240
 gcgtcgggtg ctggtgggtt gaacgtggtg gatgtgggtt ggtcgttggc ttgcacgcgt 30300
 gcggcgttcg agcaccgggc cgtcgtcgtg ggaggcgatc gggaagagct cctggggaag 30360
 20 ctttctcctg ttccgggggt cgaggtcggg gtcggggctg gtgccgggtg tgggtgtggtg 30420
 ttggtgttct ccggtcaggg gtgtcagtgg gtcggtatgg ggccggagtt gctgggttcc 30480
 25 tcgctggtgt tcgcggagtc gatgcgggag tgcgcggcgg ctctgtcgcc gtttgtggac 30540
 ttttctgtgg tggatgttct ggggtcggct ggggagttgg gtcgggtcga ggtggttcag 30600
 cctgcgttgt gggcggtgat ggtgtcgtg gcgcgggtgt ggccgtcgtg ggggtgtccg 30660
 30 gttgctgcgg tgggtgggtca ttccaggggt gagattgccg cggcgacgggt ggccgggtgcg 30720
 ttgagtgtgg gtgatgcggc gcgggtggtg gcgttcggga gccgtttgat cgcggagcgt 30780
 35 ctgtcggggc tgggtgggat ggtttcgtg gcgttgtcgc gtgagcgggt ggtgtcgttg 30840
 atcgcgggtg tgcgggtgt gtcgggtggc gcggtgaaag gttcttcgtc gacggtggtc 30900
 tcgggtgagg ccgcggggct ggagaggggt ctggcccgct gtgtgtcgtc gggggttcgg 30960
 40 gcgcgtcgta tcgatgtgga ttacgcctcg catcgggtgc aggtggagtt gatccgtgag 31020
 gagttgttgg ggggtctgga cgggatcgtc ccgcgtcgg gtgagattcc gttcgtgtcc 31080
 acggtgacgg gtgagcggat cgacaactgtc gagctggggg cggagtactg gtaccgcaat 31140
 45 ctccgtcaga cagtgaatt ccagtcggtg gtggagggtc tggtcgctca ggggtgtcgg 31200
 gtgttcctgg agtccagtcc gcatccggtg ttgacggctg gcatcgagga gtccgcggat 31260
 50 cgggtcgtgg cgttgagtc gctgcgtcgt ggcgaggggt gtctgcggcg gttggtggat 31320
 gcggccggtg aggcgtgggt gcgtgggggt cggatcgact gggcggggat gctcgcggc 31380
 ggccggcggg tcgacctgcc cacctatccc ttccaacacc agccctactg gctcgaactca 31440
 55 ccacgacacc ctgcgggaga cgtgaccgcc gtcgggtctca cagaggccgg tcacgcgttc 31500
 gtgccggcgg cggtcgacct gccggacggg cagcgggtct gacgggacg actgtcgctt 31560
 60 ccttcctacc cgtggctggc cgatcatcag gtgtcgggc aggtgtctgt ccccggcgtg 31620
 gtctgggtcg aactcgccct gcacgcgggg caccaggccg gatgcgactc tgtcgatgag 31680

65

ES 2 287 552 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

ctacacctac agtcgcgcgt cgtgctcggg gcgtccgaca ccgtacaggt gagggtcgtr 31740
 gtcacggaga ccgaagagcc cggcaccgc accgtgtcga tgcactcgcg ccgtgacgac 31800
 ggcagctggg tgactcacgc cgaggggatc ctcggggcgg gcggggcgcc gccggagccg 31860
 ctgcccgaat ggccgcccac cggcgccatg cccctcgatg tcgagggcct ctacgacgag 31920
 ctcgcggcgg gcggctacca ctacgggcct cagttccgct gcctgcggcg gcctggcggt 31980
 gccggtgagg atctcgtcgc cgagatctcg ctgccggagg gcaccgacgt cgatgcgtac 32040
 ggccctgcacc ctggactcct cgacgcggcg gtgcacagcg tggcctgcgc ccggacgagc 32100
 gcggggggcg gcgatgacgg tccccggctg ccgttcgcct tctcggacgt ccggctcttc 32160
 gcgaccgggg tgacctcgt acgggtccgg atcgatccgc agaactcctc gtggcaggcg 32220
 tgggacgaat ccgggctgcc ggtcctcacc atcgggcggc tcgcccggcg gcctgtcgac 32280
 gccgatcagt tcgcgtgcg gcgggggggc cacctcttcc gcgtcgaaac gcggcacgaa 32340
 gccctggccg gcccgcccc cgctccttg gcggtcatcg gagcggaccc ggccgggtac 32400
 gccgcagccc tggaggccac gggcgcgag gtgacgacgg ctgccgacct ggccgggtctc 32460
 acatcggcac ccgaagccgc cctgttcacg ctccccggca caaaggacgc gggggtcacc 32520
 gaggaggtgc cgaccgtgt ccgggaggcg accgctcagg tgctggaggt gctgcaggac 32580
 tggctcaccg acggacgttt cgacgatgcc cgactggtcg tcgtaagccg cgaagcggaa 32640
 gacggcgatc tctccacgg aacggcgcg gcactgctgc gcgcgcgaca ggccgagcac 32700
 ccggaccgca tcaccttgt cgacctgat gctcatcccg cctcgctcac ggcccttccc 32760
 ggttttcgcc tcggtccga accggaggtc gtcgtacgg cgggagacgg cagggcaccg 32820
 cgcttgccc gggcgaggc ccccaccgga gcgggctcac tgggcacggg cacggtcctg 32880
 atcaccggag gcacgggac cctcggggga ctgctcgccc ggcacctggt ggagacgac 32940
 ggagtcaccc ggctgtgct ggtcagccga cgaggaccg ccgcccagcg cgccgaccg 33000
 ctgcacgccg agctcacgg gcatggcgca cactcgaca tcgtggcggc cgacctcggc 33060
 gaccgcacga gcgtggccgc gctcctcgcc acggtcgacg ccgaccaccc cctgtcggcc 33120
 gtggtgcacg ccgccggagc gctggacgac ggctgtctcg gcaccgggtc cgccgactgg 33180
 ctgaccccg tctgcgccc caaggcggac gccgcttggc acctgcacga actcaccgcc 33240
 gaactgcctc tgaccgcctt cgtcatgttc tctcggccg cctcgtgct cggcgcggcg 33300
 ggacaggcca actacgcgc ggccaacgga tttctggag cactggccgc ccatcgtgcc 33360
 gcccggggac tgcccgggac ctcgctggcc tgggggctgt gggagcaccg cagcgaactg 33420
 acccggcaca cgggtcccc ctcccgacg atcgccggcg tcggcgctct gtccaccgg 33480
 gaggccttg ccgcttcga cgccggcctg gcctccgggg agccgctggc agtgccgac 33540
 cggctggagt cgacatccag tgaggaggta ccgcgatgc tcgcggccct ggtccgcgta 33600

ES 2 287 552 T3

cgccgcggg cgcgccggg cgcggaaccc ggggcgagcg cgggcgccgc gcaggaggtc 33660
 5 cggcagctgg ccgagttggg cgcgcacgag cgacagcggc gcgtgcagcg gatcgtgctc 33720
 gacaccgcgg cggccgtcct cggccatgac agccacgacg ccatccccc caccgggggc 33780
 ttcttgagc tggggttcga ctccctgaca gcggtacggc tgcgcaaccg gctcgcccg 33840
 10 cgactggggc tgcgcctgcc ggccacggtg gtgttcgacc accccagccc ggccgcctg 33900
 gccgcccacc tggtcgagca tctcgtggg accgtcgacc cgaccgcgca ggccatggag 33960
 cagctggagg ctctgcgcg cagcgtgcac gcagccacac ccgccggtgg cctggaccgc 34020
 15 gccctggtga cccaacgcct gacggccctg ctcgacgaaa tgcggcacgt cgacggcccc 34080
 ggcggcaccg aaggccccga cggctccggg gacgacctgg agaacgcgac agcggacgag 34140
 20 atctacgcc tcacgacaa cgaactggg atcgggggta cgcagtgaac ggcgacgaca 34200
 aagcactggc ctatctcaag cgggtgaccg cggacctgcg gtcggcgaga gccaggctgc 34260
 aggaactgga gtccgccgac accgaccca tcgccatcat cggcatggg tgcgtctgc 34320
 25 ccggtggcgt gcgcacccc gaggacctgt gggacctcgt ggagaagaag catgacgcga 34380
 tcggccccc cccgcgcgac cgcggatggg acctcgagaa cctgtacgac cccgacccgg 34440
 acgcgcggg caaggcctac gtccgcgaag gtgggttcgt ccacgacgtc gccggcttcg 34500
 30 acgcgggctt ctccggaatc tcgcgcgtg aggcgctggc gatggaccgc caacaccggc 34560
 ttctgctgga gtgtcgtgg gaggccttg agcgggcggg catcgacct tcctccctcg 34620
 35 agggcaccgc caccggcgtc tacaccggg tcacgaccca tgaatacgcg acccgactgc 34680
 cctcgatcga cgaggagttg gaggggtgtc tcggcatcgg caacgcggga agcgttgctt 34740
 cgggcgcggt ctctacacg ctccgcctga acggccccgc tgtcacggtc gacacggcct 34800
 40 gctcctctc gtcgtcgcc ctgcacctc ccgccaaagc cctgcgccag ggccagtgc 34860
 cccttgcgct ggccggaggt gcctcgtca tcgggcgcc gacctgttc gccacctca 34920
 gccgacagcg ggcctcgcc ccgacggcc gctgcaaggc gttctcgtcc acgacgacg 34980
 45 gcacgggctt cggcgaggg gtgggcgtac tggctctgga gcgcctctcg gacgccgctc 35040
 gcaacggaca cgaggctctg gccgtcgtac ggggctcggc ggtcaaccag gacggagcca 35100
 50 gcagcggatt caccgcccc aacggaccgt ccagcagga cgtcatccgc gaggccttg 35160
 ccgacggctg actgacctc gcggacgtgg acgtcgtgga gggtcacggt acggggacgc 35220
 ggttggtgga tccgatcag gcgcaggcgt tgctggcgac gtacgggcgg gggcgtgacg 35280
 55 cggatcgtcc gttgtggtg gggtcggtg agtcgaacat cggtcacacg caggcgccgc 35340
 cgggtgtggc aagcgtgatc aagatggtgc aggcgatgca ggcgggcgtg ctgccgcgga 35400
 60 cgtgcatgt ggacgagccg tcgggtgagg tggattggga ctcggtgcg gtcggctgc 35460
 tgaccgaggc gcgcgagtg ccgtcggggg aggggcgtgt ggcggggcg ggtgtgctc 35520

65

ES 2 287 552 T3

cggttcgggat ctccgggacg aacgcgcacg tgatccttga ggagccgccg gcggaggacg 35580
 cgctgccgga gcctgaagcg ggtgatgtgg tgccgtgggt tctttcggcg cggtcggcag 35640
 5 aggcgttgcg ggagcaggct gcccggtcg cgctcgtggc tgggtgggtg aacgtgggtg 35700
 atgtgggctg gtcgttggtc tcgacgcgtg cggcgttcga gcaccgggccc gtcgtcgtgg 35760
 10 gaggcgatcg ggaagagctc ctggggaagc tttcctcggg ttcgggggtc gaggtcgggg 35820
 tcggggtcgg tgccgtgggt ggtgtgggtg tgggtgttcg cggtcagggg tgtcagtggg 35880
 tcgggtatggg gcgggagttg ctgggttcct cgctggtgtt cgcgagtcg atgcgggagt 35940
 15 gcgcggcgcc tctgtcggc tttgtggact tttctgtggg ggatgttctg ggttcggctg 36000
 gggagttggg tcgggtcgag gtggttcagc ctgcgttgtg ggcggtgatg gtgtcgtcg 36060
 20 cgcggtgtg gcggtcgtg ggtgttccg ttgtcgggt ggtgggtcat tcgcagggtg 36120
 agattgccgc ggcgacggtg gcgggtgcgt tgagtgtggg tgatgcggcg cgggtgggtg 36180
 cgttgcggag ccgtttgatc gcggagcgtc tgtcggggct ggggtgggat gtttcggtg 36240
 25 cgttgtcgcg tgagcgggtg gtgtcgttga tcgcgggtgt gccgggtgtg tcgggtggcg 36300
 cggtgaacgg ttcttcgtcg acggtggtct cgggtgaggc cgcggggctg gagaggggtc 36360
 tggccgcgtg tgtgtcgtc ggggttcggg cgcgtcgtat cgatgtggat tacgcctcgc 36420
 30 attcgggtga ggtggagttg atccgtgagg agttgttggg ggttctggac gggatcgtcc 36480
 cgcgctcggg tgagattccg ttcgtgtcca cggtgacggg tgagcggatc gacactgtcg 36540
 35 agctgggggc ggagtactg taccgcaatc tccgtcagac agtggaaatc cagtcgggtg 36600
 tggaggggtc ggtcgtcag ggggtgcggg tgttcctgga gtccagtcg catccgggtg 36660
 tgacggtcg catcgaggag tcgcggatc gggtcgtggc gttggagtcg ctgcgtcgtg 36720
 40 gcgaggggtg tctgcggcg ttggtggatg cggccggtga ggcgtgggtg cgtgggggtc 36780
 cgatcgactg ggcggggatg ctgcggcg gcggcggggt cgacctgcc accctatcct 36840
 45 tccaacacca gccctactg ctgcactcac cagcacacc tgccggagac gtgaccggcc 36900
 cgggcgacga cgagtcttg gcggccgtg agcacggtga ggcgaccgag ttggcgacc 36960
 tgctccggag gtcggcgcg gagccggggc aggatcttca cgcaccgtc gcggccctgc 37020
 50 tgccgacgct tgcaacgtg cgtcgggacc ggcagcgag ggcggctgtg gactcctggc 37080
 ggtaccggat cgtatggcgt ccggtcgcca cgccctcgta cgacaggggt ctgtcggggc 37140
 55 gctgggctgt cgtcgtgcc gccggtcac aggacgacc cgtcgtcgac tgggtctgct 37200
 cggcgtcgc ggaccacgg ggcgagccc aacgcattgt gctgggccc cgggagagcc 37260
 gttcggcgt ggccacgcg ctggccggc atcccccg gggcgtggc tccctgctcg 37320
 60 gactgagcgg ggccgcgac ccgaccacg aggtgtgcc cagtgcgtc gccggatacc 37380
 tctgtctgc ccaggccctc tcgacggcg ccgtacgag accggtgtg accctgacc 37440

ES 2 287 552 T3

gcaacgggtgt gtccgcgacg gcgacggacc cgggtggctcc cagcacgccc gcgcaggtgt 37500
 gggccgtggc acgggtggcc ggtctggagc acccgaggc gtgggtggc ctgctcgacc 37560
 5 tgccggaccg tctcgacgac cgcgcggccg cccggttcgc cgcggtcctg tccgcgggcg 37620
 aggacgagga ccaactggca ttacgcgacg ctgggttgct ggcacgaagg ctggtgcgtg 37680
 10 cccccgttcc gcgcgacgcg gtgaccgcg gctggcagcc ccgcgacaca gcgctcgtca 37740
 cgggcggcac cggcgggtctc ggcgggcagg tcgcccgtg gctggcggcc gcgggcgtac 37800
 ggcacctcgt gctggtcagc cgtcgggggg cggaggcgga gggcgcagac cgtctgcgcg 37860
 15 acgacctcac cgcctcggc gtacaggtga cgttcggcgc gtgcgacgtc gcggaccgcg 37920
 ccgcgctctc ggcgctctc gaccgggttc aggaggacgg cccgccgatc cgcacggtcg 37980
 tgcacgcggc gggctccggt cgcgcgccca ggtgctgga caccgacgcc gaggagaccg 38040
 20 cggcgggtgct gcgggcgaag tcggccggag cccggaacct gcacgaactc ctcgatgacg 38100
 tggacgcgtt cgtgctgttc tcctccggag cgggtgtgtg ggaagcagc gccagggcg 38160
 25 cctacgcggc ggccaacgcc tacctggacg cactggccga acagcgcagg ggcaggggc 38220
 ggccggcgac ctccgtcgcc tggggcgccct gggcgggtga cggcatgaca gccgcccgcc 38280
 gcgaggaatg gtggagcagg cagggtctgc ggttcattgc cctgaggcc gccctcgacg 38340
 30 cgctgcgcca ggccgtcgac cgcgcgcgaga gcacgctcgt cgtcgcagac atcgactgga 38400
 agacgttcgc tccctcttc acgtcggccc gcagccgccc cctcatcacc gacatactg 38460
 aagcccggcc cgaaccgagg ccggaaggcg cggaccagcc tacgcagggc ctcgaggcca 38520
 35 agctggcggc gctgtccgcg gacgaacggc ggcgcgccct gctcgcgag gtgcgggcgc 38580
 aggcagcggc ggtgctcggc cccccggcg cggacgcgt accggtcgac cggccgttcc 38640
 40 gcgagctcgg attcgactcc ctacgcgcg tgaactgcg caacaggatc gttgctgcca 38700
 ccgggctcga gcttcggcc accctggtct tcgaccaccc cagttccacg gcgctcgccg 38760
 cctacctggg cgcgcggctc ggaatcgacg gcgccccgc ggggtccact ctgctggaag 38820
 45 acctcgcgcg gctcgagtc accgtcgcca cctgaccgc ggcacctctc gcagagaccg 38880
 tgccggacgc ccgggaccgc ggcgcgtca ccacacggct gcgggcgttg ctggagcggc 38940
 50 gggaccaggc cgatggcgag gaccaggccg ccgcccgaga agaactcgac gatctgagcg 39000
 acgacgacct ctctgacttc atcgacgcga agttcggccg ttcgtgacct cggtcggccg 39060
 ccgccaactc cagttacacc ccgaagacca cgatcaccac gcgaaaagga cgggcctctc 39120
 55 catgggggac gagcagaaac tccgcacctc cctcggcgcc gtcactgccg acctggccga 39180
 cgtgacggag cggttgcagc gagcagagga caagaacgcc gagccgatcg cgatcgtcgg 39240
 catggggtgc cgctaccccg gtgggtgctg gtcgcccag gagttctgga acctgctcga 39300
 60 cgaaggcgtc gacgcagtgg ccggcttccc cgaggaccgt ggtgggacc tggagaacct 39360

65

ES 2 287 552 T3

gtacgacccc gaccccgacg agccgggtaa gtgctatgcc cgcgaaggcg ggttcctcta 39420
 cgacgcgggc gagttcgacg ccgcgttctt cgggatatcg ccccgcgagg ccctgtccat 39480
 5 ggaccccgag cagcggctgc tgctggagt ctctggagt gccctcgagc gggcgggcat 39540
 cgacccgggc tcgctgcgcg gcaaagacgt cggcgtgtac gtcggcgcat ggaacagcaa 39600
 10 ctacggcagg ggcggcgggg cggagagctc cgagggccac ctgctgaccg gcaacgcctc 39660
 cagcgtggtc tcgggtcgcg tggcgtacgt gctggggctc gaaggccccg ccgtcaccat 39720
 cgacaaccgc tgttctctct ccctggtcgg cctgcacctg gccgcccagg ccctcaggtc 39780
 15 cggcagagtgc ggtcttgccg tggccggcgg cgtcacctg atgtccacc ctctgtcgt 39840
 ggtgtccttc tcccggcagc gggggctcgc acaggacggt cgttccaagg cgttctcggc 39900
 ggacgcggat ggcatgggca tggccgaagg tgtaggcgt ctggtcctgg agcgcctctc 39960
 20 ggagggcgcg cgcaacgggc acgaggctct ggccgtcctg cggagctcgg ccgtgaacca 40020
 ggatggtgcc tcgaacggtc tgagcgcccc gaacgggccc gcgcagcagc gtgtcatcca 40080
 25 gtccgccctg accgtcggcc gtctcgcccc ctccgacatc gacgtcgtcg agggccacgg 40140
 caaccggcacg gccctcggcg atccgatcga ggcgaggcg ttgctggcga cgtacgggcg 40200
 ggggctgat gcggatcgtc cgttgtggtt ggggtcgggt aagtcgaaca tcggtcacac 40260
 30 gcaggcggcc gcgggtgtgg ccggggtcat caagatggtg ctggccctgc gcaaggcggt 40320
 actgccgcgg acgttgcatg tggatgagc aaccggtgag gtggattggg actcgggtgc 40380
 35 ggtgcggctg ctgaccgagg cgcgcgagt gcccgcgggg gaggggcgtg tgcggcgggc 40440
 ggggtgtctg tcgttcggga tctccgggac gaacgcgcgt gtgacgtcg aggaggctcc 40500
 ggaggaggag ccccgcccg aggtcctctc cgtcgacgtg gtgccgtggg ttcttctggc 40560
 40 gcggtcggca gaggcgttgc gggagcaggc tgcccggctg gcgtcgggtg ctggtgggtt 40620
 gaacgtggtg gatgtgggtt ggtcgttggc ttcgacgcgt gcggcgttcg agcaccgggc 40680
 cgtggtggtg gggcgggact ccgaggaatt ggtgtcgggg ctttctctcg tttcgggggt 40740
 45 cgaggtcggg gtcggggtcg gtgccgggtg tgggtgtgtg ttggtgttcg ccggtcaggg 40800
 gtgtcagtgg gtcggtatgg ggcgggagtt gctgggttcc tcgctggtgt tcgcggagtc 40860
 50 gatgcgggag tgtgcggcgg ctctgtcgcc gtttgtggac ttttctgtgg tggatgttct 40920
 ggggttcggct ggggagttgg gtcgggtcga ggtgggttcag cctgcgttgt gggcgggtgat 40980
 ggtgtcgtcg gcgcgggtgt ggcggtcgtg ggggtgttcc gttgctgcgg tgggtgggtca 41040
 55 ttcgcagggt gagattgccg cggcgacggt ggcgggtgcg ttgagtgtgg gtgatgcggc 41100
 gcgggtggtg gcgttgcgga gccgtttgat cgcggagcgt ctgtcggggc tgggtgggat 41160
 60 ggtttcgggt gcgttgctgc gtgagcgggt ggtgtcgttg atcgcggtg tgccgggtgt 41220
 gtcggtggcg gcggtgaacg gttcttcgtc gacggtggtc tcgggtgagg ccgcggggct 41280

65

ES 2 287 552 T3

ggagaggggtg ctggccgcgt gtgtgtcgtc gggggttcgg gcgcgtcgta tcgatgtgga 41340
 ttacgcctcg cattcggtgc aggtggagtt gatocgtgag gagttgttgg gggttctgga 41400
 5 cgggatcgtc ccgcgcctcg gtgagattcc gttcgtgtcc acggtgacgg gtgagcggat 41460
 cgacactgtc gagctggggg cggagtactg gtaccgcaat ctccgtcaga cagtggaatt 41520
 10 ccaagcatcc gtgcagacgc tcctcgccca ggggcaccag gtcttcctgg agtccagtc 41580
 gcacccggtt ctacccgtcg gcacgcagga gaccgttcac gagagcggcg cacaggccgt 41640
 cgttctggga agcctgcggc gggacgaggg tgccctcacc cggctcgtca cctccgcgg 41700
 15 tgaggcatgg gcgcgcgggtg tgcccgctga ctgggcgggc atgctcgccg gcggcaggcg 41760
 ggtcgagttg ccacgtatc ccttcctccg ggagcggctg tggtcgagc cgctcgctc 41820
 ccgcaccggg aacctcaaca tggccgggct ggtcgaagcc ggacatgaaa tcctgcccgc 41880
 20 cgcagtggag ttgcccgag agcagtgggt gtggaccggc gagctgtcgc tctccgcgt 41940
 cccgtggctg gccgatcacc aggtgctcgg gcagaccctg gtgccggggc tggcgtgggt 42000
 25 cgaactcgcc ctgcacgcgg gccaccagct cggtttcgga tccgtcgagg aactcaccct 42060
 gcaggcacccg ctctgtctcg gcgagtcga cgcctgcag gtcagagtcg ttgtctccga 42120
 tctcggggag agtgatcgcc gggcagtgtc ggtgcactcg cgtggtgacg accagacgtg 42180
 30 ggtgacccat gcggagggat tcctcaccgc gaaagggggc cagccggaga ccatggccgt 42240
 gtggccggcg tccgggtcgg agccggtgga ggctgacggg ttctacgaac gcctcgccga 42300
 35 tgcgggggtac cactatggcc cggcttccca gggcgtgagc aaggtctggc gagctggcga 42360
 ggagatctac gccgaggtcg ggctgctga cgcgcgcgac gtggacggct tcggcatcca 42420
 cccgcctcg ctgcacggcg cctgcagac cgcctacgtc gcccaacggg gcccgcgaga 42480
 40 gacgaagttg ccttcgcgt tcggcgatgt acagctgttc gccaccggtg cccggtcgct 42540
 ccgcgtacgg gtctcgccgg ccgctcagca ggggatggcg tgggaggcct gggacccac 42600
 cggacttcgg gtgttctccc tcgggtacct ggcgaccgg cgggtcgacc gcggccagct 42660
 45 gaccgtgaag cggcccgagt cgtgttcaa ggtggcctgg gacgagaccg tccccgtcgt 42720
 cgggaatgcg accgcgcgc atggcgtcgt gctgggcgac gaccggttcg cctcggtgc 42780
 50 cgcgtcgcgc gcggcgggct gggaggtcgg ggcgcgcccg gaaccgcgt ccgcccacac 42840
 cgccgcgaa gtactgtgc tgccctgcac cgcgcggcg gagccggacg cggacctgc 42900
 caccgcggtc agggcgtga ctgctcgggt gctcggcgtc ctacaggagt ggctcgccga 42960
 55 cgaacggctc gccggcacc gactggcgt cgtgaccgc aacgcctgc cgggtgacct 43020
 cctgcacagc cccgtctggg gtctcgtgc ctccgccag accgagaacc ccgggcgcat 43080
 60 caccctcgtc gacctcgacg accacccgga ctggcgggc gtccttgccg agggcgtcca 43140
 gtccgacgag ccgcgcatca tggctcgcga gggccggccc accgcggccc gcctggtccg 43200

65

ES 2 287 552 T3

tgccaccgca cccgagctgg tgccgcccgc cggagccgat gcctggcgcc tcgagatcac 43260
 cgaaccgggc acgttcgaca acctcacgct gggcgctctac ccgcacgccg agaagaccct 43320
 5 cgcgcacaac gaggtccggg tcgccgtcca cgcgggcggc ctcaacttcc acgacgtggt 43380
 cgccgcactc ggcatggctg aggacgacct gacctcggc cgtgaggcgg ccggcgctgt 43440
 10 cgtcgaggtc ggagacgccg tgccggatct gacccccggc gaccacgtga tgggcatcct 43500
 gtctccggc ttcgggcccgc tcgccgtcac cgatcaccgc tacctggcac gcatgcccga 43560
 gggctggagc ttcgcccagg cggcttcggg gccgcgcggc ttcttgacgg cctactacgg 43620
 15 gctgtgcgac ctccggcgga tccgcgcggg cgaccgcgtc ctcatccacg cgcccgccgg 43680
 cgggtgcggc atggccgccg tacagatcgc ccggcacctc ggggcggagg tgttcggcac 43740
 cgccagcccg cgcaagtggg gcgcgctgcg cgccttgggg ctgcacgacg ccacactgtc 43800
 20 ctctccccc accctcgact tcgagcagga gtctctggac gccaccgacg gcaggggagt 43860
 cgacctcgtt ctgaactcgc tggcccggga gtctctgcac gcctcgtgc ggctgatgcc 43920
 25 cggcgcgggc cggltcgtgg acatgggcaa gaccgacatc cggcgggccg aacagggtggc 43980
 ggaggaccac ggcggagtcg cctaccaggc attcgacctc gtcgaggccg ggccgcagcg 44040
 caggggggag atgctcgcg agatcgtccg gctcttccaa gccggcgctg tccggccgct 44100
 30 gccgatcacc cagtgggacg tgcgccgggc gccggaggcc ttccgacaca tcagccaggc 44160
 caagcacata ggcaagatcg tctcaccgt gccccggccc atcgacaccg acggcacctg 44220
 35 catggtcacc ggcgccaccg ggaccttggg cggcttcgtc gcccggcacc tggtcaccca 44280
 tcacggcata cgacgactgc tgctggtcag ccgcagcgcg gagcgacccg acctggtgcg 44340
 ggaactcacc gagctgggcg ccgacgtcac ctgggcctcc tgcgacctag ccgacgccac 44400
 40 cgccgtcgaa gagaccgttc ggtccgtcga cgaacggcat ccgctcgtgg ccgctcgtcca 44460
 ctctgcggga gtactcgacg acggcgtcat cgacaagcag agccccgaac ggctcgacac 44520
 45 cgtgatgcgt cccaaggtcg acgcgcctg gaatctgcac cgactcctcg acaacgcccc 44580
 gctggccgac ttcgtgctct tctcctccgc cagcggcgtg ctccgtggcg ccggacagtc 44640
 caactacgcg gccgccaacg ccttctctga cgcgctcgc gagcaccgcc gtgcacaggg 44700
 50 cctcgccgga caggcgctcg cctggggact gtggtccgac cgcagcacga tgacgggaca 44760
 gctcggtccc accgaactcg cccggatcgc ccgcaacggc gtcgcccaga tgtccgagac 44820
 ggagggcctg gccctcttcg acgcccggcg ggacaccgcc gaggcggtgt tgctgcccat 44880
 55 gcacctggac gtcgcgaggc tccgcagccg caacggagag gtaccgcggg tgttcgcggc 44940
 gctgatccac gccacggccc gccgcaccgc gagcaccgcg gtccgcagcg ccggcctcga 45000
 60 acagcagctc gcctcgtgt cgggccccga acgcacggag ctgctcctgg gactggtgcg 45060
 cgacctatgc gccgcggtgc tcggccacgg cacctcgcac gccgtctcgc cggaccggcc 45120

65

ES 2 287 552 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

ctccgcgac ctgggtttcg actccctgac tgcggtggag ctgcgtaaca gggtcgcgcg 45180
 cctcacccgc ctgcgtctgc cggccacgct cgtcttcgac cacccgagcc cgacggccct 45240
 cgccggggcac ctcccgggcc tgctggggcg cgcgacgccc tccgcgccg agccggtcct 45300
 ggccgcgctc ggacggctgc gcgccgacct ccggtcgtc accccggacg ccgagggcg 45360
 cgaggacgtg acgatccagc tggaggccct cctcgccgag tggcgggagg ccgcgagaa 45420
 gcgggctccg gaggcggtcg gtgacgagga cctgtccacc gccaccgacg acgagatctt 45480
 cgcgctcgtc gacagcgaac tgggtgaggc ctgatgacg ccgaagcgtc tcaggacaag 45540
 ctgcgtgact atctgcgaaa gaccctcgcc gacctcgga ccaccaagca acggctacgc 45600
 gacaccgaac gcagggcgac cgagcctgtc gcgacgtcg gcatgagctg ccgactgccc 45660
 ggcgacgtac ggacaccgga gcggttctgg gaactcctcg acactggaac cgacgcccctg 45720
 acgcccttgc ccaccgaccg cggctggaat ctcgacacgg cgttcgacga cgaacggccg 45780
 taccggcgcg aaggcggatt cctttacgac gccggacggt tcgacgccga gttcttcggc 45840
 atctcgcccc gtgaggcgct ggccatggac cctcagcagc ggctgctcct cgaaagctcg 45900
 tgggaggcga tcgagcacgc ccgcatcgac ccaggtccc tgcacggcag tcgcaccggc 45960
 gtctggttcg gcacgatcg ccagactac ttctccctct tcgcccacg cgccggcgag 46020
 caccgcaact acttgccac cgctgctcg gccagcgtga tgcgggccg cgtctcgtac 46080
 gtgctcgcc tggagggggc cgtgtcacg gtcgacacgg cgtgctcgtc ctccctggc 46140
 gccctccact ccgccgtaca ggccctgagg tccggcgagt gcgaactggc tctcgccggg 46200
 ggccgcaacg tcatggccac ccgacggtg ttaccgcct tctcccatca gcgtggcctg 46260
 gccggtgacg gccgctgcaa ggcttcgcg gcgggtgccc acggggcggg cttcgcccag 46320
 ggggtggcg tgctggtgct ggagcgggtg tcggtggcgc ggccgaacg tcatcggtg 46380
 ttggcggtg tgccgggttc ggcggtgaac caggacggtg cgtcgaacg tctgacggcg 46440
 ccgaacggtc cgtcgcagca gcgggtgatc cgccggcgcc tggccaacgc gcgcttggcg 46500
 ccggaggacg tggacgctgt cgaaggccac ggcacgggga cttcgctggg cgacccgatc 46560
 gagcgcgagg cgttgcctgc gacgtacggg cggggccggg acgcggagcg tccgttctgg 46620
 ctggggtcgg tgaagtcgaa catcggtcat gcgcaggctg ctcccggtgt cgccgggtgc 46680
 atcaagatgg tgctggccat ggagaagggc cgtctccctc ggacgctgca tgtggatgag 46740
 ccgtcgggtg aggtggactg ggactcgggt gcggtacggc tgctgaccga ggcgcgggac 46800
 tggcgtcgg gggagggggc ggtgcggcg gcgggaggtg cgtcgttcgg gatctccggg 46860
 acgaacgcgc acgtgatcat cgaggagccg caggagggag aagcggcacc ggattcctct 46920
 gcttcgggtg ccgtgcctg ggtgctctcg gcgcgatcgg ccgaagcgtt gcaggctctg 46980
 gcttcacaac tcgcccacca cagcgccaaa tcgagtcggg tggatgtggg ttgggtcgtc 47040

ES 2 287 552 T3

gtttcgacgc gtgcggcggt cgagcatcgg gccgtggtgg tggggcgcgg gcgggacgag 47100
 5 ttggtgcggg gcttgtecca ggtcgcgcag ggtcggggcg tgaggggtgt cgcgtcttcg 47160
 gcgtcgggtg gtctcgcgtt tgtttttgct ggtcagggca gtcagcgggt ggggatgggg 47220
 cgggggttgt atgagcgggt cccggtgttt gccagggcgt tcgacgaggt gtgtgggcgg 47280
 10 gtcggtccgg ggggtcggga ggttggtttt ggttcggatg cgggtgagtt ggaccggacg 47340
 gtgtgggcgc aggcgggggt gttcgcgttg gagtgggcgc tgtttcgggt gttggagtcc 47400
 tggggtgtgc ggccgggttg tctgatcggg cattcggtcg gtgagttgtc ggccgctgt 47460
 15 gtggcgggtt tgtggtcgtt ggaggatgag tgctgggtcg tggtgcccg ggccggttg 47520
 atgcaggcgt tgccggcggg tggggtgatg gtcgcggttc ggcccgaggc gggggagctg 47580
 gccggtttcc tcggtgagga cgtggtgatc gcgtcgggtg acgcgccggg gcaggtggtg 47640
 20 atcgtcgttc ctgagggggg tgtggagcgt gtggtggctg cttgtggggc gcggtcgcgt 47700
 cgtctggcgg tctcgcacgc ttttcattcg cctttggtgg agccgatgct gggggagttc 47760
 25 cgtcgggttg tggagtcggt ggcgttcggt gtgcgctcgt tcggggtggt ttccaatgtc 47820
 acgggtgcgt ggggtgatcc ggaggagtgg gggacgcgg agtactgggt gcgtcaggtc 47880
 30 cgtgagccgg tgcttttcgc cgacggggtc gccacgttgc tcgacgcggg tgtgaggacg 47940
 ttctcgcagc tgggtccgc tgggacgctc acttcgatgg tcagccactg cgcggacgcc 48000
 accgccactt cgggtgacggc tgtacctacc ctgcgccccg atcacgacga gtcgcggacc 48060
 35 gtgttgagtg ccgcagcgtc cttgtacgtc cagggtcacc cggtcgactg ggccccgctg 48120
 ttcccgcggg cccgcacggt ggacctgcc acctaccct tccagcacca gcaactactg 48180
 40 atggaaagcg ccgcccgcc caccgtcgag gacacccgc gcgagccct cgacggctgg 48240
 acgcaccgca tcgactgggt gccgctggtg gacgaggaa cggcgccct cctggccggt 48300
 acctggctgc tcgttcgtcc cgaagaaggt cccgcgccgc tcgccgacgc cgtcgcggac 48360
 45 gcgctgaccc ggcacggcgc ctccgtcgtc gaggccgctc gtgtcccga ccaatccgac 48420
 accgagctga ccggagtcgt ctctctgctg ggcccggcg ccgacggcga cggcggcctg 48480
 gacgcgaccc tgccgctggt acaggacttg gccacgcgg ggtccaccgc gcccttggtg 48540
 50 atcgtcacca gcggagccgt ggccgtcgg acgtccgaca ccgtgccga ccccgagcag 48600
 gcgacgtctt gggggttggc ccggggcgcg gccaccgagt ggcccgccct gggggcgcg 48660
 55 cgcacgcacc tgcccgcga cctcaccgag caggtcggac gtcggctctg cggccggtg 48720
 ctccaccgga gtgagcagga gacggcggtc cgacaggccg ggtgttcgc caggcggctg 48780
 gtccgtgccc gtaccagcga cggccggtg acgcgcgcg gcaccgtgct ggtcacggc 48840
 60 gggaccggcg cgtcgcggc acacgtcgc cgatggctgg cggaggagg ggccgagcac 48900
 atcgtgctgg ccgggcgcag agggcccgc ggtcagggcg ccgaggcgt gcgcgcgcac 48960

65

ES 2 287 552 T3

ctggtcgcgc caggggtcaa ggcgacgac gtgcgctgcg acgtcgcgc cgggatgcc 49020
 gtacgtctgc tectggacgc acaccggccc agcgccatcg tgcacacggc cggggtcgtc 49080
 5 gacgacggac tgctcacctc gctgacgccc gccaggtcg agcgggtgct gcggcccaag 49140
 ctgctcggcg ccaggaacct gcatgagctg acccgggacc gggaactgga cgccttcgtg 49200
 10 ctgttctcct ccctcgcgcg agtcctcggc ggggcagggc aggccaacta cgcgcgtgcc 49260
 aacgcctact tggacgcctt ggcgcacac cgcaccgccc atgggctgcc ggcggcctcg 49320
 ctggcatggg ggccgtggga gggcgacggc atggcgcgc cgcaggaggc cgcgcaccgg 49380
 15 ctccgcgcga gcggtctcac ccgctgcgc ccgagcagg ccgtacgggc cctcggccgg 49440
 ggccacgggc cgcgtgttgt ggcgcacgc gactgggcgc ggctggccgc cgcctcgacg 49500
 20 cagcgcctgc tcgacgagct tcccgggtg cgtgcggtca ggccggcgga gcctgctgtc 49560
 ggacagcgc ccgacctacc ggcccggtt ggcgggcgtc cggccgagga gcagtccgcg 49620
 gtactgctgg aggcggtccg ggagagatc gccgcgtac tgcgttacgc cgtccggcg 49680
 25 cggatcggcg ccgacacga gttcctcgc ctcggcttcg actcgtgac atcgatcgaa 49740
 ctgcgcaaca ggcttgccac gcgcacggt ctgacgttc cgcgcacgt caccctggaa 49800
 cagcgcaccc ctgcggggt cgcgcgcac ctgcgcgag ggatcgcga ccggcccgtc 49860
 30 gggtcgggtg ccgtcccggt gcccgggagc gctgatgtcc cggagcggg cgcgggtagc 49920
 ggccctcggg agctgtggca ggaagccgac cggcacggcc ggcggtgga gttcatcgac 49980
 35 gtactcacgc cggccgcgc cttccggccc gcctaccgtg aaccggccga gctggagctg 50040
 ccgcctctac ggctcacctc cggcggggac ggcgcgcgc tgttctgcat ccctcgcac 50100
 ctcggaagg ccgaccgca caagtctctg cggttcgcgc cggccctgcg gggaacggcg 50160
 40 gacgtcttcg tectgcgca gcccggttc gtaccgggc agccctgcc cgcgggcctc 50220
 gacgtcctgc tcgacacca cgcgcgggc atggcgggc acgaccggc cgtgctgctc 50280
 ggctactcgc cggcggtct tgccgcgag gcgttgccg ccgactcgc cagctcggc 50340
 45 aggcgcgcgc cggcgtcgt gctcgtcgac acctatgcc ccgacgagc ggaggtgatg 50400
 gcccgatatc agggcgccat ggagcaggc cagcgcgac gcgacggcag gaccggtgcc 50460
 50 gccttcgggt aggcctggct caccgcgatg ggccactact tcggcttcga ctggacccc 50520
 tgtccggtc acgtgccgt gctgcacga cgcgcggcg acccatgac cggtatgcc 50580
 gtcgaagggc ggtggcaggc gcgtggaac ctgcgcaca ccgcgtoga cgtgcccgga 50640
 55 gaccacttca cgatgatgga ggatcacgc ccgcgcacc cgcacaccgt gcacgactgg 50700
 ctcggcacgc ccgtccgcg ccctgagaga accgcgact gacgactcgc cggcgacagc 50760
 60 ggcaccccgc cctgtccct tctgtcgt cgttccct tctctctc gaaacggagt 50820
 tcgttctcat gccttcctc ccgtaagcc ggtccgtgcc cgacactcc ccgcgcgagc 50880

65

ES 2 287 552 T3

acctcgaact gctcaaggag agcggcggcg tctgcccctt caccatggag gacggccgtc 50940
 cggcctggct cgcggccagc cagcagcccg tgcgtccctt gctcgccgac cgcggtatca 51000
 5 gcaacaaccc ggcaagacg ccgcccttct cgcagcggga ggccctgcag aaggagcggg 51060
 gccagttcag ccgtcacctg ttcaacatgg actcgccgga gcacgacgtg gcccgccgca 51120
 10 tgatcgcgga ggacttcact ccccggcacg ccgagggcgt cgggccgtac ttcgaggagg 51180
 tgttcggcga gatcgctgac gaagtgggcc acaaggggcc accggccgag atgatcgagt 51240
 cgttcgcctt cccggtcgcc acccgcacca tctgcaagg gctggacatt ccggaggacg 51300
 15 actcgagta cttccagaag cgcaccgagc agatcatcga gatggaccgc ggcgaggaga 51360
 acctcgaagc cgtcgctgaa ctgcgcgct acgtcgacag cgtcatgcag cagcgcaccc 51420
 gcaagcccgg cgacgacctg ctcagcagga tgatcgtaa ggcaaggcg tccaaggaga 51480
 20 tcgagctcag cgacgccgac ctggtcgaca acgcgatgtt cctgctggtg gccgggcacg 51540
 agccgtcggc caacatgctg ggccctggcg tgcctgcctt cgcgaatte ccggacgtgg 51600
 25 ccgaggaact gcgggcccag ccgcacctgt ggccggggcg gatcgacgag atgctccgct 51660
 actacaccat cgcgggggccc accaagcggg tcgcgggcgc cgacatcgag tacgaggggc 51720
 acacgatcaa ggagggggac gccgtcatcg tgctctcga caccagcaac cgcgaccgca 51780
 30 aggtgcacgc cgaaccgaac cggctcgaca tccaccgctc ggcgggcaac cacctggcct 51840
 tcagccacgg accgcaccag tgccctgggca agcacctcgt ccgggtccaa ctggagatcg 51900
 35 cgctgcgggc tgctgcgag cggtgcccg gcctgcgctt ggacatcgcc aaggaggaca 51960
 tccccttccg cggtgacgcc ctgtctcag ggccgcgcca gotgcgctc acctggtaac 52020
 agccaccatc ggcccccgcc gcggaccggg cagcacgacc cgggtccgcg cgggggcacc 52080
 40 accaccgtca acatccccag agaggtctcc ccgtggagaa gaccgacgtc gaccggctgc 52140
 gcacactcga ccgagagcac atgtgggtacc cgtggacgcc gatgaccgag tggatggccc 52200
 45 gtgatcagct cgtcgctgaa cgcgccgaag gctgctggct gatcgacgca gacggtaagc 52260
 gctacctcga cggccgctcg tcgatgggca tgaacctgca cggccacggc cgcagcgaga 52320
 tagtcgaggc cctggtcgcc caggcgcgca aggcgggtga gaccacgctc taccgctct 52380
 50 cgcacccggc ggcggtggaa ctgcgcgcc gcctggcctc gatggcgccg gccgggctcc 52440
 agcggtctt cttcgccgag tccggatcga ccgcggtgga gacggctctc aaggctgcct 52500
 55 acgcctactg ggtcgcaag ggcaaccgc agcgatccac ctctgtgtcc atggagggcg 52560
 gttaccacgg cgagacccta ggcacggta gcctgcgcg caccaacggc gaacaggtcg 52620
 acatgatccg caagacctac gagccactgt tgttccctc cctctcttc caccagcccc 52680
 60 actgctaccg gtgtcccgtc ggccagtcgt cggacagcga ctgcgggctg gaggtcaccg 52740
 attcgctgga gaacctctc acccgggaga agggccggat cgcgcggctc atcgctcagc 52800

65

ES 2 287 552 T3

cgcggttcca ggccctcgcc ggagtgatca ccgccccgga gggacacctc gcgaagggtcg 52860
 cggagatcac ccgcaggcac ggagtgtccc tcatcgtcga cgaggctcctc accggctggg 52920
 5 cccgcaccgg ccgcacgttc tcttgcgagg ccgagggcgt cacaccggat ctgatgacgg 52980
 tgggcaaggc gctgaccggc ggatatctgc cgtgtctggc caccttggcc acggaggaga 53040
 10 tcttcggagc cttccgtgag agcgtcttcc tcagcggcag cacctactcc ggatacgcgc 53100
 tcggggcggc cgtcgccctg gccagcctcg acctgttcga gaaggaggac gtaccggccc 53160
 gggccaaggc gctcgcgcac gtgtcacca ccgcactgga acccttcgcg gcgctcacc 53220
 15 acgtcgggtga cgtccggcag ctcgccctca tcgccggcgt cgagctgggtg gccgaccggg 53280
 agaccgcgc cccctaccgg ccccaggagc gcgtcgtcga tcgcatctgc accctggcca 53340
 gggacaacgg cgtgctggtc aacgcggtec ccggggacgt gatcaccatg ctgccctcac 53400
 20 cgtcgatgag ccccgacgac ctgcgcttcc tcaccggcac cctgtacacg gccgtccgag 53460
 aggtgaccga agagtgaag ggctgatgcg ggcgccgctc atccgtgcct ggggcggccc 53520
 25 cgagcggctg accctggacc gggtcgaacg gccgtcacc cgcgccgat ggatcgccgt 53580
 acgcgtcgag gcctgcgccc tgaaccacct cgacatccac gtgcgcaacg ggcttccggg 53640
 cgtagcgtg gaactcccgc acgtctccgg cggcgacgtc gtcggcgctc tcgagcaggc 53700
 30 caccgacgag gcgggggaga gactgctcgg cagccgtgtg ctgctcgacc cgatgatcg 53760
 gcgcggcatc ctcgcgagc actactgggg cgggctcgcc gagtacgtc tcgcaccgc 53820
 35 ccacaacgg ctccccgtcc ccgatcagga cgcggaccgg gcacgctacg ccgcactgcc 53880
 catctcctac ggcacggccc agcgcgtgct cttcagccgc gcccggtgc gtcccggcga 53940
 gagcgtgctg ctgttcggcg cgaccggcgg cgtcggcgtc gcctgcgccc agctcgccct 54000
 40 gcgtgccggg gcccgatca tcgcctgctc cggatcaccg gccaaagctc cccggctgcg 54060
 ccgactcggc gtgatcgaca cgatcgacac cggcaccgag gacgtacggc gcagggtccg 54120
 cgaactcacg gacggcggtg ccgacctggt cgtcgactac cagggcaagg acacctggcc 54180
 45 cgtctccctg cgtcggcgc gcgccggcgg ccgcacgtc acctgcggcg cgaccaccgg 54240
 gtacgaggcg acgaccgacc tgcgtacgt gtggtcgtc cagctggaca tctcggctc 54300
 50 caacgcgtgg caccgcgacg atctgcacac gctggctgac ctggtggcca ccgacgccct 54360
 ggaaccggtg gtgcacgcgg acttcccact ctcccgccg cccgaggcgg tcgccgaact 54420
 ggaggagcgc cgggcgttcg ggaaggctgt gatccgcacg gcgtgaactc actcatgtcc 54480
 55 cggctcgatc ccagggggaa acagcgtgac cggcaacacc acatccggcg ccttctcgcg 54540
 gcggacacag aacgcgctcg ccatgcagcg caagatatgc gccagcccg aggagaccgc 54600
 60 ggagcgcgtg ttctccgaca tctctcgggt gtcacgagac accggcttcg gccgcgaaca 54660
 cggcctcgcc ggggtccgca cccgccagga gtggcggcgt gccgtgccca tccgcaccta 54720

65

ES 2 287 552 T3

cgacgaactg gccccctacg tcgagcggca gttctccggc gaacgcgcgcg tgctcaccac 54780
 cgacgacccc cgcgccttcc tgcgcacctc gggatcgacc ggccgcgcga agctgggtacc 54840
 5 caccaccgat cactggcgcc gtgtctaccg cggaccggcg ctgtacgcgc agtgggggct 54900
 ctacttcgaa cagatcggca cgcacggct caccggcgac gaggtcctcg acctgtcctg 54960
 10 ggagcccggc cccatccggc accgactgcg cggcttcccc gtctacagca tcaccgagcg 55020
 ccccggtgctg gacgaccccg acgactggaa cccgccgtgg cgtcacgcga ggtgggttcac 55080
 ccgcgatgcc ggtgccgcga ccatggccga cctgctctac ggcaaacctgc tgcggctggc 55140
 15 cggccacgac ctgagactga tegtctcggt gaaccctcc aagatcgtcc tgctcgccga 55200
 gacactgaag gagaacgccg aacgcctgat ccaggacctg cacgacggcc acggcacgga 55260
 20 ccgggcagcc cgcgcggact tcctccggcg cctcaccgcc gccttcgacc gcaccggagg 55320
 ccgtccgctg ctcaccgacc tgtggcccg cctgcgtctg ctcgtctgtt ggaactccgc 55380
 ctccgcggcg ctgtacgggc cctggctgtc ccggctcgcg accggcgtgg cggcactgcc 55440
 25 gttcagcacc acgggcaccg agggaatcgt cacgctgccg gtcgacgacc acctctcggc 55500
 ggggccgctc gctgtcgacc aggggcattt cgaattcgtt ccgtggcagg acctggacga 55560
 30 cggcagccct ctgcccagg acacccccac cctcggctat gacgaactcg aactcggcg 55620
 cgactaccgg ctgctcatga gccaggccaa cgggctctac cgctacgacg tgggcgacgt 55680
 gtaccgcgtc gtcggagcgg tcggcgccac gccacggctg gagtttctgg gacgcgcggg 55740
 35 attccagtc tcttcaccg gcgagaagct caccgaatcc gatgtgcaca ccgccgtgat 55800
 gcgggtctc ggagcgaac gcaccgacca cccgcacttc tccggcatcc cggctctggga 55860
 caccgcgcc cactacctcg tcgccatcga atgggctgac gccacggca cgttgaaact 55920
 40 gcaggacaca gccgcgcga tcgacgcgac tctccaggaa gtcaacgtgg aatacgccga 55980
 caagcgcgc agcggacgac tgcggccct gcagatcctg cccctggtgc ccggcgcttt 56040
 45 cggccagatc gccgaacgaa ggttcgccca gggcaccgcg ggagcccaga tcaaaccacca 56100
 ctggtgcag aaggactcgg cgttctcga cacgctgcgc gacctcgacc tcgtccgcgc 56160
 ccgcccgggg acgtgacggc atgcgcatcg gattcgccgc acccatgtcc ggcccctggg 56220
 50 ccaccccgga caccgcctg cacgtcgccc gcaccgcga acagctcgga tacgcctcgc 56280
 tctggaccta ccagcgagtc ctggcgcgcc ccgacgactc ctggggcgag gccaaccgca 56340
 gcgtccacga cccctgacc accttggcct tcctggccgc gcacaccacc gggatccggc 56400
 55 tcggtgtcgc cgttctgac atgcgctgc acaccccgc ggtgctggcc aagcagctca 56460
 ccacctcga cctgtctcc ggcgccgac tcgacgtggg cctcggcaac ggtggggccg 56520
 60 ccgaggagta cgcgcgcgc ggcgtagccc ccaccgggct cagccgcgc gccgaggact 56580
 tcctcgctg tctgcgggcc ctgtggggtg agcagaccgt ggtggaacac gacggccct 56640

65

ES 2 287 552 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

tctaccgggt cccgcccgcc cgcttcgacc cgaagcccg cagtcctccg caccgcgcgc 56700
 tgctcctggg cggcgccgag cccggcgac tgcgccgag cggccgctg tgcgacggct 56760
 ggatcgcgag cagcaaggcc ggcccgccg ccatccgga cgccatcacc gtcgtacgag 56820
 acagcgctga gcgaaccgga cgcgaccccg cgacctgag gttcgtctgc cgcgccccgg 56880
 tccggctgag gaccgggtcg gcccccaac agccgcccgt gaccggcacc gcggagacga 56940
 tccggggcga tctcgccgag ctacccgaca ctggcctgac cgagatcttc ctggacccca 57000
 acttcgaccc cgagatcggc tcaccggagc cgcgacccg cgacgtgca caccgcgttg 57060
 atctgtgct gcacgaactg gcccccgaa actggtgaga ggaagagaac agtgtgctgc 57120
 gcgcgccgag ccgtcggaag agaccgaac tacgcccgc tcgacacgga cacagggctg 57180
 atccacctcc tggccggcac tccctacgac gagatccgac cgaccggcga aaccagaccg 57240
 cttgcccagg cccgcctgct cgcaccggtc gaaccagca aagtgtggt cgcaggagc 57300
 aattacggag atgtcgtaac accggacctg gtggtcttca tgaagccgtc cacctctgtc 57360
 gtcggcccca ggagcaccgt cctgctgccc gcggaggcca agcaggctcc gtacgaggga 57420
 gaactcgccc tgggtgatcg gcgcccgtgc aaagacgtcc ccgaagacac cgcggaccag 57480
 gccgtgttcg gctacacctg cgccaacgac gtcaccgctc gggacgtcgg ggaaccgaag 57540
 ggccactgga ccaaggcgaa gagcttcgac acattctgcc cgctgggacc atggatccgc 57600
 accgatctcg accccgctga cctcgtcctg cgcacaaccg tgaacggcac gctgcgccag 57660
 gacggctcca ccaaggaaat gaacaggaat gtccgcgccc tcgtgtcccg ttgcagctca 57720
 ctgatgacgc tgctgcccgg agacgtgatc ctacccgga caccggcggg cgcggcgctg 57780
 ctgctgcccg .gtgacgaggt cgtcgtcgag attgacggga tcggttcgct cgcgaatccg 57840
 atcggcgctg ccaagtagtt cactgactac actcgcgga acaacacggg cccgtctgag 57900
 gcgcttcgag ctgcgccgat ccccgaggag agattccagt gtctgtaatc cgtcccaccg 57960
 ccgaaaccga acgcgcagtc gtggtggtcc cggctgggac gacgtgcgcc gacgcggtca 58020
 ccgcggcaaa gctgccgcgc aatggcccca acgcgatcgt cgtggtgca gaccgcctcc 58080
 gcgccctgag tgacctgac tggaccccg attccgacgt cgaggtcgag gccgtcgcgt 58140
 tgtccagcga ggacggcctc acggtgctgc gccactccac ggcacacgta ctggcccagg 58200
 cggtcacgca actctggccg gaggccaggc tcggtatcgg cccgccgatc gagaacggct 58260
 tctactacga cttcgacgtg gagcgccctc tccagccaga ggacctcgag cgcgtcgagc 58320
 agcggatgaa ggagatcacc aagtccggcc agcgtctctg ccgcgcgag ttccccgatc 58380
 gggaagcggc ccgtgccgag cttgccaagg agccgtacaa gctcgagctc gttgacctca 58440
 agggcgacgt ggacgccgac gaggcaatgg aggtcggcgg gagcgacctg acgatctacg 58500
 acaacctcga cgcgagaact ggagatgtgt gctggtccga cctctgccgc ggccccact 58560

ES 2 287 552 T3

5 tgccgtcgac cgcctgac cggcggttca agctgctgcg caacgcggca gcctactggc 58620
 gcggcagcga gaagaacccc caactgcagc gcctctacgg caccggcctgg cgcacccgcg 58680
 acgagctcaa gtcccatctc gccgccttgg aggaggccgc caagcgtgac caccgcccga 58740
 tcggcgagga actcgacctc ttgcggttca acaaggagat cggccgcggc ctgcccgtgt 58800
 10 ggctgcccaa cggcgcgatc atccgcgacg aactcgagga ctgggcccgc aagaccgaac 58860
 gcaagctcgg ctacaagcgc gtcgtcacc cgcacatcac ccaggaggac ctttactacc 58920
 tctcaggcca tctgccttac tacgcggagg acctgtacgc gccgatcgac atcgacggcg 58980
 15 agaagtacta tctcaagccg atgaactgcc cgcaccacca catggtgtac aaggcgcgcc 59040
 cgcacagcta tcgcgacctg cctacaagg tcgccgaata cggcacggtg taccgattcg 59100
 agcgcagcgg tcagctgcac ggcctgatgc gtacgcggcg ttccagccag aatgacgcgc 59160
 20 acatctactg caccggcgac caggccaagg accagttcct ggaagtcctg cgcctgcacg 59220
 cggactacta ccgcactctg gggatcagcg acttctacat ggtgctcgcg ctgctgact 59280
 25 cggcgaacaa ggacaagtac cagcagcagc agcagatgtg ggaggacgct gagcggatca 59340
 cccgggaggc catggaagag tccgacatcc ccttccagat cgacctgggc ggtgcccgcg 59400
 actacggccc gaaggctgac ttcctgatcc gagccgtcac cggcaaggag ttgcgcgcc 59460
 30 ccaccaacca ggtcgacctg tacaccccgc agcgtttcgg gctgacctac cagcactccg 59520
 acggcaccga gaagcccgc gtcgtgatcc atcgctctcc gctcggtcgc caccagcgtc 59580
 35 tcaccgccta tctcaccgag caattcgcag gtgccttccc ggtgtggttg gcgcgggagc 59640
 aggtccggat cattccgac gtcggaggaa tcacggacta cgcagaggaa gtccgcgaca 59700
 tgctgctgga cgcggacgtg cgtgcgcagc tcgatgccgg cgacggccgg ctgaatgcc 59760
 40 aggtacgcgc ggcggtcacc cgggaagatc cgtcgtcgtg ggtggtcggc aggcgagagg 59820
 ctgagcagcg caccgtaacc gtgcgcgacc gctccggcga ggagacccc atgtccctgg 59880
 45 agaagtctgt ggcctatgtc actggactca tcaggaccaa gagcctggac ggcgcgggcc 59940
 acatccgtcc gctgtccaag gcctgaccca cagccacggg gcccgcggcag gtgtccgcgc 60000
 tcggaccacc ccttccgtc ctcagccgac ggcgggctca tggcagccgc ccgacctgcc 60060
 50 ggtgcccgtg ctgttcggca acccgtgggc gccgcccgcg gaggagaccg cgcgctgccg 60120
 ggggatgac tcgtccggcc cgcgccccgc gctcaccggg cggcgtgga gcgggcccga 60180
 cgggagccgg gaccggccga ccgtccgtgg tcggtctcgt cccggacgag cagggacagc 60240
 55 agtcgggcca cggtgaggcc cgttccgcc ggggtggcga ggacctgtc cgttccgacg 60300
 cggtagatgt tgctgcgcc gtcacgggtg tgggagagat agccgtcctg ctccaggtcc 60360
 60 gagatgatcc gctggacggc gcgctcgtg agtctgcagt ggcggcgat gtcgcgatc 60420
 cgcacgttcg gattgtcggc gatggccgcc agtacacggc cgtggttggg gacgaacgtc 60480

65

ES 2 287 552 T3

catccggagt gagattcagg cactgcaacc atgcacagca ttgtaggac catctttgcc 60540
 5 ggacagccaa tacatgacat acttttcgcg ttaagagtgg catgttctgt cccatgggca 60600
 actgagaagg gacccgaggg tgtctttgga tgaagcggtc gcggggtgct cgcgccacac 60660
 cggeccggct cggtcccg cgcgaggca acccacgcag gcgcagtacg aagcgacgg 60720
 10 cgctggggtc gtcagcgac ggggcgcata cgacatgaac tcggtcgagc ccttgccga 60780
 cgcgttga aa gacgcggccg agaagtctcc gaaggtggtg ctggacgcct ccggcatcac 60840
 ctteggcgac tccacctgc tgagtctgct gatcctcacc caccaggcga cggacttcgg 60900
 15 ggtggccgg cgacgtggc aagtgatgcg gctcatgcaa ctgacgggcg tcgacgcctt 60960
 cttgaaggta cgggccacgg tggaagaggc cgccaccgct taggggcacg gcgtgccggg 61020
 cctcggtga cgcaagcga tggcttgag ctgagaattc cgggcattcg acgttctgcc 61080
 20 tggcccccgg ccgtcgatgg tggccggcca ggcgtgatga agacagtcac ctccctcagg 61140
 cgcgatcgac cgcgcctcg gggaggggcg ttgacgggag aggaagggtg tccatgatc 61200
 25 tgccggcgga gaaggaactg cgtgccgtgc tggctcggtt cgctcaggcg cgcacgcacc 61260
 acgacgtacg tcccagcggc tgcaccagca ggctcctcga ggacgccacg tacacctgt 61320
 gcgtgatgac cgggtgccgt accgccgaac aggtctctcg tacggcgac gaacttctcg 61380
 30 cacagttcgc cgagcgacc gctgcccccg tggaggacga agccctggcc gcgtgagccg 61440
 acggcacaca cctgcggcgc ctgcgtggc aggtgtgtgc ctgccggcg gcggacgcga 61500
 35 gcacctgagg aaatgagaga gagtcatgag cgatacccg cttcggcagc gcgatgagac 61560
 gtgaagggg ccggccaccg agatcccgcc gccgcagtgg cgggacctct tccctgcccc 61620
 cgactggggc ggactgatg agcaggatgat cgtgccgaa gaggcgcgcg ggcccgagca 61680
 40 ctccaccgga gcgcgcgtc cgcgcggcg ccgcgatcg agtcagcgg ccgcgtgatg 61740
 cgcggccctg ccgcgacggc cgcaggtcag gtgaggcgga tccgtgcgt gatgcgttg 61800
 ccggacggct cgggccggat ctgacggcc tggctcagcg ccgccacgat ctccagaccg 61860
 45 tgctgaccca cccggcggg atcggcgggc cggggagcgg gcagggtggt gtcaccgtct 61920
 cgcacgggtga cagtcaccgc gtcgtccgtg aggtcagggt tgagttctat ggggtcggcc 61980
 50 ccgtacttga ccgcgttcgt gatcagctcg ctccaccaca ggtggacggc ctccgacgt 62040
 ctctccggca ccggcgaccg aaggtccgc tgtgaacggg tgaggtagtc ggtcgcgaag 62100
 tgacggggct cggcgatccg cagcgttcc cgccggtacg tcaccagggc cggccggga 62160
 55 tccgcacct ccggtgccc ctgtgtcgac gaaccgtcca gtccggtct acccatgtca 62220
 accgccacgt taccctcgag ccacgcacgc gcgcgacgc ctcccgcgcg atgagaacat 62280
 60 ctcatgtgtt ctaacgatag ttctgtttc cgtcggtcac cgcacctgc ggccaggag 62340
 aaagcggggg cctggacgtg atccgggct caggcgtgc tgagcacgcc tccgcgaag 62400

65

ES 2 287 552 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

cgggccggca gccgctccgg aatccgtgcg gtcgtcgteg ccatgatgtc cctccccgtt 62460
 ccgatggccc gctcgcgta cgggccactc ggcggtacc ccctgcgcgg gctcgcatgc 62520
 acggaggcgg agatcttgtc cgaggccgtt ccgctccgc ccgcgcgacg tgatcaggtc 62580
 ggcggttacg ggatcaggaa ttccagcacg ggtcgggtccc aggccaccgc gggcggttc 62640
 gcgcgggcgg ttccggtggg cagggcgcc acagcgcggt agaagccctc ggcggaggg 62700
 tgcgacacca cccggacacg gtccagccct gccgcgcggg cccgcctctt catgtgatcg 62760
 acaagcagcc gtccgatgcc gcggccctgg gtaccgtcct ggacgaacag cagggtccagc 62820
 tccgcggcgg cgaggagcag cgcgtagaag ccgagcacc gccctcggtg gtcctcggcg 62880
 tcgacgctac gaagacctgg tggttctcga tgtagtcggg cccgacgcgg tagtcggaga 62940
 ccatcgccgc gtacgggccc tcgtaggctc gtgagccgcg tacgagacgc gagagccgct 63000
 tggcgctccc ggcgaccgcc cggcggatga cgatctgcc ccgtacggac gcggaactcg 63060
 attgcacggg gagcagtcta cgcgtccggg acgggccggg acctccgcgg gtcctcgccg 63120
 agccttcctg cagtcctgcc cgatgacgac cacctcagtc cgtcccgatg acgaccacct 63180
 cgtcctcggc cgtccaccgg cgcgcctcgc gtttgggcgg attgatccgc actccgtagt 63240
 gggggcgcggt cgacgcgtcg gcgtggtcgc ggtaaccgat ggcgcactcg ccgcgacggc 63300
 gtgccgcggc cagcagcgtg gcgaaggacg tgggtgtccc tggcaacagg tagtcggtcg 63360
 ccggccgcag gcggaccccg gcgccctcgg cggagaacag ttctcgaag accgcggcca 63420
 ggtgccggtt ctgggagatc tgggacatga gcaggccgat gagcttgccg ctgatgatga 63480
 cgtcggtccc ggggccgatg ggggccagcg cccggttgcg gtcgtcgatc agttcgggtga 63540
 cgaccggcag ctcggcgccg gtggttctt ccagttggcg caggagcaga agggtgacga 63600
 gcgtgcggtt gtcgggatcg tccggcggtt ggcccggggc ggggtccgc cccagcacga 63660
 tcacgctgtc gtaggagtgg acgtccaggc gccgcagcgt ctcgggacgg gtgatgtccc 63720
 cgtggtgcag agccaggctc aggcggttc caacgttctc cccgtttccg ctgtccgctt 63780
 cggcctcgct gatctcgcg atcgtcgct cgcgcggttc cgcacgacg tcgacggccg 63840
 aaccgggccg ggcgcgtcgg tgcaactggt cgaccacgag cggcgcctcg cggttccagc 63900
 ccagtagcag aatccgctcc gccggcgcgg gcgtcgagg gcgggaggcc accgcggcct 63960
 tctcgaccga ctccgcacag tcgtccagcc gggccgtgtc gtcgtcgccg gtgatgacga 64020
 cgagcaggtc gtcggggcg accggcgctc tcggcgcgcg gttgagcaag ggggtgcagc 64080
 cgcgcacacg tccgacgacg ctgctcgctc agtaggacag gagaacctcg ccgaacgggc 64140
 ggccggtcag ggcggctcgt ctgatcagat agaactcgtc tccggcgaag tcgaggagtt 64200
 cccggtggac gagggagatc ccggggcggc gggcgccctg gacgatcagc cgggcggtga 64260
 cgggtgtcact ctccaggacg acgcgctcgg gtccggcggc gagacaggcg gccaggcggt 64320

ES 2 287 552 T3

accggtcgtc ccggacggcg gcgacgacgg gcggacgagg ttctgccccg gccagagcgg 64380
cccgcagcgc cagcagtgtc ttcaccacct ccgcgtcggc gtgcggctcg tccggaggca 64440
5 gaaccagcac gacaccggcc gtggccgggc tggtaaacgg caacacggcc gggtcgggtg 64500
tgggggccgct gcggcagatc aaccgcgtac cgccgcaggga gcccaccttc gtgcccaggg 64560
10 actcctccat gacggtcttg tcccggtcgg ccagcaccac caccgcgcgc ccgcgtgtgt 64620
tgacgttggc ggccaccagc tcgtccacca ccgtgaagac ctgttccgac catccgaggga 64680
ccacggcgtg cccctgttcc agcacgggtg aacggccccg ccgcaacgag gtgagccgct 64740
15 ccgtgagcgc cgtcgtgatc aggcgcacga gcgtggagac gtagagcagg gtgaccagcg 64800
cgagcagcac cgacaacatc gcccgcagcg gcgtacccgt ggcaccgccc agccgtagcg 64860
tctcccgggt gagacgccac acctccgcca gccgctccgc gagggacggc ggggcatccg 64920
20 ggtcgggtcca caccatcacc gcgtggcccg gcacgacgac ggccagggac agcagcgcca 64980
tccagccgac gagcgcggcg gcaccgcggg ccagggtgct gtcgaaccag taacggggccc 65040
25 tgtcaccgaa cggagtccgc cgtcgcgcca cccgtcccc ttccgtctcc ccgtactccc 65100
accaggccgc gtcggcaccg cagcggcgcc gtgcgcgcgc ggccctgccc ggagcgcgga 65160
cgtgggcacg catgtcgcag tgtgggggat cgatcagggc cgacgacgac gttcgaccag 65220
30 ccttcatccc ttccggtagc cgcgtcctcc acggcggcgg acctgcgaag acgcccggct 65280
cggacacgca gatgaaagcc gccgggcctc attcgcggaa cgccgggtat ccgcgagacc 65340
ggatcacgct cacacacagc gaggagagac cctgtgccgt ccaccgatgt cgtcgaactc 65400
35 atcctgcggg accaccgcgc tatggaggaa ctgttccgca ccctgcgcaa cgtcgaagcg 65460
gaccgtgccg cggccctgac ggagtccgcg gacctgctca tcgcgcacgc ctcgcccgag 65520
40 gaggacgagg tctaccccg cctgcgtcgg tacaagaacg togaggggtga ggatgtcgac 65580
cacagcgtcc acgagcacca cgaggccaac gaggcgctgc tggccctgct cgaggtggag 65640
gacaccgctt ccgacgagtg ggatgacaag ctcgaagagc tggtcacggc ggtcaaccac 65700
45 cacgccgacg aggaggaacg aacactcctc aacgacgccc gggagaacgt cgccgacgac 65760
cgccgcgggg aactggggca gaagtccag gaggcgctt cgcggtatct ggagaccggc 65820
50 tgcggcagtg tcgagaacgt gcgcaagctg gtcgcgcgcg ccgacgactg acccgcgctc 65880
gcgacgtccg ggcgcggagg ggagccgcgc ccgtcgggcc ccctcgccgg gcgtaccgcg 65940
gtcaggcggg tgagggtgc gtcgggtccg ggacgggttc gaggcgagc acgatgccct 66000
55 tggacgtggg ctgattgtg atgtccgcga cgctgtccag cggtagcagg acgttggtct 66060
cgggatatga ggcggcgccc gagcccttgg ccgcgggta gggaaacgacc tggaaagtct 66120
cgcccgggcg ctcggtgccg tccgcccaga cgctcacgag gtcgacgcga tcgccctggg 66180
60 cgaggccgag ttcgctcagg tcggccgggt tgacgaggac gacgtggcgg ctgccgtgga 66240

65

ES 2 287 552 T3

tgccgcggta gcgatcggtg tcggtgtagg ggacgggtgt ccactgggtc tgcgaaacga 66300
 gtgtctgcag cagcagatga ccttcggggc ccggtgggac cacgctctcg ttgcgagtga 66360
 5 acagggcctt gccgacctcg gtgttgaaga cgccttcgtt gaccgggttg ggcagttgga 66420
 agccaccggg ccgggtcacg cgtgcgttga agtcgtggaa gccgggcacg atgcgcgcga 66480
 10 tgccggtcgc gatggtgttg tagtcgcctt cgaaggcttc ccaggggatc tccacctgc 66540
 cgtccagggt gagccgggag agccggcaca ggatcgcgat ctgctcagc agcatggggg 66600
 aggcgggggc caggcggccc cgggaggtgt gcacctcgct catggagttc tcgacggtga 66660
 15 cgaactgctc gccgtcggcc tggacgtcgc gtcgggtgcg tcccagcgtc ggcaggatca 66720
 acgcggtgct accgcagacg gtgtgcgagc ggttgagctt ggtcgagatg tgggcggtca 66780
 gccggcacga gcgcatcgcc tcctcggtga cctcgctgct gggcgccgcc cggacgaagt 66840
 20 tcccgccag ggcgaggaag accttgacgc ggccctcgcg cattgccttg atcgagttca 66900
 ccgagtcag gccgtgggac cgcggcggtc cgaagccgaa ctgctccgc agggcgctca 66960
 25 ggaagggtgc cggcatctgc tcccagatgc ccattggtgc gtcgccctgg acgttgctgt 67020
 ggccgcgcac cgggcaggcg ccggtgccg cgcgtcccag gttgccgcg agcatcagga 67080
 agttgacgat ctcccgacg gtgggcacgc cgtgcttggt ctgggtgatg cccatcgccc 67140
 30 agcagacgac gacgcgttcg ctgtcgagga cctcgtcgcg taccttctcg atctcctcgc 67200
 gggtcagtc ggtcgccgag cgacgtcgt ccagtcac cgtgcgggag tgccgggaga 67260
 actcctcgaa gccggtggtg tgggcgtcga tgaagtcgtg gtccaggacg gtgccgggac 67320
 gggcgctctc ggcctccagc agcagtcggt tgaggccctg gaagagggag aggtcgccac 67380
 cgggcttgat gtgcaggaaa cggtcggcga tccgggtgac gcgcccagc accccgcgag 67440
 40 gctgctgcg gttcttgaag cgtcgacgc cggcctcgag aagcgggttc acggccacga 67500
 tccgggcgcc gttccgcttg gctcctcca gcgcgtgag ctggcgcgag tggttgctgc 67560
 cgggggtctg cccgaccag aagatcagat cggcgtggtg gaggtcgtcg agaccgacg 67620
 45 tgcccttgcc ggtcccagc gtctcgctca gggcgaagcc gctggactcg tggcacatgt 67680
 tgctgcagtc gggcaggttg ttggtgccga aggcgcgggc gaagagctgc agcacgaagg 67740
 50 cggcctcggt gctggcgcg cccgaggtgt agaacaccgc ctgctcgagg gaggccagcg 67800
 acttgagctc ctccgcgag acccccagg cgtcgctcca gccgatgggc tcgtagtgcg 67860
 cggagccggg ccgtttgatc atcggtcgg tgagccggcc ctgctggttg agccacatgt 67920
 55 cggagcggcc ggcgaggtcg gcgacgtgt gctcgcgaa gaaqtcggcg ctccaccgcc 67980
 gtgtggctcg ctgctcgttg atgtgcttg cgcggttct gcagtactcg ttgcggtggc 68040
 60 gccgtcccg ggcgggtcc gccacgcgc agccggggca gtcgatgcc cccacctggt 68100
 tcatggtcag cagatccacc ccggtcctgc gcggggacgt ctgctccaag gactactcca 68160

65

ES 2 287 552 T3

gcgcgtgcac gaccgcgggg acgcccggcg cccacttctt cggcgggtgtc acggagagggc 68220
 5 tgggtctccga ctctcaccg tgcggcttct gcatgtgttc gcctttctct cgcggtgtcc 68280
 ggccctcggc cagccgacgg gccaccgggc caccctgtct ctgggcgggt ccgggtcgatc 68340
 gtgttgacag cggccgttgc ggatgggtgc gcgccaggcg ccgccctctt gtcccagtc 68400
 10 ctccatgaac cgtttgaatc tcgtcagctc gccgtggacc agtcgggtgg tcagccggac 68460
 cgcgcgcgag gaacgggtga ggatcctcgc ggctccccgc ggttccagaa gcaccggac 68520
 ggtgatcgcg gtgccgcgg actccgtcgg cctgaactcg acctctcccc ggtgccacgg 68580
 15 cctctgtctc aggccgcgac aggccaggta ggcgtcgggg tcctgtctca ggatctcgac 68640
 cgcgaagcgg cggcgcaggg ggccgtaacc gagggtcag gcggtcacgg tgggccggac 68700
 20 ctgctcgacg tcgccgacca cggccgagaa ccgagggag gacttgaact gcgtccactg 68760
 gttgtacgcc gtccgtacgg gcaccgccac ctcgaccgtc tggtcgacgt gccgtcgcg 68820
 cagcggacga cggcccggtc tcctgtcacc gctggtgcgc gcccgcgggt tgtccggcac 68880
 25 ggcccgcggg ccctcgtgct gctccgccat cgtggtctct cctcctgtgg ggggaaccag 68940
 gcgtccaggc tctcgacgtc tcctcgtcgc tcagggtccc ctctgtcgtc aggggttttc 69000
 cctcccggtt cccacctccg gcgtttcgag tcacctcggc cccgcccgtt cctcgtgcgc 69060
 30 gtcggcctcc gacgccccgc ccgagagat ccggctcgac gtcggcctgg tcgtccggc 69120
 gggcggtgag cagcaggtac tcccagtcga tggccccgtc ccgcagcgcg ccggccgcga 69180
 35 ggtcggtgag ggcccgctcc agggcgggcg cgcgtcgggt gtcgtcgcgc atgtaccggt 69240
 agacggcgat gatcgccccg taggcggcct tgaagaactc aaggaagtc ccccttcgag 69300
 cggcgctcgg gaccgggggt gcgccggccg gcctgatgac cagcgtcgac ggggtcgtga 69360
 40 gacggtgatc gaggtgctcg acgagccgga gccaccgag gagagtcgcg ggcagcgtc 69420
 tctccagcaa cccagccttc gcacaccgt gcgcggcgca aggggcgtac gtcccctgcg 69480
 45 caaagggcg gctggaggcc ggtgccccac gcctcgtctc cgcacggcgc gtcggacggg 69540
 gggagaacgc gtccgatgaa cacggtgtgc gccctcccga ttctctccg tgacgtcaat 69600
 gatgagccca ccgcgctcgg gtcaggcgat ccgcggtccc ctgccgtcgc gtcaggggcc 69660
 50 ggtgacggaa ggagtcacg gtgttgcttc tcctctctcc ggacgggtgc gaggaagccc 69720
 tcgactgcgc gaaggcgcg gagcacctcg acatcgtcga cgtgaagaag cccgacgagg 69780
 55 gctcgtcggg cgcgaacttc ccgtgggtca tcaggagat ccgcgacgcg gtgccggcg 69840
 acaagccggt ctcgccacc gtgggggacg tccgtacaa gcccggcacg gtggcgacgg 69900
 ccgcgctggg cgcggctcgc tcgggggcca cgtacatcaa ggtcggcctc tacggatgca 69960
 60 cgacgccga acagggcatc gcggtcatgc gcgcggtggc ccgggcgggt aaggaccacc 70020
 gtcccgaagc gctcgtcgtc gcgtccggtt acgccgacgc ccaccggatc ggctgcgtca 70080

65

ES 2 287 552 T3

accggtcgc cctgcccgaac atcgccgccc gctccggcgc cgacgccgcg atgctcgaca 70140
 ccgcggtcaa ggacgggacg cggtctgtcg atcacgttcc gccggacacc tgcgccgagt 70200
 5 tcgtccgtcg cgcgccagcc gccggcctgc tcgccgccct cgcgggcagc gtcaggcaga 70260
 ccgacctcg cgggctgacc cggatcgga cggacatcgt cggggtgcgc ggagcgggtct 70320
 10 gcgagggcgg cgaccgcaac gccggacgca tccggccgca cctggtggcc gccttcggga 70380
 gcgagatgga ccggcacgcc cgcgagcacc gggccggcgt caccaccgcg agctgaccgc 70440
 cggtatgccg accccgcac ccgaccagc ccccgcacag cgggcccgc ctctcgcggt 70500
 15 cgtcgatccg gccaccgga cggctcttga cgaggcccc gaccaggac cggacgtgt 70560
 ggacgccgtc gtcgaccggg ccgcggggc ctggcacggc tggcgcgccg atcccagcgc 70620
 ccgtaccacc gcgtgcgt cgggcgccga cgcggtcgag gccgccggg acgacctcgc 70680
 20 ccgtctctc acccgggaac agggaaagcc cctggccgaa tcgcatgcgg aggtcgccc 70740
 gacggcggcc cgcctgcgt acttcgccg cctggcccc cggaaccggc gcatcaccga 70800
 25 cggcgggccg gtgcgcagc aggtccgctg ggcgccctc ggaccgcgc ccgcatcgt 70860
 gccgtggaac tccccctcc aactcgcgc ggcgaagtgc gcgccgcgc tcgcccggg 70920
 caacaccatg gtccctcaa cctccccct caccgcgc gccaccggc tgctcggtc 70980
 30 cgtcctcgc accgccctgc ccgaggacgt cctgacggtc gtcaccggcc gcgagccact 71040
 cggcgcccgc ctgcgcgc accccggcat ccgccagtc accttcaccg gatcggtgcc 71100
 35 cacggggcgg gccgtgcgc gagcggggc ggcctcgc gcccggtca ccctggaact 71160
 gggcggaac gacgccgcg tctgtctgga cgacgtgaa gtggaccgga tcgccgaccg 71220
 gctgttctgg .gccgcgttcc gcaactgcg gcaggtctgc atggcggtca agcgcgtcta 71280
 40 cgaccggcc cgtctgcac cacaggtcgt cgaagccctc accgagcgc ccaaggccgt 71340
 cgcgctcgg cccggcctc accccgcac ccggtggga ccggtcgca acgccccca 71400
 45 gctggcccgg gtcgagcaga tcaccggcg gccttgccg gacggcgccc gggcggggc 71460
 cggcggccac cggctggac ggccgggctg etttctcgc cccacgatcc tcaccgacgt 71520
 cccgcccga agcccggtg tgaccgagga gcagttcgg ccggtactgc cgtgctgcc 71580
 50 gtaccggagc ctggacgaag ccgtcgacgc ggccaacgg accggattcg ggtgggggg 71640
 ctccgtatgg ggcaccgacc tcgaccgggc cgaggcggtg gccgaccggc tggaatgcgg 71700
 55 cacggcctgg gtcaaccacc acgcgagct gtccctcgc cagcccttc cggcgacaa 71760
 ggacagcgg gtcggcgctc cggcggggc gtggggactg tacggcaacc tccgtccgtt 71820
 cgtcgtccac cgaccgggg gggagtgcg gtgagcttcc gggcgccgt actgcgcgg 71880
 60 tacgaggacc ccttcacggt cgaggaggtg accctgggga cggagcccgg cgcaggggag 71940
 atcctggctg agatcgccg ctgcggaatg tgccggaccg atctcgcggt ccggcgctcg 72000

ES 2 287 552 T3

gccggccgga gccgcgtgcc ggccggtgctc ggccacgagg gctccggggt ggtggtgcgg 72060
 acggggcggcg gcccgacac cgcgatcggc gtcggtgacc acgtggtgct gagcttcgac 72120
 5 tccctgcgggc actgccgcaa ctgccgcgcg gcggcccccg cctactgtga ttccttcgcc 72180
 tccctcaacc tcttcggggg ccgtgcggag gaccgcgcgc ggtcaccga cgggtcgggg 72240
 10 gcggcactgg ctcgccggtg gttcggacag tccgccttcg ccgagtacgc gctcgtctcc 72300
 gcccgcaacg ccgttcgggt cgaccccgcc ctgcccgtc aactgctcgg gccgctgggc 72360
 tgcggcttcc tcaccggagc cggagccgtg ctcaacacct tcgccgcgg gccgggcgac 72420
 15 accctcgtcg tgctcggcgc gggcgccgtg ggccctggcc cggatgatgg gccaccgcc 72480
 gccggcgcac cgtccgtggc cgtggaccgc aaccccgct gcctggagct ggccgagcgg 72540
 ttcggcgcg tcccgctgcc cgcgcgacg gccggactgg ccgagcggat ccggcggtc 72600
 20 acggacggcg gcgcgcggtg cgcactggac acgaccgct ccgtccact gatcaacgag 72660
 gcgctgcgc cactgcgtcc caaccgcgt ctccgcttg tggcacggct ccacaccgcg 72720
 ctgcccttg aaccgggcac gctcgaccgg gggcgacga tccgccacgt ctgcgagggg 72780
 25 gacgcggtac ccggtctgct gataccgacg ctgaccgcgc tctggcaggc cggacgcttc 72840
 cccttcgacc agctcgtccg tacctaccgc ctggccgaca tcaacgaggc ggagcgcgac 72900
 30 tgcgaaccg gccctcgtgt caaacccgtg ctgctccgc ccgcgaggag ccggtgagta 72960
 cggcgacagg caccgcggtc cgaccgcac cgacgagcag gaagctcgcg gcccacttc 73020
 cgccaacgga ggagacatga ccggcacggc gccgcagtac acggacgtgg aaggcgtgaa 73080
 35 cggagggtgtg ggccgtacgg ccttcctggt cgcgcgcgcg cgggcgatcg agaccatcg 73140
 cgacgacagt ctggccagc agtctacgc ggaacacttc gtgcgcgcg ccccgcgctg 73200
 40 cgcggactgg ccggtgcgca tcgagcaggt ccccgacggg gacggcaacc cgctgtgggg 73260
 acggttcgcc cgctatttcg gcctgcggac ccgggcccctc gacgacttcc tgcctccgtc 73320
 ggtccggacg ggcgccgac aggtggtgct gctggcgcg ggttggaaca cccgtgcctt 73380
 45 ccggtcgcac tggccgtcgc agtgcgcggt cttcgagatc gaccggacgg gcgtgctcgc 73440
 cttcaaacag caggtgctca cggacctggc ggcaacccc agagtggagc gcgtcccgct 73500
 50 tccggtcgat ctgcgcgcg actgggcccg cgcgctgacc gcggccggct tcgacccgc 73560
 ggcccagc gtctggctgg ccgagggact gctcttctat ctgccgggc ccgcgagtc 73620
 gcttctcgtc gacacggtg accggtgac caccgacggc agcgcgctgg ccttcgaggc 73680
 55 caagctggag aaggacctgc tggcgtaccg cgacagtgcg atctacacg cgacgcgcga 73740
 gcagatcggc atcgacctcc tcgcctctt cgacaagggg ccccgacccg actccgcggg 73800
 tgagctggcg gccagaggct ggtccacctc gatgcacacg cccttcgtct tcacccaccg 73860
 60 gtacggacgc ggtcccctcc ccgagccgaa cgacgcgctg gagggaacc gatgggtctt 73920

65

ES 2 287 552 T3

```

cgcccgaag cccgggccct gacgtgccgg ccgcgcttgc cgcccacgcc cggggacgcc 73980
gctgacgagc ggtgtcagac ggtccgggcg gccaccagcc ggtcgcgccg gatcccgccc 74040
5 aggtccggcg cgtagtagtc ctcgatgccg gagggccacg cggccacggt gatgcggtcc 74100
gcggcgtccg gggcgaaggc gtcgaacagc gcgtggatgt tcgcttcttc gaaaccgata 74160
gccttcatca gcgcgacgaa gaggggacgc tcgatcagcc cgtccccgtc gggatcgccg 74220
10 agcgcgagca gtgcccagac gaactcggcg atggtgggac cgaagcgctc gggatcgagt 74280
acgaacggac ggaactcctc cacggtgate acgccgtcgc cgtcggcgtc cagttcgggtg 74340
15 gccagcgtgg tccagtagcg gcggaacgcg gcccgacggc cggccttggc gctgtcgtcc 74400
gacccggccg ccgcgcgaac gaccgggtcg gtcacaggt cgaagtcgtc ggagtcgatg 74460
actccgttgc cgttggcgtc gaagaggag aagaccagct cgaccgcctt ggcggccctca 74520
20 tctcgcagtc acatcacctg tcttctacgg ccgggtcttc gcggggcccg ccccatgaat 74580
gctctgcgtg accgagcggg gcaggacgaa agcctccgag cggtcgcgtc ccagagaacc 74640
accatgaatg tccctgaact gcagatcggg catctgctgg cctggtgcgg gcgggggctg 74700
25 gcccggtcgc gcaggggagt gctctggtgc ctgggcaagg ccgtcacggg gatcatcctg 74760
ctcgccatct tcgcgtccgc gatgate 74787

```

30 <210> 2

<211> 876

<212> PRT

35 <213> *Streptomyces parvulus* Tü4055

<400> 2

```

Met Thr Gly Ser Ala Val Ser Ala Pro Phe Leu Gln Pro Pro Glu Pro
40      1              5              10              15

Val Ser Gly His Ser Glu Arg Lys Ser Asp Pro Val Leu Leu Val Gly
      20              25              30

45 Ala Gly Arg Arg Ala Arg Met Ala Asp Ala Val Arg Ala Ala Gly Ala
      35              40              45

Gln Ala Gly Ile Asp Pro Ala Val Leu Arg Arg Thr Arg Ala Thr Leu
50      50              55              60

Ile Thr Ala Gly Ser Ala Gly Ala Ala Gly Arg Leu Ala Ala Ala Leu
65      65              70              75              80

Arg Leu Thr Gly Ala Thr Ile Ser Leu Asp Thr Arg Glu Thr Pro Thr
55      85              90              95

Leu Leu Ala Leu His Leu Ala Ala Gln Ala Leu Arg Ala Gly Asp Thr
      100              105              110

60 Ser Tyr Ala Val Val Gly Ala Glu Leu Pro Asp Gly Asn Cys Ala Leu
      115              120              125

Ile Leu Ala Arg Gln Ser Ala Ala Thr Ala Glu Gly Ala Val Pro Gln
65      130              135              140

```

ES 2 287 552 T3

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

ES 2 287 552 T3

					485					490					495	
	Asn	Asp	Leu	Thr	Asp	Gln	Gly	Ile	His	Ala	Lys	Arg	Ile	Pro	Gly	Val
					500				505					510		
5																
	Asp	Thr	Val	Gly	His	Cys	Ser	Gln	Val	Glu	Val	Leu	Arg	Asp	His	Leu
			515					520					525			
	Leu	Asp	Val	Leu	Arg	Pro	Val	Ser	Pro	Arg	Pro	Ala	Ala	Val	Pro	Phe
		530					535					540				
10																
	Tyr	Ser	Thr	Val	Asp	Gly	Thr	Glu	Arg	Asp	Thr	Thr	Thr	Leu	Asp	Thr
	545					550					555					560
	Asp	Tyr	Trp	Tyr	Leu	Asn	Thr	Arg	Ser	Gln	Val	Arg	Phe	His	Gln	Ala
					565					570					575	
15																
	Val	Arg	Asn	Leu	Leu	Ala	Ala	Gly	His	Arg	Ser	Phe	Val	Glu	Val	Ser
				580					585					590		
	Pro	His	Pro	Leu	Leu	Gly	Ala	Ser	Ile	Glu	Asp	Thr	Ala	Ala	Glu	Phe
			595					600					605			
20																
	Gly	Leu	Asp	Asp	Val	Ala	Ala	Val	Gly	Thr	Leu	Arg	Arg	Gly	Gln	Gly
		610					615					620				
	Gly	Thr	Arg	Arg	Val	Leu	Thr	Ser	Val	Ala	Glu	Ala	Tyr	Val	His	Gly
	625					630					635					640
25																
	Ile	Asp	Ile	Asp	Phe	Thr	Pro	Ala	Phe	Thr	Gly	Thr	Thr	Pro	Asn	Arg
					645					650					655	
	Ile	Asp	Leu	Pro	Thr	Val	Glu	Asp	His	Gly	Ile	Glu	Gly	His	Gly	Asp
				660					665					670		
30																
	Asp	Gly	Gly	Glu	Thr	Trp	Thr	Asp	Arg	Val	Arg	Thr	Leu	Pro	Asp	Glu
			675					680					685			
	Gln	Arg	Glu	Glu	Ala	Leu	Leu	Asp	Leu	Val	Cys	Arg	Thr	Val	Ala	Ala
							695					700				
	Val	Leu	Glu	Ala	Asp	Pro	Ala	Gly	Thr	Ala	Asp	Ala	Val	Ala	Pro	Asp
	705					710					715					720
40																
	Thr	Ala	Phe	Lys	Glu	Met	Gly	Leu	Gly	Ser	Leu	Ser	Ala	Val	Arg	Leu
					725					730					735	
	Arg	Asn	Gly	Leu	Arg	Glu	Ala	Thr	Gly	Ala	His	Leu	Pro	Ala	Thr	Ile
				740					745					750		
45																
	Ala	Tyr	Asp	His	Pro	Thr	Pro	Ala	Ala	Leu	Ala	Arg	His	Leu	Ala	Met
			755					760					765			
	Thr	Leu	Phe	Asp	Ala	Thr	Gly	Ala	Ala	Pro	Ala	Val	Pro	Ala	Pro	Ser
		770					775					780				
50																
	Arg	Asp	Asp	Glu	Pro	Ile	Asp	Ala	Glu	Thr	Ala	Val	Leu	Thr	Ala	Leu
	785					790					795					800
	Glu	Arg	Ala	Asp	Glu	Ala	Leu	Glu	Arg	Leu	Arg	Ala	Pro	His	Ala	Arg
						805				810					815	
55																
	Thr	Pro	Arg	Gln	Glu	Thr	Gly	Arg	Arg	Ile	Asp	Glu	Leu	Leu	Arg	Ser
				820					825					830		
60																
	Leu	Thr	Asp	Lys	Ala	Arg	Arg	Met	Arg	Gln	Ala	Asp	Ala	Val	Asp	Asp
65																

ES 2 287 552 T3

	835	840	845
5	Val Asp Asp Pro Ala Thr	Asp Arg Phe Ala Ala	Ala Thr Asp Asp Glu
	850	855	860
10	Met Phe Glu Leu Leu Glu	Lys Arg Phe Gly Ile	Ser
	865	870	875
	<210> 3		
	<211> 1571		
	<212> PRT		
	<213> <i>Streptomyces Parvulus</i> Tü4055		
15	<400> 3		
	Met Ala His Glu Asp Lys Leu Arg His Leu Leu Lys Arg Val Ser Ala		
	1	5	10 15
20	Glu Leu Asp Asp Thr Gln Arg Arg Val Arg Glu Met Glu Glu Ser Glu		
	20	25	30
	Arg Glu Pro Ile Ala Ile Val Gly Met Ser Cys Arg Leu Pro Gly Gly		
	35	40	45
25	Val Asn Ser Pro Gly Glu Phe Trp Ser Leu Leu Glu Ala Gly Thr Asp		
	50	55	60
30	Ala Val Ser Glu Phe Pro Arg Asp Arg Gly Trp Asp Val Glu Asn Leu		
	65	70	75 80
	Tyr Asp Pro Asp Pro Asp Ala Pro Gly Arg Ser Tyr Val Arg Glu Gly		
	85	90	95
35	Gly Phe Leu Asp Gly Ala Gly Gln Phe Asp Ala Ala Phe Phe Gly Ile		
	100	105	110
	Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu		
	115	120	125
40	Glu Cys Ser Trp Glu Ala Ile Glu Arg Ser Arg Ile Asp Pro Lys Thr		
	130	135	140
45	Leu His Gly Ser Arg Thr Gly Val Phe Ala Gly Ser Asn Trp Gln Asp		
	145	150	155 160
	Tyr Asn Thr Leu Leu Leu Asn Ala Glu Glu Arg Ser Gln Ser Tyr Leu		
	165	170	175
50	Ala Thr Gly Ala Ser Gly Ser Val Leu Ser Gly Arg Val Ser Tyr Thr		
	180	185	190
	Leu Gly Met Glu Gly Pro Ala Ile Thr Val Asn Thr Ala Cys Ser Ser		
	195	200	205
55	Ser Leu Val Ala Val His Leu Ala Ala Arg Ser Leu Arg Ala Gly Glu		
	210	215	220
60	Cys Asp Leu Ala Leu Ala Gly Ala Val Thr Val Met Ser Thr Pro Gln		
	225	230	235 240
	Leu Pro Val Ala Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg		
	245	250	255
65	Ser Lys Ala Phe Ala Val Ser Ala Asp Gly Met Gly Phe Gly Glu Gly		
	260	265	270

ES 2 287 552 T3

	Val Gly Val Leu Val Leu Glu Arg Leu Ser Val Ala Arg Arg Asn Gly	275	280	285
5	His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly	290	295	300
	Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val	305	310	315
10	Ile Arg Ala Ala Leu Ala Ser Ala Gly Leu Gly Pro Ala Asp Val Asp	325	330	335
	Val Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu	340	345	350
15	Ala Gln Ala Leu Leu Ala Thr Tyr Gly Arg Gly Arg Asp Ala Glu Arg	355	360	365
	Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Ala Gln Ala	370	375	380
20	Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Glu Lys	385	390	395
	Gly Arg Leu Pro Arg Thr Leu His Val Asp Glu Pro Ser Gly Glu Val	405	410	415
25	Asp Trp Asp Ser Gly Ala Val Arg Leu Leu Thr Glu Ala Arg Asp Trp	420	425	430
30	Pro Ser Glu Glu Gly Arg Leu Arg Arg Ala Gly Val Ser Ser Phe Gly	435	440	445
	Ile Ser Gly Thr Asn Ala His Val Ile Ile Glu Glu Ala Pro Glu Glu	450	455	460
35	Gly Glu Glu Pro Glu Ser Asp Ala Gly Gly Val Val Pro Trp Val Leu	465	470	475
	Ser Ala Arg Thr Glu Gly Ala Leu Gln Ala Gln Ala Val Gln Leu Ser	485	490	495
40	Glu Phe Val Gly Glu Ser Ser Pro Val Asp Val Gly Trp Ser Leu Val	500	505	510
	Ser Thr Arg Ala Ala Phe Glu His Arg Ala Val Val Val Gly Arg Gly	515	520	525
45	Arg Asp Glu Leu Val Arg Gly Leu Ser Glu Val Ala Gln Gly Arg Gly	530	535	540
50	Val Arg Gly Val Ala Ser Ser Ala Ser Gly Gly Leu Ala Phe Val Phe	545	550	555
	Ala Gly Gln Gly Ser Gln Arg Leu Gly Met Gly Arg Gly Leu Tyr Glu	565	570	575
55	Arg Phe Pro Val Phe Ala Glu Ala Phe Asp Glu Val Cys Gly Arg Val	580	585	590
60	Gly Pro Gly Val Arg Glu Val Val Phe Gly Ser Asp Ala Gly Glu Leu	595	600	605

ES 2 287 552 T3

	Asp	Arg	Thr	Val	Trp	Ala	Gln	Ala	Gly	Leu	Phe	Ala	Leu	Glu	Val	Ala	
	610					615						620					
5	Leu	Phe	Arg	Leu	Leu	Glu	Ser	Trp	Gly	Val	Arg	Pro	Gly	Cys	Leu	Ile	
	625					630					635					640	
	Gly	His	Ser	Val	Gly	Glu	Leu	Ser	Ala	Ala	Cys	Val	Ala	Gly	Leu	Trp	
					645				650						655		
10	Ser	Leu	Glu	Asp	Ala	Cys	Arg	Val	Val	Ala	Ala	Arg	Ala	Arg	Leu	Met	
				660					665					670			
	Gln	Ala	Leu	Pro	Ala	Gly	Gly	Val	Met	Val	Ala	Val	Arg	Ala	Glu	Ala	
15		675				680				685							
	Gly	Glu	Leu	Ala	Gly	Phe	Leu	Gly	Glu	Asp	Val	Val	Ile	Ala	Ser	Val	
	690				695				700								
20	Asn	Ala	Pro	Gly	Gln	Val	Val	Ile	Ala	Gly	Pro	Glu	Gly	Gly	Val	Glu	
	705				710				715						720		
	Arg	Val	Val	Ala	Ala	Cys	Gly	Ala	Arg	Ser	Arg	Arg	Leu	Ala	Val	Ser	
				725				730						735			
25	His	Ala	Phe	His	Ser	Pro	Leu	Val	Glu	Pro	Met	Leu	Gly	Glu	Phe	Arg	
				740				745					750				
	Arg	Val	Val	Glu	Ser	Val	Ala	Phe	Gly	Val	Pro	Ser	Leu	Arg	Val	Val	
				755			760						765				
30	Ser	Asn	Val	Thr	Gly	Ala	Trp	Val	Asp	Pro	Glu	Glu	Trp	Gly	Thr	Pro	
	770				775				780								
	Glu	Tyr	Trp	Val	Arg	Gln	Val	Arg	Glu	Pro	Val	Arg	Phe	Ala	Asp	Gly	
35	785				790			795						800			
	Val	Ala	Thr	Leu	Leu	Asp	Ala	Gly	Val	Arg	Thr	Phe	Val	Glu	Leu	Gly	
				805			810						815				
40	Pro	Ala	Gly	Ala	Leu	Thr	Ser	Met	Val	Ser	His	Cys	Ala	Asp	Ala	Thr	
		820			825				830								
	Ala	Thr	Ser	Val	Thr	Ala	Val	Pro	Thr	Leu	Arg	Pro	Asp	His	Asp	Glu	
		835			840			845									
45	Ser	Arg	Thr	Val	Leu	Ser	Ala	Ala	Ala	Ser	Leu	Tyr	Val	Gln	Gly	His	
	850				855			860									
	Pro	Val	Asp	Trp	Ala	Pro	Leu	Phe	Pro	Arg	Ala	Arg	Thr	Val	Asp	Leu	
	865			870			875						880				
50	Pro	Thr	Tyr	Pro	Phe	Gln	His	Gln	His	Tyr	Trp	Leu	Asp	Val	Pro	Pro	
				885			890						895				
	Leu	Phe	Thr	Ala	Ser	Ser	Ala	Ala	Gln	Asp	Gly	Gly	Trp	Arg	Tyr	Arg	
55		900			905			910									
	Ile	His	Trp	Arg	Arg	Leu	Gly	Thr	Arg	Asp	Ser	Gly	Asp	Arg	Leu	Ser	
		915			920			925									
	Gly	Arg	Trp	Leu	Leu	Leu	Val	Pro	Glu	Ser	Asp	Gly	Thr	Glu	Pro	Trp	
60	930			935			940										
	Val	Glu	Gly	Ala	Glu	Lys	Met	Leu	Ala	Glu	Arg	Gly	Cys	Glu	Val	Val	

ES 2 287 552 T3

	945		950		955		960
	His Val Pro Ile Ala	Ala Thr Ala Asp Arg Asp Ala Met Val Gly Ala					
		965			970		975
5							
	Val Arg Glu Ser Val Glu Asp Gly Arg Val Asp Gly Val Leu Ser Leu						
		980			985		990
10	Leu Ala Leu Asp Gly Arg Pro His Pro Asp Ala Ala Ala Val Pro Thr						
		995		1000		1005	
	Gly Leu Val Ala Thr Ala Gln Val Val Gln Val Ser Asp Glu Leu Gly						
		1010		1015		1020	
15	Ile Gly Pro Leu Trp Val Ala Thr Arg Gln Ala Val Ser Val Asp Gly						
		1025		1030		1035	1040
	Ala Asp Glu Ala Asp Gly Ala Gly Arg Thr Arg Lys Ala Asp Asp Pro						
		1045		1050		1055	
20							
	Ala Asp Val Ala Gln Ala Ala Val Trp Gly Leu Gly Arg Val Ala Ala						
		1060		1065		1070	
	Leu Glu Lys Pro Arg Leu Trp Gly Gly Leu Val Asp Leu Pro Ala Arg						
25		1075		1080		1085	
	Ala Asp Glu Arg Met Arg Asp Leu Val Ala Gln Ala Leu Thr Ala Pro						
		1090		1095		1100	
	Asp Ala Glu Asp Gln Leu Ala Val Arg Ala Asp Gly Ile Ala Val Arg						
30		1105		1110		1115	1120
	Arg Leu Val Arg Ser Ala Ala Ser Ala Pro Ala Asp Asp Trp Gln Pro						
		1125		1130		1135	
35	Ser Gly Thr Val Leu Val Thr Gly Gly Thr Gly Gly Val Gly Ala Asn						
		1140		1145		1150	
	Val Ala Arg Trp Leu Val Thr Gln Asp Ile Gln His Leu Leu Leu Val						
		1155		1160		1165	
40							
	Ser Arg Arg Gly Pro Asp Ala Pro Gly Ala Ala Glu Leu Leu Ala Glu						
		1170		1175		1180	
	Leu Ser Ala Ser Gly Thr Ser Val Thr Ile Glu Pro Cys Asp Val Thr						
45		1185		1190		1195	1200
	Asp Ala Asp Ala Val Arg Arg Leu Ile Gly Ala Val Pro Ala Glu Arg						
		1205		1210		1215	
	Pro Leu Ser Thr Val Val His Ala Ala Gly Val Leu Asp Asp Cys Leu						
50		1220		1225		1230	
	Ile Asp Ala Leu Thr Pro Gln Arg Leu Ala Ala Ala Leu Glu Val Lys						
		1235		1240		1245	
	Ala Lys Gly Ala Leu Asn Leu His Glu Ala Ala Gly Glu Ala His Leu						
55		1250		1255		1260	
	Val Leu Phe Ser Ser Leu Ala Gly Thr Thr Gly Thr Lys Gly Gln Gly						
		1265		1270		1275	1280
60	Asn Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Arg Arg						
		1285		1290		1295	

65

ES 2 287 552 T3

Arg Ala Asp Gly Leu Pro Ala Thr Ser Val Ala Trp Gly Ala Trp Gln
 1300 1305 1310
 5 Gly Ala Gly Met Val Ala Asp Ala Ala Val Ala His Arg Thr Arg Arg
 1315 1320 1325
 Tyr Gly Leu Pro Leu Met Ser Pro Asp Arg Ala Val Ala Thr Leu Arg
 1330 1335 1340
 10 Gln Val Met Ala Glu Pro Val Ala Thr Gln Val Val Ala Asp Val Asp
 1345 1350 1355 1360
 Trp Gln Arg Phe Val Ala Asp Phe Thr Ala Val Arg Pro Ser Arg Leu
 1365 1370 1375
 15 Leu Ala Asp Leu Pro Glu Val Arg Ser Leu Gly Glu Gln Arg Lys Asp
 1380 1385 1390
 20 Gly Pro Gly Gly Gln Gly Glu Glu Asp Gly Leu Ala Ser Lys Leu Ala
 1395 1400 1405
 Ala Leu Pro Glu Ala Asp Arg Arg Arg Ala Val Leu Asp Leu Val Glu
 1410 1415 1420
 25 Glu Leu Val Leu Gly Val Leu Gly His Glu Thr Arg Ala Ala Ile Gly
 1425 1430 1435 1440
 Pro Asp Ser Ser Phe His Ala Ile Gly Phe Asp Ser Leu Thr Ala Val
 1445 1450 1455
 30 Glu Leu Arg Asn Leu Leu Thr Val Arg Leu Gly Met Lys Leu Pro Ala
 1460 1465 1470
 35 Thr Leu Val Tyr Asp His Pro Thr Leu Ser Ser Leu Ala Asp His Leu
 1475 1480 1485
 His Glu Gln Leu Val Ile Asp Gly Thr Pro Met Thr Asp Thr Ala Ala
 1490 1495 1500
 40 Asp Leu Leu Ala Glu Leu Asp Ala Leu Ala Ala Arg Leu Ala Ala Val
 1505 1510 1515 1520
 Gly Leu Glu Pro Glu Ala Arg Ala Arg Ile Gly Arg Arg Leu Lys Asp
 1525 1530 1535
 45 Met Gln Thr Ala Cys Glu Pro Arg Ser Glu Ser Ser Arg Asp Leu Lys
 1540 1545 1550
 50 Ser Ala Ser Arg Thr Glu Val Leu Asp Phe Leu Thr Asn Glu Leu Gly
 1555 1560 1565
 Ile Ser Arg
 1570
 55 <210> 4
 <211> 3500
 <212> PRT
 60 <213> *Streptomyces parvulus* Tü4055
 <400> 4
 65 Met Pro Asn Asp Glu Glu Leu Leu Asp Tyr Leu Lys Arg Thr Ala Ser
 1 5 10 15

ES 2 287 552 T3

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

ES 2 287 552 T3

	355	360	365
5	Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Ala Gln Ala 370 375 380		
	Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Lys Ala Met Gln Ala 385 390 395 400		
10	Gly Thr Leu Pro Arg Thr Leu His Val Asp Glu Pro Ser Gly Glu Val 405 410 415		
	Asp Trp Asp Ser Gly Ala Val Arg Leu Leu Thr Glu Ala Arg Asp Trp 420 425 430		
15	Pro Ser Glu Glu Gly Arg Leu Arg Arg Ala Gly Val Ser Ser Phe Gly 435 440 445		
	Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Glu Pro Pro Ala Glu 450 455 460		
20	Asp Ala Val Pro Glu Pro Glu Ala Gly Asp Val Val Pro Trp Val Leu 465 470 475 480		
	Ser Ala Arg Ser Ala Glu Ala Leu Arg Glu Gln Ala Ala Arg Leu Ala 485 490 495		
25	Ser Val Ala Gly Gly Leu Asn Val Val Asp Val Gly Trp Ser Leu Ala 500 505 510		
	Ser Thr Arg Ala Ala Phe Glu His Arg Ala Val Val Val Gly Arg Glu 515 520 525		
30	Arg Glu Glu Leu Leu Ala Gly Leu Phe Ala Val Ala Ala Gly Arg Pro 530 535 540		
	Ala Ala Asn Val Val Thr Gly Pro Val Ser Ser Gly Arg Pro Ala Phe 545 550 555 560		
35	Val Phe Ala Gly Gln Gly Ser Gln Arg Leu Gly Met Gly Arg Gly Leu 565 570 575		
40	Tyr Glu Arg Phe Pro Val Phe Ala Glu Ala Phe Asp Glu Val Cys Gly 580 585 590		
	Arg Val Gly Pro Gly Val Arg Glu Val Val Phe Gly Ser Asp Ala Gly 595 600 605		
45	Glu Leu Asp Arg Thr Val Trp Ala Gln Ala Gly Leu Phe Ala Leu Glu 610 615 620		
	Val Ala Leu Phe Arg Leu Leu Glu Ser Trp Gly Val Arg Pro Gly Cys 625 630 635 640		
50	Leu Ile Gly His Ser Val Gly Glu Leu Ser Ala Ala Cys Val Ala Gly 645 650 655		
	Leu Trp Ser Leu Glu Asp Ala Cys Arg Val Val Ala Ala Arg Ala Arg 660 665 670		
55	Leu Met Gln Ala Leu Pro Ala Gly Gly Val Met Val Ala Val Arg Ala 675 680 685		
60	Glu Ala Gly Glu Leu Ala Gly Phe Leu Gly Glu Asp Val Val Ile Ala 690 695 700		

65

ES 2 287 552 T3

	Ser Val Asn Ala Pro Gly Gln Val Val Ile Ala Gly Pro Glu Gly Gly	705	710	715	720
5	Val Glu Arg Val Val Ala Ala Cys Gly Ala Arg Ser Arg Arg Leu Ala	725	730	735	
	Val Ser His Ala Phe His Ser Pro Leu Val Glu Pro Met Leu Gly Glu	740	745	750	
10	Phe Arg Arg Val Val Glu Ser Val Ala Phe Gly Val Pro Ser Leu Arg	755	760	765	
	Val Val Ser Asn Val Thr Gly Ala Trp Val Asp Pro Glu Glu Trp Gly	770	775	780	
15	Thr Pro Glu Tyr Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe Ala	785	790	795	800
	Asp Gly Val Ala Thr Leu Leu Asp Ala Gly Val Arg Thr Phe Val Glu	805	810	815	
20	Leu Gly Pro Ala Gly Thr Leu Thr Ser Met Val Ser His Cys Ala Asp	820	825	830	
	Ala Thr Ala Thr Ser Val Thr Ala Val Pro Thr Leu Arg Pro Asp His	835	840	845	
25	Asp Glu Ser Arg Thr Val Leu Ser Ala Ala Ala Ser Leu Tyr Val Gln	850	855	860	
	Gly His Pro Val Asp Trp Ala Pro Leu Phe Pro Arg Ala Arg Thr Val	865	870	875	880
	Asp Leu Pro Thr Tyr Pro Phe Gln His Gln His Tyr Trp Met Met Asn	885	890	895	
35	Thr Gly Ser Ala Ala Glu Pro Ala Glu Leu Gly Leu Gly Asp Ala Arg	900	905	910	
	His Pro Leu Leu Gly Ser Val Val Thr Val Ala Gly Asp Asp Lys Val	915	920	925	
40	Val Phe Ala Gly Arg Leu Ala Leu Arg Thr His Pro Trp Leu Ala Asp	930	935	940	
	His Thr Val Leu Asp Ala Val Leu Leu Pro Ala Thr Ala Phe Leu Glu	945	950	955	960
45	Leu Ala Val Arg Ala Gly Glu Glu Val Ser Cys Pro Val Val His Asp	965	970	975	
	Leu Thr Leu His Arg Pro Leu Val Val Pro Glu Arg Gly Ala Val Gln	980	985	990	
	Val Gln Met Ala Val Gly Ala Pro Glu Ala Asp Gly Arg Arg Glu Val	995	1000	1005	
55	Arg Val Tyr Ser Arg Pro Asp Asp Asp Ala Glu His Glu Trp Thr Leu	1010	1015	1020	
	His Ala Ala Gly Leu Leu Ala Ser Ala Ala Thr Ala Glu Pro Ala Val	1025	1030	1035	1040
60					
65					

ES 2 287 552 T3

Ala Ala Gly Ala Trp Pro Pro Pro Glu Ala Gln Ala Val Asp Leu Asp
 1045 1050 1055
 5 Gly Phe Tyr Ala Gly Leu Ala Glu His Gly Tyr His Tyr Gly Pro Leu
 1060 1065 1070
 Phe Gln Gly Val Arg Ala Ala Trp Arg Leu Gly Asp Asp Val Leu Ala
 1075 1080 1085
 10 Glu Ile Val Leu Pro Glu Ala Ala Gly Ala Asp Ala Ala Arg Tyr Gly
 1090 1095 1100
 Met His Pro Ala Leu Leu Asp Ala Val Leu His Ala Ala Arg Leu Gly
 1105 1110 1115 1120
 15 Ala Phe Arg Glu Arg Ser Glu Glu Lys Tyr Leu Pro Phe Ala Trp Glu
 1125 1130 1135
 Gly Val Thr Leu Arg Thr Arg Gly Ala Thr Ala Val Arg Ala Arg Ile
 1140 1145 1150
 20 Ser Arg Ala Gly Thr Asp Ala Ile Arg Leu Asp Val Thr Asp Thr Ala
 1155 1160 1165
 Asp Arg Pro Val Leu Thr Ala Glu Ser Leu Thr Leu Arg Pro Val Ser
 1170 1175 1180
 25 Ala Gly Gln Leu Met Ala Val Pro Arg Asp Ser Leu Phe Arg Val Asp
 1185 1190 1195 1200
 Trp Val Ser Ala Pro Ala Ala Asn Gly Pro Gly Leu Arg Leu Ala Arg
 1205 1210 1215
 Ala Ala Thr Val Glu Ala Ala Leu Ala Ala Asp Ala Asp Ile Val Val
 1220 1225 1230
 35 Val Pro Cys Leu Asp Ser Glu Gly Pro His Gln Ala Thr Tyr Gln Ala
 1235 1240 1245
 Leu Glu Leu Leu Gln Arg Trp Leu Ala Ala Asp Thr Gly Thr Thr Thr
 1250 1255 1260
 40 Leu Ala Leu Leu Thr His Arg Ala Val Ala Val Gly Asp Asp Val His
 1265 1270 1275 1280
 Asp Leu His His Ala Pro Leu Trp Gly Leu Val Arg Thr Ala Gln Thr
 1285 1290 1295
 45 Glu His Pro Gly Cys Phe Arg Leu Val Asp Ser Asp Asp Pro Asp Pro
 1300 1305 1310
 Thr Thr Asp Val Leu Ala Ala Ala Leu Ala Thr Gly Glu Pro Gln Val
 1315 1320 1325
 Ala Ile Arg Asp Gly Ala Val Leu Ala Pro Arg Leu Thr Ala Ala Ser
 1330 1335 1340
 55 Ala Pro Arg Glu Pro Ala Glu Trp Asp Ala Glu Gly Thr Val Leu Ile
 1345 1350 1355 1360
 Thr Gly Gly Ser Gly Ala Leu Ala Gly Ile Val Ala Gln His Leu Val
 1365 1370 1375
 60 Ala Arg His Gly Val Arg Arg Leu Val Leu Ala Ser Arg Ser Gly Arg
 65

ES 2 287 552 T3

	1380	1385	1390
5	Pro Ala Pro Gly Ala Asp Leu Leu Asp Ala Asp Val Thr Ala Val Ser 1395 1400 1405		
	Cys Asp Val Ser Asp Arg Asp Ala Val Ala Ala Leu Leu Ala Ser Val 1410 1415 1420		
10	Pro Asp Glu His Pro Leu Thr Ala Val Val His Thr Ala Gly Val Leu 1425 1430 1435 1440		
	Asp Asp Gly Val Leu His Ala Leu Thr Thr Glu Arg Ile Asp Thr Ser 1445 1450 1455		
15	Phe Ala Ala Lys Val Asp Gly Ala Arg His Leu His Glu Leu Thr Ser 1460 1465 1470		
	His Leu Asp Leu Thr Ala Phe Val Leu Phe Ser Ser Ala Ser Ala Val 1475 1480 1485		
20	Leu Gly Ala Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Tyr Leu 1490 1495 1500		
	Asp Ala Leu Ala Ala His Arg Arg Ser Asn Asp Leu Pro Ala Val Ser 1505 1510 1515 1520		
25	Leu Ala Trp Gly Leu Trp Ala Glu His Glu Gly Met Ala Arg Gly Leu 1525 1530 1535		
	Gly Asp Ala Glu Leu Thr Arg Ile Ser Arg Ile Gly Val Thr Ala Leu 1540 1545 1550		
30	Ser Ala Glu Asp Gly Met Arg Leu Phe Asp Ala Gly Cys Ala Gly Asp 1555 1560 1565		
	Gln Ser Gln Leu Val Pro Met Arg Val Asp Thr Ala Ala Leu Arg Ala 1570 1575 1580		
35	Arg Arg Asp His Leu Pro Ala Pro Met Trp Ser Leu Val Pro Glu Arg 1585 1590 1595 1600		
40	Thr Arg Ala Ala Arg Thr Gln Pro Ala Ala Ser Leu Arg Asp Arg Leu 1605 1610 1615		
	Ala Glu Leu Thr Ala Pro Glu Arg Lys Arg Thr Val Leu Asn Leu Val 1620 1625 1630		
45	Arg Asn Ala Val Ala Asp Thr Leu Gly His Asn Ala Ala Asp Gly Val 1635 1640 1645		
	Pro Pro Asp Gln Ser Leu Asp Ala Ala Gly Phe Asp Ser Leu Thr Ala 1650 1655 1660		
50	Val Glu Phe Arg Asn Arg Leu Ser Ala Val Thr Asp Leu Arg Leu Pro 1665 1670 1675 1680		
	Ala Thr Leu Thr Tyr Asp His Pro Thr Pro Ala Ala Ile Ala Glu His 1685 1690 1695		
55	Ile Leu Thr Arg Leu Thr Leu Leu Lys Glu Thr Ala Ala Pro Ala Val 1700 1705 1710		
60	Gly Thr Ala Pro Val Ala Ala Pro Thr Glu Asp Asp Ala Ile Val Ile 1715 1720 1725		

65

ES 2 287 552 T3

	Val Gly Met Ala Gly Arg Phe Pro Gly Gly Val Arg Thr Pro Glu Gly	
	1730	1735 1740
5	Leu Trp Asp Leu Val His Ser Gly Thr Asp Ala Ile Ser Glu Trp Pro	
	1745	1750 1755 1760
	Thr Asp Arg Gly Trp Asp Val Glu Asn Leu Tyr Asp Pro Asp Pro Asp	
		1765 1770 1775
10	Ala Val Gly Lys Ser Tyr Val Arg His Gly Gly Phe Leu His Asp Val	
		1780 1785 1790
	Ala Gly Phe Asp Ala Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu	
		1795 1800 1805
15	Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Cys Ser Tyr Glu Ala	
		1810 1815 1820
	Leu Glu Arg Ala Gly Ile Asp Pro Ala Thr Leu Arg Gly Ser Arg Ser	
		1825 1830 1835 1840
20	Gly Val Tyr Ala Gly Val Met Tyr His Glu Tyr Ala Ser Arg Leu Gly	
		1845 1850 1855
	Ala Thr Pro Ala Gly Phe Glu Gly Thr Leu Gly Thr Gly Ser Ser Gly	
		1860 1865 1870
25	Ser Ile Ala Ser Gly Arg Ile Ser Tyr Thr Phe Asp Leu Thr Gly Pro	
		1875 1880 1885
	Ala Val Thr Val Asp Thr Ala Cys Ser Thr Ser Leu Val Gly Leu His	
		1890 1895 1900
	Leu Ala Val Gln Ala Leu Arg Ala Gly Glu Cys Glu Leu Ala Leu Ala	
		1905 1910 1915 1920
35	Gly Gly Val Thr Val Met His Thr Pro Arg Pro Phe Val Glu Phe Ser	
		1925 1930 1935
	Arg Gln Arg Gly Leu Ala Ala Asp Gly Arg Ser Lys Ala Phe Ala Ala	
		1940 1945 1950
40	Ser Ala Asp Gly Val Ala Trp Ala Glu Gly Ala Gly Ile Leu Val Leu	
		1955 1960 1965
	Glu Arg Leu Ser Ala Ala Arg Arg Asn Gly His Arg Val Leu Ala Val	
		1970 1975 1980
	Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr	
		1985 1990 1995 2000
50	Ala Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Ala Ala Leu Ala	
		2005 2010 2015
	Ser Ala Gly Leu Gly Pro Ala Asp Val Asp Val Val Glu Ala His Gly	
		2020 2025 2030
55	Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala	
		2035 2040 2045
	Thr Tyr Gly Arg Gly Arg Asp Ala Asp Arg Pro Leu Trp Leu Gly Ser	
		2050 2055 2060
60		
65		

ES 2 287 552 T3

	Val Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Ser	
	2065	2070 2075 2080
5	Val Ile Lys Met Val Gln Ala Met Gln Ala Gly Val Leu Pro Arg Thr	
	2085	2090 2095
	Leu His Val Asp Glu Pro Ser Gly Glu Val Asp Trp Asp Ser Gly Ala	
	2100	2105 2110
10	Val Arg Leu Leu Thr Glu Ala Arg Glu Trp Pro Ser Gly Glu Gly Arg	
	2115	2120 2125
	Val Arg Arg Ala Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala	
	2130	2135 2140
15	His Val Ile Leu Glu Glu Pro Pro Ala Glu Asp Ala Leu Pro Glu Pro	
	2145	2150 2155 2160
	Glu Ala Gly Asp Val Val Pro Trp Val Leu Ser Ala Arg Ser Ala Glu	
20	2165	2170 2175
	Ala Leu Arg Glu Gln Ala Ala Arg Leu Ala Ser Val Ala Gly Gly Leu	
	2180	2185 2190
25	Asn Val Val Asp Val Gly Trp Ser Leu Ala Ser Thr Arg Ala Ala Phe	
	2195	2200 2205
	Glu His Arg Ala Val Val Val Gly Gly Asp Arg Glu Glu Leu Leu Gly	
	2210	2215 2220
30	Lys Leu Ser Ser Val Ser Gly Val Glu Val Gly Val Gly Val Gly Ala	
	2225	2230 2235 2240
	Gly Gly Gly Val Val Leu Val Phe Ala Gly Gln Gly Cys Gln Trp Val	
	2245	2250 2255
35	Gly Met Gly Arg Glu Leu Leu Gly Ser Ser Leu Val Phe Ala Glu Ser	
	2260	2265 2270
	Met Arg Glu Cys Ala Ala Ala Leu Ser Pro Phe Val Asp Phe Ser Val	
40	2275	2280 2285
	Val Asp Val Leu Gly Ser Ala Gly Glu Leu Gly Arg Val Glu Val Val	
	2290	2295 2300
45	Gln Pro Ala Leu Trp Ala Val Met Val Ser Leu Ala Arg Val Trp Arg	
	2305	2310 2315 2320
	Ser Trp Gly Val Pro Val Ala Ala Val Val Gly His Ser Gln Gly Glu	
	2325	2330 2335
50	Ile Ala Ala Ala Thr Val Ala Gly Ala Leu Ser Val Gly Asp Ala Ala	
	2340	2345 2350
	Arg Val Val Ala Leu Arg Ser Arg Leu Ile Ala Glu Arg Leu Ser Gly	
	2355	2360 2365
55	Leu Gly Gly Met Val Ser Val Ala Leu Ser Arg Glu Arg Val Val Ser	
	2370	2375 2380
60	Leu Ile Ala Gly Val Pro Gly Val Ser Val Ala Ala Val Asn Gly Ser	
	2385	2390 2395 2400
	Ser Ser Thr Val Val Ser Gly Glu Ala Ala Gly Leu Glu Arg Val Leu	

65

ES 2 287 552 T3

	2405	2410	2415
	Ala Ala Cys Val Ser Ser Gly Val Arg Ala Arg Arg Ile Asp Val Asp		
	2420	2425	2430
5	Tyr Ala Ser His Ser Val Gln Val Glu Leu Ile Arg Glu Glu Leu Leu		
	2435	2440	2445
	Gly Val Leu Asp Gly Ile Val Pro Arg Ser Gly Glu Ile Pro Phe Val		
10	2450	2455	2460
	Ser Thr Val Thr Gly Glu Arg Ile Asp Thr Val Glu Leu Gly Ala Glu		
	2465	2470	2475
	Tyr Trp Tyr Arg Asn Leu Arg Gln Thr Val Glu Phe Gln Ser Val Val		
15	2485	2490	2495
	Glu Gly Leu Val Ala Gln Gly Cys Arg Val Phe Leu Glu Ser Ser Pro		
	2500	2505	2510
20	His Pro Val Leu Thr Val Gly Ile Glu Glu Ser Ala Asp Arg Val Val		
	2515	2520	2525
	Ala Leu Glu Ser Leu Arg Arg Gly Glu Gly Gly Leu Arg Arg Leu Val		
25	2530	2535	2540
	Asp Ala Ala Gly Glu Ala Trp Val Arg Gly Val Pro Ile Asp Trp Ala		
	2545	2550	2555
	Gly Met Leu Ala Gly Gly Arg Arg Val Asp Leu Pro Thr Tyr Pro Phe		
30	2565	2570	2575
	Gln His Gln Pro Tyr Trp Leu Asp Ser Pro Arg His Pro Ala Gly Asp		
	2580	2585	2590
	Val Thr Ala Val Gly Leu Thr Glu Ala Gly His Ala Phe Val Pro Ala		
35	2595	2600	2605
	Ala Val Asp Leu Pro Asp Gly Gln Arg Val Trp Thr Gly Arg Leu Ser		
	2610	2615	2620
40	Leu Pro Ser Tyr Pro Trp Leu Ala Asp His Gln Val Leu Gly Gln Val		
	2625	2630	2635
	Leu Leu Pro Gly Val Val Trp Val Glu Leu Ala Leu His Ala Gly His		
	2645	2650	2655
45	Gln Ala Gly Cys Asp Ser Val Asp Glu Leu Thr Leu Gln Ser Pro Leu		
	2660	2665	2670
	Val Leu Gly Ala Ser Asp Thr Val Gln Val Arg Val Val Val Thr Glu		
50	2675	2680	2685
	Thr Glu Glu Pro Gly Thr Arg Thr Val Ser Met His Ser Arg Arg Asp		
	2690	2695	2700
55	Asp Gly Ser Trp Val Thr His Ala Glu Gly Ile Leu Gly Ala Gly Gly		
	2705	2710	2715
	Pro Pro Pro Glu Pro Leu Pro Glu Trp Pro Pro Thr Gly Ala Met Pro		
	2725	2730	2735
60	Leu Asp Val Glu Gly Phe Tyr Asp Glu Leu Ala Ala Gly Gly Tyr His		
	2740	2745	2750

65

ES 2 287 552 T3

	Tyr Gly Pro Gln Phe Arg Cys Leu Arg Arg Ala Trp Arg Ala Gly Glu	
	2755	2760 2765
5	Asp Leu Val Ala Glu Ile Ser Leu Pro Glu Gly Thr Asp Val Asp Ala	
	2770	2775 2780
	Tyr Gly Leu His Pro Gly Leu Phe Asp Ala Ala Val His Ser Val Ala	
	2785	2790 2795 2800
10	Cys Ala Arg Thr Ser Ala Gly Ala Gly Asp Asp Gly Pro Arg Leu Pro	
		2805 2810 2815
	Phe Ala Phe Ser Asp Val Arg Leu Phe Ala Thr Gly Val Thr Ser Leu	
		2820 2825 2830
15	Arg Val Arg Ile Asp Pro Gln Asn Ser Ser Trp Gln Ala Trp Asp Glu	
		2835 2840 2845
	Ser Gly Leu Pro Val Leu Thr Ile Gly Arg Leu Ala Gly Arg Pro Val	
		2850 2855 2860
20	Asp Ala Asp Gln Phe Ala Val Arg Arg Ala Gly His Leu Phe Arg Val	
		2865 2870 2875 2880
	Glu Thr Arg His Glu Ala Leu Ala Gly Pro Ala Pro Ala Ser Trp Ala	
		2885 2890 2895
25	Val Ile Gly Ala Asp Pro Ala Gly Tyr Ala Ala Ala Leu Glu Ala Thr	
		2900 2905 2910
30	Gly Ala Gln Val Thr Thr Ala Ala Asp Leu Ala Gly Leu Thr Ser Ala	
		2915 2920 2925
	Pro Glu Ala Ala Leu Phe Thr Leu Pro Gly Thr Lys Asp Ala Gly Val	
		2930 2935 2940
35	Thr Glu Glu Val Pro Thr Ala Val Arg Glu Ala Thr Ala Gln Val Leu	
		2945 2950 2955 2960
	Glu Val Leu Gln Asp Trp Leu Thr Asp Gly Arg Phe Asp Asp Ala Arg	
		2965 2970 2975
40	Leu Val Val Val Ser Arg Glu Ala Glu Asp Gly Asp Leu Leu His Gly	
		2980 2985 2990
	Thr Ala Arg Gly Leu Leu Arg Ala Ala Gln Ala Glu His Pro Asp Arg	
		2995 3000 3005
45	Ile Thr Leu Val Asp Leu Asp Ala His Pro Ala Ser Leu Thr Ala Leu	
		3005 3015 3020
	Pro Gly Phe Ala Leu Gly Pro Glu Pro Glu Val Val Val Arg Ala Gly	
		3025 3030 3035 3040
50	Asp Gly Arg Ala Pro Arg Leu Ala Arg Ala Gln Ala Pro Thr Gly Ala	
		3045 3050 3055
55	Gly Ser Leu Gly Thr Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Thr	
		3060 3065 3070
	Leu Gly Gly Leu Leu Ala Arg His Leu Val Glu Thr His Gly Val Thr	
		3075 3080 3085
60		
65		

ES 2 287 552 T3

Arg Leu Leu Leu Val Ser Arg Arg Gly Pro Ala Ala Asp Gly Ala Asp
 3090 3095 3100
 5 Arg Leu His Ala Glu Leu Thr Gly His Gly Ala His Val Asp Ile Val
 3105 3110 3115 3120
 Ala Ala Asp Leu Gly Asp Arg Thr Ser Val Ala Ala Leu Leu Ala Thr
 3125 3130 3135
 10 Val Asp Ala Asp His Pro Leu Ser Ala Val Val His Ala Ala Gly Ala
 3140 3145 3150
 Leu Asp Asp Gly Val Leu Gly Thr Arg Ser Ala Asp Trp Leu Asp Pro
 3155 3160 3165
 15 Val Leu Arg Pro Lys Ala Asp Ala Ala Trp His Leu His Glu Leu Thr
 3170 3175 3180
 Ala Glu Leu Pro Leu Thr Ala Phe Val Met Phe Ser Ser Ala Ala Ser
 3185 3190 3195 3200
 Val Leu Gly Ala Ala Gly Gln Ala Asn Tyr Ala Ala Ala Asn Gly Phe
 3205 3210 3215
 25 Leu Asp Ala Leu Ala Ala His Arg Ala Ala Arg Gly Leu Pro Gly Thr
 3220 3225 3230
 Ser Leu Ala Trp Gly Leu Trp Glu His Arg Ser Glu Leu Thr Arg His
 3235 3240 3245
 30 Thr Gly Ser Pro Ser Arg Ser Ile Ala Ala Val Gly Ala Leu Ser Thr
 3250 3255 3260
 Ala Glu Ala Leu Ala Ala Phe Asp Ala Gly Leu Ala Ser Gly Glu Pro
 3265 3270 3275 3280
 35 Leu Ala Val Pro Ile Arg Leu Glu Ser Thr Ser Ser Glu Glu Val Pro
 3285 3290 3295
 Pro Met Leu Arg Gly Leu Val Arg Val Arg Arg Arg Ala Ala Thr Gly
 3300 3305 3310
 40 Thr Glu Pro Ala Ala Ser Ala Gly Ala Ala Gln Glu Val Arg Gln Leu
 3315 3320 3325
 Ala Glu Leu Gly Ala Asp Glu Arg Gln Arg Arg Val Gln Arg Ile Val
 3330 3335 3340
 45 Leu Asp Thr Ala Ala Ala Val Leu Gly His Asp Ser His Asp Ala Ile
 3345 3350 3355 3360
 50 Pro Leu Thr Arg Gly Phe Leu Glu Leu Gly Phe Asp Ser Leu Thr Ala
 3365 3370 3375
 Val Arg Leu Arg Asn Arg Leu Ala Arg Arg Leu Gly Leu Arg Leu Pro
 3380 3385 3390
 55 Ala Thr Val Val Phe Asp His Pro Ser Pro Ala Ala Leu Ala Ala His
 3395 3400 3405
 Leu Val Glu His Leu Val Gly Thr Val Asp Pro Thr Ala Gln Ala Met
 3410 3415 3420
 60 Glu Gln Leu Glu Ala Leu Arg Arg Ser Val His Ala Ala Thr Pro Ala
 65

ES 2 287 552 T3

3425 3430 3435 3440
 Gly Gly Leu Asp Arg Ala Leu Val Thr Gln Arg Leu Thr Ala Leu Leu
 3445 3450 3455
 5 Asp Glu Met Arg His Val Asp Gly Pro Gly Gly Thr Glu Gly Pro Asp
 3460 3465 3470
 10 Gly Ser Gly Asp Asp Leu Glu Asn Ala Thr Ala Asp Glu Ile Tyr Ala
 3475 3480 3485
 Leu Ile Asp Asn Glu Leu Gly Ile Gly Gly Thr Gln
 3490 3495 3500
 15 <210> 5
 <211> 1620
 <212> PRT
 20 <213> *Streptomyces parvulus* Tü4055
 <400> 5
 Met Asn Gly Asp Asp Lys Ala Leu Ala Tyr Leu Lys Arg Val Thr Ala
 1 5 10 15
 25 Asp Leu Arg Ser Ala Arg Ala Arg Leu Gln Glu Leu Glu Ser Ala Asp
 20 25 30
 30 Thr Asp Pro Ile Ala Ile Ile Gly Met Gly Cys Arg Leu Pro Gly Gly
 35 40 45
 Val Arg Thr Pro Glu Asp Leu Trp Asp Leu Val Glu Lys Lys His Asp
 50 55 60
 35 Ala Ile Gly Pro Phe Pro Ala Asp Arg Gly Trp Asp Leu Glu Asn Leu
 65 70 75 80
 40 Tyr Asp Pro Asp Pro Asp Ala Pro Gly Lys Ala Tyr Val Arg Glu Gly
 85 90 95
 Gly Phe Val His Asp Val Ala Gly Phe Asp Ala Gly Phe Phe Gly Ile
 100 105 110
 45 Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln His Arg Leu Leu Leu
 115 120 125
 Glu Cys Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Ser Ser
 130 135 140
 50 Leu Glu Gly Thr Arg Thr Gly Val Tyr Thr Gly Leu Met Thr His Glu
 145 150 155 160
 55 Tyr Ala Thr Arg Leu Pro Ser Ile Asp Glu Glu Leu Glu Gly Val Ile
 165 170 175
 Gly Ile Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ser Tyr Thr
 180 185 190
 60 Leu Gly Leu Asn Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser
 195 200 205
 Ser Leu Val Ala Leu His Leu Ala Ala Gln Ala Leu Arg Gln Gly Gln
 210 215 220
 65 Cys Thr Leu Ala Leu Ala Gly Gly Ala Ser Val Ile Ala Ala Pro Thr

ES 2 287 552 T3

	225		230		235		240
5	Val Phe Ala Thr	Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg	245	250	255		
	Cys Lys Ala Phe	Ser Ser Thr Thr Asp Gly Thr Gly Phe Gly Glu Gly	260	265	270		
10	Val Gly Val Leu	Val Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly	275	280	285		
	His Glu Val Leu	Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly	290	295	300		
15	Ala Ser Ser Gly	Phe Thr Ala Pro Asn Gly Pro Ser Gln Gln Asp Val	305	310	315		320
	Ile Arg Glu Ala	Leu Ala Asp Gly Arg Leu Thr Pro Ala Asp Val Asp	325	330	335		
20	Val Val Glu Gly	His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu	340	345	350		
	Ala Gln Ala Leu	Leu Ala Thr Tyr Gly Arg Gly Arg Asp Ala Asp Arg	355	360	365		
25	Pro Leu Trp Leu	Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala	370	375	380		
	Ala Ala Gly Val	Ala Ser Val Ile Lys Met Val Gln Ala Met Gln Ala	385	390	395		400
30	Gly Val Leu Pro	Arg Thr Leu His Val Asp Glu Pro Ser Gly Glu Val	405	410	415		
35	Asp Trp Asp Ser	Gly Ala Val Arg Leu Leu Thr Glu Ala Arg Glu Trp	420	425	430		
	Pro Ser Gly Glu	Gly Arg Val Arg Arg Ala Gly Val Ser Ser Phe Gly	435	440	445		
40	Ile Ser Gly Thr	Asn Ala His Val Ile Leu Glu Glu Pro Pro Ala Glu	450	455	460		
	Asp Ala Leu Pro	Glu Pro Glu Ala Gly Asp Val Val Pro Trp Val Leu	465	470	475		480
45	Ser Ala Arg Ser	Ala Glu Ala Leu Arg Glu Gln Ala Ala Arg Leu Ala	485	490	495		
	Ser Val Ala Gly	Gly Leu Asn Val Val Asp Val Gly Trp Ser Leu Ala	500	505	510		
50	Ser Thr Arg Ala	Ala Phe Glu His Arg Ala Val Val Val Gly Gly Asp	515	520	525		
	Arg Glu Glu Leu	Leu Gly Lys Leu Ser Ser Val Ser Gly Val Glu Val	530	535	540		
55	Gly Val Gly Val	Gly Ala Gly Gly Gly Val Val Leu Val Phe Ala Gly	545	550	555		560
60	Gln Gly Cys Gln	Trp Val Gly Met Gly Arg Glu Leu Leu Gly Ser Ser	565	570	575		

65

ES 2 287 552 T3

	Leu Val Phe	Ala Glu Ser Met Arg Glu Cys Ala Ala Ala	Leu Ser Pro
		580 585	590
5	Phe Val Asp	Phe Ser Val Val Asp Val Leu Gly Ser Ala Gly Glu Leu	
		595 600	605
	Gly Arg Val	Glu Val Val Gln Pro Ala Leu Trp Ala Val Met Val Ser	
		610 615	620
10	Leu Ala Arg	Val Trp Arg Ser Trp Gly Val Pro Val Ala Ala Val Val	
		625 630	635 640
	Gly His Ser	Gln Gly Glu Ile Ala Ala Ala Thr Val Ala Gly Ala Leu	
		645 650	655
15	Ser Val Gly	Asp Ala Ala Arg Val Val Ala Leu Arg Ser Arg Leu Ile	
		660 665	670
	Ala Glu Arg	Leu Ser Gly Leu Gly Gly Met Val Ser Val Ala Leu Ser	
		675 680	685
20	Arg Glu Arg	Val Val Ser Leu Ile Ala Gly Val Pro Gly Val Ser Val	
		690 695	700
	Ala Ala Val	Asn Gly Ser Ser Ser Thr Val Val Ser Gly Glu Ala Ala	
		705 710	715 720
	Gly Leu Glu	Arg Val Leu Ala Ala Cys Val Ser Ser Gly Val Arg Ala	
		725 730	735
30	Arg Arg Ile	Asp Val Asp Tyr Ala Ser His Ser Val Gln Val Glu Leu	
		740 745	750
	Ile Arg Glu	Glu Leu Leu Gly Val Leu Asp Gly Ile Val Pro Arg Ser	
		755 760	765
35	Gly Glu Ile	Pro Phe Val Ser Thr Val Thr Gly Glu Arg Ile Asp Thr	
		770 775	780
	Val Glu Leu	Gly Ala Glu Tyr Trp Tyr Arg Asn Leu Arg Gln Thr Val	
		785 790	795 800
	Glu Phe Gln	Ser Val Val Glu Gly Leu Val Ala Gln Gly Cys Arg Val	
		805 810	815
45	Phe Leu Glu	Ser Ser Pro His Pro Val Leu Thr Val Gly Ile Glu Glu	
		820 825	830
	Ser Ala Asp	Arg Val Val Ala Leu Glu Ser Leu Arg Arg Gly Glu Gly	
		835 840	845
50	Gly Leu Arg	Arg Leu Val Asp Ala Ala Gly Glu Ala Trp Val Arg Gly	
		850 855	860
	Val Pro Ile	Asp Trp Ala Gly Met Leu Ala Gly Gly Arg Arg Val Asp	
		865 870	875 880
55	Leu Pro Thr	Tyr Pro Phe Gln His Gln Pro Tyr Trp Leu Asp Ser Pro	
		885 890	895
	Arg His Pro	Ala Gly Asp Val Thr Gly Pro Gly Asp Asp Glu Phe Trp	
		900 905	910

ES 2 287 552 T3

Ala Ala Val Glu His Gly Glu Ala Thr Glu Leu Ala Asp Leu Leu Arg
 915 920 925
 5 Arg Ser Ala Ala Glu Pro Gly Gln Asp Leu His Ala Pro Val Ala Ala
 930 935 940
 Leu Leu Pro Thr Leu Ala Thr Trp Arg Arg Asp Arg Gln Arg Arg Ala
 945 950 955 960
 10 Ala Val Asp Ser Trp Arg Tyr Arg Ile Val Trp Arg Pro Val Ala Thr
 965 970 975
 Pro Ser Tyr Asp Arg Val Leu Ser Gly Arg Trp Ala Val Val Val Pro
 980 985 990
 15 Ala Gly His Glu Asp Asp Pro Val Val Asp Trp Val Cys Ser Ala Leu
 995 1000 1005
 Arg Asp His Gly Gly Glu Pro Glu Arg Met Val Leu Gly Pro Arg Glu
 1010 1015 1020
 20 Ser Arg Ser Ala Leu Ala Thr Arg Leu Ala Ala Asp Pro Pro Gly Gly
 1025 1030 1035 1040
 Val Val Ser Leu Leu Gly Leu Ser Gly Ala Ala His Pro Asp His Glu
 1045 1050 1055
 25 Val Leu Pro Ser Ala Val Ala Gly Thr Val Leu Leu Ala Gln Ala Leu
 1060 1065 1070
 30 Ser Asp Gly Ala Val Arg Ala Pro Val Trp Thr Leu Thr Arg Asn Gly
 1075 1080 1085
 Val Ser Ala Thr Ala Thr Asp Pro Val Ala Pro Thr His Ala Ala Gln
 1090 1095 1100
 35 Val Trp Ala Val Ala Arg Val Ala Gly Leu Glu His Pro Glu Ala Trp
 1105 1110 1115 1120
 Gly Gly Leu Leu Asp Leu Pro Asp Arg Leu Asp Asp Arg Ala Ala Ala
 1125 1130 1135
 40 Arg Phe Ala Ala Val Leu Ser Ala Gly Glu Asp Glu Asp Gln Leu Ala
 1140 1145 1150
 Leu Arg Asp Ala Gly Leu Leu Ala Arg Arg Leu Val Arg Ala Pro Val
 1155 1160 1165
 45 Pro Arg Asp Ala Val Thr Ala Gly Trp Gln Pro Arg Asp Thr Ala Leu
 1170 1175 1180
 50 Val Thr Gly Gly Thr Gly Gly Leu Gly Gly Gln Val Ala Arg Trp Leu
 1185 1190 1195 1200
 Ala Ala Ala Gly Val Arg His Leu Val Leu Val Ser Arg Arg Gly Ala
 1205 1210 1215
 55 Glu Ala Glu Gly Ala Asp Arg Leu Arg Asp Asp Leu Thr Ala Leu Gly
 1220 1225 1230
 Val Gln Val Thr Phe Gly Ala Cys Asp Val Ala Asp Arg Ala Ala Leu
 1235 1240 1245
 60 Ser Ala Leu Leu Asp Arg Val Gln Glu Asp Gly Pro Pro Ile Arg Thr
 65

ES 2 287 552 T3

	1250		1255		1260
5	Val Val His Ala Ala Gly Ser Gly Arg Ala Ala Arg Leu Leu Asp Thr				
	1265		1270		1280
	Asp Ala Glu Glu Thr Ala Ala Val Leu Arg Ala Lys Ser Ala Gly Ala				
		1285		1290	1295
10	Arg Asn Leu His Glu Leu Leu Asp Asp Val Asp Ala Phe Val Leu Phe				
		1300		1305	1310
	Ser Ser Gly Ala Gly Val Trp Gly Ser Ser Ala Gln Gly Ala Tyr Ala				
		1315		1320	1325
15	Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln Arg Arg Gly Gln				
		1330		1335	1340
	Gly Arg Pro Ala Thr Ser Val Ala Trp Gly Ala Trp Ala Gly Asp Gly				
		1345		1350	1355
20	Met Thr Ala Ala Ala Gly Glu Glu Trp Trp Ser Arg Gln Gly Leu Arg				
		1365		1370	1375
	Phe Met Ala Pro Glu Ala Ala Leu Asp Ala Leu Arg Gln Ala Val Asp				
25		1380		1385	1390
	Arg Ala Glu Ser Thr Leu Val Val Ala Asp Ile Asp Trp Lys Thr Phe				
		1395		1400	1405
30	Ala Pro Leu Phe Thr Ser Ala Arg Ser Arg Pro Leu Ile Thr Asp Ile				
		1410		1415	1420
	Pro Glu Ala Arg Pro Glu Pro Arg Pro Glu Gly Ala Asp Gln Pro Thr				
		1425		1430	1435
35	Gln Gly Leu Val Ala Lys Leu Ala Val Leu Ser Ala Asp Glu Arg Arg				
		1445		1450	1455
	Arg Ala Leu Leu Ala Glu Val Arg Ala Gln Ala Ala Val Val Leu Gly				
		1460		1465	1470
40	His Pro Gly Ala Asp Ala Val Pro Val Asp Arg Pro Phe Arg Glu Leu				
		1475		1480	1485
	Gly Phe Asp Ser Leu Ser Ala Val Lys Leu Arg Asn Arg Ile Val Ala				
45		1490		1495	1500
	Ala Thr Gly Leu Glu Leu Pro Ala Thr Leu Val Phe Asp His Pro Thr				
		1505		1510	1515
	Ser Thr Ala Leu Ala Ala Tyr Leu Gly Ala Arg Leu Gly Ile Asp Gly				
50		1525		1530	1535
	Ala Pro Ala Gly Ser Thr Leu Leu Glu Asp Leu Ala Arg Leu Glu Ser				
		1540		1545	1550
55	Thr Val Ala Thr Leu Thr Ala Ala Pro Leu Ala Glu Thr Val Pro Asp				
		1555		1560	1565
	Ala Arg Asp Arg Ala Ala Leu Thr Thr Arg Leu Arg Ala Leu Leu Glu				
		1570		1575	1580
60	Arg Trp Asp Gln Ala Asp Gly Glu Asp Gln Ala Ala Ala Arg Glu Glu				
		1585		1590	1595
	Leu Asp Asp Leu Ser Asp Asp Asp Leu Phe Asp Phe Ile Asp Ala Lys				
		1605		1610	1615
65	Phe Gly Arg Ser				
		1620			

ES 2 287 552 T3

<210> 6

<211> 2130

<212> PRT

5 <213> *Streptomyces parvulus* Tü4055

<400> 6

```

10      Met Gly Asp Glu Gln Lys Leu Arg Thr Tyr Leu Arg Arg Val Thr Ala
      1              5              10              15

      Asp Leu Ala Asp Val Thr Glu Arg Leu Gln Arg Ala Glu Asp Lys Asn
              20              25              30

15      Ala Glu Pro Ile Ala Ile Val Gly Met Gly Cys Arg Tyr Pro Gly Gly
              35              40              45

      Val Arg Ser Pro Glu Glu Phe Trp Asn Leu Leu Asp Glu Gly Val Asp
              50              55              60

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

      Tyr Asp Pro Asp Pro Asp Glu Pro Gly Lys Cys Tyr Ala Arg Glu Gly
25              85              90              95

      Gly Phe Leu Tyr Asp Ala Gly Glu Phe Asp Ala Ala Phe Phe Gly Ile
              100              105              110

30      Ser Pro Arg Glu Ala Leu Ser Met Asp Pro Gln Gln Arg Leu Leu Leu
              115              120              125

      Glu Cys Ser Trp Ser Ala Leu Glu Arg Ala Gly Ile Asp Pro Gly Ser
35              130              135              140

      Leu Arg Gly Lys Asp Val Gly Val Tyr Val Gly Ala Trp Asn Ser Asn
      145              150              155              160

      Tyr Gly Arg Gly Gly Gly Ala Glu Ser Ser Glu Gly His Leu Leu Thr
40              165              170              175

      Gly Asn Ala Ser Ser Val Val Ser Gly Arg Val Ala Tyr Val Leu Gly
              180              185              190

45      Leu Glu Gly Pro Ala Val Thr Ile Asp Thr Ala Cys Ser Ser Ser Leu
              195              200              205

      Val Gly Leu His Leu Ala Ala Gln Ala Leu Arg Ser Gly Glu Cys Gly
50              210              215              220

      Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro Leu Ser Leu
      225              230              235              240

      Val Ser Phe Ser Arg Gln Arg Gly Leu Ala Gln Asp Gly Arg Ser Lys
55              245              250              255

      Ala Phe Ser Ala Asp Ala Asp Gly Met Gly Met Ala Glu Gly Val Gly
              260              265              270

```

ES 2 287 552 T3

	Val	Leu	Val	Leu	Glu	Arg	Leu	Ser	Glu	Ala	Arg	Arg	Asn	Gly	His	Glu	
		275						280					285				
5	Val	Leu	Ala	Val	Leu	Arg	Ser	Ser	Ala	Val	Asn	Gln	Asp	Gly	Ala	Ser	
		290					295					300					
	Asn	Gly	Leu	Ser	Ala	Pro	Asn	Gly	Pro	Ala	Gln	Gln	Arg	Val	Ile	Gln	
	305					310					315					320	
10	Ser	Ala	Leu	Thr	Val	Gly	Arg	Leu	Ala	Pro	Ser	Asp	Ile	Asp	Val	Val	
					325					330					335		
	Glu	Ala	His	Gly	Thr	Gly	Thr	Ala	Leu	Gly	Asp	Pro	Ile	Glu	Ala	Gln	
				340					345					350			
15	Ala	Leu	Leu	Ala	Thr	Tyr	Gly	Arg	Gly	Arg	Asp	Ala	Asp	Arg	Pro	Leu	
		355					360						365				
	Trp	Leu	Gly	Ser	Val	Lys	Ser	Asn	Ile	Gly	His	Thr	Gln	Ala	Ala	Ala	
20		370					375					380					
	Gly	Val	Ala	Gly	Val	Ile	Lys	Met	Val	Leu	Ala	Leu	Arg	Lys	Gly	Val	
	385					390					395					400	
	Leu	Pro	Arg	Thr	Leu	His	Val	Asp	Glu	Pro	Thr	Gly	Glu	Val	Asp	Trp	
25					405					410					415		
	Asp	Ser	Gly	Ala	Val	Arg	Leu	Leu	Thr	Glu	Ala	Arg	Glu	Trp	Pro	Ser	
				420					425					430			
30	Gly	Glu	Gly	Arg	Val	Arg	Arg	Ala	Gly	Val	Ser	Ser	Phe	Gly	Ile	Ser	
		435						440					445				
	Gly	Thr	Asn	Ala	His	Val	Ile	Val	Glu	Glu	Ala	Pro	Glu	Glu	Glu	Pro	
	450					455						460					
35	Arg	Pro	Glu	Ala	Pro	Ser	Val	Asp	Val	Val	Pro	Trp	Val	Leu	Ser	Ala	
	465					470					475					480	
	Arg	Ser	Ala	Glu	Ala	Leu	Arg	Glu	Gln	Ala	Ala	Arg	Leu	Ala	Ser	Val	
				485						490					495		
40	Ala	Gly	Gly	Leu	Asn	Val	Val	Asp	Val	Gly	Trp	Ser	Leu	Ala	Ser	Thr	
				500					505					510			
	Arg	Ala	Ala	Phe	Glu	His	Arg	Ala	Val	Val	Val	Gly	Arg	Asp	Ser	Glu	
45			515				520						525				
	Glu	Leu	Val	Ser	Gly	Leu	Ser	Ser	Val	Ser	Gly	Val	Glu	Val	Gly	Val	
	530					535						540					
50	Gly	Val	Gly	Ala	Gly	Gly	Gly	Val	Val	Leu	Val	Phe	Ala	Gly	Gln	Gly	
	545				550						555					560	
	Cys	Gln	Trp	Val	Gly	Met	Gly	Arg	Glu	Leu	Leu	Gly	Ser	Ser	Leu	Val	
				565					570						575		
55	Phe	Ala	Glu	Ser	Met	Arg	Glu	Cys	Ala	Ala	Ala	Leu	Ser	Pro	Phe	Val	
				580				585						590			
	Asp	Phe	Ser	Val	Val	Asp	Val	Leu	Gly	Ser	Ala	Gly	Glu	Leu	Gly	Arg	
		595					600						605				
60	Val	Glu	Val	Val	Gln	Pro	Ala	Leu	Trp	Ala	Val	Met	Val	Ser	Leu	Ala	

ES 2 287 552 T3

	610		615		620	
5	Arg Val Trp Arg Ser Trp Gly Val Pro Val Ala Ala Val Val Gly His					
	625		630		635	640
	Ser Gln Gly Glu Ile Ala Ala Ala Thr Val Ala Gly Ala Leu Ser Val					
		645		650		655
10	Gly Asp Ala Ala Arg Val Val Ala Leu Arg Ser Arg Leu Ile Ala Glu					
		660		665		670
	Arg Leu Ser Gly Leu Gly Gly Met Val Ser Val Ala Leu Ser Arg Glu					
		675		680		685
15	Arg Val Val Ser Leu Ile Ala Gly Val Pro Gly Val Ser Val Ala Ala					
		690		695		700
	Val Asn Gly Ser Ser Ser Thr Val Val Ser Gly Glu Ala Ala Gly Leu					
		705		710		715
20	Glu Arg Val Leu Ala Ala Cys Val Ser Ser Gly Val Arg Ala Arg Arg					
		725		730		735
	Ile Asp Val Asp Tyr Ala Ser His Ser Val Gln Val Glu Leu Ile Arg					
		740		745		750
25	Glu Glu Leu Leu Gly Val Leu Asp Gly Ile Val Pro Arg Ser Gly Glu					
		755		760		765
	Ile Pro Phe Val Ser Thr Val Thr Gly Glu Arg Ile Asp Thr Val Glu					
		770		775		780
30	Leu Gly Ala Glu Tyr Trp Tyr Arg Asn Leu Arg Gln Thr Val Glu Phe					
		785		790		795
	Gln Ala Ser Val Gln Thr Leu Leu Ala Gln Gly His Gln Val Phe Leu					
		805		810		815
35	Glu Ser Ser Pro His Pro Val Leu Thr Val Gly Ile Glu Glu Thr Val					
		820		825		830
40	His Glu Ser Ala Ala Gln Ala Val Val Leu Gly Ser Leu Arg Arg Asp					
		835		840		845
	Glu Gly Ala Leu Thr Arg Leu Val Thr Ser Ala Gly Glu Ala Trp Ala					
		850		855		860
45	Arg Gly Val Pro Val Asp Trp Ala Gly Met Leu Ala Gly Gly Arg Arg					
		865		870		875
	Val Glu Leu Pro Thr Tyr Pro Phe Leu Arg Glu Arg Leu Trp Leu Glu					
		885		890		895
50	Pro Ser Arg Ser Arg Thr Gly Asn Leu Asn Met Ala Gly Leu Val Glu					
		900		905		910
	Ala Gly His Glu Ile Leu Pro Ala Ala Val Glu Leu Pro Gly Glu Gln					
		915		920		925
55	Trp Val Trp Thr Gly Glu Leu Ser Leu Ser Ala Tyr Pro Trp Leu Ala					
		930		935		940
60	Asp His Gln Val Leu Gly Gln Thr Leu Val Pro Gly Val Ala Trp Val					
		945		950		955
						960

65

ES 2 287 552 T3

	Glu Leu Ala Leu His Ala Gly His Gln Leu Gly Phe Gly Ser Val Glu	
	965	970 975
5	Glu Leu Thr Leu Gln Ala Pro Leu Val Leu Gly Glu Ser Asp Ala Val	
	980	985 990
	Gln Val Arg Val Val Val Ser Asp Leu Gly Glu Ser Asp Arg Arg Ala	
	995	1000 1005
10	Val Ser Val His Ser Arg Gly Asp Asp Gln Thr Trp Val Thr His Ala	
	1010	1015 1020
	Glu Gly Phe Leu Thr Ala Lys Gly Ala Gln Pro Glu Thr Met Ala Val	
	1025	1030 1035 1040
15	Trp Pro Pro Ser Gly Ala Glu Pro Val Glu Ala Asp Gly Phe Tyr Glu	
	1045	1050 1055
	Arg Leu Ala Asp Ala Gly Tyr His Tyr Gly Pro Val Phe Gln Gly Val	
20	1060	1065 1070
	Ser Lys Val Trp Arg Ala Gly Glu Glu Ile Tyr Ala Glu Val Gly Leu	
	1075	1080 1085
25	Leu Asp Asp Ala Asp Val Asp Gly Phe Gly Ile His Pro Ala Leu Leu	
	1090	1095 1100
	Asp Ala Ala Leu Gln Thr Ala Tyr Val Ala Gln Arg Gly Pro Ala Glu	
	1105	1110 1115 1120
30	Thr Lys Leu Pro Phe Ala Phe Gly Asp Val Gln Leu Phe Ala Thr Gly	
	1125	1130 1135
	Ala Arg Ser Leu Arg Val Arg Val Ser Pro Ala Ala Gln Gln Gly Met	
35	1140	1145 1150
	Ala Trp Glu Ala Trp Asp Pro Thr Gly Leu Pro Val Phe Ser Leu Gly	
	1155	1160 1165
40	Tyr Leu Ala Thr Arg Pro Val Asp Arg Gly Gln Leu Thr Val Lys Arg	
	1170	1175 1180
	Pro Glu Ser Leu Phe Lys Val Ala Trp Asp Glu Thr Val Pro Val Val	
	1185	1190 1195 1200
45	Gly Asn Ala Thr Ala Ala His Gly Val Val Leu Gly Asp Asp Pro Phe	
	1205	1210 1215
	Ala Leu Gly Ala Ala Leu Arg Ala Ala Gly Trp Glu Val Gly Ala Ala	
	1220	1225 1230
50	Pro Glu Pro Ala Ser Ala Asp Thr Ala Ala Glu Val Leu Leu Leu Pro	
	1235	1240 1245
	Cys Thr Ala Pro Gly Glu Pro Asp Ala Asp Leu Pro Thr Ala Val Arg	
55	1250	1255 1260
	Ala Val Thr Ala Arg Val Leu Gly Val Leu Gln Glu Trp Leu Ala Asp	
	1265	1270 1275 1280
60	Glu Arg Leu Ala Gly Thr Arg Leu Ala Val Val Thr Arg Asn Ala Leu	
	1285	1290 1295

65

ES 2 287 552 T3

Pro Gly Asp Leu Leu His Ser Pro Val Trp Gly Leu Val Arg Ser Ala
 1300 1305 1310
 5 Gln Thr Glu Asn Pro Gly Arg Ile Thr Leu Val Asp Leu Asp Asp His
 1315 1320 1325
 Pro Asp Ser Ala Ala Val Leu Ala Glu Ala Val Gln Ser Asp Glu Pro
 1330 1335 1340
 10 Arg Ile Met Val Arg Glu Gly Arg Pro Thr Ala Ala Arg Leu Val Arg
 1345 1350 1355 1360
 Ala Thr Ala Pro Glu Leu Val Pro Pro Ala Gly Ala Asp Ala Trp Arg
 1365 1370 1375
 15 Leu Glu Ile Thr Glu Pro Gly Thr Phe Asp Asn Leu Thr Leu Gly Val
 1380 1385 1390
 Tyr Pro His Ala Glu Lys Thr Leu Ala Asp Asn Glu Val Arg Val Ala
 1395 1400 1405
 20 Val His Ala Gly Gly Leu Asn Phe His Asp Val Val Ala Ala Leu Gly
 1410 1415 1420
 Met Val Glu Asp Asp Leu Thr Leu Gly Arg Glu Ala Ala Gly Val Val
 1425 1430 1435 1440
 25 Val Glu Val Gly Asp Ala Val Pro Asp Leu Thr Pro Gly Asp His Val
 1445 1450 1455
 Met Gly Ile Leu Ser Ser Gly Phe Gly Pro Leu Ala Val Thr Asp His
 1460 1465 1470
 30 Arg Tyr Leu Ala Arg Met Pro Glu Gly Trp Thr Phe Ala Gln Ala Ala
 1475 1480 1485
 35 Ser Val Pro Ala Ala Phe Leu Thr Ala Tyr Tyr Gly Leu Cys Asp Leu
 1490 1495 1500
 Gly Gly Ile Arg Ala Gly Asp Arg Val Leu Ile His Ala Ala Ala Gly
 1505 1510 1515 1520
 40 Gly Val Gly Met Ala Ala Val Gln Ile Ala Arg His Leu Gly Ala Glu
 1525 1530 1535
 Val Phe Gly Thr Ala Ser Pro Arg Lys Trp Gly Ala Leu Arg Ala Leu
 1540 1545 1550
 45 Gly Leu Asp Asp Ala His Leu Ser Ser Ser Arg Thr Leu Asp Phe Glu
 1555 1560 1565
 Gln Glu Phe Leu Asp Ala Thr Asp Gly Arg Gly Val Asp Leu Val Leu
 1570 1575 1580
 50 Asn Ser Leu Ala Arg Glu Phe Val Asp Ala Ser Leu Arg Leu Met Pro
 1585 1590 1595 1600
 55 Gly Gly Gly Arg Phe Val Asp Met Gly Lys Thr Asp Ile Arg Arg Pro
 1605 1610 1615
 Glu Gln Val Ala Glu Asp His Gly Gly Val Ala Tyr Gln Ala Phe Asp
 1620 1625 1630
 60 Leu Val Glu Ala Gly Pro Gln Arg Thr Gly Glu Met Leu Ala Glu Ile
 65

ES 2 287 552 T3

	1635	1640	1645
5	Val Arg Leu Phe Gln Ala Gly Ala Phe Arg Pro Leu Pro Ile Thr Gln 1650	1655	1660
	Trp Asp Val Arg Arg Ala Pro Glu Ala Phe Arg His Ile Ser Gln Ala 1665	1670	1675 1680
10	Lys His Ile Gly Lys Ile Val Leu Thr Val Pro Arg Pro Ile Asp Thr 1685	1690	1695
	Asp Gly Thr Val Met Val Thr Gly Ala Thr Gly Thr Leu Gly Gly Phe 1700	1705	1710
15	Val Ala Arg His Leu Val Thr His His Gly Ile Arg Arg Leu Leu Leu 1715	1720	1725
	Val Ser Arg Ser Ala Glu Arg Thr Asp Leu Val Arg Glu Leu Thr Glu 1730	1735	1740
20	Leu Gly Ala Asp Val Thr Trp Ala Ser Cys Asp Leu Ala Asp Ala Thr 1745	1750	1755 1760
	Ala Val Glu Glu Thr Val Arg Ser Val Asp Glu Arg His Pro Leu Val 1765	1770	1775
25	Ala Val Val His Ser Ala Gly Val Leu Asp Asp Gly Val Ile Asp Lys 1780	1785	1790
	Gln Ser Pro Glu Arg Leu Asp Thr Val Met Arg Pro Lys Val Asp Ala 1795	1800	1805
30	Ala Trp Asn Leu His Arg Leu Leu Asp Asn Ala Pro Leu Ala Asp Phe 1810	1815	1820
	Val Leu Phe Ser Ser Ala Ser Gly Val Leu Gly Gly Ala Gly Gln Ser 1825	1830	1835 1840
35	Asn Tyr Ala Ala Ala Asn Ala Phe Leu Asp Ala Leu Ala Glu His Arg 1845	1850	1855
40	Arg Ala Gln Gly Leu Ala Gly Gln Ala Leu Ala Trp Gly Leu Trp Ser 1860	1865	1870
	Asp Arg Ser Thr Met Thr Gly Gln Leu Gly Ser Thr Glu Leu Ala Arg 1875	1880	1885
45	Ile Ala Arg Asn Gly Val Ala Glu Met Ser Glu Thr Glu Gly Leu Ala 1890	1895	1900
	Leu Phe Asp Ala Ala Arg Asp Thr Ala Glu Ala Val Leu Leu Pro Met 1905	1910	1915 1920
50	His Leu Asp Val Ala Arg Leu Arg Ser Arg Asn Gly Glu Val Pro Ala 1925	1930	1935
	Val Phe Arg Arg Leu Ile His Ala Thr Ala Arg Arg Thr Ala Ser Thr 1940	1945	1950
55	Ala Val Arg Ser Ala Gly Leu Glu Gln Gln Leu Ala Ser Leu Ser Gly 1955	1960	1965
60	Pro Glu Arg Thr Glu Leu Leu Leu Gly Leu Val Arg Asp His Ala Ala 1970	1975	1980

65

ES 2 287 552 T3

Ala Val Leu Gly His Gly Thr Ser Asp Ala Val Ser Pro Asp Arg Pro
 1985 1990 1995 2000
 5 Phe Arg Asp Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn
 2005 2010 2015
 Arg Phe Ala Ala Leu Thr Gly Leu Arg Leu Pro Ala Thr Leu Val Phe
 2020 2025 2030
 10 Asp His Pro Ser Pro Thr Ala Leu Ala Gly His Leu Ala Gly Leu Leu
 2035 2040 2045
 Gly Ala Ala Thr Pro Ser Ala Ala Glu Pro Val Leu Ala Ala Val Gly
 15 2050 2055 2060
 Arg Leu Arg Ala Asp Leu Arg Ser Leu Thr Pro Asp Ala Glu Gly Ala
 2065 2070 2075 2080
 20 Glu Asp Val Thr Ile Gln Leu Glu Ala Leu Leu Ala Glu Trp Arg Glu
 2085 2090 2095
 Ala Ala Glu Lys Arg Ala Pro Glu Ala Val Gly Asp Glu Asp Leu Ser
 2100 2105 2110
 25 Thr Ala Thr Asp Asp Glu Ile Phe Ala Leu Val Asp Ser Glu Leu Gly
 2115 2120 2125
 Glu Ala
 30 2130
 <210> 7
 <211> 1742
 <212> PRT
 35 <213> *Streptomyces parvulus* Tü4055
 <400> 7
 40 Met Thr Ala Glu Ala Ser Gln Asp Lys Leu Arg Asp Tyr Leu Arg Lys
 1 5 10 15
 Thr Leu Ala Asp Leu Arg Thr Thr Lys Gln Arg Leu Arg Asp Thr Glu
 20 25 30
 45 Arg Arg Ala Thr Glu Pro Val Ala Ile Val Gly Met Ser Cys Arg Leu
 35 40 45
 Pro Gly Asp Val Arg Thr Pro Glu Arg Phe Trp Glu Leu Leu Asp Thr
 50 50 55 60
 Gly Thr Asp Ala Leu Thr Pro Leu Pro Thr Asp Arg Gly Trp Asn Leu
 65 70 75 80
 55 Asp Thr Ala Phe Asp Asp Glu Arg Pro Tyr Arg Arg Glu Gly Gly Phe
 85 90 95
 Leu Tyr Asp Ala Gly Arg Phe Asp Ala Glu Phe Phe Gly Ile Ser Pro
 100 105 110
 60 Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ser
 115 120 125
 Ser Trp Glu Ala Ile Glu His Ala Arg Ile Asp Pro Arg Ser Leu His
 65 130 135 140

ES 2 287 552 T3

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

65

ES 2 287 552 T3

	485	490	495
5	Ser Ala Lys 500	Ser Ser Pro Val Asp Val Gly Trp Ser Leu Val Ser Thr 505	
	Arg Ala Ala Phe Glu His Arg 515	Ala Val Val Val Gly Arg Gly Arg Asp 520	525
10	Glu Leu Val Arg Gly Leu Ser 530	Glu Val Ala Gln Gly Arg Gly Val Arg 535	540
	Gly Val Ala Ser Ser Ala Ser 545	Gly Gly Leu Ala Phe Val Phe Ala Gly 550	555
15	Gln Gly Ser Gln Arg Leu Gly Met Gly Arg Gly Leu Tyr Glu Arg Phe 565	570	575
	Pro Val Phe Ala Glu Ala Phe Asp Glu Val Cys Gly Arg Val Gly Pro 580	585	590
20	Gly Val Arg Glu Val Val Phe Gly Ser Asp Ala Gly Glu Leu Asp Arg 595	600	605
	Thr Val Trp Ala Gln Ala Gly Leu Phe Ala Leu Glu Val Ala Leu Phe 610	615	620
25	Arg Leu Leu Glu Ser Trp Gly Val Arg Pro Gly Cys Leu Ile Gly His 625	630	635
	Ser Val Gly Glu Leu Ser Ala Ala Cys Val Ala Gly Leu Trp Ser Leu 645	650	655
30	Glu Asp Ala Cys Arg Val Val Ala Ala Arg Ala Arg Leu Met Gln Ala 660	665	670
35	Leu Pro Ala Gly Gly Val Met Val Ala Val Arg Ala Glu Ala Gly Glu 675	680	685
	Leu Ala Gly Phe Leu Gly Glu Asp Val Val Ile Ala Ser Val Asn Ala 690	695	700
40	Pro Gly Gln Val Val Ile Ala Gly Pro Glu Gly Gly Val Glu Arg Val 705	710	715
	Val Ala Ala Cys Gly Ala Arg Ser Arg Arg Leu Ala Val Ser His Ala 725	730	735
45	Phe His Ser Pro Leu Val Glu Pro Met Leu Gly Glu Phe Arg Arg Val 740	745	750
	Val Glu Ser Val Ala Phe Gly Val Pro Ser Leu Arg Val Val Ser Asn 755	760	765
50	Val Thr Gly Ala Trp Val Asp Pro Glu Glu Trp Gly Thr Pro Glu Tyr 770	775	780
55	Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Ala 785	790	795
	Thr Leu Leu Asp Ala Gly Val Arg Thr Phe Val Glu Leu Gly Pro Ala 805	810	815
60	Gly Thr Leu Thr Ser Met Val Ser His Cys Ala Asp Ala Thr Ala Thr 820	825	830

ES 2 287 552 T3

	Ser Val Thr Ala Val Pro Thr Leu Arg Pro Asp His Asp Glu Ser Arg	
	835	840 845
5	Thr Val Leu Ser Ala Ala Ala Ser Leu Tyr Val Gln Gly His Pro Val	
	850	855 860
	Asp Trp Ala Pro Leu Phe Pro Arg Ala Arg Thr Val Asp Leu Pro Thr	
	865	870 875 880
10	Tyr Pro Phe Gln His Gln His Tyr Trp Met Glu Ser Ala Ala Arg Pro	
		885 890 895
	Thr Val Glu Asp Thr Pro Arg Glu Pro Leu Asp Gly Trp Thr His Arg	
		900 905 910
15	Ile Asp Trp Val Pro Leu Val Asp Glu Glu Pro Ala Pro Val Leu Ala	
		915 920 925
	Gly Thr Trp Leu Leu Val Arg Pro Glu Glu Gly Pro Arg Pro Leu Ala	
		930 935 940
20	Asp Ala Val Ala Asp Ala Leu Thr Arg His Gly Ala Ser Val Val Glu	
		945 950 955 960
	Ala Ala Arg Val Pro His Gln Ser Asp Thr Glu Leu Thr Gly Val Val	
		965 970 975
25	Ser Leu Leu Gly Pro Gly Ala Asp Gly Asp Gly Gly Leu Asp Ala Thr	
		980 985 990
30	Leu Arg Leu Val Gln Asp Leu Ala Thr Ala Gly Ser Thr Ala Pro Leu	
		995 1000 1005
	Trp Ile Val Thr Ser Gly Ala Val Ala Val Gly Thr Ser Asp Thr Val	
		1010 1015 1020
35	Pro Asn Pro Glu Gln Ala Thr Leu Trp Gly Leu Ala Arg Ala Ala Ala	
		1025 1030 1035 1040
	Thr Glu Trp Pro Gly Leu Gly Ala Ala Arg Ile Asp Leu Pro Ala Asp	
		1045 1050 1055
40	Leu Thr Glu Gln Val Gly Arg Arg Leu Cys Ala Arg Leu Leu Asp Arg	
		1060 1065 1070
	Ser Glu Gln Glu Thr Ala Val Arg Gln Ala Gly Val Phe Ala Arg Arg	
		1075 1080 1085
45	Leu Val Arg Ala Arg Thr Ser Asp Gly Arg Trp Thr Pro Arg Gly Thr	
		1090 1095 1100
50	Val Leu Val Thr Gly Gly Thr Gly Ala Leu Ala Gly His Val Ala Arg	
		1105 1110 1115 1120
	Trp Leu Ala Glu Glu Gly Ala Glu His Ile Val Leu Ala Gly Arg Arg	
		1125 1130 1135
55	Gly Pro Asp Gly Gln Gly Ala Glu Ala Leu Arg Ala Asp Leu Val Ala	
		1140 1145 1150
	Ala Gly Val Lys Ala Thr Ile Val Arg Cys Asp Val Ala Asp Arg Asp	
		1155 1160 1165

ES 2 287 552 T3

	Ala Val Arg Leu Leu Leu Asp Ala His Arg Pro Ser Ala Ile Val His	
	1170	1175 1180
5	Thr Ala Gly Val Val Asp Asp Gly Leu Leu Thr Ser Leu Thr Pro Ala	
	1185	1190 1195 1200
	Gln Val Glu Arg Val Leu Arg Pro Lys Leu Leu Gly Ala Arg Asn Leu	
		1205 1210 1215
10	His Glu Leu Thr Arg Asp Arg Glu Leu Asp Ala Phe Val Leu Phe Ser	
		1220 1225 1230
	Ser Leu Ala Gly Val Leu Gly Gly Ala Gly Gln Ala Asn Tyr Ala Ala	
		1235 1240 1245
15	Ala Asn Ala Tyr Leu Asp Ala Leu Ala Ala His Arg Thr Ala His Gly	
		1250 1255 1260
	Leu Pro Ala Ala Ser Leu Ala Trp Gly Pro Trp Glu Gly Asp Gly Met	
20		1265 1270 1275 1280
	Ala Ala Ala Gln Glu Ala Ala Asp Arg Leu Arg Arg Ser Gly Leu Thr	
		1285 1290 1295
25	Pro Leu Pro Pro Glu Gln Ala Val Arg Ala Leu Gly Arg Gly His Gly	
		1300 1305 1310
	Pro Leu Val Val Ala Asp Ala Asp Trp Ala Arg Leu Ala Ala Gly Ser	
		1315 1320 1325
30	Thr Gln Arg Leu Leu Asp Glu Leu Pro Glu Val Arg Ala Val Arg Pro	
		1330 1335 1340
	Ala Glu Pro Ala Val Gly Gln Arg Pro Asp Leu Pro Ala Arg Leu Ala	
		1345 1350 1355 1360
35	Gly Arg Pro Ala Glu Glu Gln Ser Ala Val Leu Leu Glu Ala Val Arg	
		1365 1370 1375
	Glu Glu Ile Ala Ala Val Leu Arg Tyr Ala Asp Pro Ala Arg Ile Gly	
40		1380 1385 1390
	Ala Asp His Glu Phe Leu Ala Leu Gly Phe Asp Ser Leu Thr Ser Ile	
		1395 1400 1405
	Glu Leu Arg Asn Arg Leu Ala Thr Arg Ile Gly Leu Thr Leu Pro Ala	
45		1410 1415 1420
	Thr Leu Thr Leu Glu Gln Arg Thr Pro Ala Gly Leu Ala Ala His Leu	
		1425 1430 1435 1440
50	Arg Glu Arg Ile Ala Asp Arg Pro Val Gly Ser Gly Ala Val Pro Val	
		1445 1450 1455
	Pro Gly Ser Ala Asp Val Pro Glu Ala Gly Gly Gly Ser Gly Leu Gly	
		1460 1465 1470
55	Glu Leu Trp Gln Glu Ala Asp Arg His Gly Arg Arg Leu Glu Phe Ile	
		1475 1480 1485
	Asp Val Leu Thr Ala Ala Ala Phe Arg Pro Ala Tyr Arg Glu Pro	
60		1490 1495 1500
	Ala Glu Leu Glu Leu Pro Pro Leu Arg Leu Thr Ser Gly Gly Asp Glu	

65

ES 2 287 552 T3

1505 1510 1515 1520

Pro Pro Leu Phe Cys Ile Pro Ser His Leu Gly Lys Ala Asp Pro His
1525 1530 1535

Lys Phe Leu Arg Phe Ala Ala Ala Leu Arg Gly Arg Arg Asp Val Phe
1540 1545 1550

Val Leu Arg Gln Pro Gly Phe Val Pro Gly Gln Pro Leu Pro Ala Gly
1555 1560 1565

Leu Asp Val Leu Leu Asp Thr His Ala Arg Ala Met Ala Gly His Asp
1570 1575 1580

Arg Pro Val Leu Leu Gly Tyr Ser Ala Gly Gly Leu Ala Ala Gln Ala
1585 1590 1595 1600

Leu Ala Ala Arg Leu Ala Glu Leu Gly Arg Pro Pro Ala Ala Val Val
1605 1610 1615

Leu Val Asp Thr Tyr Ala Pro Asp Glu Thr Glu Val Met Ala Arg Ile
1620 1625 1630

Gln Gly Ala Met Glu Gln Gly Gln Arg Asp Arg Asp Gly Arg Thr Gly
1635 1640 1645

Ala Ala Phe Gly Glu Ala Trp Leu Thr Ala Met Gly His Tyr Phe Gly
1650 1655 1660

Phe Asp Trp Thr Pro Cys Pro Val Asp Val Pro Val Leu His Val Arg
1665 1670 1675 1680

Ala Gly Asp Pro Met Thr Gly Met Pro Val Glu Gly Arg Trp Gln Ala
1685 1690 1695

Arg Trp Asn Leu Pro His Thr Ala Val Asp Val Pro Gly Asp His Phe
1700 1705 1710

Thr Met Met Glu Asp His Ala Pro Arg Thr Ala Asp Thr Val His Asp
1715 1720 1725

Trp Leu Gly Thr Ala Val Arg Arg Pro Glu Arg Thr Arg Asp
1730 1735 1740

<210> 8
<211> 264
<212> PRT
<213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 8

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

55 $\langle 210 \rangle$ 9

<211> 265

<212> PRT

<213> *Streptomyces parvulus* Tü4055

65

ES 2 287 552 T3

<400> 9

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

<210> 10

<211> 250

<212> PRT

60 <213> *Streptomyces parvulus* Tü4055

65

ES 2 287 552 T3

<400> 10

```

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

145          150          155          160
35
Val Ile Gly Phe Thr Lys Gln Cys Ala Tyr Glu Leu Ala Pro His Gly
          165          170          175
Val Thr Val Asn Ala Val Ala His Gly Val Ile Gly Thr Glu Arg Ile
          180          185          190
40
Arg Arg Ala Trp Glu Asp Lys Pro Pro Gln Trp Thr Ala Asp Arg Val
          195          200          205
Ser His Ile Pro Ala Gly Arg Phe Gly Ser Val Ala Glu Ala Ala Gly
          210          215          220
45
Met Val Cys His Leu Cys Gly Glu Asp Ala Gly Tyr Thr Thr Gly Thr
          225          230          235          240
50
Val Val Asp Val Asn Gly Gly Leu His Ile
          245          250

```

<210> 11

55 <211> 390

<212> PRT

<213> *Streptomyces parvulus* Tü4055

60

65

ES 2 287 552 T3

<400>11

	Met	Ile	Arg	Arg	Val	Arg	Leu	His	Thr	Ala	Val	Val	Pro	Met	Ala	Ala	
	1				5					10					15		
5	Ala	Phe	Asp	His	Ala	Thr	Arg	Ser	Arg	Arg	Ser	Ala	Ala	Ser	Leu	Leu	
				20					25					30			
	Val	Glu	Ile	Glu	Leu	Ala	Gly	Thr	Arg	Gly	Trp	Gly	Glu	Gly	Ala	Pro	
10			35					40					45				
	Arg	Asp	Tyr	Val	Thr	Gly	Glu	Thr	Leu	Asp	Gly	Ala	Val	Arg	Ala	Val	
		50					55					60					
15	Gln	Ala	Cys	Asp	Pro	Gly	Glu	Leu	Ala	Glu	Arg	Ile	Glu	Trp	Arg	Asp	
	65					70				75						80	
	Phe	Glu	Ser	Ala	Val	Ala	Ser	Ile	Ala	Gln	Leu	Pro	Leu	Thr	Gly	Leu	
					85					90					95		
20	Val	Asp	Gly	Ser	Ser	Ala	Ala	Ala	Ala	Val	Glu	Ile	Ala	Leu	Leu	Asp	
				100					105					110			
	Ala	Val	Cys	Arg	His	Phe	Ala	Arg	Pro	Leu	Ala	Asp	Val	Leu	Arg	Val	
25			115					120					125				
	Leu	Ala	Pro	Pro	Ala	Arg	Ser	Arg	Arg	Asp	Gly	Pro	Thr	Ser	Val	Ser	
		130					135					140					
30	Leu	Val	Ile	His	Leu	Ser	Arg	Asp	Val	Ala	Thr	Val	Leu	Asp	Ala	Leu	
	145					150				155						160	
	Thr	Pro	Arg	Ala	Leu	Ala	Ala	Leu	Arg	His	Val	Lys	Ile	Lys	Val	Ala	
				165					170						175		
35	Asp	Pro	Ala	Gly	Ala	Val	Asp	Arg	Leu	Thr	Ala	Ala	Gln	Asp	Arg	Leu	
			180						185					192			
	Pro	Ala	Asp	Thr	Arg	Val	Ser	Leu	Asp	Val	Asn	Gly	Ala	Trp	Thr	Ala	
40			195					200					205				

ES 2 287 552 T3

	Glu	Glu	Ala	Glu	Lys	Val	Ala	Gly	Glu	Leu	Asp	Gly	Val	Gly	Trp	Val	
	210						215					220					
5	Glu	Glu	Pro	Leu	Pro	Pro	Arg	Ser	Trp	Pro	Glu	Leu	Gly	Arg	Leu	Arg	
	225					230					235					240	
	Arg	Ala	Thr	Gly	Leu	Pro	Val	Met	Leu	Asp	Glu	Ser	Cys	Thr	Gly	Pro	
					245					250					255		
10	Ala	Asp	Leu	His	Ala	Ala	Ala	Thr	Ser	Gly	Ala	Ala	Ser	His	Ile	Asn	
				260					265					270			
	Val	Arg	Leu	Ser	Lys	Cys	Gly	Gly	Phe	Leu	Ala	Ala	Ala	Arg	Leu	Ala	
15			275					280					285				
	Leu	Arg	Ala	Asp	Glu	Leu	Gly	Val	Gly	Cys	Gln	Leu	Gly	Val	His	Val	
	290						295					300					
20	Ala	Glu	Val	Gly	Pro	Leu	Trp	Ala	Ala	Gly	Arg	Thr	Leu	Ala	Thr	Ala	
	305					310					315					320	
	Trp	Asp	Leu	Trp	Gln	Thr	Val	Glu	Ala	Gly	Arg	Ala	Asp	Glu	Trp	Phe	
					325					330					335		
25	Pro	Val	Pro	Leu	Thr	Thr	Pro	Ala	Phe	Thr	Val	Asp	Arg	Ser	Leu	His	
				340					345					350			
	Arg	Val	Glu	Pro	Leu	Thr	Gly	Pro	Gly	Thr	Gly	Ile	Glu	Pro	Thr	Glu	
			355					360					365				
30	Glu	Leu	Leu	Arg	His	Thr	Arg	Cys	Ala	Ala	Thr	Trp	Glu	Ser	Gly	Gly	
	370						375					380					
35	Gly	Trp	Arg	Arg	Asn	Thr											
	385				390												

<210> 12

<211> 272

40 <212> PRT

<213> *Streptomyces parvulus* Tü4055

45

50

55

60

65

ES 2 287 552 T3

<400> 12

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

<210> 13

<211> 539

<212> PRT

<213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 13

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

35

40

45

50

55

60

65

ES 2 287 552 T3

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

ES 2 287 552 T3

Met Gly Ala Arg Ala Leu Thr Val Arg Lys Ala Glu Glu Leu Asp Ala
 500 505 510
 Val Gly Asp Leu Leu Ser Thr Leu Glu Gly Pro Leu Leu Val Asp Val
 5 515 520 525
 Arg Leu Asp Pro Asp Val Asn Leu Val Arg Tyr
 530 535
 10 <210> 14
 <211> 683
 <212> PRT
 <213> *Streptomyces parvulus* Tü4055
 15 <400> 14
 Met Arg Gly Ala Arg Glu Asn Ser Met Thr Arg Ala Gly Pro Leu Glu
 20 1 5 10 15
 Gly Ile Ala Val Leu Met Ala Gly Arg Ser Thr Pro Ala Ala Leu Leu
 20 20 25 30
 Gly Arg Leu Leu Ala Asp Leu Gly Ala Arg Val Val Thr Leu Cys Arg
 25 35 40 45
 Ser Pro Asp His Gly Gly Pro Phe Glu Arg Trp Leu His Ser Ala Ala
 50 55 60
 30 Gln Ser Ala Ser Gly Trp Asp Gln Ala Ser Arg Leu Leu Gln Thr Ala
 65 70 75 80
 Asp Val Leu Val Cys Asp Ala Glu Gly Asp Glu Arg Leu Ala Ala Leu
 35 85 90 95
 Gly Leu Gly Ala Pro Glu Leu Pro His Arg Ser Pro Glu Leu Val Ala
 100 105 110
 Val Arg Leu Ser Ala Phe Gly Leu Thr Gly Pro Leu Arg Asp Ala Pro
 40 115 120 125
 Ala Thr Glu Arg Thr Leu Gln Ala Leu Ala Gly Leu Thr Ser Ala Thr
 45 130 135 140
 Gly Thr Glu Gly Glu Pro Ser Val Leu Ser Val Val Gly Leu Ala Ser
 145 150 155 160
 Arg Thr Ala Ala Leu Ser Gly Leu Ile Ala Val Val Ala Gly Leu Ile
 50 165 170 175
 Gly Arg Glu Arg Gly Gly Gly Gly Asp Tyr Leu Asp Ile Ala Glu Phe
 180 185 190
 55 Asp Ser Leu Phe Thr Leu Thr Gly Thr Leu Leu Pro Ser Val Ala Leu
 195 200 205
 Ala Gly Arg Pro Pro Arg Arg Thr Gly Asn Arg His Gly Met Ala Ala
 210 215 220
 Pro Trp Asn Ser Tyr Thr Cys Gln Asp Ala Pro Val Val Ile Cys Thr
 225 230 235 240
 65 Met Gly Glu Pro Ile Trp His Arg Leu Thr Ala Val Leu Gly Arg Arg
 245 250 255

ES 2 287 552 T3

	Asp	Leu	Pro	Asp	Asp	Pro	Arg	Phe	Ala	Asp	Thr	Ala	Ala	Arg	Val	Arg	
				260					265					270			
5	Asn	Ala	Asp	Glu	Leu	Asp	Glu	Ile	Leu	Gly	Lys	Trp	Thr	Ala	Gly	Gln	
			275					280					285				
	Arg	Ala	Val	Asp	Val	Val	Thr	Ala	Leu	Arg	Ala	Ala	Gly	Ile	Pro	Cys	
			290				295					300					
10	Ala	Gln	Val	Ala	Ala	Pro	Glu	Glu	Val	Arg	Asp	Gly	Ala	Ala	Ala	Arg	
						310					315					320	
	Arg	Arg	Gly	Leu	Val	Thr	Asp	Pro	Ser	Gly	Thr	Pro	Gly	Ser	Pro	Leu	
				325						330					335		
15	Arg	Ser	Leu	Ile	Pro	Ala	Val	Thr	Asp	Gly	Pro	Met	Pro	Arg	Gln	Gly	
				340					345					350			
	Gly	Leu	Trp	Glu	Pro	Ile	Ala	Arg	Gly	Thr	Pro	Pro	Leu	Arg	Gly	Val	
20			355					360					365				
	Arg	Leu	Leu	Glu	Val	Gly	Ser	Tyr	Thr	Ala	Gly	Pro	His	Ala	Gly	Arg	
							375					380					
25	Leu	Leu	Ala	Gln	Leu	Gly	Ala	Asp	Val	Leu	Lys	Val	Glu	Pro	Pro	His	
						390					395					400	
	Gly	Glu	Gly	Ser	Arg	Arg	Leu	Ala	Gln	Gln	Val	Ala	Gly	Val	Gly	Tyr	
					405				410					415			
30	Leu	Tyr	Tyr	Val	Asn	Asn	Ala	Gly	Lys	Arg	Ser	Cys	Arg	Leu	Asp	Leu	
				420					425					430			
	Ala	Asp	Ala	Glu	Asp	Arg	Ala	Gly	Phe	Glu	Arg	Leu	Leu	Ala	Gly	Cys	
35				435				440					445				
	Asp	Ile	Val	Leu	Thr	Asn	Leu	Ala	Ala	Asp	Thr	Leu	Thr	Ala	Gln	Gly	
				450			455					460					
	Leu	Ala	Pro	Asp	Gln	Ile	Leu	Ser	Arg	His	Gly	Val	Val	His	Cys	Thr	
40					470						475					480	
	Val	Thr	Gly	His	Gly	Leu	Ala	Ala	Ala	Asp	Arg	Ser	Val	Asp	Thr	Val	
					485					490				495			
45	Ile	Gln	Ala	Glu	Ser	Gly	Ile	Met	Arg	Leu	Val	Gly	Gly	Pro	Gly	Ala	
				500				505						510			
	Gly	Leu	Arg	Thr	Pro	Val	Ser	Ser	Ala	Asp	Val	Leu	Gly	Ala	Tyr	Leu	
				515				520					525				
50	Ala	Ala	Ala	Ala	Ala	Val	Val	Ser	Thr	Tyr	Val	Arg	Leu	Arg	Thr	Glu	
							535					540					
	His	Gly	Cys	Ala	Ala	Asp	Val	Ala	Leu	Phe	Asp	Ser	Ala	Val	Trp	Leu	
55						550					555					560	
	Thr	Gln	Asp	Arg	Trp	Phe	Thr	Ala	Pro	Pro	Ala	Arg	Ala	Pro	His	Leu	
					565					570				575			
60	Val	Arg	Ala	Ala	Asp	Gly	Thr	Val	Leu	Val	Asp	Ala	Glu	Gly	Pro	Pro	
				580					585				590				
	Pro	Arg	Ala	Glu	Gly	Pro	Val	Ala	Ala	Val	Leu	Asp	Ala	Ala	Ala	Ala	

65

ES 2 287 552 T3

	595	600	605
5	Val Gly Val Pro Ala Ala Pro Leu His Asp Leu Thr Arg Ala Val Arg 610 615 620		
	His Pro Gln Val Leu Ala Arg Arg Met Ala Val Ala Arg Asp Cys Ala 625 630 635 640		
10	Gly Thr Thr Val Leu Ile Thr Gly Asn His Leu Arg Ser Leu Leu Arg 645 650 655		
	Glu Asp Pro Pro Pro Thr Cys Ala Pro Val Asp Gln Asn Asp Pro Val 660 665 670		
15	Trp Leu Gln Pro Ala Pro Thr Glu Gly Gln Gln 675 680		
	<210> 15		
20	<211> 426		
	<212> PRT		
	<213> <i>Streptomyces parvulus</i> Tü4055		
25	<400> 15		
	Met Thr Thr Arg Arg Arg Gln Arg His Pro Ala Leu Ser Pro Ser Cys 1 5 10 15		
30	Pro Ser Val Pro Phe Pro Leu Leu Glu Thr Glu Phe Val Leu Met Pro 20 25 30		
	Ser Phe Pro Val Arg Arg Ser Val Pro Asp Thr Pro Pro Ala Glu His 35 40 45		
35	Leu Glu Leu Leu Lys Glu Ser Gly Gly Val Cys Pro Phe Thr Met Glu 50 55 60		
	Asp Gly Arg Pro Ala Trp Leu Ala Ala Ser His Asp Ala Val Arg Ser 65 70 75 80		
40	Leu Leu Ala Asp Arg Arg Ile Ser Asn Asn Pro Ala Lys Thr Pro Pro 85 90 95		
	Phe Ser Gln Arg Glu Ala Leu Gln Lys Glu Arg Gly Gln Phe Ser Arg 100 105 110		
45	His Leu Phe Asn Met Asp Ser Pro Glu His Asp Val Ala Arg Arg Met 115 120 125		
50	Ile Ala Glu Asp Phe Thr Pro Arg His Ala Glu Ala Val Arg Pro Tyr 130 135 140		
	Phe Glu Glu Val Phe Gly Glu Ile Val Asp Glu Val Val His Lys Gly 145 150 155 160		
55	Pro Pro Ala Glu Met Ile Glu Ser Phe Ala Phe Pro Val Ala Thr Arg 165 170 175		
	Thr Ile Cys Lys Val Leu Asp Ile Pro Glu Asp Asp Cys Glu Tyr Phe 180 185 190		
60	Gln Lys Arg Thr Glu Gln Ile Ile Glu Met Asp Arg Gly Glu Glu Asn 195 200 205		
65	Leu Glu Ala Val Val Glu Leu Arg Arg Tyr Val Asp Ser Val Met Gln 210 215 220		

ES 2 287 552 T3

	Gln	Arg	Thr	Ar	G	Lys	Pro	Gly	Asp	Asp	Leu	Leu	Ser	Arg	Met	Ile	Val	
	225						230					235					240	
5	Lys	Ala	Lys	Ala	Ser	Lys	Glu	Ile	Glu	Leu	Ser	Asp	Ala	Asp	Leu	Val		
					245					250					255			
	Asp	Asn	Ala	Met	Phe	Leu	Leu	Val	Ala	Gly	His	Glu	Pro	Ser	Ala	Asn		
				260					265						270			
10	Met	Leu	Gly	Leu	Gly	Val	Leu	Ala	Leu	Ala	Glu	Phe	Pro	Asp	Val	Ala		
			275					280						285				
	Glu	Glu	Leu	Arg	Ala	Glu	Pro	His	Leu	Trp	Pro	Gly	Ala	Ile	Asp	Glu		
15			290				295						300					
	Met	Leu	Arg	Tyr	Tyr	Thr	Ile	Ala	Arg	Ala	Thr	Lys	Arg	Val	Ala	Ala		
	305					310					315					320		
20	Ala	Asp	Ile	Glu	Tyr	Glu	Gly	His	Thr	Ile	Lys	Glu	Gly	Asp	Ala	Val		
					325					330					335			
	Ile	Val	Leu	Leu	Asp	Thr	Ser	Asn	Arg	Asp	Pro	Lys	Val	His	Ala	Glu		
25				340					345					350				
	Pro	Asn	Arg	Leu	Asp	Ile	His	Arg	Ser	Ala	Gly	Asn	His	Leu	Ala	Phe		
			355					360					365					
30	Ser	His	Gly	Pro	His	Gln	Cys	Leu	Gly	Lys	His	Leu	Val	Arg	Val	Gln		
			370				375						380					
	Leu	Glu	Ile	Ala	Leu	Arg	Ala	Val	Ala	Glu	Arg	Leu	Pro	Gly	Leu	Arg		
	385					390					395					400		
35	Leu	Asp	Ile	Ala	Lys	Glu	Asp	Ile	Pro	Phe	Arg	Gly	Asp	Ala	Leu	Ser		
					405					410					415			
	Tyr	Gly	Pro	Arg	Gln	Leu	Arg	Val	Thr	Trp								
				420					425									
40	<210> 16																	
	<211> 454																	
	<212> PRT																	
45	<213> <i>Streptomyces parvulus</i> Tü4055																	
	<400> 16																	
50	Met	Glu	Lys	Thr	Asp	Val	Asp	Arg	Leu	Arg	Thr	Leu	Asp	Arg	Glu	His		
	1				5					10					15			
	Met	Trp	Tyr	Pro	Trp	Thr	Pro	Met	Thr	Glu	Trp	Met	Ala	Arg	Asp	Gln		
				20					25					30				
55	Leu	Val	Val	Glu	Arg	Ala	Glu	Gly	Cys	Trp	Leu	Ile	Asp	Ala	Asp	Gly		
			35					40					45					
	Lys	Arg	Tyr	Leu	Asp	Gly	Arg	Ser	Ser	Met	Gly	Met	Asn	Leu	His	Gly		
		50				55					60							
60	His	Gly	Arg	Ser	Glu	Ile	Val	Glu	Ala	Leu	Val	Ala	Gln	Ala	Arg	Lys		
	65					70					75					80		
	Ala	Gly	Glu	Thr	Thr	Leu	Tyr	Arg	Val	Ser	His	Pro	Ala	Ala	Val	Glu		
65					85					90					95			

ES 2 287 552 T3

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

ES 2 287 552 T3

<210> 17

<211> 326

<212> PRT

5 <213> *Streptomyces parvulus* Tü4055

<400>17

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

65

ES 2 287 552 T3

Thr Asp Ala Leu Glu Pro Val Val His Ala Asp Phe Pro Leu Ser Arg
 290 295 300
 5 Ala Pro Glu Ala Val Ala Glu Leu Glu Glu Arg Arg Ala Phe Gly Lys
 305 310 315 320
 Val Val Ile Arg Thr Ala
 325
 10 <210> 18
 <211> 556
 <212> PRT
 <213> *Streptomyces parvulus* Tü4055
 15 <400> 18
 Met Thr Gly Asn Thr Thr Ser Ala Ala Phe Leu Arg Arg Thr Gln Asn
 1 5 10 15
 20 Ala Leu Ala Met Gln Arg Lys Ile Cys Ala Gln Pro Glu Glu Thr Ala
 20 25 30
 25 Glu Arg Val Phe Ser Asp Ile Leu Ser Val Ser Arg Asp Thr Gly Phe
 35 40 45
 Gly Arg Glu His Gly Leu Ala Gly Val Arg Thr Arg Gln Glu Trp Arg
 50 55 60
 30 Arg Ala Val Pro Ile Arg Thr Tyr Asp Glu Leu Ala Pro Tyr Val Glu
 65 70 75 80
 Arg Gln Phe Ser Gly Glu Arg Arg Val Leu Thr Thr Asp Asp Pro Arg
 85 90 95
 35 Ala Phe Leu Arg Thr Ser Gly Ser Thr Gly Arg Ala Lys Leu Val Pro
 100 105 110
 40 Thr Thr Asp His Trp Arg Arg Val Tyr Arg Gly Pro Ala Leu Tyr Ala
 115 120 125
 Gln Trp Gly Leu Tyr Phe Glu Gln Ile Gly Thr His Arg Leu Thr Gly
 130 135 140
 45 Asp Glu Val Leu Asp Leu Ser Trp Glu Pro Gly Pro Ile Arg His Arg
 145 150 155 160
 Leu Arg Gly Phe Pro Val Tyr Ser Ile Thr Glu Arg Pro Val Ser Asp
 165 170 175
 50 Asp Pro Asp Asp Trp Asn Pro Pro Trp Arg His Ala Arg Trp Phe Thr
 180 185 190
 55 Arg Asp Ala Gly Ala Ala Thr Met Ala Asp Leu Leu Tyr Gly Lys Leu
 195 200 205
 Leu Arg Leu Ala Ala His Asp Leu Arg Leu Ile Val Ser Val Asn Pro
 210 215 220
 60 Ser Lys Ile Val Leu Leu Ala Glu Thr Leu Lys Glu Asn Ala Glu Arg
 225 230 235 240
 65 Leu Ile Gln Asp Leu His Asp Gly His Gly Thr Asp Arg Ala Ala Arg
 245 250 255

ES 2 287 552 T3

	Pro	Asp	Phe	Leu	Arg	Arg	Leu	Thr	Ala	Ala	Phe	Asp	Arg	Thr	Gly	Gly	
				260					265					270			
5	Arg	Pro	Leu	Leu	Thr	Asp	Leu	Trp	Pro	Gly	Leu	Arg	Leu	Leu	Val	Cys	
			275					280					285				
	Trp	Asn	Ser	Ala	Ser	Ala	Ala	Leu	Tyr	Gly	Pro	Trp	Leu	Ser	Arg	Leu	
		290					295					300					
10	Ala	Thr	Gly	Val	Ala	Ala	Leu	Pro	Phe	Ser	Thr	Thr	Gly	Thr	Glu	Gly	
	305					310					315					320	
	Ile	Val	Thr	Leu	Pro	Val	Asp	Asp	His	Leu	Ser	Ala	Gly	Pro	Leu	Ala	
15					325					330					335		
	Val	Asp	Gln	Gly	His	Phe	Glu	Phe	Val	Pro	Trp	Gln	Asp	Leu	Asp	Asp	
				340					345					350			
20	Gly	Ser	Pro	Leu	Pro	Glu	Asp	Thr	Pro	Thr	Leu	Gly	Tyr	Asp	Glu	Leu	
			355					360					365				
	Glu	Leu	Gly	Ala	Asp	Tyr	Arg	Leu	Val	Met	Ser	Gln	Ala	Asn	Gly	Leu	
		370					375					380					
25	Tyr	Arg	Tyr	Asp	Val	Gly	Asp	Val	Tyr	Arg	Val	Val	Gly	Ala	Val	Gly	
	385					390					395					400	
	Ala	Thr	Pro	Arg	Leu	Glu	Phe	Leu	Gly	Arg	Ala	Gly	Phe	Gln	Ser	Ser	
30					405				410						415		
	Phe	Thr	Gly	Glu	Lys	Leu	Thr	Glu	Ser	Asp	Val	His	Thr	Ala	Val	Met	
				420					425					430			
35	Arg	Val	Leu	Gly	Ser	Glu	Arg	Thr	Asp	His	Pro	His	Phe	Ser	Gly	Ile	
			435					440					445				
	Pro	Val	Trp	Asp	Thr	Pro	Pro	His	Tyr	Leu	Val	Ala	Ile	Glu	Trp	Ala	
		450					455					460					
40	Asp	Ala	His	Gly	Thr	Leu	Asn	Val	Gln	Asp	Thr	Ala	Arg	Arg	Ile	Asp	
	465					470					475				480		
	Ala	Thr	Leu	Gln	Glu	Val	Asn	Val	Glu	Tyr	Ala	Asp	Lys	Arg	Arg	Ser	
45					485					490					495		
	Gly	Arg	Leu	Arg	Pro	Leu	Gln	Ile	Leu	Pro	Leu	Val	Pro	Gly	Ala	Phe	
					500				505					510			
50	Gly	Gln	Ile	Ala	Glu	Arg	Arg	Phe	Arg	Gln	Gly	Thr	Ala	Gly	Ala	Gln	
			515					520					525				
	Ile	Lys	His	His	Trp	Leu	Gln	Lys	Asp	Ser	Ala	Phe	Leu	Asp	Thr	Leu	
		530					535					540					
55	Arg	Asp	Leu	Asp	Leu	Val	Arg	Ala	Arg	Pro	Gly	Thr					
	545					550					555						
	<210>	19															
	<211>	305															
60	<212>	PRT															
	<213>	<i>Streptomyces parvulus</i>	Tü4055														
	<400>	19															
65	Met	Arg	Ile	Gly	Phe	Ala	Ala	Pro	Met	Ser	Gly	Pro	Trp	Ala	Thr	Pro	

ES 2 287 552 T3

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

<210> 20

65 <211> 248

<212> PRT

<213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 20

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

<210> 21

<211> 675

<212> PRT

<213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 21

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

ES 2 287 552 T3

	385		390		395		400
5	Gly Phe Ser Gln Asn Asp Ala His Ile Tyr Cys Thr Ala Asp Gln Ala	405	410	415			
	Lys Asp Gln Phe Leu Glu Val Met Arg Met His Ala Asp Tyr Tyr Arg	420	425	430			
10	Thr Leu Gly Ile Ser Asp Phe Tyr Met Val Leu Ala Leu Arg Asp Ser	435	440	445			
	Ala Asn Lys Asp Lys Tyr His Asp Asp Glu Gln Met Trp Glu Asp Ala	450	455	460			
15	Glu Arg Ile Thr Arg Glu Ala Met Glu Glu Ser Asp Ile Pro Phe Gln	465	470	475		480	
	Ile Asp Leu Gly Gly Ala Ala His Tyr Gly Pro Lys Val Asp Phe Met	485	490	495			
20	Ile Arg Ala Val Thr Gly Lys Glu Phe Ala Ala Ser Thr Asn Gln Val	500	505	510			
	Asp Leu Tyr Thr Pro Gln Arg Phe Gly Leu Thr Tyr His Asp Ser Asp	515	520	525			
25	Gly Thr Glu Lys Pro Val Val Val Ile His Arg Ala Pro Leu Gly Ser	530	535	540			
	His Glu Arg Phe Thr Ala Tyr Leu Thr Glu His Phe Ala Gly Ala Phe	545	550	555		560	
	Pro Val Trp Leu Ala Pro Glu Gln Val Arg Ile Ile Pro Ile Val Glu	565	570	575			
35	Glu Leu Thr Asp Tyr Ala Glu Glu Val Arg Asp Met Leu Leu Asp Ala	580	585	590			
	Asp Val Arg Ala Asp Val Asp Ala Gly Asp Gly Arg Leu Asn Ala Lys	595	600	605			
40	Val Arg Ala Ala Val Thr Arg Lys Ile Pro Leu Val Val Val Val Gly	610	615	620			
	Arg Arg Glu Ala Glu Gln Arg Thr Val Thr Val Arg Asp Arg Ser Gly	625	630	635		640	
	Glu Glu Thr Pro Met Ser Leu Glu Lys Phe Val Ala His Val Thr Gly	645	650	655			
50	Leu Ile Arg Thr Lys Ser Leu Asp Gly Ala Gly His Ile Arg Pro Leu	660	665	670			
	Ser Lys Ala	675					

<210> 22

<211> 103

60 <212> PRT

<213> *Streptomyces parvulus* Tü4055

65

ES 2 287 552 T3

<400> 22

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

<210> 23

<211> 868

25 <212> PRT

<213> *Streptomyces parvulus* T34055

<400> 23

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

ES 2 287 552 T3

	Arg	Thr	Pro	Leu	Ser	Pro	Gln	Asp	Arg	Asp	Asp	Leu	Leu	Thr	Leu	Leu	
	210						215					220					
5	Pro	Leu	Thr	Pro	Ala	Gly	Arg	Gly	Asp	Leu	Pro	Gln	Asp	Ile	Pro	Val	
	225					230					235					240	
	Arg	Glu	Thr	Lys	Ala	Leu	Val	Leu	Gly	Ala	Leu	Val	Arg	Arg	Ala	Pro	
					245					250					255		
10	Ser	Arg	Pro	Ala	Leu	Arg	Arg	Leu	Leu	Ala	Glu	Arg	Leu	Thr	Thr	Ala	
				260						265				270			
	Thr	Asp	Val	Leu	Arg	Leu	Leu	Ala	Val	Leu	Ser	Gly	Gly	Asp	Ala	Gly	
			275					280					285				
15	Leu	Val	Thr	Pro	Ala	Arg	Phe	Thr	Asn	Val	Pro	Arg	Ser	Leu	Arg	Arg	
	290						295					300					
	Asp	Leu	Leu	Ala	Val	Leu	Asp	Gly	Leu	Pro	Ala	Pro	Tyr	Leu	Val	Glu	
	305					310					315					320	
20	Asp	Met	Leu	Arg	His	Pro	Thr	Ala	Trp	Lys	Arg	Ala	Ala	Glu	Val	Leu	
					325					330					335		
	His	Pro	Phe	Glu	Gly	His	Thr	Arg	His	Pro	Arg	Ala	Ala	Leu	Ala	Thr	
25				340					345					350			
	Ala	Val	Leu	Arg	Ala	Thr	Pro	Leu	Asp	Pro	Asp	Thr	Ala	Phe	Gly	Ala	
			355					360					365				
30	Ala	Leu	Leu	Thr	Thr	Ala	Ala	Ala	His	Pro	Asp	Ala	Val	Arg	Pro	Asp	
	370					375						380					
	Gly	Thr	Arg	Val	Arg	Pro	Ala	Thr	Trp	Ala	Gly	Arg	Leu	Glu	Gln	Ala	
	385					390					395					400	
35	Met	Ala	Glu	Gly	Asp	Ala	Ala	Arg	Ala	Ala	Ala	Leu	Ala	Gly	Glu	Arg	
					405					410					415		
	Pro	Gly	Glu	Leu	Val	Arg	Arg	Leu	Asp	Val	Leu	Leu	Arg	Leu	His	Thr	
				420					425					430			
40	Asp	Glu	Ala	Leu	Val	Pro	Glu	Leu	Glu	Lys	Ala	Leu	Arg	His	Gly	Leu	
			435					440					445				
	Pro	Lys	Val	Gly	Pro	Gly	Pro	Leu	Leu	Ser	Ala	Leu	Gly	Ala	Leu	Arg	
45		450					455					460					
	Thr	Arg	Thr	Glu	Asp	Arg	Thr	Gly	Thr	Arg	Arg	Val	Phe	Phe	Pro	Arg	
	465					470					475					480	
	Gly	Asp	Val	Thr	Arg	Ala	Leu	Ser	Val	Pro	Glu	Arg	Arg	Pro	Ala	Leu	
50					485					490					495		
	Pro	Ala	Gly	Pro	Val	Ser	Glu	Val	Val	Ala	Leu	Leu	Glu	Gly	Glu	Leu	
				500					505					510			
55	Leu	Arg	Arg	Phe	Ala	Ala	Gly	Arg	Pro	Tyr	Glu	Leu	Ser	Val	Leu	Asp	
		515						520					525				
	Ala	Gly	Leu	Thr	Asp	Leu	Thr	Val	Pro	Phe	Thr	Glu	Arg	Thr	Ala	Ala	
	530					535						540					
60	Lys	Ala	Leu	Val	Thr	Val	Gly	Arg	Gly	Ser	Val	Gln	Ala	Leu	Pro	Glu	

ES 2 287 552 T3

	545		550		555		560
	Gly Ser Val	Leu Arg	Leu Phe	Leu His	Trp Thr	Glu Pro	Arg Gly Asn
		565			570		575
5	Arg Thr Asp	Leu Asp	Leu Ser	Val Ala	Phe Phe	Asp Ala	Glu Trp Thr
		580		585			590
10	Phe Thr Gly	Leu Cys	Asp Tyr	Thr Asn	Leu Val	His Gly	Pro Asp Ala
		595		600			605
	Ala Ile His	Ser Gly	Asp Leu	Thr Ser	Ala Pro	Ala Pro	Arg Gly Ala
		610		615		620	
15	Thr Glu Tyr	Val Asp	Leu Asp	Leu Glu	Arg Leu	Ala Arg	Arg Gly Asp
			630		635		640
	Thr Tyr Ala	Val Pro	Leu Val	Phe Ser	Tyr Asn	Asn Val	Pro Phe Glu
		645			650		655
20	Glu Leu Pro	Asp Ala	Phe Ala	Gly Phe	Met Ala	Leu Pro	Ala Glu Gly
		660		665			670
	Pro Arg Asp	Ala Thr	Tyr Asp	Pro Arg	Thr Val	Arg Gln	Arg Phe Asp
25		675		680		685	
	Leu Ala Gly	Asp Ser	Lys Val	Cys Leu	Pro Met	Ile Val	Asp Leu Ala
		690		695		700	
30	Arg Arg Arg	Ala Leu	Trp Thr	Asp Thr	His Leu	Pro Ser	Ala Gly Gly
			710		715		720
	Phe Gln Ser	Ile Gly	Ser His	Gly Gly	Gly Glu	Leu Ala	Ala Val Ala
		725		730			735
35	Gly Asp Leu	Trp Gln	Gln Phe	Thr Ser	Gly Gly	Arg Ala	Thr Leu Trp
		740		745			750
	Asp Leu Ala	Val Leu	Arg Ala	Ala Ala	Leu Ser	Pro Glu	Val Ala Val
40		755		760		765	
	Val Ser Arg	Glu Pro	Glu Pro	Ala Val	Leu Arg	Tyr Arg	Arg Arg Ala
		770		775		780	
45	Ala Glu Ser	Glu Ala	Ala Phe	Ala Val	Arg Val	Ala Ser	His Lys Asp
			790		795		800
	Ala Glu Glu	Arg Leu	Ala His	Thr Asp	Pro Asp	Ser Ala	Ala Ala Gly
		805			810		815
50	Leu Ala Ala	Gly Arg	Arg Val	Phe Leu	Ala Thr	Val His	Gly Asp Val
		820		825			830
	Arg Pro Pro	Gly Ala	Ser Gly	Thr Ser	Tyr Arg	Leu Phe	Pro Gly Ala
55		835		840		845	
	Gly Asp Ala	Ser Pro	Thr Leu	Thr Arg	Val Thr	Ala Gly	Asp Leu Leu
		850		855		860	
60	Ala Glu Leu	Gly					
		865					

<210> 24

<211> 212

<212> PRT

65 <213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 24

```

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

```

<210> 25

<211> 329

<212> PRT

<213> *Streptomyces parvulus* Tü4055

<400> 25

```

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

```

ES 2 287 552 T3

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

<210> 26

55 <211> 276

<212> PRT

<213> *Streptomyces parvulus* Tü4055

60

65

ES 2 287 552 T3

<400> 26

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

<210> 27

60 <211> 185

<212> PRT

<213> *Streptomyces parvulus* Tü4055

65

ES 2 287 552 T3

<400> 27

```

5      Met Gly Val Met Ile Gly Pro Ala Gly Arg Glu Arg Asp Glu Gly Asp
      1           5           10           15

      His Val Thr Gln Gln Ala Pro Val Ala Thr Asp Glu Arg Arg Val Phe
           20           25           30

10     Val Asp Lys Gln Thr Pro Gly Ala Tyr Lys Ala Phe Val Ala Ala Ala
           35           40           45

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

15

      Val Glu Leu Val Asn Ile Arg Val Ser Gln Leu Asn Ala Cys Ala Tyr
           65           70           75           80

20     Cys Leu Ser Leu His Thr Arg Ala Ala Leu Arg Ala Gly Glu Thr Thr
           85           90           95

      Gln Arg Leu Ala Val Leu Pro Ala Trp Arg Asp Thr Glu Leu Phe Thr
           100          105          110

25     Ala Arg Glu Arg Ala Ala Leu Ala Leu Ala Glu Ala Thr Thr Arg Pro
           115          120          125

      Ala Asp Ala Ala Ala Gln Ser Ala Ala Tyr Ala Gln Ala Arg Gly Val
           130          135          140

30     Leu Ser Asp Asp Glu Val Ser Ala Val Ile Trp Val Ala Ile Ser Ile
           145          150          155          160

35     Asn Ala Phe Asn Arg Val Ser Val Leu Ser Lys His Pro Val Arg Gly
           165          170          175

      Ala Ala Pro Ala Pro Val Ser Pro Ala
           180          185

```

<210> 28

<211> 324

<212> PRT

<213> *Streptomyces parvulus* Tü4055

50

55

60

65

ES 2 287 552 T3

<400> 28

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

<210> 29

<211> 126

<212> PRT

ES 2 287 552 T3

<213> *Streptomyces parvulus* Tü4055

<400> 29

```

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

      Gly Trp Thr Phe Val Thr Asn His Ala Arg Val Leu Ala Ala Ile Ala
      20          25          30

10     Asp Asn Pro Asn Val Arg Ile Arg Asp Ile Ala Ala His Cys Arg Leu
      35          40          45

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

      Tyr Leu Ser His Thr Arg Asp Gly Arg Ser Asn Ile Tyr Arg Val Glu
      65          70          75          80

20     Pro Asp Lys Val Leu Arg His Pro Ala Glu Ala Gly Leu Thr Val Ala
      85          90          95

      Ala Leu Leu Ser Leu Leu Val Arg Asp Glu Thr Asp His Gly Arg Ser
      100         105         110

25     Ala Gly Pro Gly Ser Arg Pro Ala Arg Ser Ser Ala Ala Arg
      115         120         125

```

<210> 30

30 <211> 127

<212> PRT

<213> *Streptomyces parvulus* Tü4055

35 <400> 30

```

      Met Ser Leu Asp Glu Ala Val Ala Gly Cys Ser Arg His Thr Gly Arg
      1          5          10          15

40     Arg Arg Leu Pro Ala Ala Glu Gln Pro Thr Gln Ala Gln Tyr Glu Ala
      20          25          30

      His Gly Ala Trp Val Val Ser Ala Arg Gly Ala Tyr Asp Met Asn Ser
      35          40          45

45     Val Glu Pro Leu Ala Asp Ala Leu Lys Asp Ala Ala Glu Lys Ser Pro
      50          55          60

      Lys Val Val Leu Asp Ala Ser Gly Ile Thr Phe Ala Asp Ser Thr Leu
      65          70          75          80

      Leu Ser Leu Leu Ile Leu Thr His Gln Ala Thr Asp Phe Arg Val Ala
      85          90          95

55     Ala Pro Thr Trp Gln Val Met Arg Leu Met Gln Leu Thr Gly Val Asp
      100         105         110

      Ala Phe Leu Lys Val Arg Ala Thr Val Glu Glu Ala Ala Thr Ala
      115         120         125

```

60 <210> 31

<211> 82

<212> PRT

65 <213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 31

```

5      Met Ser Met Ile Leu Pro Ala Glu Lys Glu Leu Arg Ala Val Leu Ala
      1          5          10          15

      Arg Phe Ala Gln Ala Arg Ile Asp His Asp Val Arg Pro Ser Gly Cys
              20          25          30

10     Thr Ser Arg Leu Leu Glu Asp Ala Thr Tyr Thr Leu Cys Val Met Thr
          35          40          45

      Gly Ala Arg Thr Ala Glu Gln Ala Leu Arg Thr Ala Asp Glu Leu Leu
          50          55          60

15     Ala Gln Phe Ala Glu Arg Thr Ala Ala Pro Val Glu Asp Glu Ala Leu
      65          70          75          80

20     Ala Ala

```

<210> 32

<211> 70

25 <212> PRT

<213> *Streptomyces parvulus* Tü4055

30 <400> 32

```

      Met Ser Asp Thr Arg Leu Arg Gln Arg Asp Glu Thr Ser Lys Gly Pro
      1          5          10          15

35     Ala Thr Glu Ile Pro Ala Pro Gln Trp Arg Asp Leu Phe Leu Ala Pro
          20          25          30

      Asp Trp Gly Gly Thr Asp Glu Gln Val Ile Val Ala Glu Glu Ala Arg
          35          40          45

      Gly Pro Glu His Phe Thr Gly Ala Arg Arg Pro Arg Gly Gly Arg Arg
          50          55          60

45     Ser Ser Arg Arg Ala Ala
      65          70

```

50 <210> 33

<211> 172

<212> PRT

<213> *Streptomyces parvulus* Tü4055

55

60

65

ES 2 287 552 T3

<400> 33

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

<210> 34a

<211> 105

40 <212> PRT

<213> *Streptomyces parvulus* Tü4055

<400> 34a

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

ES 2 287 552 T3

<210> 35

<211> 638

<212> PRT

5 <213> *Streptomyces parvulus* Tü4055

<400> 35

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

ES 2 287 552 T3

	260	265	270
5	Ala Arg Leu Ile Val Gln 275	Ala Arg Arg Pro Gly 280	Ile Ser Leu Val 285
	His Arg Glu Leu Leu Asp 290	Phe Ala Gly Asp Glu 295	Phe Tyr Leu Ile Ser 300
10	Glu Pro Ala Leu Thr Gly 305	Arg Pro Phe Gly Glu 310	Val Leu Leu Ser Tyr 315
	Ser Thr Thr Ser Val Val 325	Gly Leu Met Arg Gly 330	Cys Thr Pro Leu Leu 335
15	Asn Pro Pro Pro Thr Thr 340	Pro Val Ala Pro Asp 345	Asp Leu Leu Val Val 350
	Ile Thr Gly Asp Asp Asp 355	Thr Ala Arg Leu Asp 360	Asp Cys Ala Glu Ser 365
20	Val Glu Lys Ala Ala Val 370	Ala Ser Arg Pro Pro 375	Thr Pro Ala Pro Ala 380
	Glu Arg Ile Leu Leu Leu 385	Gly Trp Asn Arg Arg 390	Ala Pro Leu Val Val 395
25	Asp Gln Leu His Arg Arg 405	Ala Arg Pro Gly Ser 410	Ala Val Asp Val Val 415
	Ala Glu Pro Gly Glu Ala 420	Thr Ile Arg Glu Ile 425	Ser Glu Ala Glu Ala 430
30	Asp Ser Gly Asn Gly Glu 435	Asn Gly Gly Asn Gly 440	Leu Ser Leu Ala Leu 445
	His His Gly Asp Ile Thr 450	Arg Pro Glu Thr Leu 455	Arg Arg Leu Asp Val 460
35	His Ser Tyr Asp Ser Val 465	Ile Val Leu Gly Arg 470	Asp Pro Ala Pro Gly 475
	Gln Pro Pro Asp Asp Pro 485	Asp Asn Arg Thr Leu 490	Val Thr Leu Leu Leu 495
40	Leu Arg Gln Leu Glu Glu 500	Ala Thr Gly Arg Glu 505	Leu Pro Val Val Thr 510
45	Glu Leu Ile Asp Asp Arg 515	Asn Arg Ala Leu Ala 520	Pro Ile Gly Pro Gly 525
	Ala Asp Val Ile Ile Ser 530	Gly Lys Leu Ile Gly 535	Leu Leu Met Ser Gln 540
50	Ile Ser Gln Asn Arg His 545	Leu Ala Ala Val Phe 550	Glu Glu Leu Phe Ser 555
	Ala Glu Gly Ala Gly Val 565	Arg Leu Arg Pro Ala 570	Thr Asp Tyr Leu Leu 575
55	Pro Gly Ser Thr Thr Ser 580	Phe Ala Thr Val Val 585	Ala Ala Ala Arg Arg 590
60	Arg Gly Glu Cys Ala Ile 595	Gly Tyr Arg Asp His 600	Ala Asp Ala Ser Thr 605
	Arg Pro His Tyr Gly Val 610	Arg Ile Asn Pro Pro 615	Lys Arg Glu Arg Arg 620
65	Arg Trp Thr Ala Glu Asp 625	Glu Val Val Val Ile 630	Gly Thr Asp 635

ES 2 287 552 T3

<210> 36

<211> 165

<212> PRT

5 <213> *Streptomyces parvulus* Tü4055

<400> 36

```

10      Met Pro Ser Thr Asp Val Val Glu Leu Ile Leu Arg Asp His Arg Arg
        1          5          10          15

      Met Glu Glu Leu Phe Arg Thr Leu Arg Asn Val Glu Ala Asp Arg Ala
        20          25          30

15      Ala Ala Leu Thr Glu Phe Ala Asp Leu Leu Ile Ala His Ala Ser Ala
        35          40          45

      Glu Glu Asp Glu Val Tyr Pro Ala Leu Arg Arg Tyr Lys Asn Val Glu
        50          55          60

20      Gly Glu Asp Val Asp His Ser Val His Glu His His Glu Ala Asn Glu
        65          70          75          80

      Ala Leu Leu Ala Leu Leu Glu Val Glu Asp Thr Ala Ser Asp Glu Trp
        85          90          95

25      Asp Asp Lys Leu Glu Glu Leu Val Thr Ala Val Asn His His Ala Asp
        100          105          110

      Glu Glu Glu Arg Thr Leu Leu Asn Asp Ala Arg Glu Asn Val Ala Asp
        115          120          125

30      Asp Arg Arg Arg Glu Leu Gly Gln Lys Phe Gln Glu Ala Arg Ser Arg
        130          135          140

35      Tyr Leu Glu Thr Gly Cys Gly Ser Val Glu Asn Val Arg Lys Leu Val
        145          150          155          160

      Ala Ala Ala Asp Asp
        165

```

<210> 37

<211> 787

45 <212> PRT

<213> *Streptomyces parvulus* Tü4055

<400> 37

```

50      Met Ala Arg Arg Leu Thr Glu Gly Arg Thr Arg Arg Glu Lys Gly Glu
        1          5          10          15

      His Met Gln Lys Pro His Gly Glu Glu Ser Glu Thr Ser Leu Ser Val
        20          25          30

55      Thr Pro Pro Lys Lys Trp Ala Ala Gly Val Pro Ala Val Val His Ala
        35          40          45

      Leu Glu Tyr Ser Leu Glu Gln Thr Ser Pro Arg Arg Thr Gly Val Asp
        50          55          60

```

65

ES 2 287 552 T3

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

ES 2 287 552 T3

	405	410	415
5	Gly Arg Ala Gly Thr Gly Ala Cys Pro Val Arg Gly His Ser Asn Val 420 425 430		
	Gln Gly Asp Arg Thr Met Gly Ile Trp Glu Gln Met Pro Asp Thr Phe 435 440 445		
10	Leu Asp Ala Leu Arg Asp Glu Phe Gly Phe Glu Pro Pro Arg Ala His 450 455 460		
	Gly Leu Asp Ser Val Asn Ser Ile Lys Ala Met Arg Glu Gly Arg Val 465 470 475 480		
15	Lys Val Phe Leu Ala Leu Ala Gly Asn Phe Val Arg Ala Ala Pro Asp 485 490 495		
	Ser Glu Val Thr Glu Glu Ala Met Arg Ser Cys Arg Leu Thr Ala His 500 505 510		
20	Ile Ser Thr Lys Leu Asn Arg Ser His Thr Val Cys Gly Asp Thr Ala 515 520 525		
	Leu Ile Leu Pro Thr Leu Gly Arg Thr Glu Arg Asp Val Gln Ala Asp 530 535 540		
25	Gly Glu Gln Phe Val Thr Val Glu Asn Ser Met Ser Glu Val His Thr 545 550 555 560		
	Ser Arg Gly Arg Leu Ala Pro Ala Ser Pro Met Leu Leu Ser Glu Ile 565 570 575		
30	Ala Ile Leu Cys Arg Leu Ala Arg Leu Thr Leu Asp Gly Arg Val Glu 580 585 590		
	Ile Pro Trp Glu Thr Phe Glu Gly Asp Tyr His Thr Ile Arg Asp Arg 595 600 605		
35	Ile Ala Arg Ile Val Pro Gly Phe His Asp Phe Asn Ala Arg Val Thr 610 615 620		
	Arg Pro Gly Gly Phe Gln Leu Pro Asn Pro Val Asn Glu Gly Val Phe 625 630 635 640		
40	Asn Thr Glu Val Gly Lys Ala Leu Phe Thr Arg Asn Glu Ser Val Val 645 650 655		
45	Pro Arg Ala Pro Glu Gly His Leu Leu Leu Gln Thr Leu Arg Ser His 660 665 670		
	Asp Gln Trp Asn Thr Val Pro Tyr Thr Asp Asn Asp Arg Tyr Arg Gly 675 680 685		
50	Ile His Gly Ser Arg His Val Val Leu Val Asn Pro Ala Asp Leu Ser 690 695 700		
	Glu Leu Gly Leu Ala Gln Gly Asp Arg Val Asp Leu Val Ser Val Trp 705 710 715 720		
55	Ala Asp Gly Thr Glu Arg Arg Ala Glu Asn Phe Gln Val Val Pro Tyr 725 730 735		
60	Pro Ala Ala Lys Gly Ser Ala Ala Ala Tyr Tyr Pro Glu Thr Asn Val 740 745 750		

65

ES 2 287 552 T3

Leu Val Pro Leu Asp Ser Val Ala Asp Ile Ser Asn Gln Pro Thr Ser
 755 760 765
 5 Lys Gly Ile Val Val Arg Leu Glu Pro Val Pro Asp Arg Thr Gln Pro
 770 775 780
 Ser Pro Ala
 785
 10 <210> 38
 <211> 206
 <212> PRT
 <213> *Streptomyces parvulus* Tü4055
 15 <400> 38
 Met Ala Glu Gln His Glu Gly Pro Arg Ala Val Pro Asp Thr Pro Gly
 1 5 10 15
 20 Ala Arg Thr Ser Gly Asp Arg Ser Thr Gly Arg Arg Pro Leu Arg Glu
 20 25 30
 Arg His Val Asp Gln Thr Val Glu Val Ala Val Pro Val Arg Thr Ala
 25 35 40 45
 Tyr Asn Gln Trp Thr Gln Phe Lys Ser Phe Pro Arg Phe Ser Ala Val
 50 55 60
 30 Val Arg Asp Val Glu Gln Val Arg Pro Thr Val Thr Ala Trp Thr Leu
 65 70 75 80
 Gly Tyr Gly Pro Leu Arg Arg Arg Phe Ala Val Glu Ile Leu Glu Gln
 85 90 95
 35 Asp Pro Asp Ala Tyr Leu Ala Trp Arg Gly Leu Glu Gln Arg Pro Trp
 100 105 110
 His Arg Gly Glu Val Glu Phe Arg Pro Thr Glu Ser Gly Gly Thr Ala
 40 115 120 125
 Ile Thr Val Arg Val Leu Leu Glu Pro Arg Gly Ala Ala Arg Ile Leu
 130 135 140
 45 Thr Arg Ser Ser Arg Ala Val Arg Leu Thr Thr Arg Leu Val His Gly
 145 150 155 160
 Glu Leu Thr Arg Phe Lys Arg Phe Met Glu Gly Leu Gly Gln Glu Gly
 165 170 175
 50 Gly Ala Trp Arg Gly Thr Ile Arg Asn Gly Arg Val Gln His Asp Arg
 180 185 190
 Pro Glu Pro Pro Arg Ser Arg Val Ala Arg Trp Pro Val Gly
 195 200 205
 55
 <210> 39
 <211> 251
 60 <212> PRT
 <213> *Streptomyces parvulus* Tü4055
 65

ES 2 287 552 T3

<400> 39

```

Met Leu Leu Leu Ile Ser Pro Asp Gly Val Glu Glu Ala Leu Asp Cys
 1           5           10           15

5 Ala Lys Ala Ala Glu His Leu Asp Ile Val Asp Val Lys Lys Pro Asp
 20           25           30

Glu Gly Ser Leu Gly Ala Asn Phe Pro Trp Val Ile Arg Glu Ile Arg
 35           40           45

10 Asp Ala Val Pro Ala Asp Lys Pro Val Ser Ala Thr Val Gly Asp Val
 50           55           60

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

15 Ser Gly Ala Thr Tyr Ile Lys Val Gly Leu Tyr Gly Cys Thr Thr Pro
 85           90           95

Glu Gln Gly Ile Ala Val Met Arg Ala Val Val Arg Ala Val Lys Asp
100           105           110

His Arg Pro Glu Ala Leu Val Val Ala Ser Gly Tyr Ala Asp Ala His
115           120           125

25 Arg Ile Gly Cys Val Asn Pro Leu Ala Leu Pro Asp Ile Ala Ala Arg
130           135           140

Ser Gly Ala Asp Ala Ala Met Leu Asp Thr Ala Val Lys Asp Gly Thr
145           150           155           160

30 Arg Leu Phe Asp His Val Pro Pro Asp Thr Cys Ala Glu Phe Val Arg
165           170           175

Arg Ala His Ala Ala Gly Leu Leu Ala Ala Leu Ala Gly Ser Val Arg
180           185           190

35 Gln Thr Asp Leu Gly Arg Leu Thr Arg Ile Gly Thr Asp Ile Val Gly
195           200           205

Val Arg Gly Ala Val Cys Glu Gly Gly Asp Arg Asn Ala Gly Arg Ile
210           215           220

40 Arg Pro His Leu Val Ala Ala Phe Arg Ser Glu Met Asp Arg His Ala
225           230           235           240

Arg Glu His Arg Ala Gly Val Thr Thr Ala Ser
245           250

```

45 <210> 40

<211> 467

<212> PRT

50 <213> *Streptomyces parvulus* Tü4055

<400> 40

```

Met Pro Thr Pro Ala Pro Asp His Ala Pro Ala Gln Arg Ala Ala Pro
 1           5           10           15

55 Leu Ala Val Val Asp Pro Ala Thr Gly Thr Val Phe Asp Glu Ala Pro
 20           25           30

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

60 Ala Trp His Gly Trp Arg Ala Asp Pro Asp Ala Arg Thr Thr Ala Leu
 50           55           60

Arg Ser Ala Ala Asp Ala Val Glu Ala Ala Gly Asp Asp Leu Ala Arg
65

```

ES 2 287 552 T3

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

ES 2 287 552 T3

Glu Cys Gly Thr Ala Trp Val Asn His His Ala Glu Leu Ser Leu Ala
 420 425 430
 5 Gln Pro Phe Ala Gly Asp Lys Asp Ser Gly Val Gly Val Ala Gly Gly
 435 440 445
 Pro Trp Gly Leu Tyr Gly Asn Leu Arg Pro Phe Val Val His Arg Pro
 450 455 460
 10 Arg Gly Glu
 465
 <210> 41
 <211> 368
 15 <212> PRT
 <213> *Streptomyces parvulus* Tü4055
 20 <400> 41
 Met Ser Phe Arg Ala Ala Val Leu Arg Gly Tyr Glu Asp Pro Phe Thr
 1 5 10 15
 25 Val Glu Glu Val Thr Leu Gly Thr Glu Pro Gly Ala Gly Glu Ile Leu
 20 25 30
 Val Glu Ile Ala Gly Cys Gly Met Cys Arg Thr Asp Leu Ala Val Arg
 35 40 45
 30 Arg Ser Ala Gly Arg Ser Pro Leu Pro Ala Val Leu Gly His Glu Gly
 50 55 60
 Ser Gly Val Val Val Arg Thr Gly Gly Gly Pro Asp Thr Ala Ile Gly
 65 70 75 80
 35 Val Gly Asp His Val Val Leu Ser Phe Asp Ser Cys Gly His Cys Arg
 85 90 95
 40 Asn Cys Arg Ala Ala Ala Pro Ala Tyr Cys Asp Ser Phe Ala Ser Leu
 100 105 110
 Asn Leu Phe Gly Gly Arg Ala Glu Asp Pro Pro Arg Leu Thr Asp Gly
 115 120 125
 45 Ser Gly Ala Ala Leu Ala Pro Arg Trp Phe Gly Gln Ser Ala Phe Ala
 130 135 140
 Glu Tyr Ala Leu Val Ser Ala Arg Asn Ala Val Arg Val Asp Pro Ala
 145 150 155 160
 50 Leu Pro Val Glu Leu Leu Gly Pro Leu Gly Cys Gly Phe Leu Thr Gly
 165 170 175
 Ala Gly Ala Val Leu Asn Thr Phe Ala Ala Gly Pro Gly Asp Thr Leu
 180 185 190
 55 Val Val Leu Gly Ala Gly Ala Val Gly Leu Ala Ala Val Met Ala Ala
 195 200 205
 60 Thr Ala Ala Gly Ala Pro Ser Val Ala Val Asp Arg Asn Pro Arg Arg
 210 215 220
 Leu Glu Leu Ala Glu Arg Phe Gly Ala Val Pro Leu Pro Ala Ala Thr
 225 230 235 240
 65

ES 2 287 552 T3

	Ala Gly Leu Ala Glu Arg Ile Arg Arg Leu Thr Asp Gly Gly Ala Arg	245	250	255
5	Tyr Ala Leu Asp Thr Thr Ala Ser Val Pro Leu Ile Asn Glu Ala Leu	260	265	270
	Arg Ala Leu Arg Pro Thr Gly Ala Leu Gly Leu Val Ala Arg Leu His	275	280	285
10	Thr Ala Leu Pro Leu Glu Pro Gly Thr Leu Asp Arg Gly Arg Ser Ile	290	295	300
	Arg His Val Cys Glu Gly Asp Ala Val Pro Gly Leu Leu Ile Pro Gln	305	310	315
15	Leu Thr Arg Leu Trp Gln Ala Gly Arg Phe Pro Phe Asp Gln Leu Val	325	330	335
	Arg Thr Tyr Pro Leu Ala Asp Ile Asn Glu Ala Glu Arg Asp Cys Asp	340	345	350
20	Ala Gly Leu Val Val Lys Pro Val Leu Leu Pro Pro Ala Arg Ser Arg	355	360	365

25 <210> 42
 <211> 301
 <212> PRT
 30 <213> *Streptomyces parvulus* Tü4055

35

40

45

50

55

60

65

ES 2 287 552 T3

<400> 42

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

<210> 43

<211> 179

65 <212> PRT

<213> *Streptomyces parvulus* Tü4055

ES 2 287 552 T3

<400> 43

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

<210> 44

40 <211> 33

<212> DNA

<213> Secuencia artificial

45 <220>

<223> Descripción de secuencia artificial: oligo CM410

<400> 44

50
 aaaatgcatt cggcctgaac ggccccgctg tca

33

<210> 45

55 <211> 33

<212> DNA

<213> Secuencia artificial

60 <220>

<223> Descripción de secuencia artificial: oligo CM411

<400> 45

65
 aaatggccag cgaacaccaa caccacacca cca

33

ES 2 287 552 T3

	<210> 46	
	<211> 32	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM412	
10	<400> 46	
	aaagtcctag gcggcggccg gcgggtcgac ct	32
15	<210> 47	
	<211> 35	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM413	
25	<400> 47	
	tttagatctc gcgacgtcg acgcgccgaa cgtca	35
30	<210> 48	
	<211> 34	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM414	
40	<400> 48	
	aaactgcaga gtcgaacatc ggtcacacgc aggc	34
45	<210> 49	
	<211> 35	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM415	
55	<400> 49	
	aaaatgcatg atccacatcg atacgacgcg cccga	35
60	<210> 50	
	<211> 36	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>	
	<223> Descripción de secuencia artificial: oligo CM416	
5	<400> 50	
	taaatgcatt ccattcggtg caggtggagt tgatcc	36
10	<210> 51	
	<211> 36	
	<212> DNA	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de secuencia artificial: oligo CM417	
20	<400> 51	
	ataggatccc ctccgggtgc tccagaccgg ccaccc	36
25	<210> 52	
	<211> 35	
	<212> DNA	
	<213> Secuencia artificial	
30	<220>	
	<223> Descripción de secuencia artificial: oligo CM368	
35	<400> 52	
	tttctgcag gccatcccca cgatcgcat cggct	35
40	<210> 53	
	<211> 35	
	<212> DNA	
	<213> Secuencia artificial	
45	<220>	
	<223> Descripción de secuencia artificial: oligo CM369	
50	<400> 53	
	tttcatatga caggcagtgc tgtttcggcc ccatt	35
55	<210> 54	
	<211> 36	
	<212> DNA	
	<213> Secuencia artificial	
60	<220>	
	<223> Descripción de secuencia artificial: oligo CM370	
65	<400> 54	
	tttcatatgg cggatgccgt acgtgccgcc ggcgct	36

ES 2 287 552 T3

	<210> 55	
	<211> 32	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM371	
10	<400> 55	
	tttcatatgc cccaggcgat cgtccgcacc ac	32
15	<210> 56	
	<211> 41	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM372	
25	<400> 56	
	tttcatatgg tctcgggccc ccacacaaga gccctccggg c	41
30	<210> 57	
	<211> 20	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B1819A	
40	<400> 57	
	gtcatgcatg cggcgggctc	20
45	<210> 58	
	<211> 20	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B1819B	
55	<400> 58	
	ggtctagaac ggccgaactt	20
60	<210> 59	
	<211> 20	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>		
	<223> Descripción de secuencia artificial: oligo B1819C		
5	<400> 59		
	gttctagaac ctcggtcggc		20
10	<210> 60		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
15	<220>		
	<223> Descripción de secuencia artificial: oligo B1819D		
20	<400> 60		
	ctggatccca cgctgctgcg		20
25	<210> 61		
	<211> 19		
	<212> DNA		
	<213> Secuencia artificial		
30	<220>		
	<223> Descripción de secuencia artificial: oligo ELDA		
35	<400> 61		
	ggagacttac gggggatgc		19
40	<210> 62		
	<211> 19		
	<212> DNA		
	<213> Secuencia artificial		
45	<220>		
	<223> Descripción de secuencia artificial: oligo BLDB		
50	<400> 62		
	ctccagcagc gaccagaac		19
55	<210> 63		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
60	<220>		
	<223> Descripción de secuencia artificial: oligo B19A		
65	<400> 63		
	cccatgcatc accgacatac		20

ES 2 287 552 T3

	<210> 64	
	<211> 20	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B19B	
10	<400> 64	
	gcgatatccc gaagaacgcg	20
15	<210> 65	
	<211> 20	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B1920A	
25	<400> 65	
	gccaagcttc ctcgacgcgc	20
30	<210> 66	
	<211> 20	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B1920B	
40	<400> 66	
	cactagtgcc tcaccagtt	20
45	<210> 67	
	<211> 20	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B1920C	
55	<400> 67	
	cactagtgac ggccgaagcg	20
60	<210> 68	
	<211> 20	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>	
	<223> Descripción de secuencia artificial: oligo B1920D	
5	<400> 68	
	tcggatccgt cagaccgttc	20
10	<210> 69	
	<211> 36	
	<212> DNA	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de secuencia artificial: oligo CM384	
20	<400> 69	
	aacctgcagg taccccgggtg ggggtgcggtc gcccca	36
25	<210> 70	
	<211> 24	
	<212> DNA	
	<213> Secuencia artificial	
30	<220>	
	<223> Descripción de secuencia artificial: oligo CM385	
35	<400> 70	
	cgccgcacgc gtcgaagcca acga	24
40	<210> 71	
	<211> 24	
	<212> DNA	
	<213> Secuencia artificial	
45	<220>	
	<223> Descripción de secuencia artificial: oligo CM386	
50	<400> 71	
	tgtgggctgg tcgttggtt cgac	24
55	<210> 72	
	<211> 34	
	<212> DNA	
	<213> Secuencia artificial	
60	<220>	
	<223> Descripción de secuencia artificial: oligo CM387	
65	<400> 72	
	ggtgcctgca gcgtgagttc ctcacggat ccga	34

ES 2 287 552 T3

	<210> 73	
	<211> 26	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM388	
10	<400> 73	
	gaggaactca ccctgcaggc accgct	26
15	<210> 74	
	<211> 26	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM395	
25	<400> 74	
	cgaacgtcca gccctcgggc atgcgt	26
30	<210> 75	
	<211> 28	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM396	
40	<400> 75	
	tggcacgcat gcccgagggc tggacgtt	28
45	<210> 76	
	<211> 35	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo CM397	
55	<400> 76	
	tttctgcag gccatgccga cgatecgcac aggct	35
60	<210> 77	
	<211> 36	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>	
	<223> Descripción de secuencia artificial: oligo CM398	
5	<400> 77	
	aaacatatgg tcctggcgct gcgcaacggg gaactg	36
10	<210> 78	
	<211> 35	
	<212> DNA	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de secuencia artificial: oligo CM399	
20	<400> 78	
	tttctgcag gcgatgccga cgatggcgat gggct	35
25	<210> 79	
	<211> 43	
	<212> DNA	
	<213> Secuencia artificial	
30	<220>	
	<223> Descripción de secuencia artificial: oligo CM400	
35	<400> 79	
	aaacctgcag gttccccggc gacgtggact cgccggagtc gtt	43
40	<210> 80	
	<211> 41	
	<212> DNA	
	<213> Secuencia artificial	
45	<220>	
	<223> Descripción de secuencia artificial: oligo CM401	
50	<400> 80	
	ttttctagag cgacgtcgca ggcggcgatg gtcacgcccg t	41
55	<210> 81	
	<211> 20	
	<212> DNA	
	<213> Secuencia artificial	
60	<220>	
	<223> Descripción de secuencia artificial: oligo B25A	
65	<400> 81	
	ttctgcagcc geggcctteg	20

ES 2 287 552 T3

	<210> 82	
	<211> 20	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B25B	
10	<400> 82	
	agaattcgcc ggcgccgtg	20
15	<210> 83	
	<211> 20	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B7T1	
25	<400> 83	
	ggctgcagac gcggctgaag	20
30	<210> 84	
	<211> 20	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B7T2	
40	<400> 84	
	ccggatccca gagccacgtc	20
45	<210> 85	
	<211> 20	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo BP4501	
55	<400> 85	
	cgtatgcatg gcgccatgga	20
60	<210> 86	
	<211> 20	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>		
	<223> Descripción de secuencia artificial: oligo BP4502		
5	<400> 86		
	agccaattgg tgcactccag		20
10	<210> 87		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
15	<220>		
	<223> Descripción de secuencia artificial: oligo BNHT1		
20	<400> 87		
	gtcatgcatc agcgcacccg		20
25	<210> 88		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
30	<220>		
	<223> Descripción de secuencia artificial: oligo BNHT2		
35	<400> 88		
	gtgcaattgc cctggtagtc		20
40	<210> 89		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
45	<220>		
	<223> Descripción de secuencia artificial: oligo BTRNAS1		
50	<400> 89		
	tgtctagact cgcgcaaca		20
55	<210> 90		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
60	<220>		
	<223> Descripción de secuencia artificial: oligo BTRNAS2		
65	<400> 90		
	tgaattccga agggggtggt		20

ES 2 287 552 T3

	<210> 91	
	<211> 20	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B5B	
10	<400> 91	
	aactagtccg cagtggaccg	20
15	<210> 92	
	<211> 20	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B5A	
25	<400> 92	
	tcgatatcct caccgcccgt	20
30	<210> 93	
	<211> 20	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B6B	
40	<400> 93	
	aactagtgtg gcagacggtc	20
45	<210> 94	
	<211> 20	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B5A	
55	<400> 94	
	tcgatatcct caccgcccgt	20
60	<210> 95	
	<211> 20	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>		
	<223> Descripción de secuencia artificial: oligo B6T1		
5	<400> 95		
	cggatgcatc accggcacgg		20
10	<210> 96		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
15	<220>		
	<223> Descripción de secuencia artificial: oligo B6T2		
20	<400> 96		
	tgggatccgc ggggcggtac		20
25	<210> 97		
	<211>20		
	<212> DNA		
	<213> Secuencia artificial		
30	<220>		
	<223> Descripción de secuencia artificial: oligo BBB		
35	<400> 97		
	aactagtgcg atcccgggga		20
40	<210> 98		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
45	<220>		
	<223> Descripción de secuencia artificial: oligo BBA		
50	<400> 98		
	cgtcgatatc ctccaggggc		20
55	<210> 99		
	<211> 20		
	<212> DNA		
	<213> Secuencia artificial		
60	<220>		
	<223> Descripción de secuencia artificial: oligo BBT1		
65	<400> 99		
	tactgcagca caccggtgc		20

ES 2 287 552 T3

	<210> 100	
	<211> 20	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo BBT2	
10	<400> 100	
	tgggatccgc tgtgtcatat	20
15	<210> 101	
	<211> 20	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo BCB	
25	<400> 101	
	cactagtcct cgccgggcac	20
30	<210> 102	
	<211> 20	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo BCA	
40	<400> 102	
	gaggatcccg gtcagcggca	20
45	<210> 103	
	<211> 20	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo BCT1	
55	<400> 103	
	gcctgcagcg acctcgccgg	20
60	<210> 104	
	<211> 20	
	<212> DNA	
65	<213> Secuencia artificial	

ES 2 287 552 T3

	<220>	
	<223> Descripción de secuencia artificial: oligo BCT2	
5	<400> 104	
	cgggatcccg tggcgtggtc	20
10	<210> 105	
	<211> 20	
	<212> DNA	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de secuencia artificial: oligo B23A	
20	<400> 105	
	atctgcagcg gcatcggtgt	20
25	<210> 106	
	<211> 20	
	<212> DNA	
	<213> Secuencia artificial	
30	<220>	
	<223> Descripción de secuencia artificial: oligo B23B	
35	<400> 106	
	agaattctcc actgcggtcg	20
40	<210> 107	
	<211> 20	
	<212> DNA	
	<213> Secuencia artificial	
45	<220>	
	<223> Descripción de secuencia artificial: oligo B9A	
50	<400> 107	
	acctgcaggc cgggctcatc	20
55	<210> 108	
	<211> 20	
	<212> DNA	
	<213> Secuencia artificial	
60	<220>	
	<223> Descripción de secuencia artificial: oligo B9B	
65	<400> 108	
	agaattcggg cgagccgccg	20

ES 2 287 552 T3

	<210> 109	
	<211> 20	
	<212> DNA	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B231	
10	<400> 109	
	atcaagcttc gtgtccatgg	20
15	<210> 110	
	<211> 20	
	<212> DNA	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B232	
25	<400> 110	
	gtcatgcatc aggcgttcgg	20
30	<210> 111	
	<211> 20	
	<212> DNA	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B251	
40	<400> 111	
	cttctagatg aaccctcca	20
45	<210> 112	
	<211> 20	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: oligo B252	
55	<400> 112	
	gggcaattgc gcggcagctt	20
60	<210> 113	
	<211> 90	
	<212> PRT	
65	<213> <i>Streptomyces parvulus</i> Tü4055	

ES 2 287 552 T3

<400> 113

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