

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 556 211**

51 Int. Cl.:

C12N 15/09	(2006.01)	C07D 311/00	(2006.01)
C12N 15/52	(2006.01)	A61P 31/12	(2006.01)
C07K 14/36	(2006.01)	A61P 31/18	(2006.01)
C12P 17/08	(2006.01)	A61P 37/00	(2006.01)
C12P 21/02	(2006.01)	A61P 37/06	(2006.01)
A61K 31/35	(2006.01)	A61P 1/16	(2006.01)
A61K 31/5025	(2006.01)	C12R 1/465	(2006.01)
C07D 498/08	(2006.01)		
C07D 491/10	(2006.01)		
C07D 519/00	(2006.01)		

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **24.09.2009 E 09815612 (8)**

97 Fecha y número de publicación de la concesión europea: **16.09.2015 EP 2342335**

54 Título: **Agrupación de genes novedosa**

30 Prioridad:

24.09.2008 CN 200810200388

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

14.01.2016

73 Titular/es:

**SHANGHAI INSTITUTE OF ORGANIC
CHEMISTRY, CHINESE ACADEMY OF SCIENCES
(100.0%)
345 Lingling Road
Shanghai 200032, CN**

72 Inventor/es:

**LIU, WEN;
JIANG, NAN y
QU, XUDONG**

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 2 556 211 T3

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

DESCRIPCIÓN

Agrupación de genes novedosa

5 **Introducción**

La presente invención se refiere a secuencias de ADN aisladas novedosas que comprenden una agrupación de genes que codifica para sangliferina sintasa, genes de procesamiento y reguladores implicados en la biosíntesis de un compuesto de policétido / péptido no ribosómico mixto, o mutantes que tienen capacidad de biosíntesis alterada, polipéptidos o mutantes de los mismos codificados por ADN o los mutantes, vectores que contienen el ADN o los mutantes del mismo, células huésped transformadas con el ADN, los mutantes del mismo, o el vector, y a un método para producir compuestos de sangliferina. La presente invención también proporciona compuestos con actividad de inhibición de ciclofilina útiles como inmunosupresores, antivíricos o agentes de protección cardíaca.

15 **Campo de la invención**

La invención se refiere al campo de recursos genéticos de microorganismos e ingeniería genética, y en particular se refiere a la clonación, al análisis de secuencia, a la verificación funcional *in vivo*, y al uso de la agrupación de genes para la biosíntesis de la sangliferina A inmunosupresora y análogos relacionados.

20 **Antecedentes de la invención**

La sangliferina A (SFA) inmunosupresora es un producto natural de policétido-péptido mixto procedente de *Streptomyces sp. A92-308110*, también conocido como *Streptomyces flaveolus* o *Streptomyces sp.* DSM 9954, todos estos nombres se usan de manera intercambiable en este documento y en documentos relacionados (Sanglier *et al.*, 1999; Fehr *et al.*, 1999; documento WO 97/02285). Hasta ahora se ha publicado el aislamiento de más de veinte análogos estructurales de sangliferina, y SFA tiene una de las mayores actividades inmunosupresoras de esos análogos (Kallen *et al.*, 2005; Sanglier *et al.*, 1999). En SFA, la estructura principal de macrólido de 22 miembros consiste estructuralmente en una cadena de carbono de policétido y una cadena de tripéptido. La cadena peptídica comprende un aminoácido natural: valina, y dos aminoácidos no naturales: (S)-m-tirosina y ácido(S)-piperázico, unidos mediante un enlace amida, y es el átomo de nitrógeno β en la posición 1 del ácido piperázico el que está implicado en la formación del enlace amida, lo cual es contrario a todos los demás productos naturales que contienen ácido piperázico aislados hasta ahora. Además, una unidad espirocíclica está unida al macrólido mediante una cadena larga de policétido, formando una estructura de cesta. El resto espirocíclico contiene nueve centros quirales (SFA tiene diecisiete en total) con un centro de carbono cuaternario en la parte central, lo cual es único en los productos naturales actualmente descritos. Se ha aislado directamente una serie de análogos a partir de caldos de fermentación de *Streptomyces sp. A92-308110* (*S. flaveolus*), incluyendo sangliferina B, C, D, E, F, G, H, I, J, K y L (Fehr *et al.*, 1999, Sanglier *et al.*, 1999, documento WO98/07743). En particular, se ha mostrado que la sangliferina B (SFB) presenta una actividad inmunosupresora mayor que SFA en ensayos de MLR (Sanglier *et al.*, 1999). Aunque se logró la síntesis total de SFA y sus análogos macrocíclicos con grandes esfuerzos (Sedrani *et al.*, 2003; Nicolau *et al.*, 1999; Paquette *et al.*, 2002; Metternich *et al.*, 1999), no se ha llevado a cabo ningún estudio específico *in vivo* ni *in vitro* sobre su ruta de biosíntesis.

SFA tiene una fuerte actividad inmunosupresora (Powell y Zheng, 2006), inhibe la infección por VIH y VHC (Zander *et al.*, 2003; Sokolskaja *et al.*, 2004; Watashi *et al.*, 2005) y previene la muerte de células cardíacas grave provocada por la apertura patológica de poro de transición de permeabilidad de membrana mitocondrial (MPTP; Clarke *et al.*, 2002). En comparación con los inmunosupresores actualmente en uso clínico, tales como ciclosporina A (CsA), FK506 y rapamicina, SFA tiene un mecanismo funcional similar mientras que tiene un sitio efector diana diferente (Härtel *et al.* 2006, Zhang & Liu., 2001; Zenke *et al.*, 2001). CsA se une a ciclofilina A (CypA) (Handschumacher *et al.*, 1984), mientras que FK506 y rapamicina se unen a FKBP, para formar complejos (Schreiber, 1991); los complejos CsA-CypA y FK506-FKBP interaccionan con la misma proteína diana, calcineurina, inhibiendo así la actividad serina/treonina fosfatasa de calcineurina, y bloqueando la producción de citocinas, especialmente la transcripción de interleucina 2 (IL-2), lo que finalmente conduce a la detención de células T en la fase G_0 - G_1 (Liu *et al.*, 1991). Los complejos Rap-FKBP interaccionan con la proteína cinasa FRAP (también conocida como RAFT o mTOR) (Brown *et al.*, 1994), e impiden la fosforilación del receptor de IL-2 en células T, conduciendo a la detención del crecimiento de células T en la fase G_1 -S. Mientras que se ha mostrado que SFA se une a ciclofilinas tales como ciclofilina A y B, e inhibe sus actividades isomerasa (Zenke *et al.* 2001), actualmente sigue sin conocerse la proteína efectora para el complejo SFA-CypA.

Dado que aún falta por descubrir la proteína efectora del complejo SFA-CypA, se sugirió que la actividad inmunosupresora de SFA no está mediada directamente por medio del complejo SFA-CypA. En los últimos 3 años, estudios de muchos grupos científicos han mostrado que SFA pueden impedir de manera competitiva que NF- κ B se una al sitio de transcripción en el sentido de 5' del gen P53, para activar P53 e inhibir adicionalmente la fosforilación de ciclina E-cdk2 posterior de la ruta de señalización, inhibiendo así la alta fosforilación de Rb en respuesta a IL-2, y haciendo que las células sean insensibles a IL-2, lo cual las fuerza a permanecer en la fase G_1 -S (Zhang & Liu, 2001). En segundo lugar, mediante un mecanismo desconocido, SFA inhibe la producción de IL-12⁷⁰ mientras que

no afecta al crecimiento de células dendríticas humanas (Steinschulte *et al.*, 2003). IL-12p⁷⁰ desempeña un papel clave en la regulación de la proliferación de células Th1 y NK, y es el puente que une la inmunidad innata a la inmunidad adaptativa. Además, los fármacos inmunosupresores que están comercialmente disponibles pueden conducir a graves efectos secundarios de toxicidad renal y en el sistema nervioso central (Paquette *et al.*, 2002), por tanto sus usos en algunas enfermedades de disfunción inmunitaria se ven dificultados (por ejemplo, la calcineurina es la causa subyacente de efectos tanto inmunosupresores como tóxicos de CsA y FK506). Con el objetivo de desarrollar un inmunosupresor o modificador inmunitario alternativo, otros grupos han llevado a cabo cierto desarrollo de SFA como nueva generación de potente inmunosupresor con menor toxicidad (documento WO 97/02285).

El estudio de las relaciones estructura-actividad entre fragmentos macrocíclicos de SFA y CypA mediante difracción de rayos X mostró que la estructura de tripéptido está incrustada en el surco de CypA y es importante para la unión, mientras que el grupo hidroxilo y el grupo carbonilo de cadena lateral, respectivamente, en las posiciones 17 y 14 no son críticos para la unión; la eliminación del trans-dieno de la región saturada C18-C22 reduce la constante de unión en 7 veces, lo que sugiere que el trans-dieno estabiliza la conformación (Sedrani *et al.*, 2003). Un estudio de modelado por ordenador muestra que la unidad espirocíclica de SFA también puede contribuir a la estabilidad de la unión SFA-CypA (Pemberton *et al.*, 2003). La estructura cristalina del complejo SFA-CypA completo muestra que las regiones de unión en SFA-CypA son sustancialmente las mismas que aquéllas en CsA-CypA, y tanto SFA como CsA interaccionan principalmente con los residuos W121, R55, H126, N102 y Q63; la cadena de C24-C32 entre el macrociclo y el resto espirocíclico realizan contactos de van der Waals con los residuos I57, T119 y W121 de CypA; además, la presencia de la cadena de policétido larga de SFA impone una reorientación de cadena lateral en W121, en comparación con la estructura cristalina de CypA; dentro del espirociclo, sólo el C45 de grupo metilo realiza contactos de vdW con átomos de cadena lateral de I57 y F60 de CypA (Kallen *et al.*, 2005).

El complejo SFA-CypA puede existir en una forma dimérica estable, tal como se muestra mediante cromatografía de filtración en gel. Basándose en el análisis cristalino, con la excepción de los restos espirocíclico y α -cetobutirato, todas las demás partes de SFA están profundamente enterradas en el dímero; la región de E,E-dieno C18-C22 no está implicada en contactos directos con la CypA sino que en vez de eso forma contactos de vdW con la metatirosina de SFA vecina dentro del dímero, lo cual favorece la asociación dimérica en el complejo; las dos moléculas de SFA realizan contactos de vdW entre sí en la región C18-C22; y un enlace de hidrógeno directo une W121 de una molécula de CypA y R148 de otra molécula de CypA en el complejo dimérico.

Puede encontrarse técnica remotamente relacionada con agrupaciones de genes en la Univ. de Minnesota (documento WO 00/000620) que se refiere a una agrupación de genes de biosíntesis para metamicina, picromicina y desosamina.

Usando el estreptomiceto que se sabe que produce SFA, *Streptomyces sp. A92-308110 (S. flaveolus)*, los inventores de la presente invención clonaron la agrupación de genes de biosíntesis del mismo, y estudiaron adicionalmente la biosíntesis de SFA mediante métodos que combinaban microbiología, biología molecular, bioquímica y química orgánica. Mediante el estudio de la biosíntesis, se reveló el mecanismo enzimático que genera estructuras químicas distintivas tales como ácido piperázico. Basándose en esto, se realizaron modificaciones genéticas en la ruta de biosíntesis de SFA, y se produjeron compuestos novedosos.

La presente invención es particularmente útil ya que debe permitir la aplicación comercial de tecnología de ADN recombinante e ingeniería de biosíntesis para aumentar el rendimiento de sangliferinas y la generación de análogos de sangliferina novedosos.

Sumario de la invención

La presente invención proporciona ventajosamente secuencias de ADN y proteínas novedosas implicadas en la producción de productos génicos de biosíntesis, en particular la biosíntesis de sangliferina. Se detallan realizaciones específicas de los genes y las proteínas en la lista de secuencias adjunta y la siguiente descripción. SEQ ID No. 1 proporciona el ácido nucleico responsable de la biosíntesis de sangliferina A.

Por tanto, la invención se refiere a la clonación, al análisis de secuencia, a la verificación funcional, al análisis bioquímico *in vitro* y al uso de la agrupación de genes de biosíntesis de SFA, un producto natural de policétido-péptido no ribosómico inmunosupresor producido por *Streptomyces sp. A92-308110 (S. flaveolus)* (disponible de DSMZ, Braunschweig, Alemania como *Streptomyces sp. DSM 9954*). Además, se realizaron alteraciones dirigidas en los genes que codifican para la ruta de biosíntesis para la generación de sangliferina, conduciendo a cepas microbianas que produjeron análogos de sangliferina novedosos.

La invención permite la manipulación directa de sangliferina A y estructuras químicas relacionadas por medio de modificación por ingeniería de biosíntesis de los genes y las proteínas implicadas en la biosíntesis de la sangliferina A. Estas modificaciones químicas pueden ser o bien imposibles o bien inviables de realizar mediante metodologías químicas debido a la complejidad de las estructuras.

La agrupación de genes aislada y caracterizada de esta manera permite la optimización dirigida de la producción de sangliferina y análogos de sangliferina, por ejemplo, mediante duplicación de la agrupación de genes, o partes de la agrupación, mediante sobreexpresión de genes (en particular genes reguladores positivos) usando vectores de plásmidos y promotores no naturales, o mediante inactivación de reguladores negativos.

Además, la agrupación secuenciada y caracterizada permite la preparación por biosíntesis dirigida de análogos de sangliferina, varios ejemplos de los cuales se incluyen en este documento. Los ejemplos incluyen los siguientes:

- Inactivación de genes que codifican para proteínas implicadas en etapas de biosíntesis individuales, por ejemplo mediante interrupción génica (véanse por ejemplo los documentos WO 2004/007709; WO 2004/058976).

- Sustitución de genes que codifican para proteínas implicadas en etapas de biosíntesis individuales, mediante sustitución génica o mediante interrupción seguida por expresión separada de genes de otras rutas de biosíntesis (véanse por ejemplo Gaisser *et al.*, 2001; documentos WO 01/79520; WO 2005/054266; WO2005/054265).

- Intercambio de módulos o dominios individuales dentro de la policétido sintasa ("PKS") o péptido no ribosómico sintasa ("NRPS") por módulos o dominios de otras agrupaciones de PKS o NRPS, para permitir la generación de análogos de sangliferina novedosos (por ejemplo tal como se describe en Oliynyk *et al.*, 1996; documentos WO 98/01546; WO 00/01827; Staunton y Wilkinson, 2001; Sheehan *et al.*, 2006).

- Uso de la secuencia génica para identificar agrupaciones de biosíntesis relacionadas de otros organismos, por ejemplo mediante el uso como sonda de ADN (véanse por ejemplo Shen *et al.*, 2002; Liu *et al.*, 2002; Li *et al.*, 2004; Huang *et al.*, 2005; Jia *et al.*, 2006; Fang *et al.*, 2008).

Por tanto según un aspecto de la invención se proporciona:

i) Una molécula nucleica aislada que comprende:

(a) el ácido nucleico de agrupación de genes de biosíntesis de sangliferina A, que comprende SEQ ID No. 1;

(b) un ácido nucleico que tiene al menos el 80% de identidad de secuencia con el ácido nucleico de (a) y que codifica para polipéptidos que tienen las mismas actividades enzimáticas y reguladoras para producir un policétido o una unidad iniciadora de policétido como los codificados por el ácido nucleico de (a); o

(c) un ácido nucleico que codifica para uno o más polipéptidos que tienen al menos el 80% de identidad de secuencia de aminoácidos con uno o más polipéptidos codificados por el ácido nucleico de (a) y que codifica para uno o más polipéptidos que tienen una o más de las actividades enzimáticas o reguladoras necesarias para producir un policétido o un precursor o unidad iniciadora de policétido, en la que los polipéptidos tienen una cualquiera de SEQ ID No. 2-25.

ii) Un ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que uno o más genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaA* que son los residuos 17024-17854 de SEQ ID No. 1, *sfaB* que son los residuos 17851-19191 de SEQ ID No. 1, *sfaJ* que son los residuos 96225-97391 de SEQ ID No. 1, *sfaM* que son los residuos 103210-103929 de SEQ ID No. 1, *sfaN* que son los residuos 104001-105023 de SEQ ID No. 1, *sfaP* que son los residuos 105366-107216 de SEQ ID No. 1, *sfaQ* que son los residuos 107366-108145 de SEQ ID No. 1, *sfaL* que son los residuos 101936-103213 de SEQ ID No. 1 y *sfaO* que son los residuos 105091-105345 de SEQ ID No. 1, se han delecionado o modificado de modo que se disminuye su actividad o modificado o sustituido de modo que se altera su especificidad de sustrato.

iii) Un ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que uno o más genes de PKS seleccionados de *sfaE* que son los residuos 30707-37360 de SEQ ID No. 1, *sfaF* que son los residuos 37394-50014 de SEQ ID No. 1, *sfaG* que son los residuos 50017-60903 de SEQ ID No. 1, *sfaH* que son los residuos 60918-85823 de SEQ ID No. 1 y *sfaI* que son los residuos 85823-96040 de SEQ ID No. 1, se han modificado mediante lo cual se han delecionado uno o más dominios o módulos, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.

iv) Un ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que el gen de NRPS *sfaD* que son los residuos 19885-30714 de SEQ ID No. 1 se ha modificado mediante lo cual se han delecionado uno o más dominios o módulos, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.

v) Un ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que el gen regulador *sfaC* que son los residuos 19193-19888 de SEQ ID No. 1 se ha modificado de modo que se aumenta o disminuye su actividad, o se deleciona.

- 5 vi) Un ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que el gen de PKS iterativa *sfaK* que son los residuos 97396-101943 de SEQ ID No. 1 se ha modificado mediante lo cual se han delecionado uno o más dominios, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.
- 10 vii) Un polipéptido o pluralidad de polipéptidos codificados por un ácido nucleico según la invención reivindicada que es un ADN.
- viii) Una policétido sintasa codificada por un ácido nucleico según la invención reivindicada que es un ADN y que codifica para una o más proteínas de biosíntesis de policétido.
- 15 ix) Un vector que comprende un ácido nucleico según la invención reivindicada que es un ADN junto con uno o más promotores u otros elementos reguladores.
- x) Un método para producir un policétido que comprende cultivar una célula huésped transformada según la invención reivindicada.
- 20 xi) Una cepa modificada por ingeniería que comprende una cepa productora de sangliferina A en la que *sfaA* que son los residuos 17024-17854 de SEQ ID No. 1 y opcionalmente uno o más de otros genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaB* que son los residuos 17851-19191 de SEQ ID No. 1, *sfaJ* que son los residuos 96225-97391 de SEQ ID No. 1, *sfaM* que son los residuos 103210-103929 de SEQ ID No. 1, *sfaN* que son los residuos 104001-105023 de SEQ ID No. 1, *sfaP* que son los residuos 105366-107216 de SEQ ID No. 1, *sfaQ* que son los residuos 107366-108145 de SEQ ID No. 1, *sfaL* que son los residuos 101936-103213 de SEQ ID No. 1 y *sfaO* que son los residuos 105091-105345 de SEQ ID No. 1, se han delecionado o modificado de modo que se disminuye su actividad o modificado o sustituido de modo que se altera su especificidad de sustrato, en la que la cepa codifica para la producción de sangliferina A o un análogo de la misma.
- 25
- 30 Estos y otros aspectos de la invención explicarán mejor en la siguiente divulgación.

Definiciones

35 Los artículos “un/o” y “una” se usan en el presente documento para hacer referencia a uno o a más de uno (es decir al menos uno) de los objetos gramaticales del artículo. A modo de ejemplo “un análogo” significa un análogo o más de un análogo.

40 Tal como se usa en el presente documento, el término “análogo(s)” se refiere a compuestos químicos que son estructuralmente similares a otro pero que se diferencian ligeramente en cuanto a la composición (como en la sustitución de un átomo por otro o en la presencia o ausencia de un grupo funcional particular).

45 Tal como se usa en el presente documento, el término “policétido” se refiere a cualquier molécula generada mediante biosíntesis que implica una policétido sintasa (PKS). Esto puede incluir adicionalmente algunos elementos de dominios de péptido no ribosómico sintasa (NRPS) y/o modificación de biosíntesis adicional, tal como metilación o hidroxilación.

50 Tal como se usa en el presente documento, el término “policétido híbrido” se refiere a cualquier molécula generada mediante biosíntesis que implica una policétido sintasa (PKS), en la que la agrupación de genes que codifica para esta policétido sintasa se ha alterado mediante intervención humana para conducir a un producto de biosíntesis diferente. Esto puede incluir adicionalmente algunos elementos de dominios de péptido no ribosómico sintasa y/o modificación de biosíntesis adicional, tal como metilación o hidroxilación. Las propias alteraciones pueden incluir, pero no se limitan a, mutagénesis dirigida al sitio de dominios (por ejemplo, dominios de aciltransferasa), sustitución de dominios, módulos o genes de la misma agrupación de PKS o NRPS o una heteróloga.

55 Tal como se usa en el presente documento, el término “condiciones de alta rigurosidad” significa condiciones en las que sólo se hibridan secuencias de ADN muy estrechamente relacionadas o idénticas. Esto se realiza con frecuencia en hibridación de tipo Southern aumentando la temperatura del tampón de lavado. Para sondas oligonucleotídicas, la etapa de hibridación puede realizarse a 5°C por encima de la T_f para secuencias perfectamente apareadas, en las que se calcula T_f usando una fórmula tal como $T_f = 4 \times (\text{número de pares de bases GC}) + 2 \times (\text{número de pares de bases AT})$. Un ejemplo de condiciones de alta rigurosidad se facilita en la siguiente sección titulada “Hibridación de ácidos nucleicos”.

60

65 Tal como se usa en el presente documento, el término “huésped heterólogo” en relación con una secuencia de ácido nucleico, particularmente una agrupación de biosíntesis de sangliferina o parte de la misma, significa un huésped que no contendría de manera natural una secuencia de ácido nucleico de este tipo.

Tal como se usa en el presente documento, el término “heterólogo” en relación, por ejemplo, con un dominio o módulo de una PKS o NRPS de sangliferina, significa un dominio o módulo que no estaría presente de manera natural en esa PKS o NRPS.

5 Tal como se usa en el presente documento, el término “no nativo” significa, en relación, por ejemplo, con un dominio o módulo de una PKS o NRPS de sangliferina, un dominio o módulo que no estaría presente de manera natural en esa ubicación en esa PKS o NRPS; por ejemplo podría ser heterólogo (es decir de una PKS o NRPS diferente) o bien podría estar presente en una ubicación diferente en la misma PKS o NRPS.

10 Las sales farmacéuticamente aceptables de compuestos de la invención tales como los compuestos de fórmula (I) incluyen sales convencionales formadas a partir de ácidos o bases orgánicos o inorgánicos farmacéuticamente aceptables así como sales de adición de ácido de amonio cuaternario. Ejemplos más específicos de sales de ácido adecuadas incluyen clorhídrico, bromhídrico, sulfúrico, fosfórico, nítrico, perclórico, fumárico, acético, propiónico, succínico, glicólico, fórmico, láctico, maleico, tartárico, cítrico, palmoico, malónico, hidroximaleico, fenilacético, glutámico, benzoico, salicílico, fumárico, toluenosulfónico, metanosulfónico, naftaleno-2-sulfónico, bencenosulfónico, hidroxinaftoico, yodhídrico, málico, esteroico, tánico y similares. Las sales de ácido clorhídrico son de particular interés. Otros ácidos tales como oxálico, aunque en sí mismos no son farmacéuticamente aceptables, pueden ser útiles en la preparación de sales útiles como productos intermedios en la obtención de los compuestos de la invención y sus sales farmacéuticamente aceptables. Ejemplos más específicos de sales básicas adecuadas incluyen sales de sodio, litio, potasio, magnesio, aluminio, calcio, zinc, N,N'-dibenciletildiamina, cloroprocaina, colina, dietanolamina, etilendiamina, N-metilglucamina y procaina. Las siguientes referencias en el presente documento a un compuesto según la invención incluyen tanto compuestos de fórmula (I) como sus sales farmacéuticamente aceptables.

25 Los grupos alquilo, alqueno y alquino pueden ser de cadena lineal o ramificados.

Los ejemplos de alquilo, por ejemplo grupos alquilo C1-C4, incluyen metilo, etilo, n-propilo, i-propilo y n-butilo.

Descripción detallada de la invención

30 Toda la agrupación de genes según la invención comprende las secuencias de nucleótidos o secuencias complementarias de 24 genes (SEQ ID NO: 1), incluyendo:

35 • un gen de péptido no ribosómico sintetasa (NRPS), *sfaD*, que comprende en conjunto 3 módulos, 10 dominios funcionales, y es responsable de la formación del resto peptídico de la estructura principal macrocíclica;

• cinco genes de policétido sintasa (PKS) lineal, *sfaE*, *sfaF*, *sfaG*, *sfaH*, *sfaI*, responsables de la formación del espirociclo, la cadena larga de policétido y el resto de policétido de la estructura principal macrocíclica;

40 • un gen de policétido sintasa iterativa, *sfaK*, que comprende 4 dominios funcionales, y es responsable de la biosíntesis de un precursor de 6 carbonos especial en la ruta para la biosíntesis de la unidad de extensión de 7 carbonos poco habitual;

45 • tres genes, *sfaA*, *sfaB*, *sfaJ* responsables de la biosíntesis de elementos estructurales de precursores de aminoácidos no naturales;

• siete genes *sfaM*, *sfaN*, *sfaP*, *sfaQ*, *sfaL*, *sfaR*, *sfaO*, responsables de la biosíntesis de precursores, tales como los de la unidad iniciadora;

50 • un gen regulador, *sfaC*, asociado con el rendimiento de fermentación de SFA;

• un gen que codifica para proteína MbtH, *sfaS*, que se postula que está asociado con la regulación de la NRPS; y

• cinco genes con función desconocida, *sfaU1*, *sfaU2*, *sfaV1*, *sfaV2*, *sfaV3*.

55 Por tanto la invención también proporciona una molécula nucleica aislada que comprende uno o más de (a) un gen de PKS lineal seleccionado de *sfaE*, *sfaF*, *sfaG*, *sfaH* y *sfaI*, o (b) un gen de NRPS *sfaD*, o (c) un gen de PKS iterativa *sfaK*, o (d) un gen de biosíntesis de unidad iniciadora o precursor seleccionado de *sfaA*, *sfaB*, *sfaJ*, *sfaM*, *sfaN*, *sfaP*, *sfaQ*, *sfaL*, *sfaO*, o (e) un gen regulador *sfaC*, o (f) un gen que codifica para proteína MbtH, *sfaS*, o (g) un gen de crotonil-CoA reductasa *sfaR*, o (h) un gen de función desconocida seleccionado de *sfaU1*, *sfaU2*, *sfaV1*, *sfaV2* y *sfaV3*.

60 Los genes mencionados anteriormente se definen normalmente mediante ácidos nucleicos que codifican para las proteínas de SEQ ID No. 2-25.

65 En particular, la invención proporciona una molécula nucleica aislada que comprende uno o más de (a) un gen de

PKS lineal seleccionado de *sfaE* que son los residuos 30707-37360 de SEQ ID No. 1, *sfaF* que son los residuos 37394-50014 de SEQ ID No. 1, *sfaG* que son los residuos 50017-60903 de SEQ ID No. 1, *sfaH* que son los residuos 60918-85823 de SEQ ID No. 1 y *sfaI* que son los residuos 85823-96040 de SEQ ID No. 1 o (b) un gen de NRPS *sfaD* que son los residuos 19885-30714 de SEQ ID No. 1 o (c) un gen de PKS iterativa *sfaK* que son los residuos 97396-101943 de SEQ ID No. 1, o (d) un gen de biosíntesis de unidad iniciadora o precursor seleccionado de *sfaA* que son los residuos 17024-17854 de SEQ ID No. 1, *sfaB* que son los residuos 17851-19191 de SEQ ID No. 1, *sfaJ* que son los residuos 96225-97391 de SEQ ID No. 1, *sfaM* que son los residuos 103210-103929 de SEQ ID No. 1, *sfaN* que son los residuos 104001-105023 de SEQ ID No. 1, *sfaP* que son los residuos 105366-107216 de SEQ ID No. 1, *sfaQ* que son los residuos 107366-108145 de SEQ ID No. 1, *sfaL* que son los residuos 101936-103213 de SEQ ID No. 1, *sfaO* que son los residuos 105091-105345 de SEQ ID No. 1, o (e) un gen regulador *sfaC* que son los residuos 19193-19888 de SEQ ID No. 1, o (f) un gen que codifica para proteína MtbH, *sfaS* que son los residuos 109583-109798 de SEQ ID No. 1, o (g) un gen de crotonil-CoA reductasa *sfaR* que son los residuos 108150-109511 de SEQ ID No. 1, o (h) un gen de función desconocida seleccionado de *sfaU1* que son los residuos 14973-15413 de SEQ ID No. 1, *sfaU2* que son los residuos 15596-16063 de SEQ ID No. 1, *sfaV1* que son los residuos 109776-110312 de SEQ ID No. 1, *sfaV2* que son los residuos 111285-111743 de SEQ ID No. 1 y *sfaV3* que son los residuos 112218-112652 de SEQ ID No. 1; o que comprende una secuencia de ácido nucleico que codifica para uno o más de los mismos polipéptidos que los codificados por los genes mencionados anteriormente pero que se diferencian únicamente debido a la redundancia del código genético; o que comprende una secuencia de ácido nucleico que puede hibridarse con una o más de las secuencias génicas anteriores en condiciones de alta rigurosidad; o que comprende una secuencia de ácido nucleico que tiene al menos el 80%, el 85%, el 90%, el 95%, el 96%, el 97%, el 98% o el 99% de identidad con una o más de las secuencias génicas anteriores y que codifica para un polipéptido que tiene la misma función que el producto génico correspondiente; o que comprende una secuencia de ácido nucleico que codifica para uno o más polipéptidos que tienen al menos el 80% de identidad de secuencia de aminoácidos (por ejemplo al menos el 85 o el 90 o el 95 o el 96 o el 97 o el 98 o el 99% de identidad de secuencia) con uno o más de los polipéptidos codificados por los genes mencionados anteriormente y que tienen la misma función; o fragmentos que comprenden al menos 50, 100, 200 ó 500 nucleótidos consecutivos de uno o más de los genes mencionados anteriormente; o el complemento de cualquiera de las secuencias de ácido nucleico mencionadas anteriormente.

También se proporciona una molécula nucleica aislada que comprende uno o más de (a) un módulo o dominio de un gen de PKS seleccionado de *sfaE* que son los residuos 30707-37360 de SEQ ID No. 1, *sfaF* que son los residuos 37394-50014 de SEQ ID No. 1, *sfaG* que son los residuos 50017-60903 de SEQ ID No. 1, *sfaH* que son los residuos 60918-85823 de SEQ ID No. 1 o *sfaI* que son los residuos 85823-96040 de SEQ ID No. 1 o (b) un módulo o dominio del gen de NRPS *sfaD* que son los residuos 19885-30714 de SEQ ID No. 1, o (c) un dominio del gen de PKS iterativa *sfaK* que son los residuos 97396-101943 de SEQ ID No. 1; o que comprende una secuencia de ácido nucleico que codifica para uno o más módulos o dominios de los mismos polipéptidos que los codificados por los genes mencionados anteriormente pero que se diferencian únicamente debido a la redundancia del código genético; o que comprende una secuencia de ácido nucleico que puede hibridarse con una o más de las moléculas de ácido nucleico anteriores en condiciones de alta rigurosidad; o que comprende una secuencia de ácido nucleico que tiene al menos el 80% de identidad con una o más de las moléculas de ácido nucleico anteriores y que codifica para un polipéptido que tiene la misma función que un módulo o dominio del producto génico correspondiente; o que comprende una secuencia de ácido nucleico que codifica para uno o más polipéptidos que tienen al menos el 80% de identidad de secuencia de aminoácidos con uno o más de los polipéptidos codificados por un módulo o dominio de los genes mencionados anteriormente y que tienen la misma función; o fragmentos que comprenden al menos 50 nucleótidos consecutivos de uno o más de los módulos o dominios de los genes mencionados anteriormente; o el complemento de cualquiera de las secuencias de ácido nucleico mencionadas anteriormente.

Un módulo o dominio de un gen es la parte del gen que codifica para un módulo o dominio del producto génico correspondiente.

También se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido basado en SEQ ID No. 1 en el que uno o más (por ejemplo un) dominio(s), módulo(s) o gen(es) se han delecionado, mutado de modo que una función enzimática o reguladora se vuelve inactiva o menos activa, mutado de modo que tiene una funcionalidad alterada, o sustituido mediante una sustitución, o en el que se han insertado uno o más (por ejemplo un) dominio(s), módulo(s) o gen(es) no nativos, por ejemplo en el que uno o más (por ejemplo un) dominio(s), módulo(s) o gen(es) se han sustituido por (a) un dominio, módulo o gen de otra parte en la agrupación de genes de biosíntesis de sangliferina A o (b) un dominio, módulo o gen que es heterólogo con respecto a la agrupación de genes de biosíntesis de sangliferina A o en el que uno o más (por ejemplo un) dominio(s), módulo(s) o gen(es) se han mutado para hacer inactivar o hacer menos activa una función enzimática o reguladora.

Por ejemplo, se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido en el que uno o más genes de PKS seleccionados de *sfaE*, F, G, H y I se han modificado mediante lo cual se han delecionado uno o más dominios o módulos, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos por ejemplo mediante lo cual se han insertado uno o más dominios o módulos de otra parte en la agrupación de genes de biosíntesis de SFA o de una agrupación de genes de

biosíntesis de policétido heteróloga.

Los módulos de los genes de PKS *sfaE*, F, G, H y I pueden verse mediante referencia a la figura 7. Por tanto los módulos pueden comprender normalmente dominios KS-AT-ACP o KS-AT-DH-KR-ACP o KS-AT-KR-ACP o KS-AT-DH-ER-KR-ACP.

Por ejemplo:

- puede sustituirse un dominio AT por un dominio AT de una PKS heteróloga o de otra parte dentro de la PKS de SFA que tiene una especificidad de sustrato diferente; y/o

- puede delecionarse un dominio DH o volverse inactivo; y/o

- puede insertarse un dominio DH de la PKS de SFA o de una PKS heteróloga en un módulo;

- puede delecionarse un dominio ER o volverse inactivo; y/o

- puede insertarse un dominio ER de la PKS de SFA o de una PKS heteróloga en un módulo;

- puede sustituirse el bucle reductor de un módulo dado (que significa los dominios DH-KR o DH-ER-KR o KR cuando están presentes) por un bucle reductor que tiene elementos diferentes.

En una realización, el dominio AT del módulo 13 puede sustituirse por un dominio AT de la PKS de SFA o de una PKS heteróloga que tiene una especificidad de sustrato diferente, por ejemplo una que tiene especificidad por metilmalonilo. Esta mutación particular conduce a la introducción de un grupo metilo en la posición 14.

En una realización el dominio DH de uno o más de los módulos 1, 3, 6, 7, 8, 10, 11 y 13 se deleciona o se vuelve inactivo. Estas mutaciones particulares conducen a la introducción de uno o más grupos hidroxilo en la molécula, por ejemplo en las posiciones 21 y/o 25.

También se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido en el que el gen de NRPS *sfaD* se ha modificado mediante lo cual se han delecionado uno o más dominios o módulos, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.

También se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido en el que el gen regulador *sfaC* se ha modificado de modo que se aumenta o disminuye su actividad, o se deleciona.

También se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido en el que el gen de PKS iterativa *sfaK* se ha modificado mediante lo cual se han delecionado uno o más dominios, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios no nativos por ejemplo mediante lo cual se han insertado uno o más dominios de otra parte en la agrupación de genes de biosíntesis de SFA o de una agrupación de biosíntesis de policétido heteróloga.

El módulo del gen de PKS iterativa *sfaK* puede verse mediante referencia a la figura 7. Por tanto el módulo comprende dominios KS-AT-ACP-KR.

También se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido en el que uno o más genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaA*, B, J, M, N, P, Q, L y O se han delecionado o modificado de modo que se disminuye su actividad o modificado o sustituido de modo que se altera su selectividad de sustrato.

También se proporciona un ácido nucleico de agrupación de genes de producción de policétido híbrido en el que uno o más genes de biosíntesis de precursor o unidad iniciadora o uno o más operones que contienen uno o más genes de biosíntesis de precursor o unidad iniciadora se han delecionado o mutado de modo que son inactivos o menos activos en la producción de dicha unidad iniciadora o precursor que el gen u operón nativo.

Realizaciones más preferidas de este aspecto de la invención incluyen ácido nucleico aislado que codifica para un dominio de la PKS o NRPS de SEQ ID No. 1, residuos 19885-30714, 30707-37360, 37394-50014, 50017-60903, 60918-85823, 85823-96040. Estos ácidos nucleicos pueden usarse, solos o en combinación con ácidos nucleicos que codifican para otros dominios o módulos de PKS o NRPS como productos intermedios, por ejemplo en la construcción de vectores recombinantes.

La presente invención también proporciona un método para identificar, aislar y clonar ácido nucleico que incluye uno

cualquiera de los fragmentos de ADN descritos anteriormente. Un método preferido comprende, por ejemplo, las siguientes etapas:

- 5 a) establecer un banco de ADN genómico (por ejemplo una biblioteca de cósmidos)
 - b) examinar este banco con la ayuda de las secuencias de ADN de esta invención
 - c) aislar los clones identificados como positivos.
- 10 Un método general para identificar los fragmentos de ADN implicados en la biosíntesis de sangliferinas comprende, por ejemplo, las siguientes etapas:
- 15 a. el aislamiento de fragmentos de ADN con homología con la agrupación de genes de sangliferina puede llevarse a cabo mediante transferencia de tipo Southern de una biblioteca de cósmidos, analizando con sonda con fragmentos de ADN (por ejemplo de ~1 kb) de SEQ ID No. 1 para encontrar fragmentos clonados con homología con la agrupación de sangliferina;
 - b. entonces pueden retirarse los cósmidos que se observa que se hibridan con la sonda, y secuenciarse el ADN;
 - 20 c. entonces pueden aislarse regiones de ADN adyacentes analizando con sonda la biblioteca de cósmidos con el cósmido marcado aislado anteriormente, para cósmidos que contienen ADN solapante.

Otros métodos se describen en Maniatis *et al.*, 1998, Sambrook y Russell, 2001 y Kieser *et al.*, 1999.

- 25 La invención proporciona además una secuencia de nucleótidos, que codifica para una proteína de función desconocida. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 2, y se denomina *sfaU1*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 14973 hasta la posición de base 15413.
 - 30 La invención proporciona además una secuencia de nucleótidos que codifica para otra proteína de función desconocida. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 3, y se denomina *sfaU2*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 15596 hasta la posición de base 16063.
 - 35 La invención proporciona además una secuencia de nucleótidos que codifica para una fenilalanina meta-hidroxilasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 4, y se denomina *sfaA*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 17024 hasta la posición de base 17854.
 - 40 La invención proporciona además una secuencia de nucleótidos que codifica para una ornitina N⁵-oxigenasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 5, y se denomina *sfaB*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 17851 hasta la posición de base 19191.
 - 45 La invención proporciona además una secuencia de nucleótidos que codifica para un factor regulador de la transcripción. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 6, y se denomina *sfaC*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 19193 hasta la posición de base 19888.
 - 50 La invención proporciona además una secuencia de nucleótidos que codifica para un péptido no ribosómico sintetasa que comprende los dominios funcionales C, A, PCP, C, A, PCP, C, A, PCP, C, y es responsable de la biosíntesis del resto peptídico de la estructura principal macrocíclica. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 7, y se denomina *sfaD*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 19885 hasta la
 - 55 posición de base 30714.
- La invención proporciona además una secuencia de nucleótidos que codifica para una policétido sintasa que comprende los dominios funcionales ACP, KS, AT, DH, ER, KR, ACP, y es responsable de la biosíntesis del precursor del resto espirocíclico. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 8, y se denomina *sfaE*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 30707 hasta la posición de base 37360.
- 60
 - 65 La invención proporciona además un dominio de carga específico para iniciar la biosíntesis de PKS con un sustrato 2-etilmalonamilo-S-tioéster, que consiste en el primer ACP de *sfaE*, la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 30707 hasta la posición de base 31082.

- 5 La invención proporciona además una secuencia de nucleótidos que codifica para una policétido sintasa que comprende los dominios funcionales KS, AT, ACP, KS, AT, KR, ACP, KS, AT, KR, ACP, y es responsable de la biosíntesis de la cadena larga de policétido. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 9, denominada *sfaF*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 37394 hasta la posición de base 50014.
- 10 La invención proporciona además una secuencia de nucleótidos que codifica para una policétido sintasa que comprende los dominios funcionales KS, AT, KR, ACP, KS, AT, DH, ER, KR, ACP, y es responsable de la biosíntesis de gran parte de la cadena larga de policétido. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 10, denominada *sfaG*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 50017 hasta la posición de base 60903.
- 15 La invención proporciona además una secuencia de nucleótidos que codifica para una policétido sintasa que comprende los dominios funcionales KS, AT, DH, KR, ACP, KS, AT, DH, KR, ACP, KS, AT, KR, ACP, KS, AT, DH, KR, ACP, KS, AT, DH, KR, ACP, y es responsable de la biosíntesis de parte del resto de policétido de la estructura principal macrocíclica. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 11, denominada *sfaH*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 60918 hasta la posición de base 85823.
- 20 La invención proporciona además una secuencia de nucleótidos que codifica para una policétido sintasa que comprende los dominios funcionales KS, AT, KR, ACP, KS, AT, KR, ACP, y es responsable de la biosíntesis de parte del resto de policétido de la estructura principal macrocíclica. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 12, denominada *sfaI*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 85823 hasta la posición de base 96040.
- 25 La invención proporciona además una secuencia de nucleótidos que codifica para una dedo de zinc deshidrogenasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 13, denominada *sfaJ*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 96225 hasta la posición de base 97391.
- 30 La invención proporciona además una secuencia de nucleótidos que codifica para una policétido sintasa iterativa que comprende los dominios funcionales KS, AT, ACP, KR. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 14, denominada *sfaK*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 97396 hasta la posición de base 101943.
- 35 La invención proporciona además una secuencia de nucleótidos que codifica para una aciltransferasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 15, denominada *sfaL*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 101936 hasta la posición de base 103213.
- 40 La invención proporciona además una secuencia de nucleótidos que codifica para una deshidrogenasa/reductasa de cadena corta. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 16, denominada *sfaM*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 103210 hasta la posición de base 103929.
- 45 La invención proporciona además una secuencia de nucleótidos que codifica para una acil-cetoácido sintasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 17, denominada *sfaN*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 104001 hasta la posición de base 105023.
- 50 La invención proporciona además una secuencia de nucleótidos que codifica para una proteína transportadora de acilo libre. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 18, denominada *sfaO*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 105091 hasta la posición de base 105345.
- 55 La invención proporciona además una secuencia de nucleótidos que codifica para un análogo de asparagina sintasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 19, denominada *sfaP*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 105366 hasta la posición de base 107216.
- 60 La invención proporciona además una secuencia de nucleótidos que codifica para una tioesterasa libre. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 20, denominada *sfaQ*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 107366 hasta la posición de base 108145.
- 65

5 La invención proporciona además una secuencia de nucleótidos que codifica para una crotonil-coA reductasa. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 21, denominada *sfaR*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 108150 hasta la posición de base 109511.

10 La invención proporciona además una secuencia de nucleótidos que codifica para una proteína de la familia de MbtH, la secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 22, denominada *sfaS*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 109583 hasta la posición de base 109798.

15 La invención proporciona además una secuencia de nucleótidos que codifica para una proteína de función desconocida. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 23, denominada *sfaV1*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 109776 hasta la posición de base 110312.

20 La invención proporciona además una secuencia de nucleótidos que codifica para otra proteína de función desconocida. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 24, denominada *sfaV2*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 111285 hasta la posición de base 111743.

25 La invención proporciona además una secuencia de nucleótidos que codifica para otra proteína de función desconocida. La secuencia de aminoácidos para la que codifica es tal como se muestra en SEQ ID NO: 25, denominada *sfaV3*, y la secuencia de nucleótidos correspondiente es tal como se muestra en SEQ ID NO: 1, comenzando desde la posición de base 112218 hasta la posición 112652.

30 La secuencia complementaria de SEQ ID NO: 1 puede obtenerse según el principio de complementariedad de bases de ADN. La secuencia de nucleótidos o parte de la secuencia de nucleótidos de SEQ ID NO: 1 puede obtenerse mediante reacción en cadena de la polimerasa (PCR), o mediante digestión con enzimas de restricción adecuada de ADN correspondiente, o mediante otras técnicas adecuadas. La invención proporciona además un método para obtener un plásmido de ADN recombinante que comprende al menos parte de la secuencia de ADN en SEQ ID NO: 1.

35 La invención proporciona además un método para obtener un microorganismo que contiene genes de biosíntesis de SFA interrumpidos, en el que al menos uno de los genes comprende la secuencia de nucleótidos tal como se muestra en SEQ ID NO: 1.

40 La secuencia de nucleótidos o parte de la secuencia de nucleótidos según la invención puede obtenerse mediante un método basado en la reacción en cadena de la polimerasa (PCR), o un gen similar al gen de biosíntesis de SFA puede obtenerse a partir de otro organismos mediante un método de hibridación de tipo Southern o similar, usando un fragmento de ADN que comprende una secuencia según la invención como sonda.

45 Un ADN clonado que comprende la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención puede usarse para identificar más plásmidos de biblioteca a partir de la biblioteca genómica de *Streptomyces sp. A92-308110 (S. flaveolus)*. Estos plásmidos de biblioteca comprenden al menos parte de la secuencia según la invención, y también comprenden ADN que no se han clonado tampoco a partir de las regiones adyacentes en el genoma de *Streptomyces sp. A92-308110 (S. flaveolus)*.

50 Por tanto, por ejemplo, una secuencia de nucleótidos o ácido nucleico según cualquier aspecto de la invención es un ADN.

55 La secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención puede modificarse o mutarse. Estos métodos incluye inserción, sustitución o delección, reacción en cadena de la polimerasa, reacción en cadena de la polimerasa propensa a errores, mutagénesis específica del sitio, nuevo ligamiento de secuencias diferentes, intercambio de ADN con diferentes partes de la secuencia o con secuencias homólogas de otra fuente, o mutagénesis mediante UV o agente químico, etc.

60 La presente invención también proporciona vectores recombinantes tales como vectores de expresión de ADN que comprenden los ácidos nucleicos anteriores. Los vectores comprenden normalmente el ADN mencionado anteriormente junto con uno o más promotores u otros elementos reguladores. Estos vectores y métodos de esta invención le permiten a un experto en la técnica generar células huésped recombinantes con la capacidad de producir policétidos. Por tanto la invención proporciona un método de preparación de un policétido, tal como sangliferina A o un análogo de sangliferina A, comprendiendo dicho método cultivar una célula huésped transformada, célula huésped que se ha transformado con un vector de expresión que comprende ácidos nucleicos que codifican para la totalidad o parte de la agrupación de genes de sangliferina descrita en SEQ ID No. 1. También se proporciona un policétido que puede producirse mediante el método mencionado anteriormente y que no es

sangliferina A como un aspecto de la invención. De manera deseable, el vector comprende ácido nucleico que codifica para una PKS funcional que puede generar sangliferina A o un análogo de sangliferina A si se cultiva en un medio apropiado. En algunas realizaciones, la célula huésped transformada produce de manera natural sangliferina A. En algunas realizaciones, la célula huésped transformada no produce de manera natural sangliferina A. También se proporciona una célula huésped transformada de este tipo como un aspecto de la invención. Una realización específica es una célula huésped transformada con un vector que comprende ácido nucleico que codifica para la totalidad o parte de la agrupación de genes de biosíntesis de sangliferina A (por ejemplo codifica para la agrupación de genes de biosíntesis de sangliferina A mostrada a modo de ejemplo mediante el ácido nucleico de SEQ ID No. 1), célula huésped que no produce de manera natural sangliferina A.

Por tanto, también se proporcionan:

- un polipéptido o pluralidad de polipéptidos codificados por cualquiera de los ácidos nucleicos mencionados anteriormente;

- una policétido sintasa codificada por el ácido nucleico mencionado anteriormente que codifica para una o más proteínas de biosíntesis de policétido; y

- una proteína híbrida codificada por un ácido nucleico o agrupación de genes modular en la que al menos un dominio o módulo o gen no es nativo con respecto a la agrupación de genes de biosíntesis de sangliferina A.

En otra realización, la invención proporciona un polipéptido aislado que comprende una secuencia seleccionada del grupo que consiste en SEQ ID No. 2-25; un polipéptido aislado que consiste en al menos 10, 50, 100, 200 ó 500 aminoácidos consecutivos de los polipéptidos de SEQ ID No. 2-25; y un polipéptido aislado que tiene al menos el 80%, el 85%, el 90%, el 95%, el 97% o el 99% de homología con las secuencias anteriores tal como se determina mediante BLASTP (Altschul *et al.*, 1990) con los parámetros por defecto.

La invención también proporciona un método de preparación de un policétido híbrido, comprendiendo dicho método transformar una célula huésped con un vector recombinante que comprende ácidos nucleicos que codifican para la totalidad o parte de la agrupación de genes de sangliferina descrita en SEQ ID No. 1 en el que al menos un dominio o módulo o gen no es nativo con respecto a la agrupación de genes de biosíntesis de sangliferina A, después cultivar la célula huésped transformada.

La invención también proporciona un policétido que puede producirse mediante el procedimiento mencionado anteriormente que no es sangliferina A.

El gen clonado que comprende la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención puede expresarse en un huésped heterogéneo mediante un sistema de expresión adecuado, para obtener un rendimiento o actividad enzimática o biológica de otro tipo correspondiente o superior. Estos huéspedes heterogéneos incluyen *Streptomyces*, *Pseudomonas*, *E. coli*, *Bacillus*, levaduras, plantas y animales, etc.

Los expertos en la técnica conocen bien que las agrupaciones de genes de policétido pueden expresarse en huéspedes heterólogos (Pfeifer *et al.*, 2001). Por consiguiente, la presente invención incluye la transferencia de la agrupación de genes de biosíntesis de sangliferina con o sin genes reguladores y de resistencia, o bien completa, modificada por ingeniería, que contiene mutaciones o bien que contiene deleciones, para su complementación en huéspedes heterólogos. En la técnica se conocen bien métodos y vectores para la transferencia tal como se definió anteriormente de grandes fragmentos de ADN de este tipo (Rawlings, 2001; Staunton y Weissman, 2001) o se proporcionan en el presente documento en los métodos dados a conocer.

En este contexto una cepa de célula huésped heteróloga preferida es un procarionta, más preferiblemente un actinomiceto o *Escherichia coli*, todavía más preferiblemente incluye, pero no se limita a, *S. hygrosopicus*, *S. hygrosopicus sp.*, *S. hygrosopicus var. ascomyceticus*, *Streptomyces tsukubaensis*, *Streptomyces coelicolor*, *Streptomyces lividans*, *Saccharopolyspora erythraea*, *Streptomyces fradiae*, *Streptomyces avermitilis*, *Streptomyces cinnamonensis*, *Streptomyces rimosus*, *Streptomyces albus*, *Streptomyces griseofuscus*, *Streptomyces longisporoflavus*, *Streptomyces venezuelae*, *Micromonospora griseorubida*, *Amycolatopsis mediterranei* o *Actinoplanes sp. N902-109*.

Por tanto, en un aspecto adicional, la presente invención proporciona cepas novedosas en las que uno o más de los genes que codifican para la biosíntesis de sangliferina están inactivados o delecionados.

En un aspecto adicional, la presente invención proporciona una cepa modificada por ingeniería basada en una cepa productora de SFA en la que uno o más genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaA*, B, J, M, N, P, Q, L y O se han delecionado o modificado de modo que se disminuye su actividad, o modificado o sustituido de modo que se altera su selectividad de sustrato.

La cepa modificada por ingeniería mencionada anteriormente basada en una cepa productora de SFA puede

producir SFA o un análogo de SFA si se alimenta apropiadamente. La cepa modificada por ingeniería puede tener todas y cada una de las combinaciones de las modificaciones anteriores y puede tener modificaciones génicas adicionales. Las secuencias de aminoácidos según la invención pueden usarse para aislar una proteína deseada, y pueden usarse para la preparación de anticuerpos.

Con determinados aminoácidos delecionados o sustituidos, polipéptidos que comprenden la secuencia de aminoácidos o al menos parte de la secuencia según la invención todavía pueden tener actividad biológica o incluso tener una nueva actividad biológica o propiedades deseables tales como rendimiento aumentado o cinética de proteínas optimizada u otros.

Un gen o una agrupación de genes que comprende la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención puede expresarse en huéspedes heterogéneos, y su función en la cadena metabólica del huésped puede investigarse mediante tecnología de chip de ADN.

Puede construirse un plásmido recombinante mediante una recombinación genética a partir de un gen o una agrupación de genes que comprende la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención, para establecer la ruta de biosíntesis, o puede usarse un método de inserción, sustitución, deleción o inactivación para establecer la ruta de biosíntesis.

Puede usarse un fragmento de ADN o gen clonado que comprende la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención para generar nuevos análogos estructurales de SFA interrumpiendo una o varias etapas de la biosíntesis de SFA. El gen o fragmento de ADN puede usarse para aumentar el rendimiento de SFA o derivados de la misma, la invención proporciona un método para aumentar el rendimiento en un microorganismo modificado por ingeniería genética. Ejemplos de esto incluyen inactivación o deleción de reguladores negativos.

La policétido sintasa según la invención puede usarse para generar nuevos policétidos o policétidos híbridos, mediante deleción, inserción, alteración o inactivación de uno o más dominios, módulos o genes de la policétido sintasa del mismo sistema de policétido sintasa o de uno diferente.

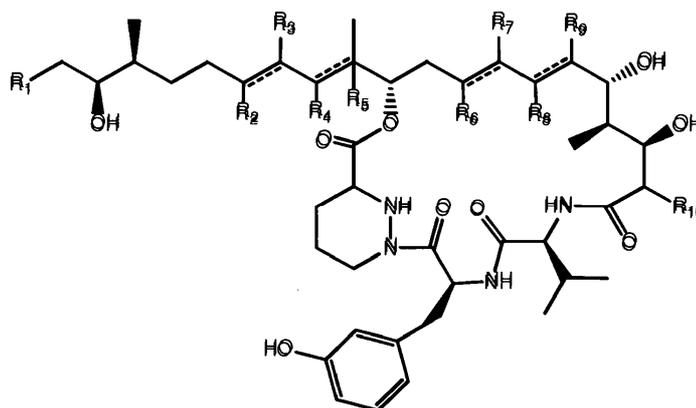
Pueden usarse fragmentos o genes que comprenden la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos según la invención para construir una biblioteca de policétido sintasas, o una biblioteca de derivados de policétido sintasas, o una biblioteca combinatoria.

En una palabra, toda la información sobre genes y proteínas asociada con la biosíntesis de SFA proporcionada en el presente documento puede facilitar el entendimiento del mecanismo de biosíntesis de productos naturales de SFA, y proporciona materiales y conocimiento para hacer avanzar la ingeniería genética. Los genes y sus proteínas proporcionados en el presente documento pueden usarse para la investigación y el descubrimiento de compuestos, genes o proteínas que pueden usarse en medicina, industria o agricultura.

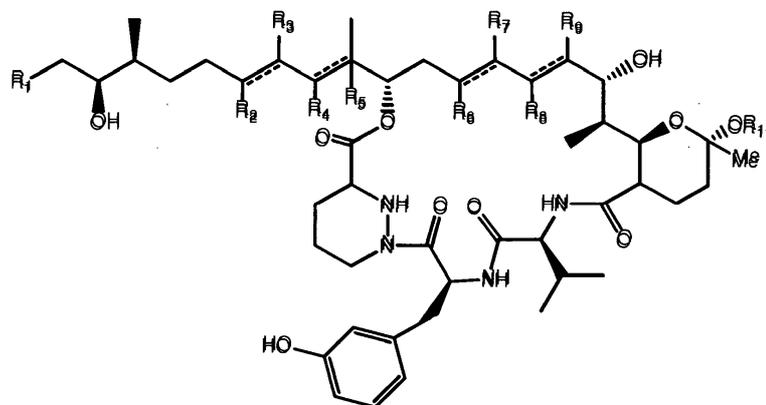
Adicionalmente, en la presente invención, se han realizado cambios específicos a la agrupación de genes de biosíntesis de sangliferina, con la intención original de entender la ruta de biosíntesis, mostrar ejemplos de las invenciones en el presente documento, y generar moléculas para examinar el mecanismo de actividad inmunosupresora.

En general, pueden determinarse porcentajes de identidad de secuencia mediante BLASTN (ácido nucleico) o BLASTP (proteína) (Altschul *et al.*, 1990) con los parámetros por defecto. La agrupación de genes según la invención puede usarse para la preparación de compuestos análogos de sangliferina de fórmula (I) o (II) o una sal farmacéuticamente aceptable de los mismos:

Fórmula (I)

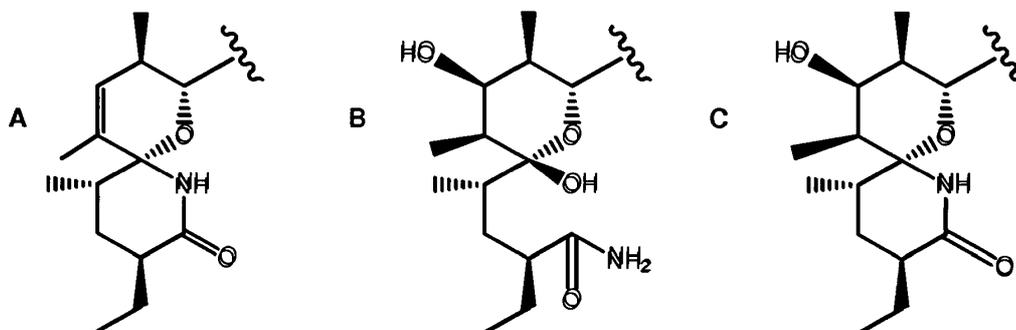


Fórmula (II)



5 en las que:

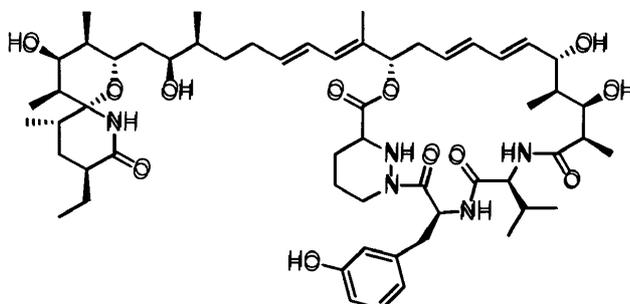
R₁ representa uno de los restos A, B o C:



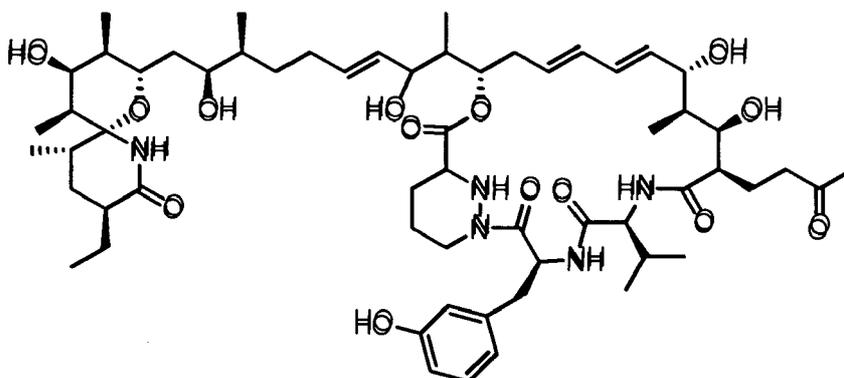
- 10 R₂ representa OH y R₃ representa H o R₂ y R₃ representan un enlace;
 R₄ representa OH y R₅ representa H o R₄ y R₅ representan un enlace;
 15 R₆ representa OH y R₇ representa H o R₆ y R₇ representan un enlace;
 R₈ representa OH y R₉ representa H o R₈ y R₉ representan un enlace;
 20 R₁₀ representa Me o CH₂CH₂C(O)CH₃;
 R₁₁ representa H o Me;
- 25 con la condición de que cuando R₁₀ representa CH₂CH₂C(O)CH₃, entonces ambos de R₂ y R₄ no pueden representar OH y/o ambos de R₆ y R₈ no pueden representar OH,
 25 y con la condición de que cuando R₁₀ representa CH₂CH₂C(O)CH₃, entonces R₂ y R₄, y R₆ y R₈ no pueden representar todos enlaces.
- 30 En una realización, R₂ representa OH, R₃ representa H y R₄ y R₅, R₆ y R₇ y R₈ y R₉ representan enlaces y R₁₀ representa CH₂CH₂C(O)CH₃.
- En una realización, R₅ representa OH, R₆ representa H y R₂ y R₃, R₆ y R₇ y R₈ y R₉ representan enlaces y R₁₀ representa CH₂CH₂C(O)CH₃.
- 35 En una realización, R₇ representa OH, R₈ representa H y R₄ y R₅, R₆ y R₂ y R₃ y R₉ representan enlaces y R₁₀ representa CH₂CH₂C(O)CH₃.
- En una realización, R₈ representa OH, R₉ representa H y R₄ y R₅, R₆ y R₇ y R₂ y R₃ representan enlaces y R₁₀ representa CH₂CH₂C(O)CH₃.
- 40 Preferiblemente R₁ representa un resto C.

Preferiblemente R₁₁ representa H.

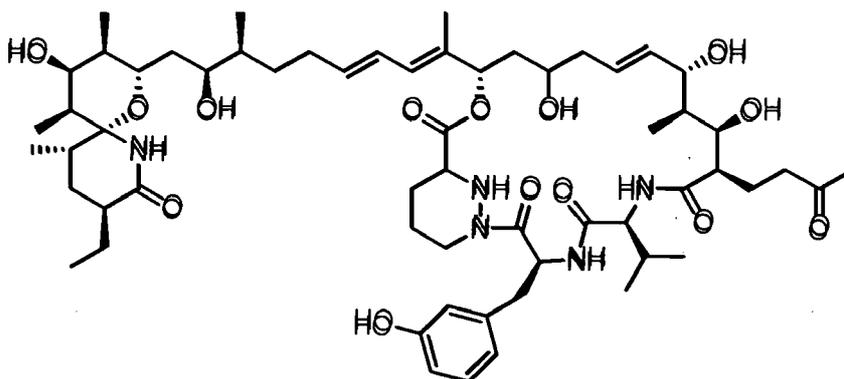
5 Un ejemplo específico de un compuesto que puede prepararse empleando la agrupación de genes de la presente invención incluye los siguientes:



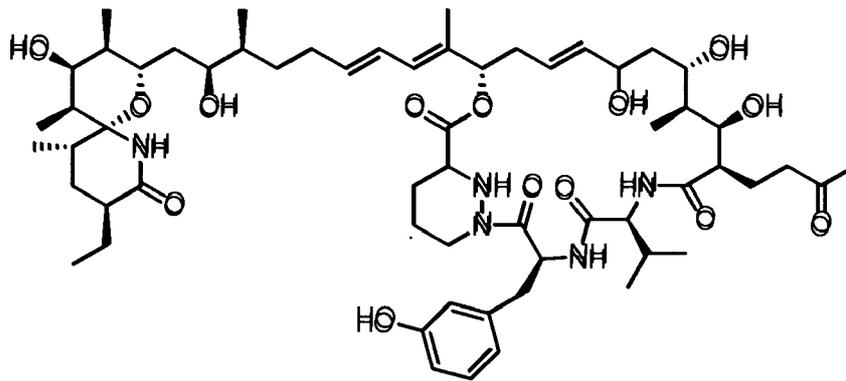
10 que puede esperarse que se produzca mediante delección/inactivación de *sfaK* opcionalmente junto con un intercambio del AT del módulo 13 por un módulo de AT que acepta malonato de metilo en la agrupación de genes de biosíntesis de SFA;



15 que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 8 en la agrupación de genes de biosíntesis de SFA;

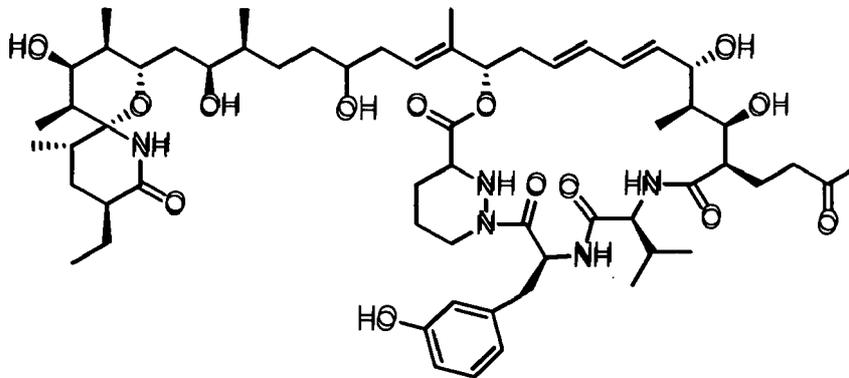


20 que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 10 en la agrupación de genes de biosíntesis de SFA;



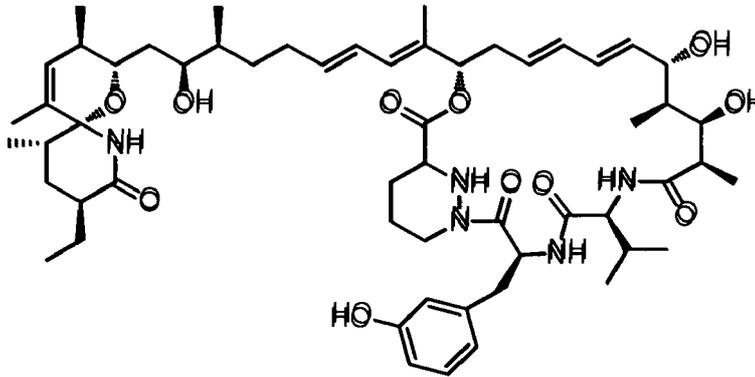
que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 11 en la agrupación de genes de biosíntesis de SFA;

5



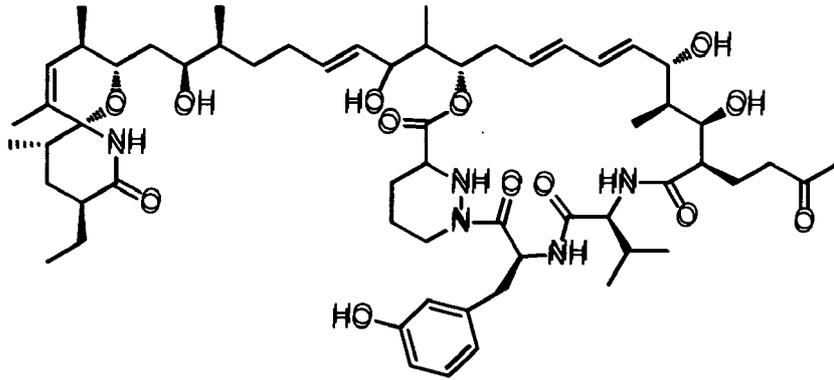
que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 7 en la agrupación de genes de biosíntesis de SFA;

10



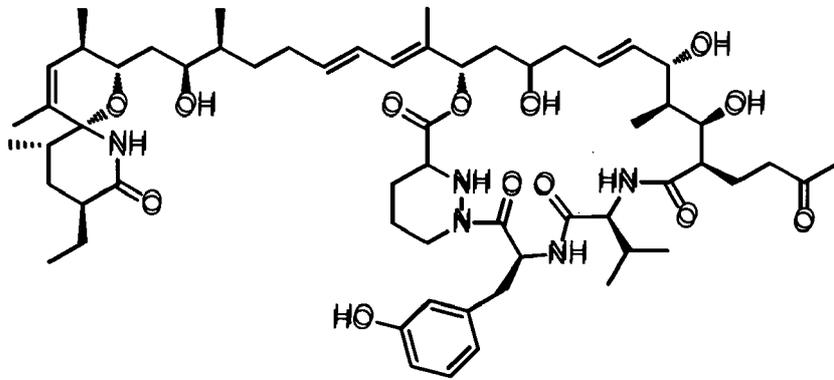
que puede esperarse que se produzca mediante delección/inactivación de *sfaK* opcionalmente junto con el intercambio del AT del módulo 13 por un módulo de AT que acepta malonato de metilo en la agrupación de genes de biosíntesis de SFA;

15



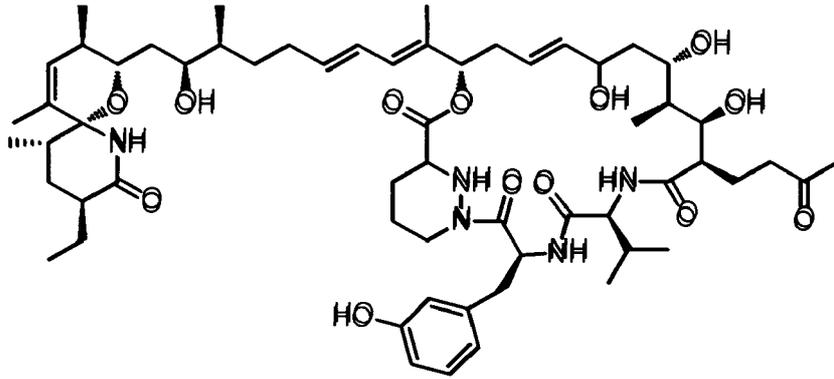
que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 8 en la agrupación de genes de biosíntesis de SFA;

5



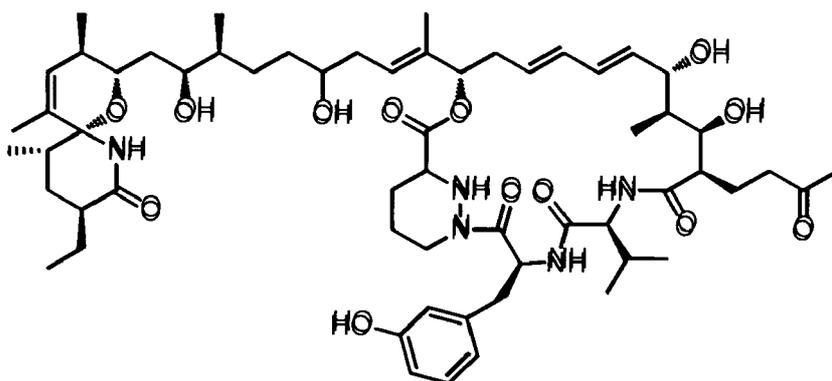
que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 10 en la agrupación de genes de biosíntesis de SFA;

10



que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 11 en la agrupación de genes de biosíntesis de SFA;

15



que puede esperarse que se produzca mediante delección/inactivación del DH del módulo 7 en la agrupación de genes de biosíntesis de SFA;

5

y sales farmacéuticamente aceptables de los mismos.

Las estructuras anteriores muestran un tautómero representativo de todos los tautómeros de los compuestos de formulas (I) y (II), por ejemplo compuestos ceto cuando se ilustran compuestos de enol y viceversa, que también pueden resultar de la agrupación de genes de la presente invención.

10

Por motivos de completitud, la invención proporciona además:

- Una agrupación de genes de biosíntesis de una sangliferina A inmunosupresora, caracterizada porque comprende 24 genes responsables de la biosíntesis de sangliferina A, que son:

15

1) un gen de estructura principal de péptido no ribosómico sintetasa (NRPS), *sfaD*, en el que:

sfaD se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 19885-30714, tiene 10830 pares de bases de longitud, y codifica para una péptido no ribosómico sintetasa de 3609 aminoácidos;

20

2) cinco genes de policétido sintasa (PKS) lineal de tipo I, *sfaE*, *sfaF*, *sfaG*, *sfaH*, *sfaI*, en los que:

sfaE se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 30707-37360, tiene 6654 pares de bases de longitud, y codifica para una policétido sintasa de 2217 aminoácidos;

25

sfaF se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 37394-50014, tiene 12621 pares de bases de longitud, y codifica para una policétido sintasa de 4206 aminoácidos;

sfaG se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 50017-60903, tiene 10887 pares de bases de longitud, y codifica para una policétido sintasa de 3628 aminoácidos;

30

sfaH se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 60918-85823, tiene 24906 pares de bases de longitud, y codifica para una policétido sintasa de 8301 aminoácidos;

35

sfaI se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 85823-96040, tiene 10218 pares de bases de longitud, y codifica para una policétido sintasa de 3405 aminoácidos;

3) un gen de policétido sintasa iterativa de tipo I, *sfaK*, en el que:

sfaK se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 97396-101943, tiene 4548 pares de bases de longitud, y codifica para una policétido sintasa de 1515 aminoácidos;

40

4) diez genes funcionales para síntesis de precursor, *sfaA*, *sfaB*, *sfaJ*, *sfaM*, *sfaN*, *sfaP*, *sfaQ*, *sfaR*, *sfaL*, *sfaO*, en los que:

45

sfaA se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 17024-17854, tiene 831 pares de bases de longitud, y codifica para una fenilalanina meta-hidroxilasa de 276 aminoácidos;

sfaB se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 17851-19191, tiene 1341 pares de bases de longitud, y codifica para una N⁵-ornitina oxigenasa de 446 aminoácidos;

50

sfaJ se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 96225-97391, tiene 1167 pares

de bases de longitud, y codifica para una dedo de zinc deshidrogenasa de 388 aminoácidos;

sfaM se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 103210-103929, tiene 720 pares de bases de longitud, y codifica para una deshidratasa/reductasa de cadena corta de 239 aminoácidos;

5 *sfaN* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 104001-105023, tiene 1023 pares de bases de longitud, y codifica para una cetosintasa de 340 aminoácidos;

10 *sfaP* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 105366-107216, tiene 1851 pares de bases de longitud, y codifica para un análogo de asparagina sintasa de 616 aminoácidos;

sfaQ se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 107366-108145, tiene 780 pares de bases de longitud, y codifica para una tioesterasa de 259 aminoácidos;

15 *sfaR* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 108150-109511, tiene 1362 pares de bases de longitud, y codifica para una crotonil-coA reductasa de 453 aminoácidos;

20 *sfaL* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 101936-103213, tiene 1278 pares de bases de longitud, y codifica para una aciltransferasa de 425 aminoácidos altamente homóloga con respecto al dominio de transacilación en policétido sintasa;

sfaO se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 105091-105345, tiene 255 pares de bases de longitud, y codifica para una proteína transportadora de acilo de 84 aminoácidos;

25 5) dos genes reguladores, *sfaC* y *sfaS*, en los que:

sfaC se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 19193-19888, tiene 696 pares de bases de longitud, y codifica para un factor regulador de la transcripción de 231 aminoácidos;

30 *sfaS* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 109583-109798, tiene 216 pares de bases de longitud, y codifica para un factor de MbtH de 71 aminoácidos;

6) cinco genes que codifican para proteínas de función desconocida, *sfaU1*, *sfaU2*, *sfaV1*, *sfaV2*, *sfaV3*, en los que:

35 *sfaU1* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 14973-15413, tiene 441 pares de bases de longitud, y codifica para una proteína de función desconocida de 146 aminoácidos;

40 *sfaU2* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 15596-16063, tiene 468 pares de bases de longitud, y codifica para una proteína de función desconocida de 155 aminoácidos;

sfaV1 se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 109776-110312, tiene 537 pares de bases de longitud, y codifica para una proteína de función desconocida de 178 aminoácidos;

45 *sfaV2* se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 111285-111743, tiene 459 pares de bases de longitud, y codifica para una proteína de función desconocida de 152 aminoácidos;

sfaV3 se ubica en la secuencia de nucleótidos de la agrupación de genes en las bases 112218-112652, tiene 435 pares de bases de longitud, y codifica para una proteína de función desconocida de 144 aminoácidos.

50 - La péptido no ribosómico sintetasa en la agrupación de genes de biosíntesis de sangliferina A, caracterizada porque comprende los siguientes módulos o dominios: dominio de enzima de condensación de peptidilo C, dominio de adenilación A, proteína transportadora de peptidilo PCP, y un dominio de terminación C para la ciclación del grupo terminal.

55 - La policétido sintasa en la agrupación de genes de biosíntesis de sangliferina A, caracterizada porque comprende los siguientes módulos o dominios: dominio de condensación de ceto KS, dominio de transacilación AT, dominio de proteína transportadora de acilo ACP, dominio de deshidratación DH, dominio de reducción de ceto KR, y dominio de reducción de enol ER.

60 - Un uso de cualquier proteína codificada por la agrupación de genes de biosíntesis de sangliferina A, para la síntesis catalítica de la sangliferina A inmunosupresora y cualquier análogo correspondiente de su familia.

- Un uso de cualquier proteína codificada por la agrupación de genes de biosíntesis de sangliferina A, para la síntesis catalítica de la estructura principal de un policétido híbrido-péptido no ribosómico.

65 - El uso de la agrupación de genes de biosíntesis de sangliferina A, en el que se realiza una modificación genética, y

la fermentación biológica del mutante obtenido produce análogos no naturales de sangliferina A, tales como los hidroxilados en la posición 21 y 25.

- 5 - El uso de la agrupación de genes de biosíntesis de sangliferina A, en el que se realiza una modificación genética, y la fermentación biológica del mutante obtenido produce productos compensatorios, análogos no naturales de sangliferina A que están sustituidos con grupo metilo en la posición 14.

Descripción de las figuras

10 Figura 1: Estructuras químicas de sangliferinas (sangliferina A y análogos B, C y D).

15 Figura 2: Organización génica y mapa de restricción de agrupación de genes de biosíntesis de SFA. (A) Cuatro cósmidos solapados que representan la región de ADN de ~150 kb del genoma de *Streptomyces sp. A92-308110* (*S. flaveolus*), B representa la enzima de restricción *Bam*HI, la línea continua representa las partes que ya se han sometido a secuenciación de ADN, Sonda P1 a P4 representan sondas marcadas; (B) la organización genética de la agrupación de genes de biosíntesis de SFA. Desconocido: gen desconocido; PKS: gen de policétido sintasa; NRPS: gen de péptido no ribosómico sintetasa; Más allá de la agrupación: gen fuera de los límites; Transposasa: transposón; Modificación: gen de síntesis de precursor.

20 Figura 3: Rutas de biosíntesis propuestas para cada elemento estructural de SFA. (A) ácido piperázico; (B) unidad iniciadora; (C) unidad de extensión especial. En las etapas marcadas con *Sfa*X, X corresponde a la proteína codificada por el gen respectivo en la agrupación de genes tal como se describe en el presente documento; marcar una etapa con el símbolo “?” significa que aún no queda claro si la etapa se realiza por una enzima codificada por un gen de esta agrupación de genes, o si se realiza *in vivo* con ayuda de una enzima implicada en el metabolismo primario; una etapa sin ninguna marca significa que puede realizarse con ayuda de un proceso de metabolismo primario.

30 Figura 4: Ruta para la formación de la estructura principal de SFA. PKS: policétido sintasa; NRPS: péptido no ribosómico sintetasa; KS: dominio funcional para la síntesis de ceto; AT: dominio funcional para la transferencia de acilo; KR: dominio funcional para la reducción de ceto; ER: dominio funcional para la reducción de enol; DH: dominio funcional para la deshidratación; ACP: proteína transportadora de acilo; C: dominio funcional para la condensación de péptido; A: dominio funcional para la adenilación; PCP: proteína transportadora de peptidilo.

35 Figura 5: Análisis mediante cromatografía de líquidos de alta resolución (HPLC) de productos de fermentación derivados de la interrupción de la secuencia de agrupación de genes de biosíntesis de SFA mediante una sonda clonada. WT: tipo natural; Patrón: patrón de SFA; Mutante: cepa mutante.

40 Figura 6: Análisis mediante cromatografía de líquidos de alta resolución (HPLC) de productos de fermentación de una cepa mutante resultante de la sustitución génica. (A) tipo natural, (B) mutante con desactivación de gen de deshidrogenasa de unión a zinc *sfaJ*, (C) mutante con desactivación de gen de policétido sintasa lineal iterativa *sfaK*, (D) mutante con desactivación de gen de aciltransferasa *sfaL*, (E) mutante con desactivación de gen de deshidrogenasa de cadena corta *sfaM*, (F) mutante con desactivación de gen de fenilalanina hidroxilasa *sfaA*, (G) mutante con desactivación de gen de ornitina oxigenasa *sfaB*, (H) mutante con desactivación de gen de ácido graso cetosintasa *sfaN*, (I) mutante con desactivación de gen de análogo de asparagina sintasa *sfaP*, (J) mutante con desactivación de gen de crotonil-coA reductasa *sfaR*, (K) mutante con desactivación de proteína desconocida de flanco izquierdo *sfaU1*, (L) mutante con desactivación de factor regulador TetR de flanco derecho, (M) mutagénesis dirigida al sitio de dominio DH del módulo 8, (N) mutagénesis dirigida al sitio de dominio DH del módulo 10. *Streptomyces sp. A92-308110* (*S. flaveolus*); mv.: cepa mutante (una marca usada para distinguirla del tipo natural); *tsekangensis*: mutante T (produce SFT); *hasangensis*: mutante H (produce SFH); *xuwengensis*: mutante X (produce SFX).

50 Figura 7: Figura resumen que contiene la ruta de biosíntesis propuesta.

55 La numeración de módulos de la figura 3 se ha reemplazado por la de la figura 7. Por tanto cuando se hace referencia a números de módulo en el presente documento, los números de módulo son a los que se hace referencia en la figura 7.

Descripción de las secuencias de la lista de secuencias

60 SEQ ID No. 1: secuencia de ácido nucleico de la agrupación de genes de biosíntesis de SFA

SEQ ID No. 2: secuencia de aminoácidos de SfaU1

65 SEQ ID No. 3: secuencia de aminoácidos de SfaU2

SEQ ID No. 4: secuencia de aminoácidos de SfaA

- SEQ ID No. 5: secuencia de aminoácidos de SfaB
- 5 SEQ ID No. 6: secuencia de aminoácidos de SfaC
- SEQ ID No. 7: secuencia de aminoácidos de SfaD
- SEQ ID No. 8: secuencia de aminoácidos de SfaE
- 10 SEQ ID No. 9: secuencia de aminoácidos de SfaF
- SEQ ID No. 10: secuencia de aminoácidos de SfaG
- 15 SEQ ID No. 11: secuencia de aminoácidos de SfaH
- SEQ ID No. 12: secuencia de aminoácidos de SfaI
- SEQ ID No. 13: secuencia de aminoácidos de SfaJ
- 20 SEQ ID No. 14: secuencia de aminoácidos de SfaK
- SEQ ID No. 15: secuencia de aminoácidos de SfaL
- SEQ ID No. 16: secuencia de aminoácidos de SfaM
- 25 SEQ ID No. 17: secuencia de aminoácidos de SfaN
- SEQ ID No. 18: secuencia de aminoácidos de SfaO
- 30 SEQ ID No. 19: secuencia de aminoácidos de SfaP
- SEQ ID No. 20: secuencia de aminoácidos de SfaQ
- 35 SEQ ID No. 21: secuencia de aminoácidos de SfaR
- SEQ ID No. 22: secuencia de aminoácidos de SfaS
- SEQ ID No. 23: secuencia de aminoácidos de SfaV1
- 40 SEQ ID No. 24: secuencia de aminoácidos de SfaV2
- SEQ ID No. 25: secuencia de aminoácidos de SfaV3
- 45 SEQ ID No. 26-29: Secuencias de cebador descritas en el ejemplo 1
- SEQ ID No. 30 y 31: Secuencias de cebador descritas en el ejemplo 11
- SEQ ID No. 32-47: Secuencias de cebador descritas en el ejemplo 12
- 50 SEQ ID No. 48-63: Secuencias de cebador descritas en el ejemplo 14
- SEQ ID No. 64-67: Secuencias de cebador descritas en el ejemplo 16
- 55 Métodos generales
- En general, se usaron métodos tal como se describe en Maniatis *et al.*, 1998, Sambrook y Russell, 2001 o en Kieser *et al.*, 1999. A continuación se detallan ejemplos o métodos alternativos.
- 60 Cepas bacterianas y plásmidos
- Se hizo crecer *Escherichia coli* DH10B (GibcoBRL) en medio TY 2x tal como se describe por Sambrook y Russell (2001) y *E. coli* ET12567(pUZ8002) tal como se describe en Paget *et al.* (1999) en medio TY 2x con kanamicina (25 mg/l) y cloranfenicol (12,5 mg/l). Se seleccionaron transformantes de *E. coli* con ampicilina 100 mg/l o apramicina 50 mg/l.
- 65 La mayoría de los plásmidos usados se describen en Kieser *et al.*, 1999 y Sambrook *et al.*, 2001. pTV1, también

conocido como pBS3030, se describe en Cheng *et al.*, 2003, pIJ773 se describe en el n.º de registro de GenBank AX657066.1. pANT841 se describe en Genbank: AF438749

- 5 *Streptomyces sp. A92-308110 (S. flaveolus)* productor de sangliferina puede obtenerse de DSMZ, Braunschweig, Alemania como *Streptomyces sp. DSM 9954*.

Recetas de medios

Medio agar AS-1:

10

	g/100 ml
Extracto de levadura	0,1
L-Alanina	0,02
L-Arginina	0,05
Almidón soluble	0,5
NaCl	0,25
Na ₂ SO ₄	1,0
Agar	2,0
Agua destilada hasta 100 ml	
Se ajusta el medio mediante NaOH a pH 7,5;	

Medio agar ISP4:

	g/l
Almidón soluble	10
K ₂ HPO ₄	1
MgSO ₄ ·7H ₂ O	1
NaCl	1
(NH ₄) ₂ SO ₄	2
CaCO ₃	2
Disolución de sales traza de ISP	1 ml
Agar	20
Agua destilada hasta 1000 ml, ajustado a pH 7,2 antes de la esterilización	

- 15 Disolución de sales traza de ISP:

	g/l
FeSO ₄ ·7H ₂ O	1
MnCl ₂ ·4H ₂ O	1
ZnSO ₄ ·7H ₂ O	1
Agua destilada hasta 1000 ml	

Medio agar IWL-4:

- 20 Se preparó medio agar IWL-4 añadiendo, a cada litro de medio ISP-4, 1 g de triptona y 0,5 g de extracto de levadura, y ajustando el medio a pH 7,2 mediante NaOH.

Método general de conjugación

- 25 Se inocularon bacterias donadoras, *E. coli* S17-1 que contenían el plásmido de interés, en tubos de ensayo en 3-4 ml de caldo LB complementado con apramicina 50 µg/ml con agitación durante la noche. Entonces se inocularon 500 µl en 50 ml de LB complementado con apramicina 50 µg/ml en matraces de 250 ml a 37°C y se hicieron hasta una DO₆₀₀ de 0,5. Se recuperaron las células en un tubo de EP de 50 ml mediante centrifugación a 3800 rpm durante 10 min. Después se resuspendieron con 20 ml de LB mediante agitación con vórtex y se recuperaron de nuevo. Esto se repitió dos veces y se resuspendieron las células recuperadas en 1 ml de LB. Se centrifugaron bacterias aceptoras, *Streptomyces sp. A92-309110 (S. flaveolus)*, a 12000 rpm durante 3 min para recuperar las esporas y se resuspendieron en 500 µl de TES 0,1 M dos veces, después se sometieron a choque térmico a 50°C durante 10 min y se añadieron 500 µl de caldo TSB. Después se incubaron las esporas a 37°C durante 4-5 h. Se recuperaron las esporas y se resuspendieron en 1 ml de LB. Se mezclaron las bacterias en razones de desde 1:99 hasta 99:1, proporcionando generalmente la razón 1:1 una transformación satisfactoria. Se extendieron las bacterias mezcladas sobre agar MS y se secaron las placas en un armario de flujo laminar. Después se incubaron las placas a 30°C durante 14-16 h. Se extendieron de nuevo las placas con 3-4 ml de agua estéril para eliminar la mayor parte de *E. coli* y se secaron las placas durante 1 h en un armario de flujo laminar. Después se recubrieron las placas con 1 ml de agua estéril que contenía apramicina a 1 mg/ml y ácido nalidíxico a 1 mg/ml y se incubaron a 30°C durante
- 30
- 35
- 40 3-5 días. Se seleccionaron uno o dos o más clones individuales y se inocularon en 3 ml de caldo TSB con

apramicina a 30 µg/ml con agitación a 30°C durante varios días. Se extendieron 100 µl de cultivo reciente sobre agar ISP-4 y se permitió que se produjera un cruzamiento doble mediante incubación durante 2-3 días.

Extracción de ADN total a partir de *Streptomyces* sp A92-309110 (*S. flaveolus*)

Se inocularon suspensiones de esporas de *Streptomyces*, sp A92-309110 (*S. flaveolus*), almacenadas a -80°C, en 3 ml de medio YEME. Se incubó el cultivo durante 12 h a 30°C, se transfirió a tubos de EP de 1,5 ml en dos porciones, se centrifugó, y se recuperaron las bacterias. Se resuspendieron las bacterias en 500 µl de disolución de STE; se añadió lisozima hasta una concentración final de 4 mg/ml, y después se incubó en un baño de agua a 37°C durante 30-45 min. Después de que las bacterias se habían vuelto se observó que se convirtieron en una sustancia transparente y de tipo gel. Después, se añadieron 250 µl de SDS al 2% y 60 µl de KAc 5 mol/l, y se mezclaron bien, se congelaron a -20°C durante 10 min, después se centrifugaron a 12000 rpm durante 10 min. Se transfirieron los sobrenadantes a nuevos tubos de EP de 1,5 ml. Se añadieron 500 µl de mezcla de fenol/cloroformo 1:1 (v/v), y se centrifugaron a 10000 rpm durante 3 min. Esta etapa se repitió varias veces hasta que no apareció más proteína desnaturizada blanca en la interfase. Después se usó cloroformo solo para otra extracción final. Se retiró el sobrenadante, se mezcló bien con un volumen igual de isopropanol, se dejó a temperatura ambiente durante 30 min para precipitar el ADN, y se centrifugó a 10000 rpm durante 5 min. Tras lavar con alcohol al 70% y secar mediante un filtro de succión, se añadieron 200 µl de tampón TE y 10 µl de ARNasa libre de ADNasa 10 mg/ml, y se congeló a -20°C durante el almacenamiento.

Fermentación de la cepa productora de SFA, *Streptomyces* sp A92-309110 (*S. flaveolus*), y cepas relacionadas, aislamiento, purificación e identificación de producto

a) Fermentación líquida

A partir de una disolución madre de esporas, se transfirieron 50 µl (concentración promedio de 1×10^8 por ml) a 3 ml de TSB (caldo de tripticasa y soja, Sigma) a presión de antibiótico apropiada (apramicina 50 µg/ml o eritromicina 50 µg/ml dependiendo del marcador resistente del mutante) y se agitó (a 27°C, 250 rpm) durante 30 h. A partir del cultivo primario, se transfirió 1 ml a 50 ml de medio de siembra en un matraz Erlenmeyer de 250 ml (diámetro, 8,5 cm en el fondo y 3,5 cm en la boca; altura, 15 cm) y se agitó (a 27°C, 250 rpm).

A partir del cultivo de siembra, se transfirieron 5 ml a 90 ml de medio de producción en un matraz Erlenmeyer de 500 ml (diámetro: 10,5 cm en el fondo y 4 cm en la boca, altura: 18 cm) y se hicieron crecer durante 24 h a 25°C (250 rpm), momento en el cual se añadieron 10 ml de medio de producción reciente que contenía 4 g de resina XAD-16 al cultivo de producción (total de 4 g de XAD-16 en 100 ml de medio de producción). Se incubó el cultivo en las condiciones anteriores (25°C, 220 rpm), durante 3 días adicionales.

Tras completarse la fermentación en el día 4, se hizo pasar el caldo de fermentación a través de un filtro de papel ayudado mediante succión. Se transfirió el filtro de papel con el material restante sobre el mismo a un vaso de precipitados, y se congeló a -80°C durante 30 min, y se liofilizó durante la noche. Se añadieron 80-100 ml de metanol al material secado, se agitó durante 40 min, y se filtró usando succión. Se extrajo el material filtrado con metanol por segunda vez, y se filtró usando succión. Se desechó la resina, y se combinaron las disoluciones metanólicas, y se secaron a vacío a 35-37°C. Tras la evaporación hasta sequedad, se transfirió el material usando dos alícuotas de 750 µl de metanol. Se combinaron las dos disoluciones metanólicas y se transfirieron alícuotas de 1,5 ml a tubos Eppendorf, y se congelaron a -20°C durante 1 h. Se centrifugaron los tubos a 12000 rpm, 4°C, durante 10 min. Se transfirió el sobrenadante a nuevos tubos Eppendorf, se secó a vacío hasta que quedaban aproximadamente 500 µl de líquido, que entonces se almacenó a 4°C.

Se purificaron aproximadamente 300 µl del extracto en bruto obtenido tal como se explicó anteriormente mediante cromatografía en columna ultrarrápida (RP-18), que se aclaró previamente con 40 ml de acetonitrilo al 40% en agua. Después se eluyó la columna con 40 ml de acetonitrilo al 50% en agua, y se recogieron fracciones de 40 ml en tubos separados, y después se concentraron. Se purificaron adicionalmente muestras previamente tratadas usando CL-EM (usando acetonitrilo al 80% en agua que contenía ácido fórmico al 1%) para el aislamiento y la identificación.

b) Fermentación sólida

Para la inoculación de varias placas de agar, se transfirieron 40 µl de disolución madre de esporas a 3 ml de medio TSB, y se incubaron durante la noche a 37°C (mientras se agitaba a 250 rpm). A partir de este cultivo durante la noche, se transfirieron 400-500 µl y se extendieron uniformemente sobre la superficie de placas de agar (150 mm) de medio agar R2YE, IWL-4, ISP-4 o AS-1. Alternativamente, pueden usarse alícuotas (50 µl) de disolución madre de esporas para inocular directamente las placas de agar. Se incubaron a 30°C las placas de agar inoculadas durante 6-7 días.

Al final del periodo de incubación, se recogió el cultivo en el medio de agar en un vaso de precipitados de 250 ml y se trituró. Se añadieron 200 ml de metanol al material de una placa de agar y se agitó la mezcla durante 2 h,

después se filtró a través de papel de filtro ayudado mediante succión. Se desecharon los materiales sólidos, y se secó el filtrado en un evaporador rotatorio. Con el fin de recuperar el producto de fermentación, se usaron dos alícuotas de 500 µl de metanol para resuspender y transferir el material secado a tubos Eppendorf de 1,5 ml. La purificación del material procedente del concentrado metanólico siguió un procedimiento similar al explicado anteriormente para el procedimiento de fermentación líquida. Sin embargo, dado que en este caso el contenido de agua en los medios sólidos es bajo, puede omitirse la etapa de liofilización. Se purificaron adicionalmente las muestras usando CL-EM (usando acetonitrilo al 80% en agua que contenía ácido fórmico al 1%) para el aislamiento y la identificación.

10 Clonación de genes de biosíntesis de SFA mediante PCR

El sistema de PCR comprende: DMSO (al 8%, v/v), MgCl₂ (25 mM), dNTP (2,5 mM), cebadores degenerados (10 µM), ADN polimerasa Taq (2,5 u), y una cantidad adecuada del ADN total de *Streptomyces* sp A92-309110 (*S. flaveolus*) como molde. En primer lugar, 95°C, 3 min, 1 ciclo; después, 94°C, 1 min, 68°C, 1 min, 72°C, 2 min, 5 ciclos; 94°C, 1 min, 65°C, 1 min, 72°C, 2 min, 30 ciclos; finalmente, 72°C, 10 min, 1 ciclo. Tras terminarse la PCR, se usó la electroforesis en gel de agarosa al 1% para examinar los resultados. Se recuperaron fragmentos de ADN de tamaños deseados a partir del gel de bajo punto de fusión, y se ligaron con vector pGEM T Easy. Se transformaron células competentes *E. coli* DH5α, y se extendieron sobre una placa de LB que contenía ampicilina, se usaron IPTG (isopropil-β-D-tiogalactopiranosido) y X-gal (5-bromo-4-cloro-3-indolil-β-D-galactósido) para la selección azul-blanco. Se escogieron las colonias bacterianas blancas para el cultivo durante la noche, y se extrajeron los plásmidos. Se evaluó mediante digestión con *Eco*RI si los plásmidos contenían fragmentos de ADN insertados de tamaños deseados. Después se secuenciaron estos plásmidos.

25 Hibridación de ácido nucleico

Se inocularon varios microlitros de micelio o esporas de la cepa mutante que iba a evaluarse en 3 ml de medio TSB, se agitaron a 220 rpm, y se cultivaron a 30°C hasta que se espesó el medio de cultivo. Después se extrajo ADN genómico. Basándose en un análisis de secuencia, se diseñó una estrategia de hibridación, se prepararon sondas, y se seleccionaron sitios de enzimas de restricción adecuados para la digestión de ADN genómico.

Se incubaron 15 µl de sonda (que contenía 0,5-3 µg de ADN) en un baño de agua en ebullición durante 10 min, después se transfirieron inmediatamente a un baño de sal-hielo para enfriar. Se añadieron secuencialmente 2 µl de mezcla de hexanucleótidos (10x), 2 µl de mezcla de dNTP marcados, 1 µl de enzima de Klenow marcada y se mezclaron bien, después se incubó la mezcla en un baño de agua a 37°C durante 16 h. Se añadieron 0,8 µl de EDTA 0,8 M (pH=8,0) para detener la reacción, y se añadieron 2,5 µl de LiCl 4 M y se mezclaron bien. Se añadieron 75 µl de alcohol anhidro previamente helado para la precipitación. Se congeló la mezcla a -80°C durante 40 min, y se centrifugó a 12000 rpm, 4°C durante 20 min. Se recogió el ADN, se lavó con alcohol previamente helado al 70%, se secó a vacío y se disolvió en 50 µl de tampón TE y se almacenó a 4°C. Tras la electroforesis, se empapó el gel de agarosa en HCl 0,25 M durante 20 min, se aclaró con agua desionizada, y después se empapó durante otros 20 min con agitación suave tras añadirse el tampón básico. Tras intercambiar el tampón una vez, se empapó el gel durante otros 20 min. Tras lavar varias veces con agua desionizada, se transfirió ADN a una membrana de nailon. Se usó un aparato de transferencia (BioRad) para la transferencia. Siguiendo las instrucciones del aparato, se colocó un trozo de papel de filtro húmedo sobre la placa inferior, después se recubrió con una membrana de nailon de tamaño adecuado. La membrana debe ser 1 cm más grande que el gel por cada lado, después se recubrió con una membrana de plástico y se fijaron las pinzas. Se trató el gel de agarosa tal como se describió anteriormente y después se colocó sobre el poro de membrana. Tras conectarse el aparato con la bomba de vacío, se ajusta la presión para mantenerla a 5-8 mmHg. Tras realizar un sello estanco al vacío, se añade aproximadamente 1 l de tampón de hibridación SSC 10x, y la superficie de líquido debe cubrir el gel de agarosa. La transferencia se realiza durante 2 h. Entonces puede realizarse la fijación cocinando durante 40 min en un horno a 120°C, o alternativamente exponiendo a luz UV (2 J/cm²).

Después se almacena la membrana en un lugar oscuro y frío antes de su uso. Se colocó la membrana de nailon fijada en tubos de hibridación, y se añadió disolución de hibridación. Se realizó una hibridación previa a 64°C durante 30 min, después se añadieron sondas, y se realizó la hibridación a la temperatura deseada durante 16 h. Se lavó secuencialmente la membrana dos veces durante 5 min con tampón de lavado riguroso SSC 2x a temperatura ambiente; después se lavó dos veces durante 15 min con tampón de lavado riguroso SSC 0,5x a 64°C. Tras el lavado riguroso, en primer lugar se equilibró la membrana de nailon durante 1-5 min con tampón de lavado, y luego se incubó con disolución de bloqueo durante 30 min, después se incubó con disolución de anticuerpo durante 30 min. Tras 1-2 lavados con tampón de lavado, se equilibró durante 2-5 min con tampón de detección. Se añadió tampón de revelado, y se dejó que se revelase la membrana en la oscuridad. Cuando se alcanzó una intensidad adecuada, se lavó la membrana con agua desionizada para detener la reacción, y se completó la hibridación.

Ejemplos

65 Se realizará una descripción adicional de la invención con referencia a las figuras adjuntas en el siguiente contexto.

Ejemplo 1: Clonación del fragmento de gen de crotonil-CoA reductasa a partir de la agrupación de genes de SFA

Hasta ahora, aunque se ha completado satisfactoriamente la síntesis total de SFA, su ruta de producción natural casi no se conoce, particularmente con respecto al mecanismo de formación del enlace N-N de la unidad de ácido piperázico, y la unidad de extensión de policétido poco habitual que probablemente usa una policétido sintasa iterativa. Durante los últimos años, estudios marcados con isótopo han mostrado que el ácido glutámico y la glutamina pueden ser los precursores de biosíntesis del ácido piperázico en determinados sistemas que producen productos naturales que comprenden unidades de ácido piperázico (Umezawa *et al.*, 2001; Miller *et al.*, 2007). Según análisis de agrupaciones de genes existentes y de otras agrupaciones de genes notificadas para productos naturales que comprenden ácido piperázico, no se excluye la posibilidad de que la ornitina sea un sustrato; además, la unidad de extensión poco habitual, como sustituyente lateral en la estructura principal macrocíclica, puede formarse basándose en una policétido sintasa iterativa poco habitual. Las dos unidades mencionadas anteriormente pueden desempeñar papeles importantes en el desarrollo de nuevos fármacos y el estudio de relaciones estructural-actividad. Por tanto, existe la necesidad de un análisis de la agrupación de genes de biosíntesis de SFA para esclarecer el mecanismo de biosíntesis de SFA, y después desarrollar su posible valor farmacéutico.

Basándose en el análisis de la estructura de PKS-NRPS híbrida para la estructura principal macrocíclica, se supuso que la agrupación de genes presentaba una región de PKS-NRPS lineal larga, sin homólogos conocidos, que conducía a un aumento en la dificultad de aislamiento y clonación de la agrupación de genes de biosíntesis. De hecho, mediante análisis con sonda con elementos de genes de PKS y NRPS, se aisló ADN de 11 agrupaciones de PKS y 7 de NRPS, después se usó para estudios de inactivación de genes. Todo esto no condujo a ningún efecto sobre la producción de sangliferina. Finalmente, los inventores consiguieron clonar la agrupación de genes usando otras unidades específicas como sondas. Mediante análisis de agrupaciones de genes de biosíntesis que se producen de manera natural para productos naturales de policétido que comprenden estructuras basadas en grupo etilo, los inventores observaron que las unidades de etilo se introducían generalmente usando etilmalonil-CoA como elemento estructural. Los inventores concibieron que la agrupación de genes de SFA puede contener un homólogo de CCR. Por tanto, se diseñaron cebadores de CCR degenerados para clonar la secuencia de la región altamente conservada de la crotonil-coA sintetasa responsable de la biosíntesis de la unidad de etilo, que se pensaba que estaba presente en el genoma de *S. sp.* A92-308110 (*S. flaveolus*). Después se marcó la secuencia clonada como sonda para realizar un examen de biblioteca. Las secuencias de cebadores degenerados fueron las siguientes:

CCR largo-dir. (SEQ ID: 26): AGGAAT TCATGG CCTCCK CSRTSA ACTACA AY,

CCR largo-inv. (SEQ ID: 27): TCGGAT CCGCCG AAGTTS GTCTCR WABCCC CA;

CCR corto-dir. (SEQ ID: 28): AGGAAT TCGACA TCGACA TSGTBW TCGAG CA,

CCR corto-inv. (SEQ ID: 29) TCGGAT CCGATG ATGCGC TTSWSB KDCATC CA.

Usando estos cebadores, se amplificaron dos secuencias de 900 pb y 300 pb a partir del genoma de *Streptomyces sp.* A92-309110 (*S. flaveolus*). Se cortó el gel, y se recuperaron los fragmentos de PCR y se digirieron mediante *EcoRI* y *BamHI*, y después se clonaron en pSP72 en el sitio *EcoRI/BamHI*. Se identificaron los plásmidos recombinantes mediante digestión enzimática y secuenciación. En total, se clonaron tres conjuntos diferentes de fragmentos de gen de CCR que eran altamente homogéneos entre sí.

Ejemplo 2: Clonación, análisis de secuencia y análisis de función de agrupación de genes de biosíntesis de SFA

Se marcaron fragmentos clonados tal como se describió anteriormente en el ejemplo 1 con digoxina, y se realizó un examen de biblioteca. Volvieron a dividirse en grupos los 3 conjuntos resultantes de cósmidos obtenidos, y se establecieron sus posiciones relativas según mapeo de restricción. Se eligió el fragmento de extremo de *BamHI* de 6,4 kb de pTL3101 (cQXD03-126-6) como sonda para realizar paseo cromosómico, y se obtuvo el cósmido pSL36. Después, se eligió el fragmento de *BamHI* de 1,4 kb para realizar el paseo cromosómico, y se obtuvieron los cósmidos pTL3104 (cQXD04-49-1~50). Después se eligió el fragmento de *BamHI* en el extremo izquierdo para realizar el paseo y se obtuvieron los cósmidos pTL3106 (cQXD04-64-1~40). Tras estas tres rondas de paseo cromosómico, se clonó un total de aproximadamente 150 kb de ADN cromosómico. Usando cebadores degenerados de PKS y NRPS respectivamente, se realizaron ensayos de PCR con algunos cósmidos, a partir de algunos de los cuales se amplificaron bandas específicas. Esto confirmó que la agrupación de genes comprende CCR y PKS-NRPS híbrido.

Tabla 1: Análisis de función basado en secuencia de la agrupación de genes de biosíntesis de SFA

gen	aminoácido	proteína similar	función de proteína homóloga	similitud / identidad
<i>sfaU1</i>	150	ZP_03193264	proteína desconocida	60/45
<i>sfaU2</i>	112	AAW49302	resolvasa/integrasa	81/75
<i>sfaA</i>	276	YP_001509705	Fenilalanina meta-hidroxilasa	70/57

<i>sfaB</i>	446	YP_001852050	ornitina N ⁵ -oxigenasa	73/61
<i>sfaC</i>	231	YP_001852059	factor regulador de la transcripción	73/60
<i>sfaD</i>	3609	YP_631823	péptido no ribosómico sintetasa	57/43
<i>sfaE</i>	2217	ABC87510	policétido sintasa	65/52
<i>sfaF</i>	4206	ABB88533	policétido sintasa	59/48
<i>sfaG</i>	3628	NP_824075	policétido sintasa	63/54
<i>sfaH</i>	8301	ZP_03174068	policétido sintasa	63/52
<i>sfaI</i>	3405	DQ_450945	policétido sintasa	73/65
<i>sfaJ</i>	332	YP_112153	deshidrogenasa de unión a zinc	67/51
<i>sfaK</i>	1515	YP_001104567	policétido sintasa lineal iterativa	58/47
<i>sfaL</i>	425	YP_480040	aciltransferasa	54/42
<i>sfaM</i>	239	YP_949278	deshidrogenasa de cadena corta	60/43
<i>sfaN</i>	340	YP_001537010	cetosintasa	66/52
<i>sfaO</i>	84	AAD_20269	proteína transportadora de acilo	61/45
<i>sfaP</i>	616	BAB12569	análogo de asparagina sintasa	77/68
<i>sfaQ</i>	259	CAQ64680	tioesterasa	70/58
<i>sfaR</i>	453	YP_002204695	crotonil-coA reductasa	91/86
<i>sfaS</i>	71	YP_882427	proteína MbtH	82/67
<i>sfaV1</i>	178	CAD18995	proteína desconocida	81/65
<i>sfaV2</i>	152	YP_002199237	proteína hipotética	63/53
<i>sfaV3</i>	144	YP_001159721	proteína hipotética	57/45

Se seleccionaron los tres cósmidos con la mayor cobertura para someterse a secuenciación de longitud completa, y los inventores obtuvieron 118.372 pb de secuencia de nucleótidos continua. Un análisis bioinformático reveló que comprendía 44 marcos de lectura abiertos (ORF), al menos 19 ORF de los cuales se sugirió que estaban relacionados con la biosíntesis de SFA, incluyendo: 1 gen de NRPS (*sfaD*) responsable de la síntesis de la estructura principal de tripéptido macrocíclica; 5 genes de PKS (*sfaE-I*) que completan la síntesis de la estructura principal de policétido macrocíclica; 11 supuestos genes de síntesis de precursores (*sfaA-B*, *sfaJ-R*), responsables de la síntesis de la unidad iniciadora y de cada unidad de condensación, y de la modificación de la estructura principal; 2 genes reguladores (*sfaC* y *sfaS*) implicados en la regulación de la biosíntesis de SFA. En la tabla anterior se mostraron resultados del análisis de función de cada gen.

Ejemplo 3: Determinación de los límites de la agrupación de genes de biosíntesis de SFA

Según el análisis de secuencia, a aproximadamente 40 kb del punto de inicio de secuenciación en el sentido de 5', hay un perfil incompleto de transposición (en el sentido de 5' de *sfaU1* y *sfaU2*); a aproximadamente 20 kb desde el punto final de secuenciación en el sentido de 3', hay obviamente dos genes de transposasa (es decir *sfaT2* y *sfaT3*), pero falta parte del gen de transposición en el sentido de 5'. Usando las dos secuencias descritas anteriormente como límites, todos los ORF pueden dividirse en tres secciones:

La secuencia de aproximadamente 120 kb en la parte central es el núcleo responsable de la biosíntesis de la estructura principal de SFA, tal como se describió anteriormente.

La sección en el sentido de 5' comprende 16 genes que supuestamente no son relevantes para la biosíntesis de SFA, incluyendo 9 genes funcionales (gen de monooxigenasa de citocromo p450 *orf16*, gen de proteína transportadora de acilo *orf14*, gen de adenilasa *orf13*, gen de metiltransferasas *orf12*, gen de hidroxilasa de citocromo p450 *orf11*, gen de desoxi-guanina cinasa *orf10*, gen de deshidrogenasa *orf9* y gen de proteinasa M23S *orf6*), 5 genes que codifican para proteínas desconocidas (*orf15*, *orf7*, *orf5*, *sfaJ1* y *sfaJ2*), 2 genes que codifican para factores reguladores (*orf4* y *orf8*, que pertenecen a la familia de TetR y LysR, respectivamente), y 1 gen de proteína acoplada a transferencia de glicosilo (*orf3*).

La sección en el sentido de 3' comprende genes que supuestamente no son relevantes para la biosíntesis de SFA, incluyendo 3 genes que codifican para proteínas desconocidas (*sfaV1*, *sfaV2* y *sfaV3*) y 2 genes de transposasa (*sfaT2* y *sfaT3*). También hay 1 gen que codifica para factor regulador de la familia de TetR (*orf23*), 1 gen que codifica para proteína de membrana acoplada a TetR (*orf24*), 1 gen de FAD monooxigenasa (*orf25*), y 1 gen de endonucleasa incompleto (*orf26*).

Un análisis preliminar de la secuencia de la agrupación de genes ayudó a los inventores a determinar los supuestos límites de la agrupación de genes de biosíntesis de SFA. El límite izquierdo está ubicado justo en el sentido de 5' de *sfaA*, incluyendo 2 genes de función desconocida (*sfaU1* y *sfaU2*). El límite derecho está ubicado justo en el sentido de 3' del gen de MbtH, incluyendo 3 genes de función desconocida y 2 genes de transposasa (*sfaV1-3*, *orf21* y *22*). Dado que las proteínas de la familia de MbtH con frecuencia están asociadas con NRPS, los inventores suponen que está implicada en la regulación de biosíntesis de precursores. Sin embargo, la actividad de proteínas de esta familia puede complementarse mediante homólogos en otra parte en la agrupación; es decir, se necesitaría desactivar todos los genes de MbtH presentes en todo el genoma, de lo contrario se complementarían mediante su

homólogo ubicado en otro lugar en el genoma. Como se desconocía la cantidad total de MbtH en las bacterias huésped, se eligieron la proteína desconocida en el sentido de 3' y el transposón para su análisis. Además, a 82 pb desde el inicio de *orf10*, hay una secuencia de promotor de 35 pb; y a 52 pb desde donde termina *orf18*, hay una parte de una secuencia de terminador independiente de p. Toda la agrupación de genes está ubicada en la unidad transcripcional grande tal como se describió anteriormente. Basándose en el siguiente análisis de las transposasas en el sentido de 3', se supone que esta agrupación de genes procede de una transferencia de genes horizontal.

3.1 Estudio del límite izquierdo

El estudio del límite izquierdo se centró en *sfaU1* y *sfaU2*. *sfaU1* tiene una longitud completa de 453 pb, codifica para una proteína de función desconocida de 150 residuos de aminoácido, y comparte el 60% de homología con una proteína desconocida en *Streptomyces sviveus*; *sfaU2* tiene una longitud completa de 468 pb, codifica para una proteína de función desconocida de 155 residuos de aminoácido, que, tras análisis, se observó que compartía el 81% de homología con una resolvasa/integrasa en *Streptomyces turgidiscabies*. Se llevó a cabo una desactivación del gen *sfaU1* y se confirmó que no era relevante para la biosíntesis de SFA. Mientras tanto, las secuencias de transposón en el sentido de 5' y de 3' de *sfaU1* también implican los límites de la transferencia horizontal de la agrupación de genes.

3.2 Estudio del límite derecho

El estudio del límite derecho se centró en *sfaV1*, *sfaV2*, *sfaV3* y *sfaT2*. *sfaV1* tiene una longitud completa de 537 pb, codifica para proteína desconocida de 178 residuos de aminoácido, y comparte el 81% de homología con una proteína desconocida en *Streptomyces cattleya*; *sfaV2* tiene una longitud completa de 459 pb, codifica para una proteína de función desconocida de 152 residuos de aminoácido, y comparte el 46% de homología con una proteína de función desconocida en *Streptomyces coelicolor*, *sfaV3* tiene una longitud completa de 435 pb, codifica para una proteína desconocida de 144 residuos de aminoácido, comparte el 45% de homología con una proteína desconocida en *S. coelicolor*, y una determinada homología con UDP-N-acetilglucosamina transferasa en *Pseudomonas syringae* pv. *tomato*. *SfaT2* tiene una longitud completa de 534 pb, codifica para una transposasa de 177 residuos de aminoácido, y es parcialmente homóloga a la secuencia de inserción de IS-4 en *Burkholderia vietnamiensis*. Basándose en el análisis de la secuencia de transposón, se determinó que el límite en el sentido de 3' estaba ubicado adyacente a la región de transposón.

Ejemplo 4: Propuesta para la biosíntesis de la unidad iniciadora de SFA

Se reduce acetoacetil-CoA endógena para dar hidroxibutilil-CoA mediante las deshidrogenasa/reductasas de cadena corta codificadas por *sfaM*, después para dar crotonil-coA mediante una deshidratasa metabólica primaria desconocida. Después, se somete a carboxilación reductora la crotonil-coA mediante crotonil-coA reductasa codificada por *sfaR*, y se amida mediante análogo de asparagina sintetasa codificado por *sfaP*. Después de eso, se completa la extensión por incorporación de la unidad de dos carbonos posterior mediante la autoacilación de la proteína transportadora de acilo iniciadora (codificada por *sfaE*-ACP1). Mientras tanto, se piensa que una aciltransferasa libre que tiene una función modificada y codificada por *sfaL* es responsable de hidrolizar otras unidades que se añaden de manera errónea mediante autoacilación.

Ejemplo 5: Propuesta para la biosíntesis de la *m*-tirosina en SFA

Todos los sistemas de hidroxilación de fenilalanina/tirosina/triptófano tradicionales se basan en una reacción de sustitución electrófila para realizar la hidroxilación en las regiones ricas en electrones de sistemas aromáticos. Por el contrario, la meta-hidroxilación de fenilalanina se produce en la posición meta relativamente pobre en electrones de la fenilalanina. Un análisis bioinformático del segundo dominio de adenilación (A2) en la NRPS muestra que su región de huella activa no comparte ninguna homología con las regiones de huella de los dominios de adenilación que son responsables de reconocer la fenilalanina en otros sistemas, y no tiene ninguna asociación aparente con la tirosina. Se expresó la supuesta meta-hidroxilasa, *SfaA*, en *E. coli* y se aisló usando métodos convencionales. Un estudio *in vitro* de esta proteína en un ensayo de AMP PPI (para los métodos, véase Garneau *et al.*, 2005) mostró que no tenía ninguna actividad sobre la fenilalanina libre, pero sí mostró actividad frente a un derivado de tioéster de acetilcisteamina (SNAC). Esto sugiere que la fenilalanina es el sustrato óptimo para la integración en la estructura principal de polipéptido, es decir tras la integración, tiene lugar modificación *in situ* para obtener la unidad de meta-tirosina.

Ejemplo 6: Propuesta para la biosíntesis de la unidad de ácido piperáxico de SFA

El tercer elemento estructural del resto peptídico corto en la estructura de SFA es una unidad de ácido piperáxico, que está implicada en la integración de la estructura principal peptídica de una manera 2,3-regioselectiva. Inicialmente se supuso que la unidad de ácido piperáxico se obtenía a partir de un precursor de prolina mediante deshidrogenación en la posición 1,2, amonólisis, oxidación en N⁵ para formar un ciclo, y después reducción. Sin embargo, ensayos marcados anteriores han mostrado que el precursor real de unidad de ácido piperáxico es el ácido glutámico (Umezawa *et al.*, 2001, Miller *et al.*, 2007). Por tanto, se sugiere que la unidad de ácido piperáxico

se forma a partir de una molécula precursora de ácido glutámico/glutamina mediante hidrogenación y deshidratación, oxidación en N⁵ y después hidrogenación. Aunque no se ha obtenido ninguna evidencia directa de que la ornitina sea el precursor en otros sistemas de ensayo marcado, los inventores consideran que es posible que la ornitina esté implicada en la formación de ácido piperázico en el sistema mediante oxidación en N⁵ directa para formar el ciclo, tal como se observa en otros sistemas (Fujimori *et al.*, 2007).

Ejemplo 7: Propuesta para la biosíntesis de la unidad de extensión poco habitual

Se piensa que la policétido sintasa iterativa codificada por *sfaK* está implicada en la biosíntesis de la unidad de extensión poco habitual de 6 carbonos. Se usa acetil-CoA o malonil-CoA como unidad iniciadora para pasar a través de dos rondas de extensión de dos carbonos, y se obtiene triceto-hexanoil-CoA. Después se reduce selectivamente el grupo ceto en la parte central mediante el dominio de oxidación-reducción poco habitual en el extremo terminal de dicha enzima, y se obtiene un derivado de CoA parcialmente reducido con un doble enlace conjugado. Después, mediante carboxilación reductora por la crotonil-coA reductasa codificada por *sfaR*, se activa para incorporarse en el proceso de extensión de cadena de policétido. En la identificación del producto de fermentación derivado del mutante desactivado de *sfaK*, los inventores detectaron una señal de m/z 1033,4 (figura 6C). Esto respalda la supuesta generación de un producto de biosíntesis compensatorio, es decir la cadena lateral de unidad de extensión especial se sustituye por el sustrato metilmalonil-CoA que no es óptimo en cuanto al dominio correspondiente para transacilación en PKS, de modo que se obtiene un producto de sustitución con una cadena lateral de metilo.

Ejemplo 8: Generación de análogos no naturales de SFA mediante medios genéticos

Se inactivaron los segmentos correspondientes a diversos dominios DH en la agrupación de genes de biosíntesis de SFA mediante mutagénesis dirigida al sitio, de modo que se codifica para una deshidratasa que carece de residuos de sitio activo. Se usó esta deshidratasa para actuar sobre un producto intermedio que había pasado por el reconocimiento de AT en el sentido de 5', transferencia sobre ACP, condensación catalizada por KS y reducción mediante KR. Dado que dicha deshidratasa no puede ejercer su función de deshidratación para formar un doble enlace, se obtiene por tanto el análogo de SFA con un grupo hidroxilo. El dominio deshidratasa (DH) es un dominio enzimático que cataliza la deshidratación del grupo hidroxilo derivado de la reducción de ceto, para formar un doble enlace. El dominio funcional tiene un motivo altamente conservado LXXHXXXGXXXXP, en el que el residuo de histidina es el centro activo catalítico; la leucina y la glicina mantienen un plegamiento normal; la prolina algunas veces no se conserva y se desconoce su función. Por tanto, se eligió el residuo de histidina como diana para la mutación. Se eligieron DH3/DH4/DH5/DH6 (los dominios DH de los módulos 7, 8, 10 y 11) en *SfaA4* (*sfaH*) (responsables respectivamente de la formación de los dos conjuntos de dos dobles enlaces dentro y fuera del grupo éster del macróido) como dominios funcionales diana para inactivarse mediante mutagénesis dirigida al sitio. Se usó la reacción en cadena de la polimerasa (PCR) para amplificar 8 fragmentos de ADN de aproximadamente 1 kb que flanqueaban los motivos de los dominios funcionales mencionados anteriormente, y se usaron estos fragmentos como brazos homólogos. Se cambió el residuo de histidina por un residuo de aminoácido apolar tal como glicina/alanina o similar, para eliminar la función de deshidratación. Se introdujeron codones alterados mediante cebadores, y se ligaron dos fragmentos que portaban la mutación (pueden introducirse sitios de restricción requeridos para tal ligamiento mediante mutación silenciosa). De esta manera, se construyeron plásmidos recombinantes homólogos necesarios para la mutagénesis dirigida al sitio. Se fermentaron los mutantes correspondientes. La cromatografía de líquidos/espectrometría de masas detectó picos de señal M+18 (CL-ESI-EM m/z 1109,3, 1131,5), mostrando que se produjeron los compuestos deseados.

Para facilitar la comprensión de la presente invención, a continuación se proporcionan adicionalmente ejemplos únicamente con fines ilustrativos y no limitativos del alcance de la invención.

Ejemplo 9: Construcción de un sistema de transferencia de genes para la cepa productora de SFA *Streptomyces* sp A92-309110 (*S. flaveolus*)

En primer lugar se transformó el plásmido diana que iba a usarse en transferencia conjugativa en *E. coli* S17-1. Se escogió una única colonia de una placa y se inoculó en tubos con 3 ml de medio LB y antibiótico correspondiente para presión selectiva, y se incubó durante la noche a 37°C. Se recogió el cultivo al día siguiente, se pipeteó 1 ml de cultivo y se inoculó en un matraz de agitación de 250 ml que contenía 50 ml de medio LB y antibiótico correspondiente para presión selectiva. Se realizó la incubación a 37°C en un agitador hasta que el valor de DO₆₀₀ alcanzó aproximadamente 0,5. Se transfirió el cultivo a tubos de EP de 50 ml, y se centrifugó a 3800 rpm durante 10 min a 16°C. Se recuperaron las bacterias, y se lavaron dos veces con 20 ml de medio LB, después se centrifugaron en las mismas condiciones. Después se recuperaron las bacterias y se resuspendieron en 1 ml de LB antes de su uso.

Se tomó un tubo de suspensión de esporas crioconservada y se centrifugó a 12000 rpm durante 3 min a temperatura ambiente. Tras retirarse el sobrenadante, se lavó dos veces con 1 ml de tampón TES, y se resuspendió en 500 µl de tampón TES. Se realizó choque térmico en un baño de agua a 50°C durante 10 min, y después se añadieron 500 µl de medio TSB. Se mezcló bien la disolución, se incubó a 37°C durante 4-5 h, y después se centrifugó para retirar totalmente el sobrenadante. Después se resuspendió el sedimento con 1,5 ml de LB antes de su uso.

Se extendieron las bacterias receptoras y las bacterias donadoras (cada una 100 μ l) sobre placas con medio MS (que contenía $MgCl_2$ 10 mM), y se cultivaron a 30°C durante 12-16 h. Además, se usó una placa, en la que sólo se extendieron bacterias receptoras, como control negativo. Tras cultivar, se añadieron 4-5 ml de ddH₂O estéril a cada placa, y se rascó suavemente la superficie con una rasqueta. Tras absorberse el agua, se secó la placa durante 1 h, después se recubrió de nuevo con 1 ml de ddH₂O estéril (que contenía Am 12,5 μ g/ml y NA 50 μ g/ml) y se cultivó a 30°C durante 3-5 d. Se escogió una única colonia que creció bien que se transformó satisfactoriamente y se inoculó en 3 ml de medio TSB, que contenía selección con antibiótico, si era necesario. Se agitó el cultivo a 30°C. Tras 1 d, se tomó aproximadamente 1 μ l del cultivo y se extendió sobre una placa de 150 mm que contenía AmR ISP-4, después se cultivó durante 2 d a 37°C. Se recubrió directamente la placa con medio de placa sólido que contenía que kanamicina ISP-4, y se cultivó durante 6-7 h a 37°C. Después se separaron los 2 fragmentos de medio de cultivo, y se cultivaron durante otros 2 d a 37°C respectivamente. Se escogieron colonias individuales que habían crecido con AmR pero no en la posición correspondiente sobre placas de kanamicina y se inocularon en 3 ml de medio TSB. Se agitó el cultivo y estaba listo para fermentación.

Ejemplo 10: Construcción de una biblioteca génica de la cepa productora de SFA *Streptomyces sp* A92-309110 (*S. flaveolus*)

10.1 Ensayo de digestión enzimática a pequeña escala

En primer lugar, se determinó la cantidad de Sau3AI a usarse mediante una serie de experimentos de dilución. Se preparó en primer lugar un sistema de reacción de 250 μ l (que contenía 40 μ l de ADN genómico (por ejemplo, gDNAQXD01-82-1), 2,5 μ l de BSA (100x)), entonces se dividió el sistema de reacción de 250 μ l en 1 x 50 μ l y 7 x 25 μ l, y se mantuvo sobre hielo. Entonces, se añadieron 2 μ l de Sau3AI, que se había diluido previamente hasta 0,5 u/ μ l, al tubo n.º 1 que contenía el sistema de reacción de 50 μ l, y se mezcló bien. Luego se transfirieron 25 μ l a los 25 μ l en el tubo n.º 2. Se repitieron estas etapas de transferencia 7 veces. Se incubaron todos los sistemas en un baño de agua a 37°C durante 15 min, luego se inactivaron a 70°C durante 10 min. Se realizó una electroforesis en gel de agarosa al 0,4% a 4°C en una sala fría, luego se tiñó con bromuro de etidio. Se examinó la calidad de la digestión enzimática bajo luz UV o usando un sistema de imágenes en gel.

10.2 Ensayo de digestión enzimática a gran escala

Según las condiciones de reacción determinadas en experimentos preliminares, se usaron 4 veces la cantidad del ADN total y la enzima a la concentración correspondiente para preparar fragmentos de ADN necesarios para construir la biblioteca: se mezcló bien el sistema sobre hielo, se dividió uniformemente en 5 alícuotas, y se incubó en un baño de agua a 37°C. A los 12, 14, 16, 18 y 20 min respectivamente, se tomaron alícuotas y se inactivaron a 70°C durante 10 min. Se realizó una electroforesis en gel de agarosa al 0,4% a 4°C en la sala fría, luego se tiñó con bromuro de etidio. Se examinó la calidad de la digestión enzimática bajo luz UV o usando un sistema de imágenes en gel.

10.3 Desfosforilación

Se extrajeron secuencialmente fragmentos de ADN digeridos de tamaño adecuado con disolución saturada de fenol, y cloroformo-alcohol isoamílico. Se añadieron 0,1 volúmenes de NaAc 3 M y 3 volúmenes de alcohol anhidro para precipitar el ADN. Entonces se lavó el sedimento de ADN con alcohol al 70%, se secó, y luego se resuspendió en 200 μ l de agua. Se retiraron 10 μ l como control, y se desfosforiló el ADN restante. Se añadieron 10 μ l de SAP (Promega, 1 unidad/ μ l) y 25 μ l de tampón 10x a 190 μ l de ADN digerido con Sau3AI, y luego se añadió agua hasta 250 μ l. Se incubó el sistema durante 1 h a 37°C, se complementó con 7 μ l de SAP, y luego se mezcló bien. Tras incubación durante 1 h, se realizó desnaturalización térmica a 70°C para inactivar la enzima. Se realizó extracción secuencialmente con disolución saturada de fenol, cloroformo:alcohol isoamílico. Se añadieron 0,1 volúmenes de NaAc 3 M y 3 volúmenes de alcohol anhidro para precipitar el ADN. Entonces se lavó el ADN con alcohol al 70%, se secó, se resuspendió en 15 μ l de disolución de TE. Entonces, se examinó para observar si la desfosforilación era completa. Se realizó una electroforesis en gel de agarosa al 0,3% a 4°C en una sala fría, y se tiñó el gel con bromuro de etidio. Se examinó la calidad de la digestión enzimática bajo luz UV o usando un sistema de imágenes en gel.

10.4 Preparación de vector SuperCos1

Se usó la enzima individual XbaI para linealizar SuperCos1 entre dos sitios *cos*, y entonces se realizó desfosforilación (para impedir el autoligamiento). Luego, para insertar fragmentos parcialmente digeridos de ADN del genoma, se realizó digestión con BamHI para obtener dos brazos de 1 kb y 7 kb. El vector preparado tenía la concentración de 1 μ g/ μ l.

10.5 Ligamiento de fragmentos derivados de ADN total en vector SuperCos1

Se mezclaron 6 μ l de ADN desfosforilado (aproximadamente 2 μ g/ μ l) y 4 μ l de SuperCos1 preparado

(aproximadamente 1 µg/µl), y se tomaron 1,2 µl como control. A los 8,8 µl restantes de la disolución de mezcla se les añadieron 1,2 µl de ADN ligasa de T4 (NEB 400 u/µl) y 1,2 µl de tampón, luego se mezcló bien. Se llevó a cabo ligamiento durante 16 h a 16°C. Se realizó una electroforesis en gel de agarosa a 4°C en una sala fría. Se tiñó el gel con bromuro de etidio, y se examinó bajo luz UV o usando un sistema de imágenes en gel.

5

10.6 Empaquetamiento de bibliotecas

Se retiró un reactivo de empaquetamiento Stratagene Gig Pack III XL del congelador a -80°C, y se descongeló rápidamente a mano. Se añadieron 4 µl de producto de ligamiento, y se pipeteó suavemente 3 veces. Se incubó la mezcla en un baño de agua a 22°C durante 2 h, luego se retiró y se añadieron 500 µl de tampón SM, después se invirtió varias veces. Se añadieron 50 µl de cloroformo y invirtió la mezcla varias veces (ahora debe ser visible cierta precipitación similar a proteína). Entonces se centrifugó la mezcla durante unos pocos segundos. Se transfirió el sobrenadante mediante pipeteo, y se almacenó a 4°C.

10

15 10.7 Valoración

En primer lugar, se escogió una colonia bacteriana individual de *E. coli* VCS257 de una placa y se inoculó en LB para el cultivo durante la noche. Se tomaron 500 µl de cultivo y se inocularon en 50 ml de LB (MgSO₄ 10 mM, maltosa al 0,2%), y se cultivaron a 37°C con agitación. Cuando la DO₆₀₀ de *E. coli* alcanzó 0,84, se añadieron 5 µl de disolución de empaquetamiento a 45 µl de tampón SM, luego se añadieron a 50 µl de *E. coli* VCS257. Se dieron golpecitos suaves al tubo para que se mezclase, luego se puso en un baño de agua a 22°C para incubar durante 30 min. Se añadieron 400 µl de LB, se invirtió el tubo varias veces, luego se incubó en un baño de agua a 37°C durante 75 min (y se invirtió varias veces cada 15 min). Se extendieron 250 µl sobre cada placa de LB (Amp 100 µg/ml), y se incubaron durante la noche a 37°C.

20

25

10.8 Confirmación de la autenticidad de la biblioteca

Para demostrar que las colonias bacterianas hechas crecer no eran falsos positivos sino que realmente contenían cósmidos recombinantes, se escogieron al azar 10 colonias, se inocularon en LB (Amp 100 µg/ml) y se cultivaron. Se extrajeron los cósmidos según el método de lisis alcalina usado para minipreparaciones de ADN de plásmido de *E. coli*, luego se digirieron con BamHI y se hicieron pasar sobre un gel de agarosa al 0,5%.

30

10.9 Amplificación de bibliotecas

Se inoculó *E. coli* VCS257 en 50 ml de LB (que contenía MgSO₄ 10 mM, maltosa al 0,2%) y se cultivó hasta DO₆₀₀ = 0,84, luego se realizó inmediatamente transfección. Se mezclaron suavemente la disolución de cultivo celular (100 µl) y de empaquetamiento (100 µl), se incubaron en un baño de agua a 22°C durante 30 min, luego se añadieron 800 µl de LB a cada uno de 5 tubos de microcentrífuga, se mezclaron suavemente, y se incubaron en un baño de agua a 37°C durante 75 min con mezclado suave cada 15 min. Mientras se realizaba la incubación, se colocaron 5 placas grandes (concentración de Amp de 100 µg/ml) a 37°C en un incubador para precalentarlas. Tras acabarse la incubación en el baño de agua, se extendió sobre las placas el contenido de un tubo de microcentrífuga para cada placa, y se incubaron durante la noche a 37°C. Tras cultivar durante 18 h (hasta que las colonias bacterianas hubieron crecido bien), se añadieron 3~4 ml de LB a cada placa. Se rascaron las colonias bacterianas con una rasqueta, y se pipeteó la disolución de cultivo y se transfirió a un tubo de centrífuga de 50 ml. Se combinaron las disoluciones de cultivo de las 5 placas grandes, y se añadieron ampicilina y glicerol estéril hasta concentraciones finales de 50 µg/ml y el 18% (V/V). Se almacenaron alícuotas de 250 µl/tubo a -80°C.

35

40

45

Ejemplo 11: Generación de una cepa *Streptomyces sp* A92-309110 (*S. flaveolus*) modificada por ingeniería que da como resultado una sustitución génica del gen *sfaK* mediante direccionamiento por PCR.

50

La generación de una cepa *Streptomyces sp* A92-309110 (*S. flaveolus*) modificada por ingeniería con una delección en marco en el gen *sfaK* de la agrupación de biosíntesis de sangliferina mediante direccionamiento por PCR implica sustituir la mayor parte de la región codificante de *sfaK* por un marcador de resistencia a apramicina y el oriT usando el procedimiento convencional de direccionamiento por PCR (Gust *et al* 2002).

55

Se diseñaron los cebadores 201-1L (SEQ ID NO: 30) y 201-1R (SEQ ID NO: 31) según el procedimiento de direccionamiento por PCR para amplificar el marcador de resistencia a apramicina junto con el oriT a partir del molde pIJ773 (n.º de registro GenBank AX657066,1). La región 5' de cada cebador (no subrayada) es idéntica a *sfaK* de manera que la sustitución de la secuencia de *sfaK* entre las regiones de unión de los oligos con el casete de sustitución de pIJ773, que contiene FRT, *oriT*, *aac(3)IV* y FRT eliminará la actividad de SfaK. La región 3' subrayada de cada oligo es idéntica a una secuencia en pIJ773 para la amplificación del casete de sustitución. Dentro del casete de sustitución, FRT es la secuencia diana de reconocimiento de FLP-recombinasa y las dos secuencias de FRT flanquean el origen de transferencia desde RK2 (*oriT*) y el gen de resistencia a apramicina (*aac(3)IV*), un experto en la técnica estará familiarizado con esta técnica y se describe en detalle en Gust *et al* 2002 y se respalda además por referencias en ese documento.

60

65

201-1L (SEQ ID NO: 30):

5'-CTCGACCGGTACTGGGCCAACGTGGTGGCCGGTGTTCGACATTCCGGGGATCCGTC
GACC-3'

5

201-1R (SEQ ID NO: 31):

5'-GGCCAGTTCGCGCAGGAAGGCCCGTACGCCGTCGTCCGGTGTAGGCTGGAGCTGC
TTC-3'

10 Se logró la amplificación del casete de ADN de direccionamiento por PCR de ~1,4 kb usando los cebadores 201-1L (SEQ ID NO: 26) y 201-1R (SEQ ID NO: 27), pIJ773 como molde y la polimerasa Primer Star (Takara Co. Ltd.) usando condiciones convencionales con una temperatura de hibridación de 50°C. El cósmido pTL3102 cubre una región apropiada de la agrupación de sangliferina y se transformó en primer lugar en *E. coli* BW25113/pIJ790 para dar la cepa *E. coli* BW25113/pIJ790/3102. Se mantiene *E. coli* BW25113/pIJ790 a 30°C en LB (medio de Luria-Bertani; Sambrook *et al.*, 1998) que contiene cloranfenicol (25 µg/ml). Para preparar el constructo de sustitución génica pTL3122, se llevaron a cabo las etapas expuestas a continuación. Se realizó direccionamiento por PCR transformando el casete de ADN purificado en gel en *E. coli* BW25113/3102 mediante electroporación, tras la electroporación se incubaron las bacterias en LB a 37 grados centígrados durante 1 h para inducir recombinación y expresar resistencia a antibióticos, luego se sembraron en placa sobre agar LB que contenía apramicina 100 µg/ml para incubación durante la noche a la misma temperatura. Se inocularon colonias resistentes a apramicina en LB complementado con apramicina 100 µg/ml y se incubaron durante la noche a 37°C. Se aisló el ADN de cósmido, y se confirmó la presencia del casete de resistencia mediante digestión con enzimas de restricción y análisis de la secuencia de ADN, el cósmido generado se nombró pTL3111. Para facilitar la transformación en *Streptomyces*, se redujo el tamaño del cósmido cortando pTL3111 con *Bg*/II. Se subclonó el fragmento de *Bg*/II de 13 kb que contenía el casete de resistencia que contenía *acc(3)IV* y *sfaK* inactivado en pKC5201 para dar el constructo de sustitución final pTK3122. Se derivó pKC5201 de pKC1139 sustituyendo el *acc(3)IV* por el gen de resistencia a neomicina de supercos1. Para generar el mutante de sustitución de *sfaK* en *Streptomyces sp* A92-309110 (*S. flaveolus*), se transformó pTL3122 en *E. coli* S17-1 mediante electroporación para generar la cepa donadora de *E. coli* para conjugación. Se transformó *Streptomyces sp* A92-309110 (*S. flaveolus*) mediante conjugación con *E. coli* S17-1/pTL3122 tal como se describió en los métodos generales anteriormente. Se homogeneizaron ex-conjugantes resistentes a apramicina y se sembraron en línea sobre el agar ISP-4 complementado con apramicina 50 µg/ml y se cultivaron a 37°C hasta que se lograron colonias que crecieron bien. Se usó una serie de clones individuales para inocular 3 ml de caldo TSB complementado con 30 µg/ml de apramicina agitando a 250 rpm, 37°C durante 4 días. Se extendieron 100 µl de cada cultivo sobre agar ISP-4 y se incubaron a 37°C durante 2-3 días para inducir el segundo acontecimiento de cruzamiento. Se sembraron en placa colonias individuales +/- neomicina para determinar la pérdida de la estructura principal del plásmido. La sensibilidad a neomicina indicó pérdida de la secuencia del plásmido y se analizaron colonias con ese fenotipo mediante PCR para confirmar que se había logrado la sustitución de *sfaK* por el casete de ADN. Una de tales colonias se denominó *Streptomyces flaveolus* mv. xuwengensis, con el nombre alternativo de *Streptomyces sp.* TL3011. Entonces se cultivaron cepas según los métodos descritos, y se observó que las cepas producían sangliferina X (véase la figura 6C).

45 Ejemplo 12: Generación de cepas *Streptomyces sp* A92-309110 (*S. flaveolus*) modificadas por ingeniería que dan como resultado sustituciones génicas de una serie de genes de biosíntesis de sangliferina mediante direccionamiento por PCR.

Usando métodos similares a los descritos en el ejemplo 11 anterior, se construyeron los mutantes de sustitución de *sfaA*, *sfaB*, *sfaC*, *sfaJ*, *sfaL*, *sfaM*, *sfaN* y *sfaP*. En cada caso, se sustituyó la mayoría de la región codificante del gen diana por un marcador de resistencia a apramicina y el oriT usando el procedimiento convencional de direccionamiento por PCR (Gust *et al* 2002). Se diseñó un par de cebadores para cada gen diana (tabla 2) según el procedimiento de direccionamiento por PCR para amplificar el marcador de resistencia a apramicina junto con el oriT a partir del molde pIJ773 (n.º de registro GenBank AX657066,1). La región 5' de cada cebador (no subrayada) es idéntica al gen diana de manera que la sustitución de la secuencia génica entre las regiones de unión de los oligos con el casete de sustitución desde pIJ773, que contiene FRT, *oriT*, *aac(3)IV* y FRT eliminará la actividad codificada por el gen. La región 3' subrayada de cada oligo es idéntica a una secuencia en pIJ773 para la amplificación del casete de sustitución. Dentro del casete de sustitución, FRT es la secuencia diana de reconocimiento de FLP-recombinasa y las dos secuencias de FRT flanquean el origen de transferencia desde RK2 (*oriT*) y el gen de resistencia a apramicina (*aac(3)IV*), un experto en la técnica estará familiarizado con esta técnica y se describe en detalle en Gust *et al* 2002 y se respalda además con las referencias en ese documento.

60 Se logró la amplificación del casete de ADN de direccionamiento por PCR de ~1,4 kb usando los cebadores

mostrados en la tabla 2, pLJ773 como molde y la polimerasa Primer Star (Takara Co. Ltd.) usando condiciones convencionales.

Tabla 2: Cebadores usados para direccionamiento por PCR

5

Cebadores		Secuencias de cebador	Gen diana
201-2L	SEQ ID NO: 32	<u>GTGGAAATCGGCTCGGGCGCGCCC</u> GAATTAACCGCGT <u>CGATTCCGGGGA</u> TCCGTCGACC	<i>sfaA</i>
201-2R	SEQ ID NO: 33	AATGGATGTATCGTCGCAGGACGCC CAGAATTCACCTGCT <u>GTAGGCTGGA</u> GCTGCTTC	
201-3L	SEQ ID NO: 34	<u>GCGCAGCAGAGCCCGGAATCAGAA</u> GTA <u>CTGGACGTCACCATTCCGGGGA</u> TCCGTCGACC	<i>sfaB</i>
201-3R	SEQ ID NO: 35	GGCGATCTCGCCCGCGCGGACCGC CACCATGGACAGCAGT <u>GTAGGCTGG</u> AGCTGCTTC	
201-4L	SEQ ID NO: 36	<u>GAGGATTGCGACGGCGTCGTCCTG</u> GCGTTTCTGCGACAC <u>ATTCCGGGGA</u> TCCGTCGACC	<i>sfaC</i>
201-4R	SEQ ID NO: 37	CTCCTCGTCGGCTTCGGTGAGTCGG CGGTCGCGCATCACT <u>GTAGGCTGGA</u> GCTGCTTC	
201-5L	SEQ ID NO: 38	<u>GGGCCGCAGGACAGGTCCGGCGGC</u> CCGGTGC GCGGCGAGATTCCGGGG ATCCGTCGACC	<i>sfaJ</i>
201-5R	SEQ ID NO: 39	CGCGCCGGAGAACAGCGGGAAGTA GGTGTGAGGTCGTCT <u>GTAGGCTGG</u> AGCTGCTTC	
201-6L	SEQ ID NO: 40	<u>GGCCGGCCCCGGAGCCCTGGGCGC</u> CGCCCTGCGTTCGGGATTCCGGGGA TCCGTCGACC	<i>sfaL</i>
201-6R	SEQ ID NO: 41	GCTCTCCACGGAGGCGCTCACCGC GGCGACGGCGGCCTCT <u>GTAGGCTG</u> GAGCTGCTTC	
201-7L	SEQ ID NO: 42	<u>GCGGTGCGTACCGGATCGTCCCGC</u> GGCATCGGCGCGGCCATTCCGGGG ATCCGTCGACC	<i>sfaM</i>
201-7R	SEQ ID NO: 43	CACGCCGTGCGCGGTCCAGCCGCC GTCGAAGCGCAGGGTT <u>GTAGGCTG</u> GAGCTGCTTC	
201-8L	SEQ ID NO: 44	<u>CTTCGGCGTCTCGCGCTCGCCAC</u> GCCCTCGGCGATCCATTCCGGGGAT CCGTCGACC	<i>sfaN</i>
201-8R	SEQ ID NO: 45	GTGCATGCCGATGGACAGGCCCGC GAGCGCGACCACGTCT <u>GTAGGCTG</u> GAGCTGCTTC	
201-9L	SEQ ID NO: 46	<u>GACGACCTCGCGCGGCACCGGTCC</u> GTCGTCCAGGCGATGATTCCGGGGA TCCGTCGACC	<i>sfaP</i>
201-9R	SEQ ID NO: 47	CATCTCGATGCCGCCCGGTCGTGC GGCAGGCTGAAGTCT <u>GTAGGCTGGA</u> GCTGCTTC	

Para preparar los constructos de sustitución génica, se ejecuta direccionamiento por PCR transformando el

fragmento de casete purificado en gel en *E. coli* Bw25113/pLJ790/3106 (para *staA*, *sfaB* y *sfaC* tal como se muestra en la tabla 3) o *E. coli* Bw25113/3102 (para *sfaJ*, *sfaL*, *sfaM*, *sfaN* y *sfaP* tal como se muestra en la tabla 3) mediante electroporación. Se preparan las cepas por adelantado transformando el cósmido pTL3106 o pTL3102 en *E. coli* Bw25113/pLJ709 mediante electroporación. Tras la transformación, se incubaron las bacterias en LB a 37°C durante 1 h para inducir recombinación y expresar resistencia a antibióticos, y luego se sembraron en placa sobre agar LB que contenía 100 µg/ml de ampicilina para incubación durante la noche a 37°C. Se usaron colonias resistentes a ampicilina para inocular LB complementado con ampicilina 100 µg/ml y se incubaron durante la noche a 37°C. Se aíslan los cósmidos recombinantes mediante un kit de aislamiento de plásmidos (Dingguo Co. Ltd.) y se confirmaron mediante digestión con enzimas de restricción y secuencia de ADN y se les dieron los nombres mostrados en la tabla 3. Debido a la baja eficiencia de transformación para ADN de cósmido más grande en *Streptomyces sp.* receptor, la transformación directa de los cósmidos recombinantes en *Streptomyces* mediante conjugación sólo fue satisfactoria para pTL3113 y pTL3114 que contenían los casetes para la inactivación de *sfaB* y *sfaC* y dieron como resultado las cepas mutantes TL3003 y TL3004, respectivamente. Para el resto, se acortó el tamaño de los insertos cortando los cósmidos recombinantes con enzimas de restricción tal como se muestra en la tabla 3. Los fragmentos de ADN resultantes se subclonaron en pKC5201 que se deriva de pKC1139 sustituyendo el *acc(3)IV* por el gen de resistencia a neomicina de supercos1 para dar los plásmidos enumerados en la tabla 3. Estos plásmidos dieron como resultado la incorporación de las mutaciones deseadas y a las cepas mutantes se les dieron los nombres designados mostrados en la tabla 3; de manera que la cepa *Streptomyces sp.* TL3002 tiene *sfaA* sustituido por el casete de resistencia que contiene *acc(3)IV*, *Streptomyces sp.* TL3005 tiene *sfaJ* sustituido por el casete de resistencia que contiene *acc(3)IV*, *Streptomyces sp.* TL3006 tiene *sfaL* sustituido por el casete de resistencia que contiene *acc(3)IV*, *Streptomyces sp.* TL3007 tiene *sfaM* sustituido por el casete de resistencia que contiene *acc(3)IV*, *Streptomyces sp.* TL3008 tiene *sfaN* sustituido por el casete de resistencia que contiene *acc(3)IV* y *Streptomyces sp.* TL3009 tiene *sfaP* sustituido por el casete de resistencia que contiene *acc(3)IV*.

Tabla 3: Los plásmidos y cósmidos recombinantes

Mutante generado	Gen diana	Cósmido	Cósmidos recombinantes	Enzima para acortar el tamaño	Tamaño de fragmentos tras digestión (kb)	Nombre de plásmido
TL3002	<i>sfaA</i>	pTL3106	pTL3112	<i>Bg/II</i>	8	pTL3123
TL3003	<i>sfaB</i>	pTL3106	pTL3113	N.A.	N.A.	N.A.
TL3004	<i>sfaC</i>	pTL3106	pTL3114	N.A.	N.A.	N.A.
TL3005	<i>sfaJ</i>	pTL3102	pTL3115	<i>Bg/III/EcoRI</i>	11	pTL3124
TL3006	<i>sfaL</i>	pTL3102	pTL3116	<i>EcoRI</i>	10	pTL3125
TL3007	<i>sfaM</i>	pTL3102	pTL3117	<i>EcoRI</i>	11	pTL3126
TL3008	<i>sfaN</i>	pTL3102	pTL3118	<i>EcoRI</i>	10,5	pTL3127
TL3009	<i>sfaP</i>	pTL3102	pTL3119	<i>EcoRI</i>	9,5	pTL3128

En primer lugar se transforman plásmidos o cósmidos para la transformación de *Streptomyces* en *E. coli* s17-1 y luego se conjugan en *Streptomyces* tal como se describió en el ejemplo 9 anteriormente. Se homogeneizan los conjugantes resistentes a ampicilina y se siembran en línea sobre medios de agar ISP-4 complementados con ampicilina 50 µg/ml y se cultivan a 37°C hasta que se lograron colonias que crecieron bien. Se usó una serie de clones individuales de cada una para inocular 3 ml de caldo TSB complementado con ampicilina 30 µg/ml agitando a 250 rpm a 37°C durante cuatro días. Se extendieron 100 µl de cultivo sobre agar ISP-4 a 37°C durante 2-3 días para inducir el segundo acontecimiento de cruzamiento. Se sembraron en placa colonias individuales con y sin neomicina para determinar la pérdida de la estructura principal del plásmido. La sensibilidad a neomicina indicó pérdida de la secuencia del plásmido y se analizaron colonias con ese fenotipo mediante PCR para confirmar que se había logrado la sustitución del gen diana por el casete de ADN. A las cepas mutantes finales se les dieron los nombres de cepa indicados en la tabla 3.

Ejemplo 13: Generación de cepas *Streptomyces sp* A92-309110 (*S. flaveolus*) modificadas por ingeniería que dan como resultado delección génica de *sfaR*

Se digirió el cósmido pTL3102 con *Bg/II* y *KpnI*. Se clonó el fragmento de ADN de 4,7 kpb resultante en pSP72 digerido con *Bg/II* / *KpnI* para crear pTL3132. Se eliminó un fragmento de ADN interno de 674 pb de pTL3132 mediante digestión con *Eco72I* y se autoligó el vector para crear pTL3133. Se clonó el fragmento de *Bg/II* y *HindIII* de pTL3133 en pKC1139 digerido con *BamHI* / *HindIII* para producir pTL3129, que se usó para modificar por ingeniería la delección génica de *sfaR*.

Se llevó a cabo la generación de la cepa recombinante doble deseada usando procedimientos similares a los descritos anteriormente. Se usó *E. coli* S17-1/pTL3129 para transformar *Streptomyces sp* A92-309110 (*S. flaveolus*) mediante conjugación. Se homogeneizaron los conjugantes resistentes a ampicilina y se sembraron en línea sobre medios de agar ISP-4 complementados con ampicilina 50 µg/ml y se cultivaron a 37°C hasta que se lograron colonias que crecieron bien. Se usó una serie de clones individuales de cada una para inocular 3 ml de caldo TSB complementado con ampicilina 30 µg/ml agitando a 250 rpm a 37°C durante cuatro días. Se extendieron 100 µl de cultivo sobre agar ISP-4 a 37°C durante 2-3 días para inducir el segundo acontecimiento de cruzamiento. Tras

inducir la recombinación de cruzamiento doble a 37°C, se escogen las colonias con fenotipo negativo para apramicina para la conformación del genotipo mediante PCR. El producto de PCR deseado es diferente en aproximadamente 700 pb del patrón de tipo natural de 1,4 kpb. La cepa final que se delecionó en *sfaR* se denominó *Streptomyces sp.* TL3010.

5

Ejemplo 14: Generación de mutantes de DH en cada uno de los 4 dominios DH de *sfaH*

10

Los cuatro dominios DH de *sfaH* están contenidos dentro de los módulos de PKS 7, 8, 10 y 11. Tal como se describe en el texto de patente anterior, la actividad de un dominio DH puede eliminarse mediante mutagénesis dirigida al sitio para alterar la histidina de sitio activo por un aminoácido apolar. Esto se llevó a cabo para los cuatro dominios DH de los módulos 7, 8, 10 y 11 amplificando 2 brazos para cada DH usando los cebadores expuestos en la tabla 4 a continuación.

Tabla 4:

15

	cebadores	SEQ ID NO	secuencia
Brazo izquierdo de DH (módulo 7)	105A-2	SEQ ID NO: 48	TTT <u>GGATCCTACACCGGCCA</u>
	105C-2	SEQ ID NO: 49	GGGCGCCC TTT <u>GGTACCGAGGACGCTAG</u> CGTCGGCCAGCCAGGGGTG C
Brazo derecho de DH (módulo 7)	105B-2	SEQ ID NO: 50	TTT <u>AAGCTTAGCACCCGTGC</u> CACCGGTCAC
	105D-2	SEQ ID NO: 51	GCTAGCGTCCTCGGTACCCC GGTGCTCCCCGGCACC
Brazo izquierdo de DH (módulo 8)	106A-2	SEQ ID NO: 52	TTT <u>GGATCCGGTGTTGTGGG</u> CGGTGATGG
	106C-2	SEQ ID NO: 53	TTT <u>GAATTCGGCGAGTACTA</u> CGGCATCGGCCGTCCAGGC GGC
Brazo derecho de DH (módulo 8)	106B-2	SEQ ID NO: 54	TTT <u>AAGCTTGCACGACGTGG</u> CCGAAGCG
	106D-2	SEQ ID NO: 55	TTT <u>GAATTCGCCGTAGTACTC</u> GCCACCACCCTGCTGCCCG GC
Brazo izquierdo de DH mód. 10	107A	SEQ ID NO: 56	TTT <u>GAATTCGGGTGGTCCGG</u> AGCTGGATCG
	107C	SEQ ID NO: 57	CGGCAGCAGGGCAGGGACT AGTATGGCGGC
Brazo derecho de DH (módulo 10)	107B	SEQ ID NO: 58	TTT <u>AAGCTTGGTGCTCGGAC</u> TGGGCGGAGC
	107D	SEQ ID NO: 59	GCCGCCATACTAGTCCCTGC CCTGCTGCCG
Brazo izquierdo de DH (módulo 11)	108A	SEQ ID NO: 60	TTT <u>GAATTCGACGACACCG</u> GATACGGGC
	108C	SEQ ID NO: 61	CGGCAGCAGGACGGTACT AGTACGGTGGC

Brazo derecho de DH (módulo 11)	108B	SEQ ID NO: 62	TTT<u>AAGCTT</u>CCACCCATGTCT
	108D	SEQ ID NO: 63	GCACCAGG GCCACCGT<u>ACTAGT</u>CACCGT

Por tanto, para generar la secuencia apropiada para mutar el DH de módulo 7, se amplificó el brazo izquierdo con los cebadores 105A-2 y 105C-2 y se clonó el fragmento de ADN de 1,2 kpb resultante en pANT841 para dar 01-44-1 y se amplificó el brazo derecho con los cebadores 105B-2 y 105D-2 y se clonó el fragmento de ADN de 1,2 kpb resultante se clonó en pANT841 (Genbank: AF438749) para dar 01-44-2.

Para generar la secuencia apropiada para mutar el DH de módulo 8, se amplificó el brazo izquierdo con los cebadores 106A-2 y 106C-2 y se clonó el fragmento de ADN de 1 kpb resultante en pANT841 para dar 01-44-3 y se amplificó el brazo derecho con los cebadores 106B-2 y 106D-2 y se clonó el fragmento de ADN de 1 kpb resultante en pANT841 para dar 01-44-4.

Para generar la secuencia apropiada para mutar el DH de módulo 10, se amplificó el brazo izquierdo con los cebadores 107A y 107C y se clonó el fragmento de ADN de 1 kpb resultante en pTLV1 para dar 01-28-1 y se amplificó el brazo derecho con los cebadores 107B y 107D y se clonó el fragmento de ADN de 1 kpb resultante en pTLV1 para dar 01-28-2.

Para generar la secuencia apropiada para mutar el DH de módulo 11, se amplificó el brazo izquierdo con los cebadores 108A y 108C y se clonó el fragmento de ADN de 1 kpb resultante en pTLV1 para dar 01-28-3 y se amplificó el brazo derecho con los cebadores 108B y 108D y se clonó el fragmento de ADN de 1 kpb resultante en pTLV1 para dar 01-28-4.

La construcción del plásmido final para cada mutación diana conlleva el ligamiento de tres fragmentos. Se digirieron los plásmidos que contenían los fragmentos de PCR que portaban las regiones de homología izquierda y derecha diana, con enzimas de restricción apropiadas y se ligaron en el vector pKC1139 digerido de manera apropiada (BamHI/KpnI para 105AC, KpnI/HindIII para 105BD, BamHI/EcoRI para 106AC, EcoRI/HindIII para 106BD, EcoRI/SpeI para 107AC, HindIII/SpeI para 107BD, EcoRI/SpeI para 108AC, HindIII/SpeI para 108BD). Se usaron reacciones de ligamiento para transformar *E. coli* DH5 α . Se analizó el ADN de plásmido mediante digestiones de restricción. En primer lugar se transforman los plásmidos finales para realizar los acontecimientos de cruzamiento doble para la transformación de *Streptomyces* en *E. coli* s17-1 y luego se conjugan en *Streptomyces* tal como se describió en el ejemplo 9 anterior. Se homogeneizaron los conjugantes resistentes a apramicina y se sembraron en línea sobre medios de agar ISP-4 complementados con apramicina 50 μ g/ml y se cultivaron a 37°C hasta que se lograron colonias que crecieron bien. Se usó una serie de clones individuales de cada una para inocular 3 ml de caldo TSB complementado con apramicina 30 μ g/ml agitando a 250 rpm a 37°C durante cuatro días. Se extendieron 100 μ l de cultivo sobre agar ISP-4 a 37°C durante 2-3 días para inducir el segundo acontecimiento de cruzamiento. Se sembraron en placa colonias individuales +/- apramicina para determinar la pérdida de la estructura principal del plásmido. La sensibilidad a apramicina indicó pérdida de la secuencia del plásmido y se analizaron colonias con ese fenotipo mediante PCR seguido por digestión con enzimas de restricción del fragmento amplificado en cada caso para confirmar la mutación de dominios DH. La digestión con enzimas de restricción de los productos de PCR de estos mutantes de DH proporciona dos fragmentos debido al sitio de restricción que se introduce junto con la mutación, lo que es contrario al de tipo natural. Se hicieron crecer las cepas en condiciones de fermentación convencionales descritas anteriormente, y se observó que producen los productos deseados (véase la figura 6).

Ejemplo 15: Caracterización de SfaR como reductasa/carboxilasa específica de ruta

Para validar el papel fundamental de SfaR para el suministro de elementos estructurales poco habituales en la ruta propuesta, en primer lugar se inactivó *sfaR* mediante delección en marco. El mutante resultante perdió por completo la capacidad para producir SFA, lo que sugiere que funciona como homólogo de *ccr* específico de ruta esencial para la biosíntesis de SFA (véase el ejemplo 16). En segundo lugar, se expresó de manera heteróloga y se purificó el SfaR con etiqueta de 6 x His N-terminal hasta cerca de la homogeneidad para estudios *in vitro*. Tal como se esperaba, en presencia de NADPH y bicarbonato, crotonil-S-CoA (figura 7, 4) se convirtió eficientemente por SfaR en el producto mayoritario etilmalonil-S-CoA (figura 7, 5) junto con el producto de derivación butiril-S-CoA (figura 7, 7). Omitiendo el bicarbonato en la mezcla de reacción, se redujo en consecuencia el crotonil-S-CoA para dar butiril-S-CoA. Para crotonil-S-CoA, $K_m = 8,4 \times 10^{-4}$ M, $K_{cat} = 6,8 \times 10^{-3}$ /min y $K_m/K_{cat} = 1,2 \times 10^{-1}$ M.min. Para excluir la posibilidad de que la reacción se produzca directamente en el ACP, se expresó SfaO y se purificó como derivado con cola de 6 x His N-terminal. Los intentos con crotonil-S-SfaO como sustrato no pudieron detectar ninguna actividad de SfaR, proporcionando la evidencia de que SfaR porta las actividades en el crotonil-S-CoA basado en CoA más que en el crotonil-S-SfaO basado en ACP (figura 7, 3). En tercer lugar, para detectar la flexibilidad de sustrato de SfaR, se sintetizó el sustrato de 5 carbonos pentenil-S-CoA. De manera intrigante, en las condiciones con y sin bicarbonato, SfaR realizó la reducción y carboxilación reductora en pentenil-S-CoA (figura 7, 8) rápidamente para generar pentanoil-S-CoA y propilmalonil-CoA (figura 7, 10), respectivamente, mostrando eficiencias comparables a las actividades en el sustrato de 4 carbonos, crotonil-S-CoA. En las condiciones usadas,

la carboxilación reductora tuvo lugar demasiado rápido como para medirse para el análisis cinético; y para pentenil-S-CoA en la producción de manera reductora de pentanoil-S-CoA, los valores de K_m , K_{cat} y K_m/K_{cat} fueron de $1,4 \times 10^{-3}$ M, $4,2 \times 10^{-2}/\text{min}$ y $= 3,4 \times 10^{-2}$ M.min, respectivamente.

- 5 Ejemplo 16: Generación de cepas *Streptomyces sp* A92-309110 (*S. flaveolus*) modificadas por ingeniería que dan como resultado delección génica de *sfaK* y producción de compuestos novedosos

Se lleva a cabo PCR usando ADN genómico o ADN de cósmido como molde, y los pares de cebadores *sfaK*-L-for (SEQ ID: 64) y *sfaK*-L-rev (SEQ ID: 65) para generar un producto de PCR (denominado *SfaK*-L), y *sfaK*-R-for (SEQ ID: 66) y *SfaK*-R-rev (SEQ ID: 67) para generar el otro (denominado *sfaK*-R). Se aíslan estos productos de PCR, y luego se digieren con enzimas de restricción (*sfaK*-L con EcoRI y BamHI, y *sfaK*-R con BamHI y HindIII), después se ligan en pKC1139 digerido previamente con EcoRI y HindIII. Se aísla el plásmido *psfaKKO*.

Se lleva a cabo la generación de la cepa recombinante doble deseada usando procedimientos similares a los descritos en los ejemplos anteriores. Se usa *E. coli* S17-1/*psfaKKO* para transformar *Streptomyces sp* A92-309110 (*S. flaveolus*) mediante conjugación. Se homogeneizan los conjugantes resistentes a apramicina y se siembran en línea sobre medios de agar ISP-4 complementados con apramicina 50 $\mu\text{g}/\text{ml}$ y se cultivan a 37°C hasta que se logran colonias que han crecido bien. Se usa una serie de clones individuales de cada una para inocular 3 ml de caldo TSB complementado con apramicina 30 $\mu\text{g}/\text{ml}$ agitando a 250 rpm a 37°C durante cuatro días. Se extienden 100 μl de cultivo sobre agar ISP-4 a 37°C durante 2-3 días para inducir el segundo acontecimiento de cruzamiento. Tras inducir la recombinación de cruzamiento doble a 37°C, se escogen colonias con fenotipo negativo para apramicina para la confirmación del genotipo mediante PCR. El producto de PCR deseado es diferente en aproximadamente 1700 pb del patrón de tipo natural. La cepa final que se deleciona en *sfaK* se denomina *Streptomyces sp. sfaKKO*.

Se hace crecer *Streptomyces sp. sfaKKO* en condiciones de fermentación convencionales descritas anteriormente. Se observa que el caldo de fermentación contiene un nuevo pico correspondiente a una sangliferrina con un peso molecular de 1034,3, que se supone que es el producto sangliferrina X. Se aíslan los productos usando métodos convencionales.

Bibliografía

- Altschul, S. F., W. Gish, *et al.* (1990). "Basic local alignment search tool." *J Mol Biol* 215(3): 403-10.
- 35 Banteli, R., J. Wagner, *et al.* (2001). "Synthesis of derivatives of the novel cyclophilin-binding immunosuppressant sangliferrin A with reduced numbers of polar functions." *Bioorg Med Chem Lett* 11 (12): 1609-12.
- Brown, E. J., M. W. Albers, *et al.* (1994). "A mammalian protein targeted by G1-arresting rapamycin-receptor complex." *Nature* 369(6483): 756-758.
- 40 Cheng, Y. Q., G. L. Tang, *et al.* (2003). "Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis." *Proc Natl Acad Sci U S A* 100(6): 3149-54.
- 45 Clarke, S. J., G. P. McStay, *et al.* (2002). "Sangliferrin A acts as a potent inhibitor of the mitochondrial permeability transition and reperfusion injury of the heart by binding to cyclophilin-D at a different site from cyclosporin A." *J Biol Chem* 277(38): 34793-9.
- Fang, J., Y. Zhang, *et al.* (2008). "Cloning and characterization of the tetrocarcin A gene cluster from *Micromonospora chalicea* NRRL 11289 reveals a highly conserved strategy for tetronate biosynthesis in spirotetronate antibiotics." *J Bacteriol* 190(17): 6014-25.
- 50 Fehr, T., J. Kallen, *et al.* (1999). "Sangliferrins A, B, C and D, novel cyclophilin-binding compounds isolated from *Streptomyces sp.* A92-308110. II. Structure elucidation, stereochemistry and physico-chemical properties." *J Antibiot (Tokyo)* 52(5): 474-9.
- 55 Fujimori, D. G., S. Hrvatin, *et al.* (2007). "Cloning and characterization of the biosynthetic gene cluster for kutznerides." *Proc Natl Acad Sci U S A* 104(42): 16498-503.
- 60 Fujimori, D. G., S. Hrvatin, *et al.* (2007). "Cloning and characterization of the biosynthetic gene cluster for kutznerides." *Proc Natl Acad Sci USA* 104(42): 16498-503.
- Gaisser, S., R. Lill, *et al.* (2001). "New erythromycin derivatives from *Saccharopolyspora erythraea* using sugar Omethyltransferases from the spinosyn biosynthetic gene cluster." *Molecular Microbiology* 41(5): 1223-1231.
- 65 Garneau, S., P. C. Dorrestein, *et al.* (2005). "Characterization of the formation of the pyrrole moiety during clorobiocin and coumestrol biosynthesis." *Biochemistry* 44(8): 2770-80.

- Handschumacher, R. E., M. W. Harding, *et al.* (1984). "Cyclophilin: a specific cytosolic binding protein for cyclosporin A." *Science* 226(4674): 544-7.
- 5 Hartel, C., P. Iblher, *et al.* (2006). "Immunosuppressive activity of the immunophilin-binding drug Sanglifehrin A in human whole blood: potent inhibition of interleukin-6 produced by lymphocytes and monocytes." *Scand J Immunol* 63(1): 26-34.
- Huang, F., S. F. Haydock, *et al.* (2005). "The neomycin biosynthetic gene cluster of *Streptomyces fradiae* NCIMB 8233: characterisation of an aminotransferase involved in the formation of 2-deoxystreptamine." *Org Biomol Chem* 3(8): 1410-8.
- 10 Jia, X. Y., Z. H. Tian, *et al.* (2006). "Genetic characterization of the chlorothricin gene cluster as a model for spirotetronate antibiotic biosynthesis." *Chem Biol* 13(6): 575-85. Kallen, J., R. Sedrani, *et al.* (2005). "Structure of human cyclophilin A in complex with the novel immunosuppressant sanglifehrin A at 1.6 Å resolution." *J Biol Chem* 280(23): 21965-71.
- 15 Kazuo Umezawa, Y. I., Osamu Kawase, Hiroshi Naganawa and Shinichi Kondo (2001). "Biosynthesis of polyoxypeptin A: novel amino acid 3-hydroxy-3-methylproline derived from isoleucine." *J. Chem. Soc., Perkin Trans. 1*: 1550-1553.
- 20 Kieser, T., M. J. Bibb, *et al.*, Eds. (1999). *Practical Streptomyces Genetics*, John Innes Foundation.
- Li, T. L., F. Huang, *et al.* (2004). "Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase." *Chem Biol* 11(1): 107-19.
- 25 Liu, J., F. JD., *et al.* (1991). "Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes." *Cell* 66(4): 807-815.
- 30 Liu, W., S. D. Christenson, *et al.* (2002). "Biosynthesis of the enediyne antitumor antibiotic C-1027." *Science* 297(5584): 1170-3.
- Maniatis, T., E. F. Fritsch, *et al.*, Eds. (1998). *Molecular Cloning. A Laboratory manual*, Cold Spring Harbor Laboratory.
- 35 Metternich, R., Denni, D., Thai, B, Sedrani, R. (1999). "Toward a Total Synthesis of the Immunosuppressant Sanglifehrin A. Preparation of Two Relay Compounds by Degradation and Their Use in the Reassembly of the Natural Product." *J. Org. Chem.* 64: 9632-9639.
- 40 Miller, E. D., C. A. Kauffman, *et al.* (2007). *J Org Chem* 72: 323-330.
- Oliynyk, M., M. J. B. Brown, *et al.* (1996). "A hybrid modular polyketide synthase obtained by domain swapping." *Chemistry & Biology* 3(10): 833-839.
- 45 Paquette, L. A., M. Duan, *et al.* (2002). "A convergent three-component total synthesis of the powerful immunosuppressant (-)-sanglifehrin a." *J Am Chem Soc* 124(16): 4257-70.
- Pemberton, T. J. and J. E. Kay (2003). "Cyclophilin sensitivity to sanglifehrin A can be correlated to the same specific tryptophan residue as cyclosporin A." *FEBS Lett* 555(2): 335-40.
- 50 Pfeifer, B. A. and C. Khosla (2001). "Biosynthesis of polyketides in heterologous hosts." *Microbiology and Molecular Biology Reviews* 65(1): 106-118.
- Powell, J. D. and Y. Zheng (2006). "Dissecting the mechanism of T-cell anergy with immunophilin ligands." *Curr Opin Investig Drugs* 7(11): 1002-7.
- 55 Rawlings, B. J. (2001). "Type I polyketide biosynthesis in bacteria (Part A-erythromycin biosynthesis)." *Nat Prod Rep* 18(2): 190-227.
- 60 Rawlings, B. J. (2001). "Type I polyketide biosynthesis in bacteria (Part B)." *Natural Product Reports* 18(3): 231-281.
- Sambrook, J. and D. Russel, Eds. (2001). *Molecular Cloning: A laboratory manual* (third edition), Cold Spring Harbor Laboratory Press.
- 65 Sanglier, J. J., V. Quesniaux, *et al.* (1999). "Sanglifehrins A, B, C and D, novel cyclophilin-binding compounds isolated from *Streptomyces sp.* A92-308110. I. Taxonomy, fermentation, isolation and biological activity." *J Antibiot*

(Tokyo) 52(5): 466-73.

Schreiber, S. (1991). "Chemistry and biology of the immunophilins and their immunosuppressive ligands." *Science* 251 (4991): 283-287.

Sedrani, R., J. Kallen, *et al.* (2003). "Sanglifehrin-cyclophilin interaction: degradation work, synthetic macrocyclic analogues, X-ray crystal structure, and binding data." *J Am Chem Soc* 125(13): 3849-59.

Sheehan, L. S., R. E. Lill, *et al.* (2006). "Engineering of the Spinosyn PKS: Directing Starter Unit Incorporation." *J Nat Prod* 69(12): 1702-10.

Sokolskaja, E., D. M. Sayah, *et al.* (2004). "Target cell cyclophilin A modulates human immunodeficiency virus type 1 infectivity." *J Virol* 78(23): 12800-8.

Staunton, J. and K. J. Weissman (2001). "Polyketide biosynthesis: a millennium review." *Natural Product Reports* 18(4): 380-416.

Staunton, J. and B. Wilkinson (2001). "Combinatorial biosynthesis of polyketides and nonribosomal peptides." *Current Opinion in Chemical Biology* 5(2): 159-164.

Steinschulte, C., T. Taner, *et al.* (2003). "Cutting edge: sanglifehrin A, a novel cyclophilin-binding immunosuppressant blocks bioactive IL-12 production by human dendritic cells." *J Immunol* 171(2): 542-6.

Umezawa, K., Y. Ikeda, *et al.* (2001). *J Chem Soc Perkin Trans 1*: 1550-1553.

Watashi, K., N. Ishii, *et al.* (2005). "Cyclophilin B is a functional regulator of hepatitis C virus RNA polymerase." *Mol Cell* 19(1): 111-22.

Zander, K., M. P. Sherman, *et al.* (2003). "Cyclophilin A interacts with HIV-1 Vpr and is required for its functional expression." *J Biol Chem* 278(44): 43202-13.

Zenke, G., U. Strittmatter, *et al.* (2001). "Sanglifehrin A, a novel cyclophilin-binding compound showing immunosuppressive activity with a new mechanism of action." *J Immunol* 166(12): 7165-71.

Zhang, L. H. and J. O. Liu (2001). "Sanglifehrin A, a novel cyclophilin-binding immunosuppressant, inhibits IL-2-dependent T cell proliferation at the G1 phase of the cell cycle." *J Immunol* 166(9): 5611-8.

Zhang, L. H., H. D. Youn, *et al.* (2001). "Inhibition of cell cycle progression by the novel cyclophilin ligand sanglifehrin A is mediated through the NFkappa B-dependent activation of p53." *J Biol Chem* 276(47): 43534-40.

Lista de secuencias

<110> Instituto de Shanghái de Química Orgánica

<120> Agrupación de genes novedosa

<130> SIO-P999PC

<160> 67

<170> Patent In versión 3.3

<210> 1

<211> 118372

<212> ADN

<213> *Streptomyces flaveolus* (*Streptomyces sp.* (DSM 9954))

<400> 1

agcacgcgga	gaacgcccgc	ctgcgggcg	cggggagagc	cgttccggtc	caactccccg	60
gcggtgtgcg	ggcgtgggcg	attccgcacc	acgaggacct	gcgccgagtg	ctgacggatc	120
ccaggggtcgc	caaggggaatc	acgcactggg	aggcagcggc	ccgcggcgaa	gtgccggacg	180
ggtggccggt	gatgggcttc	gtcgcgaccg	acagcgtcat	caattcgcac	ggcgcggatc	240
accggcggct	gcgcggactg	gtcgaccagg	cactgacccc	ggaacgggtg	gaggccatgc	300
gccccgggggt	ggaacgcctc	gtcgacgggc	tcctcgacca	gctcgccaca	cggccggaag	360
accggccggt	ggatctcogc	aaggctttcg	cctaccccat	ccccaccacc	gtcatctcgg	420
acctgctggg	gatcccgcag	cgcaagcggc	gtctgctgca	cgtgctcacc	ggcatgcaga	480
cccgtacgaa	caactcccc	gaacaggtcc	aggagatcga	cggacggatc	gaagcgctgc	540
tccgggagat	cgtcgccgaa	cgtcgcgacg	tccccggga	cgacctcatc	agcgtcctgc	600
tcacggcacg	ccaggacgga	gacgaccggc	tcaccgacag	cgaactgcac	ggcatgatcc	660
tgctgatggt	cttcgcccggc	caccagagcg	tcatcaactg	actcgtcaac	gcctgccacg	720
cgtgctcac	ccaccccag	caactggcct	ccgcccgcgc	cggggaggtc	ccgtggagtg	780
cggtggtcga	ggagacgatg	cgtggaacg	gcgcggtgaa	ccagttcccg	atgcgctacc	840
cgaccgagga	cgtcgtcatc	gccggccaga	cgatccaccg	gggagaagcc	atcctcgcgt	900
ccttcggctc	ggcgggcccgg	gacccggcgc	accacgggcc	ggacgcggac	ctcttcgacg	960
tccgccggcg	gcaggcccggc	cacctcggct	tcgggacagg	cccgcacttc	tgtatcggga	1020
cacacctggc	ccgtctccag	ctggagacgg	caactggccg	attcttcacc	cggttccccg	1080
acgtgcggct	ggccccgtcc	gacacgtgga	ccgtgcggcc	cgtgccctcg	ttcgtcagca	1140
acagcatcga	ggcgtgcc	gtactcctcg	gccctgccgc	ggagtacggg	cggctctgac	1200
gtcgcggcac	gcacgtgta	cggagccgag	ccccacaaa	cgcgcggcaa	ggtgctcgcc	1260
gacggcctgc	gcagaccgac	agagaggacg	ccctccgtgc	ccccgcacac	catcgccctg	1320
ctggcaccca	ccaccgcagt	ctccccgc	gaaaggctca	ccgtccccgc	cttcggcgac	1380
gtcacctgg	accggcccggc	acacgcgcg	tggcacccgt	acctgcacga	gttcgaaccg	1440
tcatccgagc	tcgcggcggt	gtggatcggg	gcggacgccg	agccgctgcc	cgccgtccgg	1500
gcgctgcgg	cgactcgaa	gttcaaccgg	acccgggtgt	acgtcgtcgg	gccegatcag	1560
gcgggcccgg	ttcccggctg	ggcgggaagca	ctggacgtcg	aggacatcgt	ccggcccgc	1620
gcggtggcgc	ccgaagggtt	caccacctcg	gtcgaactgg	ccctgcgggc	gcaccgcgcg	1680
ctgatgacgg	acgacctcta	caccgactac	tacctcgaca	tgaccttcaa	caagatcttc	1740
gactggttcg	agacgacgcg	ctgggactgg	accgaggtgg	acctggaccg	catcgagcac	1800
ggcatgctcg	acgaacggac	cgtggacttc	ctcaccgagg	cggccgctcat	cgagttcggc	1860
acctaccgc	gcgctcaaaa	cttcctgcgg	gaatggcagg	acgaggtcag	cttctcctcc	1920
tgggcaactgc	aatggggtgc	ggaggaatcc	cggcaactca	tgggtccaggc	ccgctacctg	1980
gaccggatcg	gggtctcact	ccgttcgaag	cacgcgctct	acaaacgaga	gccgtatccg	2040
cagggcgacg	tgcgctcgc	caccctcatg	atgaacgtga	tctcggagtc	acgcgcgtcc	2100
gccctctaca	aggcgtcgc	cgcccatgtg	tcggagccgg	tgatccgag	aatctggcgg	2160
ctcctggccc	gcgacgaggc	ccggcaactgc	cgtgcggtct	ccgtcttcat	gcgggaactg	2220
tgcgacagcg	atccggccca	tcaggccgc	gcgctgcgca	tggcctacat	ctggctcgcg	2280
gaccgtagca	acggcgtcaa	gcacccggcg	ggcatggtct	accgcattc	cacatccacg	2340
gccggcatcc	gccgtatcga	gaccatccag	cacgactcgg	tggacgccgc	ggacgccaaag	2400
gtgatgtcga	tcgtccggct	cctggcggac	gacgactcgt	tggagacccc	gcgggacatc	2460
aaagccaagc	tgcgcgcct	ggcccgcgg	cccggcgtc	gctgaccgctc	agtcccggcg	2520

tccgcctccc	gccgcgcccc	gacgcgcggy	cctcacagaa	caggacaaca	ccacgccatg	2580
tcacaggcga	gtgtcaccga	catcgaacag	accgtccgcc	gatccgtcgc	cgagcaactg	2640
cgcatcggtg	tcgaggaggt	gggcgtggac	accgatctgc	gctccttgcc	gcgcttcgac	2700
tcgatccacg	cgctccagat	cggtctggag	atcgagcagc	actacgacat	cgagggtggag	2760
gaccgcatcg	tcttcgagac	gagcaccgtg	cgggagtctg	cgcccgcgat	caccaaggtg	2820
atcgcgagg	ccggcacgga	accgctgtcg	tgaccctggt	cgcccgcgctc	ggcccgctcg	2880
ccgcacaagg	ctccggacgg	ggcatgcacg	tgctgcgccc	ggccaggac	ggcgacgcgc	2940
tgctctacag	cgacctcttc	gccgaggccg	ggcgggtggc	cgcggggctg	ctgtcgcgcg	3000
gggtgcgacc	cggcgagcgg	gtcgcctcog	tcctgccgac	gtcgggtggac	ttcgcgcgcg	3060
ccttcttcgg	cggtctggcg	gcccggcggc	tcgcggtgcc	gctgcccggt	cccgcgcctc	3120
tcgggtcgtc	cgacgcctat	ctccggcgca	cctccgcggc	cctcgcccg	ctcggggtac	3180
ggaccgtgct	gacggctccg	gccatgcgcc	cgctgctcgg	tcctggactg	tcggcggcg	3240
cccgcccggt	cgacgtactc	ctgggtgccg	aggtggccga	gccgtccgcg	gcccacgtcg	3300
cgcgggcggc	gtcggaccgg	gccgtcgttc	agtacacgtc	gggcacgagc	agcgaaccgc	3360
ggggcgctcg	gctcagccac	ggcaacgtgg	cgcccggggt	ggcggccatc	gcccattggca	3420
cacggctcgc	caggaccgac	gtcggctgca	cctggctgcc	cctcttccac	gacatgggcc	3480
tgatcggttc	cttcttgaca	ccgctgctgc	acgacgtcga	catccatctg	ctgacaccgg	3540
aggactacct	gcgcagtccg	ggtgactgga	tcggggccat	ggggcggctg	ggggcgacct	3600
tcaccatggc	tcctgactcc	ggctaccgct	acgtcctgag	gcgcgacgcg	gcgccgcccg	3660
ccggcctcga	tctctcccgc	tgccgcacog	cggtgaacgg	cgccgagccg	gtcgaccggc	3720
ggcttcagga	cgcttcgccc	gagcgtttcg	cgcccgcggg	actgtccgag	aacgtcttcc	3780
tgcccgccta	cggactggcg	gaggcgaccc	tgccggtcgc	cttccctccc	ctcggcgcgc	3840
cgccaagggt	gctcgtgccc	gaccgggacg	aactgaaccg	ggcccggtac	gtaccgggtc	3900
cgcccgcgcg	cgggccgtgt	cgcgaaactg	tcagtgtggg	aacgcccgtg	cgccgtaccg	3960
aggtccggct	caccaccgcg	agcggcgcac	cgcccggtcc	cggcacggtc	ggcgccgtcg	4020
aggtccgggg	ggcgtccgtc	acgagcaccg	ggtacgaccg	ccacccccgaa	gagagccgtc	4080
gcgtcctcct	tcggggcggc	tggttgccga	ccggtgacct	gggctgtgg	cacgacggcg	4140
agctgtacat	cggtggccgc	acgaaagagg	tgatcatcgt	cttcggcgcg	aaccactggg	4200
cgagcgacat	cgaggcggtc	gtgcgggaca	caccgcgcct	cctcgtccac	ggcgtcctgg	4260
ccgagcagat	ctggagcgac	gagggaggcc	gcctcgggct	cgctcgtcgag	acgacgcggc	4320
acgacgaggg	gagccgtcgt	gccacgacgc	aacacatccg	cgacaggtgt	gtggccgagc	4380
tgggcatcac	accggacacg	atcgagttcg	tgccggcgcg	cgggatcgca	cgcacctcca	4440
gcggcaaggc	tctgcgccc	ccgcgcctt	cggacacgta	catgtcccgc	tcgggtgctg	4500
gacatacccg	gaacaccgga	ccacaccgca	cggcagctac	agggcactca	gtgctccaga	4560
caccgcggag	ccccctggcc	acctgcccgt	gggaactccc	cgcacgaccg	gcctggagg	4620
cctcgcctgg	cgcacgacgc	atacctgctc	ggttactccg	atagcgagca	ccaaaggctg	4680
acccttcaga	acgaggcgat	ctcgcggttc	accgaacgca	ccctgcggtc	ggccggagta	4740
cgctccggga	tgagcgtcct	cgacatcggt	gcccggactg	gatacgtctc	tctgctggcc	4800
gcccactggg	tcgggcccgg	cggacacgtc	gtgggaatcg	agcgcgacc	cggaaccgtc	4860
cacaaggcgc	ggatcctggc	acaggaggcg	gggctcgcgg	actcgggtcg	gtacgaagcc	4920
gcggacctgg	ccgacttcac	ctggccgatg	aagttcgacg	tcctggtcgg	caggtacatc	4980
ctgatgtacc	tgcccgatcc	cgcatccgca	ctgcgcccgt	tgacgtccct	gctgaaaccg	5040
gacgggatcg	tcgtcatgca	cgagatggac	ttcaccaaca	cgctgcgctc	acgtccgccc	5100
ctgccggagt	gggacgagta	ctacgcgctg	tgcccggccc	ccttccctggc	ggccggcgcc	5160
gtccccgact	tcggcccccg	cctcaccggy	acgttcctcg	acgccggcct	ccggcagccg	5220
gaagtggcca	cctggacgccc	ggtcgcgggg	cgctcgggag	atcccgcggg	cctcaactgg	5280
ctcagcacca	cctgcccgca	gatcgcgacg	ccgctgcgca	acgccggcgt	cacggtgccg	5340
gaccacctcc	gctttgacga	cgccctgcca	cagcgactcc	gggacctcgt	cctgtcccgc	5400
agagcgcagg	tcgacggggc	gacgcagttc	ggtgccacg	ctcggctgag	accacaggcg	5460
tccaacgcgt	caggcccgca	cgccgggatg	cgctgagagg	cagggaacac	accgtgcgcg	5520
ctacgtgacg	aacatccggc	agccggctct	tcggccaaag	ccggtctgtc	cgccgaagcc	5580
gggcagcccg	aacagcgtgc	agaccgaaca	gcgggcccgc	ttccggggcg	tggaaaaagc	5640
ccgctcgcgg	gaccggccg	ccaccgatg	ggtcgtctcc	cacggccgtc	ggcggcaggg	5700
gagacagaca	ccccgcggcg	tttccgaccg	aacaggagta	atcgaccatg	tcgggcaact	5760
tcgatcacca	cagggteggc	cccctgccc	ccttccctggc	agcgcctca	agtgaggccg	5820
tcaaaccgggt	gcggatgccc	actggtgatc	acatgtggct	ggtcaccgac	tacgccgtgg	5880
ggcgcgccc	gctcgcgac	acacgtctca	gcagagctgc	ggccaccac	cccgatgctc	5940
cgaaatgggg	atccgtcgac	ctctcaccca	actcgatcat	gagtctggac	ggtgcggacc	6000
acgcccgggt	gcgcggggtc	gccggagcag	cgttcaccag	cgcccaggtc	accgcagag	6060
cagcgggtgat	cgaggaggcc	gccgacgcac	tgctggatcg	gctacagggt	gcggggccag	6120

gqgcccacct	ggtcagcggg	tatgettccg	cccttcccat	ggtcgcggtg	agctcactgc	6180
tcggcggtcc	ggcggcggag	cgccccgttt	tcgacgcgcg	cgtgactgcg	ctgttcgaca	6240
tgacgcctca	gggaacccgc	agaagaggcc	gccatatcct	gactcttacc	aagtacatgt	6300
cctcactgat	cgaaagaaag	agggatgcgc	ccgggcagga	tctcctgagc	gcgctggtgc	6360
gggttgagca	gtcgggtgag	atatcccgca	tcgagttgat	caacctcgga	ctggcacttc	6420
tcattggccg	ttacgagacc	accgcccatc	aattgagcct	tgccatactc	gaactgctca	6480
ccggtgacgc	gctcagcggg	caatccgtgg	aggatctggt	tgaggagctg	ctccgcaaga	6540
cgccgtccac	acctatcagt	tttccccgag	tggccatcga	ggacctgttc	ctgggcggga	6600
cgacggtgcg	ccaaggagag	gccgtcatcg	tctcgttgcg	gcacggcaat	catgacgcgc	6660
aagtattttg	cacatcgggt	gcgaagagca	gqgcccgcg	gcctgcacac	ctcacattcg	6720
gacacggcgc	ccaccgggtc	atgggcgcgc	cactggtctc	tctccagatg	aagatcgcgc	6780
tgacccgctt	gtggcagcgc	ttccccggcc	tcgactcgcg	accgggggac	gaggccctcg	6840
tatggcggga	gggcctcgcg	acgcgcggtg	tgtcaaggct	ggtcgtcgag	tggtaaccca	6900
cacgttcggc	gtccacacag	gctcggctgc	tcgctcgggg	tcgcggcgct	gttctcctga	6960
tcggtgtacg	cgctgttccg	gggaggagcg	ggcgcgctg	gcatgcgggc	gatctttcct	7020
tggtcaccgg	tgagagtga	cctgtggagg	tcggcgactc	cagcgcggct	gcggatacgt	7080
gcgtcagtgg	cccctgtccc	accactccgt	cgttccgcgt	cgttccgccc	accgcctaca	7140
tgatcacctc	cgatgatccc	gctaccgggt	gcaccagcgt	cttccatcca	gggcacgagc	7200
acgtagattc	gaacacccct	ccgtgccgat	atgatgccgc	ttcggtcgct	cgacggaggc	7260
caggctcgtc	acctgcgcct	ggggcgcttc	cacgactgcg	ctccgcggcg	gtgcgtgtgc	7320
cgtggaagag	ccggtgcgcg	ggaatcgcgc	ccgcggacgg	tgtcccggga	agaagcgcag	7380
gtatggcagt	gatctgcgtc	ggcggcatga	tcgggatcgg	gaagacaagc	gtggccgagc	7440
tgatcgccaa	agactgtggg	agcatggtct	tctacgagag	cgtggaggac	aaccgatcc	7500
tgccgctctt	ctacaccgcc	ggtcctgagg	agatagcggc	gaagcgcctat	cccttccttc	7560
tccagctcta	cttcttgacg	acgcgattcg	ccgcgatcaa	ggaggcgtac	aaacaggacg	7620
acaacgtcct	cgaccggctc	atctacgagg	actggtactt	cgccaaagtc	aatcacgacc	7680
tcggcaggat	cagttccctg	gagatgcagg	tgtacgaggg	gttgctcgcg	gagatgatgc	7740
gtgagatcga	cggcctgccc	tatcgcaagg	caccgatctc	catggtgtat	ctcagagcgg	7800
acttcgagac	ggtgctgcgc	cgcatcggac	tcggggggcg	tgacttcgag	caggaccaa	7860
gtctcgtcga	gtactaccgg	acgctttggg	ccgggtacga	cgactgggtg	cacgagcact	7920
attccgccag	cgatgtcctc	gtgatcgaca	tgaaccgcac	cgactgggtg	aacagccccg	7980
acgacgcgaa	gcgtgtcgtc	cgggaggctc	cggacgcctc	tgccgtggcg	aaagtccggc	8040
gcatctgacg	ctgcgcggcg	atcggtcgcc	gaggcggcgc	ggcatgcgca	gcacggctca	8100
gccgtccgga	gacctcgtcg	gcacggcagc	ggagcagcgt	gccgtgaaag	ccgcccgcac	8160
caggatcggg	ggagcggcgg	tcagcgtctt	ccgaagtggg	acgggtggca	caatgaccgc	8220
tgacttggcc	gacagcgacc	cgaggaaacc	ggcgtgaatc	cgacgcggca	gcgcgactgg	8280
gcctgcccg	caagatggcg	ccgcggctgc	tcgcgggcaa	ctcggcgccc	gcgctcatcg	8340
ggtcggctgg	gcaactgctg	aggtcccggg	ccggtctggg	acaccgctca	ccgagctgac	8400
acgggggatg	aagatgaagt	cgagcaccgc	cgcggccgcg	cgtccggcga	cggggcccag	8460
catcccctgt	ccgggcatcg	tgacgggtcca	ggtggtgtct	tcggccagga	ccgtttccag	8520
ctcgggtata	tgcaactcgt	ccaaggcgtg	ggcgtagttc	gccaccgcct	gctggagttt	8580
cggcgatcac	gaggccaccg	tcagcacgat	cttgccctgg	atgtggccct	gcgcggctcg	8640
cgtgtgtgcg	ctgaccgctt	cggacagcgg	gtaggtgctg	tccaccccga	cccggagtgt	8700
gccctcgtcg	aacaggcgcg	cgatctcggc	gagctggggg	ccgtgggaac	gtacctgaat	8760
gttcgagacg	gtgatgtcca	gactcgccgt	ctcttcgggg	tcgtactcgg	cgaagaacac	8820
cgggagcagg	gtgccgcgcg	gcttgagtac	ggtcaggaag	cgtgagccgt	ccgggcccgc	8880
gacggtgtcg	atcaccaggt	cgacaccgct	gaccacgtcc	gcggcctgcg	tcgtggtgta	8940
gtcagatgaac	tcacggcac	cgagctcgcg	caggaaccgc	tcgtgccgac	ttgaggccac	9000
cgcgatgacg	tgtgccccct	tccatttcgc	caattgcacc	gcgaaatggc	ccactccacc	9060
ggcgcccccg	ttgaccagca	cggtcatccc	cggcgtgatc	ggcaccggtc	ggtgcacctg	9120
gccggtgaaa	ggagacggca	cctcgtggcc	gaggtcgacc	aggtactgcc	aggccgtgag	9180
cacggccatc	ggcgcccccg	ccgcctgcac	gtggtcgata	ccggccggct	tgtgagccag	9240
gtcagaagcc	ggcgcggcca	cgtactgggc	gtactcggg	ccgtcgaatc	cggggaaccg	9300
cagcatgccg	aagacctcgt	cgccgacgcc	gaacccccgc	acgtccggag	cgaccgcctg	9360
gaccacgccc	gacatgtccg	ttccggggat	cagagggaac	tccagcgccg	gcctcatccc	9420
ggccggcatg	accttcacc	cttcacgcag	gtaccagtcc	ggcgggttga	tgcccgcgcg	9480
gtgcaccccg	acgagcacct	cgccccgggc	gatctcggga	gccggcacct	cgtcgtaccg	9540
cagaacttcc	ggcccaccct	cttcgtggat	ctggatcggc	ttcatcgcgc	ctgctgcttt	9600
ctcggggaaa	cgttgctcct	tcagtcaaag	cccgggacgg	catggccgct	caagaccggt	9660
tcggcaatct	gaccattccg	aaacggcatg	acgcgtaccc	tcgaggggtg	aacgatctcg	9720

ggcaggacct	ggaactgcgg	ctggtgcgct	acttcaccgt	ggtggcggcg	caccagcact	9780
tcggccgggc	cgccgccgac	ctgcacgtag	cccagccggc	gctgagccgt	cagatccaac	9840
ggctcgagaa	atatctcggc	acacgactgc	tggaccgcac	ccccagggc	acccggctca	9900
ctccggccgg	ccggacgttc	ctccccggg	cccaggccct	gctgcggggc	gcccgccagg	9960
ccgagctggc	cgtgctgaa	caagcccgga	ccgaacgaat	cgccatcggc	tacgtcgaag	10020
acctggtgat	cactgccgcc	gtgcgggaac	tgcgccggcg	ttacccccgac	gccgagatcg	10080
ccacccggca	tctgcaactgc	cgcgacgtcg	gggcactgtc	cgacaggcgc	gtcgacgccc	10140
tgatcgcgcg	ggccccgctg	ccgctcgtcg	ccgacgacgt	gttcaccacc	ccgctgtacg	10200
aggagccccg	gatgctcgtg	gtcccgcgcg	gccatcctct	ggccgaccgc	gcgtcgggtga	10260
ccgcggaaga	actggccggc	gaagaggcgg	cgccgtgtgc	gttcgagacc	gcaggctggg	10320
cttcctacca	gttcctcgga	accggcgtgc	cgccatcga	gagctacgag	gacaagctcg	10380
aactcgtcgc	gagcggcagg	gccatcgccg	tgctaccggg	cgccgatcgg	cgagctcac	10440
tgctcccga	cctcgtcacc	gtcccgtcg	agggcgctcc	ccccagccgg	gtcgtcctgg	10500
tcagccgcaa	ggcgaccccg	aatccgatga	tcaggaattt	ccggctggct	gccaaaggccg	10560
tcctgaccgc	cccggcatcc	tgagcgggac	caggccgggc	gatggccagg	ccgacgagcc	10620
gggcggcgcg	tggcggttca	caacggccac	caccgatccg	gagcaccttg	acgggcccgtg	10680
atgctggggc	ctctccggga	gcgggcgacc	tgctcgtcgc	gcagtactgc	gagcgcagcg	10740
aggactcgcc	ctggggggcg	ggggagtggc	ggaacagcgc	gcagtgggaa	cgtggcggca	10800
cgaggtcgcc	caccgggccc	gcgagcgacc	gcgtggccca	gaggttcggg	ctcgccttg	10860
gcgccgccgt	gatgcaggcg	ggagattctc	gggatgggccc	cgccccggtcc	ggtagcgcgc	10920
gagtgtcgcg	ggacgactga	cacggcccca	cagaatcggg	catcctggga	aaatgtctga	10980
cgtgccttgc	tccgcttgca	gcggatcccg	agtcactgag	cacaccgagc	actctgtcga	11040
agtcgataag	gacgggaacc	agaagcccgt	cgcccgtagc	tggaccggct	cctgcaatac	11100
ctgccacggc	agcggcctca	gctgacgaag	gggaagcggc	ttgcgatcaa	tctatctgtg	11160
cgtcggctgc	gcagcgatct	acctgccgcc	cgagcatccg	acggtcggac	ggccggtatc	11220
gtcggcgcac	tgctcacggc	ccgactgccg	ggccgcagtg	cgccaggccc	tcaggatgtc	11280
cggactgccc	gaggagaaga	tcggagagct	ggcccgtcgt	gccgccggca	tcgacggcga	11340
acctcgccgg	ccgagcctcc	gcttccgacc	cgcacgcggg	cgccgagcaa	ccgccgctcc	11400
gggcggcggg	cgccggcaagc	tgccgtagct	gtccgatgta	gagaagtgac	cccgctttcc	11460
tgaacctcat	caaccagcgc	actgggaagc	ggcgatgcgc	ttcgtcttcg	gtggtctgcg	11520
ctcggctcgg	tcgtcgtggg	ctttgaggcc	tggtgcacgg	tggaggtgcg	aggtcacctt	11580
ttcgaggatc	tcgctgtgcg	gaaccttgga	ctccggaatc	gtctgtatcc	gtgcgttcgg	11640
ccgggtgcgg	atgcggcgat	gaggagaggg	gcgagacgat	gggtgaaggt	gagagggaga	11700
cgtggacgac	ggaggagtcc	gggtcgtccc	acgtgggtgc	ggtcgggggtg	ttgctggccg	11760
atggcaccgt	cccggatcct	gtcctcttc	tcagctcttc	cgagcgggag	gccaggtggc	11820
tatcggagcg	gagcgtctac	gacggccggc	ctcatggcgg	accgcggggc	gcagcgttgc	11880
gcgcggtgtg	ctcgtgcgga	tgggcgggac	ccggccgcgc	tctcgactgg	aacgagagcg	11940
acgaggaaga	ggcgaggagc	ccgagggagg	ccgccgaagc	cgcatcgccc	gcgtccttgc	12000
ggtactcgtt	ccccacgcg	agcgacaccc	gggcagcatg	ggccagcggc	cacggcgtgt	12060
acttcccgga	cgggctccga	agccacgacg	actccaacga	acaggctgcc	ggcaccgcgcg	12120
cgaacagcgg	aagcagcgca	ctgggcccgat	gcctccacac	ccgcgctaca	tcatccccat	12180
cccgtaccgc	ccgacccccct	gatggccacg	tcgtgcactc	cagggatctc	gtgtccacc	12240
ctctgcgctg	tctgctcat	tgacgcccgt	ccaggggacc	ggcaccgatg	ccccccgacc	12300
gcaatgccct	ctgcgggcaa	ccgccgcgct	ccccgtaggg	gtcaatggcg	cgcccttggc	12360
gccacacccc	taatcagaca	catgtgtcaa	ctgctcatcc	gcccactcgg	cgaaggtggt	12420
gagcgggatg	cccaacttcc	gggcgaatgt	gggcccggcg	ggctggagga	ccagttgtt	12480
ccactcgtgt	ccggtcctcc	acttcggcat	gcccgcggca	aggcctctt	ccacgctcag	12540
ggagggcgcg	gtcacgggca	cgccccaggc	ggcggaacaag	gtctgcgcga	tctgctccat	12600
cgtgcgaagg	tcaccagcca	gttccagttc	cacctggtgg	aaccggtcgg	ggtctcggag	12660
ggcgtgtgcc	gcggccgtgc	cgatgtcccg	cacgccacc	aaggccagtt	cggtgtccgg	12720
cttcagtacc	gtcagcagtc	cgccgcgggg	ccccttgggt	gccagcaggg	gcaggttctc	12780
catgaagaag	gcgggcttga	tcaccgtcca	gcgggcgaaa	cccgcaccac	gcaccgcctc	12840
catgatcgcc	tgcttggtgt	ggaagtactc	cgccatcgcc	gccagcggc	cctcggccca	12900
gccggcgacc	cgggtgtgtt	caccgactcc	actggtcgag	gactgtacga	actgccgtac	12960
tccccctatc	ttcgcccgct	ccaccaggtt	ggtggcctgg	gcgagttcgc	tcgcgaagtc	13020
cacgctggtc	tcggtcatgg	gcggcatctg	caccgagaac	accgcgcgga	ccccctcgac	13080
cgccgggtcg	agggaggccc	ggtcggaaag	atccgcgctg	accagttccg	cgcccagcgc	13140
ctcgatcgcc	cgggcggact	tcgacgaggg	atcgcgtacg	agcgcacgta	cgggcacctt	13200
ggcggccaaa	agcgcgcgag	ccgtggcccc	gecttgcctg	ccggtggcgc	gggtgaccag	13260
gacgggtatcg	ctgctgctgg	gcatggtgct	gctcctcaag	ggccggagag	acaaaacggc	13320

ggggcccgcc	gtttcgtgtg	ccgctaagat	atggcggggc	ccgccactta	gcaacgagga	13380
gaccatgact	ggccagcgct	cggacgcccg	acgcaactat	cagcgcaccc	tcgccgtcgc	13440
cgaagccgaa	gtcgccgcgc	acggcgccga	ggcatcccag	gagcagatcg	cccgcacgcg	13500
aggcgtcgg	tcggcgaccg	tgcgcccgtca	cttccccacc	cgccggggcg	tcctcgaagc	13560
cgtcttccag	gagcgcattg	cgggcctgtg	cgagcgcgcc	catcggctga	gtgctcga	13620
ggacggccgc	accgcactgc	tggagtgggt	ccacgccttc	gtccgctacg	ccgtctcagc	13680
ccgcggattc	gccatatacc	tcagctacga	gccgcccacc	gaagaaccct	ccccgaagag	13740
ctgcccggggc	gtaatcgaag	cagccgcaac	tccttctgct	cagcgggcca	tcggcgacaa	13800
ggcgggtgca	ccgcacgtca	cctttcacga	cctgctgacc	ctcgccgtcg	gcatcgccct	13860
ggccaccgag	caccacacag	accccgcac	ccaggccgac	cgctcttcc	gcctcgccgt	13920
ggaagggctc	agccctcagc	cgggcccagc	ggggagccc	gggaactccg	gctcgtcggc	13980
gcccgggtgaca	ggaagtccgc	cgaatgagca	cgcgtgagaa	gtgggatgcc	tcctcggatga	14040
caaccgaact	gcctcggcac	gccctcctcg	ccagctcatc	cgccgtcggc	gtgagaactg	14100
cctgagcttc	ggcttatcca	taacaccgag	ggcctcccc	gggtccccag	caccctgtga	14160
ggcagctggg	gatccggagg	aatggacccg	cctgacctct	ggtggacgcc	ttccgagggg	14220
cgacctgagg	gactgagggc	ggatgacctg	tgacagcgaa	ccttgatcga	gcgccatcga	14280
aggttgacca	gtcaccgaat	ccgccgtcga	ccggatcggc	cgaccgatcg	gatgccgggt	14340
gggtcggctc	gaggaacgg	gacctttgac	ccgtgtggag	tggtccggag	cggatatccg	14400
tatgtcggcg	gtcgtccgcg	cggagcgacc	ggcggggcagc	cccggcggag	cctcgtcggg	14460
ccttctcgg	caccacctac	agcaaaggac	gcacaccgtg	ggcatcatcg	cctggattct	14520
catcggcctc	atcgccggcg	ccatcgccaa	ggcactcatg	cccggcaagg	acccgggccc	14580
ctgcctcgtc	acgatgctca	tcggcatcgt	cggcggtctg	ctcggcggct	ggctcggcaa	14640
ggatgatcttc	ggcgtgcaact	ccatcaacgg	cttcttccac	ctctcgacat	ggatcgccgc	14700
ggctcgtcggc	tcggtgatcg	tgctcctgct	gtaccgcctc	ttcaccggac	agcgttcggc	14760
ctgaccgcga	ccggcgatcc	acgtctgtcc	gtagccggcg	tcgagccctg	aaccgactgt	14820
gctcggggcg	ggcactgcgc	ccctgagctg	cggcggtccg	tccttctgct	caagaactgc	14880
attcggcca	ccttcgaggc	cagtcgctg	tcacaggtca	gcgggcccagc	tcggaattgg	14940
tgtcacgatc	accgaacgtc	acaatctggt	tgtcacgggc	tcctgtcggc	ccgccgcccga	15000
acgggtgaca	gtagccgctg	ctgctggggg	cgagctgtcc	ggcctgtcct	gtgggatcgg	15060
tgtgggaggg	ggcgaatcgg	cgttttggga	cgccttctg	gggatgccgc	gcgtcgtgca	15120
gcccggcggtg	ctggattctc	tgctcgtcgg	gccgcctggt	ggagctggac	cggctgcgcc	15180
gcccggccggt	acgggtgacg	ggcgccggca	gagtggtcgc	ggtagccgggc	ggacgacaag	15240
gtgtcgttgc	cgcgcggca	ctcggactct	cgcagcctca	caacgctgct	ggccgcccggc	15300
gtgtacctca	cgtcgcgggc	ggtggacgac	gcctggatc	tgctgaaggt	cctgatcggc	15360
acgaagcttg	cagaccggct	gcacatcctg	cgccgggtg	ctggacgcca	cattcgaaaa	15420
gcccctgctc	aatacccgcc	cccaaccgag	acgaggttga	acgacaagct	cagggcgtct	15480
gagtacttcg	gttgagtagt	tctgccatg	tgaggggag	ggccgtcggc	acaggctgtg	15540
tacacaacgc	cgcgggtgcg	agagcgtcgc	gcgaagagac	cagtagcggg	catctatgag	15600
gggatacttc	atgtccaagc	gccttctgct	ctcgtcgtc	gccgtggccg	cagccgtcgt	15660
tgccggcacc	gtggtgttcg	tctcctcggc	cgacgcctgc	gtgccggcca	agccggagat	15720
ctcaaaggcc	accgcccact	acaccagtac	ggccgggtgg	agcgcctcgc	tcaccttcag	15780
cgccaccgtg	gccgacaact	ccggaatcaa	gagcctgcgg	gtgctcgcct	ggccggcgag	15840
ttcgggccc	gcgcccacgg	cgggcgagat	gcccggatgtc	gaggaagcca	cgtgcaaggc	15900
gacttccgcg	acggcctcgg	tgtgcacctc	caccgtgaag	agctcggcca	aggaagccgc	15960
cgcgttgcgc	aagggcgtct	ggcacgtatc	cgtgctggcc	accgccaagg	accacgcac	16020
gaccttcgcg	ccccaaaggc	ccacgttcac	cgtcaagcac	tgaccggctcc	gccccccgg	16080
aatgatgatc	gcggcccggc	cgcggggcgg	accgtatcgc	gacctccatt	ggacccttg	16140
aaccgcacga	cgcgggtgac	tcgcccacc	ctgccccagg	gcggacggag	ttcacctggg	16200
ggcgacggcg	cgcatctgac	gcgcgctgcc	acgatgccgc	accactccgc	gcgcggggcg	16260
atgcaggcga	agtgactccc	agctcgacgc	gctcgccaaa	cacggcatct	cgcgcgacta	16320
catcttcggc	gagaagatca	gcacccgggc	gcccgggagc	ccgaagtcc	gggaggaggc	16380
gctgagggcg	gcgcgggagg	tcaaggcgca	cgcgccacac	tgccgtgtca	tcctcacggg	16440
gtacgagcgg	aagcggctcg	gtcgaacgc	cgcggaactc	accgcccctcg	ccgaccacct	16500
caccgcccac	ggcttggctc	tggagatatt	cgcggggccc	tgctcgaagga	ctcccggagc	16560
cgtggaaccc	ggcggcccacc	cgacgcgaca	gcccgcagcg	gcagttgggg	cgctcccggc	16620
cggcctgccc	ggacaccgaa	acgcccggca	ccacaagcga	aagagcgtcc	gtcggcaagc	16680
tgaccgggtcc	tcatgaagga	tttagggcag	tgatttggga	cacacccgaa	cgcgcccggc	16740
ggactctgagg	aatcgcttag	ggcccgcctc	tatcgggaac	ttgaagccgc	cctcggcagc	16800
caacgcttga	ctcgggttcc	ggcgggtcgg	atgacgataa	tttccgggtga	gctgcgcaa	16860
aagggtagat	agcgggcccga	tagaaaactc	ttgcgagtg	tgccgggtggc	ttgtagggtc	16920

ctaataaatc	ggctggacaa	gggaaggttg	atgcgggcgt	ccgaacccaaa	atagcttcgg	16980
acagcaactg	ctgccttctg	tcgatggaag	tagggggaag	ttcgtggaaa	tcggctcggg	17040
cgcgcccgaa	ttaaccgcgt	cgctcgggtgta	tcagcagcgg	cgtgacccaaa	tcgcccgaag	17100
cgctgcccgc	tatgtgcccg	gcgagcccat	tcagagggtc	gagtacacgg	acgccgagca	17160
cgctctgtgg	cgccctggttt	ccaagcggct	cgcggaccgg	caccggcaca	tggcggcgcc	17220
ggagttcgtc	gagggcggcg	agcggctcga	ggtggggcggc	gacggcgtcc	cgcagttgcg	17280
tgaagtgtcc	gaccggctcg	accagctgac	cggattccgc	ctacgccccg	cgtccggcgt	17340
ggttcccttc	gccctgttct	gcggtctctct	ggccgacggg	tacttccact	ccaccagta	17400
cctccgcgac	agcgcgacac	ccttctactc	gacggaaccg	gacatcctgc	acgaggtcat	17460
cgggcacggc	agcgccttg	ccgacgaccg	gttcgccaac	ctgtaccgcc	tggcccggca	17520
ggcgtgctg	aggtggagt	ccgaggaccg	cgtccagttc	gtcgccaaga	ccttctggtt	17580
cacgctcgaa	tggcggcctgc	tggacgcgc	cgaaggggccc	cgggcctacg	gcgccagcgt	17640
cgtctcctcc	tacggcgagc	tggagcactt	ccgttccgcg	gagatccggc	cgtggacat	17700
cgccgacatg	gcccacgtcg	actacgacat	aaccagtac	cagaccacct	tttacagcgc	17760
gcgttccctc	acgcatctgg	aggacgtcgc	aggtgaattc	tgggcgtcct	gcgacgatac	17820
atccattgaa	aaactaatgg	cggttgatat	atgaccggc	tcgagagca	atcatccact	17880
gcgagcaga	gcccggaaatc	agaagtactg	gacgtcaccg	gaatcggatt	cggtgcccg	17940
aatctcgccc	tggcgggtggc	gctccatgaa	tccgaagccg	ccgggaaggc	ccttttccctg	18000
gagaagcaga	aggaattcgg	ctggcatcgg	gggatgctcc	tggggggctc	ctcgtccag	18060
gtgtcctttc	tcaaggacat	cgccacgatg	cgcaatccca	ccagtgattt	cggttccctg	18120
tcctatctcc	aggagaagga	ccggctggtc	gacttcatca	accagcacac	cctgctgccc	18180
tcccggatcg	agtaccacga	ctacctccag	tgggcgcgcg	accggctgaa	ccacctggtc	18240
gagtagggc	tggagccac	cggtgtcgg	ccggtgaccg	aagccggtga	ggtcgtcgcg	18300
ctcgacgtgc	tcgcccggga	ccgggtggtc	gcccggacca	gaaacctcgt	cctcgcctcc	18360
ggcctgcgcc	cccggctgcc	cgagggcgcg	gagaccggcg	aacgcgtctg	gcacagctcc	18420
cagttgctgc	accggctgcc	cgcgttcgac	gaacgcccgc	cccgccgggc	cgtcgtggtc	18480
ggcgccggcc	agagcgcggc	cgaggtcgcc	gcgacctca	tggaccgcta	cccgcaggcc	18540
gaggtgtgcg	cggtgttcgc	ccgctacggc	tacagcgtcg	ccgactccag	cccgttcgcc	18600
aaccgcgtct	tcgaccggc	cgccgtggac	gacttctact	tcgccccgcc	cgaggtcaag	18660
caggccatca	tgcgctacca	cgccggcacc	aactacggcg	tcgtcgacga	ggacgtcctc	18720
cagggcctct	accgccgcca	gtacgagcag	aaggtgtccg	gcgccccgcg	gctgcggggtg	18780
atgaacgcct	cccgcctggt	gtccgtcgaa	ccgcgccagg	aatccgccgc	cgtacgcgtg	18840
gagttcctgc	ccacggggcga	acacaccgac	ctggacggcg	acctggtcgt	gtacgccacc	18900
gggtacgact	ccaccgacc	ggccgaactg	ctcggcggcg	tctccggcgc	cctccggcgg	18960
gacgagcgg	ggttctggg	gatcggccgc	gactaccggc	tcggcaccac	cggggatctc	19020
cggtgcccga	tctacgtcca	gggcgccacc	gagggcacc	acggcatcgc	ctccaccctg	19080
ctgtccatgg	tggcgggtccg	cgcgggcgag	atcgcccgg	cgatcaccgg	cgccgggtgc	19140
gaccgggacc	gctccaccgg	aagcaaggca	gcagcgggga	acaggggctg	aagtgtacga	19200
acgtccgctg	taccgggagg	attgcgacgg	cgtcgtcctg	gcgtttctgc	gacacaacc	19260
actggcaatg	gtcgtcacct	cgcacgacga	cgtcccgggtg	gccaccacag	cgccgggtgct	19320
gttccggcac	ggaccgacg	gcgcccagc	cgaggccgtc	gccgcgggca	ccgtcccgt	19380
cgccggctcc	accctgatcg	gccacatgaa	cgtcgagaac	ccgcagtggc	gccgatgcg	19440
ctccggcgac	cgggcgctca	tcgtcttcca	gggcccgcac	ggctatgtct	cgccgacggt	19500
ctacggggtc	acgcccgcgg	ccccacctg	ggacttcatc	gccgtccacg	tgaacggcac	19560
agtggagccc	accgccgacc	ccgcccggc	gctggacatc	gtctccgaca	ccgcccggcg	19620
gctggagctc	ggcttcggg	gcggctggga	ccaggagtc	tcctcgcact	acttccgcca	19680
gatcgccc	ggcgtggcg	ccttaccct	gcggtcgat	tcggtgcaga	cgatgttcaa	19740
gctcagccag	gagaagccc	ccccgatgcg	gcggcgcgtg	gtcgagcagt	tcaagcaag	19800
cgagtccggc	accaccgcg	ccctggccag	cgtgatgcg	gaccgcggac	taccgaagc	19860
cgacgaggag	cgggagacag	ccggatgacg	caggggagct	cgcgacggcc	gggactggac	19920
tcggcactca	accggctccg	ggagcgcacg	gcgaggaccg	gcacgacacc	cgcgggcatc	19980
agccgggtcc	cccgcgaccg	gccgctggag	ctgtcgttcg	cccagcagcg	gctgtggttc	20040
ctcgaccggc	tgatgccgga	cagcgccttc	tacaacttcg	gcaccgccc	gcgctccgc	20100
ggcccggctc	gcgcccggc	gctggcccgg	gcgctgagct	gtgtcgtcgc	ccggcacgag	20160
accctgcgca	ccgtcttcgc	cgaccagag	ggcgtggccg	gcgcccgtc	cctgccgccg	20220
gagccgggtg	ccgttcccgt	gaccgacgcc	gtggacgagg	ccgacgccga	gcggctggcc	20280
ggcgaggagg	ccgcgcgcc	cttcgacctc	accaagggcc	ccctgctccg	cgcccgcctg	20340
ctgcccggctc	ccgacgacga	ccacgtcctg	ctgctgacgg	tgcatacat	cgccaccgac	20400
ggctggtcgc	acggcctgct	gtgggcccag	ctgaccgcgg	cctacaccgc	cctgcccgac	20460
ggcccggcagc	ccgcgctgcc	ggaactgccg	gtgcagtagc	ccgacttcgc	cgccggcag	20520

cgccgtaccc	tgtcgcccgc	cgtgctggag	cggggctgg	cgtaactggca	ggaccggctg	20580
gccggcctgc	ccccgctgga	cctgcegetc	gaccggcccc	gcccggccgt	cgectccgcc	20640
gaggcgccgc	tcgtcacctg	gcggctgccg	goggacgccg	tcgccgccgc	ccgggcccgc	20700
gccgcaacgc	agggcgccgc	cctgcacatg	accctgctgg	ccgcgttcag	cgcggtcctc	20760
ggggcgccgc	cccgcaccga	ggacgtcgcc	gtcgcccagc	cggtcgccgg	ccggcccctg	20820
gccgaggtcg	agcaactcat	cggattcttc	gtcaacaccg	tcgtcacccg	caccgacctc	20880
ggcgccgacc	ccaccttcgc	ggaactcgtc	gaacgcgtcc	gcgccgcctc	ggtggacgag	20940
atggcgacc	aggacgtgcc	gttcgagtac	ctggtggagc	ggctggtccc	cgagcgcgac	21000
ctgtcccgca	accgctcgc	ccaggtcgtc	ttccagtagc	tgccccggcc	cgccgcgcgc	21060
cccgcgccgt	tccccggcac	caccgcccga	cccttcgcgc	gcgaccgcgc	cttcaccgc	21120
atggacctgg	aggtctacct	cggcgaggac	gcccagggcg	gcgtcgaggg	actgatcaac	21180
tacagccgcg	ccctgttcga	ccgggagacc	gtcgagcggc	tggcccccca	cctgacggcc	21240
ctgctgcgcg	ccgctgcgc	cgagcccagc	cgcccgctgt	cccggctgac	gatgaccgac	21300
gcgggcgagg	acgcggccct	ggaccggggc	gcccggggca	ccggcgtacc	gctgcccgag	21360
gcctcgctgc	ccgagctgtt	cgccgcccag	gcccggccga	ccccggacgc	cgctgcctgt	21420
gccgacggga	cggaacacct	cacctacgcc	cagctggacc	gcgccgcca	ccggctcgcc	21480
catgtcctgg	ccggccacgg	agccggacc	gagagcgtcg	tcgcgctggc	gacggaacct	21540
tccgcccacc	tggctcgtcgc	cgtgctcgcg	gtcctcaagg	ccggcggcgc	ctacctgccg	21600
ctcgacgccc	gcaaccggc	cgcccggacc	cgcccgctcc	tcgccgacac	cggtgccgcg	21660
ctgctgctga	ccgacggcgg	tcccgcaccc	gcccggcaccg	agcacctgcc	ggccgtagac	21720
ctgctgcgcg	tccccggcac	agcccccgac	accgcttgc	cgaaacacct	cgccccggac	21780
ggcctcgcgt	acgtgatgtc	cacctccggt	tcccagggca	cgccccaggc	cgctccacag	21840
accacccgcg	ccgtggccgc	cctcgccctc	caccggcgct	ggtccggcgg	cgcccagaa	21900
cgggtcctgc	tgcactcccc	gcaggcgttc	gacgcctcca	cctacgaact	gtggtccccg	21960
ctgctgtccg	gccgcccgcg	cgtggtggcc	ccgcccggcg	ccctcggacc	cgccgcctc	22020
gcccgtgtcg	tcgccgacca	gggctcacc	gcgctgtggc	tgacctccgg	cctcttcgac	22080
ctcgtggtcg	aggaggacgt	cacctgcctg	gcccggctcc	gcgaaactcgt	cgctggcggc	22140
gacaccgtgt	ccccggcgac	cgtggcccgc	gtgcgcggcg	cccacccgga	cctgacggctc	22200
gtcaacggct	acggccccac	cgagaccacc	accttcgcca	ccctccacc	gatcgcgccg	22260
gccgaccccg	cccccgccgg	ccgggtgcc	atcggcagcc	ccctggacaa	caccgcgcgc	22320
cacgtcctcg	acgaccggct	gcggcccgtc	ccgttcggcg	tgcccgccga	actgtacgtc	22380
ggcgccccgc	gcctggcccc	cggtacgcc	ggccggccgg	cggccaccgc	cgagcgcctc	22440
ctgcccgacc	cgtccggacc	ggcgggcagc	cgcatgtacc	gcaccggcga	cggtggtgcgc	22500
cgccgccccg	acggcgtcct	ggagttcctc	ggccgcgcgc	acgaccaggc	gaagctgcgc	22560
ggcctccggg	tcgagcccgg	cgaggtcgag	gcgctcctgg	ccgcccacc	cgccgtcgcg	22620
cacgcccggg	tcgtggtgcg	cggtgacggc	ccggcccggca	agcggctcgt	cgcccacgtc	22680
gtcccgcgcg	ccggccgcac	caccgacacc	gcccgcctgc	gcgcccacgc	cgcgccggcc	22740
ctgcccgact	acctggtgcc	ctcgcccttc	gtcctcgccg	acgcgctgcc	gctgaccgcc	22800
accggcaagg	tcgaccgcgc	cgcgctgccc	gccccggcgg	agacggcggga	cgccggcctg	22860
gcccgccccg	gcaccgacgc	cgagcggggc	ctgtgcgaga	tcttcgccga	gctgctcgac	22920
gccgacgcgt	tcggcgccga	ggacgacttc	ttcgtacgcg	gcggccactc	gctgctcgcc	22980
accgcctcgt	tcgcccgaat	cgcccgcgcc	ttcggcaccg	aggctcccgt	gcgcgaggtg	23040
ttcgagcacc	gcaccccgcg	agccctggcc	tcggtggtcg	ccggcgcgcg	gctcccggcc	23100
gacccgccgc	cgccgctggt	gcccgcggac	cgcgaccggc	tgctgccgct	gtccttcgcc	23160
cagcagcggg	tgtggttcc	ggaccagctc	gccccggcga	gcgctcgtg	cacctccggc	23220
ggagccctgc	gcgtgcgcgg	accgctggac	cccagcggc	tggccggcgc	cctgtcggcg	23280
gtcgtcgccc	ggcacgagac	cctgcgcacc	accttcacgg	tcgccgacgg	cgtgccccgc	23340
gcggtcatcg	gcgcccgcgc	ccggctcgcc	ccgcccgatcg	tcgacgtacc	cgacgcgcgac	23400
gcgccacgcg	ccgcccctc	cgccgagctg	tccaccgggt	tcgacctgac	gcggggcccg	23460
ctgctgcgcg	ccacgctgct	ccggctcgcc	cccgaacgac	acgtcctcgt	cgctgcctgt	23520
caccacatcg	cgaccgacgg	ctggtcccag	gcccctgctgt	gggcccagat	cgccgcccgcg	23580
tacgacggcg	ccccgctgcc	cgaactgccc	gtccagtagc	gcgaccacgc	cgatggcgag	23640
cggctcctggc	tgaccggcga	ggtgctggag	cgcccggccg	gctactggac	cgccggctg	23700
gccggcctcg	ccccgctgga	actgcccgtg	gacaaggccc	ggcccgcctg	cgccaccggg	23760
cgggcgggca	ccctgcccctg	gcagctgccc	gcccagctga	tcccggacgc	ccgggcccgc	23820
gcccgcccgcg	agggagccac	gctctacatg	gtgctcctgg	ccgcgttcac	cctcgtcctg	23880
tcccggtagc	cccgcaccga	ggacatcgcc	gtgggctcgc	cgacggcccgg	ccggaccgcg	23940
gcccgagacc	agggcctgat	cggtctcttc	gtcaacgtcg	tcgccgtccg	caccgacctg	24000
tccggcgacc	cgacctctcg	ggaactcctg	ggcccggtag	gggagtcggt	ggtcggccgc	24060
gtcgagcacc	aggacgtccc	cttcgagcac	ctggtggaac	ggctgcgccc	cgagcgcgac	24120

ctgtcccgca	acccgctcgt	gcaggtggcg	ttccagctgc	tcgccgacgc	gccccggcgg	24180
ccctggtggc	agggcgcccc	ggccgagccg	ttcgacatcg	accacgcgta	cacccggatg	24240
gacctggagg	tgcacgccgt	cgagaccggc	gacgaggtcg	gcgcgaccgt	cctgtacgcg	24300
gccgacctgt	tcgacgccga	caccgtccgc	cagctgatgc	accacgtgtc	ggtggtcctc	24360
ggcgaggtgc	tcgccgaccc	cgaccggccg	gtctccgcgg	cgaccatgct	cgacgagacc	24420
gaccggcacc	gcaccctggt	cgccctggaac	gacacggccg	caccgctgcc	ggacggctgc	24480
gtgccccggc	tctacgccga	gcaggccgc	cgacccccg	acgccgtcgc	cctgatctgc	24540
ggcgacgagc	gggtcaccta	cgccgaactg	gaccggcggg	ccaaccgctt	cgcccacctc	24600
ctgctggccc	acggcgtcgg	cgccgacgaa	ccggtcggcg	tcgccaccgg	acgctccacc	24660
ggcatggtgg	ccgccgtgct	cggcgtgctc	aacgcggggc	ccgcctacgt	gccgtcgcac	24720
ccgcgcaacc	ccccgggccc	caccgagcgc	atcgtggcca	cctcggggct	gcgcgtcgtc	24780
atcgccgacc	gcccgtgccc	cggcaccgac	ggcatcaccg	tcctcgacgt	caccgaccgg	24840
ggccccgggc	ccgacaccga	ccccggtatc	gaccgcacc	ccgacaccac	ggcgtacgtc	24900
atctacacct	ccgggtccag	cgccgagccc	aagggcgtgg	ccgtgaccca	ccgcaacgtc	24960
gtcgtgctgg	ccgccgaccg	ccgctggagc	aacggcaacc	acgaacgggt	gctgctgcac	25020
tatccgctgg	ccaccgacat	ctccacctac	gagctgtggc	cgttcctgct	gaccggcaag	25080
cagatcgtcg	tcgccaccga	cgagcacgtc	gaaccgcaca	ccttcgaccg	gctcatccgc	25140
gagcacggcg	tcaccgcgat	gtgectgccc	gcacccctgt	tcagcctcct	ggccgaggag	25200
tgcatggagt	gcttcggcgg	gctgcgcgag	gtgctcaccg	gcggcgaggc	cgtgtccggc	25260
gagaccgtcg	cccaggtgat	ggccgcccac	ccgcacctga	ccgtggccga	cgectacggc	25320
ccgaccgagg	cgaccgcgtt	caccacgctg	ttcccgatgg	aacccggctt	ccggctcgcc	25380
ggctcccggg	tgcccatcgg	cgcgccgatc	gacaacacc	gggtgtacgt	cctcgacgac	25440
accctccgcc	cggcggcctc	gggcgtggcg	ggcgaactgc	tcacggcggg	gccccgggtg	25500
gcgcgcggct	acctgggccc	gcccgacctg	accgcgaga	agttcgtgcc	ggacccgtgg	25560
ggcgggcccg	gcgagcggat	gtaccgtacg	ggcgacgtcg	tacggtggct	gcccggcggc	25620
gcgctggagt	tcctcggccc	cgccgacgtc	caggtcaaga	tcgcggctt	ccgggtggag	25680
cccggcgagg	tcgaggccgc	gctgctgcgg	cacccggcgg	tctcgcaggt	caccgtcgcc	25740
gtgcgcgagg	acatcccggg	cgacaagcgc	ctggtgcct	atgtggtgct	ggagccggag	25800
gccggcgcg	gcgtgctgcc	cgccctgcgc	gcccacggcg	cgggctccgt	acccgactac	25860
atggtgccgt	cggcgttcgt	cgcgctcgac	gcgttcccgc	tgtccaccac	cggaagatc	25920
gaccgccggg	ccctgcccgc	gcccgacacc	cgctcggctc	ccgagagcgg	gtacgtcccg	25980
cccgcgaccg	aggcggagcg	ggtgctgtgc	gagatcttcg	ccgaggtgct	ggacgtccac	26040
ccggtggggc	ccgacgacga	cttcttcgcc	ctgggcgggc	actcgtgct	ggcaccgcg	26100
gccatcgccc	ggatccgcgc	ccgcttcggg	cccgacgtgc	cgctccaggc	ggtgttcgaa	26160
cggcgtccc	gcgcggcct	ggccgagacc	ctcggcgagc	ccggcacggc	cacggacgtc	26220
atccggcccg	cccgcgggga	cggcgcgcgc	cttccgctgt	cgtcctccca	gcggcgccctg	26280
tggttcctgg	accggctcac	cccggacagc	ggcttctgga	acgtggcgat	gggcgtacgc	26340
gcccacggcc	cgctcgacgt	cgacgccctc	ggccgggcac	tgtccctggt	ggtctcccgc	26400
cacgaggcgc	tgccgaccgt	cttcgcccgc	gacgcggcgc	agcctatggc	cgtcgtacgg	26460
cccgcgacc	cgctccgcct	ggaggtcacg	gacgtcgcgc	acgaggccga	ggtacgcgcc	26520
ctggccgagg	cggacgcgc	gcgccccttc	gacctggccc	ggggaccgct	gctgcgcgcc	26580
cgcgtgctgc	gcctcgcggc	cgaggaccac	gccgtcctga	tcaccgcgca	ccacgcggtc	26640
acggacggct	ggtcgcacgc	cgtggtctgg	ggcgagctgg	ccgaggccta	ccgcgccgaa	26700
ctgtccggcg	accccgcgga	actgcccgaa	ctgectgtcc	agtacggcga	cttcgccgtc	26760
tggcagcagg	gccggctcac	cggcgccgaa	ctggagcggc	acctcaccta	ctggcggggc	26820
cggctcgccg	ggctgcgcgc	cctggaactt	cccctggacc	ggccccgccc	cgcgctcgcc	26880
ggctcggcgg	gcgctccca	gccgtgggaa	ctgcccgagg	acctggtccg	ggccgcgcgc	26940
gcgttcggcg	acaccgaggg	cgccaccctc	tacatgacgc	tgctgaccgc	cttcaccgtg	27000
gtcctggcac	ggttcgcccg	caccgaggac	gtcgcgctcg	gcgccccctg	ggcaggccgt	27060
acccgccccg	aggtggagcg	gctgatcggc	ttcttcgtca	acatgctggt	gctgcgcacc	27120
gacgtctccg	gcgacccgac	cttcgcccgc	ctgctgggcc	gggtgcgcga	gacggtggtg	27180
ggcgccatgg	accaccagga	cctgcccctc	gagcacctgg	tggagaccct	ggcaccggag	27240
cgggacctgt	cccgcaacc	gctggtccag	gtggtcttc	agctgatgcg	ggcggcccgg	27300
gacaagggcg	accgcctcgg	cgccgcccgc	ctggaaccgc	tgctggacga	gcacgccttc	27360
acccgcgtcg	acctcgaagt	gcacctcacc	gaggacggcg	accgggtgcg	cggcaccgtg	27420
ctgcaactca	ccgcgctggt	cgaggcggac	acggtccgcc	gcctgctgca	tcaccacacg	27480
gtgctgctgc	gcgcgcctc	cgccgaccgg	gacgcgcgc	tgctccgcct	ctcgtgctg	27540
gacgacgacg	accgggcct	gctgctggag	cggtggaacg	acaccgccct	cccgtaccgg	27600
gacgtcccgc	tcgtggagct	gttcgcccag	caggtcgccc	gcacccccgg	cgccaccgcc	27660
gtggagtgcg	aggacgacgt	cctcacctac	gcccgcctgg	accacgaggc	cgagcggatc	27720

gccgcccggc	tgcgggccc	ggcggtggg	cggagcacc	tcgtcggcct	ctgcctggag	27780
cgcggcaccg	tgcagatggc	cgctctgac	ggcatcctca	aggccggcgc	cgcctatgtg	27840
ccgatcgacc	cgagccacc	ccgggaccgg	atccggctca	tcgtcgacga	cgcccggatg	27900
accgtcggcg	tcaccgaccg	ggctcacgcg	gacgtcttcg	ccgagggcac	cgccctcgtc	27960
ctcgtggacg	ccccggccgg	gcaggagccg	cccgcggcaa	cgcccgccgg	acgaccggc	28020
cccgactccc	tcgcctacgt	cgtctacacc	tccggctcca	ccggcgtgcc	caagggcatc	28080
gccatgcccg	cccgggtgtg	ggtcaacatg	ctcgcctggc	agaagaagac	ggtaccgggc	28140
ggacccggca	cgcgcaccgc	ccagttcacc	gcccctacct	tcgacgtgca	tgtgcaggag	28200
gtgctctccg	cgctcctgta	cggcgagacg	ctggtcatcc	ccaccgagga	gacccgccc	28260
gacccggccc	gcttcgccc	ctggctcgac	gaacgggccc	tcgagcagat	attcgtgccc	28320
aacctgatga	tccgcgcgct	cgccgaggag	gcccggccc	gcccggccc	gctcacctcg	28380
ctgcggcaca	tctcccaggt	ggcgagccg	ctgtcgctgg	acaccgtgct	gcgcgagttc	28440
tgcgcccggc	ggcccgtct	gcgtctgcac	aaccactacg	gctccaccga	gatccaggtc	28500
gtcacctcct	tcaccctgcc	cgctgacgtc	gcccactggc	cccggaccgc	acacctgggc	28560
gagccggctg	acaactccc	cgccctacgtc	ctggacgacc	ggctgcggcc	ggtccccgtg	28620
ggcgtggccg	gogaactctg	cttcgcccgg	cccggactcg	cccgcggcta	cgtcggcaag	28680
ccggaactga	cggcgcagaa	gttcgtgccc	gacccttcg	gcccggccc	ctcccgcctc	28740
taccgcaccg	gggacctggg	ccgctggcgg	cccgaacggc	gcctggaata	cctgggccc	28800
ctcgaccacc	aggtcaagat	ccggggcttc	cgggtggagc	tgggcgaggt	cgaggccgtc	28860
ctgctgcgcc	accccagggt	caccgcgcgc	gtcatcgtcg	cccgcgagga	cgccccggc	28920
gtcaagcggc	tcgtcggeta	cgctcgtgccc	gtccccggtg	ccgacggagg	gctgcccggc	28980
cggctgcgcg	cccacctggc	cgacgcgta	ccggactaca	tggtgccgtc	ggcgtcgtc	29040
gcgctggacg	cgctcccgt	gaccaccacc	ggcaagatcg	accggcccgc	gctgcccgc	29100
ccggacctgc	gcaccacgct	cgacaccggc	ttcacccgac	cccgtaccgg	tgcggagcag	29160
acgctgtgcg	aggtgttcgc	cgaggtgctg	gagaccggcg	gggtcggcat	cgacgacgac	29220
ttcttcgcg	tcggcgggca	ctcgtggctc	gcccggccga	cggtcggccc	gatccgcgag	29280
gcgctgggcg	ccgaggtgtc	cctgcccggg	ctggtccagc	accgacccc	gcgcgccctc	29340
gcccaggtcg	tcgcccgtgg	cccgcgcacg	gcccgtaccg	ccctggtgcc	ggtgtcccgg	29400
caggagcccc	tgccgctgtc	gctgggccc	ctgcccgtgt	ggcggctgca	cgaggcccac	29460
cccggcgacc	cggtctggac	gatcccgtcg	gcccgtcggg	tcaccggcga	actggacgcg	29520
gacctgctgg	gcccgcgcc	gaccgaggtc	gtgcccggc	acgagggcgt	gcgcacggtc	29580
ttcgtgccc	gcgacgagcc	ggcgtcgggt	atcctgccc	ccacggacat	cgtgctggac	29640
ccggtggacg	tggccgacga	gaccgcgcgc	cgcgcctgg	ccgacgagge	agcggcccgg	29700
ccgttcgacc	tcgtcccggg	accggtgctg	cggcccctgc	tgctgcggat	cgccccgcac	29760
gaccacgtcc	tgtgctgac	ggtgcaccac	atcgccaccg	accgctggtc	ccagggcgtg	29820
ctctggaccg	aactgtccgg	cgccctacgcg	gcactacgag	agaaccgtcc	ggccgagctg	29880
cccgaactgc	ccgtccagta	cggcgacttc	gccttctggc	agcggctcctg	gctgaccgga	29940
gcccgcgctg	acgcccagct	cgccactgg	cgcgcggc	tggacggcct	gcggcccctc	30000
gcccgtccc	gcgtcccggc	cgacgcggcc	cacgacgcca	ccggcgtgct	gaccgagtg	30060
cggctgccc	ccggcctggt	cgccaccgcc	cgcgggctcg	gcgcccagca	cgacgccacc	30120
ctctacatga	ccctgctggc	cgcgttcacg	gcccacgctcg	cccgtgggg	gggacggac	30180
gacctcggcg	tcggctccc	ggtggcgggc	cgcacccggg	ccgaggtcga	ggggtgatc	30240
ggcttcttcg	ccaacttcgt	gcccctgcgc	gtcgacctca	gcggcgacc	gtccttcgcc	30300
ggcctgctgg	agcgggtccg	ggacaccgcc	ctcgacgct	acgcccacca	ggcctgccc	30360
tgggaacgcg	tcgtcgaagg	cctcggcctg	gaccccgaac	agccgctggt	cgacgtcgtg	30420
ttccagctcg	tcaacgtcga	actcggcgaa	ctcggcctgc	ctggcggccc	cgtcgagcag	30480
ttcacggggc	aacaggcgta	cgcccgtgg	cacctcgagc	tgcacctggt	cgaggacccc	30540
gacggcggcc	tcaccggaca	cgctcgtacac	cgcgcggccc	ccctcggccc	acgcgtcgtg	30600
gacggcctgc	tggccggcac	cgccgcgctg	ctacaggccg	cactcgccga	acccggcctg	30660
cccgtctccg	cactgcacgc	accggacccc	ctggagggga	ccacccatga	ctgagccgct	30720
gaccgacccg	aaggctgctc	agccgcccgc	gctcgcgcac	cggtcggccc	gtctgcccga	30780
acccgagcgc	cggcgcgctg	tcctggacct	ggtccggaca	caggtcgcgg	aggtgctcga	30840
cctcggcacc	gccgcccagc	taccgcccga	ccgggctgatc	cgggaaactcg	gactgcgctc	30900
gctgaccgcc	gtacaactgc	tgtcgggct	cagccgcggc	accggcgtga	agctgcccac	30960
caccgcccac	tacgaccacc	ccaccgccc	cgcctggcg	gacgtcctgg	tcgacgccc	31020
ctccggccc	cagctcccgg	tcgcccagga	cgacgagccc	gtacagcgcg	cggcccagca	31080
cgacccggtc	gcccgtggtg	gcatggcctg	ccgcttccc	ggcgggtgta	ccaccccga	31140
cgagctgtgg	cggctcgtgc	tggaggagcg	cgacgccatc	acgcccgtcc	cgcccagccg	31200
cggctgggac	ctcgcgcgcc	tcgcccagcc	ggaacggccc	tcgcccagcc	gcacccggca	31260
cggcggattc	ctcgacgacg	tcgcccgtt	cgacgcccgc	ttcttcggca	tcagcccgcg	31320

cgaggcccag	ctgatggacc	cccagcagcg	gctgctgctg	gagaccagct	gggaggcgct	31380
ggagcgggcc	ggcgtcgcac	ccggctcgtg	gcggggcgga	cgcgtcggcg	tgttcgtcgg	31440
cgccaacgcc	cagtcctact	cctcgtcgtc	cgccggcatc	ccggaaggcg	gcgacgggca	31500
cgccctgacc	ggccgcctgg	ccagcgtggt	ggccggccgc	atctcctacg	tcctcggcct	31560
ggaaggcccc	gccttcacgg	tggacaccgc	ctgctcgtcc	tcctcgtcgg	ccctgcacca	31620
ggcgggtgcg	tcgctgcgct	cgggcgagag	caccctcgcg	ctggccggcg	gtgtgaccgt	31680
catgccgacg	ccggagctgt	tcgtcgactt	caccaagcag	ggcggcctgt	ccgaggacgg	31740
tcgctgccgc	tcgtacgcca	aggccgcgga	cggactcggc	tggccgagg	gcgtcggcat	31800
gctcctgctg	gagcggctct	ccgacgcccc	gcgaaacggc	caccgcgtgc	tcgccctgct	31860
gccccggcacc	gccgtcaacc	aggacggcgc	ctccaacggc	ctcaccgcgc	cgagcggcgc	31920
cgcccagcag	cgcgtggctc	ggcaggcact	ggccagcggc	ggactgcgtc	ccgccgctg	31980
ggacgcctg	gagggccacg	gcaccggcac	cgccctgggc	gaccccatcg	aggcccaggc	32040
actcctctcc	tcctacggcc	aggaccgcga	ccggccgctg	tggctgggct	cgctgaagt	32100
caacctcggc	cacgcccagg	ccgccgcggg	cgtcggcgga	gtgatcaaga	cggtgctcgc	32160
gctccggcac	ggactgctcc	cgaagacgct	gcacgtcgac	gagcccaccc	cgcacgtgga	32220
ctgggtgtcc	ggcgacgtac	ggctgctgac	cgaggccccg	ccgtggcccc	ggggcgagcg	32280
tccccgccc	gccggcgtgt	cctcgttcgg	cgtcagcggc	accaacgccc	atgtgatcgt	32340
ggcggaggcg	cccgaaccgg	aggaagcgac	ggcgcggccg	gccggtgcgc	tccccgtgcc	32400
ctggcagctg	tccgcgcgca	ccgaggccgc	gctgctcgaa	caggccgccc	ggctgcgcga	32460
gttcgtcggc	gccgaccccc	gcctcgagcc	cgcctccgtc	gggttcaccc	tggccacccg	32520
ccgcaccggc	ttcgagcacc	gcgccgcctg	cgtgggcccg	gaccgcgccc	aactgctcgc	32580
cgcgctcgac	gtgctcgccc	cgggcggcgc	cgaccggccc	gtcgtccgcg	gcaccgagg	32640
agccgacggc	agcgtcgtct	tcgtcttcgc	cggccagggc	ggacagtggc	tcggcatggg	32700
cgtggaactg	ctcgacaccc	acccggtggt	cgccgcggcg	atggccgagt	gcgagcgcgc	32760
cctcgccccc	tacctggact	ggtcggctgt	ggacgtgctg	cgcggagccg	aggacgcgcc	32820
gccgctcagc	cgcgtcgacg	tggtcgagcc	ggtgctgttc	gcgctcatgg	tctcgtcggc	32880
cgcggtgtgg	cgtcccacg	gcgtcgtgcc	ggcggccgtg	gtcggccact	cccagggcga	32940
gatcgccggc	gcctgcgtcg	ccggcgcctt	caccctggac	gacgccgcga	agaccgtcgc	33000
gctgcgcgcc	aaggccgtcg	ccgacctgcc	cggccgctgc	ggcatggcct	tcgtcgcggc	33060
ccccgcggcg	aacgtcgagc	ggatgctgga	gcgctggccc	ggccggctcg	gcatcgccgc	33120
tgccaaacagc	ccccgctcgc	tggtcgtcgc	gggogaccgc	gatgccctgg	aggagctgct	33180
cgacctctgc	gacgacgagg	ggctgcgcgc	cggccgcgtc	gccgccgact	acgcctcgca	33240
ctccccgcag	gtggaggccg	tacggcagcg	gctgctcggc	gacctcaagg	gcatccgccc	33300
ccgggacggc	gacgtgcccc	tctactccac	cgtcaccgcc	gactggatcg	acggcagcga	33360
actgggcgcc	cgctactggt	accgaacct	gcgogagccc	gtgctcttcc	agaacgcgat	33420
cagcgatctg	ctcgccggcg	gccaccacgg	cttcgtggag	atcagcccgc	accccgctgt	33480
cacggctcggc	atgcagcaga	ccgccgaggc	ggcgggcacc	gaactgcggg	tcgccgccac	33540
cctgcgccgc	gacgagagcg	accggctgcg	cctgaccacc	gcaactggccg	aggtggccgc	33600
cgacggcgtg	ccggtcgact	ggcgcctcct	gttcgacgcc	accggcggcc	gcccggctga	33660
actgcccggc	tacccttcc	agcgcgagcg	ctactggctc	acccccagg	ccgccgcggg	33720
cgacctgacc	gccgtcggcc	tcaccgcccg	agggcatccg	ctgctcgccg	ccgaggccga	33780
actccccgac	gaggacggcc	tgctcctcac	cggccggatc	agccccgaga	cgcacccctg	33840
gctgaccgag	cacaccgtgc	tgggcaccgc	cctgctgccc	ggcaccgccc	tcctggagat	33900
ggtggcccac	gccggccacc	ggctcgaccg	cgcccagatc	gaggagctga	ccctggccgc	33960
gcccacggc	gtccccggcg	acggcctgac	cctgcgcgtc	accgtgcacg	gcgccgacgg	34020
gtccggccgc	cgcacggctg	ccgtccactc	ccactccggc	gacggctgga	cccggcagct	34080
cacgggcgtc	ctggcaccgg	ccgagccggc	cgagccggcg	gagcccgaga	ccggcgcgtg	34140
gcccccgggc	ggcgtgagc	cggtcgggt	ggaecggcgtg	tacgaccgtt	tcgccgcccg	34200
cggctacggc	tacgggcccc	ccttcagggg	cctgcgcgcc	ctgtggcgcc	gggacaccga	34260
ggtctacgcc	gaggtcgaac	tccccgacac	cgtcgacacc	gccggccatc	tcgtgcaccc	34320
ggcgtgctc	gacgccgtca	cccaggccgt	gcccgcggcc	gagcccgagg	acgcgccgct	34380
gctgctgccc	ttcacctgga	ccggcgtgac	cgtgcacccc	ggcccggccc	gggtcctgcg	34440
ggtgcgcgcc	gcccgggaga	gcgccgacac	gctgagcctc	accgcgaccg	accgcgaggg	34500
ccggcccgtc	gtggagctgg	ccgcgctgcg	gctgcggccc	gcgtcggccc	ggcagctccg	34560
cgccgtggcc	gccgcgggca	cgcgcgacgc	gctgttccgc	gtgacctggc	agacccccga	34620
cgccgagaca	cccgcggccg	cgcgctgcgc	cgtcctgggc	gacggaccgg	aaggcctcgc	34680
cgcccccctg	agcaccgccc	tcgccgacat	cgaecggggc	gcggtggacc	tcgtcctcgc	34740
tcccgtcagc	acaggagccc	atgtggtcgc	cgcgcggccc	cgggccaccg	cccaggtgct	34800
ggaactgctg	cacgagtggc	tcgccgacga	ccggttcggg	caggcccggc	tcgcgatcgt	34860
cacccggcac	gccgtcggcg	cccggcccgg	cgaggaaacc	gacccggccc	ccgccgcggt	34920

ctggggcctg	gtgcgctccg	cgagaccga	gcaccccgac	cggttcctgc	tcgtcgacac	34980
cgacggcaca	cccgcctccc	tggacgctgt	gcccggccgtc	ggcgacgagc	cgagaccgc	35040
gctccgcgac	ggcgaacgcc	tcgtcgcccc	gctcaccggg	gccgctgaga	ccgcgctgcg	35100
cccgcctcgc	ggcgcggcag	cctggcgcgt	cgacgtggtc	cgcccccggca	gcatcgacgg	35160
cgtcgacgcc	gtcgccgctc	ccgacgccac	cgccccctg	gcacccggcc	aggtccgcat	35220
cgccgtccgc	gcggcggggc	tgaacttccg	cgacgtgctg	tgcgcgctcg	acatgtacce	35280
ggacgaggtc	gacgcgatcg	gctccgaggc	cgccgggacc	gtggtcgcg	tcgccccgga	35340
cgtcaccgac	ctcgccgtcg	gcgaccgagt	cctcggcatg	gtgccggggc	gcttcggcac	35400
cctcgcctgt	gtggaccggc	ggctcgtggt	gcccgttccc	gccggctggt	cctgggtccg	35460
cgccgcggcc	ctgccgtccg	tgttcgccac	cgctgggttc	gcgctgcgcg	atgtggcccg	35520
ggtgcggggc	ggggagcggg	tgtggtgca	cgccggggcc	ggtggtgtgg	gcagggccgc	35580
ggtgcgggtg	gcccggctgc	tgggcgccga	ggtgtatgcg	acggcagatc	ccggcaagca	35640
cgaggtgctg	cgggcggccg	gtctggacga	ggcgcgtgtg	gcgtcgtcgc	gggatacggg	35700
gttcgcgcag	cggttcccgg	agatggacgt	cgtgctgaac	tcctcacgg	gtgagttcgt	35760
ggacgcgctg	ctgcgactgc	tccgtcccgg	cggaagggtt	gtggaactcg	gcaagaccga	35820
ccgcgcgac	cccgcgggcc	tgccgggtgt	cgactacctc	ccgttcgacc	tgtgtctgga	35880
cgccggcccc	gaccgcgtcc	agagcctgct	gaccgaggtc	gtcgcccacg	ccgaggcggg	35940
cgagctgacc	gggctgccca	cccggacctg	gcccctcgcc	gacgcgcgca	ccgcgttccg	36000
gttcatggcc	caggcccggc	acaccggcaa	gatcgtgctc	accgtcgccc	cgtacgccga	36060
cggcaccgtc	ctgatcaccg	gcgcggcgt	gctcggcggc	atgctcgccc	ggcatctggt	36120
gtccgaacac	ggcgcctcgt	acctggtgct	ggccagccgg	cgccggcggc	ccgccccggg	36180
cagcgcggag	ctggtcgccc	aactggccgc	ggcggggccc	accgtccgct	tcgagaccct	36240
cgagctcacc	gacctcgccc	ccctcgacgc	cctcgtcgcc	aaagtgaccg	ccgagggccc	36300
gctgaccgcg	gtggtgcaca	ccgcggggcg	cctcgacgac	ggcgtgctga	ccgaactcgg	36360
cgccggcccg	ctgccggggc	tactgcgccc	gaaggccgac	gccgcctggc	acctgcacga	36420
actcaccgag	gacaaggacc	tgtccgcctt	cgtgctgttc	tcctccgccc	cggccaccct	36480
cggcaccccc	gcgcaggcca	actacgcggc	cgccaacgcc	ttcctcgacg	cgctcgccga	36540
acggcggccc	gccgcgggcc	tgccgggcac	cgccgcggcc	tggggcctgt	gggcccagcg	36600
caccggcctc	accgggcacc	tggacgccgc	cgacgtcgcc	cgcccccggc	gcaaccgcat	36660
cgtgccgatg	gccgcggccc	aaggactggc	cctgttcgac	acggccaccg	ccaccggcga	36720
cgccgtgacc	gtcaccgcgc	gcctggacct	cgccaccgcc	tcggcccggc	ccaccccggc	36780
gctgctgctg	ggcctggtcg	ccaccccggc	cgcccccggc	gccgcccggc	cggtgcccgg	36840
tgcggcggcc	ctggtggcgc	ggatcaccgg	gctgcccggc	cccgagcggc	ccccgcctct	36900
gctcgacgtc	gtccgcggcc	aggtcgccga	cgtcctggga	caccggggag	gcagcgggtg	36960
ggcaccggac	cgcggttca	aggaactggg	cttogactcg	ctgaccggcg	tggaaactcg	37020
caaccgcctc	ggcaccgcca	ccggcctgcg	actgcccacc	accatcgtgt	tcgaccaccc	37080
gaacccggcc	gcgctcgccc	accacctcct	cgacgcgctc	ctgccgagcg	agtccggcga	37140
ggcgagcgcc	gaccggatca	tcgcccgaact	ggcccgcgtg	gagagcgcgc	tcggcgcact	37200
gccggccgac	ggcaccgacc	gcgcccgggt	ggcggcgcac	ctgcgggacc	tcgcccggcc	37260
ctgggacgcc	ggcacgaccg	acggcacacc	cgagcggggc	gctctggacg	gcgtcaccgc	37320
cgacgaactc	ttcgacctga	tcgaccgggg	cgtctcgtga	cccgcgcccga	acgaaccgca	37380
cggggggtgca	ccgatggctg	acgaggccaa	actgctcgac	tacctcaagc	aggtcacggg	37440
cgatctccag	gtcgcccggc	gccggtgctg	cgaggccgag	gcccgcgacc	gcgaaccgat	37500
cgogatcgtg	ggcatggcct	gccgctaccc	cgggcggcgtg	acctcgcccg	aggacctgtg	37560
gcggtcgtg	accgatggga	ccgacgcggt	ctcggcgttc	cccggccgacc	gcggctggga	37620
catggcctc	gtctacgacc	cggaccccga	ccggcccggg	acctcctacg	cccgcgaagg	37680
cggattcctc	gacggcggcc	cggacttcga	cgccggcttc	tcggcatca	gcccgcgca	37740
ggcgctggcc	atggaccggc	agcagcggct	ggtgctggag	acctcatggg	aggccctgga	37800
gcgggctggc	atcgaccggg	ccggcctgcg	ctccaccggc	accggcgtgt	tcgtcggctt	37860
cagcagcgag	gactactccg	acatcaccgg	cccgggtggc	accgagctgg	aggggtacgt	37920
cgtcaccggc	acctcgccc	gcgtgctctc	cgcccccggc	gcctacacc	tcggcctgga	37980
gggccccggc	ctctccgtgg	acaccgcctg	ctcctcgtcc	ctggtcgccc	tgcacctcgc	38040
ggtccgctcg	ctgcgcggcc	gggagtgcac	gctggcgtc	accggcgggg	cgaccgtgct	38100
gtccacgccc	ggcgtgttca	ccgagtacag	ccggcagcgg	gccctcgcg	ccgacggccg	38160
ctgcaaggcg	ttcgccggcg	cgcccgaccg	gttcggcttc	gccgagggcg	cgggcatgct	38220
cgtcctggaa	cggtctcgg	acgcccggcg	caacggccac	cccgtgctgg	ccgtggtccg	38280
cggcaccggc	atcaaccagg	acggcgcgag	cagcggactc	accgccccca	acggactcgc	38340
ccagcagcgc	gtgatccggc	aggcgctggc	cgacgcccgc	ctggccgcca	gcaggctcga	38400
cgccgtggag	gcgcacggta	cgggaccccg	gctgggcgac	ccgatcgagg	ccgagccct	38460
cctcgccacc	tacggccagg	agcgggacga	gcccgtgctg	ctcggttcgg	tgaagtccaa	38520

catcggggcac	acgcaggccg	ccgccgggtgt	ggccgggtgtg	atcaagatgg	tgcaggccat	38580
gcgccacggc	acgctgccgc	gcacgctcca	cgtggacgag	gcgtcaccgc	acgtcgactg	38640
gaccgccggc	gccgtggaac	tgctcaccga	ggagcgggac	tggccgggcg	gggagcagcc	38700
gcgccgcgcg	ggcgtgtcgt	cgttcggcgt	gtccggcacc	aacgccccacg	ccatcctgga	38760
gcaggcacc	gcccggcccc	agcccaccga	cgaggccacg	gaccgcacc	tgccggtcgt	38820
cccgtggacg	ctgtcctccc	ggacggcggc	cggactgcgc	gcccaggccg	aacgcctcct	38880
cgcccaccgg	cccgcgcacg	acgccgcccc	ccacgacgtg	gccctcgccc	tcgccaccac	38940
ccgtacggcc	ttcgaacacc	gcgtggtcgt	cctcggcgcg	gaccacgaca	ccctgctcgc	39000
cggcctgacc	gccgtggcgg	aaggcgccga	gtcggccgac	gtggtgcgcg	gccgcgccgt	39060
cggcgacggc	cgcgccgtct	tctgttccc	gggtcagggg	gcgcagtggg	tggggatggc	39120
ggtggatctg	ctggattcgt	cgccgggtgtt	cgccgggccc	atggctgagt	gtgccggggc	39180
gttggagccg	ttcgtgggct	ggtcgttgcg	tgggtgctg	ggtgatccgg	tggcgttggg	39240
gcgggtggat	gtcgttcagc	cggtgttgtg	ggcgggtgatg	gtgtcgttgg	cggaggtgtg	39300
gcgttcgtac	ggtgtcgtgc	cgtcggctgt	ggtcggtcac	tcgcaggggg	agatcgcggc	39360
tgcttgtgtg	gcgggtgtgt	tgtcgtcggc	ggacggcgcc	cgggtcgtcg	ccctgcgcag	39420
cagggctttg	acggcgctgg	cgggcagtg	gggatgggtg	tcggtcgctg	ccgggccctc	39480
cgggtgcgag	gagctgttgg	tgggctgggc	gggccgggtg	gcggtggctg	cggtaaccgg	39540
tcctgagtcg	gtcgtcgtcg	cgggtgaggg	ggtggcgctg	gaggagtcc	tggctcattg	39600
cgggggtcgg	ggggtgcggg	cccggcgtat	cgctgtcgc	tatgcctcgc	attcgggtgt	39660
ggtggagccg	gtgcgtgagg	ctctgctggc	tgatctggag	ggcgtgcggc	cgggtgaggg	39720
cacggttccg	ctgttctcca	ccgtgacggg	ggagtggggc	gacggcacgg	cgtcggacgc	39780
cggctactgg	taccgaacc	tgcgcgaacc	cgtccgctac	gccgacgccc	tcgccccct	39840
gctcggccag	ggcacaacc	gcttcacoga	ggtcagtcgg	caccccgtac	tgaccgtcgg	39900
caccagggag	accgtcgagc	gcacggagac	cgagaccggc	gtcgtcggca	ccctccagcg	39960
cgaccggggc	ggcctgccga	ccctgctgac	caacctggcc	gaggcgcaca	cgcacggcgt	40020
gcgcatcgac	tggaccgect	tcttccgggg	caccggagcc	cacccggccc	acctgccgac	40080
gtacgccttc	cagcgcgacc	ggtactggcc	gcccgtggac	gcagcggccc	gcccgtgccc	40140
cccggcccgc	accccctcgc	cgctccccgc	gcagcggacc	gccgacgagg	gcacgccctg	40200
gcgcgacgag	ctggccgggc	tgacccccgc	cgaacgcccg	caccgcgtca	tggacctggt	40260
acggctgcgg	accgccggcg	tgctcggcca	cgccgacccc	gccgcgatcg	aaccgcaccg	40320
cggtttcgcc	gccctcggct	tcgactccct	cgccctccctg	cggctgcgca	ccgcgctgac	40380
cgaggcgacc	ggcctgacgc	tggccagcag	cgtcgtcttc	gaccaccccga	acccggccgc	40440
cctcaccgac	catctgetca	cccggctcga	cggcgcgccc	gtcccggccc	gcccccgct	40500
ccgcgccacc	ccggccgacg	agccgctggc	cgtggtcggc	atggcctgcc	gcttccccg	40560
cggagtggcc	tcgcccgacg	acctgtggca	gctcctcgcc	gaaggccgtg	acgtcatcgg	40620
cgacttcccc	gccgaccggc	gctgggacct	cgacggcctc	tacgaccccg	acccggaccg	40680
ggccggacac	acctacctgc	gccagggcgg	cttccctgac	gacgccgccc	gcttcgacgc	40740
cgacttcttc	ggcatcagcc	cgcgcgaggc	gctggcgatg	gacccgcagc	agcggctggt	40800
cctcgaagtc	gcctgggagg	cgttcgagca	cgccggcgtc	gatccgcacg	gcctgcgagg	40860
cagcggccacc	ggcgtcttcg	cgggcgtcac	cgaccagcgc	tacgactcac	ggcacggcgc	40920
ggtcggccggc	gtggacgagg	ggctgctcgg	caccgggaac	tacgccagcg	tgctctccgg	40980
ccgggtggcc	tacaccctcg	gcctggaagg	cccggccgtc	agcgtggaca	ccgcctgctc	41040
gtcctcgtc	gtcgcctgc	acctggccgg	ccagtccctg	cgcagcggcg	aatgctcgt	41100
ggccctcgcc	ggcggcgtga	tggatcatgtc	gaccccgctg	gccttcgtgg	agttctccc	41160
gcagcgcggc	ctcgcccccg	acggccgctg	caaggcggtc	gccgccgccc	cggacggcat	41220
cggctggtcc	gagggcgccg	gcgtggtcgt	gctggagcgg	ctgtccgacg	cccgccgcaa	41280
cggccaacccc	gtgctggccg	tggctccgcg	ctcggcgatc	aaccaggacg	cgccctcctc	41340
cgggctcacc	gcccccaacg	gctccgcccc	gcagcgggtc	atccgctccg	cgctggccgc	41400
cgccggcctc	accgccggcg	acgtcgacgt	cgtggaggca	cacggcaccg	gcaccacct	41460
cggcgacccc	atcgaggcca	acgccctgct	ggccacctac	ggccaggacc	gcgcacagcc	41520
gctgcggctg	ggctcgtga	agtccaacat	cggccacacc	ggccccgagg	cgggcgtggc	41580
cggcgtgatc	aagacgggtc	tcgcgctgcg	cgccggcgaa	gtgccccgca	ccctgcacgt	41640
ggacgcccc	tcaccgcaca	tcgactgggc	ctccggctcc	gtcgaactcg	tcaccggcgc	41700
cgccgtctgg	cccccgaccg	accggccggc	ccgcgcgggc	gtgtcgtcgt	tcggcgccag	41760
cggcaccaac	gcccacctca	tccctggagg	ggccccggcg	gaaccggccg	cgcccgcggc	41820
cggcgccggc	ccgcgctggg	tgccgtggct	gctctccgcc	cgtcccagg	ccgccctcgc	41880
cgagcaggcc	accggctcgc	ccgccacct	cgacacccac	cccggcctgg	acccgggtga	41940
cctctcctgg	tcccttgccg	cgtecggggc	ggtgttgccg	catcgtgccc	tggtgggtggc	42000
cggatgatgtg	gccggtgcgc	gggcttctct	ggggcggttg	gccgcggggg	agccgggtga	42060
gggtgtggtg	tcgggtggtg	ccgggttgc	gggtgagggc	cgggtggtgt	tcgtgttccc	42120

gggtcagggg	gcgcagtggg	tggggatggc	ggtggatctg	ctggattcgt	cgccggtggt	42180
cgcggggccc	atggctgagt	gtgcggtggc	ggtggagccg	ttcgtgggct	ggtcgttgcg	42240
tgggtgtgctg	ggtgatccgg	tggcggttga	gcggggtgat	gtcgttcagc	cggtggttgg	42300
ggcggtgatg	gtgtcgttgg	cggaggtgtg	gcgttcgtac	ggtgtcgtgc	cgtcggctgt	42360
ggtcggtcac	tgcagggggg	agatcgcggc	tgcttgtgtg	gcggtgtgtc	tgtcgtggtc	42420
ggacggcgcc	cggatcgtcg	ccctgcggtc	gcggtccatc	gccgagcggc	tggccggggc	42480
cgggcgcatg	gtctccgtcg	ccctgcccga	ggccaccgtc	acacggcggc	tgcaccctg	42540
gtccggacgc	gtttgctgctg	ccgccctcaa	cggcccgtcg	tcggtggtcc	tacggcgga	42600
ccccgatgcg	ctggacgagg	tgatggccgc	ctggtcggcc	gacggggtac	gactgcgccc	42660
catcgcggtc	gactacgctc	cccactcggc	ccatgtcgag	tccctcgaag	ccgagtgtcg	42720
cgccgcgctc	gcccgaactgc	ggcccgggtga	ggccggcatc	gccttccact	ccaccctgct	42780
gggcggcccc	ctcggcgaga	cacggctcga	cgccggctac	tggtaaccga	acctgcggga	42840
gcccgtacgg	ttcgagcccc	tgggtgcgcg	cctgctcgac	tccggccacg	ccgtgttctg	42900
ggagatcagc	ccgcacccgg	tgctgacggc	cgccgtgcag	gagacggcgg	aggccaccga	42960
gcggaaccgccc	gtcgtcgtcg	gcaccctccg	ccgocagcag	gacggcccgc	gcccctcct	43020
cacctcgtc	gcccaggccc	ccgtccacgg	cgtgtccgtc	gactggacgg	tgggctgccc	43080
cgacggacgc	cacgtcgacc	tgccaccta	cgccttccag	cgccgcccgt	tctggcccag	43140
cggccccggtc	gcctccgacg	ccgcggcct	gggectgacc	ggcggccggc	accgctgct	43200
cgggcggctc	accccgtcctg	ccggatccgg	cggacaggtg	ttcaccgggc	gcgtgcccgc	43260
cggcaccgag	ctgcccggccg	gcgcccact	cgacctggcc	ctgcaacggc	ccgacggcag	43320
gacgctcggc	gaactgaccg	aggaggcacc	gctcgacacc	gtcggcgaag	cccggcgcct	43380
ccaggtcacc	ctgggocggg	agaacgagga	cgggcggccg	ccggtcggcc	tccctcctcc	43440
ccccgcggc	gcccgaactgc	accagccctg	gaccggccac	gcgaccggcg	tcctcctccc	43500
gcacaccggc	accgccccgg	ccgcggcgc	aggggaccgg	gacaccgacg	tgaccgtcga	43560
gctgaccgac	gtgaccgttg	agctgaccga	cgagaccggc	ggcggctggg	gcgtacacc	43620
cgccctgctc	accgacgccc	tgcgggccc	acaccccggc	ctggtccccg	gcgcccggca	43680
cggcgtcacc	ctgcacgccc	tgggcccacc	ccggctgccc	gtacgcccgc	accccgggg	43740
cgagcacacg	gtcaccctgc	acgccaccga	cgacagcggg	gccgcccgtc	taccgctcga	43800
cgcggtgacc	ctgcgcccgc	cggccacggc	cggccccggc	cgcgaccacg	agctgtaccg	43860
ggtcgagtg	acgcccgtgc	cgtgcccgc	ggccgacgtc	gacggcatcg	ccgtgctcgg	43920
cggccttccg	ctgcccggat	accctcctg	ccccgacctc	gccgcccgtc	ccgccctgga	43980
caccgtgccc	ggcaccctcg	tgctgcccgtg	tgggcagcag	tccggccgacg	acaccgcccg	44040
cgcggcccac	gcccggcggc	ggcgcgtgct	gaccgctctt	caggagtggc	tccgcccagca	44100
gagactggcc	ggcaccctgc	tggccgctcct	caccacggc	gcggtcccgg	tcgagcacga	44160
ggagctcacc	gacctcggcc	acgcggcccgt	gtggggcctg	atccgctccg	cccggcccga	44220
acaccccggc	aggctggtgc	tggtcgacat	cgacggcccc	gacgcaactgc	aacagctgcc	44280
cgtgtcctc	gccaccggcg	aaccgagat	cgccgtgccc	tccggcccggg	tcctcgcacc	44340
ccggctggtc	cgcgccccgc	gcaccgacga	gcccacgagc	caggcggccc	tgtggcccgtc	44400
cgacggcgtc	gtcctcgtca	ccggcggcac	cgccacgctc	ggcggcccctc	gcgcccggca	44460
cctggtcacc	gagcacggcg	tacgcccggct	ggtcctggcc	ggacggcggc	gcgacgccc	44520
ccccgacgccc	gtgaccctcg	ccggcgaact	gcgcccctc	ggcggccgag	tgaccgtcgc	44580
cgcctgcgac	gccgcccacc	gcaccgaact	ggcccggcctg	ctcgaccgga	taccgcccga	44640
gcaccccctg	accggagtgc	tccacgcccgc	cggcgtcctg	gacgacggcg	tcatcgcctc	44700
ccagaccccc	gagcggctcg	ccaaggtgct	ccgccccaa	gtcgacgccc	cctggaacct	44760
gcacgaactc	accgccccgc	accggcccggc	cgtcttctg	ctgttctcct	ccacctccgg	44820
cctcttcggg	gcgcccggac	agggcaacta	cgccggggc	aacgccttcg	tggacgccc	44880
cgcggcacac	cgccgcccgc	agggcctgccc	gcgccctcc	caggtgtggg	gctgtgggc	44940
gcacggccagc	ggcatgaccg	gccacctggg	cgccgcccgc	ctgcccggcg	cgcccggga	45000
cggcgtggtg	ccgctgcccga	cgcccgacgc	gctcggccc	ttcgaccgcg	ccaccgccc	45060
ctccgcccc	gtcgtcgtcc	ccgcccggct	ggacctcacc	tctgttcgcca	ccggcgtcac	45120
cgccgtaccc	gcgctgatgc	gcccggctggt	ccgtggccc	gtcccggcgg	ccgccaccgc	45180
gggcgcccga	ccggacaccc	tggcccga	gctcaccgga	ctcaccgccc	cggaacgcga	45240
acggaccctg	ctcacgctcg	tccgctcgca	cgccggcgtg	gtcctcggcc	acaccgacga	45300
caccgcccgtg	accccggccc	ggcgcttcaa	ggaactgggc	ttcgactcgc	tgaccgccc	45360
cgaactccgc	aaccggctgt	ccgcccacc	cgggctgccc	ctgcccggca	ccctcgtggt	45420
cgaccacccc	aaccgcccgt	ccctggccc	ccacctgctg	gccgaactcc	tccggcgaacg	45480
cgccgaggag	acggcggccc	tgcccggcgt	cgcccggccc	gcccacgacg	atcccacgc	45540
gatcgtcggc	atggcgtgccc	ggtaccccgg	cggcgtgacc	tcccccggg	acctgtggcg	45600
gctgctcggc	gacgagcggc	acgcgctcac	ccgctcccc	gacgaccggg	gctgggacct	45660
cgcgggccc	ttccaccggg	atccggagca	cgcgggccc	tctgactgctc	gggagggcgg	45720

cttcctcgcg	gacgtcgccg	ggttcgacgc	ggaacttcttc	gggatctcgc	cgcgcgaggc	45780
gctggcgatg	gacccgcagc	agcggctggc	cctggagacc	gcctgggagg	ccgtcgaacg	45840
cgccggcatc	gacccgaagt	ccctgcgcg	cgcggaagtc	ggcgtctacc	tcggcaccaa	45900
cggccaggac	tatgccagc	tcgtgcgag	gaccgtcgag	agcgcggagg	gctacgtcgg	45960
catcggaac	tccgccagcg	tgctgtcggg	ccgcacgcc	tacgtcctcg	gcctggaagg	46020
ccctgcggtc	accgtggaca	ccgcgtgctc	ggcgtcgctg	gtcgccctgc	actgggcgat	46080
ccaggcgctg	cgagcggcg	agtgtcgat	ggcgtggcc	ggcgtgtca	ccgtcatgtc	46140
cgccccgac	gtcttcgtcg	acttctccc	ccagcgcggc	ctggccgtgg	acggccgctg	46200
caagtccctc	gccgccgcg	ccgacggcac	cggctggctc	gagggcgccg	gcatgtcctc	46260
cgtggaacgg	ctctccgacg	cccgcgcga	cggccaccag	gtactggccg	tcgtacgcgg	46320
caccgcgatc	aaccaggacg	gcgcgtcgaa	cggctcgacg	gccccgaacg	gtccttcgca	46380
gcagccgggtg	atccgtcagg	cgctctccgg	tgccgggctc	acgcccgcgg	acgtggacgc	46440
ggtggaggcg	cacggtacgg	gcacgcggct	gggcgaccgg	atcgaggcgc	aggcgtgct	46500
cgccacctac	gggcaggagc	gggacgagcc	gctgtggctc	ggttcggtga	agtccaacat	46560
cgggcacacg	caggccgccc	ccggtgtggc	cggtgtgatc	aagatggtgc	aggccatgcg	46620
gcacggcacg	ctgccgcgca	cgctccacgt	ggacgagccg	tcacctgagg	tcgactgggc	46680
ctcgggcgcg	gtggagctgc	tgaccgaggc	ccgggagtgg	gcgcgtgccg	gtcggcccg	46740
ccgtgccccg	gtgtcgtcgt	tcgggtgtgc	cggcacgaa	gcgcatgtgg	tgctggaaga	46800
ggctcccag	gagaccgggc	gcaccgcgcc	ggctgccccg	cccgccgtgc	cctggctggt	46860
gtcgccccgc	tccgagaccg	cgctgccccg	acaggtcgag	cggtgccccg	cgtacgtcac	46920
ggagcaccgg	gagggcgccc	cggccgacat	cggcctcagc	ctcctcaccg	cacgtccccg	46980
cttcgagcac	agcgcctcgc	tcgtcggcac	cgaccacgac	gcactcctcg	ccgcgtggc	47040
ggaccgggag	ccgctcagcg	cacgtgacgg	cctgaccgcg	ttcgtgttcg	cggggcagg	47100
tgccgagcgg	gtggggatgg	gggcggaact	ggctgccccg	tatccggtgt	tcgctcaggt	47160
gttcgccccg	gtgtgtgccc	cgttcgacgg	tggtgccccg	cgcccgctgg	gtgaggtggt	47220
cgccgagggg	ggtccggagc	tggtcggac	ggtttatcg	cagggcgggt	tgttcgcttt	47280
cgaggtggcg	ctgttccccg	tgctggagtc	gtgggggtgc	gcgccgatg	tggtgctcgg	47340
gcattccgtg	ggtgagttga	cggctgcctg	tggtgggggt	gtgtggtcgc	tcgaggacgc	47400
ggtgccccg	gtggctgccc	gtggggcggt	gatgcagggc	tgcccgagg	gcggtgcat	47460
ggtcgctctc	gaggtgtcgg	cgggtgagct	ggagtgtccc	gtgggtgtgg	agctggccgg	47520
ggtcaatggc	ccgacgtcgg	tggtgctgtc	cgggtgaggg	gagggcgtcc	tggtgaggc	47580
cgccccgctg	ccggaccgtc	gtgccaaagg	tctgccccg	agtccagcct	tccactcgca	47640
tcggatggat	ccgatgctgg	aggacttccc	gcgggtactg	gagtccgtga	ccttccaggc	47700
tcgggaggtt	gccttcgtgt	cgacggtgac	cggctgctcg	gtgacggacg	agctgtgca	47760
ccccgggat	tggtgcccga	acgtccccga	gaccgtcccc	ttcgccgacg	ccgtccccg	47820
cgtgccccg	accggtgccc	acgccctggt	cgagatcggt	cccgatgccc	ccctcgctcc	47880
gctgccccg	ggcgggtgtc	cgctgctgcg	caggggcccc	ccggaggccc	tcgctgctgg	47940
cgaggggctc	gccccgtccc	agatggccgt	cgactgggag	aagttcttcg	gcgccccggg	48000
gagccccgtc	gacctgccc	cgatccggt	ccagcacagc	cggttctggc	cggaggtccc	48060
ggtgccccg	gagggcgatc	cggccgcccc	gtggcgctac	cgggtgcccc	gggagccccg	48120
ccgtccccg	gtggacgccc	tgcccccccc	gcctgccccg	gtggccccg	ccgccccggg	48180
cgccggatg	gctctggtgc	gggagtgcct	gaaggggctc	gccggccccg	gagtggagg	48240
ggccgagctc	gcgggtggcg	gtgccccgga	ccgtgccccg	ctgccccgag	cggtgtcggc	48300
cacgggcccc	gtggacggcg	tgctgtccct	gctgccccg	ggtcctgacg	cggccggccc	48360
caccctcgtc	ctgccccagg	ccctcggcga	cgccccgacc	gacgcgccc	tggtgtcgg	48420
gaccggcgga	gcccgtcgcg	ccgccccagg	agaacccccg	gacccggagc	aggccccggt	48480
ctggggcctc	ggccgggtgc	tggtcctgga	ggccgccccg	cgctggggcg	gcctcgtcga	48540
cctgccccg	gagccccagc	agcaggcggt	cgtaacagctg	tgcccccgtc	tcgccccgca	48600
cgacgacgag	gaccaggtcg	ccgtgccccg	ggccccgggtc	ttcgggcggc	ggctggtccg	48660
ccccggccag	ggcggccccg	cggagggctg	ggccccgtcc	ggcaccgtgc	tgatcaccgg	48720
cggcacgggc	gcgctggggc	ggcacaccgc	ccgctgggca	gccccgctcg	gcgccccggc	48780
ggtggtgctg	gtgagccccg	gcggaccccc	cgcggacggc	gcggccgga	tcgtggcggg	48840
actcgccgaa	ctgggctgcc	aggccgtcgc	ggaggcgtgc	gacgtggccc	accgggcccc	48900
gctcgccgcg	ctgctcgaca	agatccccg	cgacggcccc	ccgctgaccg	cagtgggtga	48960
caccgcaggc	gtcctggacg	acgggggtga	cggctcctcg	acccccgaa	ggctggccc	49020
ggtgctgccc	gccaaggccc	acggccccac	cgccctgat	gaactcaccg	cccacctgcc	49080
cctggaggcg	ttcgtcctgt	tcgccccgct	cggccccggg	gtcggcggcg	cgggcccaggc	49140
caactacgcc	gcccccaacg	cccacctcga	cgcgctcccc	gcgagccccg	gggccccgg	49200
gctgccccg	accgcggtgg	cgtggggcgc	ctgggccccg	ggcgggatgg	ccgacgccc	49260
cgaggtacgc	cggccctggg	accgccccg	gctgctgccc	atggaccccc	ggcgggctc	49320

ggacgctctc	ggccgggaga	tcgcgcccg	cgaccggcc	gtggtgctcg	ccgacgtcga	49380
ctggaccccg	ctcgcgccca	acctgcacgc	cgtccggccc	agcccgctga	tctcgaccgt	49440
gcccgaggca	cgccgggccc	tcgcgcccga	acccggcacg	gcccggggcg	agtccgatcc	49500
gcggcagcgc	ctggcgggccc	tgcccgcagc	cgaaacgcgc	cgcgtgctgc	tcgacctcgt	49560
ccgcgacgcc	atcgccgccc	tgctcgggta	cggcggtccc	ggcgcggctc	acatcacgcg	49620
gggcctcgtc	gacctgggct	tcgactcgct	gaccgcccgc	gaactgcgca	accggctcgc	49680
ccggggccacc	ggactcgccc	tgccgctcac	gctggtcttc	gaccaccccg	acggggaggc	49740
gctcgcccgc	cacctcgccc	ccgcaactgc	gcccgcggcg	gaccggcccgc	agaccaccga	49800
cctcgaccgg	ttcgcccggc	tcgaccgggc	cacggccgac	gagccgaccc	gggtccgggt	49860
cgccgcggag	ctcgccgggc	tgctcgacgc	gtggacggcg	ccaccggccc	cgcccggccg	49920
gtccgcccgg	ttcgccgggc	accccagccc	gctgaccacg	gcgtccgccc	acgagatctt	49980
cgacttcatt	cggaacgaac	tcgggaagtc	ctgaacgtgg	agaacgagaa	caagctcctc	50040
gaccacctgc	ggtgggtcac	cggagaactg	gcccaggcac	ggcagacgct	gcgcgagacg	50100
gccgaacgtg	ccaccgagcc	gatcgcgatc	gtcggcatgg	cctgccgggt	cccggggcggc	50160
gtgagcacgc	ccgaggacct	gtggcggtcg	ctcgccgacg	agcgtgacgc	ggtcggcgag	50220
ttccccggcg	accgcggtcg	ggacctggcc	tcgctgcacc	acccggaccc	cgagcacgcg	50280
ggcacgtcgt	acgtgcggga	gggcggttc	ctcgcgacg	tcgcccgggt	cgacgcggac	50340
ttcttcggga	tctcgccggc	cgaggcgctg	gcgatggacc	cgcagcagcg	gctcgccctg	50400
gagaccgcct	gggaggccgt	cgaaacgcgc	ggccttgacc	cgaaagtccct	gcggggcgcg	50460
gacgtcggcg	tgtacctggg	caccaacgga	caggactaca	tctccgcccgc	gcgcccgctg	50520
ctgcaccagg	tgaggggcca	cgggcgccacc	ggcatcgccc	gaagcgtgct	gtccggacgc	50580
tcgcctaca	ccctcgccct	gcgcccggccc	gcggccaccg	tcgacacggc	gtgctccgcc	50640
tcctcgtcgc	ccctgcactg	ggcgatgcgg	gcgctcgccc	gcggcgagtg	cccgatggca	50700
ctcgcgggcg	gtgtgacggt	catgtcgtcc	ccggcgacct	tcgtggagtt	ctcccggcag	50760
cgcgccctcg	cctccgacgg	ccgctgcaag	ccgttcgccc	ccgcgccgga	cgccaccggc	50820
tggggcgagg	gcgcccgaat	gctgctgctg	gagcggtctt	ccgacgcccg	gcgctcggc	50880
caccgggtgc	tcgcggtgct	gcgcccgtcc	gcccgtcaacc	aggacggcgc	cagcagcgc	50940
ctcaccgccc	ccaacggccc	ggcgcaggtc	cgcgtcatcc	aggcggcgct	cgccgacgcc	51000
cggctcgcgc	ccgcccagct	cgacctgctg	gagggcgacg	gcaccggcac	cgtcctcggc	51060
gacccgatcg	aggcccaggc	cctgctcacc	gcctacggcc	agaaccgcac	cgcccccgcc	51120
tggtcggct	cggtcaagtc	caacgtcgga	cacacccagg	ccgcgcccgg	ggtcgccgga	51180
gtgctcaaga	ccgtgctcgc	cctgctcacc	ggggtgctgc	cccggaccct	gcacgtggac	51240
gcgcccaccc	cgaaggtcga	ctggtcggcg	ggcgcggtac	ggctgctcac	cgaggcccgg	51300
ccgtggcccg	ccggcgagcg	gcccgcggcgc	gcccgtgctt	ccgcttcggg	tgtcagcggc	51360
accaacgcgc	acgtcatcgt	ggagcaggca	cccagacccg	ccgcccaggc	gtccggcgcg	51420
gcccgggggg	cggtgccgtg	ggtgctgtcc	ggccgtaccg	agagcgcgct	gcggggccag	51480
gcccgcgcgc	tcgcccggca	tctggcgagg	cgcccggcg	accggcccgg	ggacgtcgcc	51540
ctgtccctgg	ccaccacccc	cccggcgctt	gaacaccgcg	ccgtcgtggt	cgggacggac	51600
gtcgacgacc	tgctgcggcg	cgtggcccgc	gtcgccggcg	gagagcccac	gcccgggggtg	51660
gtccgcggca	ccgcccggca	cctcggcccg	atcgcttcg	tgttccccgg	gcagggctcc	51720
cagtgagccg	gcatggcccg	cgaaactcgc	gacgcctcaa	cggagttcgc	cgcccgcctg	51780
gacgagtgcg	ccgcccgcct	cgccccgcac	gtcgactggt	cgctgcgcga	cgtcctggcc	51840
gacccgggcg	ccctggagcg	cgtcgatgct	gtccagcccg	ccctgtgggc	cgtgatggct	51900
tccttgcccg	ccctctggca	ggcgcacgga	gtggagccgg	ccgcccgtcgt	cggccactca	51960
cagggcgaga	tcgcccggcg	gtgctgcgcc	ggcccccctt	ccctggagga	cgcccgcctg	52020
ctggtcaccg	cgcgcccggc	ggcgtgcgc	gcccgtatcg	gcccggggcg	catggtgtcc	52080
ctgcccgtgc	cgagggcccg	cgcccgcgaa	gtcatcgacg	cgtggggcga	ccggtgtcgc	52140
gtggcccggc	tcaacggccc	cgcccgcgaca	gtcgtgtccg	gcgacgcccg	cgccctggac	52200
gaactcctcg	cccaggccga	gcgggacggc	ctgcccggca	agcgcctccc	cgtcgactac	52260
gcctcgcact	ccgcgcacgt	cgaagccgta	cgcccggagg	tactggcccg	gaccgcccgc	52320
gtcacccccg	gcgcccacgg	caccgccttc	gtgtcctcgg	tgaccggcgg	attcctggac	52380
accaccggtc	tggacgcccg	ctactggtac	cgcaacctgc	gcgagccggg	ccgtttcgac	52440
cgggcggggc	gggcccctgct	ggacgcccgg	ttcacgacgt	tcacgaggt	cagcgcgcac	52500
cccgtgctca	ccgcccgaact	ccaggagagc	gaccggggcg	tcctcgcccg	cggtcgtcgt	52560
cgcccgcagc	acggcgggccc	cgcccgcttc	ctcgcccctc	tcgcccgaggc	cgcggtccgc	52620
ggcgtaccgg	tcgactggcg	accggcccctg	accgcccggc	gcaccgtgcc	tcttcccacc	52680
tacgccttcc	agcgcgaacg	cctctggctg	gaggaggacg	ccggggccgat	caccccggaa	52740
cgggcccggc	aggacagcgc	gttctggggc	gcgggtccacg	gctccgcccga	cgacctggcc	52800
gcccgtgctcg	gcgcccgaacc	cgcccggccc	gagccgatcg	ccgcccgtca	cccgctgctc	52860
gcccgcctggc	acgagcggcg	cgcccgaagc	tcgcccgtgg	actcctggcg	ctaccgctc	52920

gactggaccc	cgctcacgc	gtcgccctcc	gocggcgcgc	ccaccggccg	ctggctggcg	52980
gtcgcggccg	acgacccctg	gtccgacgcc	gtcgtcacgg	cactggcggg	ccgcatggac	53040
ctcaccocgc	tcccggccta	cgacccccgc	accgtcgcgg	aactcacgc	cgggaccgcg	53100
ggcgtcgtcg	ccgtgctccg	ccccgacgac	cccgacccgc	tcccggcct	cctcgccctc	53160
ctccgggccc	acgaccgcgc	cgacggcggc	ctgcccgtgt	ggtgctcac	ccgcggcgc	53220
gtggcgaccg	gcccggcggg	caccccgccg	gacccggcca	ccgcgagat	ctggggcttc	53280
ggccgcgctg	ccgctgga	acgaccgcag	agctggggcg	ggctgatcga	cctgcccggc	53340
ggcgccaccg	tgaccgggga	cgctgctggc	ggcacggccc	gggcccagcg	cccaccggac	53400
gcccgcacatg	ccgacgacgt	cctcggcgcc	cgtctcgcgg	accggctcgt	cgcccgctg	53460
gcgggtccgg	aggaccaggt	ggcggtcgcg	gectcggcg	ccttcggccc	ccggctgccc	53520
cgcgcgcccc	cagacgggac	cacccccgac	tggaggccca	ccggaaccat	cctcgtcac	53580
ggcgccaccg	gtggtctcgg	cgggcacgtg	gcccgtggc	tcgcccgtgc	cgggcgggag	53640
cacctcgtgc	tcaccgggcg	acgcggcccc	gaggcgcccc	gcgcgcgga	actggccgc	53700
gaactcaccc	cgctcggcgc	cgaggtcgac	atcgtcgcct	gcgacgcgc	cgaccgtgac	53760
gcaactggccc	gggtgctggc	cgagcaccac	gtcgacgcgg	tgttccacct	cgccggcatc	53820
gagcgtacc	ggccgctgga	cgagctgacc	ccggccgatc	tcaccgaggt	cgccggccgc	53880
aaggtcaccg	gcgccctcct	gctggacgaa	ctgacccgcg	accggcggct	gtcccgcttc	53940
gtgctgttca	cctcgggccc	cggagtgtgg	ggcagcagcg	gccaggcccg	ctacgccccc	54000
gccaacaccc	gcctcgacgc	gctcgcgcgc	cggcgcgcgc	ccgagggcct	gcccggcacc	54060
gccgtcgcct	ggggacactg	ggacggcgcc	ggcatgtccg	acggcccggc	ccgcgcccag	54120
atgctccgct	ggggcctgcc	cgccatggcg	ccccaccgcg	ccgtcgcgc	cctcgcgggc	54180
gcctcgcct	gggacgagc	ctcgtggtc	gtcgccgaca	tcgactggcc	ggtgttcgcg	54240
ccgacgttca	ccctggcccg	gcgcagcccg	ctcctcgcgc	gcctgcccga	ggccaccgac	54300
ccggccgaca	ccgccccgcc	gcccgcgcgc	ccggccggcc	ggctggaccg	ggcccgctg	54360
ctggagctgg	tgctcgcgca	gaccggcgcc	gtcctcggcc	acgcctccgg	caccggactg	54420
cccgcgcgcc	gcgccttcca	ggaaactggc	ttcgactcgc	tcaccgccct	ggaaactgcg	54480
aaccggctgc	gcgacgcgac	cggactgccc	ctgcccgcca	ccctcgtctt	cgaccatccc	54540
accccgacgg	cgctcgcgca	gcgcctggcc	ggggaactca	ccggcgaaca	gcccggcgac	54600
gttccggaca	ccacgcgcgg	gcccggccgc	gcccgcgacg	acgacccggt	cgatcatcga	54660
ggcatgggct	gccgcttccc	cggcggcgcc	gactcgcgcc	aacggctgtg	ggacctcgtc	54720
gccgcggggc	gggacgcgat	gaccggcttt	cccgcgcgac	gcggctggga	cgatcaccgac	54780
accggctaca	ccaaggccgg	cggattcctc	gacgcgcgcg	gtgagttcga	cgccgggttc	54840
ttcggcatca	gcccccgca	ggccaccagc	atggaccgc	agcagcggct	gctgctggag	54900
acggcgtggg	agacggtcga	gcgcgcggcg	atcgaccgcg	cctcgtcgcg	cggcagcccg	54960
accggctgtg	tcgtcggcta	cggcggccag	gactacctca	ccagcctgta	cgccacccc	55020
gaggaactac	agggccacct	cctcacccgc	acctccggca	gcgtggtctc	cgcccggttc	55080
gcctacgtcc	tcggcctgga	gggccccgcg	gtcaccatcg	acaccgcgtg	ctcctcgtcc	55140
ctggtcgcgc	tgcaactggc	gatccgcgcg	ctgocgtccg	gcgagtgca	cctggccctg	55200
gcaggcggcg	tgaccgtgat	ggccaccccc	ggcgtcttcg	tcgagttcgg	ccgacagggc	55260
ggcctggccg	gcgacggccg	ctgcaaggcg	ttcgccgacg	ccgcccagcg	caccggctgg	55320
ggagagggcg	ccggcctgct	cctggtggag	cggtgtccg	acgcccggcg	aaacggccac	55380
cgctcctcgc	ccgtcgtccg	cggtccgc	gtgaaactccg	acggcgcctc	caacggcctc	55440
accgcgccc	acggcccctc	ccagcagcgg	gtcatccggg	ccgcgctggc	cgacggcggc	55500
ctcaccocgc	ccgacatcga	cgccgtggag	gcacacggca	ccggcaccgc	cctcggcgac	55560
cccacgcagg	cccaggccct	acaggtggtc	taocggcagg	accgcgagcg	cccgtgtgg	55620
atcggtcgg	tcaagtccaa	cctgggccc	acccagcccg	cctccggcgc	cgccggcctc	55680
atcaagacgg	tccctggca	gcggcacggc	gtgctgcccg	ccaccctgca	cgctcgaccg	55740
ccctccgcac	aggtcgactg	gaccagggga	gggtgtccg	tcctcaccga	gagcactccg	55800
tgccccggga	ccgacgtgcc	gcgcgcgcgc	gccgtgtcct	ccttcggggg	cagcggcacc	55860
aacgcccatg	tcgtcctcga	acaggccccg	cccgcggggg	aagccggacc	ggccggggac	55920
ggcggaaccg	tcccgtggct	ggtctccggc	cggacgcccg	aggccgtacg	cgcccaggtg	55980
gagcggctgc	gcgccagcct	cgccggacac	cccgaccccg	tcgcccgtcg	ccgcgcccctc	56040
gccaccacgc	gcaccgcgct	cgagtaccgt	gtcgcgcgcg	cgggcggcga	caccgagggc	56100
ctgctcgacg	ccctcgcgca	ggcccgaacc	gtcaccgccc	gccagggccg	taccgcgttc	56160
ctgtgcaccg	gccagggcgc	ccagcgcgc	ggcatgggtg	ccggcctgta	cgccgcccac	56220
cccgtctacg	ccgacgcctt	cgacgcgcgtc	tgcgcgaggt	tcgaccgcct	gctggaccgc	56280
ccgctgcgcg	acctggtgct	gtccggcccc	gocgacgtac	tcgaccgcac	cgectacgcg	56340
cagcccgcgc	tgctcgcctg	cgaggccgc	ctcgcgcgca	tgctgcgcca	ctggggcggtg	56400
acccccgacc	tgctggccgg	ccactcgtc	ggcgagatca	ccgcccgcga	cctggccgga	56460
gtcctctccc	tgcccggacgc	cgccggcctg	gtcgcgcgc	gcggccggct	gatggacgcg	56520

ctgccggccg	gcgggcgcgat	ggtcgccgtg	gaggccgacg	aggaccgtgt	acagcccctg	56580
ctgggcgacg	acgtctgcct	cgccgccgtg	aacggccccg	gcgccctggt	gctgtccggc	56640
cgcgaggagg	ccgtggacgg	ggtggcccg	cggtcgccg	ccgaggggtg	ccggaccgcc	56700
cgctgccccg	tgtcccacgc	gttccactcc	gcaactgatg	agccgatgct	ggacgagttc	56760
cgcgccaccg	tggccgccct	cgacctgcgg	gcgccctccg	tccccgtggt	gtccgcgctg	56820
accggccgcc	cgctgacggc	cgacgaggcc	cgctcaccgg	agcaactgggt	acggcacgtc	56880
cgcgaggccg	tccgcttcca	cgacgccgtg	cggggacttg	ccgccgaggg	cgccgtacgg	56940
tacctggaac	tcgggcccga	cggtgtgctc	accgccctcg	cccagagcag	cctgccgccc	57000
gccgacaccg	acgccgacgg	gcgcgacccc	ctcgccgtgc	ccctgctgcg	ggccggccgg	57060
ccggaaccgg	agacgctcac	cgacgcgctg	gccccccccg	ccgccgacgg	cctgaccgtc	57120
gactgggccc	gtacttccac	gggcccgggc	ggcgcccccg	tcgagctgcc	cacctacgcc	57180
ttccagcgccg	agcactactg	gctgcccgtc	gactccggcg	ccggcacggc	cccggccggg	57240
caccccctgc	tgtccgcccg	ggtcgacctg	ccggacggcg	gcctcgtact	caccggcggg	57300
ctctcgcccc	ccgcgcgccc	ctggctcgcc	cagcacaccg	tgcgcggcag	cgccctgctg	57360
cccggcaccg	ccctgctgga	cctggccctc	gcccggcccg	gccaggccgg	cgcgcccggc	57420
gtcgccgaac	tgatcctcga	agccccctc	gtgctgcccg	ccgagggcgc	cgtggagggtg	57480
cgctcaccg	tcggcgcccg	cgggaccgac	ggccgcggcg	cgatcgccct	gcacaccgcc	57540
gcccggcgacg	gcgactggac	ccggcacgcc	accggagccc	tgggcgagggt	gcccggcgag	57600
cccacggccg	ccggcgccctg	gccgcccccg	gatgcccggc	ccgccgacct	cgccgccctg	57660
tacggcgggc	tggccgacgc	cggtctcggc	tacgggcccg	cctaccaggg	cctccgtgcc	57720
gcctggcggc	gcggcgaaag	cccggcgggc	gaggtgttcg	ccgaggccga	actgcccggc	57780
gccatccccg	acgccgaccg	ctgcccgtta	cacccggcgc	tgctggacgc	cgtcctgcac	57840
gccatcgggc	tgggcgggct	gatcaccgac	ccggcgacgc	gcggaactcc	gttcgcctgg	57900
accggggctac	gggtcttcgc	ccccggcgcc	cgggcggttc	gcgcccggct	ctcccgggcc	57960
ggcgccgaag	gcgccctcgc	cgtcgacctg	ttcgacggcg	acgggctgcc	agtggccgcc	58020
atcggtccc	tgcggctgcg	cccgcccgcc	gctcccggcg	tgcccgacgc	cctcttcgag	58080
accgctgga	cgctgtcga	gcagggcacg	gccccggccc	gccgctcgc	gctgctcggc	58140
gccgacaccg	ccctcgcggc	cgccctcacc	gcggccggcg	ccgccctcgc	cgacgccacg	58200
gaccgctccg	ccgaggtgct	cgctcctgcc	atcgtcaccg	accccggcgc	ggccccctc	58260
accgagaccc	accgggcgac	cgccgcgctc	ctgacggccc	tccgtgacgt	cctggccgac	58320
gaggagagca	ccgcccgcct	cgccgtggtc	acccgcggcg	ccctcgcgct	gtccgccgag	58380
gagtccccgg	acccggccgc	ccgcgcgctc	tggggtctgg	tgcggtccgc	ccagaccgag	58440
caccccgacc	ggatcgtcct	cgccgacctg	gacgcggccg	acgcctcggc	ccgcgccctg	58500
cccccgccg	tgacctgccc	ggaaccgcag	ctcgccgtgc	ggtccggcgc	ggtcagcgca	58560
ccccggctca	cccgcgcggg	cgccgacggc	ctggtcctgc	ccgacggcgg	ctggcgccctg	58620
cgggcccggcg	ccaccggcac	cgtcgacggc	atgacggccg	tgccgcaccc	cgacgcgccc	58680
ctcgccgacg	gcgaggtacg	ggtcgccgtc	cgcgcggtcg	gcgtcacctt	ccgggacgtc	58740
ctcagcgtgc	tcgcccteta	ccccggggca	ccccagccgc	tcggcataca	ggcggcgggc	58800
gtggtgaccg	ggaccggccc	cggcgtgagc	gacctcgccc	ccggcgaccg	ggtgttcgga	58860
ctgctgcccc	gctccatggg	ctcctccgcc	gtcgccgacc	ggcgcgtgct	cgcgcccgtc	58920
cccgacggct	ggggcttcac	ccgggcccgc	tcggtgccct	ccgcgttctt	caccgcctgg	58980
ttcgcgctgc	gcgatgtggc	cggggtgcgg	gcgggggagc	gggtgctggt	gcacgcggcg	59040
gccggtggtg	tgggcatggc	cgcggtgcgg	gtggcgccgg	tgctgggccc	cgaggtgtat	59100
gcgacggcga	gtcccggcaa	gcatgggggtg	ctgcccggcg	ccggtctgga	cgagggcgct	59160
gtggcgtcgt	cgcgggacac	ggagttcgcg	cagcggttcc	ccgagatgga	cgtcgtgctg	59220
aactccctca	cgggtgagtt	cgtaggacgc	tcgctcgcac	tgctccgtcc	cgggcgaccg	59280
ttcgtggaac	tcggcaagac	cgacctgcgc	accgacaccg	ccggcatcac	ctaccgggcc	59340
gtcgacctcg	cggacgcggg	ccccgaccgc	atccaggaga	tgctcaccga	actcctggac	59400
cgctcgcgg	ccggcgacct	cgcccacctg	cccgtccgca	gcatgcccac	gggcccgcgc	59460
cgcgaggcgt	tccgcttcat	ggcccaggcc	cggcacaccg	ggaagctcgt	cctcaccacc	59520
gccccgtacg	gcgacggcac	cgctcctcgtc	accggcgcca	ccggcgccct	cggcggcctc	59580
gtggcccggc	acctcgtcac	cgaacacggc	atccgcgacc	tgggtgctcgt	cggacgacag	59640
ggcgccgagc	ccccctcac	cgccgaactg	cgcgccggcg	gcgcccgggt	ccgcgtggcg	59700
gcctgcgacg	tgtccgaccg	ggccgcgctc	gcccgcgctg	tcgcgacat	cgagccgccc	59760
ctgaccgcgg	tgggtcacgc	ggccggcgctc	ctcgacgacg	gcacgctcac	ctcgctgacc	59820
cccgaaccgg	tcgcccggct	actgcgcccg	aaggccgacg	ccgcctggca	cctgcacgaa	59880
ctcaccgagg	acagggacct	gtccgccttc	gtgctgttct	cctcggcggc	cggcacgttc	59940
ggcgcccccg	gccagggcaa	ctacgcggcc	gccaacgcct	ccctggacgc	gctcgccgag	60000
caccgcccgt	cccgcggcct	gcccgcgctc	tcctcgccct	gggggcccgtg	ggccgcggag	60060
agcgccatga	ccggcgccct	cagcagcggc	gaccgcgcca	ggatgaccgg	ggccggcgctc	60120

cggccctg	ccgccaccga	ggcactcgcc	gtgctcgacg	ccgcctgccc	caccggagcc	60180
ggcgccctcg	ccgcgctccg	tctcgacacc	gcgccgctca	ccgcccgcac	cgccgccccg	60240
caccgctgc	tgcgagacct	gggtccgccgt	ccggccggcc	ccgcccgcga	cgacgcccgc	60300
accagcccg	cgctgccccca	acggctggcc	ggactgggcg	aggagcagcg	ccgcccggcc	60360
gtgctcgacg	tcgtagcccg	caacgcggca	gcggtcctcg	gccacgcccg	ggcgtccgcc	60420
gtggacaccg	cgcgcggtt	cctggacctc	ggcttcgact	cgctgaccgc	cgtcgaactg	60480
cgcaaccggc	tcaccgaggc	caccggcctg	cggtgtccg	cgaccgccgt	cttcgaccac	60540
ccgaccccgg	ccgcctggc	ccgccatctc	ctgaccgaac	tggaaaccgt	ggtccgggcc	60600
gcgcgagcg	ccctgcccag	cgcgccggac	cccgaagccg	acctgcccgg	cgccatcgcg	60660
gccatcccgc	tcgaacgcct	ccggcaggcc	ggactcctgg	acgaactggc	ccgcctggca	60720
ggcgtcgccg	taccggccaa	ggacgccccg	gccgagcagg	acgcagcccc	cgacgaccac	60780
cccgacgacc	ccgcggcagg	acccgaggac	gacggttccg	acgacctcat	ggacgcgctg	60840
gacgacatga	gcatcgacga	cctgatcagg	atcgcgacag	acgagcgcgc	ccgcggaaac	60900
tgaggatcac	gaacagcatg	agcaaacacca	acgaagaact	cgtagaggcg	ctgcgagct	60960
cgctgcccga	gaccgagcgg	ctgcccgcgc	acaaccgctc	gctgacggcc	gcccgcgacg	61020
agccgatcgc	catcgtccgc	accgcctgcc	ggttccccgg	cgcatcgac	tcgcccgagc	61080
ggctgtggga	cgccgtccgc	gcccgtccgg	acctgatcac	cggttcccc	gacgaccgcg	61140
gctgggacct	cggcgtccac	gaccccagcc	ccgagcgggc	cgcccgagc	tacaccgacc	61200
gcccggcgtt	cctcaccggc	gcccgcgact	tcgacccggc	gttcttcggt	atctccccgc	61260
gcgaggcccg	cgccatggac	ccccagcagc	gggtgctgct	ggagaccgcc	tgggaggcgt	61320
acgaacaggc	cgccatcgac	ccgcacgccc	agcgcggcag	ccgcaccggc	gtgttcgtcg	61380
gcacctggag	ccagggtac	ggcatcggcg	cccgcctccc	cgaggacgcc	gagggtacc	61440
tcgtcaccgg	cggcggcacc	gcccgtcgtc	ccggccggat	ctcctacgtg	ctcggcctgg	61500
agggcccggc	cgtcacggtc	gacaccgctc	gctcctcgtc	gctggtcgcc	ctgcaactgg	61560
cggtgcccgc	gctgcccgc	ggggaatgct	cgatggcgct	cgccggcggc	gtgaccgtca	61620
tgcccggccc	cgggggtgtc	gtggagtctc	cccgggagcg	cgccctcgcc	cccgcggccc	61680
gctgcaaggc	ctactcggcc	gacgcggacg	gcaccggctg	gggcgagggc	gtcggcgtgc	61740
tcctgctgga	gcccgtctcc	gacgcctccc	gcaaccggca	ccgcgtcctc	gcccgtcgtc	61800
gcccgtccgc	cgtcaactcc	gacggcgcca	gcagcggcct	caccgccccg	aacggcccct	61860
cccagcagga	ggtcatccgg	caggccctcg	ccgacgcccg	gctcaccccc	tcggacgtcg	61920
acgtcgtgga	gggcccaggg	accggcacc	ggctcggcga	ccccatcgag	gcccaggccc	61980
tgctcggccg	ctacggcccag	gaccgcccc	gcccgtgct	gctcggctcc	gtgaagtcca	62040
acctcggcca	caccagggcc	gcccggggcg	tcgcccggct	catcaagatg	gtcagggcca	62100
tgcccggcgg	catcgcacc	cgccacctcc	acgccaccga	gcccaccccc	caggtggact	62160
ggctcccggg	agcccgtcga	ctcctcaccg	acaaccgctc	ctggcccggc	accggggcac	62220
cccgcgggtc	cgcgggtgtc	gcccgtcggcg	tcagcggcac	caaccgcccac	atcgtcctcg	62280
aacaggcccc	cgagcccggc	gacaccgaca	ccaccgcccg	caccgcacac	ccggtcgtgc	62340
cctggctgct	gtcccggcac	acaccggccg	ccctgcccgc	ccaggcccga	cggtctcctg	62400
ccggcctgcc	cgacgacgcc	gaccccctcg	acgtcggcgc	cgccctggcc	accaccgccc	62460
ccgcctgcc	gctgcccgc	gcccgtcctcg	gcccgcagcg	gactcagctg	cgcccggccc	62520
tcgcccggct	cgcccgggc	gcaccgccc	ccctctcccg	cgaaagcccg	tcggaccgca	62580
ccaccggcct	cctctacacc	ggccagggcg	cccagcggcg	cgccatgggc	gaggaaactcg	62640
ccgcgccta	cccggcgttc	gcccgcgct	ggaccgaggt	ctgcccggag	ttcgacaccg	62700
tgctgccccg	ccccctgccc	caggtgatca	ccgaaggcgg	ccccgacctc	gaccgcacc	62760
tgtacgcgca	ggcccgggtg	ttcgccctcg	agaccgcctc	gaccgcgctg	ctcggctcct	62820
gggggatagc	ccccgacctc	gtcctcggcc	actcctcgg	cgaactcgcc	gcccaccaca	62880
ccgcgggagt	gctgtccctg	cggcacgccc	tcgctcgtgt	cgcccggccc	ggcccggctca	62940
tggaggcgct	gcccggaggc	ggcgcctatg	tcgcccgtcca	ggcgagcag	gacgagatcg	63000
agctgcccga	gggctgggcc	ctggcggccc	tcaaccggccc	gtcgtcgggt	gtcctgtccg	63060
gcgacgaggg	cgctgtcctg	gccaccgccc	cccactgggc	cgagcgcgga	tggccgacca	63120
agcggctcac	cgtcagccac	gccttcact	cccaccgcat	ggaccccgtc	ctgaacggct	63180
tcgcccggct	gctgggccc	gtcaccctca	atacgcggcg	catcccgttc	gtctccacgg	63240
tcaccggcgc	ccccgtcga	gcccggactgc	gagaccccga	gtactggctg	cgaaacgtcc	63300
gagacaccgt	ccggttcgccc	gacggcgtcc	ggaccctggc	cgacgagggc	gcccgacacc	63360
tcgctgaggt	cggcccggac	gcccgtgctcg	gcccctcgt	cgcccgacgg	ctgcccgacg	63420
accccagaccg	gcccgaata	ctcggcgtac	ccaccgcccg	cgccgaccgg	cccgagcccg	63480
agaccctggt	cgccgcgctc	gcccgcactc	acgcccacgg	cgccgacgctc	gactgggccc	63540
ccttcttcgg	ctcccggccc	cgcccggctg	acctgcccac	ctacgccttc	cagcaccagc	63600
actactggct	ggcccggca	cccgcgacg	acctccccgc	cgcccggcctc	ggcaccgtcg	63660
gccaccccgt	gctgcccgc	gcccgtcgaac	tgcccggcca	cgccaccggc	gaccacgccc	63720

cccgccccgg	cgccgccgga	gccgtcgtct	tcaaccggcac	cctgtccgcc	cacacgcacc	63780
cctggctggc	cgaccacagc	gtcctcggca	cgccgggtgt	ccccggcacc	gccctcgccg	63840
aactagccgc	cgccgcgggc	gaccgcctcg	gctgcgccac	cgtcgccgaa	ctcgtcctca	63900
ccgcgccgct	cgccgtcccc	gcctccggcg	ccgtccggct	ccgcgtccac	gtcgacgccc	63960
ccgaccgcga	cggccaccgc	gccgtcaccg	tcgactcccg	gcccgcgcac	cagctcgccg	64020
acggcgccgc	ctggaccgcg	cacgccaccg	gacgcctcgc	ccccaccggt	cccgaaccgg	64080
cccaggcgcc	caccgcctgg	ccgcccgccg	acgccgaacc	cctccccgtc	gacggcctct	64140
acgaccggct	caccgcggcc	ggcttcgggt	acggccccgc	cttccgcggc	gtccgcgcgg	64200
cctggcgctc	cggcaccgac	ctgctcgccg	acgtcgaact	gcccgcggac	accgacgagt	64260
cgggctttct	gctgcacccc	gcctctgttcg	acgcgccctc	gcacgccctc	ggcctcggcg	64320
gctcggtcga	gcacggcgga	ctgcccttca	cctggacggg	cgtccgcctg	cacgcgcacc	64380
gcgcccgcctc	cctgcgggtg	cggctgaccc	cgaccggccc	cgacgcctgc	gcgctcaccg	64440
cggccgacgc	cacgggcccg	cccgtcgtca	ccgtcaccga	cctccgcctg	cgcccggcca	64500
ccgaggtccg	cggcaccacc	acggccgacg	ccctgcacca	catacgggtg	gaacccccgc	64560
cggcagctgc	ggccaccgac	gtcaccggca	ccgacttcac	cgtgttcacc	ctgccggccg	64620
ccgggtccga	cccgcaggcc	gtccgggagg	ccacgcgcgc	gtccctcgcc	cggttgacgg	64680
agcatctcgc	cgccgtcggg	ccggccggcc	cgctcgtcgt	ggtcaccac	ggcgcctcgc	64740
ccgcagaacc	cggcgacacc	gtccccgacc	tggccggcgc	ggcgtctgg	ggcctggtgc	64800
gctccgccc	gtccgagcac	cccggccgct	tcgtcctgat	cgacctggcc	gcttccgagg	64860
atgccgctgt	cgttccggcc	gcgatcgcca	ccggagagcc	tcagatcgct	gtacggaatg	64920
gagcgtgta	cgcgccccgg	ctggtgcgcg	gggcttctcg	gccgggtggt	gaggtgccgt	64980
tcggtgcggg	cgacgtggtg	ctggtgaccg	gtggcacggg	tgctctgggc	cggttggtgg	65040
cccggcacct	ggtggccgag	cacggtgtgc	gtcggctggt	gctggcgcgc	cgccgcggag	65100
gcgcccggga	gctggtcgcg	gagttgggtg	agtgggtgct	cgctgtggac	gtcgtggcct	65160
gcgatgtgtc	cgaccgggac	gcgctggagc	gtctggtcgc	ggcccatccg	ctgacgggtg	65220
tggtaggggt	gctgcggggc	ctggacgacg	gcacggtgga	gtcgtgacc	cctgaccggg	65280
ccctggacct	ccgtgccttt	aaggtcgacg	gtgcctggca	tctgcacgag	ctgacggccg	65340
gtcaggggtg	ttacgcggcg	gtgctgttct	cgtegtcgc	gggtgtggtg	ggcagtgccg	65400
ccgagggcct	gcccgccacc	tcacgcgcct	cgctggacgc	gctggcggag	caccggcggg	65460
gcctcgaccg	ggtggacgcg	gcccggctcg	ggggaccctg	gaccgacggc	atggcgagcg	65520
ccgacggcct	ggagctgttc	gacgcctct	cccgcgggtg	agtctgccc	ttcgcccacg	65580
cgcgcctgga	cctcgccgtg	ctgcgcgccc	gcgccaccga	cgacccgctg	accgtcgccg	65640
gctcgggtgc	ggcgcgccga	cgccgtaccg	aggcggaggc	gctgccgccc	gtgctgcgcg	65700
agcggctcgc	cgcgctcccc	gagaccgagc	cgctggcgcg	ccggcccggc	ggcttcgccc	65760
agaccgcggc	gaccgtgctc	ggccacgtcg	ggccggaggc	cgccctggaa	ctggtccggg	65820
tcctcgacct	cggttcgac	tcactggccg	cccgggaggc	cgctcgcccc	gaccgtcct	65880
cgaccggcct	gcgcctggcc	gcgacgatca	ccgtcgaact	ccgcaaccgg	ctcaccgccc	65940
cacgccacct	cctggacgcg	gcctggacg	ccttcgacca	ccccacctcc	gccgcctcgc	66000
tgcccgcgcc	caccgcctga	cgcgcgcgac	ccggcccgtc	ggccacgagc	ggcgcgcacc	66060
gctaccccgg	cgggtgtagc	tcgcccgagg	agccgatcgc	gatcgtgggc	atggcctgcc	66120
acgccgtcac	cggttcccc	gaggaccgtg	acctgtggcg	actcgtgctc	gcgggcgccg	66180
accagtcccg	caccgggacg	agctacaccc	gctgggacct	cacctccgtc	tacgaccccc	66240
acttcgacgc	cgagttcttc	ggcatcagcc	gogagggcgg	attcctcacc	ggggcggccc	66300
agaggctgct	gctggagacc	tcctgggagg	cgcgtgaggc	gctggccatg	gaccgcgagc	66360
cgctgcgcgg	cagcgcacc	ggtgtcttcg	cgatcgagcg	cgccggcatc	gaccgaccg	66420
gcaccggcac	cctcccggag	ggcgtcgagg	ccggcctgat	gtaccacgac	tacggcagcg	66480
tgctctcggg	ccgggtcgcc	tacacgctcg	gctacctggg	tctcggcacc	gccggcagcg	66540
ccgcgtgctc	ctcctcgctg	gtcgcctcgc	gctggagggg	ccccgcggtc	accgtggaca	66600
agtgctcgat	ggcgtggcc	ggcgggtgca	actgggcgat	ccaggcgcctg	cgcagcggcg	66660
agttctcccg	gcagcggggc	ctcgcccccg	ccgtcatggc	cacaccgggc	acgttcgtgg	66720
ccgacggagt	cggctggtcc	gagggcgtcg	acggccgctg	caagtcttc	gccgcgcggg	66780
cccgccgcaa	cggccaccaa	gtgctggccg	gcatgctcct	cgtggaacgc	ctgtcggacg	66840
gcgcgtcgaa	cggtctgacc	gccccgaacg	tcgtacgcgg	atccgcggtg	aaccaggacg	66900
cgctccaggg	cgccgggctg	acgacggccg	gtccctcgca	gcagcgggtg	atccgccagg	66960
gcaccaaact	gggcgacccg	atcgaggcgc	acgtcgacgc	cgtggaggcg	cacggtacgg	67020
gggacgagcc	gctgtggctc	ggttcgggtg	aggccctcct	cgccacctac	ggccaggagc	67080
ccggtgtggc	cggtgtgatc	aagatggtgc	agtccaacat	cgggcacacg	caggccgccc	67140
cgctccacgt	ggacgaggcg	tcaccgcacg	aggccatgcg	gcacgggacg	ctgcccgcga	67200
tcaccgagga	acgcgagtg	accgcggacc	tcgactggac	cgccggcgcg	gtggaactgc	67260
			gtcggccgcg	ccgcgcccgc	gtgtcgtcct	67320

tccgctgtgc	cggcaccaac	gcccacgtcg	tccctgaaga	ggcaccocgag	gagaccgcac	67380
cccccgatc	ggccgaactg	cccctcgtgc	cgctgctgct	gtccggtcac	acctccacgg	67440
cgctcgccgc	ccaggcacgc	cgactgcacg	accacctcgc	cgagtccggc	acaccggccc	67500
tggaaacggg	cggccgctcc	cttgcccgct	cgcgggctgt	gttgccgcat	cgtagcgggg	67560
tggtagccgc	tgatgtggcc	ggtgcccggg	cttctctggc	ggcgttggcc	gcgggggagc	67620
cggtggaggg	tgtggtgtcg	ggtgctgccg	ggttgcocgg	tgagggccgg	gtggtgttcg	67680
tgttcccggg	tcagggggcg	cagtgggtgg	ggatggcggg	ggatctgctg	gattcgtcgc	67740
cggtgttcgc	ggggcggatg	gctgagtggt	cggtggcgtt	ggagccgttc	gtggagtggg	67800
cgttgcgtgg	tgtgctgggt	gatccgggtg	cgttggagcg	ggtggatgtc	gttcagccgg	67860
tgttgtgggc	ggtgatggtg	tcgttggcgg	aggtgtggcg	ttcgtacggt	gtcgtgccgt	67920
cgctgtggg	cggtcattcg	cagggggaga	tcgcgctgc	ttgtgtggcg	ggtgtgctgt	67980
cgctggcggg	cggcgcccgg	gtcgtcgccc	tgcgcagcag	ggctctgacg	gcgctggcgg	68040
gcagcggggg	gatggtgtcg	gtcgcgcgg	ggcctccgg	tgtagaggag	ttgctggtgg	68100
gctgggcggg	ccggttggcg	gtggctgccc	tcaacggctc	tgagtccgtc	gtcgtcgcgg	68160
gtgagggggg	ggcgtgag	gagttcctgg	ctcattgccc	gggtcggggg	gtgcccgtc	68220
ggcgtatcgc	tgtcactat	gcctcgcatt	cggtgctggt	ggagccggtg	cgtaggctc	68280
tgctggctga	tctggagggc	gtgcccggg	gtgagggcac	ggttccgctg	ttctccaccg	68340
tgacggggga	gtgggcggac	ggcacggcgc	tggacggccg	ctactggtac	cgcaacctgc	68400
gcgaaccggg	gggcttcgaa	ccggcgggtc	gcccctgct	cgactccggc	cacgccgtgt	68460
tcgtggagat	cagcccgcac	ccggtgctga	cgcccccgt	gcaggagacg	gcggatgcca	68520
ccgagcggac	cgccgtcgtg	gtcggcaccc	tccgcccga	ccacgagggt	cagcggcaac	68580
tgctcaccga	cctcggcgtc	ctgcacacca	cgggcggcga	catcgactgg	accggtgct	68640
tcacggggcg	caccggccgt	gcccactgc	ccacctcgc	cttccagcac	acgcgctact	68700
ggctgacccc	ctccggcccg	tccgcccgg	aactggcccg	agcgggcctg	accgcccccg	68760
ggcaccggct	gctcggcgt	gccgtcgacc	tccggaggga	cgccggcctc	gtcctgaccg	68820
gacgcctcgc	cgccgacccg	gccgcctgga	cgcccgatca	cgtagctcct	ggcaccaccc	68880
tgctgcccgg	cgccgcccgc	gccgaactcg	ccctcggcgc	cgccgacggc	gtcggctgcg	68940
gcaccctgga	cgaactcgtc	ctcggcgcac	ccctggcact	gcccgaacgc	ggtgccctgc	69000
acctccaggt	acgggtgggc	cgcgccgagg	ccgaccaccg	ccgtaccgtc	agcgtccacc	69060
cgcgccccga	ggacggtgac	gcgccctgga	cccggcacgc	cgagggcgtc	ctcgtgcccg	69120
gggacaccgc	cgccggcacg	ccgctgaccg	agtggccgcc	cgccgacgcc	gaaccgggtg	69180
acgtgtccgc	gctctacgac	agcctcggcg	accgcccct	cgactacggg	ccggtcttcc	69240
gtggcgtacg	cgccgcccgt	cgccacggcg	acgacatcct	ggcggagggt	gaactgcccg	69300
cggagggccg	cgagtcgggc	ttcctgctgc	accgcccct	gctggacggg	gccctgcacc	69360
cgatcggcct	cggcggcctg	gtcggcgcag	cgccgctgcc	cttcgcccgt	cacggcctgc	69420
gtgtgcaacg	cgtcggcggc	cgcgcgcgcg	gggtacggct	cacccccctg	ggcgcagaga	69480
ccgtcgcctg	agacctcggc	gacggcggcg	gcccggcccct	ggccggccatt	gtctccctgc	69540
gcctgcgagc	cgtcacggcc	gcacaggtcg	ccgcccggcg	caccctcacc	gaccgggccc	69600
tgcacctgga	ctggctcccg	gtggcggggg	ccgctcggc	gccctcggac	gacaccggcg	69660
tggaaactgct	gcgcttggac	gacgcaccgc	acgggacggg	cgccgggtccc	gggggctgct	69720
ggcagggcct	gcgtacggcc	ctcaccgcga	tccaggagcg	gatcgccgac	gaccgggccc	69780
agcggctggt	cgtcgtcacc	cgcgacaccg	acctcggccc	cgccggcctc	ggcggctgct	69840
tgcggctccg	ccaggcccag	caccggggcc	gcttcggcca	cgtagctgct	gacggacacc	69900
ccgactccga	gcgcccctg	cccaccgcca	ccgcccgtgac	cgacgagccc	tgggtcggcg	69960
tgcgcgcccg	ggagacatac	gtaccgcccg	tgcgcccgtc	cggtgcgccc	ccgggcccgg	70020
acgtgcccgt	cgggcggcag	gacgtggtgc	tggtagccgg	cgccaccggg	gtgctggggc	70080
ccctggtggc	ccggcacctg	gtgaccgagc	acggtagcgg	caggctggtg	ctggcgagcc	70140
ggcgcggagg	cgccggggag	ctggtcggcg	agttggggcg	attgggcccg	gctgtggatg	70200
tggtagccct	tgatgtgtcc	gaccgggagg	cgctggagcg	tctggtcggc	gcccattccg	70260
tgacgggtgt	ggtgcacggc	gccggtgtgc	tggacgacgg	cacggtggag	tcgctgacc	70320
ctgaccgggt	ggacgggggt	ctgcccggcg	aggtcgacgg	tgctggcat	ctgcacgagc	70380
tgacggcccg	cctggacctc	cgtgcctttg	tgctgttctc	gtcgtcggcg	ggtgtggtgg	70440
gcagtgcccg	tcaggtggtg	tacgcccggg	cgaacggccc	gttggacggc	ctggcggagt	70500
accggcgggc	cgagggcctg	cccgccacct	ccatcgcctg	gggctgtggt	gcgcccggcg	70560
gcgccatgac	ctccggcggc	gacaccggcg	gcctggcccg	gagcggcatc	ctcccgtccc	70620
ccgcccggcg	ggccctggaa	ctgttcgaca	ccgcccggcg	cgccgagacg	ccgctgaccg	70680
tcgcccggcg	cctggacctc	acggcgttcc	gcccgggggg	cacccgtatg	cccgtcgtcc	70740
tgcggagcct	cgggggcccc	gcccggcccg	gcacggcagc	cgccggcgac	gcccggacct	70800
tgcgcgaccg	cctcggcggc	cagacccccg	ccgaacggac	ccgaccgtg	ctcgacctgg	70860
tccgagggca	ggccggggcc	gttctcggac	acgagtcggc	cgccggccatc	gcccggagcc	70920

gggccttct	cgaactcggc	ttcgactccc	tgaccgcgct	cgacctgcgc	aaccgcctcg	70980
gcacggtgac	gggcctccgc	ctgcccacca	ccaccgtgtt	cgaccacccc	aaccccgcg	71040
cactgaccgg	gcacatcctc	gccgagctgc	tggcgccac	ggccggtctg	gcccaggccg	71100
ccccaccgc	cgtgcggacc	gacgagccga	tcgogatcgt	gggcatggcc	tgccggtacc	71160
ccggcggagt	ggcctctccc	gaggacctgt	ggcgcgctgt	cgccgagggc	cgtgacgtca	71220
tctccccctt	cccggaggac	cggggctggg	acctcggcgc	cctgtaccac	gccgaccccg	71280
accacacggg	gaccagctac	gcccgcgagg	gcggttctct	gcacgacgcc	gccgggtctg	71340
acgccgagtt	cttcggcatc	agcccgcgctg	aggcgcctgg	catggacccc	cagcagcggc	71400
tgctcctgga	ggcgtcctgg	gaggcgatcg	agcgcgcgg	catcgacccc	accgcgctgc	71460
gcggcagcgc	caccggtgtc	ttcgccggcc	tgatgtacca	cgactacgcg	gcccggctcg	71520
gcaccacccc	ggagggtacc	gaaggctacc	tgggcatggg	caactccggc	agcctgcctt	71580
ccggccgcat	ctcctacacc	ctgggcctgg	aaggcccggc	ggtcaccgtg	gacaccgctt	71640
gctcctcgtc	cctggtcggc	ctgcaactgg	cgatccaggc	gctgcgctcg	ggggagtgcg	71700
acctcgccct	ggccggcggc	gtgtccgtga	tggcgacccc	gggcacgttc	gtggagttct	71760
cccggcagcg	gggcctcggc	cccgacggcc	gctgcaagtc	cttcgcccgc	gcccgcgacg	71820
gggccaagctg	gtccgagggc	gtcggcatgc	tgctcgtgga	gcgctgtctg	gacgcccgcc	71880
gcaacgggtca	ccgtgtgctg	gccgtcgtgc	gcccgtcggc	ggtgaaccag	gacggtgcgt	71940
cgaacgggtct	gacggccccg	aacggtcctt	cgcagcagcg	ggtgatccgt	caggcgtctg	72000
ccggtgccgg	gctcacgtcc	gcggacgtgg	acgcggtgga	ggcgcacggt	acgggcacgc	72060
ggctgggcga	cccgatcgag	gcgcaggccc	tcctcgccac	ctacgggcag	gagcgggacg	72120
agccgctgtg	gctcggctcg	gtcaagtcca	acatcgcca	tgccgagggc	gcccgcgggtg	72180
tggccggtgt	gatcaagatg	gtccaggcga	tgcggaacgg	cgtgctcccg	cgcacgctcc	72240
acgtcgatga	gccgtccccg	cacgtcgaact	ggaccgtcgg	cgcggtcgag	ctgctgaccg	72300
gggaacagga	gtggcccgcg	caggaccgcc	cgccgcgctg	gggtgtgtcg	tcgttcggtg	72360
tgtccggcac	gaacgcgcat	gtggtgctgg	aagaggctcc	cgaggagacc	gagggcaccg	72420
cgccggtctg	gctgcccggc	gtgccctggg	tgggtcggc	ccgctccgag	accgcgctgc	72480
gtgctcaggg	cgcacggctc	gcggactggc	tgcaacggcg	caccgacgtc	ctcggcaccg	72540
cctactccct	cgccaccggc	cgggcggcgc	tgccgcaccg	cgcggtcgtg	gtcgggaccg	72600
accgggccga	actctccgac	ggactggccg	cgctggcggc	cggtcgcgcc	gccgcgcacg	72660
tcgaatcggg	ccgtgcccgc	gacaaccgtg	tcaccgcggt	cgtgttcgcg	gggcagggtg	72720
cgcagcgggc	ggggatgggg	gcggaactgg	ctgccgcgta	tcgggtgttc	gctcaggtgt	72780
tcgctcaggt	gtgtgcccgc	ttcgacgggtg	tgctggagcg	gccgctgggt	gaggtgatcg	72840
ccgaggggtg	tccggagctg	gatcggacgg	tgtacgcgca	ggcggggttg	ttcgctttcg	72900
aggtggcgtt	gttccggctg	ctggagtcgt	gggggtcgc	gcccgacgtc	gtactcgggc	72960
attccgtggg	tgagtggcc	gctgcctgtg	tggcgggtgt	gtggtcgctc	gaggacggcg	73020
tgccgggtggt	ggctgcgctg	gggcggttga	tgcaaggcgt	gccgcagggc	ggtgcgatgg	73080
tcgctctcga	ggtgtcggcg	ggtgagctgg	agttgccgga	gggtgtggag	ctggcggcgg	73140
tcaatggtcc	gtcgtcggtg	gtgctgtccg	gtgaggagga	ggccgtcctg	gctgaggcgg	73200
cccgtggcc	ggaccgtcgt	gccaaagcgtc	tgcgggtcag	tcacgccttc	cactcgcgtc	73260
ggatggatcc	gatgctggag	gacttccggc	gggtactgga	gtcgggtggc	ttccacgctc	73320
cggagcttgt	tttcgtgtcg	acggtgaccg	gtgctgtggt	gacggacgag	ttgtgtgacc	73380
cgggggtattg	ggtgcgcaat	gtgcgggaga	cggtccggtt	cgccgatgcg	gtggtggcgg	73440
cggaggcccg	tgtcttcgtg	gagttggcgc	cggacgcggt	gttgtcgggg	ctggtgggcg	73500
agtccgtcga	gggtgtgctg	tgtgttcccg	cgccagcggc	gggcaggccc	gcccgcgggg	73560
ttctcgtgtc	ggcgtggggc	acgctgcaca	cgcaacgggt	cgacgtcgcc	tggggccggt	73620
tcttcgcccg	tagcgggtgcc	cgtcgtgtcg	acctgcccac	ctacgccttc	caaccgcgac	73680
gctactggct	ggacgcccc	gcgccgcaga	ccggtggcgc	ctcggacgac	gccttctggg	73740
ccgcccgtcca	gagcggtgac	ctcgccggtc	tgctgggtgt	cgccgaggac	gccgggctgg	73800
acgccgtact	gcccacgctc	gcgtcctggc	acgacagggg	acgtgccgac	gcggtcgtgg	73860
acggatggcg	gcacaaggtc	cgctggaccc	cgctccccga	ggccggcggc	gccgtgctga	73920
gcggccggtg	gctgctggtc	ggcccggagg	gggacgaggc	gctcgtcgcc	gacgtcgcca	73980
ccgcccctgcg	cgagcacggg	gccgaggtca	cgcaacctgc	gctcccgcgg	gacgcccgac	74040
gggagacgac	cgccgaactg	ctccgcggca	ccgacgacac	cggtctcacc	gcagtgtgtg	74100
ccctgctcgc	gcgggcccgg	cgcccgggtg	acgccaccct	cgccctcgtc	caggcgtctg	74160
gcgacgccga	agtaccctg	ccgctgtggt	gcgctacctc	cgagtccgtc	gccgtcgctt	74220
ccacggacgt	cgtaccggaa	gccgcccgtg	acgcggccgg	cctctggggg	ctcggcccgg	74280
tcgtacgcct	ggaggcaccg	gaccgctggg	gcccactggt	cgaccttccc	ggcgcgctgg	74340
acgcgcgggc	ccgcccgcgg	ctcgccgcgg	tectggccgg	cgccgaggac	gagtgccgctg	74400
tgccgcgagaa	cggcgccttc	gccgctcgtc	tcgctccgagc	cgccgacgca	ccccgcggcg	74460
agtggcggcc	ccagggcacg	gtcctcgtga	ccggcggcac	cgccgcgctc	ggcgcgcccg	74520

tgcgccagcg	gctcgcggag	cgcggtgcgc	ggcacgtgct	gctggtgagc	aggcgcggtc	74580
cggcggccga	cggcgcggcc	gaactggtcc	gcgccatcga	ggcggcgggc	gccacggcga	74640
ccgtcgccgc	gtgcgatgtc	gccgatccgg	aggcggtcgc	ggccctgctg	gagcggatac	74700
ccgccagcgc	cccgtcgacc	gcgggtggtgc	acacggccag	cgtcctcgcc	gacgccccgc	74760
tggacacgct	caccccggac	cggatcaccg	ccgtcctccg	ggccaaggcc	gatgccgcac	74820
gaatcctgca	caccgcgacg	gccagcctcg	acctggacgc	gttcgtgctg	ttctcctcgc	74880
tgcgccggcac	cctcggcaac	cccgggcagg	ccgcctacgc	cgccgcgaac	gccgtactcg	74940
acacgctcgc	cgcgcacccg	cgggccctcg	gactgcccgg	caccgcccgtc	gcctggggcc	75000
cctgggcccgg	cggcggcatg	ctcgacgaca	cggtggcggg	gcgactgcgc	cgcgcccggg	75060
tcatcccgtc	cgaccccagag	cacgccttgg	tcgccctcga	ccgtgccgtg	gcccgccggg	75120
acgcgcacag	cgtcatcgcc	gacgcgcgact	ggaccggcct	caccgcccggc	tccatgctgg	75180
ccgaactctc	cggcccggca	cccgtcgagg	cgcccacggc	gctcaccgggt	cccgaaccgcg	75240
agcgcgccgc	actggccctc	gtccgctcct	gcgcgcgccg	cgtcctcggc	cgctcggccg	75300
ccgtggacgt	cgaaccccag	accgccttcc	gcgagctggg	cttcgactcc	atggccgccc	75360
tccagctcgc	caaccggctc	aacgcgcgca	cgggggtcgt	gctcaccgcg	accgcccgtgt	75420
tgcaccacc	gagcgcgccg	gccctggcca	cccacctgct	cgctctggcg	accggggagg	75480
gcgcccggctc	ggcccggcgcg	accgaactcc	cgggtacggc	cgtccacacg	gacgagccga	75540
tgcgcgatcgt	cggcatggcc	tgcgccttcc	ccggcgacgt	gacctcgccc	gaggacctgt	75600
ggcggctgct	cgccgacggc	gtggacgccg	tcggccccc	cccggcggac	cgccagtggg	75660
gacccaccga	cgccctacgcc	gaaggcggct	tcctgcgcgg	ggccggcggg	ttcgacgcgg	75720
acttcttcgg	gatcagcccg	cgcgaggcgc	tggtcatgga	cccgcagcag	cggtcgcgcc	75780
tggagaccgg	ctgggaggtg	ttcgaacgcg	cgggcatcaa	cccgcacacg	gtgcgcggca	75840
cgccgctcgg	cgtcttcctc	ggcaccaacg	gccaggacta	cgtgtccctc	ctcgcggcg	75900
ccaccgaggc	ccacgcggga	cacatcggca	ccggcaactc	cgccagcgtg	ctgtcgggcc	75960
gcatcgcta	cgtcctcggc	ctggaggggc	ccgcggtcac	cgtggacacc	gcgtgctcct	76020
cctcgtcgtg	cgccctgcac	tgggcgatcc	aggcgtcgcg	cagcggcggg	tgtctccatgg	76080
ccctggcccgg	cggtgtcacc	gtcatggcca	ccccgggcgc	cttcgcccag	ttctcccacc	76140
agcggggact	ggcggaggac	ggccgctgca	agtccctcgc	ggcatcggcg	gacggcaccg	76200
gctggggcga	gggcgtcggc	atgctcctgg	tggagcgtct	gtcggacgcc	cgccgcaacg	76260
gtcaccgtgt	gctggccgtc	gtgcgcggct	cggcggtgaa	ccaggacggt	gcgtcgaacg	76320
gtctgacggc	tccgaacggt	ccttcgcagc	agcgggtgat	ccgtcaggcg	ctcgcgggtg	76380
ccgggctcac	gcccgtggac	gtggacgcgg	tggaggcgca	cggcacgggc	acgcggctgg	76440
gcgacccgag	cgaggcgcag	gcgctcctcg	ccacctacgc	gcaggagcgg	gacgagccgc	76500
tgtggctcgg	ttcggtcaag	tccaacatcg	ggcacacgca	ggccgcccgc	ggtgtggccg	76560
gtgtgatcaa	gatggtggag	gcgatgcgga	acggaacgct	gcccgccacc	ctccacgtgg	76620
acgagccgctc	cccgcacgtc	gactggtccg	ccggcgcggg	cgagctgctg	acggaagcgc	76680
gggagtggaa	gcgtgcccgt	cgccgcgcgc	gtgcgggtgt	gtcgtcgttc	ggtgtgtccg	76740
gcacgaacgc	gcatgtggtg	ctggaagaag	cgccggagga	gaccgaaccg	gcggcttcgg	76800
acgaactgcc	ggtcgcaccc	tggctgttgt	cgcccccgtc	agagaaggcg	ctgctggcac	76860
aggctcagcgc	gctgcgcgcg	tacgtcacgg	agcaccgcga	ggcgcggccg	gccgacatcg	76920
gcctcagcct	ggccaccggc	cgagccgcgc	tggcccaccg	tctgagcggg	gccggcgaga	76980
ccaccggagga	actgctcgcc	gcccttgaca	ccgccttcc	cgccgtggct	cgcgaaacc	77040
cgaccgcggt	cgtgttcgcg	gggcagggtg	cgcagcgggt	ggggatgggg	gcggaactgg	77100
ctgcctgta	tccggtgttc	gctcaggtgt	tcgcgcagg	gtgtgcggcg	ttcgacgggtg	77160
tgctggagcg	gccgttgggt	gaggtggtcg	ccgagggtgg	tccggagctg	gatcggacgg	77220
tgtacgcgca	ggcgggggtg	ttcgctttcg	aggtggcgtt	gttccggctg	ctggagctgt	77280
gggggtcgc	gcccggacgtc	gtactcgggc	attecgtggg	tgagttggcg	cctgcctgtg	77340
tggcgggtgt	gtggtcgtc	gaggacgcgg	tgcgggtggt	ggctgcgcgt	ggcgggtga	77400
tgcaggcgc	gccgcagggc	ggtgcgatgg	tcgctctcga	ggtgtcggcg	ggtgagttgg	77460
agttgccgga	gggtgtggag	ctggcggcgg	tcaatggtcc	gacgtcgggtg	gtgctgtccg	77520
gtgaggagga	cgccgtcctg	gctgaggcgg	cccgtggcc	ggaccgtcgt	gccaagcgtc	77580
tgcgggtcag	tcacgccttc	cattcgcgtc	ggatggatcc	gatgctggag	gacttcccggc	77640
gggtactgga	gtccgtcacc	ttccacgctc	cgcagcttgc	cttcgtgtcg	accgtgaccg	77700
gtgctgcggg	gacggacgag	ttgtgtgacc	cggggatttg	ggtgcgcaat	gtgcgggaga	77760
cggtcccgtt	cgccgatgcg	gtggtggcgg	caggggcccg	tgtcttcgtg	gagttggcgc	77820
cggacgcggg	gttgtcgggg	ctggtgggcg	agtccgtcga	ggcgtgctg	tgtgttcccg	77880
cgcagcgggc	gggcaggccg	gccgcgcggg	ctctcgtgtc	ggcgtgggc	acgctgcaca	77940
cgcacgggtc	cgacgtcggc	tgggaccggg	tcttcgcccg	tagcgggtg	cgctcgcgtc	78000
acctgcccac	ctacgccttc	caacgcgaac	gctactggct	cgcccccggc	cgccgcggcc	78060
ccgtcggcct	cgcgggcgtc	ggtctcacgg	ccaccgggtca	ccccctgctc	ggcgtcagcg	78120

tcgagctgcc	cggcacggac	gcggtggcgt	tcaccgggag	cgtgtcgtg	agcaccacc	78180
cgtggctggc	cgaccacgcc	atcctggggc	ctgccctgct	gccgggcacg	gcgttcctgg	78240
acctcgccct	cgcgggcggc	gaacacgtcg	gctgcccgcg	ggtggacgac	ctcgcgtgc	78300
acgccccgct	ggcgtccccg	gcccacggct	cgggtgcgcat	gcaggtcagg	gtcaggggga	78360
ccgacccccg	cgggagccgc	caggtcggca	tctactcacg	gcccgaggac	gacgaggacg	78420
cctggaccca	gcacgcgacc	ggcgtgctgg	cgcccgaag	cggccccgcc	gccgaggcgc	78480
tgaccgagtg	gccgccccag	ggcgccgaac	ccgtgtccgt	ggacggcctg	tacgacgacc	78540
tggccgccac	cgggttcgcg	tacggcccc	tcttcgcgcg	cctgcgtgcc	gcctgggtgc	78600
gggacggcgc	cgtctacgcc	gacgtcgccc	ttccggacga	gacggcctcc	gtcgcgggct	78660
acagcgtgca	cccggcgctg	ctcgacgcgc	ccctgcacgc	gctcggctgc	gcgaacctgg	78720
tcggcgactg	ggcggacgga	cagctgccct	tccgctggac	cggcgcgcgg	atccacgccg	78780
tgggcccag	ggcgtgccc	gtccggctcc	gggcccgaag	cggcgggatc	gcgctgaccg	78840
cggccgatgc	cggcggacag	cccgtggccg	ggatcgaagg	agtccggctg	cggcccgtcg	78900
gcgacgcgcg	tgccgtccgt	gcggcggcgg	gcgcgcgccc	ctgccgggtc	gagtggtgc	78960
ccgtcgtacc	ggcggacgac	cccgaggacg	acgacatcac	cgtcgtcccc	gtgaccggcc	79020
gtaccgcccg	cggcgacgtc	gtggccgagg	tgcgtgccgc	cgtcgcgcc	gccctggccc	79080
ggctccagga	atggctggcc	gacgaccggt	ccgagcgcct	ggtcctcgtc	accggggcg	79140
ccgtcgcggc	gctgccggac	gaggcaccgg	acccggtggc	cggcgcggtc	tggggcctgg	79200
tgcgctccgc	ccagtccgag	caccccggcc	ggttcgtcct	ggccgacttg	gccgcttccg	79260
aggatgccgc	tgtcgtttcg	gccgcggtcg	ccaccggaga	gcctcagatc	gccgtacgga	79320
atggagcgct	gtacgcgccc	cggctggtga	ggggggcttc	tcggccgggt	ggtgaggtgc	79380
cgttccggtc	ggcgacgctg	gtgctggtg	ccggtgggac	gggtgctctg	ggcggggtgg	79440
tggcccggca	tctggtggcc	gagcacggtg	gtcgtcggct	ggtgctggcg	agccggcgag	79500
gggggtgccg	ggagctggtc	gccgagttgg	gtgagttggg	cgcggtctg	gatgtggtgg	79560
cgtgtgatgt	gtccgaccgg	gaggcgtgg	agggtctggt	cgcggcctat	ccgctgacgg	79620
gtgtggtgca	cgcggccggt	gtgctggacg	acggcacggt	ggagtgcgtg	accgcgagc	79680
gggttgacgg	ggtgctgcg	gcgaaggtcg	acggtgcctg	gcactgacac	gagctgacgt	79740
ccggactgga	tctccgtgcg	ttcgtgctgt	tctcgtcgct	cgcgggtgtg	gtgggcagtg	79800
cgggtcaggg	tggttacgcg	gcggcgaacg	ccgcttggga	cgcgctggcg	gagtaccggc	79860
gggcccaggg	cctgcccgcc	acctccatcg	cctggggacc	ctggaccgac	ggcatgacca	79920
cagggtcggg	gcgggcccgg	cgtgcccgta	tcagccggtc	gggtaccggg	gccctcggca	79980
ccgaggacgg	actcgccctc	ttcgaccagg	ccgtacgcgg	cggccagcgc	ctcgcctcg	80040
ccgcgctgtg	ggacctgtcc	gcgctgcgcg	ccgcgcagtc	gctgccgect	ctgttccagg	80100
ggctcgccgg	acgaccggtc	cggcggaccg	tcgcccagcg	cggccggcgg	ggcgccgagt	80160
gggcccagcg	gttaccgggg	ctgtcgcgg	ccgaaagcga	acgcgcccgc	gtcagtggtg	80220
tccgtgagca	ggcggcggcc	gtgctggggc	acgcgtcctc	cggccgctc	gccgctgacc	80280
gggccttctc	cgacctgggc	ttcgactcgc	tcaccgccgt	cgaaactgcg	aaccggctcg	80340
ccacggcgac	cggcctgcg	ctcggcacga	ccgcccgtgt	cgactacccc	accccggccg	80400
ggctggccgc	gcacctgctg	gagcgcgtgc	tcggcgccac	gaccggctcg	gcgagccgg	80460
ctcccgcgc	cgtaccggcg	gacgagccga	tcgcgatcgt	cggcatggcc	tgccgctacc	80520
ccggcggaat	cacctccccg	gaggagctgt	ggcgcgaggt	cggcgaaggc	cgcgacgcga	80580
tctccggggt	cccgaccgac	cgcggtggg	acctgcgggc	cctgttcgcc	gacgaccccg	80640
ggcggccccg	caccagccac	acccgcgaag	goggattcct	gcacgacgcg	ggcgagttcg	80700
acgccgactt	cttcggcatc	agcccgcgag	aggccctggc	catggacccc	cagcagcggc	80760
tgctgctgga	ggcgtcctgg	gaggccattg	agcgcgcggg	catcgacccc	gcactcgtgc	80820
gcggcagccg	gaccggcgtg	tacgcgcggc	tgatgtacca	cgactacgcg	gcccgcgtgg	80880
acgtcctgcc	cgacggcgtc	gagggtacc	tcggcaccgg	caactcgggc	agcatgcct	80940
ccggccggat	cgctacgcc	ctcggcctgg	agggccaggc	cgtcaccgctc	gacaccgect	81000
gctcctcgtc	gctggtcgcc	ctgcaactgg	cggctccggtc	gctgcgctcc	ggcgagagcg	81060
acctcgcgct	cgccggcggt	gtcaccgtga	tggccaacccc	gggtgtcttc	gtcgacttct	81120
cccgacagcg	cggcctcgcc	accgacggac	gctgcaagtc	ctacggcgcc	ggcgccgacg	81180
gcaccggctg	gtccgagggc	gtcggcatgc	tcctggtcga	acggctctcc	gacgcccgtc	81240
gcaacgggca	ccgggtgctc	gccgtcatcc	gcggcaccgc	cgtcaaccag	gacggcgcgt	81300
ccaacgggct	caccgcgcc	aacgggccc	cgcagcagcg	ggtgatccgg	caggcgtgg	81360
ccgacgccag	cctgcggccg	tccgacgtgg	acgccgtgga	gggcccacggc	accggcacct	81420
cgctcggcga	cccgatcgag	gtggagggcg	tgctcggcac	ctacggccag	gagcgggacg	81480
agccgctgtg	gctcggctcg	atcaagtcca	acatcgggca	cacgcaggcc	gcccgcggcg	81540
tggccgggtg	gatcaaaatg	gtcgaggcca	tcgggcacgg	tgtgctgccc	cgcacgctgc	81600
acggccagca	gccgtccccg	cacatcgact	gggcccggcg	cgcggtggaa	ctcctggccg	81660
agcaacggca	gtggccgcgc	acggaccggc	cgcgcggggc	cggcgtgtcc	tccttcggtc	81720

tgtccggcac	caacgccccat	gtggtgctgg	aacacacgga	ccacggcgac	caccactcgg	81780
ccgggaccgg	cgaacggccc	gccgtccccg	tgcccgtecc	cgtgccggtg	gcgctgtcgg	81840
cgcgtaccga	cgcgggactg	cgggcccagg	cggaccggct	cgccgccgcg	ctgaccgccg	81900
acccggacct	gatcccgcct	gacatcgcct	actccgcggt	gaccggggcg	gcccggctgg	81960
agcgcggggc	cagcgtcgtc	gcggcgaccc	gcgaggaact	gctggccgga	ctcggcgacc	82020
tgggcccggc	cacggtggcc	ggcgccgggc	acaccgcggt	cctgttcacc	gggcagggcg	82080
cccagcgccc	cggcaccggc	gaggaactgg	ccgccgccc	cccggtgttc	gccgccgct	82140
acgccgaggt	gtgcaccggc	ttcgacggcg	ttctcgaccg	gccgctgcgc	gaggtcgtcg	82200
ccaccggcga	cgtctcgcac	gacaccggat	acgggcagcc	ggcgtgttc	gcgctggagg	82260
tcgcgttcgg	acggctcttc	gagtctctgg	gcgtcgtacc	ggacttcctg	ctcggctact	82320
ccgtcggcga	actggccgcc	gcccactctgg	ccggcgtgtg	gtcgtgccc	gacgtgggtc	82380
gcgtcgtcgc	cgcccgcagc	cggctcatgg	ccgccctgcc	ggcgggcggc	gcgatggccg	82440
ccgtcagggc	gtccgcggac	gaagtggccg	ccgaactcgc	cgacggcgcg	gtgctcggcg	82500
cggccaacgg	cccccggtcg	gtggtcgtct	ccggggcagc	ggacgcggtg	ctggccaccg	82560
ccggtctgtg	ggccgcgcgc	ggctgccgca	cccgggaact	gaaggtgtcg	cacgccttcc	82620
actcgcacct	gatggagccc	atgctcagcg	acttcgcggc	ggccctggcc	gacgtggagt	82680
tccgcgcgcc	gcggatcccc	ctggtgtcca	cggtgaccgg	cgccgtcgcc	ggcgacgagc	82740
tgtgcacccc	cggctactgg	gtccggcacg	tccgggacac	cgtccggttc	gccgacggcg	82800
tccgcgccct	cggccaggcc	ggcgtggaca	ccgtcgtgga	actcggcccc	gacggcgtcc	82860
tcaccgcgat	ggccgctccg	ctgctgccgg	acaccgccgt	ggccctgccc	acgctgcgcg	82920
ccggacgccc	cgaggcgccc	gccgtcgccg	ccgcgctcgg	cgccctgcac	gaccggggca	82980
cggcggtgga	ctggccggcc	ttcttcaggg	gcaccggagc	ccggaccgtc	gaactgcgga	83040
cgaccgcgct	ccagcgaccc	cgctactggc	tggagtccgc	cgcacggacc	ggcgacctct	83100
ccgccgccgg	gctcgccgcc	gccgggcacc	cgctgctcgg	ggccgccgctc	gactcgcccg	83160
acgggatgct	gctcaccggc	cggtcgaca	cgcccaacca	tccttggtg	gccgaccaca	83220
ccgtcctgga	caccgtcctg	ctgccgggta	cggcctcgt	ggaactcgcc	cgcgcgggcg	83280
gcgaacgcgt	cggcctgccc	cgggtacggc	aactcaccct	ggccgcaccg	ctcgtactgc	83340
ccgcggacgg	ctccgtcctg	gtccaggtgc	acgtcggcgc	cgccgtcgac	ggcgagcggc	83400
cggtcaccgt	gtccgccggc	accgacgacg	gccaggactg	ggcacggcac	gccaccggcg	83460
tcctcgcccc	cgcggcggcg	gcgcccgggc	ccgacccgct	gccctggccg	ccgcgggacg	83520
ccgaaccggg	ggcgacggcc	ggcactacg	acgagctggc	cgccgccggc	ctcggctacg	83580
gccccgcctt	ccgcgccttg	cgcgcggtct	ggcggcgcg	tgacggcccc	gccgccgagg	83640
tgttcgccga	gatcggctcc	gcgccgggca	ccgatcccgc	gggcttcggc	gtgcaccggg	83700
cgctgctcga	cacggccctg	cacgcggccg	ctgtcggcgc	actcggcgtc	gcgggctgtc	83760
cgttcacctg	gaacgaggtg	gcggtgcaca	ccgcggcgcg	ccgtccctg	cgggtcggga	83820
tcgccccgga	cggaaagcggc	gggctgtccc	tgcgcgccac	ggacgacgag	ggccgcgccg	83880
tgatcgacgt	cggagcgctg	cggctgcgtg	cgatcgcccc	ggaggacctg	cgcacccgcg	83940
cggccgccgc	cgaggcgctg	ttcgacctgg	actgggcgcc	cgtctccgtg	gcgcggggcg	84000
cgcggccccg	cggacgctgg	gcggtgctcg	cccccgccga	cgcgagctg	accgacgtac	84060
tgggagccgg	gatcgagacc	gtcaaggacc	tgcgcacggg	gccggacgac	ctggacatcg	84120
tgctcgcggc	cgtgtccgac	aggggcgggg	ccggggtcga	tgcgcccacg	gctgcctccg	84180
tcggcgccgc	cgccgagacc	gacgtgcccg	aggccgtccg	ggccatcctg	cgcggcgcc	84240
tcggcctggt	gcagacatgg	gtggcggggg	agtccgccgc	ccggtggtc	ctgctcacc	84300
gcggcgcggt	ggcgtcgcgc	ggtgagcggc	ccgacctggc	cggtgccgcc	gcctggggcc	84360
tggtgcgcgc	ggcgcagtc	gagcaccccc	gcccgtcgt	gctggccgac	gtggacgacg	84420
acccggcgtc	cctggccggc	ctgcccgcgc	ccctggcgac	cggcgagccg	cagctcgtga	84480
tccgcgccgg	tgcggtccgc	gccgcccgte	tggtcgcgcg	cgcccccgcc	gaaccgcgcc	84540
taccgcgccg	cccgttgggc	gcacgcggca	cgggtgctgg	caccggagcc	accggcagcc	84600
tcggcaccct	ggtggtccgg	cacctggtcg	ccgagcacgg	cgtccggcgg	ctgctgctgg	84660
tcagccggca	gggacggcag	ccggagccgg	ccgcgggaact	cgccgcggcc	ggagccgaag	84720
tcogcttcgc	ggcctgcgac	gtcgcgcgac	gggacgcgct	ggccgcgctc	ctggcgagcg	84780
tcgacccccg	gcaccccgtg	acggccgtcg	tgcacgccgc	gggctgtctc	gacgacggcg	84840
tcgtcgcggc	gctcaccccc	gaccggctcg	acaccgtcct	gcggcccagg	gcggacgccg	84900
cctggcacct	gcacgaactc	accggcgagc	tggacgcggt	cgtgctgttc	tcgtccgccg	84960
cgggactcct	cggcgcaccc	ggacaggcca	actacgcggc	cgccaaacgcg	ttcctggacg	85020
ccctcgcggc	gcaccgcggg	gccgccggcc	tgcccgcggt	gtcgatcgcc	tggggccct	85080
gggccgacgg	catggccgca	cagctcgaca	cccggcgcg	ctcccgggcc	ggcctgctgc	85140
cgctcgacgc	cgactcggc	ctggccctgt	tcgacaccgc	ccgtaccggg	gcgccgaccg	85200
ccccgctcgc	cgcccgtctc	gacctgccgg	gactgcgggc	cgcggcggcg	gaactcccgc	85260
ccatcctcgc	gacctgggtg	cccgcccg	cccggacggc	gcccgagccg	gccgagccgc	85320

tcgcgacac	cctggccg	ctgcccgc	aggagcggga	acgccgggca	ctcgacgcc	85380
tcctgcgca	caccgcccag	gtgctcggcc	acgccaccgc	cgacggcgtc	gaccgcgaac	85440
gcggttcca	gcagctcggc	ttcgactccc	tgatgtccgt	cgagctgcgc	aaccggctgg	85500
gcgccgccc	gggcctgccc	ctgcccgc	cggtgatctt	cgatcacc	acgcccgcc	85560
cgctcggc	ccacctggtg	gccgaactgg	ccccggccg	ccccaccgtg	gcgtccctgc	85620
gcaccagcac	gctcgccgaa	ctggaggccg	cgccgaacga	gttcgcccgc	gaccccagac	85680
tgccgaggg	cctgcccacc	cggttgccc	cactgctgcg	cacctggac	gacccggcgc	85740
cggacgagcc	gctcgaggag	accggcgagg	agagcctcgc	cgaaactgctc	gacctcgccg	85800
accgggagct	gggggacttc	tgatgaccga	gcacgaccgc	acgaccgacc	gcgccaccgc	85860
acgaggagcc	accgcgttgt	cccaggaaca	gaccgacgac	cgcaaggtcg	tggagaccgt	85920
caggcgtctg	accaccgacc	tgccgcccgc	caagcagcgg	ctgcccggag	cggagaccgc	85980
ggcgcacgag	ccgatcgcga	tcgtcggcat	ggcctgccgg	tacccccgtg	gcgtcggctc	86040
gcccggaggac	ctgtggccc	tcgtcctgga	cggcccggac	gccatgggcg	cctccccac	86100
cgaccgccc	tgggacctgg	cgccctctt	cgccgacgac	cccgagcgtc	ccggcaccag	86160
ccacacccc	gagggcggat	tcctgcacga	tgccggcgag	ttcgacccc	gcctgttcgg	86220
gatcagcccc	cgcgaggcgc	tcgccatgga	cccgcagcag	cggtcctcc	tggagaccgc	86280
ctgggaggcc	gtcgaacgcg	ccggcatcga	ccccacctcg	ctgcccggca	gccgcaccgg	86340
cgtctacgcc	ggcgtgatgt	accacgacta	cgccaccggc	gccgaccctc	tgcccgaggg	86400
cgtcgagggc	tacctcggcc	tcggcaccgc	gggcagcgtc	gccagcggcc	ggatcgccta	86460
cacactgggc	ctggaggggc	cgcccgctcac	cgtcgacacc	gcctgctcct	cgtccctggt	86520
cgccctgcac	tcgctgtcc	gggcgctgcg	cgccggggag	tgcacgatgg	ccctggcggg	86580
cggcgcgac	gtgctgtcca	ccccggcctc	cttcgtggac	ttctcccgtc	agggcgactc	86640
cgccgcccag	ggagctgca	agtcctactc	cgccgagccc	gacggcaccg	gctggtccga	86700
gggcgctcgg	atgctcctcg	tcgaacggct	cgccgaccgc	gaacggctcg	gccaccccgt	86760
cctcgcctg	ctgcccggct	cagcggctca	ccaggaccgt	gccagcagcg	gcctcaccac	86820
ccccaacggg	ccggcccagc	agcgcgtcat	ccggcagcgc	ctcggccgac	cccggctcac	86880
ccccgcccag	ctggacctcg	tggaggggcca	cgccaccggc	accccgtcgc	gcgaccgat	86940
cgaggtgcag	gccctgctcg	ccacctacgg	ccaggaccgc	gccgaaccgc	tctggctggg	87000
ctcgggtgaag	tccaacatcg	gccacaccca	ggccgcccgc	ggcgtcgccg	gagtcatcaa	87060
ggccgctcctg	gccctccggc	acgggtgact	gcccggcacc	gcccacctga	ccgagccgac	87120
cccgcaggtc	gactggaccg	ccggcgccgt	ggaaccgctg	cgggagacgc	gcgcctggcc	87180
cgagaccggc	aggccgcgcc	gcgcccggct	gtcctcgttc	ggcatcagcg	gcaccaacgc	87240
ccacatcgtc	ctggaacagg	ccccggcccc	cgccggcggc	caggcggccc	gagcccaggc	87300
gcccgcggcg	ccgcccggcc	tcgggaacca	ggccaccgc	gcgcccgggt	ccatggagga	87360
ccggaccggc	gcccgcctca	gcgcccggcg	agaccggacc	ctcaccgcgc	cgccccctc	87420
cgcccggcgc	cccggcccgc	ccgcccctcc	cgtcccgcgt	tcgcccgcga	ccgagcccgc	87480
tgctcgtgcc	caggccctcc	ggctggccgc	ccacctcacc	gagcaccccc	aactcgcccc	87540
gcaggacatc	gcgttcagcg	ccgcccaccac	gcgcccggcg	ctggcgtccc	gggcccgtcgt	87600
gctcggccgac	gaccgggccc	ggctgctgga	cgccctcacc	gcgctggccc	agggacggcc	87660
cggccccggc	gtcgtcaccg	gcgcccggcg	ggcccggcgc	cgccggatca	ccttcgtctt	87720
ccccggccag	ggcggcccag	gggcccggcat	ggccgtaccc	ctgctggaga	cctcggccgt	87780
gttcgcccgg	aagtgggccc	aatgcgccc	cgtgctcggc	ccctgggtgg	actggtcgcc	87840
cgacgaggcg	ctgcccctcac	cgcaggcact	ggaaccgggtc	gacgtcgtcc	agcccgtgct	87900
gtgggcccgtc	atggtcagcc	tcgcccagct	gtggcggggc	gcgggctac	ggcccgacgc	87960
cgtactcggc	cattcgcagg	gcgagatcgc	cgcccctgc	gtcggcccgg	ccctgtcctc	88020
ggaggacggc	gccaaggtcg	tcgcccgtcg	cgccaaagggc	ctgctcgcgc	tcgcccggcc	88080
cggcggcatg	ctctccgtcc	cgtgcccga	ggccgaggtc	cgcccggcgc	tcgacagccc	88140
gcccggcctc	ggcatcggcg	ccgtcaaccg	gcccggcacc	gtggtggtct	ccggcgagac	88200
ggccgcccctc	gacgaggccc	aggccgccc	ggaggccgag	ggcgtccggg	tgcccgcgat	88260
ccccgctcag	tacgcccctc	actccccgca	cgtcggccgag	gtgcaggacc	gcctcggccc	88320
cgacctcgcc	ggcatcggcc	cgcccggcgc	cgaggtgacc	ttcctgtcca	cgtcaccggg	88380
ggaacccttc	gacaccaccg	gactcgacgc	cggtactgg	taccgcaacc	tgcccgagca	88440
ggtccgcttc	gagggcgcca	cccggcgcgc	cctggagcag	ggccaccgcg	tgttcacgca	88500
ggtcggccc	caccccgtgc	tcacgctcgg	cgtccagcag	accgcccagg	ccatggacgt	88560
gcccgcccag	gcgatcggca	ccctcggccc	cgaccagggc	gacctgctcc	gcttcggcac	88620
cgcgctcggc	gagggcggcg	tcctcggcgc	ccccgtcgac	tgggcccggc	aactcgcccc	88680
gtacgccc	cgcccgggtcg	acctgcccac	ctacgcccctc	cagcggcggc	gctactggct	88740
caccccgcag	cggcggccc	ccctggccc	cagcaccggc	accgaccctg	gggagggccc	88800
gttctgggac	atcgtggacc	gcgcccagc	cgaggaaactc	tcccgcgcgc	tcggcgtcga	88860
cacggacgac	ccgctcaccg	cggtcgtgcc	ggccctcggc	cgtggcggcc	gcctcctgcg	88920

cgagcgggtcc	gccgtggaca	gctggcgcta	caccgtcgcc	tgggagcggc	tcgccgtccc	88980
cgacagcgcc	cggtgaccg	ggcgtggct	gctcgtcgta	cccagagccc	gggcccgtga	89040
cgccctcgcc	gcgcggtgcg	ccgccgcgct	caccgggac	ggcggcggg	tcaccacgct	89100
caccctgggt	gcggacgaca	ccgaccgcgc	ggcgtcgcg	gcccggctga	ccggcctccg	89160
cccggcccggc	gtgctgtccc	tgctcgccct	ggacgacgcc	ccgcaccccg	cacaccggc	89220
cctgcccacc	ggactggccc	tgaccgtcgc	cctcgtacag	gccctcggcg	acgccggcgt	89280
caccgcgccc	ctgtggtgcg	ccaccgggg	cgccgtcgcc	accggccccg	ccgaccccgt	89340
caccgcaccc	gtccaggcgc	agatctgggg	cctgggcccg	gtcgtcgcg	tggagcacc	89400
cgaccgctgg	ggcggcctga	tcgacctgcc	cgccgagtgg	gacgaccgcg	ccgcaggccg	89460
gctggcccgc	ctgctcgccg	caggcggcga	cgaggaccag	acggtgctgc	gcccaccctg	89520
cgtgtccgcc	cgccgcctga	cccgcgcgcc	gctcggcacg	gagcccggcc	cccggcccgt	89580
ggcgtgccc	ggcaccgtcc	tgctcaccgg	ggcaccggc	gccctcgggc	cgcacctggt	89640
gcgctggctg	gccggacagg	gcgccgagcg	catcgtgctg	cccggccgcc	gcggcgccga	89700
cgccccgatc	gcggccgagc	tgaccgccga	actggcccgc	accggcaccg	aactccactt	89760
cccggctctg	gacgtcaccg	accgcgcaga	actggcccgc	ctcatcgacg	gactcgacac	89820
ggccggcacc	ccggtgcgca	gcgtcctgca	cgccgccacc	gccctcgaac	tgcgcccgct	89880
ggccgacacc	cccgtcgaca	cgttcgccgg	ccagacggcc	gccaaggtgc	tcggcgcccg	89940
gcacctcgac	gaactcttcg	ccgaccggga	cctggacgcc	ttcgtgctgt	tctcctcggg	90000
cgccggcgtc	tggggcagcg	gactgcacgc	cccctacgcc	gccgccaacg	cctacctcga	90060
cgccctcgcc	gaggaccgcc	gggcccgcgg	cctgaccgcg	acctccgtcg	cctggggcat	90120
ctgggcccgc	gtcaacgagt	gggacggcgt	gaacgggac	gtcgaccccg	aacgcgtcgg	90180
ccggcagggc	ctgcccttcc	tcgaccggga	cctcgcctgc	gcccggcctg	ggcaggcgtt	90240
cgaccacgac	gagacgaccg	tggtgatcgc	cgactcgac	tggaccgggt	tcgtcccggg	90300
gttctgctcc	gcgcgcccgc	ggcgtgctgt	ggagtccgtg	cccgaggccg	ccgccgtcct	90360
gcgcaccgcc	accgtcgaca	ccggaacggt	gtccgcgctg	cgcgagcggc	tcggctccct	90420
gcccgaggcc	ggcccgcggc	gggcccgtcac	cgaactggtc	cgcgagcacg	ccgccgcctg	90480
cctcggccac	gactcgcccg	ccgcgtccc	cgccgaccgc	gccttccggg	acgtcggtt	90540
cgactccatc	accgcggctg	agctgcgcaa	ccggtccgc	tcggccaccg	gcctggccct	90600
gcccgccacc	ctcgtcttcg	accaccgctc	gcccaccgcg	ctggcccggc	acctgctcgc	90660
gctcgccttc	gacaccgccc	cgccggacct	cgccgcgccc	gccgcccgcg	ccgccgacga	90720
tgacgacgac	ccgatcgccc	tcgtcgccct	cagctgccgc	taagcccggc	gcgtgcctc	90780
cccggacgag	ctgtggcggc	tcgtcgtggc	cggtcaggac	gcggtgggcg	ccctgcccac	90840
cgaccgtggc	tgggacctcg	actcgtctca	cgactccgac	cccgacgccc	gcggtcgag	90900
ctacgtcccc	caaggggccc	tcctcaccga	cccgcggc	ttcgacgccc	ccttcttcgg	90960
catcgcccgc	gcggaggccc	gggcccaccga	cccgcagcag	cggtgctccc	tggaagccgc	91020
ctgggaggcg	ttcgagcacg	ccggcatcga	cgccaccggc	ctgcgcggct	cgcgctcgg	91080
cgtcttcgcc	ggcgccaacg	tcggcgacta	cgccccagc	cgcgccctg	gcgcccggcg	91140
ctccgacgga	cagctgctca	ccggcaacgt	cccagtggtg	atctccggcc	ggatctccta	91200
caccttcggt	ttcgaggggc	cgccgcgtcac	cgtggacacc	gcctgctcgt	ccgccctggt	91260
cgccctccac	ctggcctgcc	ggtcgggtgcg	cgccggcgag	agcgacatgg	ccctggccgg	91320
cggcgtcgcg	ctcatgtcca	gcccggcccgc	cctgatcggc	ttctccgcgc	agcgccgct	91380
gtccggcgac	ggccgctgca	aggccttcgc	cgacgcggcc	gacggcaccg	gtctcgccga	91440
gggcgtcgga	ctgctgctgg	tggaacgccc	ctcccgggccc	cgcgcccagg	gccaccgct	91500
cctcgccttc	gtacgcggct	cgccgatcaa	ccaggacggc	gcctccaacg	gactcaccgc	91560
cccagcggga	cccgccccagc	agcgctcat	caccgcggcg	ctcgccgacg	ccgggctgcg	91620
gcccggcccag	gtcgacgccc	tggaggcccga	tggcaccggc	accgcctcgc	gcgaccgat	91680
cgaggcccag	gccctgctcg	ccacctacgg	acaggaccgc	gccgaaccgc	tctggctcgg	91740
ctcgggtgaag	tccaacatcg	gccactccca	ggccgcgctc	ggcgcgcccg	gcgtgatcaa	91800
gaccgtgcag	gcgctgcggc	acggcctgct	gcccgccaccg	ctccacgtgg	accggcccac	91860
caccaggtc	gactggaccg	ccggcgcctg	cgaggtgctg	accgagccc	gggactggcc	91920
ggcgtggac	cgccctcggc	gggcccgcctg	gtcggcgttc	ggcctgtccg	gcaccaacgc	91980
gcacgtgatc	ctcgaacagg	ccccgcggga	agacgcccac	ccggcccccg	aaccggcccc	92040
gggcgaggac	tcccaccgga	ccccgaaac	ggcccagggc	gaggacgccc	cgcgaccgc	92100
gcccgagccc	gcgcggcccg	tggtgtggcc	ggtgcaacggc	cgtaccgggg	acgccctgcg	92160
cgcccaggcg	gcgcggctgc	gcaccacct	ggagaccccgc	cccgacgccc	gcccggccga	92220
cgtcggctgg	acctcgcgg	ccggtcgggc	cgtgttcgac	caccgcgccc	tggtgctcgg	92280
cgccgaccgc	gccgagctgc	tgccgggact	cgacgcctgc	gcccgggca	ccccgacc	92340
cgccgtcgc	gacggcgccg	cccagggcgc	cgaccgggccc	gtgttcgtct	tccccggaca	92400
cgccgcccag	tggcccggca	tggcccggcg	gctcttcgac	gacttcccgg	tgctccggga	92460
aagcgtgctc	cagtcgcggc	acgccttcgc	cgagttcgtc	gactggtccc	tgctcgacgt	92520

gctccgcgac	gaggagggcg	cgccgcgcgt	gcaccgcgtg	gacgtcgtcc	agcccgcctt	92580
gttcaccatg	atggtgtcgc	tggccgcgcgt	gtggcgctcg	tacggcgtgg	agccgtcggc	92640
cgtggtcggc	cacagccagg	gcgagatcgc	cgccgcctac	gtggcggggc	ccctcgacct	92700
gcgggacgcc	gcccggatcg	tcgccacccg	cggcaaggcc	tggctcacc	tggccggcac	92760
cgggcgcatg	gcctccgtgg	cgctgccccg	cgccgaggcg	gccgagcggc	tgcgcccggt	92820
cggccaccgc	ctggacatcg	cgcccgctcaa	cgaccgcgcg	tcggtcaccg	tcgccggcga	92880
cctcgacgcg	ctggaggagt	tcctgaccgg	cctggagacc	gagggcgtcc	gcgtccgccg	92940
ggtccgccag	atcgtcggcg	cgggccacac	cgcgcacgtc	gacgcgctgc	gcgaccagct	93000
gatcgagacc	ctcgcgccga	ccgcgcgcgc	ctccgcgcgc	atcgcgttct	gctccaccgt	93060
caccggcgga	ctgctcgaca	ccgcgcggcct	cgaccaccac	tactggtacc	gcaacgcctg	93120
ccggacggtc	ctgttcgaac	aggcgggtccg	cacgctggcg	gagcagggct	gagccccctt	93180
cctggagatc	agcgcgcacc	cgatgttcac	cgtcgccgtg	caggagacc	tggaggacgc	93240
cgggcgtcggc	gccgcggctc	tcgccacgct	gcgcgcgcgac	gagggcggcc	ccgaccggtt	93300
cctgcgcgcc	gccgccgagg	cccacacggc	cgggcgtcacc	gtcgactggc	gcccggcctt	93360
cgccggcgcc	ggagccccta	cgaccgacct	gccgacgtac	gccttccagc	gcaccgggca	93420
ctggctggag	ccgggcgggc	acagcggcga	cgtgaccgcc	gcaggcctgg	caccggccgg	93480
ccatccgctg	ctgggagccg	tggtcgaact	cgccgacgga	gccatggtcc	tgaccggccg	93540
gctctccgcc	accgcgcagc	cctggccggg	gcccgaggcg	ccggacgccc	cgctggtcga	93600
cctcgtcctg	gccgccgccc	accgggccga	ctgcgcgggc	gtcggcgaac	tgaccgtgta	93660
cgaaccgctg	atggcgcggg	ccggcggcgc	cgacaccggg	gtgaccgtcg	gcgaccccga	93720
cgcggtgggc	agccgcacgg	ccgccgtcca	caccgcacc	gcggacagcc	cgtgggtgcg	93780
ccacgcccag	gccacgctcg	tcgccccggc	ggccccggcg	gagtccctga	ccgactggcc	93840
gcccggccggc	cgggaaccga	tcgaccggcc	cgcgaggcgc	ggccgcgccc	gtgtcgccgt	93900
gaccgcccgc	tggcagcgcg	gtgacgacct	tttcacagag	gtggcgcctcg	acgacggcgc	93960
cgccgagcgg	gccgacgcgt	tcgcgctgca	tcgcgctgctg	ctggacgccc	ccctgagccc	94020
gctgctcgac	ggcgagctga	cgcccaccgc	gtggtcgggc	gtccgcctgc	acgcgaccgg	94080
ggcgaggacc	ctgagagcgc	gcgcggaacg	catcggcccc	gacaccgtgg	cgctcaccct	94140
ggccgacccc	gagggcggcc	cggtgctgac	cgccgacacc	gtccgtctcg	ccgccgcgcc	94200
cgccgcggga	gccggccggc	gcaccggcaa	ggaecgccctg	ttccggatgg	agtgggtgcc	94260
ggcgccgctc	gccccggcca	cgcccggccc	ctgggcccgtc	ctcggggccc	accgcctcgg	94320
cgccgcggac	accctgcgct	ccctcggcca	caccgtgcac	ggcgccgacg	gcccggccgg	94380
cctgaccgag	gtgcccgacg	cggtgctgat	caccgcgcgtc	ggcgccccgg	gcgaggcgcc	94440
cgccgaggca	cgaccctgct	tgacggcac	gctggcggcc	ctccagaccc	tcctcgccga	94500
cgaccggttc	accgcctgac	cgctcgtggt	cctcacccgg	ggcgccgccc	ccgaccgggc	94560
cgccgcccac	ggcctcccgc	cgccctgggg	cctggtgcgc	tcggcaccagt	cgagcacc	94620
gggcccgttc	gtcctggccc	acatcgacga	cgacccccgc	tcctggcggg	cgctggcgcg	94680
cgtaccggcc	accggggagc	cgcagctggc	gctgcgggccc	ggtgcggcca	ccgtcccgcg	94740
gctggcccgg	ctcgcccccg	cggacgcccc	ggcgccgtgg	gacccggacc	gcaccgtcct	94800
cgtcaccggc	gcctccggag	acctcggcgc	gctggtggcc	cggcacctgg	tggccgccc	94860
cggggtacgg	cacctgatcc	tcgcctcccg	gcgcggcccc	gccgccccgg	gcgccgcggg	94920
actcggcgcc	gagctgcggg	cctcggggcg	cgccacggtc	acgatcgcg	cctgcgacac	94980
cgccgaccgc	aaggcactcg	cggaactgat	cgccgcgcgt	cccgcagagc	accgcctcac	95040
cgcggtcgtg	cacagcgtct	ccgtcctcga	cgacgggggtg	atcgccgccc	tcgaccggga	95100
ccggctggac	accgcgctgc	gcccgaaggc	cgacgcggcc	tggcacctgc	acgaactgac	95160
ccggcatctg	aacctgtccg	ccttcgtgct	gttctcgtcc	gtggcgggca	ccttcggcgg	95220
cctcggccaag	ggcaactacg	cggcgggcaa	cgcctcctcg	gacgcgctgg	cccgccaccg	95280
cgcgccccac	ggcctcccgc	cgacctccgt	ggcctggggc	tgggtggcg	aacgcgcgcg	95340
gaagagcggc	cacggcactg	ccgcgcaggc	gcccgtcgcc	gtcaacggca	tgaccccgt	95400
gaccgaggac	cacggcctcg	ccctgttcga	cgcggcctgc	cgcggcgacg	agccgttcgt	95460
ggtggcccgg	gcgctgcacc	tcgcctcgct	gcgcgcgcgc	gccgacgaac	tgcccggcgc	95520
gctgcgcggc	ctcgtgcggg	cgcccgcgcc	caaggccggc	gccagggccg	ccgggaccgg	95580
cacgcccacg	gtggccgggg	aactggccgg	ccggccgccc	gcgcggcgcg	aggcgttcc	95640
cgtggacctg	gtgcgcgacg	agaccgcgct	cgctcctcgg	cacaccggcc	gggacgacgt	95700
accggccccac	caccggttcc	tggaccaggg	attcgactcc	ctggcggcgc	tgaagctgcg	95760
caaccggctc	gccgccgcca	ccggactgcg	gctgcccgcg	acgctcgtct	tcgaccacce	95820
gaccccgcag	gagctggccc	gccatctgct	ggccgaactc	gtcccgcggg	cggacgcgga	95880
accgcccggc	ttcgacgacg	aggacgcccg	cacggccgtc	ctggtcctgg	aggaactcgg	95940
ccagttggac	gagggcctca	cacggctgcc	cgcgcgaggg	cccgaacgca	cgcgtatcgc	96000
gggcctgttg	acagacctgc	tggccaggtg	gggtcgatga	ggttcgcgac	aacctggccc	96060
acaggccctg	gaccgtccta	tagcatcacc	gtgatcgcta	cggcaggccc	ggccgaccac	96120

cgccccggcc	gcaccgcgcc	acggggggcg	gtacgactga	gggcgatcga	tacgcgtgag	96180
caggcaggca	cgtccgccgc	acgcacgcac	acaaggggga	actgatggag	aacgttcagg	96240
cctgggttct	gccccgccgg	ccgcaggaca	ggtccggcgg	cccgttgccg	ggcgagctga	96300
ccctcaccgg	catccgggtc	gcctccccgg	gcgagcacga	ggtactggtc	gagtcctcgc	96360
tggctgctg	ggaggccaac	atggagcacg	ccctcgccgg	cagccccgtc	gacatcgtcc	96420
ggcagcgccg	cgaggagcag	atcgtcatcg	gcacctgccc	cgtcgtccgc	gtgctcagca	96480
ccggttcggc	cgtgcgcgga	ctgctgagg	gccaggaatg	cctgtggatc	ccgttcggcc	96540
acatcgaccg	caacggctac	gcggagacca	tctgcgccta	cgacgcgccg	ggttcgcccg	96600
gtctgctggc	cgagcgcgat	gtcgtccccg	ccgaccggct	cgtgccgctg	ccggccgacg	96660
gacgctacc	gctggagcgc	tgggcgcgct	tcccccggta	cttcaccgcg	tgggacaact	96720
ggcgggtcac	cagccgggtc	tggcgcagcc	aggtcgacga	cgccctgggac	gagcgcgcgc	96780
tcgctcctcgg	ctggggcggc	ggggtcgtct	tcgcccgaact	cgaaactggcc	cgccgcgagg	96840
gcttcgccac	ggcgtgggtg	tccggccggc	agagccggct	gaaggagatc	gccgcgtccg	96900
gcgcgatcgc	cgtcgaccgc	cgtgagttcc	gcgacctgga	ctacctccgg	gccaaaggagt	96960
ccaccgaccc	cgacgccatg	gaccgctacc	gcgcctccga	ggccccggtc	ctggagatcg	97020
tggcgaact	ctccggcggc	cacggcgtct	cggctctcct	ggacaacctc	ggcggcgggc	97080
tgcaccgggc	cacgatgaag	agcctcgccc	gcgagggcgt	ggtgtccacc	gtcggctgga	97140
agaccggcat	gcggctgtgg	aacctgcgcg	ccaccgagtg	catcagccgc	cacatccacg	97200
tccacaccca	cgcttgccgc	caccaggacg	ccccccgcat	ccgtgacgtc	atgcaggaga	97260
ccggtggct	gccccacata	gccgacgacc	cggtgaccgc	ggttcgcccgc	gtccccgaac	97320
tggccgactc	ctaccgccgg	gacgacctcg	acacctactt	cccgtgttcc	tccggcgccg	97380
ggaggggtag	acccgatgag	cccgcgttcc	tacgacgtgt	cggtcgtcgg	catggcctgc	97440
gcgttccccg	gtgcaccgga	cctcgaccgg	tactgggcca	acgtggtggc	cggtgtcgac	97500
tccgtcaccg	aagtaccccc	ggaccgctgg	ccggccgacc	gccactgggg	cggcgccgac	97560
gcggcccccg	gcgagcagtc	gccgtcgaaa	tggggcggat	tcctgccggc	cctgcccttc	97620
gacgccctcg	cccacggcgt	cccgcccaac	tcctcggcgg	gcatcgagac	cgcccagctg	97680
ctcgccctgc	acaccgccga	ccgggcaactg	gccgacgccg	ggtaccccga	gcggccgttc	97740
gaccgggaga	ccacctcggg	gatcttcgcc	gcggaggcgg	gggccgacct	cgctccgcgc	97800
tacgcggcgc	gctcgatgct	cggccagcac	ctggagtcgg	tgccggagga	actggacgcc	97860
cggctgcccc	ggctgaccga	ggactccttc	cccggcacc	tcgccaaactg	catcgccggc	97920
cgcatcgcca	accggctcga	cctcggcggc	cccacctaca	ccgtcgacgc	cgctgcgcgc	97980
tcctcgctgg	ccgcgctcga	ccaggcctgc	aaggaaactcc	tcgccgaaac	aagcgacatg	98040
gtgctgtgcg	gcgccgtcga	caccacaac	gcactccacg	accacctget	gttcggctcc	98100
gcgcggggac	tctgcgccac	cgcccgctgc	cgcctctcgc	acgcctccgc	cgacggcctc	98160
gtcctcgccg	aaggcgtcgc	ctgcctcgtg	ctgaagcggc	tcgcggacgc	cgaacgggac	98220
ggcgaccgcg	tctacgccgt	gatcaaggcg	gtcggcgccg	gcagcgacgg	ccgcggcctc	98280
ggcctgaccg	cgccccgccc	cgagggccag	cggcgcgcc	tggaacgcgc	ctacgccctc	98340
gcgggcctgt	ccccccgcga	ggtcggcctt	gtcgaggcgc	acggcaccgg	cacggtgctc	98400
ggcgaccgca	ccgaactgga	gacgctcacc	gaggtgttcc	gcacggccgg	cgccggtccc	98460
ggagcatgct	cgctcggctc	ggtgaagtcc	ctcatcgccc	acgccaagtg	cgcgccgggc	98520
atggcgggac	tgatcaaggc	cgtactcgcc	gtgcacaccg	gcgtccggcc	ccccaccggc	98580
ctcaccgagc	ccaacgcggc	ctgggacccg	gcgaccagcc	cattcgccctt	cgacaccacg	98640
gcccccccct	gggcccagcc	ccgacgcgtc	gcaggagtca	gcgcgttcgg	cttcggcggg	98700
gtcaactacc	acgccgtcgt	ggccgcccc	gccccggacc	cgtcgcacga	gagcgccccg	98760
aggtccgccc	gactgctcct	gttccgcggc	accgagacg	aggtccaaac	cgccctgccc	98820
gacctggccc	gcccggctggc	cgagggcacc	gtacggccgt	ccgccctcac	ggggtcacc	98880
ggcgcggccc	cgccccgggt	cgccctcgtc	accgagggcg	ccgccacgct	gggagaacag	98940
ctcgacctgg	ctctgcgcct	ggaggaccgc	cccgagcgcg	gcgtccacgt	gggagacggc	99000
gccccgcagc	ccgtcgcgct	gctcttcccc	ggccagggca	gccagcggcc	cggcgatgctg	99060
gccgatctct	tcgtcgcctt	cccgcgctg	caccggttcc	tgcgcgccga	accggcactg	99120
accggcgccc	tcttcccgcc	cgccgccttc	ggcggcacac	ccccgcgctt	ggtcgacacc	99180
cggctcgccc	agcccgcctt	cggtctcgga	gcactcgccg	ccctggacct	cctgcggggc	99240
tgcggcgtcg	aaccgacat	gaccgcccgg	cacagctacg	gcgaaactgcc	cgccctcgcc	99300
gccgcccggc	ccctccccga	gacggccgta	atgccgctca	gcggcgcccc	ggccgaggcc	99360
atgcacaccg	cggcgggcca	ggaccccggc	gcgatggccg	cggtcgcccg	gcccgcgggc	99420
acggtcgccc	aactgctcgc	ccggcacggc	ctggcggacg	acgtacccc	ggccaaccgc	99480
aacgcccccg	gccaagtggg	gctcgcgggc	tcgcccggcg	ggatcgacgc	cgccgtcgcc	99540
gtactccggg	aggccgggca	caccgccaag	cggctcccgg	tcgcccgggc	cttccacagc	99600
ccgcggatgg	ccccggcgtg	cgacgccttc	gcccggcgcg	tggcccggca	cgacctgcgc	99660
gcgccccggc	tgcgggtctg	gtcggggccc	accgcccggc	cgtaccccgg	ggaccccag	99720

ES 2 556 211 T3

gcgatacgc	cccagctcgc	cgggcgaactc	gcccgcgccg	tgcggttcgc	cgagcagatc	99780
gaggacatgt	acgcggcccg	cgctcgcgctc	ttcgtggagg	ccggcccggg	cacggtcctg	99840
acccgcctgg	ccggcgaggt	gctcgcocggc	cgcccgcaca	ccgccctgtc	cgtggaggcc	99900
ccgggcccgc	cgggcctcgc	ccatctgctc	ggcaccctcg	ccgcgctcgc	cgtacgcggc	99960
acacccttac	ggctggacga	gctgaccgctc	gcccgcgacg	agccggtgtc	cgcgcccccc	100020
ggatggaccg	tcgacggcca	cctggtecgc	accgcggacg	gcacgcccgt	ccccggcgga	100080
ctccagcccg	ccggcccgc	cgtgaccctc	accggcggcg	gcgtgcccg	cagccgcgac	100140
gagaccgtca	tcgaattcct	gcgcggcacc	cagcggatcg	tggccgcaca	gcaggaggtc	100200
atgctccgct	acctcggcac	gcaacccccg	gccccggagc	cgccccccgc	gcacgaccca	100260
ccggaatccg	ccgcgaccgc	cgcgcccacc	gtcctggacc	tggaggccgc	gcgcaccggc	100320
tacccgcccg	ccatgctgga	cccggactcg	gacctggaag	ccgccctcgg	cgtcgactcg	100380
ctcaagcgca	cggagatcgc	ctccgcctcg	gtccgcggca	ctccggccgc	ggacgcgctc	100440
ggcgaactcg	ccgacctgcg	caccatccgg	gcgatgaccg	actggctcgc	caccgcgacc	100500
ggccccgacg	gccccggccc	cgacggcccc	ggccgccaga	cggccgtgcc	ggagcccgaa	100560
accatgcccg	ccggcgctc	ggacaccgtg	cgccacgtcg	tcgaaccctg	gcccgaggaa	100620
ccgcccgcg	gtccgctgcc	gaccctcacc	cgggcccgcg	tcagcggcgg	cgggtgcgctc	100680
ggcgacgtgc	tggccacgct	gctcaaggaa	cgcggcaccg	aggtcgtcag	cgaccccgc	100740
ggctgcgacg	ccctcctgct	gctcgatgcc	ctcgacggcg	gcgactaac	cctgcccggc	100800
cggttcaccg	acataagggc	ggccgtcctg	ggcggtttac	gcaccctgct	gctggccacc	100860
tgccacgacg	cgggaccgc	cggcagcggg	gtgcacggcc	tcgcccgcgc	cctctcccgc	100920
gaacaccccg	gcctcgcctg	gaccgcctc	gacctgcccg	ccggacagcc	ggccgaggag	100980
gcggcgcgca	ccctgctcgc	cgaactcggc	ggcagcccgc	cgtccgtgac	ccacaccgac	101040
ggccgcgggg	ccgtctggcg	caccgcggcg	gccccgtcgc	ccgcccgcga	cgtcaccgcc	101100
ggcgacctcg	gcctggaccg	cgactccgtg	gtgctcctca	ccggcggcgc	gcgcggcatc	101160
accgcccgcg	tcgcccgcgc	cctggcgacc	ctcaccggtc	gccacctcga	actcgtcggc	101220
cgcagcgaac	ccgtcaccgg	gacggtcctc	accgacgccc	atctgcggac	ccgcctgatc	101280
gcccagggcg	ggcgcgacc	ggccggcatc	gaacgcgcgg	tgcgtgcca	cgccgcgggc	101340
cggcaggtcc	ggcagtgctc	ggacgacctc	gcccggcccg	ccgcctccgt	gcgctaccac	101400
cgcgccgacg	tacgcgacc	cgagcggctg	ggcgccgtcc	tggacgacgt	gtacgcccgg	101460
cacggcccggc	tggacaccgt	cgtccacgcg	gcccggccagg	tcgcccgacc	gctgctgcgg	101520
gacaagagcc	ccgacgagtt	cgccgaggtg	tacgacacga	aggtcgcggg	cgcccggggc	101580
ctggccgccc	ggctgcggcc	gggcctgcgc	cacctggtgc	tcttcggcag	catcgccggg	101640
gtgaccggca	accgcggcca	gaccgactac	gcccggcca	acgacgcctc	ggacaccatg	101700
gcccggcagt	ggtccggccg	ggtcgcgcg	cgggtcctcg	ccctggactg	gggcccctgg	101760
gcggccgacg	cgggcccgat	ggtagccgcc	gaaactggaac	gggcctacgc	ccgcaaccgg	101820
atcggcctga	tcgaccggga	cgacggcgta	ggggccttcc	tgcgcgaact	ggccttcggc	101880
cgcgacccgc	aggtgctgct	caccgtcggg	gaccccgcgc	ggttcgggag	cgcccttgac	101940
tgacatactc	ctcgcgcgg	cgccgggccc	cggagccctg	ggcgcggccc	tgcgttcggg	102000
ccggcgcggc	ggcgcgggac	cctgcccggg	cgccgtgctc	gaccccacgg	ccgagcggat	102060
cgcacaggcc	ctcgcctcga	tcgaccgggg	cgaaccgtgg	cacggcggca	acgacatccg	102120
gttcgcgccc	cggcccgggc	cgccgggccc	caccgtcctg	ctcttccccg	ggctcgacgc	102180
ccgtaccggc	ccgccaccgc	acgacgtcct	gcgctggctg	ggccgccccg	ccggaccggc	102240
gaccggcgcc	gccaccctcg	ccgagcgcac	ccacgccacc	ctgcggtca	accgggtgct	102300
ccacgcggcg	ctgctgcggg	ccgggatcgt	cccggacgcc	atggcggggc	acagcgcggg	102360
ggagtggagc	gcctgttcg	cggcgggact	cttccccgag	acggcgttcg	acgagttcgc	102420
ccggaccatg	ggcgcggccc	cgggtccccg	cgtggtcttc	gcccgcgtcg	gectgcccgc	102480
cgaagcggcg	gcccggctca	ccgagggaga	gcccgcggctc	gtgctgtccc	acgacaaccg	102540
gcccgaccag	agcgtgctgt	gcggcgacga	ggacgcggctc	acacgctgcc	tggaccggat	102600
cgcgcacccc	gaggtgccgc	gcagagtgct	gccgttcggg	tcgggggttc	actcgccgat	102660
gctcgccccc	tacctgccgg	gcttctcggc	cgtgttcgac	cgggtgccgc	tcgcccggcc	102720
cgcgctcccc	gtctggtccg	cgaccaaccgt	cgccccctac	ccggacgacg	agaccgagct	102780
gcgcgcctctg	ctcggccggc	acctggtgga	acccgtccgg	ttccgtgccg	tgatcgagcg	102840
gctgtacgag	cagggcgccc	gcgtcttcgt	gcaggcccggc	tgcggcagcc	tcaccgcctt	102900
cgtcacggcc	accctcgcgc	gcccggcccc	cctcacggtc	agcgcgggct	cggcgacccc	102960
gccccaaactc	gaccagctgc	gctgggcccgc	cgccgcgctg	tggaccggcg	gtcacacccc	103020
ggacctggcc	gcctcggcc	tcaccggggcc	cgcggcggac	ccggtgccga	gcaccccgcgc	103080
cgtgaccggc	gcgcacccga	gaaccgggccc	gtaccccgcg	acccaccggg	tggccggcgc	103140
ctacgagggc	gccgtcgcgc	cggtagcgcgc	ctccgtggag	agcgtctacc	ggacctggaa	103200
ggagcggaca	tgagcagccg	gcgacacgcg	gtcgtgaccg	gatcgtcccg	cggcatcggc	103260
gcggccgctcg	ccgcccggct	ggcccggggcg	ggctgggcgg	tcaccggctg	ggaccggaca	103320

cccggcgaca	ccaccggcct	cgccgactgg	cgcgaggtgg	acgtctccga	cgccgagtcc	103380
gtccaggccg	ccgcccgcga	ggtgaccgcg	accgacctcc	tcgtcaacag	cgcgggcgtc	103440
ggcgccatcg	ccccgtccgc	ggagctgaag	ccccgcacct	gggaacgggt	cgctggcgtc	103500
aacctcagcg	gcacgttcta	ctgcgcccag	gcaactgttcc	ccgccctgtc	cgcgcgccgc	103560
ggactggtgg	tgaacctcgc	ctccgtcatg	gcacaccggg	ccgtccccgg	ccgcgcgcgc	103620
tactgcgcgt	ccaaggccgg	cgtaggtcatg	ctcaccgagg	cgctcggcgt	cgaatgggcg	103680
gagcacggcg	tccgcgtggt	ggccgtcagc	cccgcctacg	tccgcacccc	cctcgtcgcg	103740
gagggtctcg	ccaacggcaa	cctggaocgag	tccgccatcg	ccgcccgcac	cccgtgggc	103800
cggtcgcgcg	agcccagagga	gatcgccgac	ctcgtcctgt	ccctcaccgg	cgacgagttc	103860
gcctatctga	ccgccaccac	cctgcgcttc	gacggcggct	ggaccgccga	cggcgtgttc	103920
ccccgctgaa	ccggcccggg	cctcgtgtaa	ccggcccggg	ccccgctgaa	cccccccacg	103980
gacaaggaga	accgaccccg	atgcccgtca	cggaactcgg	cgctcctcgcg	ctcgcgccacg	104040
ccctcggcga	tccgcgcgac	gtcgcgcgca	ccgcagccga	ccacgtcgac	gaccccgcac	104100
gcgtcctgct	ctggggctac	cgcggtacc	accgcgcccc	gcagggcacc	acctccaccg	104160
cgctcgcgcg	ccgcgcgcgc	gagaaggccc	tcgccaagge	gggctcgac	ccccgcgacc	104220
tcgacctcgt	cgtaggtcgc	gacagcagcg	tccccgagta	cctgctgtgg	gacacctccg	104280
ccgtcgtcgc	ccgcgcgatc	ggcgcgcgca	cggcgcgcac	cctgctgctc	accaggggct	104340
gcgcgtccgc	ggtcaccgcc	ttccagcaga	tcgcccggcat	cttcgcgacc	cgccccgcag	104400
ccgagacggt	cctgctggtc	gccgtcaacc	aggtcagcga	ggcgcacacc	aaccgcgatc	104460
gcttcaaacac	cctgctcggc	agcgacggcg	ccgcggcggc	cgtagtgcgc	cgaggtcacg	104520
accggtccg	ctggtggcc	accgagcagc	tcaccgaccc	cacgtacgcc	gacttctacc	104580
gggtcgagta	cggtggcgcc	gccgtcccgc	ccgccccga	gggcgcggc	aacctggacg	104640
tcgaccgctg	gtccctggtg	taccggcact	tccgcgggga	gcccgcgac	ctggcggaag	104700
tcgtccgcac	cctcaacagc	cgctcgcga	ccgtgttcga	ggacgcgtgc	gacggggccg	104760
gagtcgaccc	cggccaggtg	aagcggttcc	tggtcctcaa	cgacaaccag	gactccctcg	104820
ccgacgtcgc	caaggccgtc	ggcgtgcccc	tgaccgcac	caacgccgaa	ctggccaagg	104880
acctcggtea	ctgcggcggc	gccgaccagc	tgatctgctc	ggacaccctg	ctcgaacgcg	104940
gtgagctcgc	cgagggcgac	gtggtcgcgc	tcgcccggcct	gtccatcggc	atgcaactgt	105000
actgcaccct	gctcaccgtc	tgaccgcgca	ccctcaccga	gccctcaccg	acagcccgcg	105060
acaccgagag	aaacagagag	aagaacgacc	atggctgagc	tgaccatcca	ggacctgatc	105120
gacatcttca	accgcagcat	cgccgaccag	gacccggctc	agcccgaggg	cgacccgtcg	105180
gacgtcacgt	tcgcccgcct	cggtctcgac	tcgctcacca	cgctgaacgc	cgtaacgcgg	105240
atcgagcgca	agcacggcgt	cgaactcggc	gagaacgtga	tcagcgaggg	ccgtacgcgc	105300
cgccagctgc	tcgaccgcgt	caacgcgcgc	ctcgcgcgga	cctgaagcgc	gaaaagagac	105360
acgacatgct	cggaatcacc	ggatgggtgg	cctacggcga	cgacctcgcg	ggcaccgggt	105420
ccgtcgtcca	ggcgtgacc	gacaccatga	tctgccgcgg	cccggacgcc	gagggcgtct	105480
ggatcgaggg	accgctcgcg	ctcggccacc	gccggtgtc	catcatcgac	ccggcgggcg	105540
gcgctcagcc	catggtcgc	acacgcgacg	gccggaccct	ggcgccctg	accttctgcg	105600
gagagatcta	caacttcggg	gagctgocgg	ccgaactggc	cgccctcggg	cacacgttcc	105660
gcaccgacag	cgacaccgag	gtcgtcctca	acgcctacct	ccagtggggc	gccgacttgc	105720
ccgcccggct	caacggcatg	ttcgcctatc	gcctgtggga	cgcgaccacc	cgcgaactgc	105780
tgctcgtccg	cgaccgcgac	ggggtcaagc	cgctctaacta	cgcccgcacc	tccgacggcg	105840
tcgtgttcgg	ctcggaaaccg	aaggccgtcc	tgcccaccgc	cggcgtgccc	cgccgcgtcg	105900
acgcccaggg	cctcgcgcgag	atcctggaca	tggtccgcac	cccggaggtc	acgcggttca	105960
ccgacctgtc	cgaggtccgc	cccggacagg	tggtcggggt	gaccgagggc	ggcctgaccc	106020
ggaccccgta	ctggcagctc	accgctcgcg	agcacaccga	cgacctggac	accacgatcg	106080
ctaccgtccg	ggagctgctg	gaggacatcg	tctcccgccca	gctcatcgcg	gacgtcccgg	106140
tcgccaccct	gctccecggc	ggcctcgaact	cctcggccat	caccgcgctg	gcccagcgca	106200
cgctgaccgc	cgagggacgc	ggcccgggtcc	gctcgtactc	cgtaggactc	cagggcgcg	106260
ccgacgggtt	cgagccggac	gcccgtccgcg	gcaccgccga	cgcgccgtac	gtgcccggact	106320
gcgcccgccca	tgtgggcgccc	gaccacagcg	aggtcctgct	cgacagcacc	gagctggccg	106380
accggcgggt	ccgggcggcg	atcctgggcg	ccaccgacct	cccgcggcg	tactggggcg	106440
acctgtggcc	gtcgtcttac	ctgctgtgcc	gcgagatccg	caagcggggc	accgtagtgc	106500
tgtccggcga	gtccgcgcgac	gagctgttcg	gcccgtaccg	ctggtaccag	cgccccgagg	106560
cggtcgcgac	gccgaccttc	ccctggctca	cccccggtc	ggcgcgcac	ttcggcgcca	106620
gctcgtgat	cgaccagggc	ctgctggaga	agctggacct	ggagggctac	cgccgggacc	106680
gctacgccga	ggccctggcc	gaagtgcocg	tgctggccgg	ggagtccgcg	gtcgaccggc	106740
gcatgcgcga	ggtcacctac	ctcaacctca	cccggttcat	gcagccctg	ctgaccgcga	106800
aggaccgcac	gagcatggcc	gtcggcctgg	aggtccgggt	gccgttctgc	gaccaccgcc	106860
tgatggacta	cgtgttcaac	gtcccgtggg	cgatgaagtc	gttcgacggc	agggagaaga	106920

gcctgctgcg	cgccgccgtc	cgcgacgtgc	tgccgaactc	ggtcctggag	cggaccaaga	106980
ccccgttccc	cgccaccag	gacaccgggt	acgagcaggc	gctgcgcgcc	gagctgcgcg	107040
aggtgctggc	tgaccgggac	gcgccgggtgc	ggccgttget	caacaccgcc	cgggtcaacc	107100
gcgtcctcgg	ccgtgagctg	gacgacttca	gcctgccgca	cgaccggggc	ggcatcgaga	107160
tggcgctgtg	gctcaaccgg	tggctcgect	cctacgaggt	caccgtgacc	gtctgagccc	107220
ctggagcgac	gcccgcacgg	gcccggcgtc	gctccgctcc	gacgcacggg	ccgcgtcaca	107280
tgcgccggccg	aaacacctcc	cggcctcaac	cacgctccga	acaagtccat	ttcgcacaca	107340
gcatggtcca	gcaaggggga	ttacatgcc	cgaccgggtcc	agcgcagagc	accaattgtg	107400
gctgcgcggt	ttccaccagc	cgaagccggg	cgcgcccgcg	ctgatctgcc	tgccgcacgc	107460
gggaggctcg	gcgagcttct	acttcccggg	ctcccaggcg	ctgtcgcccg	cgggtggacgt	107520
gctggcggtg	cagtaccccc	gcccgcagga	ccgcggcgag	gagtccccgg	cagcctccgt	107580
ccaggaactc	cgaggggcg	tgttccgcgc	ggtggacgat	caggaggaca	cgcccctggc	107640
cctgttccggc	cacagcatgg	gcgccatggt	cggcttcgag	ctggcccggc	tccctggaggc	107700
cgcggggcg	ccaccgcgg	tgctcttcat	ctccggccgc	cgaggcccgt	ccatcgtgtg	107760
gacggaaacc	gtccacacga	tggacgacga	acggctgatc	gcgagggtcg	ccaagctcga	107820
aggcaccgac	gcccgcctgc	tccaggacga	ggaggtactg	cggatgatcc	tgcccgcgct	107880
gcccggccgac	taccgtgccc	tggagacctc	ccagcgcacc	cccggacccc	ggctgagttg	107940
ccccttcgctc	gtgatgacgg	gcgacgccga	cccccggtc	accccgcagc	aggcgcggac	108000
ctgggcagag	gagacggacg	gggccttcga	actcgaggtg	taccccggcg	cgcaacttcta	108060
cctgggtggcg	cagcagcagg	ccgtgctggc	ccgatcgag	gccacgatgc	ggcggtcctg	108120
ttccaccacc	ggagcgggtg	tctgagcctg	tgaacgagat	actcagcgcc	gtcctgtcgg	108180
ccgagggcgac	cgccgcggac	ttcgcggccc	tgcgctccc	cgagtctat	cgcgccatca	108240
ccgtgcacaa	ggagatgttc	gagatgttca	ccggtctcgg	cccggccgac	aaggaccgc	108300
gcaaatacact	gcatctggac	gaggtgccgg	tgcgccgaact	gggcccgggc	gaagccctgg	108360
tggcggtgat	ggcctcctcg	gtcaactaca	actccgtgtg	gtcctccatc	ttcgagccgc	108420
tgccgacctt	cgctttcctg	gagcgtacg	gcccggcgcg	cgaactggcg	aagcggcacg	108480
acctgccgta	ccacgtgatc	ggctccgacc	tggctggcgt	ggtgctgccc	accggccccg	108540
gggtgcaggc	atggaagccc	ggtgacgagg	tcgtagcgca	ctgtctgagc	gtggagctgg	108600
agtcgagcga	cggccacgac	gacaccatga	tggaccccga	gcagcgcac	tggggcttcg	108660
agaccaactt	cggcggtctg	gcggagctcg	cgttggtgaa	gtccaaccag	ctgatgccga	108720
aaccccggca	cttgacctgg	gaggaggccg	ccgcgccggg	actcgtcaac	tccaccgcct	108780
accgacaact	ggtgtcgcgc	aacggcgcgc	agatgaagca	ggcgacaac	gtgctgatct	108840
ggggcgcgag	cggcgccctg	ggttcgtacg	ccacacagct	ggcgctggca	ggcggtgcca	108900
accgggtctg	tgtggtgtcc	acgcccagca	aggcggcgat	ctgccgctcc	atggcgctcg	108960
aggcggatcat	gcaccgtgat	gcccaggact	accggtctcg	gcgggacggg	aacaccaggg	109020
accccgggga	atggaagcgt	ttcggcaagc	gcatccgcga	gctgaccggc	ggtgaggacg	109080
tggacatcgt	cttcgagcac	ccgggcccgc	agaccttcgg	cgcgtccgtg	ttcgttgccg	109140
ggaagggcgg	cacgatcgtc	acgtgcccgt	ccacctcggg	ctatgagcat	cagtagcaca	109200
atcggtatct	gtggatgtcg	ctgaagagga	tcatcggttc	ccacttcgcg	aactaccgtg	109260
aggcctggga	ggccaaccgg	ctcgtcgcga	agggcaggat	ccaccccacg	ctctcgaagg	109320
tctaccctt	ggaggagacc	ggacaggccg	ccttcgacgt	gtacggcaac	gtgcaccacg	109380
gcaaggtggg	ggtcctgggc	ctggcgccca	cggaaaggcct	gggtgtgccc	gacgaggaga	109440
tgcgcgccag	gcacgtagga	gccatcaacc	gcttccggcc	gtccgcggag	ccgcagacag	109500
agccggtctg	agggctcgac	ccgcaaaaaga	ccgcacacga	tgcccacttc	tgcaacagag	109560
cagctttgaa	aggcaggacg	tcatgagcaa	cccgttcgac	aaggaagacg	gcagcttctt	109620
cgtggtggtg	aacatggagg	ggcagcactc	cctctggccc	gccttcgccc	agatccctgc	109680
gggctggtcg	atcgtccacg	gcgagaccga	caaggcggcc	tgcctggagt	acgtcgccgc	109740
tcaactggacc	gacatgagcg	ctcgttcgct	gaccgatgcc	gcgccgacc	cagagttagaa	109800
ctccccttcg	gaaacagtc	gttgacaaca	aagtgttcaa	atgccagggt	tgccgcatgt	109860
catcgattcc	gccgcagcca	gccgatgccc	tcaccgcccc	gagcgcgcac	gaagccagag	109920
tgcagcgtgc	gtacacctct	cttttcogca	tgcgccaacg	gcacgcggcc	accgacgggc	109980
agcgtcgtcg	gcaggcccag	tctcacatga	tcagccccta	tgaggcgggtg	aggcttgtgt	110040
cgttcctcct	cagcggggcg	gcgcagctcg	agggagagga	accggagggtg	gatcgcgcgg	110100
acatcacggc	cgcgttgagt	cttctcccgc	tcgctcgtgc	cgagatggac	gaagtggagg	110160
ccggaattct	caagatggcg	catggtcgtg	ggatgacgtg	gtcggagatc	gcgttcggcc	110220
ttggcctcgc	aaccccgcag	gcggcgaggc	agcgcctatga	gcgectggcc	agccgcagtg	110280
aatccggcgc	cgaaaaaaag	gccgacggat	aggagctcgc	catcggcatg	gcggcagccg	110340
accgcccga	tgccgatgcg	tcatatccgt	cagccgagaa	tcctggcgaa	tccgatgtgg	110400
acgcaccggc	ccgacggccg	cgagggcggtg	cccgggctcc	gaacggggcg	tcgggcggcg	110460
ggaagtctgc	ggccttgctc	gacgtgctca	cagccgcgtg	aacccgacgg	ccacgacgcc	110520

gatgccacc	aggaggcaga	ggagccacaa	ggcgacctgg	gagcccgtgc	cgctggcata	110580
gcggcgggcg	gtcgcgcgcc	cgagggcgcc	tatcgcgcat	ccgaagacga	tgaggaccac	110640
acccacggcc	acggcgtcat	cccttctct	tcttgttcca	tatcccgtac	gcgaacatgg	110700
cgacgatccc	caccaggatc	agaatcgcga	cgacggaccc	cgcggtgccg	gccgggatcc	110760
agcttccacg	accgtcagcg	gccagcagca	tcgctgcact	ccttactcag	tacgagcact	110820
gggtgtgggt	tggcccgcga	gcccccaagg	gcaccccggg	ccggccgatt	gtctcaggat	110880
caggattcgg	acgccgcgtc	ctccacctga	ctacccgcga	aatcggcga	gatgccgccc	110940
ccggcaccgg	ccaccgcgcg	gtggagccac	tgtccgcagc	ccctgacccg	ctcagccact	111000
caccagtgcc	tgggcgtgtc	gctcagcgac	ggacttgggc	gccaaccggc	cggtggggga	111060
ggtgggctac	tcggcccgcg	gcgcggggcg	tacgggcgtc	ttgcaccga	tgactcgctc	111120
cgatagccaa	ccccgcccc	agagcagcgg	cccgctccgg	gccgcgtcca	cgtcgtccag	111180
gacggcgacg	ctcactccgg	ctcaccgcc	ccggcggccg	ggagatctga	ctacgtccgg	111240
acacaccacg	ggctccgagg	acgggcagcg	gtgctcctcc	ttcgtcagat	gtacggctcc	111300
gcagcgggtg	cacctctgga	ggtcatcgaa	gccggtcgcg	caggcgaagc	agaagtccac	111360
cggcggttca	gccgcttccg	ggtccgttcc	ggggtgcagc	tggacagcga	gccgctcctcc	111420
aaggccgtac	gggtggccgc	tcgggggtcg	gtgaacacca	ggctccagtg	ctccatgagc	111480
agccgggctt	tgggaaggac	ctcggccccg	gcggccaggt	gcagcaccgc	gtacttcaga	111540
tcccggggcg	tgcccgcgcc	tttctcgcgc	tgcaagaagct	cgacgacgct	gacaaggtag	111600
tcgagaccgt	tggcgtatgg	cgggaaagtgc	gggtctccgg	cgaaccgggt	ttcagcctcc	111660
gaggtctctg	cgaatgccgg	ccagttgggg	gttctgccag	gccgggaggt	agtccggggtg	111720
gtctcgccac	tgetcgectg	caccggctcag	caggtgccaa	gcccgtgtca	cgctgtcggg	111780
gtcgccttgg	gcgagcaggt	acgagacgga	ctcggcgtg	cccgcgaggc	cgtacttgct	111840
ggcattggag	acgcggaagt	accgggctcg	agctcctcc	tccagggtgt	cgtcggcgag	111900
cgcgcccgcg	tgctcctcgt	cgatgcgcgc	gctgatgaag	tcgtacagct	ctcggacatc	111960
cgcagcgggtg	ggcacgttcg	tggatcatggg	gtgcgctcct	cctctgaggg	cggtcaggc	112020
cggcagcctc	atgtaaggga	atltatacga	gggtgatgaa	ggcttggtg	cgtaggttag	112080
accgaagagc	acttgtcggg	gagagctggg	aacagcagge	tgccgccggg	cctccgaatg	112140
ttgacgagaa	gtagcagcgg	tcccatgcc	cggaaacgtg	ctgtcgtcaa	cggtgtgacg	112200
tccgtatcct	ttggcgggtg	agtctgatcg	aggtgggtcc	agggcaggtc	gagctcgtcg	112260
tgcggggggcc	gggggtgctg	gcgacttctg	tccggttggt	cgactgggtcg	cggtcgggatg	112320
agtacgagac	cgtcttcgcg	gtggaggcca	ttgctgatgg	tgtgcgtgca	cggtcogaga	112380
acgtgaccgt	caccgtctgg	gatgacatga	gcgcgttctt	cgacggcctt	gcacgtgatt	112440
tccgtggttg	ggatggggag	cgggtttgga	tcaacaacca	cctggtcgtg	acggcgactt	112500
tcggctcggg	tggatcatgtg	tatgtggact	ggacgcttca	gtcccggctt	tttcccgggtg	112560
actggaagtg	cacggtgacg	accgtgatcg	aagccgtgga	agggatgacg	gctgtagctg	112620
cggacctcgc	ggaatttctg	cgccaggagt	aggtccccac	gtcatccggc	aggcacgccc	112680
gtccggtcca	gccgacggct	tactggctcag	cgcgggcgcg	ggtggtggtg	agacggtagg	112740
agtctgtccg	gtatcgatga	tgttgccgcc	ccggtccgag	cggtcacgcg	gacccaagcc	112800
ggtccccgac	cggtctgct	tgcagggcat	cctgttcgtg	ctccacaccg	ggatcccgtg	112860
gcagcagctt	ccgctggagc	tgggtttcgg	ctccgacggc	atcccgcccg	tcgcccggccg	112920
agtccggccac	ccccgcaagc	gccccgacgc	gctacttggc	gacaagggct	acgacagcaa	112980
cgcggaaccgc	cgagagttgc	gtcggcgcg	gatcctgccg	gtcatctcct	cccgcaagggt	113040
gcagcccagc	atccatggcc	ccgacatcca	tggcctgggc	aggatccgct	acgtggtgga	113100
gcagatcttc	gtcctgctcc	accagttcat	gcgcctcgct	gttcgggtggg	aacgcogttt	113160
ggctctgcac	gacgccttcg	tctcgggtcgg	gtgctcgctc	atctgctgga	ggaggctcaa	113220
gagcccgctc	acctgcag	cccggcgctc	agcccgtgg	cccgtcgggc	attctcacgc	113280
cgccaccggg	cagtgatctg	atgaagtgc	tgccccgcag	accgatgacc	cagccgaagg	113340
catcggagga	aacgggcacg	tccagacgcc	atccggtggc	caagcggact	cgaggctgc	113400
ctgaatcggc	tgcgtgaacc	tcggcgatca	cctgatcgag	cagctgcggc	agccggtcaa	113460
ggttggtccg	tgtggatgcg	tggacggaca	cggttcgtg	gtccgctcgtc	agttcggcg	113520
gactgtgaac	ctcgatccgg	ttctcacccg	tgtgaagccg	caactcgttg	ccgatgatca	113580
gctctccgat	gccgctccca	cgcacgcaga	ggatgtccca	cccgtcgtcc	acacggttgg	113640
ttccggtgtc	catgccgcaa	gactacggac	cccggctacc	acgctgccgc	acaagatcgt	113700
gtgacgaaact	ctaacgagct	ggtgtgtgac	tccggttga	gggtgatcgc	gcgaagggtg	113760
gttgatccag	tccgtgttga	tcatggctgt	cgaccgggat	atcctcgcgc	atcagctgtt	113820
caccggcgtc	tcccggcagc	atctggcttg	ctcggctcag	gagttggccg	aaccgtggca	113880
ggcgatcaca	ccgcccagaa	agccgaacaa	gctcgtttg	ccgcaggtac	acgcagccag	113940
agaagaagcc	cggcaccgac	actcatcgaa	gcgcatcacc	gtcgagcagc	cgctcaccga	114000
ccacaaacgc	tcgaagcagc	tgaccgcgtg	gacgcaccgg	tgagagggcc	ggcccgcac	114060
ctaccgggccc	atcgcggccc	tcgtctccga	ccgcaccgcc	aacgcgtgac	catgaccccc	114120

atcaggcgaa	cacgcctcac	ggtcaagagc	cggccgctcg	ggcaggtgag	gcccgtctcc	114180
tcctgcgtct	ctcgtgtcgc	ggcgggtgcg	ggggactcgt	ggtggtcgag	gttcccgcgc	114240
gggaactgcc	agatctgcct	gccagtgcgg	cggccgcgca	cagcgcgtcg	ggcacggtga	114300
gcgtgtacct	ggcctggccg	gcgggtccgga	tcagaacgcc	gctgggatcg	ctgccgccgc	114360
tgtgagcgtg	cggacggacg	tcatcggggg	agtgaccgga	gcgaccgtct	ggctcttgct	114420
gctggtgagg	catttctgga	aggaatccgg	ttgcctcgtc	gttctcggca	gtgccaggtg	114480
gccccgccgt	cgctgacgac	cgcggttca	cgccccgtcc	tgcgggtgga	gtcggcgggc	114540
acctgcaaag	ccgtcgcgca	gtcggtgccc	tgcgcggccg	acaccgccag	gcccgtctcg	114600
ccgcggtccg	tgccgcggtg	cccggcaact	cgcacgcgga	cacggctgca	gggcccggaga	114660
cgaccgtcgc	cgatacggcc	atcaggcctc	ccagccacta	cccaacttcc	ccgagtaacta	114720
ccctgagccga	ggtcgcactc	gcgtacgcac	aggcacgctg	tccgcccctc	agtggacggc	114780
cctcaacgcc	ctgctcgcgc	cagccctgtg	ctcgtcgcgc	gactcggggt	acgacaagat	114840
ccaacagcac	ctcgacgcgg	acgatcacct	tgcgccgcac	ggcgggtggc	acgcgtacgg	114900
gcgcggcaac	ttcttcgctc	ccttcctcag	cacaccccgc	gacaccggga	tctgggagct	114960
gcagttcggc	ggccatcacc	tgcgggtcgc	ccacacctac	gccgacggac	gcctgatcgg	115020
tgccacaccc	gcgttcgcgc	gcatcgagcc	gttccggctc	ttcaccgcgc	acggcatccc	115080
gaccagcccg	aacgagccga	acagcaggcc	ctctccgccc	tgctccagtc	cctcgacgcc	115140
tcgcaggcgg	cggccggact	cagggagcgg	cacgacgacc	ttctgctcgg	cccgcaccgg	115200
gactggaagt	tcccccgaa	gtccgaaggc	atcgccgccc	acgatctgac	agccgcgcaa	115260
cgtgacctgc	tcgtcgcggc	catcgcccgt	taegtgcggc	acttgcccga	agcggacgcc	115320
cgcaggtttc	tcgaccgcta	ccgggcccga	cttcccgcga	cccacgtcgg	cttcaccgga	115380
tcgacactgc	tgaccgacc	cggtgactac	gtccgcatcg	acggctcgtc	ggtgtggatc	115440
gagttctcca	tgcacgacgg	catcgtcctg	gcccgatccg	acccccacgc	cgtctggcgc	115500
gacaagcaca	ccgactacgg	cggactcaag	ccctgacgcg	gtcccgcgcg	cccgtatccg	115560
gggcccgacc	acgccccttc	cccagctcac	accgcgcggc	ggtcgcgcgc	ggggctcacc	115620
gctggacggc	gcgcagtccg	tccaggagca	gccgcagtcc	gaactcgaag	tgtgccgcga	115680
agtccgtgct	ggtgagcact	ggcaggggtg	cggcgaggtg	tggatatgtc	cccgcgcgca	115740
cggcctcgtg	cagggctccc	gcatcggtgg	gctccccgtc	ccccggcccg	agggcggcct	115800
gttcctccag	ggtgtggccg	acgggtgaagt	tggtcaccga	gtagagcgcc	cgcgcgcgct	115860
cgccgtcacc	gaaacccgcc	gcacgcagca	caccgacgag	gctgtcggcg	aagcccaggg	115920
tgttcggacc	ggtcgagtgc	gtaccgggtg	agacgcgtgc	gccgtcgcga	tgggcgagca	115980
gggcccgtacg	cagcgcgcgg	gcgagccgcg	ccgctctctc	gctccagtcg	ccgtccccgc	116040
ccggcccggc	ggcgtcggcg	acgcgcgcga	ccatccgctc	cgccatcgcg	gtgagcaggt	116100
cctgtttggt	cgcgaagtag	cggtacagcg	caccggcctg	gacgcccgat	gcgtcggcga	116160
gcccgcgcat	ggtcaacgcg	tcgagaccgg	actcgttcag	caggtcgagg	cggtctgga	116220
cgggtcggac	ctggctgata	cgggcggggc	ggcctctggt	cgggcttgac	gggggtctca	116280
tggacctcac	tatactgaac	accgttcacg	tgaacgctgt	tcactaaatg	tgggaggaca	116340
tctccgatga	acaccgatgt	gatcgtcata	ggcgccggcc	cgaccgggct	gatgctcgcc	116400
tgcgagctgg	cactggcggg	cgtacggacc	cgcgtcctgg	agcgcgcgcg	cgaaccgcag	116460
cgcaattcgc	gggcggtgag	cctgcacccg	cgcagcgtgg	aactgatgga	ccagcggagg	116520
ctgctcgacc	ggttcctccc	gctcggccgg	accgtgccgg	gctggcactt	cgcagggctg	116580
cgcacccggc	tggacttcgg	cgccctcgac	tcccggcacg	gctacaccct	cttcctggct	116640
caggagcgca	ccgagggcat	cctggaggag	cgggcacacg	agctgggggc	ggagatcagc	116700
cggggatacgc	agaccctcgg	tgtgagccag	gacggcggcg	gggtcgaggt	gcaggtgcgc	116760
gtccccggcg	gcggcaccga	gaccgtccgc	ggcctctacg	cggtcgggtg	cgcagggagg	116820
cgcagcgtcg	tgcggcaggc	cgcagggatc	gacttccccg	gcaccgacga	gaccctgacc	116880
gggggtgctc	gggacttcgc	ggtcgtcgat	ccgcggcccg	gcgccctcga	tgcggcacga	116940
gccccgcggc	tgatcatcgc	gcccgtggag	ggagggctca	cccggttcgt	ctacctggat	117000
cccgagcgga	tacgggtgcc	gtcccgggag	cccgtgaccc	tggaggagtt	ccgtacgtcg	117060
ctcaccggga	tcaccggctc	cgactgcggg	atcgccgagc	cccgttggtc	gtcgcggttc	117120
ggcaatgccca	ctcggctcgc	ggagagttac	cgtgcgggac	gcgtcctgct	ggcgggogac	117180
gccgcccaca	tccacttccc	cgcggcgggc	caggggctca	acaccgggct	ccaggacgcc	117240
atgaacctgg	ggtggaagct	ggtcgcgcgc	gtgggcggct	gggcgcggcc	cgggctgctg	117300
gacagctacg	acggggaacg	caggccggtc	ggccggctcg	tgaccgagaa	caccgaggtg	117360
cagactctgc	tggcggagct	gacgctggtc	gcgcagtacc	agcgtccggc	cgcggccctg	117420
cgcgagctcc	tcgaccaact	gctgggcatg	acggaggtca	accgccggct	cgcgcgaccg	117480
gtctccgcgc	tcggcaccca	ctatccgcgc	gcggactcgg	acgctgatcc	gctggtgggg	117540
cggcgtatgc	ccgacatcgg	cctgaccgcc	gcccgttcgg	cggcgacacg	ggtgtacgaa	117600
ctgctcacc	ggggcgcctt	cgtcctgctc	gcccgtgcgc	gagaccggc	ctctcgcgag	117660
gcccgtcgacg	cgggctgggg	ctcgcgggct	tccggccgctca	ccgtggacaa	gtgcgacgag	117720

```

cattcggacc tcgacggggt ggccgaggtt ctcgtacgcc cgcacggcca cgtcgcgtgg 117780
gccacccgta ccgccgacgc cgacccgcgc ggcgacgagc ggggccagge gctgaccgct 117840
tgggcgggca cgccctcacc gaggagctga tcatgcccggt gaaccgcgtc agtccaccgt 117900
ccgtgcacaa cccgtccgac cagatcagcc aggtcgtcac cgtcgcacgga acccggctcg 117960
cgcattcttc cgggcaggtc gcctgggatg cggaaggatg gcccggtggg ctcggcgacc 118020
acgaggccca ggcggcacag atcgcaacga acctcgacgc cgcaactcgc gcggtcggcg 118080
ccacccgcga cgacatcatc gaagagaccg tctacgtagt cgactacacc ccgaaactgt 118140
tatccgcgat cttcggggcg ctacggggcg gtgtctcagc agcacctgcg agcacgtggt 118200
tgggcgtgcc cgcgctcttc gcaccggaat acctcctgga ggtccaggtg gtcgccgcgc 118260
tccccgtacc agggggccat gaccgcgcca caccgaccga cgacaccgcg gcgaacgagc 118320
gtgtcagggc acgggagccc cgcaggatcc gtgaaacacc ccccacgggc ga 118372

```

<210> 2

<211> 146

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 2

```

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

```

<210> 3

<211> 155

<212> PRT

15 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 3

ES 2 556 211 T3

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

<210> 4

5 <211> 276

<212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 4

10

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

<210> 5

<211> 446

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 5

Met Thr Arg Leu Ala Glu Gln Ser Ser Thr Ala Gln Gln Ser Pro Glu
 1 5 10 15
 Ser Glu Val Leu Asp Val Thr Gly Ile Gly Phe Gly Ala Ala Asn Leu
 20 25 30

10

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

ES 2 556 211 T3

<211> 231
 <212> PRT
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

5 <400> 6

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

10 <210> 7
 <211> 3609
 <212> PRT
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

15 <400> 7

ES 2 556 211 T3

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

ES 2 556 211 T3

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

ES 2 556 211 T3

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

ES 2 556 211 T3

Ala Val Ile Gly Ala Ala Ala Pro Val Ala Pro Arg Ile Val Asp Val
1155 1160 1165
Pro Asp Ala Asp Ala Ala Arg Ala Ala Ala Ser Ala Glu Leu Ser Thr
1170 1175 1180
Gly Phe Asp Leu Thr Arg Gly Pro Leu Leu Arg Ala Thr Leu Leu Arg
1185 1190 1195 1200
Leu Ala Pro Asp Asp His Val Leu Val Val Ala Val His His Ile Ala
1205 1210 1215
Thr Asp Gly Trp Ser Gln Ala Leu Leu Trp Ala Glu Ile Ala Ala Ala
1220 1225 1230
Tyr Asp Gly Ala Pro Leu Pro Glu Leu Pro Val Gln Tyr Gly Asp His
1235 1240 1245
Ala Val Trp Gln Arg Ser Trp Leu Thr Gly Glu Val Leu Glu Arg Arg
1250 1255 1260
Ala Gly Tyr Trp Thr Gly Arg Leu Ala Gly Leu Ala Pro Leu Glu Leu
1265 1270 1275 1280
Pro Leu Asp Lys Ala Arg Pro Ala Val Ala Thr Gly Arg Ala Gly Thr
1285 1290 1295
Leu Pro Trp Gln Leu Pro Ala Glu Leu Ile Arg Asp Ala Arg Ala Val
1300 1305 1310
Ala Ala Arg Glu Gly Ala Thr Leu Tyr Met Val Leu Leu Ala Ala Phe
1315 1320 1325
Thr Leu Val Leu Ser Arg Tyr Ala Arg Thr Glu Asp Ile Ala Val Gly
1330 1335 1340
Ser Pro Thr Ala Gly Arg Thr Arg Ala Glu Thr Glu Ala Leu Ile Gly
1345 1350 1355 1360
Phe Phe Val Asn Val Val Ala Val Arg Thr Asp Leu Ser Gly Asp Pro
1365 1370 1375
Thr Phe Arg Glu Leu Leu Gly Arg Val Arg Glu Ser Val Val Gly Ala
1380 1385 1390
Val Glu His Gln Asp Val Pro Phe Glu His Leu Val Glu Arg Leu Arg
1395 1400 1405
Pro Glu Arg Asp Leu Ser Arg Asn Pro Leu Val Gln Val Ala Phe Gln
1410 1415 1420
Leu Leu Ala Asp Ala Pro Arg Arg Pro Trp Trp Gln Gly Ala Arg Ala
1425 1430 1435 1440
Glu Pro Phe Asp Ile Asp His Ala Tyr Thr Arg Met Asp Leu Glu Val
1445 1450 1455
His Ala Val Glu Thr Gly Asp Glu Val Gly Ala Thr Val Leu Tyr Ala
1460 1465 1470
Ala Asp Leu Phe Asp Ala Asp Thr Val Arg Gln Leu Met His His Val
1475 1480 1485
Ser Val Val Leu Gly Glu Val Leu Ala Asp Pro Asp Arg Pro Val Ser
1490 1495 1500
Ala Ala Thr Met Leu Asp Glu Thr Asp Arg His Arg Thr Leu Val Ala
1505 1510 1515 1520
Trp Asn Asp Thr Ala Ala Pro Leu Pro Asp Gly Cys Val Pro Arg Leu
1525 1530 1535
Tyr Ala Glu Gln Ala Ala Arg Thr Pro Asp Ala Val Ala Leu Ile Cys
1540 1545 1550
Gly Asp Glu Arg Val Thr Tyr Ala Glu Leu Asp Arg Arg Ala Asn Arg
1555 1560 1565
Phe Ala His Leu Leu Leu Ala His Gly Val Gly Ala Asp Glu Pro Val
1570 1575 1580
Gly Val Ala Thr Gly Arg Ser Thr Gly Met Val Ala Ala Val Leu Gly
1585 1590 1595 1600
Val Leu Lys Ala Gly Ala Ala Tyr Val Pro Leu Asp Pro Arg Asn Pro
1605 1610 1615
Pro Gly Arg Thr Glu Arg Ile Val Ala Thr Ser Gly Leu Arg Val Val
1620 1625 1630

ES 2 556 211 T3

Ile Ala Asp Arg Pro Val Pro Gly Thr Asp Gly Ile Thr Val Leu Asp
1635 1640 1645
Val Thr Asp Pro Gly Pro Gly Pro Asp Thr Asp Pro Gly Ile Asp Pro
1650 1655 1660
His Pro Asp Thr Thr Ala Tyr Val Ile Tyr Thr Ser Gly Ser Ser Gly
1665 1670 1675 1680
Glu Pro Lys Gly Val Ala Val Thr His Arg Asn Val Val Val Leu Ala
1685 1690 1695
Ala Asp Arg Arg Trp Ser Asn Gly Asn His Glu Arg Val Leu Leu His
1700 1705 1710
Tyr Pro Leu Ala Thr Asp Ile Ser Thr Tyr Glu Leu Trp Pro Phe Leu
1715 1720 1725
Leu Thr Gly Lys Gln Ile Val Val Ala Thr Asp Glu His Val Glu Pro
1730 1735 1740
His Thr Phe Asp Arg Leu Ile Arg Glu His Gly Val Thr Ala Met Cys
1745 1750 1755 1760
Leu Pro Ala Pro Leu Phe Ser Leu Leu Ala Glu Glu Cys Met Glu Cys
1765 1770 1775
Phe Gly Gly Leu Arg Glu Val Leu Thr Gly Gly Glu Ala Val Ser Gly
1780 1785 1790
Glu Thr Val Ala Gln Val Met Ala Ala His Pro His Leu Thr Val Ala
1795 1800 1805
Asp Ala Tyr Gly Pro Thr Glu Ala Thr Ala Phe Thr Thr Leu Phe Pro
1810 1815 1820
Met Glu Pro Gly Phe Arg Leu Ala Gly Ser Arg Val Pro Ile Gly Ala
1825 1830 1835 1840
Pro Ile Asp Asn Thr Arg Val Tyr Val Leu Asp Asp Thr Leu Arg Pro
1845 1850 1855
Ala Pro Leu Gly Val Ala Gly Glu Leu Val Ile Gly Gly Pro Arg Val
1860 1865 1870
Ala Arg Gly Tyr Leu Gly Arg Pro Asp Leu Thr Ala Glu Lys Phe Val
1875 1880 1885
Pro Asp Pro Trp Gly Gly Pro Gly Glu Arg Met Tyr Arg Thr Gly Asp
1890 1895 1900
Val Val Arg Trp Leu Pro Gly Gly Ala Leu Glu Phe Leu Gly Arg Ala
1905 1910 1915 1920
Asp Val Gln Val Lys Ile Arg Gly Phe Arg Val Glu Pro Gly Glu Val
1925 1930 1935
Glu Ala Ala Leu Leu Arg His Pro Ala Val Ser Gln Val Thr Val Ala
1940 1945 1950
Val Arg Glu Asp Ile Pro Gly Asp Lys Arg Leu Val Ala Tyr Val Val
1955 1960 1965
Leu Glu Pro Glu Ala Gly Ala Ser Val Leu Pro Ala Leu Arg Ala His
1970 1975 1980
Ala Ala Gly Ser Val Pro Asp Tyr Met Val Pro Ser Ala Phe Val Ala
1985 1990 1995 2000
Leu Asp Ala Phe Pro Leu Ser Thr Thr Gly Lys Ile Asp Arg Arg Ala
2005 2010 2015
Leu Pro Ala Pro Asp Thr Arg Ser Val Ala Glu Ser Gly Tyr Val Pro
2020 2025 2030
Pro Ala Thr Glu Ala Glu Arg Val Leu Cys Glu Ile Phe Ala Glu Val
2035 2040 2045
Leu Asp Val His Pro Val Gly Ala Asp Asp Asp Phe Phe Ala Leu Gly
2050 2055 2060
Gly His Ser Leu Leu Ala Thr Arg Ala Ile Ala Arg Ile Arg Ala Arg
2065 2070 2075 2080
Phe Gly Pro Asp Val Pro Leu Gln Ala Val Phe Glu Arg Arg Ser Pro
2085 2090 2095
Arg Arg Leu Ala Glu Thr Leu Gly Glu Pro Gly Thr Ala Thr Asp Val
2100 2105 2110

ES 2 556 211 T3

Ile Arg Pro Ala Arg Arg Asp Gly Ala Ala Leu Pro Leu Ser Ser Ser
 2115 2120 2125
 Gln Arg Arg Leu Trp Phe Leu Asp Arg Leu Thr Pro Asp Ser Gly Phe
 2130 2135 2140
 Trp Asn Val Ala Met Gly Val Arg Ala His Gly Pro Leu Asp Val Asp
 2145 2150 2155 2160
 Ala Leu Gly Arg Ala Leu Ser Leu Val Val Ser Arg His Glu Ala Leu
 2165 2170 2175
 Arg Thr Val Phe Ala Ala Asp Ala Gly Glu Pro Met Ala Val Val Arg
 2180 2185 2190
 Pro Ala Thr Pro Leu Arg Leu Glu Val Thr Asp Val Ala Asp Glu Ala
 2195 2200 2205
 Glu Val Arg Ala Leu Ala Glu Ala Asp Ala Ala Arg Pro Phe Asp Leu
 2210 2215 2220
 Ala Arg Gly Pro Leu Leu Arg Ala Arg Val Leu Arg Leu Ala Ala Glu
 2225 2230 2235 2240
 Asp His Ala Val Leu Ile Thr Ala His His Ala Val Thr Asp Gly Trp
 2245 2250 2255
 Ser His Ala Val Phe Trp Gly Glu Leu Ala Glu Ala Tyr Arg Ala Glu
 2260 2265 2270
 Leu Ser Gly Asp Pro Ala Glu Leu Pro Glu Leu Pro Val Gln Tyr Gly
 2275 2280 2285
 Asp Phe Ala Val Trp Gln Gln Gly Arg Leu Thr Gly Ala Glu Leu Glu
 2290 2295 2300
 Arg Tyr Leu Thr Tyr Trp Arg Ala Arg Leu Ala Gly Leu Arg Pro Leu
 2305 2310 2315 2320
 Glu Leu Pro Leu Asp Arg Pro Arg Pro Ala Val Ala Gly Ser Ala Gly
 2325 2330 2335
 Ala Ser Gln Pro Trp Glu Leu Pro Glu Asp Leu Val Arg Ala Ala Arg
 2340 2345 2350
 Ala Phe Gly Asp Thr Glu Gly Ala Thr Leu Tyr Met Thr Leu Leu Thr
 2355 2360 2365
 Ala Phe Thr Val Val Leu Ala Arg Phe Ala Gly Thr Glu Asp Val Ala
 2370 2375 2380
 Val Gly Ala Pro Val Ala Gly Arg Thr Arg Pro Glu Val Glu Arg Leu
 2385 2390 2395 2400
 Ile Gly Phe Phe Val Asn Met Leu Val Leu Arg Thr Asp Val Ser Gly
 2405 2410 2415
 Asp Pro Thr Phe Arg Asp Leu Leu Gly Arg Val Arg Glu Thr Val Val
 2420 2425 2430
 Gly Ala Met Asp His Gln Asp Leu Pro Phe Glu His Leu Val Glu Thr
 2435 2440 2445
 Leu Ala Pro Glu Arg Asp Leu Ser Arg Asn Pro Leu Val Gln Val Val
 2450 2455 2460
 Phe Gln Leu Met Arg Ala Pro Gly Asp Lys Gly Asp Arg Leu Gly Ala
 2465 2470 2475 2480
 Ala Arg Leu Glu Pro Leu Leu Asp Glu His Ala Phe Thr Arg Val Asp
 2485 2490 2495
 Leu Glu Val His Leu Thr Glu Asp Gly Asp Arg Val Arg Gly Thr Val
 2500 2505 2510
 Leu His Ser Thr Ala Leu Phe Glu Ala Asp Thr Val Arg Arg Leu Leu
 2515 2520 2525
 His His His Thr Val Leu Leu Arg Ala Ala Leu Ala Asp Pro Asp Ala
 2530 2535 2540
 Pro Leu Ser Ala Leu Ser Leu Leu Asp Asp Asp Asp Arg Arg Leu Leu
 2545 2550 2555 2560
 Leu Glu Arg Trp Asn Asp Thr Ala Leu Pro Tyr Arg Asp Val Pro Leu
 2565 2570 2575
 Val Glu Leu Phe Ala Glu Gln Val Ala Arg Thr Pro Gly Ala Arg Ala
 2580 2585 2590

ES 2 556 211 T3

Val Glu Cys Glu Asp Asp Val Leu Thr Tyr Ala Ala Leu Asp His Glu
 2595 2600 2605
 Ala Glu Arg Ile Ala Ala Gly Leu Arg Ala Gln Gly Val Gly Pro Asp
 2610 2615 2620
 Asp Leu Val Gly Leu Cys Leu Glu Arg Gly Thr Val Gln Met Ala Ala
 2625 2630 2635 2640
 Leu Ile Gly Ile Leu Lys Ala Gly Ala Ala Tyr Val Pro Ile Asp Pro
 2645 2650 2655
 Ser His Pro Arg Asp Arg Ile Arg Leu Ile Val Asp Asp Ala Arg Met
 2660 2665 2670
 Thr Val Ala Val Thr Asp Arg Ala His Ala Asp Val Phe Ala Glu Gly
 2675 2680 2685
 Thr Ala Leu Val Leu Val Asp Ala Pro Ala Gly Gln Glu Pro Pro Ala
 2690 2695 2700
 Ala Thr Ala Ala Gly Arg Pro Gly Pro Asp Ser Leu Ala Tyr Val Val
 2705 2710 2715 2720
 Tyr Thr Ser Gly Ser Thr Gly Val Pro Lys Gly Ile Ala Met Pro Ala
 2725 2730 2735
 Arg Cys Val Val Asn Met Leu Ala Trp Gln Lys Lys Thr Val Pro Gly
 2740 2745 2750
 Gly Pro Gly Thr Arg Thr Ala Gln Phe Thr Ala Leu Thr Phe Asp Val
 2755 2760 2765
 His Val Gln Glu Val Leu Ser Ala Leu Leu Tyr Gly Glu Thr Leu Val
 2770 2775 2780
 Ile Pro Thr Glu Glu Thr Arg Arg Asp Pro Ala Arg Phe Ala Arg Trp
 2785 2790 2795 2800
 Leu Asp Glu Arg Ala Val Glu Gln Ile Phe Val Pro Asn Leu Met Ile
 2805 2810 2815
 Arg Ala Leu Ala Glu Glu Ala Gly Ala Gly Arg Ala Arg Leu Thr Ser
 2820 2825 2830
 Leu Arg His Ile Ser Gln Ala Gly Glu Pro Leu Ser Leu Asp Thr Val
 2835 2840 2845
 Leu Arg Glu Phe Cys Ala Ala Arg Pro Arg Leu Arg Leu His Asn His
 2850 2855 2860
 Tyr Gly Ser Thr Glu Ile Gln Val Val Thr Ser Phe Thr Leu Pro Ala
 2865 2870 2875 2880
 Asp Val Ala Asp Trp Pro Arg Thr Ala His Leu Gly Glu Pro Val Asp
 2885 2890 2895
 Asn Ser Arg Ala Tyr Val Leu Asp Asp Arg Leu Arg Pro Val Pro Val
 2900 2905 2910
 Gly Val Ala Gly Glu Leu Cys Phe Ala Gly Pro Gly Leu Ala Arg Gly
 2915 2920 2925
 Tyr Val Gly Lys Pro Glu Leu Thr Ala Gln Lys Phe Val Pro Asp Pro
 2930 2935 2940
 Phe Gly Pro Pro Gly Ser Arg Leu Tyr Arg Thr Gly Asp Leu Gly Arg
 2945 2950 2955 2960
 Trp Arg Pro Asp Gly Ser Leu Glu Tyr Leu Gly Arg Leu Asp His Gln
 2965 2970 2975
 Val Lys Ile Arg Gly Phe Arg Val Glu Leu Gly Glu Val Glu Ala Val
 2980 2985 2990
 Leu Leu Arg His Pro Glu Val Thr Arg Ala Val Ile Val Ala Arg Glu
 2995 3000 3005
 Asp Ala Pro Gly Val Lys Arg Leu Val Gly Tyr Val Val Pro Val Pro
 3010 3015 3020
 Gly Thr Asp Gly Gly Leu Pro Ala Arg Leu Arg Ala His Leu Ala Asp
 3025 3030 3035 3040
 Ala Val Pro Asp Tyr Met Val Pro Ser Ala Leu Val Ala Leu Asp Ala
 3045 3050 3055
 Phe Pro Leu Thr Thr Gly Lys Ile Asp Arg Ala Ala Leu Pro Ala
 3060 3065 3070

ES 2 556 211 T3

Pro Asp Leu Arg Thr Thr Leu Asp Thr Gly Phe Thr Ala Pro Arg Thr
 3075 3080 3085
 Gly Ala Glu Gln Thr Leu Cys Glu Val Phe Ala Glu Val Leu Glu Thr
 3090 3095 3100
 Gly Gly Val Gly Ile Asp Asp Asp Phe Phe Ala Leu Gly Gly His Ser
 3105 3110 3115 3120
 Leu Val Ala Ala Arg Thr Val Ala Arg Ile Arg Glu Ala Leu Gly Ala
 3125 3130 3135
 Glu Val Ser Leu Arg Glu Leu Phe Gln His Arg Thr Pro Arg Ala Leu
 3140 3145 3150
 Ala Glu Val Val Ala Val Ala Pro Arg Thr Ala Val Pro Pro Leu Val
 3155 3160 3165
 Pro Val Ser Arg Gln Glu Pro Leu Pro Leu Ser Leu Gly Gln Leu Arg
 3170 3175 3180
 Leu Trp Arg Leu His Glu Ala Asp Pro Gly Asp Pro Val Trp Thr Ile
 3185 3190 3195 3200
 Pro Leu Ala Val Arg Ile Thr Gly Glu Leu Asp Ala Asp Leu Leu Gly
 3205 3210 3215
 Arg Ala Leu Thr Glu Val Val Arg Arg His Glu Ala Leu Arg Thr Val
 3220 3225 3230
 Phe Val Pro Gly Asp Glu Pro Ala Ser Val Ile Leu Pro Ala Thr Asp
 3235 3240 3245
 Ile Val Leu Asp Pro Val Asp Val Ala Asp Glu Thr Ala Ala Arg Ala
 3250 3255 3260
 Leu Ala Asp Glu Ala Ala Ala Arg Pro Phe Asp Leu Val Arg Gly Pro
 3265 3270 3275 3280
 Val Leu Arg Pro Ala Leu Leu Arg Ile Ala Pro Asp Asp His Val Leu
 3285 3290 3295
 Leu Leu Thr Val His His Ile Ala Thr Asp Gly Trp Ser Gln Gly Val
 3300 3305 3310
 Leu Trp Thr Glu Leu Ser Gly Ala Tyr Ala Ala Leu Arg Glu Asn Arg
 3315 3320 3325
 Pro Ala Glu Leu Pro Glu Leu Pro Val Gln Tyr Gly Asp Phe Ala Phe
 3330 3335 3340
 Trp Gln Arg Ser Trp Leu Thr Gly Ala Ala Leu Asp Ala Gln Leu Gly
 3345 3350 3355 3360
 His Trp Arg Arg Arg Leu Asp Gly Leu Arg Pro Leu Ala Leu Pro Gly
 3365 3370 3375
 Val Pro Ala Asp Ala Ala His Asp Ala Thr Gly Val Leu Thr Glu Trp
 3380 3385 3390
 Arg Leu Pro Ala Gly Leu Val Ala Thr Ala Arg Arg Val Gly Ala Glu
 3395 3400 3405
 His Asp Ala Thr Leu Tyr Met Thr Leu Leu Ala Ala Phe Thr Ala Thr
 3410 3415 3420
 Leu Ala Arg Trp Ala Gly Thr Asp Asp Leu Ala Val Gly Ser Pro Val
 3425 3430 3435 3440
 Ala Gly Arg Thr Arg Ala Glu Val Glu Gly Leu Ile Gly Phe Phe Ala
 3445 3450 3455
 Asn Phe Val Pro Leu Arg Val Asp Leu Ser Gly Asp Pro Ser Phe Ala
 3460 3465 3470
 Gly Leu Leu Glu Arg Val Arg Asp Thr Ala Leu Asp Ala Tyr Ala His
 3475 3480 3485
 Gln Ala Leu Pro Trp Glu Arg Val Val Glu Gly Leu Gly Leu Asp Pro
 3490 3495 3500
 Glu Gln Pro Leu Val Asp Val Val Phe Gln Leu Val Asn Val Glu Leu
 3505 3510 3515 3520
 Gly Glu Leu Gly Leu Pro Gly Ala Arg Val Glu Gln Phe Thr Gly Gln
 3525 3530 3535
 Gln Ala Tyr Ala Arg Trp His Leu Glu Val His Leu Val Glu Asp Pro
 3540 3545 3550

ES 2 556 211 T3

Asp Gly Gly Leu Thr Gly His Val Val His Arg Ala Ala Ala Leu Gly
3555 3560 3565
Arg Arg Val Val Asp Gly Leu Leu Ala Gly Thr Ala Ala Leu Leu Gln
3570 3575 3580
Ala Ala Leu Ala Glu Pro Gly Leu Pro Val Ser Ala Leu His Ala Pro
3585 3590 3595 3600
Asp Pro Leu Glu Gly Thr Thr His Asp
3605

<210> 8

<211> 2217

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 8

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

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

ES 2 556 211 T3

	1315		1320		1325
Ser Thr Ala Leu Ala Asp Ile Asp Ala Gly Ala Val Asp Leu Val Leu					
1330			1335		1340
Ala Pro Val Ser Thr Gly Ala Asp Val Val Ala Ala Ala His Arg Ala					
1345			1350		1355
Thr Ala Gln Val Leu Glu Leu Leu His Glu Trp Leu Ala Asp Asp Arg					
			1365		1370
Phe Gly Gln Ala Arg Leu Ala Ile Val Thr Arg His Ala Val Ala Ala					
			1380		1385
Arg Pro Gly Glu Glu Pro Asp Pro Ala Ala Ala Ala Val Trp Gly Leu					
			1395		1400
Val Arg Ser Ala Gln Thr Glu His Pro Asp Arg Phe Leu Leu Val Asp					
			1410		1415
Thr Asp Gly Thr Pro Ala Ser Leu Asp Ala Val Pro Ala Val Gly Asp					
			1425		1430
Glu Pro Gln Thr Ala Leu Arg Asp Gly Glu Arg Leu Val Ala Arg Leu					
			1445		1450
Thr Arg Ala Ala Glu Thr Ala Leu Arg Pro Pro Val Gly Ala Asp Ala					
			1460		1465
Trp Arg Val Asp Val Val Arg Pro Gly Ser Ile Asp Gly Val Asp Ala					
			1475		1480
Val Ala Ala Pro Asp Ala Thr Ala Pro Leu Ala Pro Gly Gln Val Arg					
			1490		1495
Ile Ala Val Arg Ala Ala Gly Leu Asn Phe Arg Asp Val Leu Cys Ala					
			1505		1510
Leu Asp Met Tyr Pro Asp Glu Val Asp Ala Ile Gly Ser Glu Ala Ala					
			1525		1530
Gly Thr Val Val Ala Val Ala Pro Asp Val Thr Asp Leu Ala Val Gly					
			1540		1545
Asp Arg Val Leu Gly Met Val Pro Gly Gly Phe Gly Thr Leu Ala Val					
			1555		1560
Val Asp Arg Arg Leu Val Val Pro Val Pro Ala Gly Trp Ser Trp Val					
			1570		1575
Arg Ala Ala Ala Leu Pro Ser Val Phe Ala Thr Ala Trp Phe Ala Leu					
			1585		1590
Arg Asp Val Ala Gly Val Arg Ala Gly Glu Arg Val Leu Val His Ala					
			1605		1610
Ala Ala Gly Gly Val Gly Met Ala Ala Val Arg Val Ala Arg Leu Leu					
			1620		1625
Gly Ala Glu Val Tyr Ala Thr Ala Ser Pro Gly Lys His Glu Val Leu					
			1635		1640
Arg Ala Ala Gly Leu Asp Glu Ala Arg Val Ala Ser Ser Arg Asp Thr					
			1650		1655
Glu Phe Ala Gln Arg Phe Pro Glu Met Asp Val Val Leu Asn Ser Leu					
			1665		1670
Thr Gly Glu Phe Val Asp Ala Ser Leu Arg Leu Leu Arg Pro Gly Gly					
			1685		1690
Arg Phe Val Glu Leu Gly Lys Thr Asp Arg Arg Asp Pro Ala Gly Leu					
			1700		1705
Pro Gly Val Asp Tyr Leu Pro Phe Asp Leu Leu Leu Asp Ala Gly Pro					
			1715		1720
Asp Arg Val Gln Ser Leu Leu Thr Glu Val Val Ala His Ala Glu Ala					
			1730		1735
Gly Glu Leu Thr Gly Leu Pro Thr Arg Thr Trp Pro Leu Ala Asp Ala					
			1745		1750
Arg Thr Ala Phe Arg Phe Met Ala Gln Ala Arg His Thr Gly Lys Ile					
			1765		1770
Val Leu Thr Val Ala Pro Tyr Ala Asp Gly Thr Val Leu Ile Thr Gly					
			1780		1785
Ala Gly Val Leu Gly Gly Met Leu Ala Arg His Leu Val Ser Glu His					

	1795		1800		1805														
Gly	Ala	Arg	Asp	Leu	Val	Leu	Ala	Ser	Arg	Arg	Gly	Ala	Ala	Ala	Pro				
	1810						1815				1820								
Gly	Ser	Ala	Asp	Leu	Val	Ala	Glu	Leu	Ala	Ala	Ala	Gly	Ala	Thr	Val				
1825					1830					1835					1840				
Arg	Phe	Glu	Thr	Cys	Asp	Val	Thr	Asp	Arg	Ala	Ala	Leu	Asp	Ala	Leu				
				1845					1850					1855					
Leu	Ala	Lys	Leu	Thr	Ala	Glu	Ala	Pro	Leu	Thr	Ala	Val	Val	His	Thr				
			1860					1865					1870						
Ala	Gly	Ala	Leu	Asp	Asp	Gly	Val	Leu	Thr	Glu	Leu	Gly	Ala	Gly	Arg				
	1875					1880						1885							
Leu	Pro	Gly	Val	Leu	Arg	Pro	Lys	Ala	Asp	Ala	Ala	Trp	His	Leu	His				
	1890					1895					1900								
Glu	Leu	Thr	Glu	Asp	Lys	Asp	Leu	Ser	Ala	Phe	Val	Leu	Phe	Ser	Ser				
1905					1910					1915					1920				
Ala	Ala	Ala	Thr	Leu	Gly	Thr	Pro	Ala	Gln	Ala	Asn	Tyr	Ala	Ala	Ala				
				1925					1930					1935					
Asn	Ala	Phe	Leu	Asp	Ala	Leu	Ala	Glu	Arg	Arg	Arg	Ala	Ala	Gly	Leu				
			1940					1945					1950						
Pro	Gly	Thr	Ala	Ala	Ala	Trp	Gly	Leu	Trp	Ala	Asp	Ala	Thr	Gly	Leu				
	1955						1960					1965							
Thr	Arg	His	Leu	Asp	Ala	Ala	Asp	Val	Ala	Arg	Ala	Gly	Arg	Asn	Arg				
	1970						1975				1980								
Ile	Val	Pro	Met	Ala	Ala	Ala	Glu	Gly	Leu	Ala	Leu	Phe	Asp	Thr	Ala				
1985					1990					1995					2000				
Thr	Ala	Thr	Gly	Asp	Ala	Val	Thr	Val	Thr	Ala	Arg	Leu	Asp	Leu	Ala				
				2005					2010					2015					
Thr	Ala	Ser	Ala	Ala	Pro	Thr	Pro	Pro	Leu	Leu	Arg	Gly	Leu	Val	Ala				
			2020				2025						2030						
Thr	Pro	Ala	Gly	Pro	Ala	Ala	Ala	Arg	Pro	Val	Pro	Gly	Ala	Ala	Ala				
	2035						2040					2045							
Leu	Val	Ala	Arg	Ile	Thr	Gly	Leu	Pro	Ala	Pro	Glu	Arg	Ala	Pro	Ala				
	2050					2055					2060								
Leu	Leu	Asp	Val	Val	Arg	Gly	Gln	Val	Ala	Asp	Val	Leu	Gly	His	Arg				
2065					2070					2075					2080				
Gly	Arg	Asp	Ala	Val	Ala	Pro	Asp	Arg	Gly	Phe	Lys	Glu	Leu	Gly	Phe				
			2085						2090					2095					
Asp	Ser	Leu	Thr	Ala	Val	Glu	Leu	Arg	Asn	Arg	Leu	Gly	Thr	Ala	Thr				
			2100					2105					2110						
Gly	Leu	Arg	Leu	Pro	Thr	Thr	Ile	Val	Phe	Asp	His	Pro	Asn	Pro	Ala				
	2115						2120				2125								
Ala	Leu	Ala	Asp	His	Leu	Leu	Asp	Ala	Leu	Leu	Pro	Ser	Glu	Ser	Gly				
	2130					2135					2140								
Glu	Ala	Ser	Ala	Asp	Arg	Ile	Ile	Ala	Glu	Leu	Ala	Arg	Val	Glu	Ser				
2145					2150					2155					2160				
Ala	Leu	Gly	Ala	Leu	Pro	Ala	Asp	Gly	Thr	Asp	Arg	Ala	Arg	Val	Ala				
			2165						2170					2175					
Ala	His	Leu	Arg	Asp	Leu	Ala	Ala	Arg	Trp	Asp	Ala	Gly	Thr	Thr	Asp				
			2180					2185					2190						
Gly	Thr	Pro	Glu	Arg	Ala	Ala	Leu	Asp	Gly	Val	Thr	Ala	Asp	Glu	Leu				
	2195						2200				2205								
Phe	Asp	Leu	Ile	Asp	Arg	Gly	Val	Ser											

<210> 9
 <211> 4206

<212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 9

5

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

ES 2 556 211 T3

Ser Ser Arg Thr Ala Ala Gly Leu Arg Ala Gln Ala Glu Arg Leu Leu
485 490 495
Ala His Arg Pro Ala His Asp Ala Ala Pro His Asp Val Ala Leu Ala
500 505 510
Leu Ala Thr Thr Arg Thr Ala Phe Glu His Arg Val Val Val Leu Gly
515 520 525
Ala Asp His Asp Thr Leu Leu Ala Gly Leu Thr Ala Val Ala Glu Gly
530 535 540
Ala Glu Ser Ala Asp Val Val Arg Gly Arg Ala Val Gly Asp Gly Arg
545 550 555 560
Ala Val Phe Val Phe Pro Gly Gln Gly Ala Gln Trp Val Gly Met Ala
565 570 575
Val Asp Leu Leu Asp Ser Ser Pro Val Phe Ala Gly Arg Met Ala Glu
580 585 590
Cys Ala Val Ala Leu Glu Pro Phe Val Gly Trp Ser Leu Arg Gly Val
595 600 605
Leu Gly Asp Pro Val Ala Leu Glu Arg Val Asp Val Val Gln Pro Val
610 615 620
Leu Trp Ala Val Met Val Ser Leu Ala Glu Val Trp Arg Ser Tyr Gly
625 630 635 640
Val Val Pro Ser Ala Val Val Gly His Ser Gln Gly Glu Ile Ala Ala
645 650 655
Ala Cys Val Ala Gly Val Leu Ser Leu Ala Asp Gly Ala Arg Val Val
660 665 670
Ala Leu Arg Ser Arg Ala Leu Thr Ala Leu Ala Gly Ser Gly Gly Met
675 680 685
Val Ser Val Ala Ala Gly Pro Ser Gly Val Glu Glu Leu Leu Val Gly
690 695 700
Trp Ala Gly Arg Leu Ala Val Ala Ala Val Asn Gly Pro Glu Ser Val
705 710 715 720
Val Val Ala Gly Glu Gly Val Ala Leu Glu Glu Phe Leu Ala His Cys
725 730 735
Gly Gly Arg Gly Val Arg Ala Arg Arg Ile Ala Val Asp Tyr Ala Ser
740 745 750
His Ser Val Leu Val Glu Pro Val Arg Glu Ala Leu Leu Ala Asp Leu
755 760 765
Glu Gly Val Arg Pro Gly Glu Gly Thr Val Pro Leu Phe Ser Thr Val
770 775 780
Thr Gly Glu Trp Ala Asp Gly Thr Ala Leu Asp Ala Gly Tyr Trp Tyr
785 790 795 800
Arg Asn Leu Arg Glu Pro Val Arg Tyr Ala Asp Ala Val Ala Ala Leu
805 810 815
Leu Ala Gln Gly His Arg Gly Phe Ile Glu Val Ser Pro His Pro Val
820 825 830
Leu Thr Val Gly Thr Gln Glu Thr Val Glu Arg Thr Glu Thr Gly Ala
835 840 845
Ala Val Val Gly Thr Leu Gln Arg Asp Arg Ala Gly Leu Pro Thr Leu
850 855 860
Leu Thr Asn Leu Ala Glu Ala His Thr His Gly Val Arg Ile Asp Trp
865 870 875 880
Thr Ala Phe Phe Arg Gly Thr Gly Ala His Pro Ala Asp Leu Pro Thr
885 890 895
Tyr Ala Phe Gln Arg Asp Arg Tyr Trp Pro Pro Val Asp Ala Ala Ala
900 905 910
Arg Pro Leu Pro Pro Ala Ala Thr Pro Val Ala Leu Pro Ala Gln Arg
915 920 925
Thr Ala Asp Glu Gly Thr Pro Trp Arg Asp Glu Leu Ala Gly Leu Thr
930 935 940
Pro Ala Glu Arg Arg His Arg Val Met Asp Leu Val Arg Leu Arg Thr
945 950 955 960

Ala Ala Val Leu Gly His Ala Asp Pro Ala Ala Ile Glu Pro His Arg
 965 970 975
 Gly Phe Ala Ala Leu Gly Phe Asp Ser Leu Ala Ser Leu Arg Leu Arg
 980 985 990
 Thr Ala Leu Thr Glu Ala Thr Gly Leu Thr Leu Ala Ser Ser Val Val
 995 1000 1005
 Phe Asp His Pro Asn Pro Ala Ala Leu Thr Asp His Leu Leu Thr Arg
 1010 1015 1020
 Leu Asp Gly Ala Pro Val Pro Ala Arg Pro Ala Val Arg Ala Thr Pro
 1025 1030 1035 1040
 Ala Asp Glu Pro Leu Ala Val Val Gly Met Ala Cys Arg Phe Pro Gly
 1045 1050 1055
 Gly Val Ala Ser Pro Asp Asp Leu Trp Gln Leu Leu Ala Glu Gly Arg
 1060 1065 1070
 Asp Val Ile Gly Asp Phe Pro Ala Asp Arg Gly Trp Asp Leu Asp Gly
 1075 1080 1085
 Leu Tyr Asp Pro Asp Pro Asp Arg Ala Gly His Thr Tyr Leu Arg Gln
 1090 1095 1100
 Gly Gly Phe Leu Asn Asp Ala Ala Gly Phe Asp Ala Asp Phe Phe Gly
 1105 1110 1115 1120
 Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Phe
 1125 1130 1135
 Leu Glu Val Ala Trp Glu Ala Phe Glu His Ala Gly Val Asp Pro His
 1140 1145 1150
 Gly Leu Arg Gly Ser Ala Thr Gly Val Phe Ala Gly Val Thr Asp Gln
 1155 1160 1165
 Arg Tyr Asp Ser Arg His Gly Ala Val Ala Gly Val Asp Glu Gly Leu
 1170 1175 1180
 Leu Gly Thr Gly Asn Tyr Ala Ser Val Leu Ser Gly Arg Val Ala Tyr
 1185 1190 1195 1200
 Thr Leu Gly Leu Glu Gly Pro Ala Val Ser Val Asp Thr Ala Cys Ser
 1205 1210 1215
 Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ser Leu Arg Ser Gly
 1220 1225 1230
 Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Met Val Met Ser Thr Pro
 1235 1240 1245
 Arg Ala Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly
 1250 1255 1260
 Arg Cys Lys Ala Phe Ala Ala Ala Ala Asp Gly Ile Gly Trp Ser Glu
 1265 1270 1275 1280
 Gly Ala Gly Val Val Val Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn
 1285 1290 1295
 Gly His Pro Val Leu Ala Val Val Arg Gly Ser Ala Ile Asn Gln Asp
 1300 1305 1310
 Gly Ala Ser Ser Gly Leu Thr Ala Pro Asn Gly Ser Ala Gln Gln Arg
 1315 1320 1325
 Val Ile Arg Ser Ala Leu Ala Ala Ala Gly Leu Thr Ala Ala Asp Val
 1330 1335 1340
 Asp Val Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile
 1345 1350 1355 1360
 Glu Ala Asn Ala Leu Leu Ala Thr Tyr Gly Gln Asp Arg Ala Gln Pro
 1365 1370 1375
 Leu Arg Leu Gly Ser Leu Lys Ser Asn Ile Gly His Thr Gly Pro Ala
 1380 1385 1390
 Ala Gly Val Ala Gly Val Ile Lys Thr Val Leu Ala Leu Arg Ala Gly
 1395 1400 1405
 Glu Val Pro Arg Thr Leu His Val Asp Ala Pro Ser Pro His Ile Asp
 1410 1415 1420
 Trp Ala Ser Gly Ser Val Glu Leu Val Thr Gly Ala Ala Val Trp Pro
 1425 1430 1435 1440

Pro Thr Asp Arg Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Ala Ser
 1445 1450 1455
 Gly Thr Asn Ala His Leu Ile Leu Glu Glu Ala Pro Ala Glu Pro Ala
 1460 1465 1470
 Ala Pro Ala Arg Gly Ala Ala Pro Arg Trp Val Pro Trp Leu Leu Ser
 1475 1480 1485
 Ala Arg Ser Gln Ala Ala Leu Ala Glu Gln Ala Thr Arg Leu Ala Ala
 1490 1495 1500
 His Leu Asp Thr His Pro Gly Leu Asp Pro Val Asp Leu Ser Trp Ser
 1505 1510 1515 1520
 Leu Ala Ala Ser Arg Ala Leu Leu Pro His Arg Ala Val Val Val Ala
 1525 1530 1535
 Ala Asp Val Ala Gly Ala Arg Ala Ser Leu Ala Ala Leu Ala Ala Gly
 1540 1545 1550
 Glu Pro Val Glu Gly Val Val Ser Gly Val Ala Gly Leu Pro Gly Glu
 1555 1560 1565
 Gly Arg Val Val Phe Val Phe Pro Gly Gln Gly Ala Gln Trp Val Gly
 1570 1575 1580
 Met Ala Val Asp Leu Leu Asp Ser Ser Pro Val Phe Ala Gly Arg Met
 1585 1590 1595 1600
 Ala Glu Cys Ala Val Ala Leu Glu Pro Phe Val Gly Trp Ser Leu Arg
 1605 1610 1615
 Gly Val Leu Gly Asp Pro Val Ala Leu Glu Arg Val Asp Val Val Gln
 1620 1625 1630
 Pro Val Leu Trp Ala Val Met Val Ser Leu Ala Glu Val Trp Arg Ser
 1635 1640 1645
 Tyr Gly Val Val Pro Ser Ala Val Val Gly His Ser Gln Gly Glu Ile
 1650 1655 1660
 Ala Ala Ala Cys Val Ala Gly Val Leu Ser Leu Ala Asp Gly Ala Arg
 1665 1670 1675 1680
 Ile Val Ala Leu Arg Ser Arg Leu Ile Ala Glu Arg Leu Ala Gly Arg
 1685 1690 1695
 Gly Gly Met Val Ser Val Ala Leu Pro Glu Ala Thr Val Thr Arg Arg
 1700 1705 1710
 Leu Ala Pro Trp Ser Gly Arg Val Cys Val Ala Ala Leu Asn Gly Pro
 1715 1720 1725
 Ser Ser Val Val Leu Ser Gly Asp Pro Asp Ala Leu Asp Glu Val Met
 1730 1735 1740
 Ala Ala Trp Ser Ala Asp Gly Val Arg Leu Arg Arg Ile Ala Val Asp
 1745 1750 1755 1760
 Tyr Ala Ser His Ser Ala His Val Glu Ser Leu Glu Ala Glu Leu Arg
 1765 1770 1775
 Ala Ala Leu Ala Glu Leu Arg Pro Gly Glu Ala Gly Ile Ala Phe His
 1780 1785 1790
 Ser Thr Leu Leu Gly Gly Pro Leu Gly Glu Thr Arg Leu Asp Ala Gly
 1795 1800 1805
 Tyr Trp Tyr Arg Asn Leu Arg Glu Pro Val Arg Phe Glu Pro Val Val
 1810 1815 1820
 Arg Gly Leu Leu Asp Ser Gly His Ala Val Phe Val Glu Ile Ser Pro
 1825 1830 1835 1840
 His Pro Val Leu Thr Ala Ala Val Gln Glu Thr Ala Glu Ala Thr Glu
 1845 1850 1855
 Arg Thr Ala Val Val Val Gly Thr Leu Arg Arg Asp Glu Asp Gly Pro
 1860 1865 1870
 Arg Arg Leu Leu Thr Ser Leu Ala Glu Ala Ala Val His Gly Val Ser
 1875 1880 1885
 Val Asp Trp Thr Val Gly Cys Pro Asp Gly Arg His Val Asp Leu Pro
 1890 1895 1900
 Thr Tyr Ala Phe Gln Arg Arg Arg Phe Trp Pro Ser Gly Pro Val Ala
 1905 1910 1915 1920

ES 2 556 211 T3

Ser Asp Ala Ala Gly Leu Gly Leu Thr Gly Ala Gly His Pro Leu Leu
1925 1930 1935
Gly Ala Val Thr Pro Leu Ala Gly Ser Gly Gly Gln Val Phe Thr Gly
1940 1945 1950
Arg Val Pro Ala Gly Thr Glu Leu Pro Ala Gly Ala Val Leu Asp Leu
1955 1960 1965
Ala Leu His Ala Ala Asp Gly Arg Thr Leu Gly Glu Leu Thr Glu Glu
1970 1975 1980
Ala Pro Leu Asp Thr Val Gly Glu Ala Arg Arg Leu Gln Val Thr Leu
1985 1990 1995 2000
Gly Ala Glu Asn Glu Asp Gly Ala Arg Pro Val Ala Val His Ser Arg
2005 2010 2015
Pro Ala Asp Ala Asp Asp Asp Gln Pro Trp Thr Arg His Ala Thr Gly
2020 2025 2030
Val Leu Leu Pro His Thr Gly Thr Ala Pro Ala Ala Ala Pro Gly Asp
2035 2040 2045
Pro Asp Thr Asp Val Thr Val Glu Leu Thr Asp Val Thr Val Glu Leu
2050 2055 2060
Thr Asp Glu Thr Ala Gly Gly Trp Gly Val His Pro Ala Leu Leu Thr
2065 2070 2075 2080
Asp Ala Leu Ala Ala Val His Pro Gly Leu Val Pro Gly Ala Trp His
2085 2090 2095
Gly Val Thr Leu His Ala Val Gly Ala Thr Arg Leu Arg Val Arg Leu
2100 2105 2110
His Pro Ala Gly Glu His Thr Val Thr Leu His Ala Thr Asp Asp Ser
2115 2120 2125
Gly Ala Ala Val Leu Thr Val Asp Ala Val Thr Leu Arg Pro Pro Ala
2130 2135 2140
Thr Ala Gly Pro Gly Arg Asp His Glu Leu Tyr Arg Val Glu Trp Thr
2145 2150 2155 2160
Pro Val Pro Leu Pro Ala Ala Asp Val Asp Gly Ile Ala Val Leu Gly
2165 2170 2175
Gly Leu Pro Leu Pro Gly Tyr Pro Ser Cys Pro Asp Leu Ala Ala Val
2180 2185 2190
Ala Ala Leu Asp Thr Val Pro Gly Thr Leu Val Leu Pro Cys Arg Gln
2195 2200 2205
Gln Ser Ala Asp Asp Thr Ala Gly Ala Ala His Ala Gly Ala Arg Arg
2210 2215 2220
Val Leu Thr Ala Leu Gln Glu Trp Leu Ala Asp Glu Arg Leu Ala Gly
2225 2230 2235 2240
Thr Arg Leu Ala Val Leu Thr His Gly Ala Val Pro Val Glu His Glu
2245 2250 2255
Asp Val Thr Asp Leu Ala His Ala Pro Val Trp Gly Leu Ile Arg Ser
2260 2265 2270
Ala Arg Ala Glu His Pro Gly Arg Leu Val Leu Val Asp Ile Asp Gly
2275 2280 2285
Pro Asp Ala Leu Gln Gln Leu Pro Ala Val Leu Ala Thr Gly Glu Pro
2290 2295 2300
Glu Ile Ala Val Arg Ser Gly Arg Val Leu Ala Pro Arg Leu Val Arg
2305 2310 2315 2320
Ala Pro Arg Thr Asp Glu Pro Thr Ser Gln Ala Ala Leu Trp Pro Ser
2325 2330 2335
Asp Gly Val Val Leu Val Thr Gly Gly Thr Gly Thr Leu Gly Ala Leu
2340 2345 2350
Cys Ala Arg His Leu Val Thr Glu His Gly Val Arg Arg Leu Val Leu
2355 2360 2365
Ala Gly Arg Arg Gly Asp Ala Ala Pro Asp Ala Val Thr Leu Ala Gly
2370 2375 2380
Glu Leu Arg Ala Leu Gly Ala Glu Val Thr Val Ala Ala Cys Asp Ala
2385 2390 2395 2400

ES 2 556 211 T3

Ala Asp Arg Thr Glu Leu Ala Ala Leu Leu Asp Arg Ile Thr Ala Glu
2405 2410 2415
His Pro Leu Thr Gly Val Val His Ala Ala Gly Val Leu Asp Asp Gly
2420 2425 2430
Val Ile Ala Ser Gln Thr Pro Glu Arg Leu Ala Lys Val Leu Arg Pro
2435 2440 2445
Lys Val Asp Ala Ala Trp Asn Leu His Glu Leu Thr Ala Pro His Arg
2450 2455 2460
Pro Ala Val Phe Val Leu Phe Ser Ser Thr Ser Gly Leu Phe Gly Ala
2465 2470 2475 2480
Pro Gly Gln Gly Asn Tyr Ala Ala Gly Asn Ala Phe Val Asp Ala Leu
2485 2490 2495
Ala Ala His Arg Arg Ala Gln Gly Leu Pro Ala Thr Ser Gln Val Trp
2500 2505 2510
Gly Leu Trp Ala His Ala Ser Gly Met Thr Gly His Leu Gly Gly Ala
2515 2520 2525
Asp Leu Arg Arg Ala Ala Arg Asp Gly Val Val Pro Leu Pro Thr Pro
2530 2535 2540
Asp Ala Leu Ala Leu Phe Asp Arg Ala Thr Ala Gly Ser Ala Pro Val
2545 2550 2555 2560
Val Val Pro Ala Trp Leu Asp Leu Thr Ser Phe Ala Thr Gly Val Thr
2565 2570 2575
Ala Val Pro Ala Leu Met Arg Arg Leu Val Arg Gly Pro Val Arg Arg
2580 2585 2590
Ala Ala Thr Ala Gly Ala Gly Pro Asp Thr Leu Ala Gly Lys Leu Thr
2595 2600 2605
Gly Leu Thr Ala Ala Glu Arg Glu Arg Thr Leu Leu Thr Leu Val Arg
2610 2615 2620
Ser His Ala Ala Val Val Leu Gly His Thr Asp Asp Thr Ala Val Thr
2625 2630 2635 2640
Pro Gly Arg Ala Phe Lys Glu Leu Gly Phe Asp Ser Leu Thr Ala Val
2645 2650 2655
Glu Leu Arg Asn Arg Leu Ser Ala Ala Thr Gly Leu Arg Leu Pro Ala
2660 2665 2670
Thr Leu Val Phe Asp His Pro Asn Pro Arg Ser Leu Ala Gly His Leu
2675 2680 2685
Leu Ala Glu Leu Leu Gly Glu Arg Ala Glu Glu Thr Ala Pro Val Pro
2690 2695 2700
Ala Val Ala Arg Pro Ala Asp Asp Asp Pro Ile Ala Ile Val Gly Met
2705 2710 2715 2720
Ala Cys Arg Tyr Pro Gly Gly Val Thr Ser Pro Glu Asp Leu Trp Arg
2725 2730 2735
Leu Leu Ala Asp Glu Arg Asp Ala Leu Thr Pro Phe Pro Asp Asp Arg
2740 2745 2750
Gly Trp Asp Leu Ala Gly Leu Phe His Pro Asp Pro Glu His Ala Gly
2755 2760 2765
Thr Ser Tyr Val Arg Glu Gly Gly Phe Leu Ala Asp Val Ala Gly Phe
2770 2775 2780
Asp Ala Asp Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp
2785 2790 2795 2800
Pro Gln Gln Arg Leu Ala Leu Glu Thr Ala Trp Glu Ala Val Glu Arg
2805 2810 2815
Ala Gly Ile Asp Pro Lys Ser Leu Arg Gly Ala Asp Val Gly Val Tyr
2820 2825 2830
Leu Gly Thr Asn Gly Gln Asp Tyr Ala Gln Leu Val Arg Arg Thr Val
2835 2840 2845
Glu Ser Ala Glu Gly Tyr Val Gly Ile Gly Asn Ser Ala Ser Val Leu
2850 2855 2860
Ser Gly Arg Ile Ala Tyr Val Leu Gly Leu Glu Gly Pro Ala Val Thr
2865 2870 2875 2880

ES 2 556 211 T3

Val Asp Thr Ala Cys Ser Ala Ser Leu Val Ala Leu His Trp Ala Ile
 2885 2890 2895
 Gln Ala Leu Arg Ser Gly Glu Cys Ser Met Ala Leu Ala Gly Gly Val
 2900 2905 2910
 Thr Val Met Ser Ala Pro Asp Val Phe Val Asp Phe Ser Arg Gln Arg
 2915 2920 2925
 Gly Leu Ala Val Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp
 2930 2935 2940
 Gly Thr Gly Trp Ser Glu Gly Ala Gly Met Leu Leu Val Glu Arg Leu
 2945 2950 2955 2960
 Ser Asp Ala Arg Arg His Gly His Gln Val Leu Ala Val Val Arg Gly
 2965 2970 2975
 Thr Ala Ile Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn
 2980 2985 2990
 Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ser Gly Ala Gly
 2995 3000 3005
 Leu Thr Pro Ala Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr
 3010 3015 3020
 Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly
 3025 3030 3035 3040
 Gln Glu Arg Asp Glu Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile
 3045 3050 3055
 Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val
 3060 3065 3070
 Gln Ala Met Arg His Gly Thr Leu Pro Arg Thr Leu His Val Asp Glu
 3075 3080 3085
 Pro Ser Pro Glu Val Asp Trp Ala Ser Gly Ala Val Glu Leu Leu Thr
 3090 3095 3100
 Glu Ala Arg Glu Trp Ala Arg Ala Gly Arg Pro Arg Arg Ala Gly Val
 3105 3110 3115 3120
 Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Val Leu Glu Glu
 3125 3130 3135
 Ala Pro Glu Glu Thr Gly Arg Thr Ala Pro Ala Ala Leu Pro Ala Val
 3140 3145 3150
 Pro Trp Leu Leu Ser Ala Arg Ser Glu Thr Ala Leu Arg Gly Gln Val
 3155 3160 3165
 Glu Arg Leu Arg Ala Tyr Val Thr Glu His Pro Glu Ala Arg Pro Ala
 3170 3175 3180
 Asp Ile Gly Leu Ser Leu Leu Thr Ala Arg Ser Arg Phe Glu His Ser
 3185 3190 3195 3200
 Ala Val Val Val Gly Thr Asp His Asp Ala Leu Leu Ala Ala Leu Ala
 3205 3210 3215
 Asp Pro Glu Pro Leu His Ala Arg Asp Gly Leu Thr Ala Phe Val Phe
 3220 3225 3230
 Ala Gly Gln Gly Ala Gln Arg Val Gly Met Gly Ala Glu Leu Ala Ala
 3235 3240 3245
 Ala Tyr Pro Val Phe Ala Gln Val Phe Ala Gln Val Cys Ala Ala Phe
 3250 3255 3260
 Asp Gly Val Leu Glu Arg Pro Leu Gly Glu Val Val Ala Glu Gly Gly
 3265 3270 3275 3280
 Pro Glu Leu Asp Arg Thr Val Tyr Ala Gln Ala Gly Leu Phe Ala Phe
 3285 3290 3295
 Glu Val Ala Leu Phe Arg Leu Leu Glu Ser Trp Gly Val Ala Pro Asp
 3300 3305 3310
 Val Val Leu Gly His Ser Val Gly Glu Leu Thr Ala Ala Cys Val Ala
 3315 3320 3325
 Gly Val Trp Ser Leu Glu Asp Ala Val Arg Val Val Ala Ala Arg Gly
 3330 3335 3340
 Arg Leu Met Gln Ala Leu Pro Glu Gly Gly Ala Met Val Ala Leu Glu
 3345 3350 3355 3360

ES 2 556 211 T3

Val Ser Ala Gly Glu Leu Glu Leu Pro Val Gly Val Glu Leu Ala Ala
 3365 3370 3375
 Val Asn Gly Pro Thr Ser Val Val Leu Ser Gly Glu Glu Glu Ala Val
 3380 3385 3390
 Leu Ala Glu Ala Ala Arg Trp Pro Asp Arg Arg Ala Lys Arg Leu Arg
 3395 3400 3405
 Val Ser His Ala Phe His Ser His Arg Met Asp Pro Met Leu Glu Asp
 3410 3415 3420
 Phe Arg Arg Val Leu Glu Ser Val Thr Phe Gln Ala Pro Glu Val Ala
 3425 3430 3435 3440
 Phe Val Ser Thr Val Thr Gly Ala Ala Val Thr Asp Glu Leu Cys Asp
 3445 3450 3455
 Pro Gly Tyr Trp Val Arg Asn Val Arg Glu Thr Val Arg Phe Ala Asp
 3460 3465 3470
 Ala Val Ala Ala Val Arg Ala Thr Gly Ala Asp Ala Leu Val Glu Ile
 3475 3480 3485
 Gly Pro Asp Ala Ala Leu Ala Pro Leu Ala Glu Gly Gly Val Pro Leu
 3490 3495 3500
 Leu Arg Arg Gly Arg Pro Glu Ala Val Ala Leu Val Glu Gly Leu Ala
 3505 3510 3515 3520
 Arg Ser Gln Met Ala Val Asp Trp Glu Lys Phe Phe Gly Ala Gly Val
 3525 3530 3535
 Ser Pro Val Asp Leu Pro Thr Tyr Pro Phe Gln His Ser Arg Phe Trp
 3540 3545 3550
 Pro Glu Val Pro Val Arg Ala Glu Ala Asp Pro Ala Ala Arg Trp Arg
 3555 3560 3565
 Tyr Arg Val Arg Trp Glu Pro Leu Arg Pro Ala Val Asp Ala Val Pro
 3570 3575 3580
 Ala Gly Arg Trp Leu Val Ala Val Pro Ala Ala Gly Ala Asp Ala Ala
 3585 3590 3595 3600
 Leu Val Arg Glu Cys Leu Lys Gly Leu Ala Gly Arg Gly Val Glu Val
 3605 3610 3615
 Ala Glu Leu Ala Val Ala Gly Ala Pro Asp Arg Ala Arg Leu Ala Glu
 3620 3625 3630
 Ala Val Ser Ala Thr Gly Pro Val Asp Gly Val Leu Ser Leu Leu Ala
 3635 3640 3645
 Ala Gly Pro Asp Ala Ala Ala Gly Thr Leu Val Leu Ala Gln Ala Leu
 3650 3655 3660
 Gly Asp Ala Gly Thr Asp Ala Pro Leu Trp Cys Val Thr Gly Gly Ala
 3665 3670 3675 3680
 Val Ala Ala Ala Asp Gly Glu Pro Ala Asp Pro Glu Gln Ala Gln Val
 3685 3690 3695
 Trp Gly Leu Gly Arg Val Leu Cys Leu Glu Ala Pro His Arg Trp Gly
 3700 3705 3710
 Gly Leu Val Asp Leu Pro Ala Glu Pro Asp Glu Gln Ala Val Val Gln
 3715 3720 3725
 Leu Cys Ala Val Leu Ala Gly His Asp Asp Glu Asp Gln Val Ala Val
 3730 3735 3740
 Arg Ala Ala Gly Val Phe Gly Arg Arg Leu Val Arg Pro Ala Thr Gly
 3745 3750 3755 3760
 Ala Ala Ala Glu Gly Trp Arg Pro Ser Gly Thr Val Leu Ile Thr Gly
 3765 3770 3775
 Gly Thr Gly Ala Leu Gly Gly His Thr Ala Arg Trp Ala Ala Arg Ser
 3780 3785 3790
 Gly Ala Gly Arg Val Val Leu Val Ser Arg Arg Gly Pro Asp Ala Asp
 3795 3800 3805
 Gly Ala Ala Glu Leu Val Ala Glu Leu Gly Glu Leu Gly Cys Gln Ala
 3810 3815 3820
 Val Ala Glu Ala Cys Asp Val Ala Asp Arg Ala Ala Leu Gly Ala Leu
 3825 3830 3835 3840

Leu Asp Lys Ile Ala Ala Asp Gly Pro Pro Leu Thr Ala Val Val His
 3845 3850 3855
 Thr Ala Gly Val Leu Asp Asp Gly Val His Gly Ser Leu Thr Pro Glu
 3860 3865 3870
 Arg Leu Ala Thr Val Leu Arg Ala Lys Ala Asp Gly Ala Thr Ala Leu
 3875 3880 3885
 His Glu Leu Thr Ala His Leu Pro Leu Glu Ala Phe Val Leu Phe Ala
 3890 3895 3900
 Ala Leu Gly Gly Val Val Gly Gly Ala Gly Gln Ala Asn Tyr Ala Ala
 3905 3910 3915 3920
 Ala Asn Ala His Leu Asp Ala Leu Ala Ala Ser Arg Arg Ala Ala Gly
 3925 3930 3935
 Leu Pro Ala Thr Ala Val Ala Trp Gly Ala Trp Ala Gly Gly Gly Met
 3940 3945 3950
 Ala Asp Ala Asp Glu Val Arg Arg Arg Leu Asp Arg Asp Gly Leu Leu
 3955 3960 3965
 Pro Met Asp Pro Arg Arg Ala Leu Asp Ala Leu Gly Arg Glu Ile Ala
 3970 3975 3980
 Ala Gly Asp Pro Ala Val Val Leu Ala Asp Val Asp Trp Thr Arg Leu
 3985 3990 3995 4000
 Ala Pro Asn Leu His Ala Val Arg Pro Ser Pro Leu Ile Ser Thr Val
 4005 4010 4015
 Pro Glu Ala Arg Arg Ala Val Ala Pro Glu Pro Gly Thr Ala Gly Gly
 4020 4025 4030
 Glu Ser Asp Pro Arg Gln Arg Leu Ala Ala Leu Pro Asp Gly Glu Arg
 4035 4040 4045
 Ala Arg Val Leu Leu Asp Leu Val Arg Asp Ala Ile Ala Ala Val Leu
 4050 4055 4060
 Gly Tyr Gly Gly Pro Gly Ala Val Asp Ile Thr Arg Gly Leu Val Asp
 4065 4070 4075 4080
 Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Ala
 4085 4090 4095
 Arg Ala Thr Gly Leu Ala Leu Pro Leu Thr Leu Val Phe Asp His Pro
 4100 4105 4110
 Asp Gly Glu Ala Leu Ala Ala His Leu Ala Ala Ala Leu Ala Pro Ala
 4115 4120 4125
 Ala Asp Arg Pro Gln Thr Thr Asp Leu Asp Arg Phe Ala Arg Leu Asp
 4130 4135 4140
 Pro Ala Thr Ala Asp Glu Pro Thr Arg Val Arg Val Ala Ala Glu Leu
 4145 4150 4155 4160
 Arg Arg Leu Leu Asp Ala Trp Thr Pro Pro Ala Ala Pro Ala Gly
 4165 4170 4175
 Ser Ala Gly Ser Ala Ala Asp Pro Ser Arg Leu Thr Thr Ala Ser Ala
 4180 4185 4190
 Asp Glu Ile Phe Asp Phe Ile Arg Asn Glu Leu Gly Lys Ser
 4195 4200 4205

<210> 10

<211> 3628

<212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 10

ES 2 556 211 T3

Val	Glu	Asn	Glu	Asn	Lys	Leu	Leu	Asp	His	Leu	Arg	Trp	Val	Thr	Gly
1				5					10					15	
Glu	Leu	Ala	Gln	Ala	Arg	Gln	Thr	Leu	Arg	Glu	Thr	Ala	Glu	Arg	Ala
			20					25					30		
Thr	Glu	Pro	Ile	Ala	Ile	Val	Gly	Met	Ala	Cys	Arg	Phe	Pro	Gly	Gly
		35					40					45			
Val	Ser	Thr	Pro	Glu	Asp	Leu	Trp	Arg	Leu	Leu	Ala	Asp	Glu	Arg	Asp

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

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

1010						1015										1020
Ala Gly Thr	Ala Gly Val	Val Ala Val	Leu Arg Pro	Asp Asp Pro	Asp											
1025		1030							1035							1040
Pro Leu Pro	Gly Leu Leu	Ala Leu Leu	Arg Ala His	Asp Arg Ala	Asp											
		1045							1050							1055
Gly Gly Leu	Pro Leu Trp	Cys Val Thr	Arg Gly Ala	Val Ala Thr	Gly											
		1060							1065							1070
Pro Ala Asp	Thr Pro Ala	Asp Pro Ala	Thr Ala Gln	Ile Trp Gly	Phe											
		1075							1080							1085
Gly Arg Val	Ala Ala Leu	Glu Arg Pro	Gln Ser Trp	Gly Gly Leu	Ile											
		1090							1095							1100
Asp Leu Pro	Ala Gly Ala	Thr Val Thr	Gly Asp Ala	Leu Ala Gly	Thr											
1105		1110							1115							1120
Ala Arg Ala	Asp Gly Pro	Pro Asp Ala	Arg Ile Ala	Asp Asp Val	Leu											
		1125							1130							1135
Gly Ala Arg	Leu Ala Asp	Arg Leu Val	Ala Ala Leu	Ala Gly Pro	Glu											
		1140							1145							1150
Asp Gln Val	Ala Val Arg	Ala Ser Gly	Ala Phe Gly	Arg Arg Leu	Arg											
		1155							1160							1165
Arg Ala Pro	Ala Asp Gly	Thr Thr Pro	Asp Trp Arg	Pro Thr Gly	Thr											
		1170							1175							1180
Ile Leu Val	Thr Gly Gly	Gly Thr Gly	Gly Gly Leu	Gly Gly His	Val											
		1185							1190							1195
Trp Leu Ala	Arg Ala Gly	Ala Glu His	Leu Val Leu	Thr Gly Arg	Arg											
		1205							1210							1215
Gly Pro Glu	Ala Pro Gly	Ala Ala Glu	Leu Ala Ala	Glu Leu Thr	Ala											
		1220							1225							1230
Leu Gly Ala	Glu Val Asp	Ile Val Ala	Cys Asp Ala	Ala Asp Arg	Asp											
		1235							1240							1245
Ala Leu Ala	Arg Val Leu	Ala Glu His	Pro Val Asp	Ala Val Phe	His											
		1250							1255							1260
Leu Ala Gly	Ile Glu Arg	Tyr Arg Pro	Leu Asp Glu	Leu Thr Pro	Ala											
		1265							1270							1275
Asp Leu Thr	Glu Val Ala	Ala Ala Lys	Val Thr Gly	Ala Leu Leu	Leu											
		1285							1290							1295
Asp Glu Leu	Thr Arg Asp	Arg Arg Leu	Ser Ala Phe	Val Leu Phe	Thr											
		1300							1305							1310
Ser Gly Ala	Gly Val Trp	Gly Ser Ser	Gly Gln Ala	Ala Tyr Ala	Ala											
		1315							1320							1325
Ala Asn Thr	Arg Leu Asp	Ala Leu Ala	Ala Arg Arg	Arg Ala Glu	Gly											
		1330							1335							1340
Leu Pro Ala	Thr Ala Val	Ala Trp Gly	His Trp Asp	Gly Ala Gly	Met											
		1345							1350							1355
Ser Asp Gly	Pro Ala Arg	Ala Gln Met	Leu Arg Trp	Gly Leu Pro	Ala											
		1365							1370							1375
Met Ala Pro	His Arg Ala	Val Ala Ala	Leu Ala Gly	Ala Leu Ala	Arg											
		1380							1385							1390
Asp Glu Ala	Ser Leu Val	Val Ala Asp	Ile Asp Trp	Pro Val Phe	Ala											
		1395							1400							1405
Pro Thr Phe	Thr Leu Ala	Arg Arg Ser	Pro Leu Leu	Ala Gly Leu	Pro											
		1410							1415							1420
Glu Ala Thr	Asp Pro Ala	Asp Thr Ala	Pro Pro Pro	Ala Ala Pro	Ala											
		1425							1430							1435
Gly Arg Leu	Asp Arg Ala	Ala Leu Leu	Glu Leu Val	Leu Ala Glu	Thr											
		1445							1450							1455
Gly Ala Val	Leu Gly His	Ala Ser Gly	Thr Gly Leu	Pro Ala Ala	Arg											
		1460							1465							1470
Ala Phe Gln	Glu Leu Gly	Phe Asp Ser	Leu Thr Ala	Val Glu Leu	Arg											
		1475							1480							1485
Asn Arg Leu	Arg Asp Ala	Thr Gly Leu	Pro Leu Pro	Ala Thr Leu	Val											

1490						1495										1500
Phe	Asp	His	Pro	Thr	Pro	Gln	Ala	Leu	Ala	Glu	Arg	Leu	Ala	Gly	Glu	
1505					1510					1515					1520	
Leu	Thr	Gly	Glu	Gln	Pro	Ala	Asp	Val	Pro	Asp	Thr	Thr	Pro	Ala	Pro	
				1525						1530				1535		
Ala	Ala	Ala	Ala	Asp	Asp	Asp	Pro	Val	Val	Ile	Val	Gly	Met	Gly	Cys	
			1540				1545						1550			
Arg	Phe	Pro	Gly	Gly	Ala	Asp	Ser	Pro	Glu	Arg	Leu	Trp	Asp	Leu	Val	
	1555						1560						1565			
Ala	Ala	Gly	Arg	Asp	Ala	Met	Thr	Gly	Phe	Pro	Ala	Asp	Arg	Gly	Trp	
	1570					1575						1580				
Asp	Val	Thr	Asp	Thr	Gly	Tyr	Thr	Lys	Ala	Gly	Gly	Phe	Leu	Asp	Ala	
1585					1590					1595				1600		
Ala	Gly	Glu	Phe	Asp	Ala	Gly	Phe	Phe	Gly	Ile	Ser	Pro	Arg	Glu	Ala	
				1605					1610					1615		
Thr	Ser	Met	Asp	Pro	Gln	Gln	Arg	Leu	Leu	Leu	Glu	Thr	Ala	Trp	Glu	
			1620						1625					1630		
Thr	Val	Glu	Arg	Ala	Gly	Ile	Asp	Pro	Ala	Ser	Leu	Arg	Gly	Ser	Arg	
	1635						1640							1645		
Thr	Gly	Val	Phe	Val	Gly	Tyr	Gly	Gly	Gln	Asp	Tyr	Leu	Thr	Ser	Leu	
	1650					1655								1660		
Tyr	Gly	Thr	Pro	Glu	Glu	Leu	Gln	Gly	His	Leu	Leu	Thr	Gly	Thr	Ser	
1665					1670					1675				1680		
Gly	Ser	Val	Val	Ser	Gly	Arg	Leu	Ala	Tyr	Val	Leu	Gly	Leu	Glu	Gly	
				1685					1690					1695		
Pro	Ala	Val	Thr	Ile	Asp	Thr	Ala	Cys	Ser	Ser	Ser	Leu	Val	Ala	Leu	
			1700					1705						1710		
His	Trp	Ala	Ile	Arg	Ala	Leu	Arg	Ser	Gly	Glu	Cys	Asp	Leu	Ala	Leu	
	1715						1720							1725		
Ala	Gly	Gly	Val	Thr	Val	Met	Ala	Thr	Pro	Gly	Val	Phe	Val	Glu	Phe	
	1730					1735						1740				
Gly	Arg	Gln	Gly	Gly	Leu	Ala	Gly	Asp	Gly	Arg	Cys	Lys	Ala	Phe	Ala	
1745					1750					1755				1760		
Asp	Ala	Ala	Asp	Gly	Thr	Gly	Trp	Gly	Glu	Gly	Ala	Gly	Leu	Leu	Leu	
			1765						1770					1775		
Val	Glu	Arg	Leu	Ser	Asp	Ala	Arg	Arg	Asn	Gly	His	Arg	Val	Leu	Ala	
			1780					1785						1790		
Val	Val	Arg	Gly	Ser	Ala	Val	Asn	Ser	Asp	Gly	Ala	Ser	Asn	Gly	Leu	
	1795						1800							1805		
Thr	Ala	Pro	Asn	Gly	Pro	Ser	Gln	Gln	Arg	Val	Ile	Arg	Ala	Ala	Leu	
	1810					1815								1820		
Ala	Asp	Gly	Gly	Leu	Thr	Pro	Ala	Asp	Ile	Asp	Ala	Val	Glu	Ala	His	
1825					1830					1835				1840		
Gly	Thr	Gly	Thr	Ala	Leu	Gly	Asp	Pro	Ile	Glu	Ala	Gln	Ala	Leu	Gln	
				1845						1850				1855		
Val	Val	Tyr	Gly	Gln	Asp	Arg	Glu	Arg	Pro	Leu	Trp	Ile	Gly	Ser	Val	
			1860					1865						1870		
Lys	Ser	Asn	Leu	Gly	His	Thr	Gln	Ala	Ala	Ser	Gly	Ala	Ala	Gly	Leu	
	1875						1880							1885		
Ile	Lys	Thr	Val	Leu	Ala	Leu	Arg	His	Gly	Val	Leu	Pro	Ala	Thr	Leu	
	1890					1895								1900		
His	Val	Asp	Arg	Pro	Ser	Ala	Gln	Val	Asp	Trp	Thr	Arg	Gly	Ala	Val	
1905					1910					1915				1920		
Ser	Val	Leu	Thr	Glu	Ser	Thr	Pro	Trp	Pro	Gly	Thr	Asp	Val	Pro	Arg	
			1925							1930				1935		
Arg	Ala	Ala	Val	Ser	Ser	Phe	Gly	Val	Ser	Gly	Thr	Asn	Ala	His	Val	
			1940					1945						1950		
Val	Leu	Glu	Gln	Ala	Pro	Pro	Ala	Gly	Glu	Ala	Gly	Pro	Ala	Gly	Asp	
	1955						1960							1965		
Gly	Gly	Pro	Val	Pro	Trp	Leu	Val	Ser	Gly	Arg	Thr	Pro	Glu	Ala	Val	

ES 2 556 211 T3

1970		1975		1980
Arg Ala Gln Val Glu Arg Leu Arg Ala Ser Leu Ala Gly His Pro Asp				
1985		1990		2000
Pro Val Ala Val Ala Arg Ala Leu Ala Thr Thr Arg Thr Ala Phe Glu				
		2005		2015
Tyr Arg Val Ala Ala Ala Gly Gly Asp Thr Glu Ala Leu Leu Asp Ala				
		2020		2030
Leu Ala Glu Ala Arg Pro Val Thr Ala Arg Gln Gly Arg Thr Ala Phe				
		2035		2045
Leu Cys Thr Gly Gln Gly Ala Gln Arg Ala Gly Met Gly Ala Gly Leu				
		2050		2060
Tyr Ala Ala His Pro Val Tyr Ala Asp Ala Phe Asp Ala Val Cys Ala				
2065		2070		2080
Glu Phe Asp Arg Leu Leu Asp Arg Pro Leu Arg Asp Leu Val Leu Ser				
		2085		2095
Gly Pro Ala Asp Val Leu Asp Arg Thr Ala Tyr Ala Gln Pro Ala Leu				
		2100		2110
Phe Ala Val Glu Ala Ala Leu Ala Ala Leu Leu Arg His Trp Gly Val				
		2115		2125
Thr Pro Asp Leu Leu Ala Gly His Ser Leu Gly Glu Ile Thr Ala Ala				
		2130		2140
His Leu Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Ala Leu Val Ala				
2145		2150		2160
Ala Arg Gly Arg Leu Met Asp Ala Leu Pro Ala Gly Gly Ala Met Val				
		2165		2175
Ala Val Glu Ala Asp Glu Asp Arg Val Gln Pro Leu Leu Gly Asp Asp				
		2180		2190
Val Cys Leu Ala Ala Val Asn Gly Pro Arg Ala Leu Val Leu Ser Gly				
		2195		2205
Arg Glu Glu Ala Val Asp Gly Val Ala Ala Arg Leu Ala Ala Glu Gly				
		2210		2220
Cys Arg Thr Arg Arg Leu Arg Val Ser His Ala Phe His Ser Ala Leu				
2225		2230		2240
Met Glu Pro Met Leu Asp Glu Phe Arg Ala Thr Val Ala Ala Leu Asp				
		2245		2255
Leu Arg Ala Pro Ser Val Pro Val Val Ser Ala Leu Thr Gly Arg Pro				
		2260		2270
Leu Thr Ala Asp Glu Ala Arg Ser Pro Glu His Trp Val Arg His Val				
		2275		2285
Arg Glu Ala Val Arg Phe His Asp Ala Val Arg Gly Leu Ala Ala Glu				
		2290		2300
Gly Ala Val Arg Tyr Leu Glu Leu Gly Pro Asp Gly Val Leu Thr Ala				
2305		2310		2320
Leu Ala Gln Ser Ser Leu Pro Pro Ala Asp Thr Asp Ala Asp Gly Arg				
		2325		2335
Asp Pro Leu Ala Val Pro Leu Leu Arg Ala Gly Arg Pro Glu Pro Glu				
		2340		2350
Thr Leu Thr Asp Ala Leu Ala Arg Ala Ala Ala Asp Gly Leu Thr Val				
		2355		2365
Asp Trp Ala Gly Tyr Phe Thr Gly Arg Gly Gly Ala Pro Val Glu Leu				
		2370		2380
Pro Thr Tyr Ala Phe Gln Arg Glu His Tyr Trp Leu Pro Val Asp Ser				
2385		2390		2400
Gly Ala Gly Thr Ala Pro Ala Gly His Pro Leu Leu Ser Ala Ala Val				
		2405		2415
Asp Leu Pro Asp Gly Gly Leu Val Leu Thr Gly Arg Leu Ser Pro Ala				
		2420		2430
Ala Arg Pro Trp Leu Ala Gln His Thr Val Arg Gly Ser Ala Leu Leu				
		2435		2445
Pro Gly Thr Ala Leu Leu Asp Leu Ala Leu Ala Ala Ala Gly Gln Ala				

2450					2455					2460					
Ala Ala Pro Gly Val	Ala Glu Leu Ile Leu	Glu Ala Pro Leu Val Leu													
2465		2470			2475					2480					
Pro Ala Glu Gly Ala Val	Glu Val Arg Val Thr Val	Gly Ala Ala Gly													
		2485			2490					2495					
Thr Asp Gly Arg Arg	Ala Ile Ala Leu His Thr Arg	Ala Gly Asp Gly													
		2500			2505					2510					
Asp Trp Thr Arg His	Ala Thr Gly Ala Leu Gly Glu Val	Pro Gly Glu													
		2515			2520					2525					
Pro Thr Ala Ala Gly	Ala Trp Pro Pro Pro Asp	Ala Arg Pro Ala Asp													
		2530			2535					2540					
Leu Ala Ala Leu Tyr	Gly Gly Leu Ala Asp Ala Gly Phe	Gly Tyr Gly													
2545		2550			2555					2560					
Pro Ala Tyr Gln Gly	Leu Arg Ala Ala Trp Arg Arg	Gly Glu Gly Pro													
		2565			2570					2575					
Ala Ala Glu Val Phe	Ala Glu Ala Glu Leu Pro Ala Ala	Val Pro Asp													
		2580			2585					2590					
Ala Asp Arg Cys Pro	Val His Pro Ala Leu Leu Asp	Ala Val Leu His													
		2595			2600					2605					
Ala Ile Gly Val Gly	Gly Leu Ile Thr Asp Pro Ala His	Gly Gly Leu													
		2610			2615					2620					
Pro Phe Ala Trp Thr	Gly Val Arg Val Phe Ala Pro Gly	Ala Arg Ala													
2625		2630			2635					2640					
Val Arg Ala Arg Leu	Ser Arg Ala Gly Ala Glu Gly	Ala Leu Ala Val													
		2645			2650					2655					
Asp Leu Phe Asp Ala	Asp Gly Leu Pro Val Ala Ala	Ile Gly Ser Leu													
		2660			2665					2670					
Arg Leu Arg Pro Pro	Ala Ala Pro Ala Val Pro Asp	Ala Leu Phe Glu													
		2675			2680					2685					
Thr Ala Trp Thr Pro	Val Glu Gln Gly Thr Ala Pro Ala	Arg Arg Leu													
		2690			2695					2700					
Ala Leu Leu Gly Ala	Asp Thr Ala Leu Ala Ala Gly	Leu Thr Ala Ala													
2705		2710			2715					2720					
Gly Ala Ala Leu Ala	Asp Ala Thr Asp Arg Ser Ala	Glu Val Leu Val													
		2725			2730					2735					
Leu Pro Ile Val Thr	Asp Pro Gly Ala Ala Pro Val Thr	Glu Thr His													
		2740			2745					2750					
Arg Ala Thr Ala Ala	Val Leu Thr Ala Leu Arg Asp	Val Leu Ala Asp													
		2755			2760					2765					
Glu Glu Ser Thr Ala	Arg Leu Ala Val Val Thr Arg	Gly Ala Leu Ala													
		2770			2775					2780					
Leu Ser Ala Glu Glu	Ser Pro Asp Pro Ala Ala Arg	Ala Val Trp Gly													
2785		2790			2795					2800					
Leu Val Arg Ser Ala	Gln Thr Glu His Pro Asp Arg	Ile Val Leu Ala													
		2805			2810					2815					
Asp Leu Asp Ala Ala	Asp Ala Ser Ala Arg Ala Leu	Pro Ala Ala Leu													
		2820			2825					2830					
Thr Cys Gly Glu Pro	Gln Leu Ala Val Arg Ser Gly	Ala Val Ser Ala													
		2835			2840					2845					
Pro Arg Leu Thr Arg	Ala Gly Ala Asp Ala Leu Val	Leu Pro Asp Gly													
		2850			2855					2860					
Gly Trp Arg Leu Arg	Pro Gly Ala Thr Gly Thr Val	Asp Gly Met Thr													
2865		2870			2875					2880					
Ala Val Pro His Pro	Asp Ala Pro Leu Ala Asp	Gly Glu Val Arg Val													
		2885			2890					2895					
Ala Val Arg Ala Val	Gly Val Thr Phe Arg Asp	Val Leu Ser Val Leu													
		2900			2905					2910					
Gly Leu Tyr Pro Gly	Ala Pro Gln Pro Leu Gly Ile	Glu Ala Ala Gly													
		2915			2920					2925					
Val Val Thr Gly Thr	Gly Pro Gly Val Ser Asp	Leu Ala Pro Gly Asp													

3410						3415						3420					
Asp	Asp	Ala	Asp	Thr	Gln	Pro	Ala	Leu	Pro	Gln	Arg	Leu	Ala	Gly	Leu		
3425						3430						3435					3440
Gly	Glu	Glu	Gln	Arg	Arg	Arg	Ala	Val	Leu	Asp	Val	Val	Arg	Arg	Asn		
				3445						3450					3455		
Ala	Ala	Ala	Val	Leu	Gly	His	Ala	Arg	Ala	Ser	Ala	Val	Asp	Thr	Ala		
			3460							3465				3470			
Arg	Gly	Phe	Leu	Asp	Leu	Gly	Phe	Asp	Ser	Leu	Thr	Ala	Val	Glu	Leu		
		3475						3480						3485			
Arg	Asn	Arg	Leu	Thr	Glu	Ala	Thr	Gly	Leu	Arg	Leu	Ser	Ala	Thr	Ala		
	3490						3495							3500			
Val	Phe	Asp	His	Pro	Thr	Pro	Ala	Ala	Leu	Ala	Arg	His	Leu	Leu	Thr		
3505					3510							3515					3520
Glu	Leu	Glu	Pro	Leu	Val	Arg	Ala	Ala	Arg	Ser	Ala	Leu	Pro	Ser	Ala		
				3525						3530					3535		
Pro	Asp	Pro	Asp	Ala	Asp	Leu	Arg	Gly	Ala	Ile	Ala	Ala	Ile	Pro	Leu		
			3540						3545						3550		
Glu	Arg	Leu	Arg	Gln	Ala	Gly	Leu	Leu	Asp	Glu	Leu	Ala	Arg	Leu	Ala		
		3555					3560						3565				
Gly	Val	Ala	Val	Pro	Ala	Lys	Asp	Ala	Pro	Ala	Glu	Gln	His	Ala	Ala		
	3570						3575						3580				
Pro	Asp	Asp	His	Pro	Asp	Asp	Pro	Ala	Asp	Gly	Pro	Glu	Asp	Asp	Gly		
3585					3590						3595				3600		
Ser	Asp	Asp	Leu	Met	Asp	Ala	Leu	Asp	Asp	Met	Ser	Ile	Asp	Asp	Leu		
			3605							3610					3615		
Ile	Arg	Ile	Ala	His	Asp	Glu	Arg	Pro	Arg	Gly	Asn						
3620	3625																

<210> 11
 <211> 8301
 <212> PRT
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

5

<400> 11

ES 2 556 211 T3

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

ES 2 556 211 T3

Leu Val Ala Leu His Trp Ala Val Arg Ser Leu Arg Ala Gly Glu Cys
 210 215 220
 Ser Met Ala Leu Ala Gly Gly Val Thr Val Met Ala Gly Pro Gly Val
 225 230 235 240
 Phe Val Glu Phe Ser Arg Glu Arg Gly Leu Ala Pro Asp Gly Arg Cys
 245 250 255
 Lys Ala Tyr Ser Ala Asp Ala Asp Gly Thr Gly Trp Gly Glu Gly Val
 260 265 270
 Gly Val Leu Leu Leu Glu Arg Leu Ser Asp Ala Val Arg Asn Gly His
 275 280 285
 Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Ser Asp Gly Ala
 290 295 300
 Ser Ser Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Glu Val Ile
 305 310 315 320
 Arg Gln Ala Leu Ala Asp Ala Arg Leu Thr Pro Ser Asp Val Asp Val
 325 330 335
 Val Glu Gly His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala
 340 345 350
 Gln Ala Leu Leu Ala Ala Tyr Gly Gln Asp Arg Pro Arg Pro Leu Leu
 355 360 365
 Leu Gly Ser Val Lys Ser Asn Leu Gly His Thr Gln Ala Ala Ala Gly
 370 375 380
 Val Ala Gly Val Ile Lys Met Val Glu Ala Met Arg Arg Gly Ile Ala
 385 390 395 400
 Pro Arg Thr Leu His Ala Thr Glu Pro Thr Pro Gln Val Asp Trp Ser
 405 410 415
 Arg Gly Ala Val Glu Leu Leu Thr Asp Asn Arg Ser Trp Pro Asp Thr
 420 425 430
 Gly Ala Pro Arg Arg Ser Ala Val Ser Ala Phe Gly Val Ser Gly Thr
 435 440 445
 Asn Ala His Ile Val Leu Glu Gln Ala Pro Glu Pro Ala Asp Thr Asp
 450 455 460
 Thr Thr Pro Arg Thr Ala His Pro Val Val Pro Trp Leu Leu Ser Ala
 465 470 475 480
 His Thr Pro Ala Ala Leu Arg Ala Gln Ala Glu Arg Leu Ser Ala Gly
 485 490 495
 Leu Pro Asp Asp Ala Asp Pro Leu Asp Val Ala Ala Ala Leu Ala Thr
 500 505 510
 Thr Arg Ala Ala Leu Pro Leu Arg Ala Ala Val Leu Gly Ala Asp Ala
 515 520 525
 Thr Gln Leu Arg Ala Gly Leu Ala Ala Leu Ala Ala Gly Ala Pro Ala
 530 535 540
 Ala Leu Ser Gly Glu Ala Arg Ser Asp Arg Thr Thr Ala Phe Leu Tyr
 545 550 555 560
 Thr Gly Gln Gly Ala Gln Arg Ala Gly Met Gly Glu Glu Leu Ala Ala
 565 570 575
 Ala Tyr Pro Ala Phe Ala Ala Ala Trp Thr Glu Val Cys Ala Glu Phe
 580 585 590
 Asp Thr Val Leu Pro Arg Pro Leu Arg Gln Val Ile Thr Glu Gly Gly
 595 600 605
 Pro Asp Leu Asp Arg Thr Leu Tyr Ala Gln Ala Ala Val Phe Ala Phe
 610 615 620
 Glu Thr Ala Leu Thr Ala Leu Leu Gly Ser Trp Gly Ile Arg Pro Asp
 625 630 635 640
 Leu Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala His Thr Ala
 645 650 655
 Gly Val Leu Ser Leu Arg His Ala Val Val Val Val Ala Ala Arg Gly
 660 665 670
 Arg Leu Met Glu Ala Leu Pro Glu Gly Gly Ala Met Val Ala Val Gln
 675 680 685

Ala	Ser	Glu	Asp	Glu	Ile	Glu	Leu	Pro	Glu	Gly	Val	Ala	Leu	Ala	Ala
690						695					700				
Val	Asn	Gly	Pro	Ser	Ser	Val	Val	Leu	Ser	Gly	Asp	Glu	Ala	Ala	Val
705					710					715					720
Leu	Ala	Thr	Ala	Ala	His	Trp	Ala	Glu	Arg	Gly	Cys	Arg	Thr	Lys	Arg
				725					730						735
Leu	Thr	Val	Ser	His	Ala	Phe	His	Ser	His	Arg	Met	Asp	Pro	Val	Leu
			740					745					750		
Asn	Gly	Phe	Arg	Arg	Val	Leu	Gly	Ala	Val	Thr	Leu	Asn	Thr	Pro	Arg
		755					760					765			
Ile	Pro	Phe	Val	Ser	Thr	Val	Thr	Gly	Ala	Pro	Val	Glu	Ala	Gly	Leu
770						775						780			
Arg	Asp	Pro	Glu	Tyr	Trp	Leu	Arg	Asn	Val	Arg	Asp	Thr	Val	Arg	Phe
785					790					795					800
Ala	Asp	Gly	Val	Arg	Thr	Leu	Ala	Asp	Glu	Gly	Ala	Asp	Thr	Phe	Val
				805					810					815	
Glu	Val	Gly	Pro	Asp	Ala	Val	Leu	Gly	Ala	Leu	Val	Ala	Asp	Ala	Leu
			820					825					830		
Pro	Asp	Asp	Pro	Asp	Arg	Pro	Glu	Ile	Leu	Ala	Val	Pro	Thr	Ala	Arg
		835					840					845			
Ala	Asp	Arg	Pro	Glu	Pro	Glu	Thr	Leu	Val	Gly	Ala	Leu	Ala	Arg	Ile
850						855					860				
His	Ala	His	Gly	Ala	His	Val	Asp	Trp	Ala	Ala	Phe	Phe	Gly	Ser	Gly
865					870					875					880
Ala	Arg	Arg	Val	Asp	Leu	Pro	Thr	Tyr	Ala	Phe	Gln	His	Gln	His	Tyr
				885					890						895
Trp	Leu	Ala	Pro	Ala	Pro	Ala	Asp	Asp	Leu	Pro	Ala	Ala	Gly	Leu	Gly
			900					905					910		
Thr	Val	Gly	His	Pro	Leu	Leu	Arg	Ala	Ala	Val	Glu	Leu	Pro	Ala	Asp
		915						920				925			
Ala	Thr	Ala	Asp	His	Ala	Ala	Arg	Pro	Gly	Ala	Ala	Gly	Ala	Val	Val
930						935					940				
Phe	Thr	Gly	Thr	Leu	Ser	Ala	His	Thr	His	Pro	Trp	Leu	Ala	Asp	His
945					950					955					960
Ser	Val	Leu	Gly	Thr	Pro	Val	Leu	Pro	Gly	Thr	Ala	Leu	Ala	Glu	Leu
				965					970					975	
Ala	Ala	Ala	Ala	Gly	Asp	Arg	Leu	Gly	Cys	Ala	Thr	Val	Ala	Glu	Leu
			980					985					990		
Val	Leu	Thr	Ala	Pro	Leu	Ala	Val	Pro	Ala	Ser	Gly	Ala	Val	Arg	Leu
		995					1000					1005			
Arg	Val	His	Val	Asp	Ala	Pro	Asp	Arg	Asp	Gly	His	Arg	Ala	Val	Thr
1010						1015					1020				
Val	Asp	Ser	Arg	Pro	Asp	Asp	Gln	Leu	Ala	Asp	Gly	Ala	Ala	Trp	Thr
1025					1030					1035					1040
Arg	His	Ala	Thr	Gly	Arg	Leu	Ala	Pro	Thr	Gly	Pro	Glu	Pro	Ala	Gln
				1045						1050					1055
Ala	Pro	Thr	Ala	Trp	Pro	Pro	Ala	Asp	Ala	Glu	Pro	Leu	Pro	Val	Asp
				1060					1065					1070	
Gly	Leu	Tyr	Asp	Arg	Leu	Thr	Ala	Ala	Gly	Phe	Gly	Tyr	Gly	Pro	Ala
		1075						1080					1085		
Phe	Arg	Gly	Val	Arg	Ala	Ala	Trp	Arg	Leu	Gly	Thr	Asp	Leu	Leu	Ala
1090						1095					1100				
Asp	Val	Glu	Leu	Pro	Ala	Asp	Thr	Asp	Glu	Ser	Gly	Phe	Leu	Leu	His
1105					1110						1115				1120
Pro	Ala	Leu	Phe	Asp	Ala	Ala	Leu	His	Ala	Leu	Gly	Leu	Gly	Gly	Leu
				1125						1130					1135
Val	Glu	His	Gly	Gly	Leu	Pro	Phe	Thr	Trp	Thr	Gly	Val	Arg	Leu	His
			1140						1145					1150	
Ala	Thr	Gly	Ala	Arg	Ser	Leu	Arg	Val	Arg	Leu	Thr	Pro	Thr	Gly	Pro
		1155						1160						1165	

ES 2 556 211 T3

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

ES 2 556 211 T3

Val Ala Pro Asp Arg Ser Phe Leu Asp Leu Gly Phe Asp Ser Leu Ala
 1650 1655 1660
 Ala Val Glu Leu Arg Asn Arg Leu Thr Ala Ala Thr Gly Leu Arg Leu
 1665 1670 1675 1680
 Ala Ala Thr Ile Thr Phe Asp His Pro Thr Ser Ala Ala Leu Ala Arg
 1685 1690 1695
 His Leu Leu Asp Ala Ala Leu Asp Ala Gly Pro Ser Ala Thr Ser Gly
 1700 1705 1710
 Ala His Leu Pro Ala Ala Thr Ala Val Arg Ala Asp Glu Pro Ile Ala
 1715 1720 1725
 Ile Val Gly Met Ala Cys Arg Tyr Pro Gly Gly Val Thr Ser Pro Glu
 1730 1735 1740
 Asp Leu Trp Arg Leu Val Leu Ala Gly Ala Asp Ala Val Thr Gly Phe
 1745 1750 1755 1760
 Pro Glu Asp Arg Gly Trp Asp Leu Thr Ser Val Tyr Asp Pro Asp Gln
 1765 1770 1775
 Ser Arg Thr Gly Thr Ser Tyr Thr Arg Glu Gly Gly Phe Leu Thr Gly
 1780 1785 1790
 Ala Ala Asp Phe Asp Ala Glu Phe Phe Gly Ile Ser Pro Arg Glu Ala
 1795 1800 1805
 Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr Ser Trp Glu
 1810 1815 1820
 Ala Ile Glu Arg Ala Gly Ile Asp Pro Thr Ala Leu Arg Gly Ser Ala
 1825 1830 1835 1840
 Thr Gly Val Phe Ala Gly Leu Met Tyr His Asp Tyr Gly Ser Gly Thr
 1845 1850 1855
 Gly Thr Leu Pro Glu Gly Val Glu Gly Tyr Leu Gly Leu Gly Thr Ala
 1860 1865 1870
 Gly Ser Val Leu Ser Gly Arg Val Ala Tyr Thr Leu Gly Leu Glu Gly
 1875 1880 1885
 Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu
 1890 1895 1900
 His Trp Ala Ile Gln Ala Leu Arg Ser Gly Glu Cys Ser Met Ala Leu
 1905 1910 1915 1920
 Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Thr Phe Val Glu Phe
 1925 1930 1935
 Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala
 1940 1945 1950
 Ala Ala Ala Asp Gly Val Gly Trp Ser Glu Gly Val Gly Met Leu Leu
 1955 1960 1965
 Val Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Gln Val Leu Ala
 1970 1975 1980
 Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu
 1985 1990 1995 2000
 Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu
 2005 2010 2015
 Gln Gly Ala Gly Leu Thr Thr Ala Asp Val Asp Ala Val Glu Ala His
 2020 2025 2030
 Gly Thr Gly Thr Lys Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu
 2035 2040 2045
 Ala Thr Tyr Gly Gln Glu Arg Asp Glu Pro Leu Trp Leu Gly Ser Val
 2050 2055 2060
 Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val
 2065 2070 2075 2080
 Ile Lys Met Val Gln Ala Met Arg His Gly Thr Leu Pro Arg Thr Leu
 2085 2090 2095
 His Val Asp Glu Ala Ser Pro His Val Asp Trp Thr Ala Gly Ala Val
 2100 2105 2110
 Glu Leu Leu Thr Glu Glu Arg Glu Trp Thr Arg Thr Arg Arg Pro Arg
 2115 2120 2125

ES 2 556 211 T3

Arg Ala Ala Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val
 2130 2135 2140
 Val Leu Glu Glu Ala Pro Glu Glu Thr Ala Pro Pro Val Ser Ala Glu
 2145 2150 2155 2160
 Leu Pro Leu Val Pro Leu Leu Leu Ser Gly His Thr Ser Thr Ala Leu
 2165 2170 2175
 Ala Ala Gln Ala Arg Arg Leu His Asp His Leu Ala Glu Ser Gly Thr
 2180 2185 2190
 Pro Ala Leu Glu Thr Val Gly Arg Ser Leu Ala Ala Ser Arg Ala Leu
 2195 2200 2205
 Leu Pro His Arg Ala Val Val Val Ala Ala Asp Val Ala Gly Ala Arg
 2210 2215 2220
 Ala Ser Leu Ala Ala Leu Ala Ala Gly Glu Pro Val Glu Gly Val Val
 2225 2230 2235 2240
 Ser Gly Ala Ala Gly Leu Pro Gly Glu Gly Arg Val Val Phe Val Phe
 2245 2250 2255
 Pro Gly Gln Gly Ala Gln Trp Val Gly Met Ala Val Asp Leu Leu Asp
 2260 2265 2270
 Ser Ser Pro Val Phe Ala Gly Arg Met Ala Glu Cys Ala Val Ala Leu
 2275 2280 2285
 Glu Pro Phe Val Glu Trp Ser Leu Arg Gly Val Leu Gly Asp Pro Val
 2290 2295 2300
 Ala Leu Glu Arg Val Asp Val Val Gln Pro Val Leu Trp Ala Val Met
 2305 2310 2315 2320
 Val Ser Leu Ala Glu Val Trp Arg Ser Tyr Gly Val Val Pro Ser Ala
 2325 2330 2335
 Val Val Gly His Ser Gln Gly Glu Ile Ala Ala Ala Cys Val Ala Gly
 2340 2345 2350
 Val Leu Ser Leu Ala Asp Gly Ala Arg Val Val Ala Leu Arg Ser Arg
 2355 2360 2365
 Ala Leu Thr Ala Leu Ala Gly Ser Gly Gly Met Val Ser Val Ala Ala
 2370 2375 2380
 Gly Pro Ser Gly Val Glu Leu Leu Val Gly Trp Ala Gly Arg Leu
 2385 2390 2395 2400
 Ala Val Ala Ala Val Asn Gly Pro Glu Ser Val Val Val Ala Gly Glu
 2405 2410 2415
 Gly Val Ala Leu Glu Glu Phe Leu Ala His Cys Gly Gly Arg Gly Val
 2420 2425 2430
 Arg Ala Arg Arg Ile Ala Val Asp Tyr Ala Ser His Ser Val Leu Val
 2435 2440 2445
 Glu Pro Val Arg Glu Ala Leu Leu Ala Asp Leu Glu Gly Val Arg Pro
 2450 2455 2460
 Gly Glu Gly Thr Val Pro Leu Phe Ser Thr Val Thr Gly Glu Trp Ala
 2465 2470 2475 2480
 Asp Gly Thr Ala Leu Asp Ala Gly Tyr Trp Tyr Arg Asn Leu Arg Glu
 2485 2490 2495
 Pro Val Gly Phe Glu Pro Ala Val Arg Gly Leu Leu Asp Ser Gly His
 2500 2505 2510
 Ala Val Phe Val Glu Ile Ser Pro His Pro Val Leu Thr Ala Ala Val
 2515 2520 2525
 Gln Glu Thr Ala Asp Ala Thr Glu Arg Thr Ala Val Val Val Gly Thr
 2530 2535 2540
 Leu Arg Arg Asp His Glu Gly Gln Arg Gln Leu Leu Thr His Leu Gly
 2545 2550 2555 2560
 Val Leu His Thr Thr Gly Ala Asp Ile Asp Trp Thr Gly Cys Phe Thr
 2565 2570 2575
 Gly Val Thr Gly Arg Ala Asp Leu Pro Thr Tyr Ala Phe Gln His Thr
 2580 2585 2590
 Arg Tyr Trp Leu Thr Pro Ser Gly Pro Ser Ala Gly Glu Leu Ala Gly
 2595 2600 2605

Ala Gly Leu Thr Ala Ala Gly His Pro Leu Leu Gly Ala Ala Val Asp
 2610 2615 2620
 Leu Ala Glu Asp Gly Gly Leu Val Leu Thr Gly Arg Leu Ala Ala Asp
 2625 2630 2635 2640
 Pro Ala Ala Trp Thr Ala Asp His Val Val Leu Gly Thr Thr Leu Leu
 2645 2650 2655
 Pro Gly Ala Ala Leu Ala Glu Leu Ala Leu Ala Ala Gly Asp Gly Val
 2660 2665 2670
 Gly Cys Gly Thr Leu Asp Glu Leu Val Leu Gly Ala Pro Leu Ala Leu
 2675 2680 2685
 Pro Glu Arg Gly Ala Leu His Leu Gln Val Arg Val Gly Arg Ala Glu
 2690 2695 2700
 Ala Asp His Arg Arg Thr Val Ser Val His Ala Arg Pro Glu Asp Gly
 2705 2710 2715 2720
 Asp Ala Pro Trp Thr Arg His Ala Glu Gly Val Leu Val Pro Gly Asp
 2725 2730 2735
 Thr Ala Ala Gly Thr Pro Leu Thr Glu Trp Pro Pro Ala Asp Ala Glu
 2740 2745 2750
 Pro Val Asp Val Ser Ala Leu Tyr Asp Ser Leu Ala Asp Arg Gly Leu
 2755 2760 2765
 Asp Tyr Gly Pro Val Phe Arg Gly Val Arg Ala Ala Trp Arg His Gly
 2770 2775 2780
 Asp Asp Ile Leu Ala Glu Val Glu Leu Pro Ala Glu Ala Asp Glu Ser
 2785 2790 2795 2800
 Gly Phe Leu Leu His Pro Ala Leu Leu Asp Ala Ala Leu His Pro Ile
 2805 2810 2815
 Gly Leu Gly Gly Leu Val Gly Asp Gly Gly Leu Pro Phe Ala Trp His
 2820 2825 2830
 Gly Leu Arg Val His Ala Val Gly Ala Arg Ala Ala Arg Val Arg Leu
 2835 2840 2845
 Thr Pro Leu Gly Asp Glu Thr Val Ala Val Asp Leu Ala Asp Gly Ala
 2850 2855 2860
 Gly Ala Pro Leu Ala Ala Ile Val Ser Leu Arg Leu Arg Ala Val Thr
 2865 2870 2875 2880
 Ala Ala Gln Val Ala Ala Ala Arg Thr Leu Thr Asp Pro Ala Leu His
 2885 2890 2895
 Leu Asp Trp Leu Pro Val Ala Gly Ala Ala Ser Ala Pro Ser Asp Asp
 2900 2905 2910
 Thr Ala Val Glu Leu Leu Arg Leu Asp Asp Ala Pro His Gly Thr Asp
 2915 2920 2925
 Ala Gly Pro Gly Gly Val Arg Gln Ala Val Arg Thr Ala Leu Thr Ala
 2930 2935 2940
 Ile Gln Glu Arg Ile Ala Asp Asp Arg Ala Glu Arg Leu Val Val Val
 2945 2950 2955 2960
 Thr Arg Asp Thr Asp Leu Ala Gly Ala Ala Val Gly Gly Leu Leu Arg
 2965 2970 2975
 Ser Ala Gln Ala Glu His Pro Gly Arg Phe Gly His Val Val Leu Asp
 2980 2985 2990
 Gly His Pro Asp Ser Glu Arg Ala Leu Pro Thr Ala Thr Ala Leu Thr
 2995 3000 3005
 Asp Glu Pro Trp Val Ala Val Arg Ala Gly Glu Thr Tyr Val Pro Arg
 3010 3015 3020
 Leu Ala Arg Ser Gly Ala Arg Pro Gly Gly Asp Val Pro Phe Gly Ala
 3025 3030 3035 3040
 Asp Asp Val Val Leu Val Thr Gly Gly Thr Gly Val Leu Gly Ala Leu
 3045 3050 3055
 Val Ala Arg His Leu Val Thr Glu His Gly Val Arg Arg Leu Val Leu
 3060 3065 3070
 Ala Ser Arg Arg Gly Gly Ala Gly Glu Leu Val Ala Glu Leu Gly Glu
 3075 3080 3085

ES 2 556 211 T3

Leu Gly Ala Ala Val Asp Val Val Ala Cys Asp Val Ser Asp Arg Glu
 3090 3095 3100
 Ala Leu Glu Arg Leu Val Ala Ala His Pro Leu Thr Gly Val Val His
 3105 3110 3115 3120
 Ala Ala Gly Val Leu Asp Asp Gly Thr Val Glu Ser Leu Thr Pro Asp
 3125 3130 3135
 Arg Val Asp Gly Val Leu Arg Ala Lys Val Asp Gly Ala Trp His Leu
 3140 3145 3150
 His Glu Leu Thr Ala Ala Leu Asp Leu Arg Ala Phe Val Leu Phe Ser
 3155 3160 3165
 Ser Leu Ala Gly Val Val Gly Ser Ala Gly Gln Gly Gly Tyr Ala Ala
 3170 3175 3180
 Ala Asn Ala Ala Leu Asp Ala Leu Ala Glu Tyr Arg Arg Ala Glu Gly
 3185 3190 3195 3200
 Leu Pro Ala Thr Ser Ile Ala Trp Gly Leu Trp Ala Pro Ala Ser Ala
 3205 3210 3215
 Met Thr Ser Gly Ala Asp Thr Ala Arg Leu Ala Arg Ser Gly Ile Leu
 3220 3225 3230
 Pro Leu Pro Ala Asp Arg Ala Leu Glu Leu Phe Asp Thr Ala Cys Ala
 3235 3240 3245
 Ala Glu Thr Pro Leu Thr Val Ala Ala Arg Leu Asp Leu Thr Ala Phe
 3250 3255 3260
 Arg Ala Gln Gly Thr Arg Met Pro Val Val Leu Arg Ser Leu Ala Gly
 3265 3270 3275 3280
 Pro Ala Ala Arg Arg Thr Ala Arg Ala Gly Asp Ala Gly Thr Leu Arg
 3285 3290 3295
 Asp Arg Leu Ala Ala Gln Thr Pro Ala Glu Arg Thr Arg Thr Val Leu
 3300 3305 3310
 Asp Leu Val Arg Gly Gln Ala Ala Ala Val Leu Gly His Glu Ser Ala
 3315 3320 3325
 Ala Ala Ile Ala Glu Asp Arg Ala Phe Leu Glu Leu Gly Phe Asp Ser
 3330 3335 3340
 Leu Thr Ala Val Asp Leu Arg Asn Arg Leu Gly Thr Val Thr Gly Leu
 3345 3350 3355 3360
 Arg Leu Pro Thr Thr Thr Val Phe Asp His Pro Asn Pro Ala Ala Leu
 3365 3370 3375
 Thr Arg His Ile Leu Ala Glu Leu Leu Gly Ala Thr Ala Gly Ser Ala
 3380 3385 3390
 Gln Ala Ala Pro Thr Ala Val Arg Thr Asp Glu Pro Ile Ala Ile Val
 3395 3400 3405
 Gly Met Ala Cys Arg Tyr Pro Gly Gly Val Ala Ser Pro Glu Asp Leu
 3410 3415 3420
 Trp Arg Val Val Ala Glu Gly Arg Asp Val Ile Ser Pro Phe Pro Glu
 3425 3430 3435 3440
 Asp Arg Gly Trp Asp Leu Gly Ala Leu Tyr His Ala Asp Pro Asp His
 3445 3450 3455
 Thr Gly Thr Ser Tyr Ala Arg Glu Gly Phe Leu His Asp Ala Ala
 3460 3465 3470
 Gly Phe Asp Ala Glu Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala
 3475 3480 3485
 Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ala Ser Trp Glu Ala Ile
 3490 3495 3500
 Glu Arg Ala Gly Ile Asp Pro Thr Ala Leu Arg Gly Ser Ala Thr Gly
 3505 3510 3515 3520
 Val Phe Ala Gly Leu Met Tyr His Asp Tyr Ala Ala Arg Leu Gly Thr
 3525 3530 3535
 Thr Pro Glu Gly Leu Glu Gly Tyr Leu Gly Met Gly Asn Ser Gly Ser
 3540 3545 3550
 Val Ala Ser Gly Arg Ile Ser Tyr Thr Leu Gly Leu Glu Gly Pro Ala
 3555 3560 3565

Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Trp
 3570 3575 3580
 Ala Ile Gln Ala Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly
 3585 3590 3595 3600
 Gly Val Ser Val Met Ala Thr Pro Gly Thr Phe Val Glu Phe Ser Arg
 3605 3610 3615
 Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala
 3620 3625 3630
 Ala Asp Gly Ala Ser Trp Ser Glu Gly Val Gly Met Leu Leu Val Glu
 3635 3640 3645
 Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val
 3650 3655 3660
 Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala
 3665 3670 3675 3680
 Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Gly
 3685 3690 3695
 Ala Gly Leu Thr Ser Ala Asp Val Asp Ala Val Glu Ala His Gly Thr
 3700 3705 3710
 Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr
 3715 3720 3725
 Tyr Gly Gln Glu Arg Asp Glu Pro Leu Trp Leu Gly Ser Val Lys Ser
 3730 3735 3740
 Asn Ile Gly His Ala Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys
 3745 3750 3755 3760
 Met Val Gln Ala Met Arg Asn Gly Val Leu Pro Arg Thr Leu His Val
 3765 3770 3775
 Asp Glu Pro Ser Pro His Val Asp Trp Thr Val Gly Ala Val Glu Leu
 3780 3785 3790
 Leu Thr Gly Glu Gln Glu Trp Pro Arg Gln Asp Arg Pro Arg Arg Ala
 3795 3800 3805
 Gly Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Val Leu
 3810 3815 3820
 Glu Glu Ala Pro Glu Glu Thr Glu Gly Thr Ala Pro Ala Ala Leu Pro
 3825 3830 3835 3840
 Ala Val Pro Trp Val Val Ser Ala Arg Ser Glu Thr Ala Leu Arg Ala
 3845 3850 3855
 Gln Ala Ala Arg Leu Ala Asp Trp Leu His Gly Asp Thr Asp Val Leu
 3860 3865 3870
 Gly Thr Ala Tyr Ser Leu Ala Thr Gly Arg Ala Ala Leu Pro His Arg
 3875 3880 3885
 Ala Val Val Val Gly Thr Asp Arg Ala Glu Leu Ser Asp Gly Leu Ala
 3890 3895 3900
 Ala Leu Ala Ala Gly Arg Ala Ala Ala His Val Glu Ser Gly Arg Ala
 3905 3910 3915 3920
 Arg Asp Asn Arg Val Thr Ala Phe Val Phe Ala Gly Gln Gly Ala Gln
 3925 3930 3935
 Arg Ala Gly Met Gly Ala Glu Leu Ala Ala Tyr Pro Val Phe Ala
 3940 3945 3950
 Gln Val Phe Ala Gln Val Cys Ala Ala Phe Asp Gly Val Leu Glu Arg
 3955 3960 3965
 Pro Leu Gly Glu Val Ile Ala Glu Gly Gly Pro Glu Leu Asp Arg Thr
 3970 3975 3980
 Val Tyr Ala Gln Ala Gly Leu Phe Ala Phe Glu Val Ala Leu Phe Arg
 3985 3990 3995 4000
 Leu Leu Glu Ser Trp Gly Val Ala Pro Asp Val Val Leu Gly His Ser
 4005 4010 4015
 Val Gly Glu Leu Ala Ala Ala Cys Val Ala Gly Val Trp Ser Leu Glu
 4020 4025 4030
 Asp Ala Val Arg Val Val Ala Ala Arg Gly Arg Leu Met Gln Ala Leu
 4035 4040 4045

ES 2 556 211 T3

Pro Gln Gly Gly Ala Met Val Ala Leu Glu Val Ser Ala Gly Glu Leu
 4050 4055 4060
 Glu Leu Pro Glu Gly Val Glu Leu Ala Ala Val Asn Gly Pro Ser Ser
 4065 4070 4075 4080
 Val Val Leu Ser Gly Glu Glu Glu Ala Val Leu Ala Glu Ala Ala Arg
 4085 4090 4095
 Trp Pro Asp Arg Arg Ala Lys Arg Leu Arg Val Ser His Ala Phe His
 4100 4105 4110
 Ser Arg Arg Met Asp Pro Met Leu Glu Asp Phe Arg Arg Val Leu Glu
 4115 4120 4125
 Ser Val Ala Phe His Ala Pro Glu Leu Val Phe Val Ser Thr Val Thr
 4130 4135 4140
 Gly Ala Val Val Thr Asp Glu Leu Cys Asp Pro Gly Tyr Trp Val Arg
 4145 4150 4155 4160
 Asn Val Arg Glu Thr Val Arg Phe Ala Asp Ala Val Val Ala Ala Glu
 4165 4170 4175
 Ala Gly Val Phe Val Glu Leu Ala Pro Asp Ala Val Leu Ser Gly Leu
 4180 4185 4190
 Val Gly Glu Ser Val Glu Gly Val Leu Cys Val Pro Ala Gln Arg Ala
 4195 4200 4205
 Gly Arg Pro Ala Ala Arg Val Leu Val Ser Ala Leu Gly Thr Leu His
 4210 4215 4220
 Thr His Gly Val Asp Val Ala Trp Gly Arg Phe Phe Ala Gly Ser Gly
 4225 4230 4235 4240
 Ala Arg Arg Val Asp Leu Pro Thr Tyr Ala Phe Gln Arg Glu Arg Tyr
 4245 4250 4255
 Trp Leu Asp Ala Pro Ala Pro Gln Thr Gly Gly Ala Ser Asp Asp Ala
 4260 4265 4270
 Phe Trp Ala Ala Val Gln Ser Gly Asp Leu Ala Gly Leu Leu Gly Val
 4275 4280 4285
 Ala Glu Asp Ala Gly Leu Asp Ala Val Leu Pro Thr Leu Ala Ser Trp
 4290 4295 4300
 His Asp Arg Glu Arg Ala Asp Ala Val Val Asp Gly Trp Arg His Lys
 4305 4310 4315 4320
 Val Arg Trp Thr Pro Leu Pro Glu Ala Gly Gly Ala Val Leu Ser Gly
 4325 4330 4335
 Arg Trp Leu Leu Val Gly Pro Glu Gly Asp Glu Ala Leu Val Ala Asp
 4340 4345 4350
 Val Ala Thr Ala Leu Arg Glu His Gly Ala Glu Val Thr His Leu Ala
 4355 4360 4365
 Leu Pro Arg Asp Ala Asp Arg Glu Thr Thr Ala Glu Leu Leu Arg Gly
 4370 4375 4380
 Thr Asp Asp Thr Gly Leu Thr Ala Val Leu Ser Leu Leu Ala Arg Ala
 4385 4390 4395 4400
 Asp Arg Pro Val His Ala Thr Leu Ala Leu Val Gln Ala Leu Gly Asp
 4405 4410 4415
 Ala Glu Val Thr Val Pro Leu Trp Cys Ala Thr Ser Glu Ser Val Ala
 4420 4425 4430
 Val Ala Ser Thr Asp Val Val Pro Glu Ala Ala Val Asp Ala Ala Gly
 4435 4440 4445
 Leu Trp Gly Leu Gly Arg Val Val Arg Leu Glu Ala Pro Asp Arg Trp
 4450 4455 4460
 Gly Gly Leu Val Asp Leu Pro Gly Ala Leu Asp Ala Arg Ala Arg Arg
 4465 4470 4475 4480
 Arg Leu Ala Ala Val Leu Ala Gly Ala Glu Asp Glu Cys Ala Val Arg
 4485 4490 4495
 Glu Asn Gly Ala Phe Ala Ala Arg Leu Val Arg Ala Ala Asp Ala Pro
 4500 4505 4510
 Arg Arg Glu Trp Arg Pro Gln Gly Thr Val Leu Val Thr Gly Gly Thr
 4515 4520 4525

ES 2 556 211 T3

Gly Ala Leu Gly Ala Arg Val Ala Gln Arg Leu Ala Glu Arg Gly Ala
 4530 4535 4540
 Arg His Val Leu Leu Val Ser Arg Arg Gly Pro Ala Ala Asp Gly Ala
 4545 4550 4555 4560
 Ala Glu Leu Val Arg Ala Ile Glu Ala Ala Gly Ala Thr Ala Thr Val
 4565 4570 4575
 Ala Ala Cys Asp Val Ala Asp Pro Glu Ala Val Ala Ala Leu Leu Glu
 4580 4585 4590
 Arg Ile Pro Ala Asp Ala Pro Leu Thr Ala Val Val His Thr Ala Ser
 4595 4600 4605
 Val Leu Ala Asp Ala Pro Leu Asp Thr Leu Thr Pro Asp Arg Ile Thr
 4610 4615 4620
 Ala Val Leu Arg Ala Lys Ala Asp Ala Ala Arg Ile Leu His Thr Ala
 4625 4630 4635 4640
 Thr Ala Ser Leu Asp Leu Asp Ala Phe Val Leu Phe Ser Ser Leu Ala
 4645 4650 4655
 Gly Thr Leu Gly Asn Pro Gly Gln Ala Ala Tyr Ala Ala Ala Asn Ala
 4660 4665 4670
 Val Leu Asp Thr Leu Ala Ala His Arg Arg Ala Leu Gly Leu Pro Gly
 4675 4680 4685
 Thr Ala Val Ala Trp Gly Pro Trp Ala Gly Gly Gly Met Leu Asp Asp
 4690 4695 4700
 Thr Val Ala Glu Arg Leu Arg Arg Ala Gly Val Ile Pro Leu Asp Pro
 4705 4710 4715 4720
 Glu His Ala Leu Val Ala Leu Asp Arg Ala Val Ala Ala Ala Asp Ala
 4725 4730 4735
 His Ser Val Ile Ala Asp Ala Asp Trp Thr Gly Leu Thr Ala Gly Ser
 4740 4745 4750
 Met Leu Ala Glu Leu Ser Gly Pro Ala Pro Val Glu Ala Pro Thr Ala
 4755 4760 4765
 Leu Thr Gly Pro Asp Arg Glu Arg Ala Ala Leu Ala Leu Val Arg Ser
 4770 4775 4780
 Cys Ala Ala Ala Val Leu Gly Arg Ser Ala Ala Val Asp Val Glu Pro
 4785 4790 4795 4800
 Asp Thr Ala Phe Arg Glu Leu Gly Phe Asp Ser Met Ala Ala Val Gln
 4805 4810 4815
 Leu Arg Asn Arg Leu Asn Ala Ala Thr Gly Val Val Leu Thr Ala Thr
 4820 4825 4830
 Ala Val Phe Asp His Pro Ser Ala Arg Ala Leu Ala Thr His Leu Leu
 4835 4840 4845
 Ala Leu Ala Thr Gly Glu Gly Ala Gly Ser Ala Gly Ala Thr Glu Leu
 4850 4855 4860
 Pro Gly Thr Ala Val His Thr Asp Glu Pro Ile Ala Ile Val Gly Met
 4865 4870 4875 4880
 Ala Cys Arg Phe Pro Gly Asp Val Thr Ser Pro Glu Asp Leu Trp Arg
 4885 4890 4895
 Leu Leu Ala Asp Gly Val Asp Ala Val Gly Pro Ile Pro Ala Asp Arg
 4900 4905 4910
 Gln Trp Gly Pro Thr Asp Ala Tyr Ala Glu Gly Gly Phe Leu Arg Gly
 4915 4920 4925
 Ala Gly Glu Phe Asp Ala Asp Phe Phe Gly Ile Ser Pro Arg Glu Ala
 4930 4935 4940
 Leu Val Met Asp Pro Gln Gln Arg Leu Ala Leu Glu Thr Gly Trp Glu
 4945 4950 4955 4960
 Val Phe Glu Arg Ala Gly Ile Asn Pro His Thr Val Arg Gly Thr Ser
 4965 4970 4975
 Val Gly Val Phe Leu Gly Thr Asn Gly Gln Asp Tyr Val Ser Leu Leu
 4980 4985 4990
 Ala Gly Ala Thr Glu Ala His Ala Gly His Ile Gly Thr Gly Asn Ser
 4995 5000 5005

ES 2 556 211 T3

Ala Ser Val Leu Ser Gly Arg Ile Ala Tyr Val Leu Gly Leu Glu Gly
5010 5015 5020
Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu
5025 5030 5035 5040
His Trp Ala Ile Gln Ala Leu Arg Ser Gly Glu Cys Ser Met Ala Leu
5045 5050 5055
Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Ala Phe Ala Glu Phe
5060 5065 5070
Ser His Gln Arg Gly Leu Ala Glu Asp Gly Arg Cys Lys Ser Phe Ala
5075 5080 5085
Ala Ser Ala Asp Gly Thr Gly Trp Gly Glu Gly Val Gly Met Leu Leu
5090 5095 5100
Val Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala
5105 5110 5115 5120
Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu
5125 5130 5135
Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu
5140 5145 5150
Ala Gly Ala Gly Leu Thr Pro Val Asp Val Asp Ala Val Glu Ala His
5155 5160 5165
Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu
5170 5175 5180
Ala Thr Tyr Gly Gln Glu Arg Asp Glu Pro Leu Trp Leu Gly Ser Val
5185 5190 5195 5200
Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val
5205 5210 5215
Ile Lys Met Val Glu Ala Met Arg Asn Gly Thr Leu Pro Pro Thr Leu
5220 5225 5230
His Val Asp Glu Pro Ser Pro His Val Asp Trp Ser Ala Gly Ala Val
5235 5240 5245
Glu Leu Leu Thr Glu Ala Arg Glu Trp Lys Arg Ala Gly Arg Pro Arg
5250 5255 5260
Arg Ala Gly Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val
5265 5270 5275 5280
Val Leu Glu Glu Ala Pro Glu Glu Thr Glu Pro Ala Ala Ser Asp Glu
5285 5290 5295
Leu Pro Val Ala Pro Trp Leu Leu Ser Ala Arg Ser Glu Lys Ala Leu
5300 5305 5310
Leu Ala Gln Val Glu Arg Leu Arg Ala Tyr Val Thr Glu His Pro Glu
5315 5320 5325
Ala Arg Pro Ala Asp Ile Gly Leu Ser Leu Ala Thr Gly Arg Ala Ala
5330 5335 5340
Leu Ala His Arg Leu Ser Gly Ala Gly Glu Thr Thr Glu Glu Leu Leu
5345 5350 5355 5360
Ala Ala Leu Asp Thr Ala Leu Pro Ala Val Ala Arg Glu Thr Pro Thr
5365 5370 5375
Ala Phe Val Phe Ala Gly Gln Gly Ala Gln Arg Val Gly Met Gly Ala
5380 5385 5390
Glu Leu Ala Ala Val Tyr Pro Val Phe Ala Gln Val Phe Ala Gln Val
5395 5400 5405
Cys Ala Ala Phe Asp Gly Val Leu Glu Arg Pro Leu Gly Glu Val Val
5410 5415 5420
Ala Glu Gly Gly Pro Glu Leu Asp Arg Thr Val Tyr Ala Gln Ala Gly
5425 5430 5435 5440
Leu Phe Ala Phe Glu Val Ala Leu Phe Arg Leu Leu Glu Ser Trp Gly
5445 5450 5455
Val Ala Pro Asp Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala
5460 5465 5470
Ala Cys Val Ala Gly Val Trp Ser Leu Glu Asp Ala Val Arg Val Val
5475 5480 5485

ES 2 556 211 T3

Ala Ala Arg Gly Arg Leu Met Gln Ala Leu Pro Gln Gly Gly Ala Met
5490 5495 5500
Val Ala Leu Glu Val Ser Ala Gly Glu Leu Glu Leu Pro Glu Gly Val
5505 5510 5515 5520
Glu Leu Ala Ala Val Asn Gly Pro Thr Ser Val Val Leu Ser Gly Glu
5525 5530 5535
Glu Asp Ala Val Leu Ala Glu Ala Ala Arg Trp Pro Asp Arg Arg Ala
5540 5545 5550
Lys Arg Leu Arg Val Ser His Ala Phe His Ser Arg Arg Met Asp Pro
5555 5560 5565
Met Leu Glu Asp Phe Arg Arg Val Leu Glu Ser Val Thr Phe His Ala
5570 5575 5580
Pro Gln Leu Ala Phe Val Ser Thr Val Thr Gly Ala Ala Val Thr Asp
5585 5590 5595 5600
Glu Leu Cys Asp Pro Gly Tyr Trp Val Arg Asn Val Arg Glu Thr Val
5605 5610 5615
Arg Phe Ala Asp Ala Val Val Ala Ala Gly Ala Gly Val Phe Val Glu
5620 5625 5630
Leu Ala Pro Asp Ala Val Leu Ser Gly Leu Val Gly Glu Ser Val Glu
5635 5640 5645
Gly Val Leu Cys Val Pro Ala Gln Arg Ala Gly Arg Pro Ala Ala Arg
5650 5655 5660
Ala Leu Val Ser Ala Leu Gly Thr Leu His Thr His Gly Val Asp Val
5665 5670 5675 5680
Ala Trp Asp Arg Phe Phe Ala Gly Ser Gly Ala Arg Arg Val Asp Leu
5685 5690 5695
Pro Thr Tyr Ala Phe Gln Arg Glu Arg Tyr Trp Leu Ala Pro Pro Ala
5700 5705 5710
Ala Gly Pro Val Gly Leu Ala Gly Val Gly Leu Thr Ala Thr Gly His
5715 5720 5725
Pro Leu Leu Gly Val Ser Val Glu Leu Pro Gly Thr Asp Ala Val Ala
5730 5735 5740
Phe Thr Gly Ser Val Ser Leu Ser Thr His Pro Trp Leu Ala Asp His
5745 5750 5755 5760
Ala Ile Leu Gly Pro Ala Leu Leu Pro Gly Thr Ala Phe Leu Asp Leu
5765 5770 5775
Ala Leu Ala Ala Gly Glu His Val Gly Cys Pro Arg Val Asp Asp Leu
5780 5785 5790
Ala Leu His Ala Pro Leu Ala Leu Pro Ala His Gly Ser Val Arg Met
5795 5800 5805
Gln Val Arg Val Glu Gly Thr Asp Pro Asp Gly Ser Arg Gln Val Gly
5810 5815 5820
Ile Tyr Ser Arg Pro Glu Asp Asp Glu Asp Ala Trp Thr Gln His Ala
5825 5830 5835 5840
Thr Gly Val Leu Ala Pro Glu Ser Gly Pro Ala Ala Glu Ala Leu Thr
5845 5850 5855
Glu Trp Pro Pro Gln Gly Ala Glu Pro Val Ser Val Asp Gly Leu Tyr
5860 5865 5870
Asp Asp Leu Ala Ala Thr Gly Phe Ala Tyr Gly Pro Leu Phe Arg Gly
5875 5880 5885
Leu Arg Ala Ala Trp Val Arg Asp Gly Ala Val Tyr Ala Asp Val Ala
5890 5895 5900
Leu Pro Asp Glu Thr Ala Ser Val Ala Gly Tyr Ser Val His Pro Ala
5905 5910 5915 5920
Leu Leu Asp Ala Ala Leu His Ala Leu Gly Cys Ala Asn Leu Val Ala
5925 5930 5935
Asp Trp Ala Asp Gly Gln Leu Pro Phe Ala Trp Thr Gly Ala Arg Ile
5940 5945 5950
His Ala Val Gly Ala Arg Ala Leu Arg Val Arg Leu Arg Ala Glu Asn
5955 5960 5965

Gly Gly Ile Ala Leu Thr Ala Ala Asp Ala Gly Gly Gln Pro Val Ala
 5970 5975 5980
 Gly Ile Glu Gly Val Arg Leu Arg Pro Val Gly Asp Ala Arg Ala Val
 5985 5990 5995 6000
 Arg Ala Ala Ala Gly Ala Arg Pro Cys Arg Val Glu Trp Leu Pro Val
 6005 6010 6015
 Val Pro Ala Asp Asp Pro Glu Asp Asp Asp Ile Thr Val Val Pro Val
 6020 6025 6030
 Thr Gly Arg Thr Ala Gly Gly Asp Val Val Ala Glu Val Arg Ala Ala
 6035 6040 6045
 Val Ala Ala Ala Leu Ala Arg Leu Gln Glu Trp Leu Ala Asp Asp Arg
 6050 6055 6060
 Ser Glu Arg Leu Val Leu Val Thr Arg Gly Ala Val Ala Ala Leu Pro
 6065 6070 6075 6080
 Asp Glu Ala Pro Asp Pro Val Ala Ala Ala Val Trp Gly Leu Val Arg
 6085 6090 6095
 Ser Ala Gln Ser Glu His Pro Gly Arg Phe Val Leu Ala Asp Leu Ala
 6100 6105 6110
 Ala Ser Glu Asp Ala Ala Val Val Ser Ala Ala Val Ala Thr Gly Glu
 6115 6120 6125
 Pro Gln Ile Ala Val Arg Asn Gly Ala Leu Tyr Ala Pro Arg Leu Val
 6130 6135 6140
 Arg Gly Ala Ser Arg Pro Gly Gly Glu Val Pro Phe Gly Ala Gly Asp
 6145 6150 6155 6160
 Val Val Leu Val Thr Gly Gly Thr Gly Ala Leu Gly Arg Val Val Ala
 6165 6170 6175
 Arg His Leu Val Ala Glu His Gly Val Arg Arg Leu Val Leu Ala Ser
 6180 6185 6190
 Arg Arg Gly Gly Ala Gly Glu Leu Val Ala Glu Leu Gly Glu Leu Gly
 6195 6200 6205
 Ala Ala Val Asp Val Val Ala Cys Asp Val Ser Asp Arg Glu Ala Leu
 6210 6215 6220
 Glu Gly Leu Val Ala Ala Tyr Pro Leu Thr Gly Val Val His Ala Ala
 6225 6230 6235 6240
 Gly Val Leu Asp Asp Gly Thr Val Glu Ser Leu Thr Ala Glu Arg Val
 6245 6250 6255
 Asp Gly Val Leu Arg Ala Lys Val Asp Gly Ala Trp His Leu His Glu
 6260 6265 6270
 Leu Thr Ser Gly Leu Asp Leu Arg Ala Phe Val Leu Phe Ser Ser Leu
 6275 6280 6285
 Ala Gly Val Val Gly Ser Ala Gly Gln Gly Gly Tyr Ala Ala Ala Asn
 6290 6295 6300
 Ala Ala Leu Asp Ala Leu Ala Glu Tyr Arg Arg Ala Glu Gly Leu Pro
 6305 6310 6315 6320
 Ala Thr Ser Ile Ala Trp Gly Pro Trp Thr Asp Gly Met Thr Thr Gly
 6325 6330 6335
 Leu Glu Arg Ala Asp Arg Ala Arg Ile Ser Arg Ser Gly Thr Arg Ala
 6340 6345 6350
 Leu Gly Thr Glu Asp Gly Leu Ala Leu Phe Asp Gln Ala Val Arg Gly
 6355 6360 6365
 Ala Asp Ala Leu Ala Val Ala Ala Leu Trp Asp Leu Ser Ala Leu Arg
 6370 6375 6380
 Ala Ala Gln Ser Leu Pro Pro Leu Phe Gln Gly Leu Ala Gly Arg Pro
 6385 6390 6395 6400
 Val Arg Arg Thr Val Ala Asp Gly Ala Gly Arg Gly Ala Glu Trp Ala
 6405 6410 6415
 Asp Arg Phe Thr Gly Leu Ser Pro Ala Glu Ser Glu Arg Ala Ala Val
 6420 6425 6430
 Glu Trp Val Arg Glu Gln Ala Ala Ala Val Leu Gly His Ala Ser Ser
 6435 6440 6445

Ala Ala Val Ala Ala Asp Arg Ala Phe Leu Asp Leu Gly Phe Asp Ser
6450 6455 6460
Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Ala Thr Ala Thr Gly Leu
6465 6470 6475 6480
Arg Leu Gly Thr Thr Ala Val Phe Asp Tyr Pro Thr Pro Ala Gly Leu
6485 6490 6495
Ala Ala His Leu Leu Glu Arg Val Leu Gly Ala Thr Thr Gly Ser Ala
6500 6505 6510
Gln Pro Ala Pro Ala Ala Val Pro Ala Asp Glu Pro Ile Ala Ile Val
6515 6520 6525
Gly Met Ala Cys Arg Tyr Pro Gly Gly Ile Thr Ser Pro Glu Glu Leu
6530 6535 6540
Trp Arg Glu Val Ala Glu Gly Arg Asp Ala Ile Ser Gly Phe Pro Thr
6545 6550 6555 6560
Asp Arg Gly Trp Asp Leu Arg Ala Leu Phe Ala Asp Asp Pro Gly Arg
6565 6570 6575
Pro Gly Thr Ser His Thr Arg Glu Gly Gly Phe Leu His Asp Ala Gly
6580 6585 6590
Glu Phe Asp Ala Asp Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala
6595 6600 6605
Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ala Ser Trp Glu Ala Ile
6610 6615 6620
Glu Arg Ala Gly Ile Asp Pro Ala Ser Leu Arg Gly Ser Arg Thr Gly
6625 6630 6635 6640
Val Tyr Ala Gly Val Met Tyr His Asp Tyr Ala Ala Arg Val Asp Val
6645 6650 6655
Leu Pro Asp Gly Val Glu Gly Tyr Leu Gly Thr Gly Asn Ser Gly Ser
6660 6665 6670
Ile Ala Ser Gly Arg Ile Ala Tyr Ala Leu Gly Leu Glu Gly Gln Ala
6675 6680 6685
Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Trp
6690 6695 6700
Ala Val Arg Ser Leu Arg Ser Gly Glu Ser Asp Leu Ala Leu Ala Gly
6705 6710 6715 6720
Gly Val Thr Val Met Ala Thr Pro Gly Val Phe Val Asp Phe Ser Arg
6725 6730 6735
Gln Arg Gly Leu Ala Thr Asp Gly Arg Cys Lys Ser Tyr Gly Ala Gly
6740 6745 6750
Ala Asp Gly Thr Gly Trp Ser Glu Gly Val Gly Met Leu Leu Val Glu
6755 6760 6765
Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Ile
6770 6775 6780
Arg Gly Thr Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala
6785 6790 6795 6800
Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asp
6805 6810 6815
Ala Ser Leu Arg Pro Ser Asp Val Asp Ala Val Glu Gly His Gly Thr
6820 6825 6830
Gly Thr Ser Leu Gly Asp Pro Ile Glu Val Glu Ala Leu Leu Ala Thr
6835 6840 6845
Tyr Gly Gln Glu Arg Asp Glu Pro Leu Trp Leu Gly Ser Ile Lys Ser
6850 6855 6860
Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys
6865 6870 6875 6880
Met Val Glu Ala Met Arg His Gly Val Leu Pro Arg Thr Leu His Ala
6885 6890 6895
Asp Glu Pro Ser Pro His Ile Asp Trp Ala Ser Gly Ala Val Glu Leu
6900 6905 6910
Leu Ala Glu Gln Arg Gln Trp Pro Arg Thr Asp Arg Pro Arg Arg Ala
6915 6920 6925

ES 2 556 211 T3

Ala Val Ser Ser Phe Gly Leu Ser Gly Thr Asn Ala His Val Val Leu
6930 6935 6940
Glu His Thr Asp His Gly Asp His His Ser Ala Gly Thr Gly Glu Arg
6945 6950 6955 6960
Pro Ala Val Pro Val Pro Val Pro Val Pro Val Ala Leu Ser Ala Arg
6965 6970 6975
Thr Asp Ala Gly Leu Arg Ala Gln Ala Asp Arg Leu Ala Ala Ala Leu
6980 6985 6990
Thr Ala Asp Pro Asp Leu Ile Pro Leu Asp Ile Ala Tyr Ser Ala Val
6995 7000 7005
Thr Gly Arg Ala Arg Leu Glu Arg Arg Ala Ser Val Val Ala Ala Thr
7010 7015 7020
Arg Glu Glu Leu Leu Ala Gly Leu Gly Asp Leu Gly Pro Ala Thr Val
7025 7030 7035 7040
Ala Gly Ala Gly His Thr Ala Phe Leu Phe Thr Gly Gln Gly Ala Gln
7045 7050 7055
Arg Pro Gly Thr Gly Glu Glu Leu Ala Ala Ala His Pro Val Phe Ala
7060 7065 7070
Ala Ala Tyr Ala Glu Val Cys Thr Ala Phe Asp Ala Val Leu Asp Arg
7075 7080 7085
Pro Leu Arg Glu Val Val Ala Thr Gly Asp Gly Leu Asp Asp Thr Gly
7090 7095 7100
Tyr Gly Gln Pro Ala Val Phe Ala Leu Glu Val Ala Phe Gly Arg Leu
7105 7110 7115 7120
Phe Glu Ser Trp Gly Val Val Pro Asp Phe Leu Leu Gly His Ser Val
7125 7130 7135
Gly Glu Leu Ala Ala Ala His Leu Ala Gly Val Trp Ser Leu Pro Asp
7140 7145 7150
Ala Val Arg Val Val Ala Ala Arg Ser Arg Leu Met Ala Ala Leu Pro
7155 7160 7165
Ala Gly Gly Ala Met Ala Ala Val Glu Ala Ser Ala Asp Glu Val Ala
7170 7175 7180
Ala Glu Leu Ala Asp Gly Ala Val Leu Ala Ala Val Asn Gly Pro Arg
7185 7190 7195 7200
Ser Val Val Val Ser Gly Ala Arg Asp Ala Val Leu Ala Thr Ala Gly
7205 7210 7215
Leu Trp Ala Ala Arg Gly Cys Arg Thr Arg Glu Leu Lys Val Ser His
7220 7225 7230
Ala Phe His Ser Pro Leu Met Glu Pro Met Leu Ser Asp Phe Ala Ala
7235 7240 7245
Ala Leu Ala Asp Val Glu Phe Arg Ala Pro Arg Ile Pro Leu Val Ser
7250 7255 7260
Thr Val Thr Gly Ala Val Ala Gly Asp Glu Leu Cys Thr Pro Gly Tyr
7265 7270 7275 7280
Trp Val Arg His Val Arg Asp Thr Val Arg Phe Ala Asp Gly Val Arg
7285 7290 7295
Ala Leu Gly Gln Ala Gly Val Asp Thr Val Val Glu Leu Gly Pro Asp
7300 7305 7310
Gly Val Leu Thr Ala Met Ala Ala Pro Leu Leu Pro Asp Thr Ala Val
7315 7320 7325
Ala Leu Pro Thr Leu Arg Ala Gly Arg Pro Glu Ala Pro Ala Val Ala
7330 7335 7340
Ala Ala Leu Gly Ala Leu His Asp Arg Gly Thr Ala Val Asp Trp Pro
7345 7350 7355 7360
Ala Phe Phe Arg Gly Thr Gly Ala Arg Thr Val Glu Leu Pro Thr Thr
7365 7370 7375
Ala Phe Gln Arg Thr Arg Tyr Trp Leu Glu Ser Ala Ala Arg Thr Gly
7380 7385 7390
Asp Leu Ser Ala Ala Gly Leu Ala Ala Ala Gly His Pro Leu Leu Gly
7395 7400 7405

ES 2 556 211 T3

Ala Ala Val Asp Ser Pro Asp Gly Met Leu Leu Thr Gly Arg Leu Asp
7410 7415 7420
Thr Ala Thr His Pro Trp Leu Ala Asp His Thr Val Leu Asp Thr Val
7425 7430 7435 7440
Leu Leu Pro Gly Thr Ala Phe Val Glu Leu Ala Arg Ala Ala Gly Glu
7445 7450 7455
Arg Val Gly Leu Pro Arg Val Arg Glu Leu Thr Leu Ala Ala Pro Leu
7460 7465 7470
Val Leu Pro Ala Asp Gly Ser Val Leu Val Gln Val His Val Gly Ala
7475 7480 7485
Ala Val Asp Gly Glu Arg Pro Val Thr Val Ser Ala Arg Thr Asp Asp
7490 7495 7500
Gly Gln Asp Trp Ala Arg His Ala Thr Gly Val Leu Ala Pro Ala Ala
7505 7510 7515 7520
Pro Ala Pro Gly Ala Asp Pro Leu Pro Trp Pro Pro Arg Asp Ala Glu
7525 7530 7535
Pro Val Ala Thr Ala Gly His Tyr Asp Glu Leu Ala Ala Ala Gly Leu
7540 7545 7550
Gly Tyr Gly Pro Ala Phe Arg Ala Leu Arg Ala Val Trp Arg Arg Gly
7555 7560 7565
Asp Gly Pro Ala Ala Glu Val Phe Ala Glu Ile Gly Pro Ala Pro Gly
7570 7575 7580
Thr Asp Pro Ala Gly Phe Gly Val His Pro Ala Leu Leu Asp Thr Ala
7585 7590 7595 7600
Leu His Ala Ala Ala Val Gly Gly Leu Gly Val Ala Gly Val Pro Phe
7605 7610 7615
Thr Trp Asn Glu Val Ala Val His Thr Ala Gly Ala Arg Ser Leu Arg
7620 7625 7630
Val Arg Ile Ala Pro Asp Gly Ser Gly Gly Leu Ser Leu Arg Ala Thr
7635 7640 7645
Asp Asp Glu Gly Arg Ala Val Ile Asp Val Gly Ala Leu Arg Leu Arg
7650 7655 7660
Ala Ile Ala Pro Glu Asp Leu Arg Thr Ala Pro Ala Ala Ala Glu Ala
7665 7670 7675 7680
Leu Phe Asp Leu Asp Trp Ala Pro Val Ser Val Ala Arg Gly Ala Arg
7685 7690 7695
Pro Ala Gly Arg Trp Ala Val Leu Ala Pro Ala Asp Ala Glu Leu Thr
7700 7705 7710
Asp Val Leu Gly Ala Gly Ile Glu Thr Val Lys Asp Leu Ala Thr Val
7715 7720 7725
Pro Asp Asp Leu Asp Ile Val Leu Ala Ala Val Ser Asp Arg Gly Gly
7730 7735 7740
Ala Gly Val Asp Ala Pro Thr Ala Ala Ser Val Gly Ala Ala Ala Glu
7745 7750 7755 7760
Thr Asp Val Pro Glu Ala Val Arg Ala Ile Leu Arg Gly Ala Leu Gly
7765 7770 7775
Leu Val Gln Thr Trp Val Ala Gly Glu Ser Ala Ala Arg Leu Val Leu
7780 7785 7790
Leu Thr Arg Gly Ala Val Ala Leu Ala Gly Glu Arg Pro Asp Leu Ala
7795 7800 7805
Gly Ala Ala Ala Trp Gly Leu Val Arg Ala Ala Gln Ser Glu His Pro
7810 7815 7820
Gly Arg Leu Val Leu Ala Asp Val Asp Asp Asp Pro Ala Ser Leu Ala
7825 7830 7835 7840
Ala Leu Pro Ala Ala Leu Ala Thr Gly Glu Pro Gln Leu Leu Ile Arg
7845 7850 7855
Ala Gly Ala Val Arg Ala Ala Arg Leu Val Arg Ala Ala Pro Ala Glu
7860 7865 7870
Pro Ala Val Pro Ala Ala Pro Leu Gly Ala Arg Gly Thr Val Leu Val
7875 7880 7885

Thr Gly Ala Thr Gly Ser Leu Gly Thr Leu Val Val Arg His Leu Val
 7890 7895 7900
 Ala Glu His Gly Val Arg Arg Leu Leu Leu Val Ser Arg Gln Gly Arg
 7905 7910 7915 7920
 Gln Pro Glu Pro Ala Ala Glu Leu Ala Ala Ala Gly Ala Glu Val Arg
 7925 7930 7935
 Phe Ala Ala Cys Asp Val Ala Asp Arg Asp Ala Leu Ala Ala Leu Leu
 7940 7945 7950
 Ala Ser Val Asp Pro Glu His Pro Val Thr Ala Val Val His Ala Ala
 7955 7960 7965
 Gly Val Leu Asp Asp Gly Val Val Ala Ala Leu Thr Pro Asp Arg Leu
 7970 7975 7980
 Asp Thr Val Leu Arg Pro Lys Ala Asp Ala Ala Trp His Leu His Glu
 7985 7990 7995 8000
 Leu Thr Gly Glu Leu Asp Ala Phe Val Leu Phe Ser Ser Ala Ala Gly
 8005 8010 8015
 Leu Leu Gly Ala Pro Gly Gln Ala Asn Tyr Ala Ala Ala Asn Ala Phe
 8020 8025 8030
 Leu Asp Ala Leu Ala Ala His Arg Arg Ala Ala Gly Leu Pro Ala Val
 8035 8040 8045
 Ser Ile Ala Trp Gly Pro Trp Ala Asp Gly Met Ala Ala Gln Leu Asp
 8050 8055 8060
 Thr Arg Arg Ala Ser Arg Ala Gly Leu Leu Pro Leu Asp Ala Ala Leu
 8065 8070 8075 8080
 Gly Leu Ala Leu Phe Asp Thr Ala Arg Thr Gly Ala Pro Thr Ala Pro
 8085 8090 8095
 Leu Ala Ala Arg Leu Asp Leu Pro Gly Leu Arg Ala Ala Ala Ala Glu
 8100 8105 8110
 Leu Pro Pro Ile Leu Arg Thr Leu Val Pro Ala Pro Ala Arg Thr Ala
 8115 8120 8125
 Pro Glu Pro Ala Glu Pro Leu Ala Asp Thr Leu Ala Ala Leu Pro Ala
 8130 8135 8140
 Glu Glu Arg Glu Arg Arg Ala Leu Asp Ala Val Leu Arg His Thr Ala
 8145 8150 8155 8160
 Glu Val Leu Gly His Ala Thr Ala Asp Gly Val Asp Arg Glu Arg Gly
 8165 8170 8175
 Phe Gln Gln Leu Gly Phe Asp Ser Leu Met Ser Val Glu Leu Arg Asn
 8180 8185 8190
 Arg Leu Gly Ala Ala Ala Gly Leu Arg Leu Pro Ala Thr Val Ile Phe
 8195 8200 8205
 Asp His Pro Thr Pro Ala Ala Leu Ala Ala His Leu Val Ala Glu Leu
 8210 8215 8220
 Ala Pro Gly Arg Pro Thr Val Ala Ser Leu Arg Thr Ser Thr Leu Ala
 8225 8230 8235 8240
 Glu Leu Glu Ala Ala Ala Asn Glu Phe Ala Gly Asp Pro Glu Leu Arg
 8245 8250 8255
 Glu Gly Leu Arg Thr Arg Leu Arg Ala Leu Leu Arg Thr Leu Asp Asp
 8260 8265 8270
 Pro Ala Pro Asp Glu Pro Leu Glu Glu Thr Gly Glu Glu Ser Leu Ala
 8275 8280 8285
 Glu Leu Leu Asp Leu Ala Asp Arg Glu Leu Gly Asp Phe
 8290 8295 8300

<211> 3405

<212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

5 <400> 12

Met Thr Glu His Asp Arg Thr Thr Asp Arg Ala Thr Ala Arg Gly Ala

ES 2 556 211 T3

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

				485					490					495	
Pro	Arg	Pro	Val	Gly	Asn	Gln	Ala	Thr	Ala	Ala	Pro	Arg	Ser	Met	Glu
			500					505					510		
Asp	Arg	Thr	Ala	Ala	Ala	Pro	Ser	Ala	Gly	Gly	Asp	Pro	Thr	Leu	Thr
		515					520				525				
Ala	Pro	Ala	Pro	Ser	Ala	Pro	Arg	Pro	Ala	Pro	Ala	Ala	Leu	Pro	Val
	530					535					540				
Pro	Leu	Ser	Ala	Ala	Thr	Glu	Pro	Gly	Val	Arg	Ala	Gln	Ala	Leu	Arg
545					550					555					560
Leu	Ala	Ala	His	Leu	Thr	Glu	His	Pro	Glu	Leu	Ala	Pro	Gln	Asp	Ile
			565						570					575	
Ala	Phe	Ser	Ala	Ala	Thr	Thr	Arg	Ala	Ala	Leu	Ala	Ser	Arg	Ala	Val
			580					585					590		
Val	Leu	Ala	Asp	Asp	Arg	Ala	Gly	Leu	Leu	Asp	Ala	Leu	Thr	Ala	Leu
		595					600				605				
Ala	Glu	Gly	Arg	Pro	Gly	Pro	Ala	Val	Val	Thr	Gly	Ala	Ala	Ala	Ala
	610				615						620				
Gly	Ala	Arg	Arg	Ile	Thr	Phe	Val	Phe	Pro	Gly	Gln	Gly	Ala	Gln	Trp
625				630						635					640
Ala	Gly	Met	Ala	Val	Pro	Leu	Leu	Glu	Thr	Ser	Pro	Val	Phe	Ala	Ala
			645						650					655	
Lys	Trp	Ala	Glu	Cys	Ala	Arg	Val	Leu	Ala	Pro	Trp	Val	Asp	Trp	Ser
		660						665					670		
Pro	Asp	Glu	Ala	Leu	Arg	Ser	Pro	Gln	Ala	Leu	Glu	Arg	Val	Asp	Val
		675				680						685			
Val	Gln	Pro	Val	Leu	Trp	Ala	Val	Met	Val	Ser	Leu	Ala	Glu	Leu	Trp
	690					695					700				
Arg	Ala	Ala	Gly	Val	Arg	Pro	Asp	Ala	Val	Leu	Gly	His	Ser	Gln	Gly
705					710						715				720
Glu	Ile	Ala	Ala	Ala	Cys	Val	Ala	Gly	Ala	Leu	Ser	Leu	Glu	Asp	Gly
			725						730					735	
Ala	Lys	Val	Val	Ala	Leu	Arg	Ala	Lys	Ala	Leu	Leu	Ala	Leu	Ala	Gly
		740						745						750	
Arg	Gly	Gly	Met	Leu	Ser	Val	Pro	Leu	Pro	Glu	Ala	Glu	Val	Arg	Ala
		755					760					765			
Arg	Leu	Asp	Ser	Arg	Pro	Gly	Leu	Gly	Ile	Ala	Ala	Val	Asn	Gly	Pro
	770					775					780				
Ala	Thr	Val	Val	Val	Ser	Gly	Glu	Thr	Ala	Ala	Leu	Asp	Glu	Ala	Gln
	785				790						795				800
Ala	Ala	Trp	Glu	Ala	Glu	Gly	Val	Arg	Val	Arg	Arg	Ile	Pro	Val	Asp
			805						810						815
Tyr	Ala	Ser	His	Ser	Pro	His	Val	Ala	Glu	Val	Gln	Asp	Arg	Leu	Ala
		820						825				830			
Ala	Asp	Leu	Ala	Gly	Ile	Ala	Pro	Arg	Pro	Ala	Glu	Val	Thr	Phe	Leu
		835					840					845			
Ser	Thr	Leu	Thr	Gly	Glu	Pro	Phe	Asp	Thr	Thr	Gly	Leu	Asp	Ala	Gly
	850					855					860				
Tyr	Trp	Tyr	Arg	Asn	Leu	Arg	Glu	Gln	Val	Arg	Phe	Glu	Ala	Ala	Thr
865				870						875					880
Arg	Arg	Ala	Leu	Glu	Gln	Gly	His	Arg	Val	Phe	Ile	Glu	Val	Gly	Pro
			885						890					895	
His	Pro	Val	Leu	Thr	Leu	Gly	Val	Gln	Gln	Thr	Ala	Glu	Ala	Met	Asp
		900						905						910	
Val	Pro	Ala	Glu	Ala	Ile	Ala	Thr	Leu	Arg	Arg	Asp	Gln	Gly	Asp	Leu
		915					920					925			
Leu	Arg	Phe	Arg	Thr	Ala	Leu	Ala	Glu	Ala	Ala	Val	Leu	Gly	Ala	Pro
	930					935					940				
Val	Asp	Trp	Ala	Ala	Glu	Leu	Ala	Pro	Tyr	Ala	Pro	Arg	Arg	Val	Asp
945				950						955					960
Leu	Pro	Thr	Tyr	Ala	Phe	Gln	Arg	Glu	Arg	Tyr	Trp	Leu	Thr	Pro	Gln

ES 2 556 211 T3

				1445					1450					1455	
Leu	Pro	Phe	Leu	Asp	Pro	Asp	Leu	Ala	Val	Ala	Gly	Leu	Arg	Gln	Ala
			1460					1465						1470	
Leu	Asp	His	Asp	Glu	Thr	Thr	Val	Val	Ile	Ala	Asp	Val	Asp	Trp	Thr
		1475					1480							1485	
Arg	Phe	Val	Pro	Val	Phe	Cys	Ser	Ala	Arg	Arg	Arg	Pro	Leu	Leu	Glu
	1490					1495						1500			
Ser	Val	Pro	Glu	Ala	Ala	Ala	Val	Leu	Arg	Thr	Ala	Thr	Val	Asp	Thr
1505					1510						1515				1520
Gly	Thr	Val	Ser	Ala	Leu	Arg	Glu	Arg	Leu	Gly	Ser	Leu	Pro	Glu	Ala
				1525					1530						1535
Gly	Arg	Arg	Arg	Ala	Val	Thr	Glu	Leu	Val	Arg	Glu	His	Ala	Ala	Ala
			1540					1545						1550	
Val	Leu	Gly	His	Asp	Ser	Pro	Ala	Ala	Leu	Pro	Ala	Asp	Arg	Ala	Phe
		1555					1560						1565		
Arg	Asp	Val	Gly	Phe	Asp	Ser	Ile	Thr	Ala	Val	Glu	Leu	Arg	Asn	Arg
	1570					1575							1580		
Leu	Arg	Ser	Ala	Thr	Gly	Leu	Ala	Leu	Pro	Ala	Thr	Leu	Val	Phe	Asp
1585					1590					1595					1600
His	Pro	Ser	Pro	Thr	Ala	Leu	Ala	Gly	His	Leu	Leu	Ala	Leu	Ala	Phe
				1605					1610						1615
Asp	Thr	Ala	Ala	Ala	Asp	Leu	Ala	Ala	Pro	Ala	Ala	Arg	Ala	Ala	Asp
			1620					1625						1630	
Asp	Asp	Asp	Asp	Pro	Ile	Ala	Val	Val	Gly	Leu	Ser	Cys	Arg	Tyr	Ala
		1635					1640						1645		
Gly	Gly	Val	Ala	Ser	Pro	Asp	Glu	Leu	Trp	Arg	Leu	Val	Val	Ala	Gly
	1650					1655							1660		
Gln	Asp	Ala	Val	Gly	Ala	Leu	Pro	Thr	Asp	Arg	Gly	Trp	Asp	Leu	Asp
1665					1670						1675				1680
Ser	Leu	Tyr	Asp	Ser	Asp	Pro	Asp	Ala	Arg	Gly	Arg	Ser	Tyr	Val	Arg
				1685					1690						1695
Gln	Gly	Ala	Phe	Leu	Thr	Asp	Pro	Ala	Gly	Phe	Asp	Ala	Ala	Phe	Phe
		1700						1705							1710
Gly	Ile	Ala	Pro	Ala	Glu	Ala	Arg	Ala	Thr	Asp	Pro	Gln	Gln	Arg	Leu
	1715						1720							1725	
Leu	Leu	Glu	Ala	Ala	Trp	Glu	Ala	Phe	Glu	His	Ala	Gly	Ile	Asp	Ala
	1730					1735						1740			
Thr	Gly	Leu	Arg	Gly	Ser	Arg	Val	Gly	Val	Phe	Ala	Gly	Ala	Asn	Val
1745					1750					1755					1760
Gly	Asp	Tyr	Ala	Ser	Ser	Arg	Gly	Pro	Gly	Ala	Gly	Gly	Ser	Asp	Gly
				1765					1770						1775
Gln	Leu	Leu	Thr	Gly	Asn	Val	Pro	Ser	Val	Ile	Ser	Gly	Arg	Ile	Ser
			1780					1785							1790
Tyr	Thr	Phe	Gly	Phe	Glu	Gly	Pro	Ala	Val	Thr	Val	Asp	Thr	Ala	Cys
	1795						1800						1805		
Ser	Ser	Ala	Leu	Val	Ala	Leu	His	Leu	Ala	Cys	Arg	Ser	Val	Arg	Gly
	1810					1815						1820			
Gly	Glu	Ser	Asp	Met	Ala	Leu	Ala	Gly	Gly	Val	Ala	Leu	Met	Ser	Ser
1825					1830					1835					1840
Pro	Ala	Ala	Leu	Ile	Gly	Phe	Ser	Ala	Gln	Arg	Gly	Leu	Ser	Gly	Asp
				1845					1850						1855
Gly	Arg	Cys	Lys	Ala	Phe	Ala	Asp	Ala	Ala	Asp	Gly	Thr	Gly	Leu	Ala
			1860					1865							1870
Glu	Gly	Val	Gly	Leu	Leu	Leu	Val	Glu	Arg	Leu	Ser	Arg	Ala	Arg	Ala
	1875						1880								1885
Gln	Gly	His	Arg	Val	Leu	Ala	Leu	Val	Arg	Gly	Ser	Ala	Ile	Asn	Gln
	1890					1895						1900			
Asp	Gly	Ala	Ser	Asn	Gly	Leu	Thr	Ala	Pro	Ser	Gly	Pro	Ala	Gln	Gln
1905					1910					1915					1920
Arg	Val	Ile	Thr	Ala	Ala	Leu	Ala	Asp	Ala	Gly	Leu	Arg	Pro	Ala	Asp

ES 2 556 211 T3

				1925					1930				1935			
Val	Asp	Ala	Val	Glu	Ala	His	Gly	Thr	Gly	Thr	Arg	Leu	Gly	Asp	Pro	
			1940						1945				1950			
Ile	Glu	Ala	Gln	Ala	Leu	Leu	Ala	Thr	Tyr	Gly	Gln	Asp	Arg	Ala	Glu	
			1955						1960				1965			
Pro	Leu	Trp	Leu	Gly	Ser	Val	Lys	Ser	Asn	Ile	Gly	His	Ser	Gln	Ala	
			1970				1975					1980				
Ala	Ser	Gly	Ala	Ala	Gly	Val	Ile	Lys	Thr	Val	Gln	Ala	Leu	Arg	His	
1985						1990					1995				2000	
Gly	Leu	Leu	Pro	Ala	Thr	Leu	His	Val	Asp	Arg	Pro	Thr	Thr	Gln	Val	
				2005					2010					2015		
Asp	Trp	Thr	Ala	Gly	Ala	Val	Glu	Val	Leu	Thr	Glu	Ala	Arg	Asp	Trp	
			2020					2025					2030			
Pro	Ala	Val	Asp	Arg	Pro	Arg	Arg	Ala	Ala	Val	Ser	Ala	Phe	Gly	Leu	
			2035				2040					2045				
Ser	Gly	Thr	Asn	Ala	His	Val	Ile	Leu	Glu	Gln	Ala	Pro	Ala	Glu	Asp	
			2050			2055					2060					
Ala	His	Pro	Ala	Pro	Glu	Pro	Ala	Pro	Gly	Glu	Asp	Ser	His	Pro	Thr	
2065					2070					2075					2080	
Pro	Glu	Thr	Ala	Pro	Gly	Glu	Asp	Ala	Pro	Arg	Thr	Ala	Pro	Glu	Pro	
				2085					2090					2095		
Ala	Arg	Pro	Val	Val	Trp	Pro	Val	His	Gly	Arg	Thr	Arg	Asp	Ala	Leu	
			2100					2105					2110			
Arg	Ala	Gln	Ala	Ala	Arg	Leu	Arg	Thr	His	Leu	Glu	Thr	Arg	Pro	Asp	
			2115				2120					2125				
Ala	Arg	Pro	Ala	Asp	Val	Gly	Trp	Thr	Leu	Ala	Ala	Gly	Arg	Ala	Val	
			2130			2135					2140					
Phe	Asp	His	Arg	Ala	Val	Val	Leu	Gly	Ala	Asp	Arg	Ala	Glu	Leu	Leu	
2145					2150					2155					2160	
Arg	Gly	Leu	Asp	Ala	Val	Ala	Ala	Gly	Thr	Pro	Asp	Pro	Ala	Val	Ala	
				2165					2170					2175		
Asp	Gly	Ala	Ala	Gln	Gly	Ala	Asp	Arg	Ala	Val	Phe	Val	Phe	Pro	Gly	
			2180					2185					2190			
His	Gly	Ala	Gln	Trp	Pro	Gly	Met	Ala	Arg	Arg	Leu	Phe	Asp	Asp	Phe	
			2195				2200					2205				
Pro	Val	Phe	Arg	Glu	Ser	Val	Leu	Gln	Cys	Ala	Asp	Ala	Phe	Ala	Glu	
			2210			2215					2220					
Phe	Val	Asp	Trp	Ser	Leu	Leu	Asp	Val	Leu	Arg	Asp	Glu	Glu	Gly	Ala	
2225					2230					2235					2240	
Pro	Pro	Leu	His	Arg	Val	Asp	Val	Val	Gln	Pro	Ala	Leu	Phe	Thr	Met	
				2245					2250						2255	
Met	Val	Ser	Leu	Ala	Ala	Leu	Trp	Arg	Ser	Tyr	Gly	Val	Glu	Pro	Ser	
			2260					2265					2270			
Ala	Val	Val	Gly	His	Ser	Gln	Gly	Glu	Ile	Ala	Ala	Ala	Tyr	Val	Ala	
			2275				2280					2285				
Gly	Ala	Leu	Asp	Leu	Arg	Asp	Ala	Ala	Arg	Ile	Val	Ala	Thr	Arg	Gly	
			2290			2295					2300					
Lys	Ala	Trp	Leu	Thr	Leu	Ala	Gly	Thr	Gly	Gly	Met	Ala	Ser	Val	Ala	
2305					2310					2315					2320	
Leu	Pro	Arg	Ala	Glu	Ala	Ala	Glu	Arg	Leu	Arg	Pro	Phe	Gly	His	Arg	
				2325					2330					2335		
Leu	Asp	Ile	Ala	Ala	Val	Asn	Asp	Pro	Arg	Ser	Val	Thr	Val	Ala	Gly	
			2340					2345					2350			
Asp	Leu	Asp	Ala	Leu	Glu	Glu	Phe	Leu	Thr	Gly	Leu	Glu	Thr	Glu	Gly	
			2355				2360					2365				
Val	Arg	Val	Arg	Arg	Val	Arg	Gln	Ile	Val	Gly	Ala	Gly	His	Thr	Ala	
			2370			2375					2380					
His	Val	Asp	Ala	Leu	Arg	Asp	Gln	Leu	Ile	Glu	Thr	Leu	Ala	Pro	Thr	
2385					2390					2395					2400	
Ala	Pro	Arg	Ser	Ala	Pro	Ile	Ala	Phe	Cys	Ser	Thr	Val	Thr	Gly	Gly	

ES 2 556 211 T3

				2405					2410				2415		
Leu	Leu	Asp	Thr	Ala	Gly	Leu	Asp	His	His	Tyr	Trp	Tyr	Arg	Asn	Ala
			2420					2425					2430		
Arg	Arg	Thr	Val	Leu	Phe	Glu	Gln	Ala	Val	Arg	Thr	Leu	Ala	Glu	Gln
		2435						2440					2445		
Gly	Tyr	Gly	Pro	Phe	Leu	Glu	Ile	Ser	Ala	His	Pro	Met	Phe	Thr	Val
	2450					2455						2460			
Ala	Val	Gln	Glu	Thr	Leu	Glu	Asp	Ala	Gly	Val	Gly	Ala	Ala	Val	Leu
2465					2470						2475				2480
Ala	Thr	Leu	Arg	Arg	Asp	Glu	Gly	Gly	Pro	Asp	Arg	Phe	Leu	Arg	Ala
				2485					2490						2495
Ala	Ala	Glu	Ala	His	Thr	Ala	Gly	Val	Thr	Val	Asp	Trp	Arg	Pro	Ala
			2500					2505						2510	
Phe	Ala	Gly	Ala	Gly	Ala	Arg	Thr	Thr	Asp	Leu	Pro	Thr	Tyr	Ala	Phe
		2515					2520						2525		
Gln	Arg	Thr	Arg	His	Trp	Leu	Glu	Pro	Gly	Gly	Asp	Ser	Gly	Asp	Val
	2530					2535					2540				
Thr	Ala	Ala	Gly	Leu	Ala	Pro	Ala	Gly	His	Pro	Leu	Leu	Gly	Ala	Val
2545					2550					2555					2560
Val	Glu	Leu	Ala	Asp	Gly	Ala	Met	Val	Leu	Thr	Gly	Arg	Leu	Ser	Ala
				2565						2570					2575
Thr	Ala	Gln	Pro	Trp	Pro	Gly	Pro	Glu	Ala	Pro	Asp	Ala	Ala	Leu	Val
			2580					2585						2590	
Asp	Leu	Val	Leu	Ala	Ala	Ala	Asp	Arg	Ala	Asp	Cys	Ala	Gly	Val	Gly
		2595					2600						2605		
Glu	Leu	Thr	Val	Tyr	Glu	Pro	Leu	Met	Ala	Pro	Ala	Gly	Gly	Ala	Asp
	2610					2615						2620			
Thr	Arg	Val	Thr	Val	Gly	Ala	Pro	Asp	Ala	Val	Gly	Ser	Arg	Thr	Ala
2625					2630						2635				2640
Ala	Val	His	Thr	Arg	Thr	Ala	Asp	Ser	Pro	Trp	Val	Arg	His	Ala	Glu
				2645					2650						2655
Ala	Thr	Leu	Val	Ala	Arg	Pro	Ala	Pro	Gly	Glu	Ser	Leu	Thr	Asp	Trp
			2660						2665						2670
Pro	Pro	Ala	Gly	Ala	Glu	Pro	Ile	Asp	Pro	Pro	Ala	Glu	Ala	Gly	Arg
		2675					2680							2685	
Ala	Gly	Val	Ala	Val	Thr	Ala	Ala	Trp	Gln	Arg	Gly	Asp	Asp	Leu	Phe
	2690					2695						2700			
Thr	Glu	Val	Ala	Leu	Asp	Asp	Gly	Ala	Ala	Glu	Arg	Ala	Asp	Ala	Phe
2705					2710						2715				2720
Ala	Leu	His	Pro	Leu	Leu	Leu	Asp	Ala	Ala	Leu	Ser	Pro	Leu	Leu	Asp
				2725						2730					2735
Gly	Glu	Leu	Thr	Pro	Thr	Ala	Trp	Ser	Gly	Val	Arg	Leu	His	Ala	Thr
			2740					2745						2750	
Gly	Ala	Arg	Thr	Leu	Arg	Ala	Arg	Ala	Glu	Arg	Ile	Gly	Pro	Asp	Thr
		2755					2760						2765		
Val	Ala	Leu	Thr	Leu	Ala	Asp	Pro	Glu	Gly	Gly	Pro	Val	Leu	Thr	Ala
	2770					2775						2780			
Asp	Thr	Val	Arg	Leu	Ala	Ala	Ala	Pro	Ala	Ala	Gly	Ala	Gly	Arg	Arg
2785					2790						2795				2800
Thr	Gly	Lys	Asp	Ala	Leu	Phe	Arg	Met	Glu	Trp	Val	Pro	Ala	Pro	Leu
				2805						2810					2815
Ala	Pro	Ala	Thr	Pro	Gly	Arg	Trp	Ala	Val	Leu	Gly	Ala	Asp	Pro	Leu
			2820					2825						2830	
Gly	Ala	Ala	Asp	Thr	Leu	Arg	Ser	Leu	Gly	His	Thr	Val	His	Gly	Ala
		2835					2840							2845	
Asp	Gly	Pro	Ala	Gly	Leu	Thr	Glu	Val	Pro	Asp	Ala	Val	Leu	Ile	Thr
	2850					2855						2860			
Ala	Val	Gly	Ala	Pro	Gly	Glu	Ala	Pro	Ala	Glu	Ala	Arg	Thr	Val	Leu
2865					2870						2875				2880
His	Gly	Thr	Leu	Ala	Ala	Leu	Gln	Thr	Leu	Leu	Ala	Asp	Asp	Arg	Phe

ES 2 556 211 T3

				2885					2890				2895		
Thr	Ala	Val	Pro	Leu	Val	Val	Leu	Thr	Arg	Gly	Ala	Ala	Ala	Asp	Arg
			2900						2905				2910		
Ala	Asp	Asp	Leu	Ala	Gly	Ala	Ala	Ala	Trp	Gly	Leu	Val	Arg	Ser	Ala
			2915						2920				2925		
Gln	Ser	Glu	His	Pro	Gly	Arg	Phe	Val	Leu	Ala	Asp	Ile	Asp	Asp	Asp
			2930						2935				2940		
Pro	Arg	Ser	Trp	Arg	Ala	Leu	Gly	Gly	Val	Pro	Ala	Thr	Gly	Glu	Pro
			2945						2950				2955		2960
Gln	Leu	Ala	Leu	Arg	Ala	Gly	Ala	Ala	Thr	Val	Pro	Arg	Leu	Ala	Arg
									2965						2975
Leu	Ala	Pro	Pro	Asp	Ala	Pro	Ala	Pro	Trp	Asp	Pro	Asp	Arg	Thr	Val
									2980						2990
Leu	Val	Thr	Gly	Ala	Ser	Gly	Asp	Leu	Gly	Ala	Leu	Val	Ala	Arg	His
									2995						3005
Leu	Val	Ala	Ala	His	Gly	Val	Arg	His	Leu	Ile	Leu	Ala	Ser	Arg	Arg
									3010						3015
Gly	Pro	Ala	Ala	Pro	Gly	Ala	Ala	Gly	Leu	Gly	Ala	Glu	Leu	Arg	Ala
									3020						3025
Ser	Gly	Ala	Ala	Thr	Val	Thr	Ile	Ala	Ala	Cys	Asp	Thr	Ala	Asp	Arg
									3030						3035
Lys	Ala	Leu	Ala	Glu	Leu	Ile	Ala	Ala	Val	Pro	Asp	Glu	His	Pro	Leu
									3045						3050
Thr	Ala	Val	Val	His	Ser	Ala	Ser	Val	Leu	Asp	Asp	Gly	Val	Ile	Ala
									3055						3060
Ala	Leu	Asp	Pro	Asp	Arg	Leu	Asp	Thr	Ala	Leu	Arg	Pro	Lys	Ala	Asp
									3065						3070
Ala	Ala	Trp	His	Leu	His	Glu	Leu	Thr	Arg	His	Leu	Asn	Leu	Ser	Ala
									3075						3080
Phe	Val	Leu	Phe	Ser	Ser	Val	Ala	Gly	Thr	Phe	Gly	Gly	Leu	Gly	Gln
									3085						3090
Gly	Asn	Tyr	Ala	Ala	Gly	Asn	Ala	Phe	Leu	Asp	Ala	Leu	Ala	Arg	His
									3095						3100
Arg	Arg	Ala	His	Gly	Leu	Pro	Ala	Thr	Ser	Val	Ala	Trp	Gly	Trp	Trp
									3105						3110
Ala	Glu	Arg	Ala	Ala	Lys	Ser	Gly	His	Gly	Thr	Ala	Ala	Glu	Ala	Pro
									3115						3120
Val	Ala	Val	Asn	Gly	Met	Thr	Pro	Leu	Thr	Glu	Asp	His	Gly	Leu	Ala
									3125						3130
Leu	Phe	Asp	Ala	Ala	Cys	Arg	Gly	Asp	Glu	Pro	Phe	Val	Val	Ala	Gly
									3135						3140
Ala	Leu	His	Leu	Arg	Ser	Leu	Arg	Ala	Ala	Ala	Asp	Glu	Leu	Pro	Ala
									3145						3150
Pro	Leu	Arg	Gly	Leu	Val	Arg	Ala	Pro	Ala	Arg	Lys	Ala	Ala	Ala	Gln
									3155						3160
Ala	Ala	Gly	Thr	Gly	Thr	Pro	Thr	Val	Ala	Gly	Glu	Leu	Ala	Gly	Arg
									3165						3170
Pro	Pro	Ala	Arg	Arg	Glu	Ala	Phe	Leu	Val	Asp	Leu	Val	Arg	Asp	Glu
									3175						3180
Thr	Ala	Leu	Val	Leu	Gly	His	Thr	Gly	Arg	Asp	Asp	Val	Pro	Ala	His
									3185						3190
His	Arg	Phe	Leu	Asp	Gln	Gly	Phe	Asp	Ser	Leu	Ala	Ala	Leu	Lys	Leu
									3195						3200
Arg	Asn	Arg	Leu	Ala	Ala	Ala	Thr	Gly	Leu	Arg	Leu	Pro	Pro	Thr	Leu
									3205						3210
Val	Phe	Asp	His	Pro	Thr	Pro	Thr	Glu	Leu	Ala	Arg	His	Leu	Leu	Ala
									3215						3220
Glu	Leu	Val	Pro	Pro	Ala	Asp	Ala	Glu	Pro	Pro	Ala	Phe	Asp	Asp	Glu
									3225						3230
Asp	Ala	Ala	Thr	Ala	Val	Leu	Val	Leu	Glu	Glu	Leu	Gly	Gln	Leu	Asp
									3235						3240

ES 2 556 211 T3

					3365					3370					3375
Glu	Ala	Ile	Thr	Arg	Leu	Pro	Ala	Asp	Gly	Pro	Glu	Arg	Thr	Arg	Ile
					3380					3385					3390
Ala	Gly	Leu	Leu	Thr	Asp	Leu	Leu	Ala	Arg	Trp	Gly	Arg			
					3395					3400					3405

<210> 13

<211> 388

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 13

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

<210> 14

5 <211> 1515

<212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 14

5

```

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

```

ES 2 556 211 T3

Ala Ala Trp Asp Pro Ala Thr Ser Pro Phe Ala Phe Asp Thr Thr Ala
405 410 415
Arg Pro Trp Ala Glu Pro Arg Arg Val Ala Gly Val Ser Ala Phe Gly
420 425 430
Phe Gly Gly Val Asn Tyr His Ala Val Val Ala Ala Pro Ala Pro Asp
435 440 445
Pro Ser His Glu Ser Ala Pro Arg Ser Ala Gly Leu Leu Leu Phe Arg
450 455 460
Gly Thr Glu Asp Glu Val Gln Thr Ala Leu Arg Asp Leu Ala Gly Arg
465 470 475 480
Leu Ala Glu Gly Thr Val Arg Pro Ser Ala Leu Thr Gly Val Thr Gly
485 490 495
Ala Gly Pro Ala Arg Val Ala Leu Val Thr Glu Gly Ala Ala Thr Leu
500 505 510
Gly Glu Gln Leu Asp Leu Ala Leu Arg Leu Glu Asp Arg Pro Glu Arg
515 520 525
Gly Val His Val Gly Asp Gly Ala Arg Gln Pro Val Ala Leu Leu Phe
530 535 540
Pro Gly Gln Gly Ser Gln Arg Pro Gly Met Leu Ala Asp Leu Phe Val
545 550 555 560
Ala Phe Pro Arg Leu His Arg Phe Leu Arg Ala Glu Pro Ala Leu Thr
565 570 575
Gly Ala Leu Phe Pro Pro Ala Ala Phe Gly Gly Thr Pro Pro Arg Leu
580 585 590
Val Asp Thr Arg Leu Ala Gln Pro Ala Leu Gly Leu Gly Ala Leu Ala
595 600 605
Ala Leu Asp Leu Leu Arg Ala Cys Gly Val Glu Pro Asp Met Thr Ala
610 615 620
Gly His Ser Tyr Gly Glu Leu Pro Ala Leu Ala Ala Ala Gly Ala Leu
625 630 635 640
Pro Glu Thr Ala Val Met Pro Leu Ser Gly Ala Arg Ala Glu Ala Met
645 650 655
His Thr Ala Ala Gly Gln Asp Pro Gly Ala Met Ala Ala Val Ala Ala
660 665 670
Pro Ala Gly Thr Val Ala Glu Leu Ala Arg His Gly Leu Ala Asp
675 680 685
Asp Val Thr Pro Ala Asn Arg Asn Ala Pro Gly Gln Val Val Leu Ala
690 695 700
Gly Ser Ala Ala Gly Ile Asp Ala Ala Val Ala Val Leu Arg Glu Ala
705 710 715 720
Gly His Thr Ala Lys Arg Leu Pro Val Ala Ala Ala Phe His Ser Pro
725 730 735
Arg Met Ala Pro Ala Val Asp Ala Phe Ala Gly Ala Leu Ala Gly His
740 745 750
Asp Leu Arg Ala Pro Arg Leu Pro Val Trp Ser Gly Ala Thr Ala Arg
755 760 765
Pro Tyr Pro Ala Asp Pro Glu Ala Ile Arg Thr Gln Leu Ala Gly Ala
770 775 780
Leu Ala Arg Pro Val Arg Phe Ala Glu Gln Ile Glu Asp Met Tyr Ala
785 790 795 800
Ala Gly Val Arg Val Phe Val Glu Ala Gly Pro Gly Thr Val Leu Thr
805 810 815
Arg Leu Ala Gly Glu Val Leu Ala Gly Arg Pro His Thr Ala Leu Ser
820 825 830
Val Glu Ala Pro Gly Arg Pro Gly Leu Ala His Leu Leu Gly Thr Leu
835 840 845
Ala Ala Leu Ala Val Arg Gly Thr Pro Val Arg Leu Asp Glu Leu Thr
850 855 860
Val Ala Arg Asp Glu Pro Val Ser Ala Pro Pro Gly Trp Thr Val Asp
865 870 875 880

ES 2 556 211 T3

Gly His Leu Val Arg Thr Ala Asp Gly Thr Pro Val Pro Gly Gly Leu
 885 890 895
 Gln Pro Ala Gly Pro Pro Val Thr Leu Thr Gly Gly Gly Val Pro Asp
 900 905 910
 Ser Arg Asp Glu Thr Val Ile Glu Phe Leu Arg Gly Thr Gln Arg Ile
 915 920 925
 Val Ala Ala Gln Gln Glu Val Met Leu Arg Tyr Leu Gly Thr Gln Pro
 930 935 940
 Pro Ala Pro Glu Pro Ala Pro Ala His Asp Pro Pro Glu Ser Ala Ala
 945 950 955 960
 Thr Ala Ala Pro Thr Val Leu Asp Leu Val Ala Ala Arg Thr Gly Tyr
 965 970 975
 Pro Pro Ala Met Leu Asp Pro Asp Leu Asp Leu Glu Ala Ala Leu Gly
 980 985 990
 Val Asp Ser Leu Lys Arg Thr Glu Ile Ala Ser Ala Leu Val Arg Gly
 995 1000 1005
 Thr Pro Ala Ala Asp Ala Val Gly Glu Leu Ala Asp Leu Arg Thr Ile
 1010 1015 1020
 Arg Ala Met Thr Asp Trp Leu Ala Thr Arg Thr Gly Pro Asp Gly Pro
 1025 1030 1035 1040
 Gly Pro Asp Gly Pro Gly Arg Gln Thr Ala Val Pro Glu Pro Glu Thr
 1045 1050 1055
 Met Pro Ala Gly Ala Ser Asp Thr Val Arg His Val Val Glu Pro Val
 1060 1065 1070
 Pro Glu Glu Pro Pro Ala Gly Pro Leu Pro Thr Leu Thr Arg Ala Ala
 1075 1080 1085
 Val Ser Gly Gly Gly Ala Val Gly Asp Val Leu Ala Thr Leu Leu Lys
 1090 1095 1100
 Glu Arg Gly Thr Glu Val Val Ser Asp Pro Ala Gly Cys Asp Ala Leu
 1105 1110 1115 1120
 Leu Leu Leu Asp Ala Leu Asp Gly Gly Asp Tyr Thr Leu Pro Gly Arg
 1125 1130 1135
 Phe Thr Asp Ile Arg Ala Ala Val Leu Gly Gly Leu Arg Thr Leu Leu
 1140 1145 1150
 Leu Ala Thr Cys His Asp Ala Gly Pro Ala Gly Ser Gly Val His Gly
 1155 1160 1165
 Leu Ala Arg Ala Leu Ser Arg Glu His Pro Gly Leu Ala Val Thr Ala
 1170 1175 1180
 Val Asp Leu Pro Ala Gly Gln Pro Ala Glu Glu Ala Ala Arg Thr Leu
 1185 1190 1195 1200
 Leu Ala Glu Leu Gly Gly Ser Arg Pro Ser Val Thr His Thr Asp Gly
 1205 1210 1215
 Arg Arg Ala Val Trp Arg Thr Arg Pro Ala Pro Leu Pro Ala Ala Asp
 1220 1225 1230
 Val Thr Ala Gly Asp Leu Gly Leu Asp Arg Asp Ser Val Val Leu Leu
 1235 1240 1245
 Thr Gly Gly Ala Arg Gly Ile Thr Ala Arg Val Ala Arg Ala Leu Ala
 1250 1255 1260
 Thr Leu Thr Gly Cys His Leu Glu Leu Val Gly Arg Ser Glu Pro Val
 1265 1270 1275 1280
 Thr Gly Thr Val Leu Thr Asp Ala Asp Leu Arg Thr Arg Leu Ile Ala
 1285 1290 1295
 Glu Gly Gly Arg Asp Pro Ala Gly Ile Glu Arg Ala Val Arg Ala His
 1300 1305 1310
 Ala Ala Gly Arg Gln Val Arg Gln Cys Leu Asp Asp Leu Ala Gly Pro
 1315 1320 1325
 Ala Ala Ser Val Arg Tyr His Arg Ala Asp Val Arg Asp Pro Glu Arg
 1330 1335 1340
 Leu Gly Ala Val Leu Asp Asp Val Tyr Ala Arg His Gly Arg Leu Asp
 1345 1350 1355 1360

Thr Val Val His Ala Ala Gly Gln Val Arg Asp Arg Leu Leu Arg Asp
 1365 1370 1375
 Lys Ser Pro Asp Glu Phe Ala Glu Val Tyr Asp Thr Lys Val Ala Gly
 1380 1385 1390
 Ala Arg Ala Leu Ala Ala Arg Leu Arg Pro Gly Leu Arg His Leu Val
 1395 1400 1405
 Leu Phe Gly Ser Ile Ala Gly Val Thr Gly Asn Arg Gly Gln Thr Asp
 1410 1415 1420
 Tyr Ala Ala Ala Asn Asp Ala Leu Asp Thr Met Ala Arg Gln Trp Ser
 1425 1430 1435 1440
 Gly Arg Val Ala Asp Arg Val Leu Ala Leu Asp Trp Gly Pro Trp Ala
 1445 1450 1455
 Ala Asp Ala Gly Gly Met Val Thr Ala Glu Leu Glu Arg Ala Tyr Ala
 1460 1465 1470
 Arg Asn Gly Ile Gly Leu Ile Asp Pro Asp Asp Gly Val Arg Ala Phe
 1475 1480 1485
 Leu Arg Glu Leu Ala Phe Gly Arg Asp Pro Gln Val Leu Leu Thr Val
 1490 1495 1500
 Gly Asp Pro Ala Gly Phe Gly Ser Ala Leu Asp
 1505 1510 1515

<210> 15

<211> 425

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 15

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

10

Val	Pro	Leu	Ala	Arg	Pro	Ala	Val	Pro	Val	Trp	Ser	Ala	Thr	Thr	Val
			260					265					270		
Ala	Pro	Tyr	Pro	Asp	Asp	Glu	Thr	Glu	Leu	Arg	Ala	Leu	Leu	Gly	Arg
		275					280					285			
His	Leu	Val	Glu	Pro	Val	Arg	Phe	Arg	Ala	Val	Ile	Glu	Arg	Leu	Tyr
	290					295					300				
Glu	Gln	Gly	Ala	Arg	Val	Phe	Val	Gln	Ala	Gly	Cys	Gly	Ser	Leu	Thr
305					310					315					320
Arg	Phe	Val	Thr	Ala	Thr	Leu	Ala	Gly	Arg	Pro	His	Leu	Thr	Val	Ser
				325					330					335	
Ala	Gly	Ser	Ala	Thr	Arg	Pro	Gln	Leu	Asp	Gln	Leu	Arg	Trp	Ala	Ala
			340					345					350		
Ala	Ala	Leu	Trp	Thr	Ala	Gly	His	Thr	Pro	Asp	Leu	Ala	Ala	Leu	Gly
		355					360					365			
Leu	Thr	Gly	Pro	Ala	Ala	Asp	Pro	Val	Pro	Ser	Thr	Arg	Ala	Val	Thr
	370					375					380				
Gly	Ala	His	Pro	Arg	Thr	Gly	Ala	Tyr	Pro	Arg	Thr	His	Pro	Val	Ala
385					390					395					400
Gly	Ala	Tyr	Glu	Ala	Ala	Val	Ala	Ala	Val	Ser	Ala	Ser	Val	Glu	Ser
				405					410					415	
Val	Tyr	Arg	Thr	Trp	Lys	Glu	Arg	Thr							
420	425														

5 <210> 16
 <211> 239
 <212> PRT
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

10 <400> 16

ES 2 556 211 T3

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

<210> 17

<211> 340

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 17

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

<210> 18

<211> 84

<212> PRT

5 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 18

ES 2 556 211 T3

```

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

```

<210> 19

<211> 616

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 19

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

305					310					315					320
Ala	His	Val	Gly	Ala	Asp	His	Ser	Glu	Val	Leu	Leu	Asp	Ser	Thr	Glu
				325					330					335	
Leu	Ala	Asp	Pro	Ala	Val	Arg	Ala	Ala	Ile	Leu	Gly	Ala	Thr	Asp	Leu
			340					345					350		
Pro	Pro	Ala	Tyr	Trp	Gly	Asp	Leu	Trp	Pro	Ser	Leu	Tyr	Leu	Leu	Cys
		355					360					365			
Arg	Glu	Ile	Arg	Lys	Arg	Ala	Thr	Val	Val	Leu	Ser	Gly	Glu	Ser	Ala
	370					375					380				
Asp	Glu	Leu	Phe	Gly	Gly	Tyr	Arg	Trp	Tyr	Gln	Arg	Pro	Glu	Ala	Val
385				390						395					400
Asp	Ala	Pro	Thr	Phe	Pro	Trp	Leu	Thr	Pro	Gly	Ser	Ala	Arg	Ile	Phe
				405					410					415	
Gly	Gly	Ser	Ser	Leu	Ile	Asp	Gln	Gly	Leu	Leu	Glu	Lys	Leu	Asp	Leu
			420					425					430		
Glu	Gly	Tyr	Arg	Arg	Asp	Arg	Tyr	Ala	Glu	Ala	Leu	Ala	Glu	Val	Pro
		435				440						445			
Val	Leu	Ala	Gly	Glu	Ser	Ala	Val	Asp	Arg	Arg	Met	Arg	Glu	Val	Thr
	450					455					460				
Tyr	Leu	Asn	Ile	Thr	Arg	Phe	Met	Gln	Ala	Val	Leu	Asp	Arg	Lys	Asp
465				470						475					480
Arg	Met	Ser	Met	Ala	Val	Gly	Leu	Glu	Val	Arg	Val	Pro	Phe	Cys	Asp
				485					490					495	
His	Arg	Leu	Met	Asp	Tyr	Val	Phe	Asn	Val	Pro	Trp	Ala	Met	Lys	Ser
			500					505					510		
Phe	Asp	Gly	Arg	Glu	Lys	Ser	Leu	Leu	Arg	Ala	Ala	Val	Arg	Asp	Val
		515					520					525			
Leu	Pro	Asn	Ser	Val	Leu	Glu	Arg	Thr	Lys	Thr	Pro	Phe	Pro	Ala	Thr
	530					535					540				
Gln	Asp	Thr	Arg	Tyr	Glu	Gln	Ala	Leu	Arg	Ala	Glu	Leu	Arg	Glu	Val
545				550						555					560
Leu	Ala	Asp	Pro	Asp	Ala	Pro	Val	Arg	Pro	Leu	Leu	Asn	Thr	Ala	Arg
				565					570					575	
Val	Asn	Arg	Val	Leu	Gly	Arg	Glu	Leu	Asp	Asp	Phe	Ser	Leu	Pro	His
			580					585					590		
Asp	Arg	Gly	Gly	Ile	Glu	Met	Ala	Leu	Trp	Leu	Asn	Arg	Trp	Leu	Ala
		595					600					605			
Ser	Tyr	Glu	Val	Thr	Val	Thr	Val								
	610					615									

<210> 20

<211> 259

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 20

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

<210> 21

5 <211> 453

<212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 21

10

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

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

<210> 22

<211> 71

5 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

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

10 <400> 22

<210> 23

<211> 178

<212> PRT

15 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 23

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

Asp Gly
 <210> 24
 5 <211> 152
 <212> PRT
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 24

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

<210> 25
 <211> 144
 15 <212> PRT

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 25

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

<210> 26

<211> 32

10 <212> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 26

aggaattcat ggcctcckcs rtsaactaca ay 32

15 <210> 27

<211> 32

<212> ADN

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

20 <400> 27

tcggatccgc cgaagtsgt ctcwabccc ca 32

<210> 28

<211> 31

25 <212> ADN

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 28

aggaattcga catcgacats gtbwtcgagc a 31

30 <210> 29

<211> 32

<212> ADN

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

35 <400> 29

tcggatccga tgatgcgctt swsbkdcac ca 32

<210> 30

<211> 59

<212> ADN

<213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

<400> 30

45 ctcgaccggt actgggcaa cgtggtggcc ggtgctgaca ttccgggat cgcgaccc 59

ES 2 556 211 T3

<210> 31
 <211> 58
 <212> ADN
 5 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 <400> 31
 ggccagttcg cgcaggaagg cccgtacgcc gtcgtccggt gtaggctgga gctgcttc 58

 10 <210> 32
 <211> 59
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 15 <400> 32
 gtggaaatcg gctcgggagc gcccaatta accgctcga ttccgggat ccgtcgacc 59

 <210> 33
 <211> 58
 20 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 <400> 33
 25 aatgatgta tcgtcgagg acgcccagaa ttcacctgct gtaggctgga gctgcttc 58

 <210> 34
 <211> 59
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))
 30
 <400> 34
 gcgcagcaga gcccggaatc agaagtactg gacgtcacca ttccgggat ccgtcgacc 59

 <210> 35
 <211> 58
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))
 35
 <400> 35
 40 ggcatctcg cccgcgga ccgccacat ggacagcagt gtaggctgga gctgcttc 58

 <210> 36
 <211> 59
 <212> ADN
 45 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 <400> 36
 gaggattgag acggcgtcgt cctggcgttt ctgcgacaca ttccgggat ccgtcgacc 59

 50 <210> 37
 <211> 58
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 55 <400> 37
 ctctctgctg gcttcggtga gtccggtgc ggcacatcact gtaggctgga gctgcttc 58

 <210> 38
 <211> 59
 60 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 <400> 38
 65 gggccgagg acaggtccgg cggccggtg cgccgagaga ttccgggat ccgtcgacc 59

 <210> 39

ES 2 556 211 T3

<211> 58
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

5 <400> 39
 cgccgaggag aacagcggga agtaggtgic gaggctgtct gtaggctgga gctgcttc 58

<210> 40
 <211> 59
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

10

<400> 40
 ggccggcccc ggagccctgg gcgccgcct gcttcggga ttccgggat ccgtcgacc 59

15

<210> 41
 <211> 58
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

20

<400> 41
 gctctccacg gaggcgtca ccgcggcgac ggcggcctct gtaggctgga gctgcttc 58

<210> 42
 <211> 59
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

25

<400> 42
 gcggtcgtga ccggatcgtc ccgcggcac gcgcggcca ttccgggat ccgtcgacc 59

30

<210> 43
 <211> 58
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

35

<400> 43
 cacgccgtcg gcggtccagc ccgctcgaa gcgcagggtt gtaggctgga gctgcttc 58

40

<210> 44
 <211> 59
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

45

<400> 44
 cttcggctc ctcgctcg cccagccct cggcgatcca ttccgggat ccgtcgacc 59

<210> 45
 <211> 58
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

50

<400> 45
 gtgatgccg atggacaggc ccgagcgc gaccacgtct gtaggctgga gctgcttc 58

55

<210> 46
 <211> 59
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

60

<400> 46
 gacgacctc gcggcaccg gtccgtcgtc caggcgatga ttccgggat ccgtcgacc 59

<210> 47
 <211> 58
 <212> ADN

65

ES 2 556 211 T3

	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 47	
5	catctcgatg ccgccccggt cgtgcggcag gctgaagtct gtaggctgga gctgcttc	58
	<210> 48	
	<211> 28	
	<212> ADN	
10	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 48	
	tttggatcct acaccggcca gggcgccc	28
	<210> 49	
15	<211> 40	
	<212> ADN	
	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 49	
20	tttggatcct aggacgctag cgtcggccag ccaggggtgc	40
	<210> 50	
	<211> 30	
	<212> ADN	
25	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 50	
	tttaagctta gcaccctgac caccggtcac	30
30	<210> 51	
	<211> 36	
	<212> ADN	
	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
35	<400> 51	
	gctagcgtcc tcggtacccc ggtgctcccc ggcacc	36
	<210> 52	
	<211> 29	
40	<212> ADN	
	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 52	
45	tttggatcct ggtgtggtgg cggatgatg	29
	<210> 53	
	<211> 42	
	<212> ADN	
50	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 53	
	tttgaattcg gcgagtacta cggcatcggc cgtccaggcg gc	42
	<210> 54	
55	<211> 28	
	<212> ADN	
	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
	<400> 54	
60	tttaagcttg caccagctgg ccgaagcg	28
	<210> 55	
	<211> 42	
	<212> ADN	
65	<213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	

ES 2 556 211 T3

	<400> 55 tttgaattcg ccgtagtact cgccaccacc ctgctgcccg gc	42
5	<210> 56 <211> 30 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
10	<400> 56 tttgaattcg ggtggtccgg agctggatcg	30
15	<210> 57 <211> 30 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
20	<400> 57 cggcagcagg gcagggacta gtatggcggc	30
25	<210> 58 <211> 30 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
30	<400> 58 tttaagcttg gtgctcggac tggcgggagc	30
35	<210> 59 <211> 30 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
40	<400> 59 gccgccatac tagtccctgc cctgctgccg	30
45	<210> 60 <211> 29 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
50	<400> 60 tttgaattcc gacgacaccg gatacgggc	29
55	<210> 61 <211> 30 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
60	<400> 61 cggcagcagg acggtgacta gtacgggtggc	30
65	<210> 62 <211> 29 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
70	<400> 62 tttaagcttc cacccatgtc tgcaccagg	29
75	<210> 63 <211> 20 <212> ADN <213> <i>Streptomyces flaveolus</i> (<i>Streptomyces</i> sp. (DSM 9954))	
80	<400> 63 gccaccgtac tagtcaccgt	20

ES 2 556 211 T3

<210> 64
 <211> 31
 <212> ADN
 5 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 <400> 64
 atgaattcgt ggcggtcgtc ggcattgcct g 31

 10 <210> 65
 <211> 31
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 15 <400> 65
 atggatcccc tcggtgccgc ggaacaggag c 31

 <210> 66
 <211> 32
 20 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

 <400> 66
 atggatccgg cgatgaccga ctggctcgcc ac 32

 25 <210> 67
 <211> 31
 <212> ADN
 <213> *Streptomyces flaveolus* (*Streptomyces* sp. (DSM 9954))

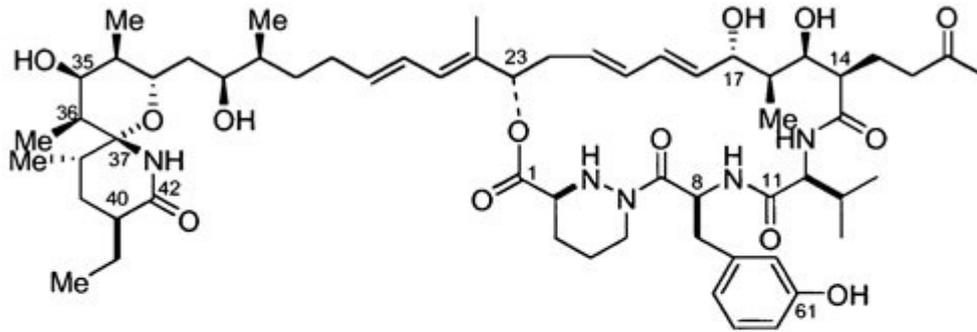
 30 <400> 67
 ataagcttgg tccccgacgg tgagcagcac c 31

REIVINDICACIONES

1. Molécula nucleica aislada, que comprende:
 - 5 (a) el ácido nucleico de agrupación de genes de biosíntesis de sangliferina A que comprende SEQ ID No. 1;
 - (b) un ácido nucleico que tiene al menos el 80% de identidad de secuencia con el ácido nucleico de (a) y que codifica para polipéptidos que tienen las mismas actividades enzimáticas y reguladoras para producir un policétido o una unidad iniciadora de policétido que los codificados por el ácido nucleico de (a); o
 - 10 (c) un ácido nucleico que codifica para uno o más polipéptidos que tienen al menos el 80% de identidad de secuencia de aminoácidos con uno o más polipéptidos codificados por el ácido nucleico de (a) y que codifica para uno o más polipéptidos que tienen una o más de las actividades enzimáticas o reguladoras necesarias para producir un policétido o un precursor o unidad iniciadora de policétido, en la que los polipéptidos tienen una cualquiera de SEQ ID No. 2-25.
2. Ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1, en el que uno o más genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaA* que son los residuos 17024-17854 de SEQ ID No. 1, *sfaB* que son los residuos 17851-19191 de SEQ ID No. 1, *sfaJ* que son los residuos 96225-97391 de SEQ ID No. 1, *sfaM* que son los residuos 103210-103929 de SEQ ID No. 1, *sfaN* que son los residuos 104001-105023 de SEQ ID No. 1, *sfaP* que son los residuos 105366-107216 de SEQ ID No. 1, *sfaQ* que son los residuos 107366-108145 de SEQ ID No. 1, *sfaL* que son los residuos 101936-103213 de SEQ ID No. 1 y *sfaO* que son los residuos 105091-105345 de SEQ ID No. 1, se han deletado o modificado de modo que se disminuye su actividad o modificado o sustituido de modo que se altera su especificidad de sustrato.
3. Ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que uno o más genes de PKS seleccionados de *sfaE* que son los residuos 30707-37360 de SEQ ID No. 1, *sfaF* que son los residuos 37394-50014 de SEQ ID No. 1, *sfaG* que son los residuos 50017-60903 de SEQ ID No. 1, *sfaH* que son los residuos 60918-85823 de SEQ ID No. 1 y *sfaI* que son los residuos 85823-96040 de SEQ ID No. 1, se han modificado mediante lo cual se han deletado uno o más dominios o módulos, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.
4. Ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que el gen de NRPS *sfaD* que son los residuos 19885-30714 de SEQ ID No. 1 se ha modificado mediante lo cual se han deletado uno o más dominios o módulos, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.
5. Ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que el gen regulador *sfaC* que son los residuos 19193-19888 de SEQ ID No. 1 se ha modificado de modo que se aumenta o disminuye su actividad, o se deleta.
6. Ácido nucleico de agrupación de genes de producción de policétido híbrido que comprende SEQ ID No. 1 en el que el gen de PKS iterativa *sfaK* que son los residuos 97396-101943 de SEQ ID No. 1 se ha modificado mediante lo cual se han deletado uno o más dominios, mutado de modo que una función enzimática se vuelve inactiva o menos activa o de modo que tiene una funcionalidad alterada o sustituido mediante una sustitución o mediante lo cual se han insertado uno o más dominios o módulos no nativos.
7. Ácido nucleico de agrupación de genes de producción de policétido híbrido según la reivindicación 2, en el que *sfaA* que son los residuos 17024-17854 de SEQ ID No. 1, y opcionalmente uno o más de otros genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaB* que son los residuos 17851-19191 de SEQ ID No. 1, *sfaJ* que son los residuos 96225-97391 de SEQ ID No. 1, *sfaM* que son los residuos 103210-103929 de SEQ ID No. 1, *sfaN* que son los residuos 104001-105023 de SEQ ID No. 1, *sfaP* que son los residuos 105366-107216 de SEQ ID No. 1, *sfaQ* que son los residuos 107366-108145 de SEQ ID No. 1, *sfaL* que son los residuos 101936-103213 de SEQ ID No. 1 y *sfaO* que son los residuos 105091-105345 de SEQ ID No. 1, se han deletado o modificado de modo que se disminuye su actividad o modificado o sustituido de modo que se altera su especificidad de sustrato.
8. Polipéptido o pluralidad de polipéptidos codificados por un ácido nucleico según una cualquiera de las reivindicaciones 1 a 7 que es un ADN.
9. Policétido sintasa codificada por un ácido nucleico según una cualquiera de las reivindicaciones 1 a 7 que es un ADN y que codifica para una o más proteínas de biosíntesis de policétido.

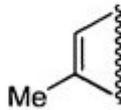
10. Vector que comprende un ácido nucleico según una cualquiera de las reivindicaciones 1 a 7 que es un ADN junto con uno o más promotores u otros elementos reguladores.
- 5 11. Célula huésped transformada con un vector según la reivindicación 10.
12. Célula huésped según la reivindicación 11, célula huésped que no produce de manera natural sangliferina A.
- 10 13. Método para producir un policétido, que comprende cultivar una célula huésped transformada según la reivindicación 11 o la reivindicación 12.
14. Proteína híbrida codificada por un ácido nucleico o agrupación de genes modular según cualquiera de las reivindicaciones 1 a 7, en la que al menos un dominio o módulo o gen no es nativo para la agrupación de genes de biosíntesis de sangliferina A.
- 15
15. Cepa modificada por ingeniería que comprende cepa productora de sangliferina A en la que *sfaA* que son los residuos 17024-17854 de SEQ ID No. 1 y opcionalmente uno o más de otros genes de biosíntesis de precursor o unidad iniciadora seleccionados de *sfaB* que son los residuos 17851-19191 de SEQ ID No. 1, *sfaJ* que son los residuos 96225-97391 de SEQ ID No. 1, *sfaM* que son los residuos 103210-103929 de SEQ ID No. 1, *sfaN* que son los residuos 104001-105023 de SEQ ID No. 1, *sfaP* que son los residuos 105366-107216 de SEQ ID No. 1, *sfaQ* que son los residuos 107366-108145 de SEQ ID No. 1, *sfaL* que son los residuos 101936-103213 de SEQ ID No. 1 y *sfaO* que son los residuos 105091-105345 de SEQ ID No. 1, se han deletado o modificado de modo que se disminuye su actividad o modificado o sustituido de modo que se altera su especificidad de sustrato, en la que la cepa codifica para la producción de sangliferina A o un análogo de la misma.
- 20
- 25

Figura 1

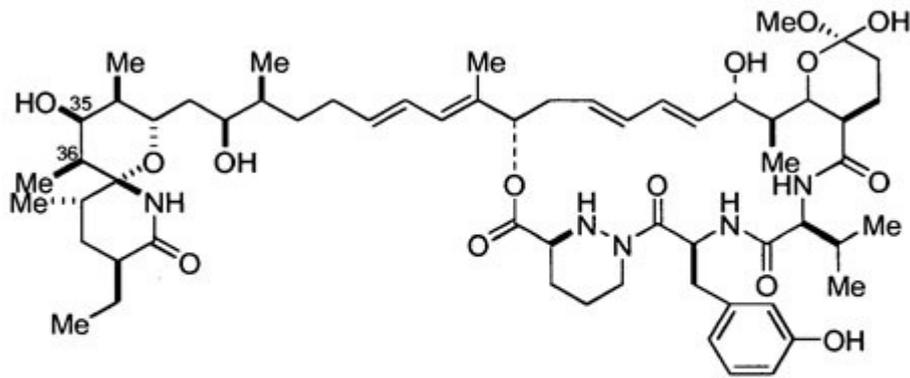
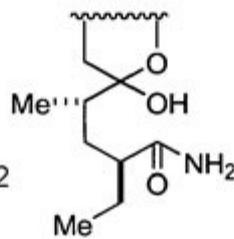


Sangliferina A

Sangliferina B C35/C36



Sangliferina E C38/C42



Sangliferina C

Sangliferina D C35/36

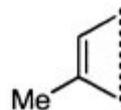


Figura 2

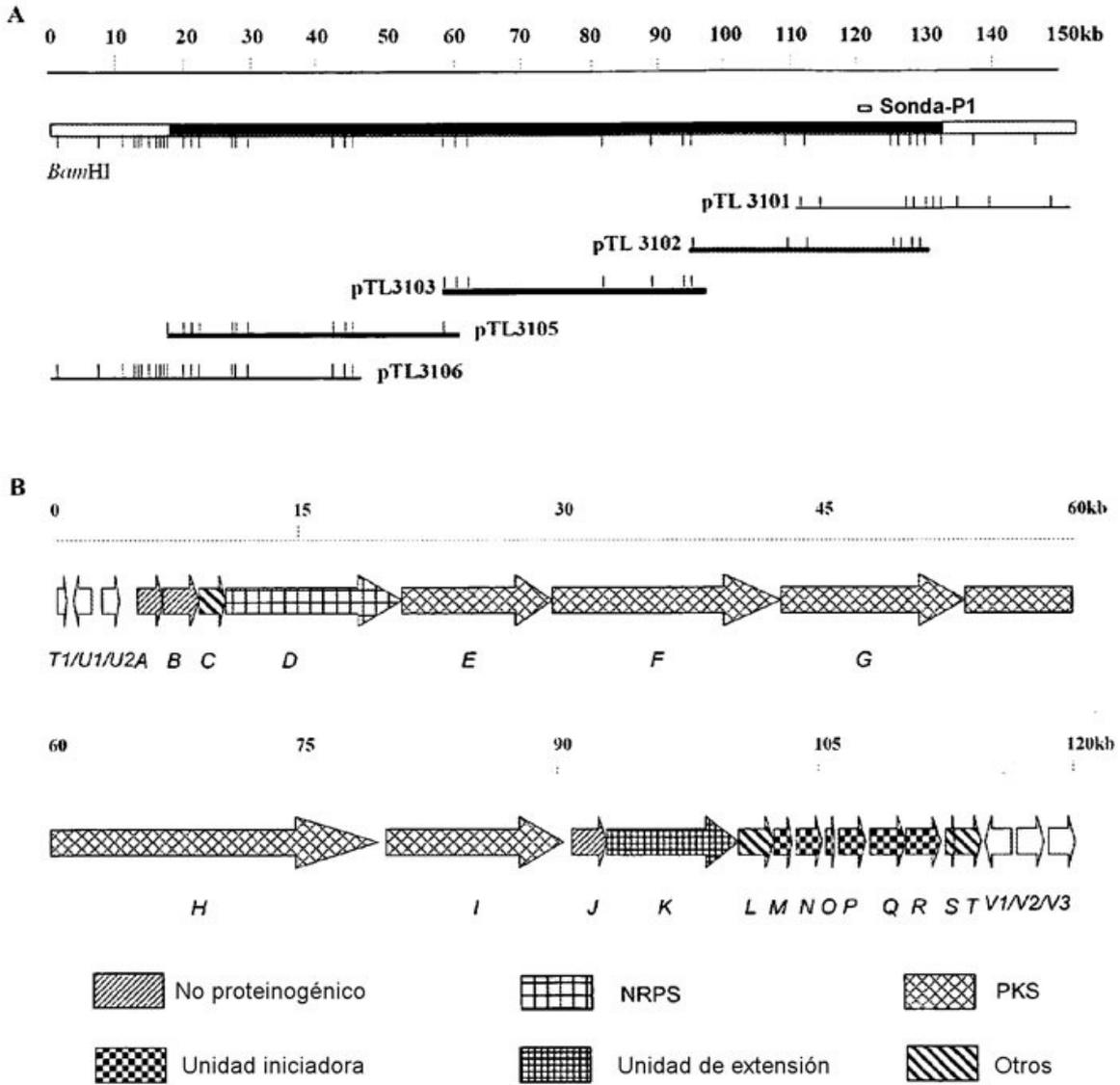


Figura 3

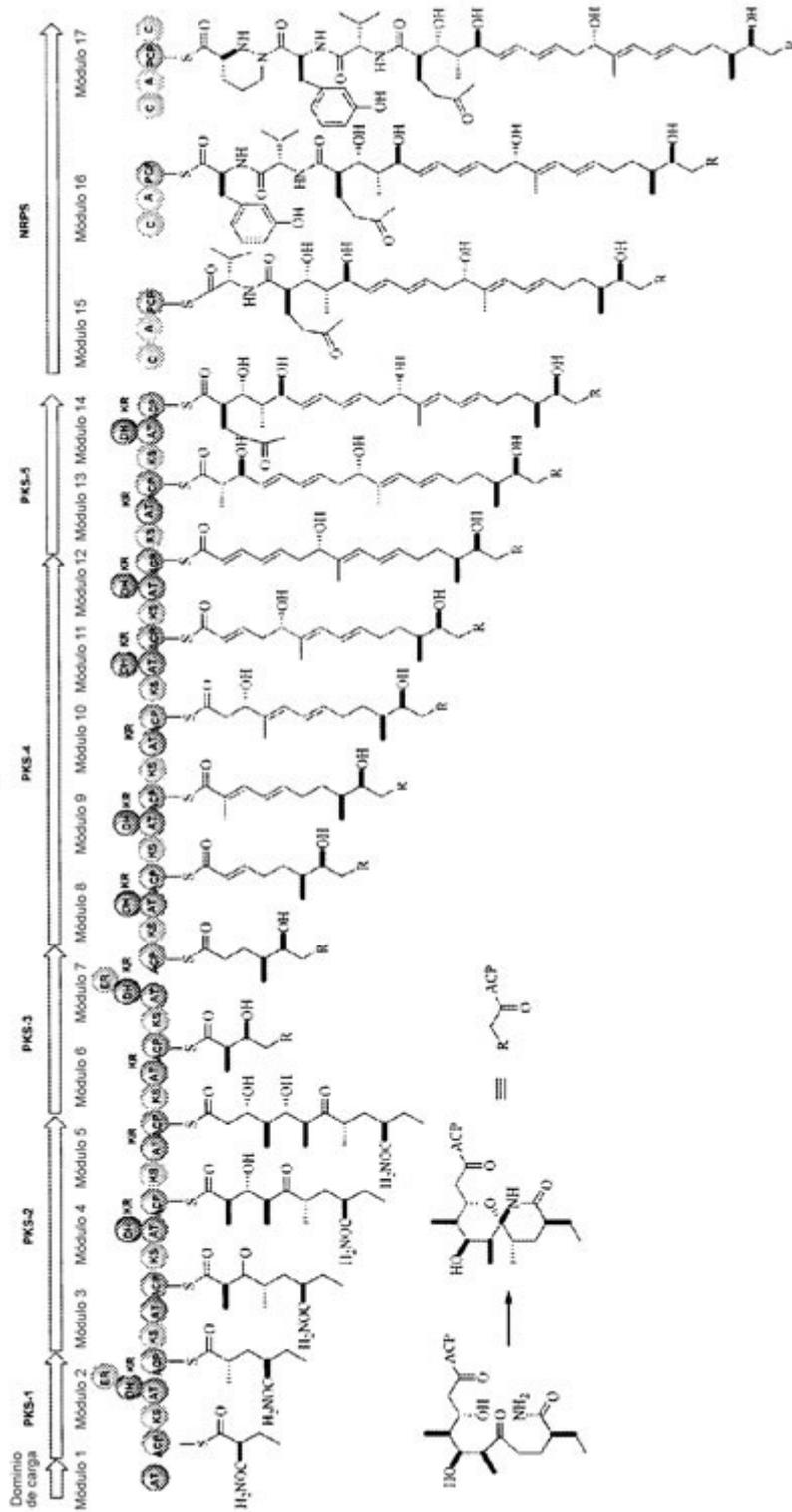


Figura 4

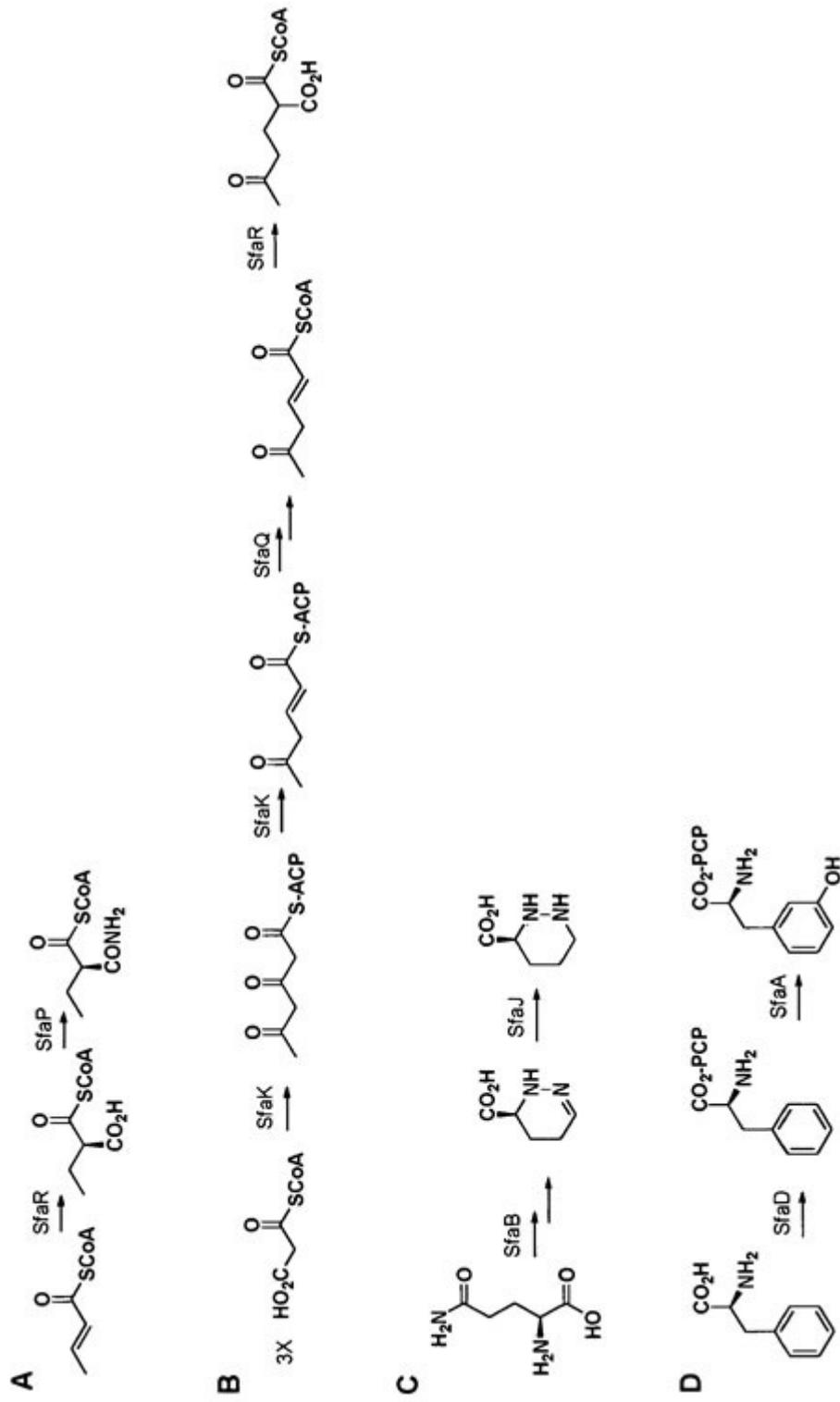


Figura 5

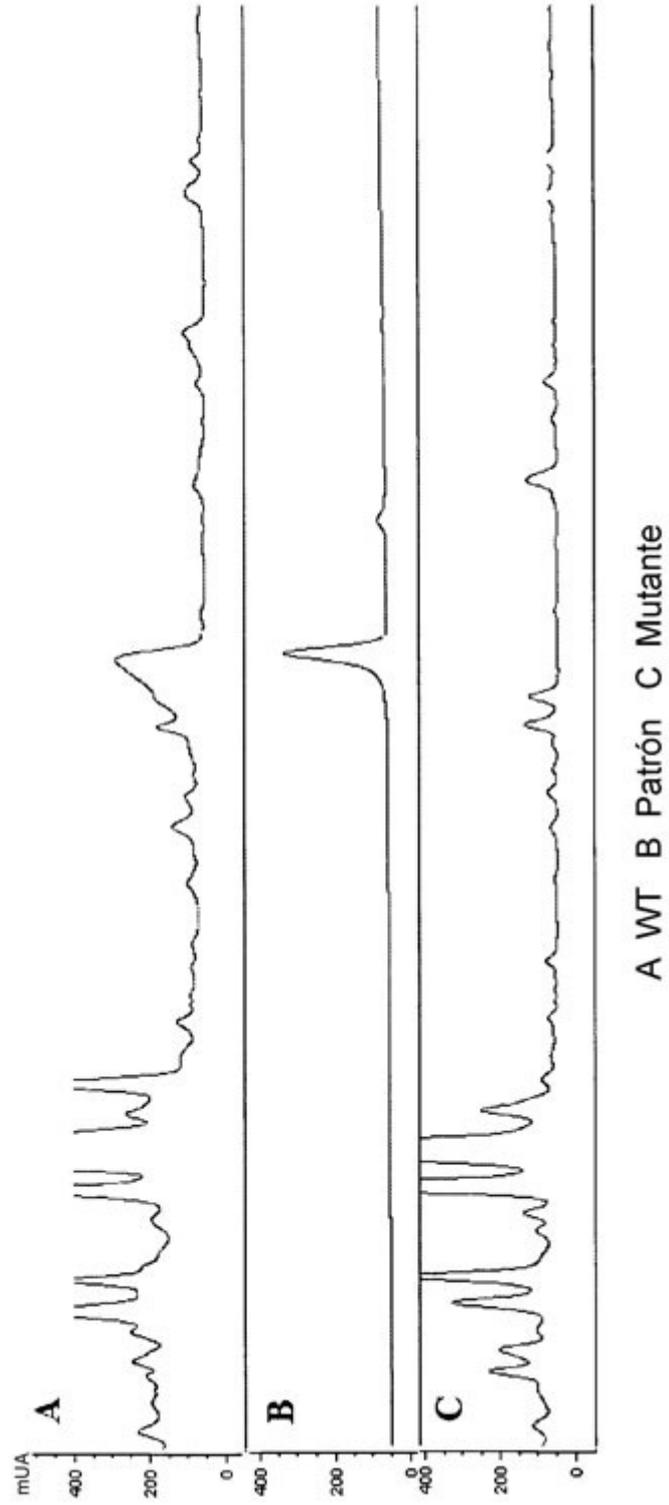


Figure 6A

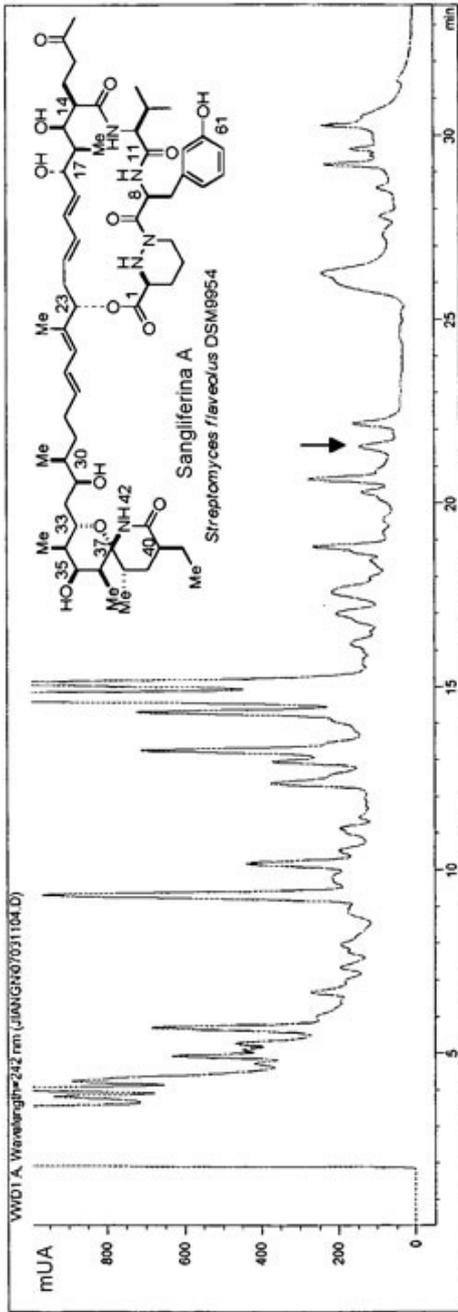


Figure 6B

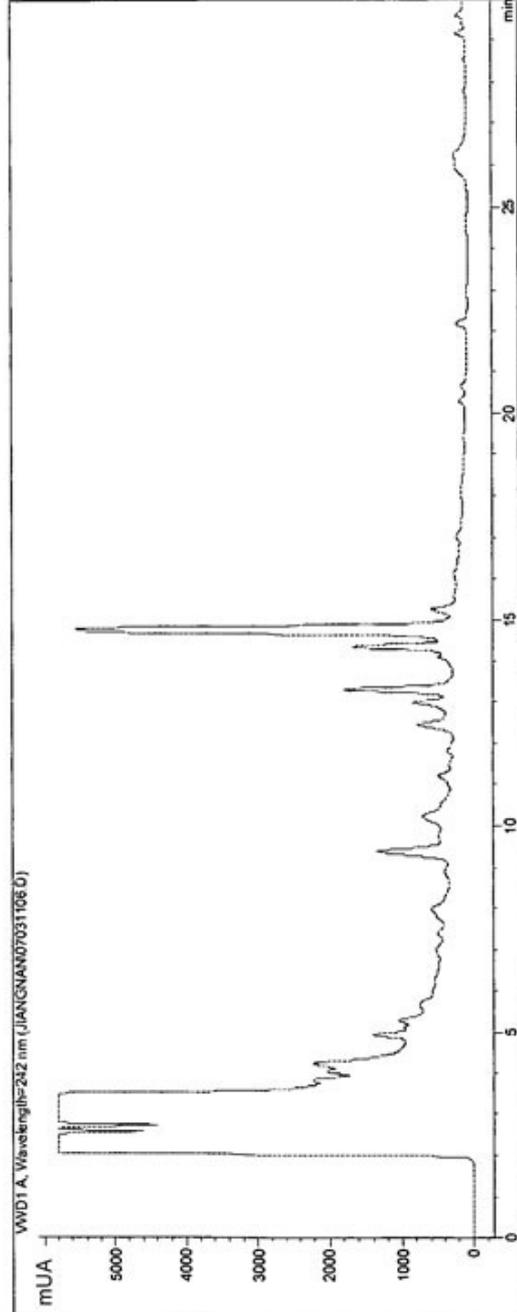


Figura 6C

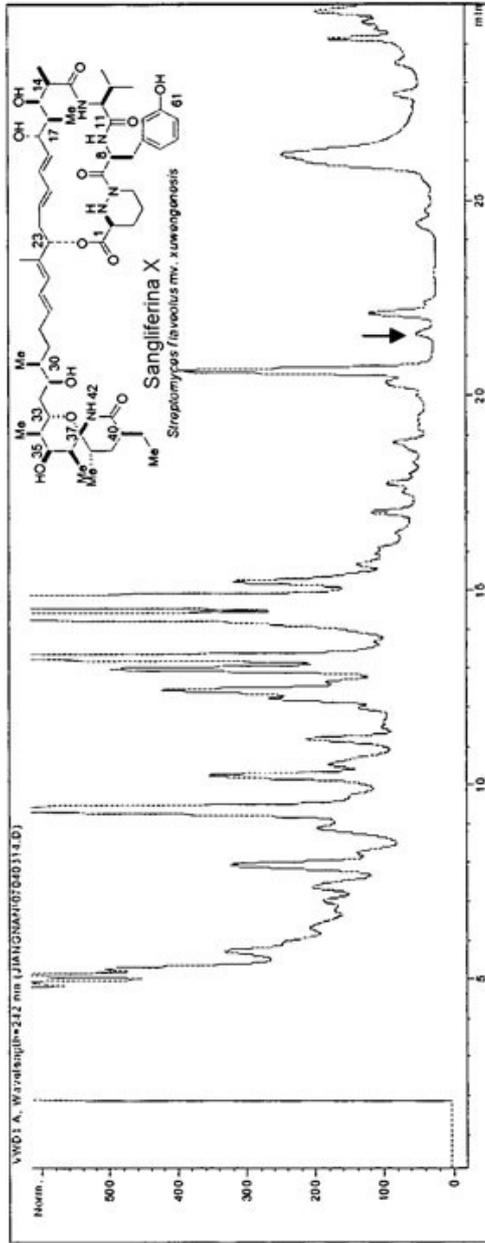


Figura 6D

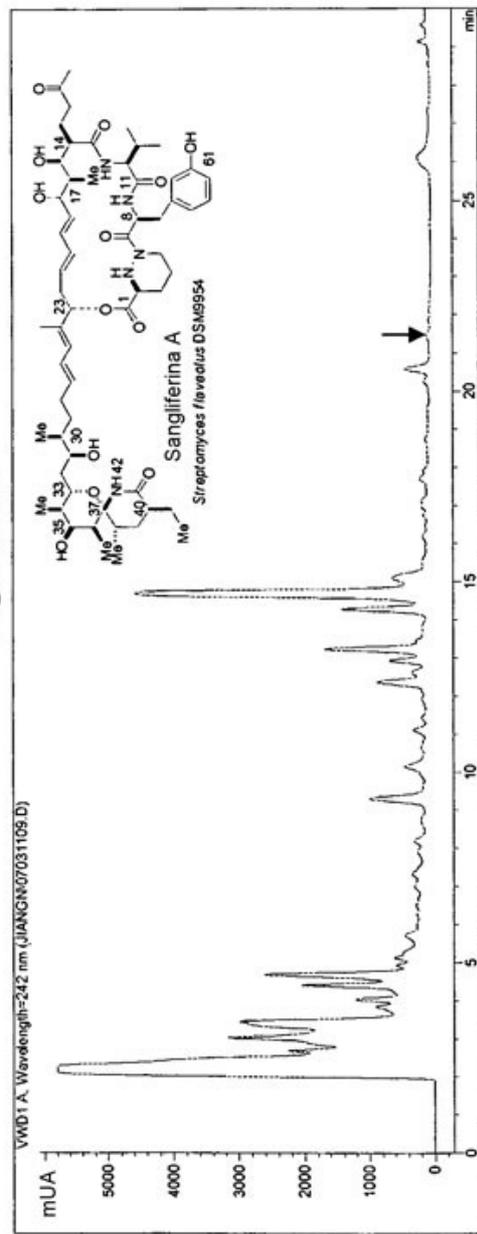


Figura 6E

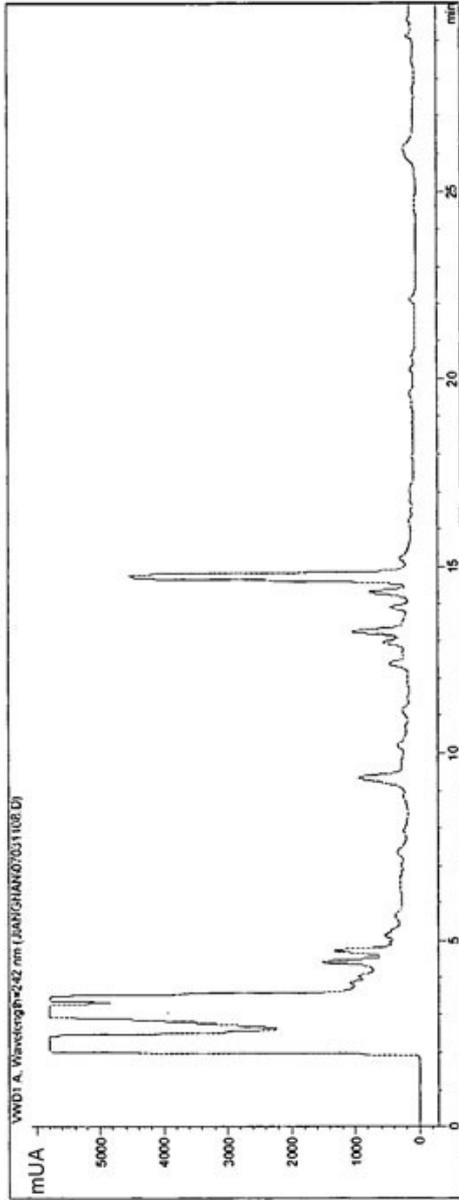


Figura 6F

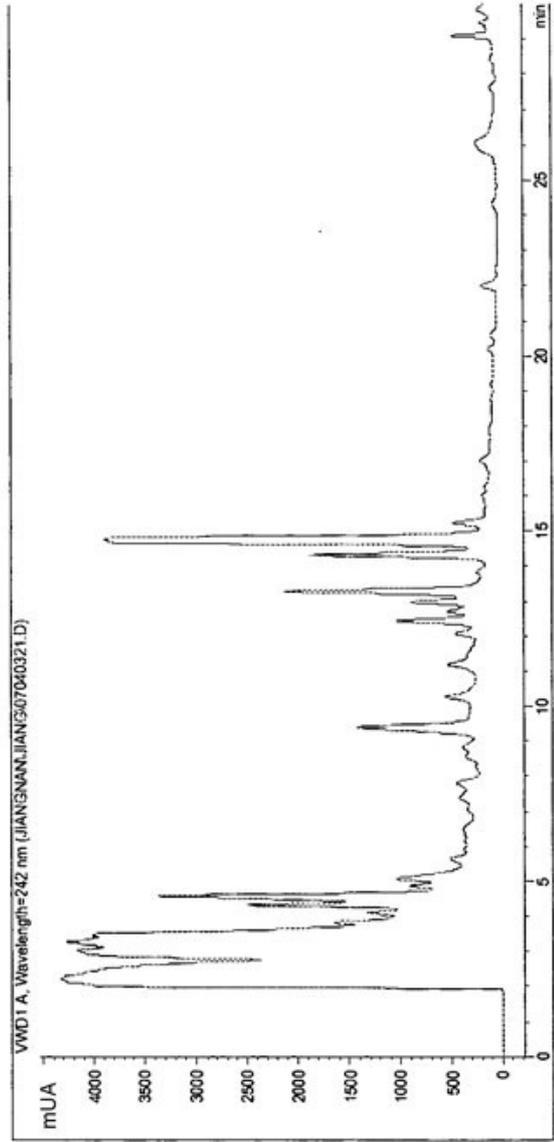


Figura 6G

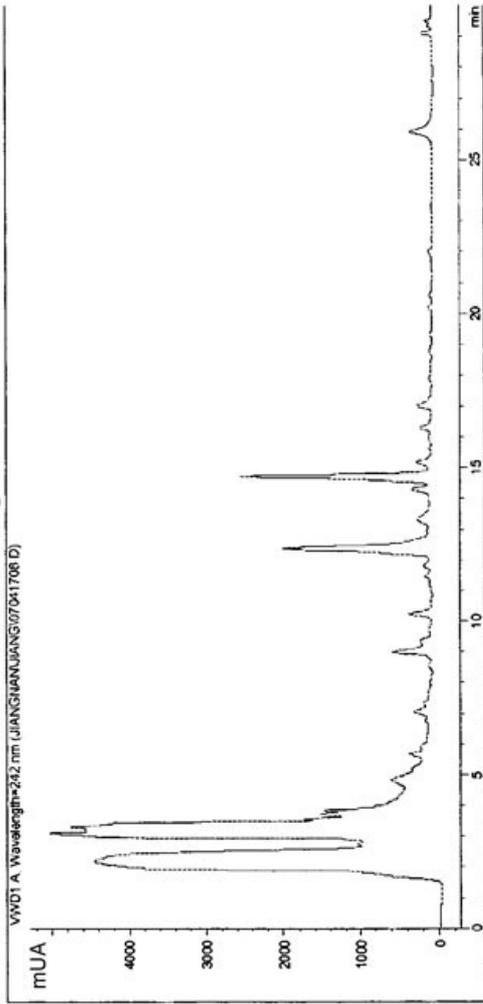


Figura 6H

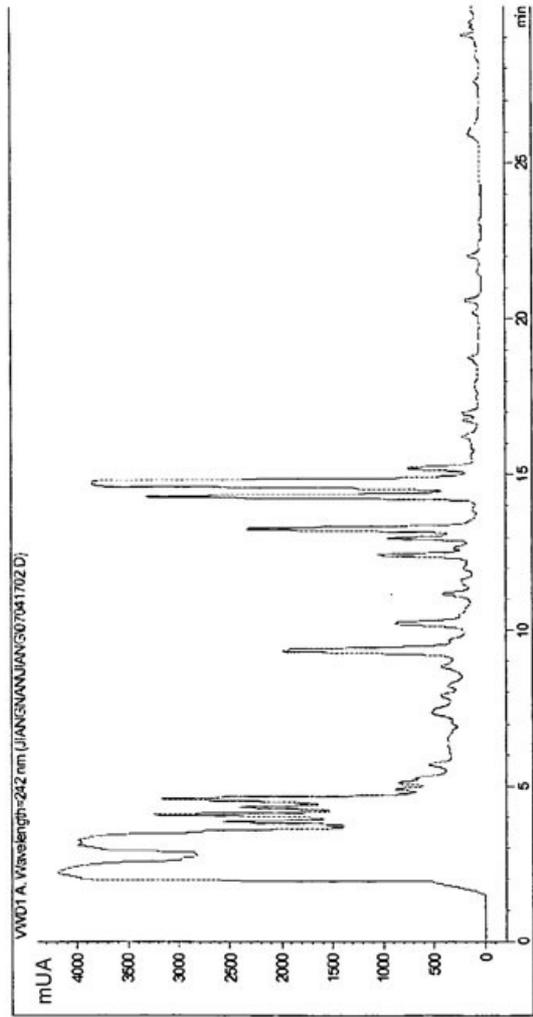


Figura 6I

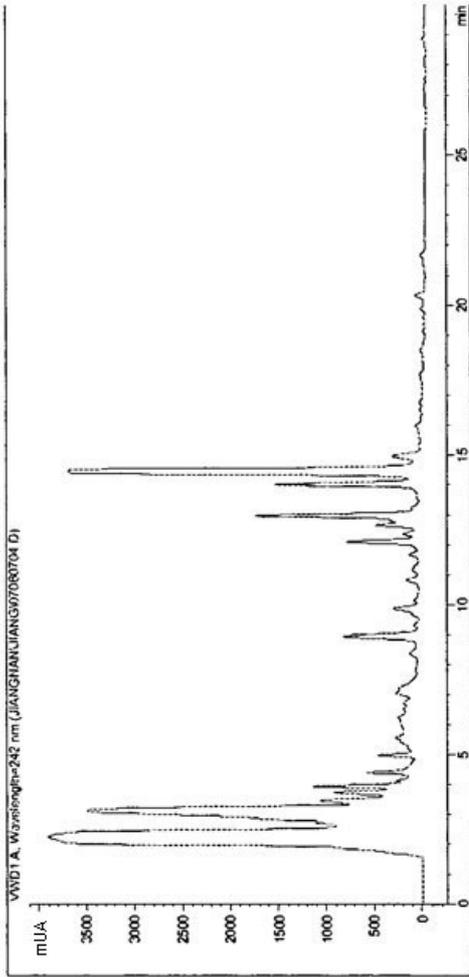


Figura 6J

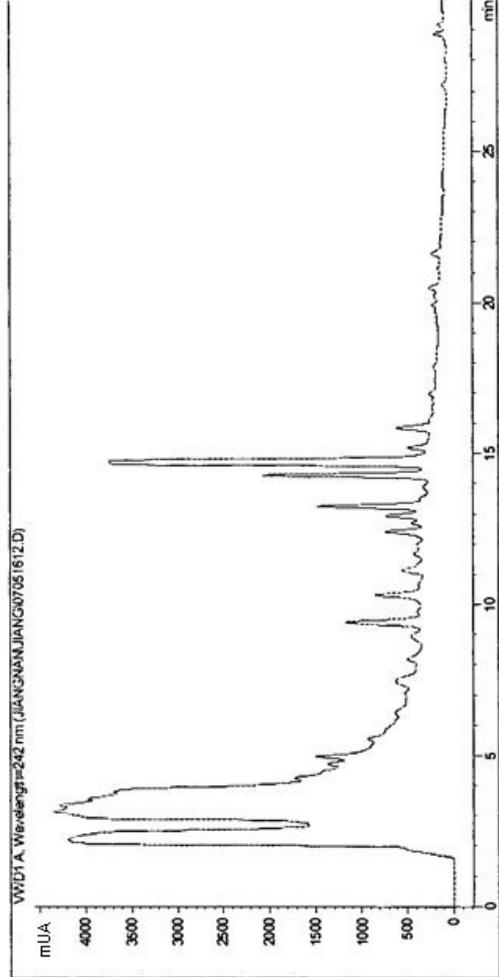


Figura 6K

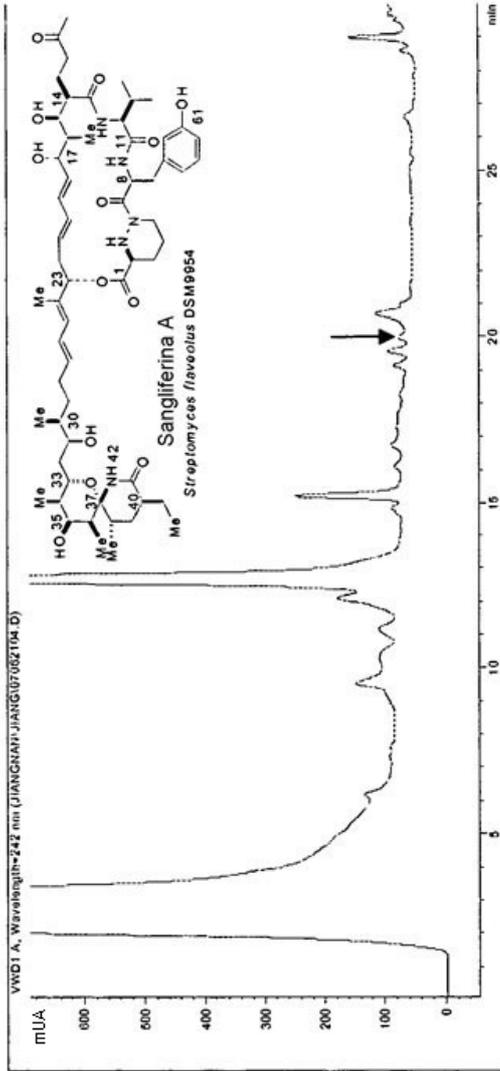


Figura 6L

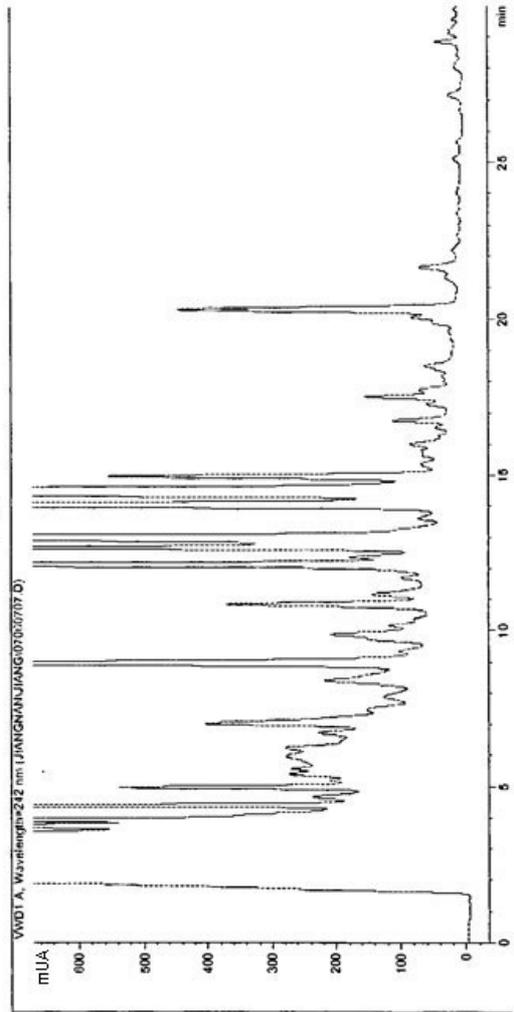


Figura 6M

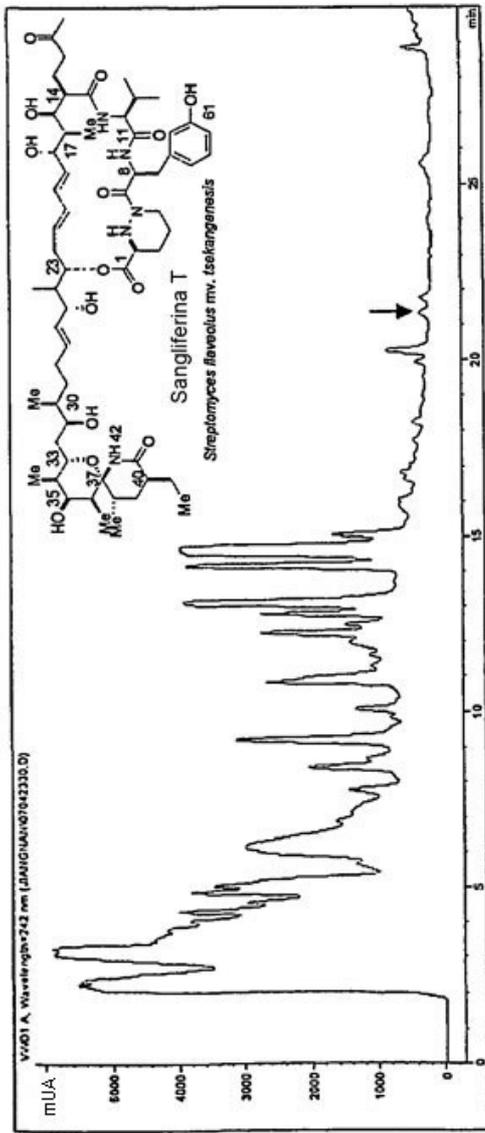


Figura 6N

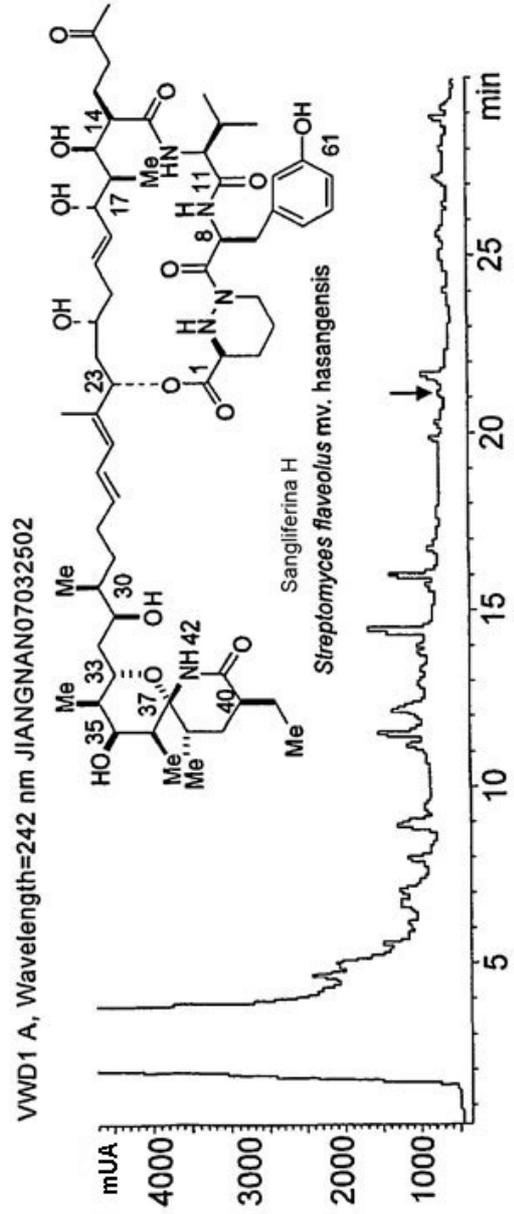


Figura 7

