



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 283 323**

51 Int. Cl.:
C12N 15/61 (2006.01)
C12N 1/21 (2006.01)
C12P 11/00 (2006.01)
C12P 19/62 (2006.01)
C12N 9/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **00972384 .2**
86 Fecha de presentación : **27.10.2000**
87 Número de publicación de la solicitud: **1224300**
87 Fecha de publicación de la solicitud: **24.07.2002**

54 Título: **Producción heteróloga de policétidos.**

30 Prioridad: **27.10.1999 US 161703 P**

45 Fecha de publicación de la mención BOPI:
01.11.2007

45 Fecha de la publicación del folleto de la patente:
01.11.2007

73 Titular/es: **Kosan Biosciences, Inc.**
3832 Bay Center Place
Hayward, California 94545, US

72 Inventor/es: **Santi, Daniel;**
Peck, Larry;
Dayem, Linda y
Kealey, James

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 283 323 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

Producción heteróloga de policétidos.

5 **Campo de la invención**

La presente invención proporciona métodos y materiales recombinantes para producir policétidos por tecnología de ADN recombinante. La invención está relacionada con los campos de la agricultura, cría de animales, química, química médica, medicina, biología molecular, farmacología, y tecnología veterinaria.

10 **Fundamentos de la invención**

Los policétidos representan una extensa familia de compuestos diversos sintetizados a partir de unidades de 2 carbonos a través de una serie de condensaciones y modificaciones subsiguientes. Los policétidos se encuentran en muchos tipos de organismos, incluyendo hongos y bacterias miceliales, en particular, los actinomicetos. Existe una amplia variedad de estructuras de policétidos, y la clase de los policétidos abarca numerosos compuestos con diversas actividades. La eritromicina, FK-506, FK-520, megalomicina, narbomicina, oleandomicina, picromicina, rapamicina, espinocina, y tilosina son ejemplos de tales compuestos. Dada la dificultad en la producción de compuestos policetídicos mediante metodología química tradicional, y la producción típicamente baja de policétidos en células salvajes, ha habido un considerable interés en encontrar medios mejorados o alternativos de producir compuestos policetídicos. Véanse las publicaciones PCT Nos. WO 93/13663; WO 95/08548; WO 96/40968; 97/02358; y 98/27203; Patentes de los EE.UU. Nos. 4.874.748; 5.063.155; 5.098.837; 5.149.639; 5.672.491; 5.712.146; y 5.962.290; y Fu *et al.*, 1994, *Biochemistry* 33: 9321-9326; McDaniel *et al.*, 1993, *Science* 262: 1546-1550; y Rohr, 1995, *Angew. Chem. Int. Ed. Engl.* 34(8): 881-888.

Los policétidos son sintetizados en la naturaleza por las enzimas policétido sintasas (PKS). Estas enzimas, que son complejos de múltiples proteínas grandes, son similares a las sintasas que catalizan la condensación de unidades de 2 carbonos en la biosíntesis de ácidos grasos. Las enzimas PKS son codificadas por genes de PKS que consisten usualmente en tres o más marcos de lectura abiertos (ORFs, del inglés "open reading frames"). Se conocen dos tipos principales de enzimas PKS; éstas difieren en su composición y modo de síntesis. Estos dos tipos principales de enzimas PKS se denominan comúnmente de Tipo I o "modulares" y de Tipo II o "iterativas". Un tercer tipo de PKS encontradas fundamentalmente en células fúngicas tiene características de ambas enzimas de Tipo I y Tipo II, y se denominan enzimas PKS "fúngicas".

Las PKS modulares son responsables de la producción de un gran número de antibióticos macrolídicos de 12, 14 y 16 miembros incluyendo eritromicina, megalomicina, metimicina, narbomicina, oleandomicina, picromicina y tilosina. Cada ORF de una PKS modular puede comprender uno, dos o más "módulos" de actividad cetosintasa, consistiendo cada módulo en al menos dos (si es un módulo de carga) y más típicamente tres (para el módulo de elongación más simple) o más actividades enzimáticas o "dominios". Estas grandes enzimas multifuncionales (>300.000 kDa) catalizan la biosíntesis de macrolactonas policetídicas a través de rutas en varios pasos que implican condensaciones descarboxilativas entre los tioésteres acílicos y a continuación ciclos de actividades de procesamiento de carbonos β variables (véase O'Hagan, D. *The polyketide metabolites*; E. Horwood: Nueva York, 1991).

Durante la pasada mitad de década, el estudio de la función y especificidad de las PKSs modulares se ha visto enormemente facilitado por el sistema de expresión en *Streptomyces coelicolor* basado en un plásmido, desarrollado con los genes de la 6-desoxieritronolida B (6-dEB) sintasa (DEBS) (véanse Kao *et al.*, 1994, *Science* 265: 509-512, McDaniel *et al.*, 1993, *Science* 262: 1546-1557, y las Patentes de los EE.UU. Nos. 5.672.491 y 5.712.146). Las ventajas de este sistema genético basado en un plásmido para DEBS son que supera las tediosas y limitadas técnicas para manipular el organismo natural hospedante de DEBS, *Saccharopolyspora erythraea*, permite una más fácil construcción de PKSs recombinantes, y reduce la complejidad del análisis de PKS proporcionando un fondo "limpio" del hospedante. Este sistema aceleró también la construcción de la primera biblioteca combinatoria de policétidos modulares en *Streptomyces* (véanse las publicaciones PCT Nos. WO 98/49315 y 00/024907).

La capacidad para controlar aspectos de la biosíntesis de policétidos, tales como la selección de monómeros y el grado de procesamiento del carbono β , por manipulación genética de PKSs, ha estimulado un gran interés en la ingeniería genética combinatoria de nuevos antibióticos (véanse Hutchinson, 1998, *Curr. Opin. Microbiol.* 1: 319-329; Carreras y Santi, 1995, *Curr. Opin. Biotech.* 9: 403-411; y las Patentes de los EE.UU. Nos. 5.712.146 y 5.672.491). Este interés ha dado como resultado la clonación, análisis y manipulación por tecnología de ADN recombinante de genes que codifican enzimas PKS. La tecnología resultante permite manipular un conjunto ("cluster") de genes de PKS conocido para producir el policétido sintetizado por esa PKS en niveles más altos que los presentes en la naturaleza, o bien en hospedantes que de otra manera no producen el policétido. La tecnología permite también producir moléculas que están estructuralmente relacionadas, pero son distintas, con los policétidos producidos por "clusters" de genes de PKS conocidos.

Ha habido un gran interés en expresar los policétidos producidos por las enzimas PKS de Tipo I y de Tipo II en células hospedantes que no expresan normalmente tales enzimas. Por ejemplo, se ha publicado la producción del policétido fúngico ácido 6-metilsalicílico (6-MSA) en células de *E. coli*, levaduras y plantas heterólogas. Véanse Kealey *et al.*, Jan. 1998, "Production of a polyketide natural product in nonpolyketide-producing prokaryotic and

eukaryotic host”, *Proc. Natl. Acad. Sci. USA* 95:505-9, Patente de los EE.UU. N° 6.033.883, y Publicaciones de Patentes PCT Nos. 98/27203 y 99/02669. La producción heteróloga de 6-MSA requirió, o aumentó considerablemente mediante la coexpresión de una proteína sintasa portadora de acilo (ACPS, del inglés “acyl carrier protein synthase”) heteróloga y que, para *E. coli*, los suplementos del medio fueran de ayuda para aumentar el nivel del sustrato malonil-CoA utilizado en la biosíntesis del 6-MSA. Véase también la Publicación de Patente PCT N° 97/13845.

Khosla *et al.*, *Annu. Rev. Biochem.* 1999 68:219-253 describen la expresión heteróloga de genes *phs* seleccionados en un hospedante heterólogo adecuado para estudiar las policétido sintasas. La producción de ciertos metabolitos en *E. coli* está impedida debido a la ausencia de ciertos precursores tales como metilmaloni-CoA. Se sugiere una solución a este problema como la de alimentar las células con compuestos miméticos sintéticos de los sustratos necesarios.

Stassi *et al.*, *Proc. Natl. Acad. Sci. USA*, Vol. 95. pp7305-7309 junio 1998, describen la producción en *Saccharopolyspora erythraea* de un derivado de eritromicina por modificación de un dominio AT en una policétido sintasa. Debido a que el sustrato necesario era limitante, se expresó crotonil-CoA reductasa en la cepa.

La biosíntesis de otros policétidos requiere sustratos distintos o adicionales a maloni-CoA. Tales sustratos incluyen, por ejemplo, propionil-CoA, 2-metilmaloni-CoA, 2-hidroximaloni-CoA y 2-etilmaloni-CoA. De las miles de células hospedantes posibles para utilizar como hospedantes productores de policétidos, muchas no producen de manera natural tales sustancias. Dado el potencial para producir policétidos valiosos y útiles en grandes cantidades en células hospedantes heterólogas, existe la necesidad de células hospedantes capaces de producir los sustratos requeridos para la biosíntesis de los policétidos. La presente invención ayuda a satisfacer esa necesidad proporcionando células hospedantes recombinantes, vectores de expresión, y métodos para producir policétidos en diversas células hospedantes.

25 Sumario de la invención

La presente invención proporciona células hospedantes recombinantes y vectores de expresión para producir productos en células hospedantes, que son de otra forma incapaces de producir esos productos debido a la falta de una ruta biosintética para producir un precursor requerido para la biosíntesis del producto. La presente invención proporciona también métodos para aumentar las cantidades de un producto producido en una célula hospedante proporcionando rutas biosintéticas recombinantes para la producción de un precursor utilizado en la biosíntesis de un producto.

En una realización, la célula hospedante no produce el precursor, y la célula hospedante se modifica por introducción de un vector de expresión recombinante de tal forma que pueda producir el precursor. En otra realización, el precursor se produce en la célula hospedante en pequeñas cantidades, y la célula hospedante se modifica por introducción de un vector de expresión recombinante de tal forma que pueda producir el precursor en mayores cantidades. En una realización preferida, el precursor es un metabolito primario que se produce en una primera célula pero no en una segunda célula heteróloga. Conforme a los métodos de la invención, los genes que codifican las enzimas que producen el metabolito primario en la primera célula se transfieren a la segunda célula. La transferencia se realiza usando un vector de expresión de la invención. El vector de expresión dirige la expresión de los genes y la producción del metabolito en la segunda célula.

En una realización preferida, el producto es un policétido. El policétido es un policétido sintetizado por una PKS modular, iterativa, o bien fúngica. El precursor se selecciona del grupo que consiste en maloni-CoA, propionil-CoA, metilmaloni-CoA, etilmaloni-CoA, e hidroximaloni-CoA o metoximaloni-CoA. En una realización especialmente preferida, el policétido utiliza metilmaloni-CoA en su biosíntesis. En una realización preferida, el policétido es sintetizado por una PKS modular que requiere metilmaloni-CoA para sintetizar el policétido.

En una realización, la célula hospedante es una célula hospedante procarionta, o bien eucariota. En una realización, la célula hospedante es una célula hospedante de *E. coli*. En otra realización, la célula hospedante es una célula hospedante de levaduras. En otra realización, la célula hospedante es una célula hospedante de Actinomycetos, incluyendo, pero sin limitarse a ella, una célula hospedante de *Streptomyces*. En otra realización, la célula hospedante es una célula hospedante de plantas. En una realización preferida, la célula hospedante es una célula hospedante de *E. Coli* o de levaduras, el producto es un policétido, y el precursor es metilmaloni CoA.

En una realización, la invención proporciona un vector de expresión recombinante que comprende un promotor situado para dirigir la expresión de uno o más genes que codifican las enzimas necesarias para la biosíntesis de un precursor. En una realización preferida, el promotor se deriva de un gen de PKS. En una realización afín, la invención proporciona células hospedantes recombinantes que comprenden uno o más vectores de expresión que dirigen la expresión de las enzimas que producen el precursor.

En otra realización, la invención proporciona una célula hospedante recombinante que comprende no sólo un vector de expresión de la invención, sino también un vector de expresión que comprende un promotor colocado para dirigir la expresión de una PKS. En una realización afín, la invención proporciona células hospedantes recombinantes que comprenden el vector que produce la PKS y su correspondiente policétido. En una realización, la célula hospedante es una célula hospedante de *E. coli* o de levaduras.

ES 2 283 323 T3

Estas y otras realizaciones de la invención se describen con más detalle en la siguiente descripción, los ejemplos, y las reivindicaciones expuestas más abajo.

Breve descripción de las figuras

La Figura 1 muestra los módulos y dominios de DEBS y la biosíntesis de 6-dEB a partir de propionil-CoA y metilmalonil-CoA.

La Figura 2 muestra la construcción de pSK-MUT, en el que cuatro fragmentos de PCR se secuenciaron y juntaron para formar el gen mutasa completo en pSK-bluescript.

La Figura 3 muestra el análisis de acil-CoAs en cepas de BL21(DE3) panD *in vivo*.

La Figura 4 muestra los resultados del análisis de CoA de *E. coli* que sobre-expresa metilmalonil-CoA mutasa. Se muestran los niveles de ³H detectados en las fracciones recogidas del análisis por HPLC de los extractos libres de células de *E. coli* alimentadas con ³H β-alanina que hospedaban el vector de control pET cultivadas sin hidroxocobalamina (línea sólida), pET cultivadas con hidroxocobalamina (línea de puntos-trazos), pET que sobreexpresa la mutasa y cultivadas sin hidroxocobalamina (línea de puntos), o bien pET que sobreexpresa la mutasa y cultivadas con hidroxocobalamina (línea de trazos).

La Figura 5 muestra las tres vías o rutas biosintéticas para la síntesis de metilmalonil-CoA que pueden introducirse por ingeniería genética en levaduras.

La Figura 6 muestra el análisis de acil-CoAs de *E. coli* que sobreexpresa metilmalonil-CoA mutasa. Se muestran los niveles de ³H detectados en las fracciones recogidas del análisis por HPLC de los extractos libres de células de *E. coli* alimentadas con ³H β-alanina que hospedaban el vector de control pET (línea sólida), o bien pET que sobreexpresa la mutasa (línea de trazos).

La Figura 7 muestra el análisis de acil-CoAs en *S. cerevisiae*. Se muestran los niveles de ³H detectados en las fracciones recogidas del análisis por HPLC de los extractos libres de células de *S. cerevisiae* alimentadas con ³H β-alanina después de crecer durante 24 horas (línea sólida), 48 horas (línea de trazos) y 66 horas (línea de puntos).

La Figura 8 muestra el Casete de Clonación Común.

La Figura 9 muestra un método general para clonar genes en vectores de expresión en levaduras.

Descripción detallada de la invención

La presente invención proporciona células hospedantes recombinantes y vectores de expresión para producir productos en células hospedantes, que son de otra forma incapaces de producir esos productos debido a la falta de una ruta biosintética para producir un precursor requerido para la biosíntesis del producto. Tal como se usa en la presente memoria, el término “recombinante” se refiere a una célula, compuesto o composición producido al menos en parte mediante intervención humana, particularmente por modificación del material genético de una célula. La presente invención proporciona también métodos para aumentar las cantidades de un producto producido en una célula hospedante proporcionando rutas biosintéticas recombinantes para la producción de un precursor utilizado en la biosíntesis de un producto.

En una realización, la célula hospedante no produce el precursor, y la célula hospedante se modifica por introducción de un vector de expresión recombinante de tal forma que pueda producir el precursor. En otra realización, el precursor se produce en la célula hospedante en pequeñas cantidades, y la célula hospedante se modifica por introducción de un vector de expresión recombinante de tal forma que pueda producir el precursor en mayores cantidades. En una realización preferida, el precursor es un metabolito primario que se produce en una primera célula pero no en una segunda célula heteróloga. Conforme a los métodos de la invención, los genes que codifican las enzimas que producen el metabolito primario en la primera célula se transfieren a la segunda célula. La transferencia se realiza usando un vector de expresión de la invención. El vector de expresión dirige la expresión de los genes y la producción del metabolito en la segunda célula.

La invención, en su forma más general, concierne a la introducción, completa o en parte, de una ruta metabólica de una célula en una célula hospedante heteróloga. La invención abarca también la modificación de una ruta metabólica existente, completa o en parte, en una célula, a través de la introducción de material genético heterólogo en la célula. En todas las realizaciones, la célula resultante es diferente con respecto a su fisiología celular y bioquímica de una manera tal que la biosíntesis, biodegradación, transporte, modificación bioquímica, o niveles de metabolitos intracelulares permiten la producción o mejoran la expresión de productos deseados. La invención se ejemplifica aumentando el nivel de policétidos producidos en un hospedante heterólogo y limitando la composición química de los productos a las estructuras deseadas.

Por consiguiente, en una realización preferida, el producto producido por la célula es un policétido. El policétido es un policétido sintetizado por una PKS modular, iterativa, o bien fúngica. El precursor se selecciona del grupo que

consiste en malonil-CoA, propionil-CoA, metilmalonil-CoA, etilmalonil-CoA, e hidroximalonil-CoA o metoximalonil-CoA. En una realización especialmente preferida, el policétido utiliza metilmalonil-CoA en su biosíntesis. En una realización preferida, el policétido es sintetizado por una PKS modular que requiere metilmalonil-CoA para sintetizar el policétido.

5

La clase de productos naturales de tipo policétido incluye miembros que tienen propiedades estructurales y farmacológicas diversas (véase Monaghan y Tkacz, 1990, *Annu. Rev. Biochem.* 44:271). Los policétidos son ensamblados por policétido sintetas a través de condensaciones sucesivas de monómeros de tioésteres de coenzima-A activados derivados de ácidos orgánicos pequeños tales como acetato, propionato y butirato. Los sitios activos requeridos para la condensación incluyen una aciltransferasa (AT), proteína transportadora de grupos acilo (ACP, del inglés “acyl carrier protein”), y β -cetoacilsintasa (KS, del inglés “ β -ketoacylsynthase”). Cada ciclo de condensación da como resultado un grupo β -ceto que experimenta todas, algunas o ninguna de una serie de actividades de procesamiento. Los sitios activos que realizan estas reacciones incluyen una ceto-reductasa (KR, del inglés “ketoreductase”), deshidratasa (DH) y enoyl-reductasa (ER). Por consiguiente, la ausencia de cualquiera de los dominios de procesamiento del grupo β -ceto da como resultado la presencia de una cetona, una KR sola da origen a un hidroxilo, una KR y una DH dan como resultado un alqueno, mientras que una combinación de KR, DH y ER conduce a la reducción completa hasta un alcano. Tras el ensamblaje de la cadena del policétido, la molécula experimenta típicamente ciclación(es) y modificación post-PKS (p.ej., glicosilación, oxidación, acilación) para lograr el compuesto activo final.

20

Los macrólidos tales como la eritromicina y megalomicina son sintetizados por PKSs modulares (véase Cane *et al.*, 1998 *Science* 282:63). A modo ilustrativo, la PKS que produce el policétido eritromicina (6-desoxieritronolida B sintasa o DEBS; véase la Patente de los EE.UU. N° 5.824.513), se muestra en la Figura 1. DEBS es el sistema de PKS modular más caracterizado y ampliamente usado. DEBS sintetiza el policétido 6-desoxieritronolida B (6-dEB) a partir de propionil-CoA y metilmalonil-CoA. En las enzimas PKS modulares tales como DEBS, los pasos enzimáticos para cada ciclo de condensación y reducción son codificados en un único “módulo” del polipéptido (esto es, un módulo distinto para cada ciclo de condensación). DEBS consiste en un módulo de carga y 6 módulos de elongación y un dominio tioesterasa (TE) de terminación de la cadena, en tres polipéptidos sumamente grandes codificados por tres marcos de lectura abiertos (ORFs, designados *eryAI*, *eryAII* y *eryAIII*).

30

Cada una de las tres subunidades polipeptídicas de DEBS (DEBSI, DEBSII y DEBSIII) contiene 2 módulos de elongación, DEBSI contiene adicionalmente el módulo de carga. Colectivamente, estas proteínas catalizan la condensación y reducción apropiada de 1 unidad iniciadora de propionil-CoA y 6 unidades de elongación de metilmalonil-CoA. Los módulos 1, 2, 5 y 6 contienen dominios KR; el módulo 4 contiene un juego completo, KR/DH/ER, de dominios reductores y deshidratasa; y el módulo 3 no contiene ningún módulo reductor funcional. Después de la condensación y reacciones de deshidratación y reducción apropiadas, el compuesto intermedio unido a la enzima es convertido en lactona por la TE al final del módulo de elongación 6 para formar 6-dEB.

35

Más particularmente, el módulo de carga de DEBS consiste en dos dominios, un dominio aciltransferasa (AT) y un dominio de proteína transportadora de grupos acilo (ACP). En otras enzimas PKS, el módulo de carga no está compuesto de un AT y un ACP, sino que utiliza en cambio un KS parcialmente inactivado, un AT y un ACP. Este KS parcialmente inactivado se denomina en la mayoría de los casos KS^Q, donde el superíndice es la abreviatura del aminoácido, glutamina, que está presente en lugar de la cisteína del sitio activo requerida para la actividad completa. El dominio AT del módulo de carga reconoce un acil-CoA en particular (propionilo para DEBS, que también puede aceptar acetilo) y lo transfiere como un tiol-éster al ACP del módulo de carga. Simultáneamente, el AT de cada uno de los módulos de elongación reconoce un elongador-CoA (metilmalonilo para DEBS) y lo transfiere al ACP de ese módulo para formar un tioéster. Una vez que la PKS se ceba con acil- y malonil-ACPs, el grupo acilo del módulo de carga migra para formar un tiol-éster (trans-esterificación) en el KS del primer módulo de elongación; en este momento, el módulo de elongación 1 posee un acil-KS y un metilmalonil-ACP. El grupo acilo derivado del módulo de carga se une entonces covalentemente al carbono alfa del grupo malonilo para formar un enlace carbono-carbono, dirigido por la concomitante descarboxilación, y generando un nuevo acil-ACP que tiene una cadena principal dos carbonos más larga que la unidad de carga (elongación o extensión). La cadena creciente del policétido se transfiere del ACP al KS del siguiente módulo, y el proceso continúa.

50

La cadena del policétido, que crece en dos carbonos en cada módulo, pasa secuencialmente como un tiol-éster unido covalentemente de módulo a módulo, en un proceso de ensamblaje de tipo lineal. La cadena de carbonos producida por este proceso solo tendría una cetona en un átomo de carbono sí y otro no, produciendo una policetona, de donde viene el nombre de policétido. Comúnmente, sin embargo, el grupo beta-ceto de cada unidad de dos carbonos se modifica justo después de haber sido añadida a la cadena creciente del policétido pero antes de ser transferida al siguiente módulo por un KR, un KR más un DH, o bien un KR, un DH y un ER. Como se señala anteriormente, los módulos pueden contener asimismo actividades enzimáticas adicionales.

55

Una vez que una cadena del policétido atraviesa el módulo de elongación final de una PKS, se encuentra con el dominio de liberación o tioesterasa que se halla en el extremo carboxílico de la mayoría de PKSs. Aquí, el policétido se escinde de la enzima y típicamente se cicla. El policétido resultante puede modificarse adicionalmente por enzimas de adaptación o modificación; estas enzimas añaden grupos carbohidrato o grupos metilo, o realizan otras modificaciones, esto es, oxidación o reducción, sobre la molécula central del policétido. Por ejemplo, los pasos finales en la conversión de 6-dEB en eritromicina A incluyen las acciones de varias enzimas modificadoras, tales como: hidroxilación en C-6,

65

ES 2 283 323 T3

unión de los azúcares micarosa y desosamina, hidroxilación en C-12 (que produce la eritromicina C), y conversión de micarosa en cladinosa por medio de *O*-metilación.

5 Con esta perspectiva general de PKS y enzimas de modificación post-PKS y sus sustratos, se puede apreciar mejor los beneficios proporcionados por la presente invención. DEBS se produce de forma natural en *Saccharopolyspora erythraea* y se ha transferido a una variedad de especies de *Streptomyces*, tales como *S. coelicolor* CH999 y *S. lividans* K4-114 y K4-155, en las que funciona sin modificación adicional de la célula hospedante para producir 6-dEB. Por consiguiente, *S. erythraea*, *S. coelicolor* y *S. lividans* producen los precursores requeridos para la síntesis de 6-dEB. Sin embargo, muchas otras células hospedantes que no son *Saccharopolyspora* o *Streptomyces* no producen todos
10 los precursores requeridos o los producen sólo en niveles suficientes para respaldar sólo cantidades muy pequeñas de biosíntesis de policétidos.

La presente invención proporciona vectores de expresión de ADN recombinante y métodos para producir un policétido y sus precursores requeridos en cualquier célula hospedante. En una realización, la célula hospedante es una
15 célula hospedante procarionta, o bien eucariota. En una realización, la célula hospedante es una célula hospedante de *E. coli*. En otra realización preferida, la célula hospedante es una célula hospedante de levaduras. En otra realización, la célula hospedante es una célula hospedante de plantas. En una realización preferida, la célula hospedante es una célula hospedante de *E. coli* o bien de levaduras, el producto es un policétido y el precursor es metilmalonil-CoA.

20 Los vectores de expresión recombinantes de la invención comprenden un promotor colocado para dirigir la expresión de uno o más genes que codifican las enzimas requeridas para la biosíntesis de un precursor. En una realización preferida, el promotor deriva de un gen de PKS. En otra realización preferida, el promotor es uno derivado de un gen de una célula hospedante o de un virus o fago que infecta normalmente la célula hospedante y es heterólogo con respecto al gen que codifica la enzima biosintética.

25 En otra realización, la invención proporciona una célula hospedante recombinante que comprende no sólo un vector de expresión de la invención, sino también un vector de expresión que comprende un promotor colocado para dirigir la expresión de una PKS. En una realización afín, la invención proporciona células hospedantes recombinantes que comprenden el vector que produce la PKS y su correspondiente policétido. En una realización, la célula hospedante es
30 una célula hospedante de *E. coli* o de levaduras.

Ni *E. coli* ni las levaduras producen metilmalonil-CoA suficiente para respaldar la biosíntesis de grandes cantidades de policétidos que requieren metilmalonil-CoA en su biosíntesis, y la mayoría de las especies no producen el sustrato metilmalonil-CoA en absoluto. En una realización, la presente invención proporciona *E. coli*, levaduras, y otras células
35 hospedantes que producen metilmalonil-CoA en cantidades suficientes para respaldar la biosíntesis del policétido. En realizaciones preferidas, las células producen cantidades suficientes de metilmalonil-CoA para respaldar la biosíntesis de policétidos que requieren metilmalonil-CoA para su biosíntesis en niveles en los intervalos de 1 µg/L a 1 mg/L, a 10 mg/L, a 100 mg/L, a 1 g/L, a 10 g/L.

40 En una realización, las células hospedantes de la invención se han modificado para expresar una metilmalonil-CoA mutasa heteróloga. Esta enzima, que convierte succinil-CoA en metilmalonil-CoA (aunque la reacción inversa está 20 veces más favorecida) se ha expresado en *E. coli* usando un gen clonado de propionibacterias pero fue inactiva debido a la falta de vitamina B12. Conforme a los métodos de la presente invención, esta enzima puede producirse en forma activa en *E. coli* y otras células hospedantes, bien expresando (constitutivamente o de otra forma) un gen
45 transportador de B12, tal como en gen endógeno de *E. coli*, y/o utilizando un medio que facilite la absorción de B12 (tal como se usa en la presente memoria, B12 puede hacer referencia al precursor hidroxocobalamina, que se convierte en B12). Aunque ciertas metilmalonil-CoA mutasas producen el isómero R, incluyendo las metilmalonil-CoA mutasas derivadas de propionibacterias, el isómero R puede convertirse en el isómero S usando una epimerasa. Por ejemplo, pueden emplearse genes de epimerasa de propionibacterias o *Streptomyces* para este propósito.

50 En otra realización, las células hospedantes de la invención se han modificado para expresar una propionil-CoA carboxilasa heteróloga que convierte propionil-CoA en metilmalonil-CoA. En esta realización, se puede aumentar adicionalmente la cantidad del precursor de metilmalonil-CoA cultivando las células en un medio suplementado con propionato. En una realización preferida, las células hospedantes son células hospedantes de *E. coli*.

55 Por consiguiente, conforme a los métodos de la invención, la producción heteróloga de ciertos policétidos en *E. coli*, levaduras, y otros organismos hospedantes requiere tanto la expresión heteróloga de una PKS deseada como también las enzimas que producen al menos algunos de las moléculas sustrato requeridas por la PKS. Estas moléculas sustrato, llamadas precursores, no se encuentran normalmente como metabolitos intracelulares en el organismo
60 hospedante o están presentes en baja abundancia. La presente invención proporciona un método para producir o modificar la composición o cantidades de metabolitos intracelulares dentro de un organismo hospedante donde tales metabolitos no están presentes de forma natural o están presentes en cantidades no óptimas.

Una realización específica de la presente invención concierne a la introducción y modificación de rutas bioquímicas
65 para la biosíntesis de metilmalonil-CoA. El metilmalonil-CoA, como se señala anteriormente, es un sustrato utilizado para la biosíntesis de policétidos por muchas policétido sintasas. Algunas de las rutas bioquímicas conocidas para la producción intracelular de metilmalonil-CoA emplean enzimas y sus correspondientes genes encontrados en ciertos organismos. Estas enzimas y genes no se han encontrado, o son de otra forma no óptimos, en otros organismos.

Estos otros organismos incluyen aquéllos que podrían por lo demás ser muy útiles como hospedantes heterólogos para la producción de policétidos. La presente invención proporciona métodos para modificar por ingeniería genética un organismo hospedante para que contenga una capacidad nueva o modificada para producir metilmalonil-CoA y/o aumentar o disminuir los niveles de metilmalonil-CoA en el hospedante.

5

Como se señala anteriormente, hay dos rutas bioquímicas que implican al metilmalonil-CoA particularmente relevantes para este aspecto de la presente invención. Estas rutas son la ruta de la metilmalonil-CoA mutasa, denominada en lo sucesivo ruta MUT, y la ruta de la propionil-CoA carboxilasa, denominada en lo sucesivo ruta PCC.

10

La ruta MUT incluye las enzimas metilmalonil-CoA mutasa (5.4.99.2, usando el sistema de numeración creado por el Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular), metilmalonil-CoA epimerasa (5.4.99.1) y malonil-CoA descarboxilasa (4.1.1.9). La ruta bioquímica incluye la conversión de succinil-CoA en (*R*)-metilmalonil-CoA a través de la acción de la metilmalonil-CoA mutasa (5.4.99.2) y a continuación la conversión de (*R*)-metilmalonil-CoA en (*S*)-metilmalonil-CoA a través de la acción de la metilmalonil-CoA epimerasa (5.4.99.1). El (*S*)-metilmalonil-CoA es un sustrato utilizado por varias policétido sintasas. La enzima malonil-CoA

15

descarboxilasa (4.1.1.9) cataliza la descarboxilación de malonil-CoA, pero se ha publicado también que cataliza la descarboxilación de (*R*)-metilmalonil-CoA para formar propionil-CoA. El propionil-CoA es un sustrato utilizado por algunas policétido sintasas.

20

La ruta PPC incluye las enzimas propionil-CoA carboxilasa (6.4.1.3) y propionil-CoA sintetasa (6.2.1.17). La ruta bioquímica incluye la conversión de propionato en propionil-CoA a través de la acción de la propionil-CoA sintetasa (6.2.1.17) y a continuación la conversión de propionil-CoA en (*S*)-metilmalonil-CoA a través de la acción de la propionil-CoA carboxilasa (6.4.1.3). El (*S*)-metilmalonil-CoA es un sustrato utilizado por muchas policétido sintasas.

25

Una realización ilustrativa de la presente invención emplea enzimas específicas de estas rutas. Como podrán reconocer los expertos en la técnica tras la contemplación de esta descripción de la invención, la invención puede ponerse en práctica usando enzimas adicionales y/o alternativas implicadas en las rutas MUT y PCC. Además, la invención puede ponerse en práctica usando rutas adicionales y alternativas para el metilmalonil-CoA y otros metabolitos intracelulares.

30

Los métodos de la invención implican la introducción de material genético en una cepa hospedante elegida para modificar o alterar la fisiología celular y bioquímica del hospedante. A través de la introducción de material genético, la cepa hospedante adquiere nuevas propiedades, p. ej., la capacidad para producir un nuevo metabolito intracelular, o mayores cantidades del mismo. En una realización ilustrativa de la invención, la introducción de material genético en la cepa hospedante da como resultado una capacidad nueva o modificada para producir metilmalonil-CoA. El material genético introducido en la cepa hospedante contiene un gen(es), o partes de genes, que codifican una o más de las enzimas implicadas en la biosíntesis/biodegradación de metilmalonil-CoA y pueden incluir también elementos adicionales para la expresión y/o regulación de la expresión de esos genes, p.ej., secuencias promotoras. Las secuencias

35

génicas específicas que codifican las enzimas implicadas en la biosíntesis/biodegradación de metilmalonil-CoA se enumeran más abajo.

40

Un gen de metilmalonil-CoA mutasa (5.4.99.2) adecuado puede aislarse de *Streptomyces cinnamonensis*. Véase Birch *et al.*, 1993, *J Bacteriol.* 175: 3511-3519, titulado "Cloning, sequencing, and expression of the gene encoding methylmalonyl-coenzyme A mutase from *Streptomyces cinnamonensis*." Esta enzima es una enzima con dos subunidades: las secuencias que codifican las subunidades A y B están disponibles bajo la entrada de GenBank L10064. Otro gen de metilmalonil-CoA mutasa adecuado puede aislarse de *Propionibacterium shermanii*. Véase Marsh *et al.*, 1989, *Biochem. J.* 260: 345-352, titulado "Cloning and structural characterization of the genes coding for adenosylcobalamin-dependent methylmalonyl CoA mutase from *Propionibacterium shermanii*." Alternativamente, puede aislarse un gen de metilmalonil-CoA mutasa adecuado de *Porphyromonas gingivalis*. Véase Jackson *et al.*, 1995, *Gene* 167: 127-132, titulado "Cloning, expression and sequence analysis of the genes encoding the heterodimeric methylmalonyl CoA mutase of *Porphyromonas gingivalis* W50." Alternativamente, pueden aislarse genes de metilmalonil-CoA mutasa adecuados de cualquiera de las fuentes señaladas en la siguiente tabla de un informe de búsqueda parcial con BLAST o de análisis adicionales con BLAST.

55

Resultados de la búsqueda con BLAST de la Base de Datos NCBI para Metilmalonil-CoA Mutasa mutA

gb|L10064|STMMUTA *Streptomyces cinnamonensis* 931 0.0 (secuencia problema)

60

gb|AD000015|MSGY175 Secuencia de *Mycobacterium tuberculosis* 300 7e-80

emb|Z79701|MTCY277 *Mycobacterium tuberculosis* H37Rv 300 7e-80

gb|AD000001|MSGY456 Secuencia de *Mycobacterium tuberculosis* 238 8e-76

65

emb|X14965|PSMUTAB *Propionibacterium shermanii* mutA 268 5e-70

gb|L30136|POYMCMAB *Porphyromonas gingivalis* 137 9e-31

ES 2 283 323 T3

gb|AE000375|AE000375 *Escherichia coli* K-12 MG1655 134 1e-29
gb|U28377|ECU28377 Genoma de *Escherichia coli* K-12; 134 1e-29
5 emb|X66836|ECSERAICI Genes serA, iciA, sbm de *E. coli* 133 1e-29
gb|AF080073|SMPCAS2 *Sinorhizobium meliloti* 130 2e-28
ref|NM_000255.1|MUT| *Homo sapiens* 113 2e-23
10 dbj|AP000006|AP000006 *Pyrococcus horikoshii* OT3 110 2e-22
emb|AJ248285.1|CNSPAX03 *Pyrococcus abyssi* 109 3e-22
15 emb|X51941|MMMMCOAM mARN de ratón 109 3e-22
gb|AE000952|AE000952 Sección 155 de *Archaeoglobus fulgidus* 104 9e-21
emb|AJ237976.1|SCO237976 Gen icmA de *Streptomyces coelicolor* 103 2e-20
20 dbj|AP000062.1|AP000062 ADN genómico de *Aeropyrum pernix* 102 3e-20
gb|U67612|SCU67612 Coenzima B12 de *Streptomyces cinnamonensis* 98 7e-19
25 gb|AE001015|AE001015 Sección 92 de *Archaeoglobus fulgidus* 97 1e-18
emb|X59424|BFOF4 Genes OF4 de *Bacillus firmus* para la unión de ATP 82 7e-14
30
mutB
gb|L10064|STMMUTA *Streptomyces cinnamonensis* 1379 0.0 (secuencia problema)
35 gb|AD000001|MSGY456 *Mycobacterium tuberculosis* 1018 0.0
emb|Z79701|MTCY277 *Mycobacterium tuberculosis* H37Rv 1017 0.0
gb|AD000015|MSGY175 Secuencia de *Mycobacterium tuberculosis* 1017 0.0
40 emb|X14965|PSMUTAB *Propionibacterium shermanii* 996 0.0
gb|L30136|POYMCMA B Metilmalonilo de *Porphyromonas gingivalis* 882 0.0
45 ref|NM_000255.1|MUT| Metilmalonil-Coenzima A de *Homo sapiens* 855 0.0
emb|X51941|MMMMCOAM mARN de ratón 32 0.0
gb|U28377|ECU28377 Genoma de *Escherichia coli* K-12 798 0.0
50 gb|AE000375|AE000375 *Escherichia coli* K-12 MG1655 798 0.0
emb|X66836|ECSERAICI Genes serA, iciA, sbm de *E. coli* 797 0.0
55 gb|AF080073|SMPCAS2 *Sinorhizobium meliloti* 782 0.0
gb|AE001015|AE001015 *Archaeoglobus fulgidus* 516 e-145
dbj|AP000062.1|AP000062 ADN genómico de *Aeropyrum pernix* 408 e-139
60 emb|AJ248285.1|CNSPAX03 Genoma completo de *Pyrococcus abyssi* 486 e-135
dbj|AP000006|AP000006 ADN genómico de *Pyrococcus horikoshii* OT3 480 e-133
65 gb|AE000952|AE000952 Sección 155 de *Archaeoglobus fulgidus* 467 e-130
emb|Z35604.1|CEZK1058 Cósmido ZK1058 de *Caenorhabditis elegans* 316 e-109

ES 2 283 323 T3

emb|AJ237976.1|SCO237976 *Streptomyces coelicolor* icmA 377 e-103

gb|U67612|SCU67612 Coenzima de *Streptomyces cinnamonensis* 372 e-101

5 emb|AL035161|SC9C7 Cósmido 9C7 de *Streptomyces coelicolor* 359 2e-97

gb|U28335|MEU28335 *Methylobacterium extorquens* 351 4e-95

gb|AF008569|AF008569 Coenzima de *Streptomyces collinus* 337 8e-91

10

gb|U65074|ECU65074 Cromosoma de *Escherichia coli* 275 3e-72

gb|M37500|HUMMUT03 Metilmalonil-CoA mutasa humana 202 3e-50

15

gb|AF178673.1|AF178673 *Streptomyces cinnamonensis* 183 1e-44

ermb|Z49936.1|CEF13B10 Cósmido F13B10 de *Caenorhabditis elegans* 138 2e-41

gb|M37499|HUMMUT02 Metilmalonil-CoA mutasa humana 112 4e-23

20

dbj|AP000001.1|AP000001 *Pyrococcus horikoshii* UT3 genómico 106 2e-21

emb|AJ248283.1|CNSPAX01 Genoma completo de *Pyrococcus abyssi* 106 2e-21

25

gb|M37503|HUMMUT06 Metilmalonil-CoA mutasa humana 101 7e-20

gb|M37508|HUMMUT11 Metilmalonil-CoA mutasa humana 86 3e-15

gb|M37509|HUMMUT12 Metilmalonil-CoA mutasa humana 80 3e-13

30

gb|M37501|HUMMUT04 Metilmalonil-CoA mutasa humana 77 2e-12

La metilmalonil-CoA mutasa requiere vitamina B12 (adenosilcobalamina) como cofactor esencial para su actividad. Una de las dificultades de expresar metilmalonil-CoA mutasa activa en un hospedante heterólogo es que el organismo hospedante puede no proporcionar suficientes, si acaso alguna, cantidades de este cofactor. El trabajo sobre la expresión de metionina sintasa, una enzima dependiente de cobalamina, en *E. coli*, un hospedante que no sintetiza cobalamina, ha demostrado que es posible expresar una enzima dependiente de cobalamina activa aumentando la tasa de transporte de cobalamina. Véase Amaratunga *et al.*, 1996, *Biochemistry* 35: 2453-2463, titulado "A synthetic module for the *metH* gene permits facile mutagenesis of the cobalamin-binding region of *Escherichia coli* methionine synthase: initial characterization of seven mutant proteins".

35

40

Los métodos de la presente invención incluyen el paso de aumentar la disponibilidad de cobalamina para la expresión heteróloga de metilmalonil-CoA mutasa activa en ciertos hospedantes, p.ej., *E. coli*. En particular, estos métodos incorporan células que crecen en un medio que contiene hidroxocobalamina y/o otros nutrientes, como se describe en Amaratunga *et al.*, *ut supra*. Otros métodos adicionales para aumentar la disponibilidad de cobalamina incluyen la expresión constitutiva y/o sobreexpresión de proteínas transportadoras de vitamina B12 y/o sus reguladores.

45

Un gen de metilmalonil-CoA epimerasa (5.1.99.1) adecuado para los propósitos de la presente invención puede aislarse de *Streptomyces coelicolor* como se publica en el lugar SC5F2A de GenBank como gen SC5F2A.13 (denominado aquí EP5) o de *S. coelicolor* como se publica en el lugar SC6A5 de GenBank como gen SC6A5.34 (denominado aquí como EP6). Véase Redenbach *et al.*, 1996, *Mol. Microbiol.* 21 (1), 77-96, titulado "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome". Hasta la fecha, no se ha llevado a cabo caracterización bioquímica alguna de las proteínas codificadas por los genes EP5 y EP6; por consiguiente, la presente invención proporciona un método para usar estos genes para proporcionar la actividad metilmalonil-CoA epimerasa a un hospedante. Que estos genes codifican proteínas con actividad metilmalonil-CoA epimerasa está respaldado por su homología con la secuencia de una 2-arilpropionil-CoA epimerasa de rata. Véanse Reichel *et al.*, 1997, *Mol. Pharmacol.* 51: 576-582, titulado "Molecular cloning and expression of a 2-arylpropionyl-coenzyme A epimerase: a key enzyme in the inversion metabolism of ibuprofen," y Shieh & Chen, 1993, *J. Biol. Chem.* 268: 3487-3493, titulado "Purification and characterization of novel "2-arylpropionyl CoA epimerases" from rat liver cytosol and mitochondria." Ambas 2-arilpropionil-CoA epimerasa de rata y metilmalonil-CoA epimerasa catalizan la misma inversión estequiométrica, pero con diferentes grupos químicos unidos.

50

55

60

La caracterización bioquímica de una enzima metilmalonil-CoA epimerasa purificada de *Propionibacterium shermanii* se ha completado. Véanse Leadlay, 1981, *Biochem. J.* 197: 413-419, titulado "Purification and characterization of methylmalonyl CoA epimerase from *Propionibacterium shermanii*", Leadlay & Fuller, 1983, *Biochem. J.* 213: 635-642, titulado "Proton transfer in methylmalonyl CoA epimerase from *Propionibacterium shermanii*: Studies with specifically tritiated (2R)-methylmalonyl CoA as substrate; Fuller & Leadlay", 1983, *Biochem. J.* 213: 643-650, titulado "Proton transfer in methylmalonyl CoA epimerase from *Propionibacterium shermanii*: The reaction of (2R)-

65

ES 2 283 323 T3

5 methylmalonyl CoA in tritiated water". La secuencia de ADN del gen que codifica esta enzima de *Propionibacterium shermanii* se proporciona en la presente invención como SEC ID N° 1 en forma aislada y recombinante y se incorpora a vectores de expresión y células hospedantes de la invención. Los genes de metilmalonil-CoA epimerasa adecuados pueden aislarse de una búsqueda con BLAST usando la secuencia de *P. shermanii* proporcionada en el Ejemplo 1, más abajo. Las epimerasas preferidas además de la epimerasa de *P. shermanii* incluyen genes identificados por homología con la secuencia de *P. shermanii* localizada en el cósmido 8F4 del proyecto de secuenciación del genoma de *S. coelicolor* y la epimerasa de *B. subtilis* descrita por Haller *et al.*, 2000, *Biochemistry* 39 (16). 4622-4629.

10 Se puede preparar también S-metilmalonil-CoA a partir de R-metilmalonil-CoA utilizando una actividad de malonil-CoA descarboxilasa A, que convierte R-metilmalonil-CoA en propionil-CoA. Como se describe anteriormente, el propionil-CoA puede convertirse seguidamente en S-metilmalonil-CoA por la propionil-CoA carboxilasa. Un gen de malonil-CoA descarboxilasa (4.1.1.9) adecuado para los propósitos de la presente invención puede aislarse de *Saccharopolyspora erythraea* como se publica en Hsieh & Kolattuduky, 1994, *J. Bacteriol.* 176: 714-724, titulado "Inhibition of erythromycin synthesis by disruption of malonyl-coenzyme A decarboxylase gene *eryM* in *Saccharopolyspora erythraea*." Alternativamente, pueden aislarse genes de malonil-CoA descarboxilasa adecuados de cualquiera de las fuentes señaladas en la siguiente tabla de informes de búsqueda con BLAST o por búsquedas con BLAST adicionales.

20 *Resultados de la Búsqueda con BLAST de la Base de Datos NCBI para Malonil-CoA Descarboxilasa Malonil-CoA Descarboxilasa (DC)*

20 gb|L05192|SERMALCOAD Malonilo de *S. erythraea* 664 0.0 (secuencia problema)
emb|AL022268|SC4H2 Cósmido 4H2 de *Streptomyces coelicolor* 128 3e-28
25 emb|Z75555|MTCY02B10 *Mycobacterium tuberculosis* H37Rv 109 1e-22
gb|AD000018|MSGY151 Secuencia de *Mycobacterium tuberculosis* 109 1e-22
gb|AF141323.1|AF141323 *Shigella flexneri* SHI-2 95 5e-18
30 emb|X76100|ECIUC Genes *iucA*, *iucB* e *iucC* plasmídicos de *E. coli* 92 3e-17
emb|AL116808.1|CNS01DGW cADN de la Cepa T4 de *Botrytis cinerea* 88 5e-16
35 gb|AF110737.1|AF110737 Cepa 2011 de *Sinorhizobium meliloti* 84 9e-15
emb|AL109846.1|SPBC17G9 Cósmido c17G9 del cromosoma II de *S. pombe* 71 7e-11
gb|L06163|PSEAAC Aminoglucósido de *Pseudomonas fluorescens* 70 1e-10
40

Un gen de propionil-CoA carboxilasa (6.4.1.3) adecuado para los propósitos de la presente invención puede aislarse de *Streptomyces coelicolor* como se publica en los lugares AF113605 (*pccB*), AF113604 (*accA2*) y AF113603 (*accA2*) de GenBank por H. C. Gramajo y colaboradores. El producto del gen de la propionil-CoA carboxilasa requiere biotina para su actividad. Si la célula hospedante no produce biotina, entonces pueden transferirse los genes para el transporte de biotina a la célula hospedante. Aunque la célula hospedante produzca o transporte biotina, la enzima biotina transferasa endógena puede no tener suficiente actividad (bien debido a restricciones de especificidad o a otras razones) para biotinilar la propionil-CoA carboxilasa a la tasa requerida para un alto nivel de síntesis del precursor. En este caso, se puede simplemente proporcionar a la célula hospedante un gen de la enzima biotina transferasa suficientemente activo, o si hay un gen de transferasa endógeno, tal como el gen *birA* en *E. coli*, se puede simplemente sobreexpresar ese gen por métodos recombinantes. Se han publicado muchos otros genes que codifican propionil-CoA carboxilasas, o acetyl-CoA carboxilasas con una especificidad de sustrato relajada que incluye propionato, y pueden usarse como fuentes para este gen, como se muestra en la siguiente tabla.

55 *Resultados de la Búsqueda con BLAST de la Base de Datos NCBI para Propionil-CoA Carboxilasa Propionil-CoA Carboxilasa (pccB)*

60 gb|AF113605.1|AF113605 Propionilo de *S. coelicolor* 1035 0.0 (secuencia problema)
emb|X92557|SEPCCBBCP *pccB*, *bcpA2* y *orfX* de *S. erythraea* 800 0.0
emb|Z92771|MTCY71 *Mycobacterium tuberculosis* H37Rv 691 0.0
dbj|AB018531|AB018531 *Corynebacterium glutamicum* dtsR1 686 0.0
65 gb|U00012|U00012 Cósmido B1308 de *Mycobacterium leprae* 686 0.0
dbj|AB018530|AB018530 Gen *dtsR* de *Corynebacterium glutamicum* 612 e-174

ES 2 283 323 T3

- gb|AE001742.1|AE001742 Sección de *Thermotoga maritima* 54 610 e-173
emb|AJ002015|PMAJ2015 *Propionigenium modestum* mmdD 589 e-167
5 dbj|AB007000|AB007000 Gen MxppcB de *Myxococcus xanthus* 588 e-166
gb|L48340|MTBKATA Catalasa de *Methylobacterium extorquens* 588 e-166
gb|AE000952|AE000952 Sección de *Archaeoglobus fulgidus* 155 572 e-162
10 dbj|AP000005|AP000005 *Pyrococcus horikoshii* OT3 genómico 570 e-161
emb|AJ248285.1|CNSPAX03 Genoma completo de *Pyrococcus abyssi* 570 e-161
15 emb|AL031124|SC1C2 Cósmido 1C2 de *Streptomyces coelicolor* 563 e-159
gb|L22208|VEIMCDC Metilmalonil CoA de *Veillonella parvula* 558 e-157
gb|AF080235|AF080235 Landomicina de *Streptomyces cyanogenus* 552 e-155
20 emb|AJ235272|RPXX03 Cepa Madrid E de *Rickettsia prowazekii* 545 e-153
dbj|AB000886|AB000886 mARN para el Propionil CoA de *Sus scrofa* 539 e-152
25 ref|NM_000532.1|PCCB| Propionil Coenzima A de *Homo sapiens* 538 e-151
emb|X73424|HSPCCBA Gen para el propionil CoA de *Homo sapiens* 538 e-151
gb|M14634|RATPCCB Propionil CoA mitocondrial de rata 535 e-150
30 gb|S67325|S67325 Subunidad beta de la pronionil CoA carboxilasa 531 e-149
gb|U56964|CELF52E4 Cósmido F52E4 de *Caenorhabditis elegans* 367 e-143
35 emb|Z99116|BSUB0013 Genoma completo de *Bacillus subtilis* 494 e-138
dbj|D84432|BACJH642 Región 283 kb del ADN de *Bacillus subtilis* 494 e-138
gb|AF042099|AF042099 *Sulfolobus metallicus* supuesto 486 e-136
40 emb|AL022076.1|MTV026 *Mycobacterium tuberculosis* H37Rv 483 e-135
gb|L04196|PRSTRANSC *Propionibacterium shermanii* 383 e-104
45 emb|AL023635.1|MLCB1243 Cósmido B1243 de *Mycobacterium leprae* 356 1e-96
emb|Z70692.1|MTCY427 *Mycobacterium tuberculosis* H37Rv 353 1e-95
gb|L78825|MSGB1723CS ADN del cósmido B1723 de *Mycobacterium leprae* 319 4e-93
50 gb|M95713|RERCOABETA *Rhodococcus erythropolis* 340 5e-92
emb|Z99113|BSUB0010 Genoma completo de *Bacillus subtilis* 325 2e-87
55 gb|U94697|CCU94697 Topoisomerasa de ADN de *Caulobacter crescentus* 270 6e-71
emb|Z95556|MTCY07A7 *Mycobacterium tuberculosis* H37Rv 253 9e-66
emb|Y07660|MTACCBC Gen accBC de *M. tuberculosis* 231 6e-59
60 emb|Z79700|MTCY10D7 *Mycobacterium tuberculosis* H37Rv 229 2e-58
dbj|AB018557.1|AB018557 Gen cyaA de *Streptomyces griseus* 228 5e-58
65 gb|U46844|MSU46844 Catalasa de *Mycobacterium smegmatis* 209 2e-52
emb|Z19555.1|CEF02A9 Cósmido F02A9 de *Caenorhabditis elegans* 105 9e-51

ES 2 283 323 T3

gb|M13573|HUMPCCB Propionil CoA carboxilasa beta humana 194 5e-48
gb|AF030576|AF030576 *Acidaminococcus fermentans* 170 9e-41
5 emb|Y13917|BSY13917 *Bacillus subtilis* ppsE, yngL, yngK 149 2e-34
emb|X69435|AFGCDA Gen GCDA para *A. fermentans* 107 1e-21
emb|Z82368|RPZ82368 Fragmento de ADN genómico de *R. prowazekii* 93 2e-17
10 gb|AF025469|CELW09B6 Cósmido W09B6 de *Caenorhabditis elegans* 78 5e-13
gb|U87980|MRU87980 Supuesto elemento IS de *Malonomonas rubra* 78 7e-13
15 gb|AE001518|AE001518 Cepa J99 de *Helicobacter pylori* 75 6e-12
gb|AE000604.1|AE000604 Sección 26695 de *Helicobacter pylori* 82 75 8e-12
gb|U89347|ACU89347 Malonato de *Acinetobacter calcoaceticus* 74 1e-11
20 emb|AL021961|ATF28A23 ADN de *Arabidopsis thaliana* 61 2e-11
gb|AE001591|AE001591 Sección de *Chlamydia pneumoniae* 73 2c-11
25 emb|Z46886|UMACCGEN Gen ACC para el acetil CoA de *U. maydis* 71 1e-10
gb|U86128|SSPCCB1 Propionil CoA carboxilasa B de *Sus scrofa* 70 2e-10
emb|AJ006497|HSA006497 Exones 11 del gen PCCB de *Homo sapiens* 70 2e-10
30 gb|AE001301|AE001301 Sección de *Chlamydia trachomatis* 28 69 5e-10
gb|U32724|U32724 Sección 39 de *Haemophilus influenzae* Rd 68 8e-10
35 gb|U04358|PSU04358 *Pseudomonas syringae* pv. *syringae* Y30 68 8e-10

Propionil-CoA carboxilasa (accA2)

40 gb|AF113604.1|AF113604 *S. coelicolor* supuesto 1101 0.0 (secuencia problema)
gb|AF113603.1|AF113603 *Streptomyces coelicolor* supuesto 1090 0.0
45 gb|AF126429.1|AF126429 *Streptomyces venezuelae* JadJ 967 0.0
emb|Z92771|MTCY71 *Mycobacterium tuberculosis* H37Rv 758 0.0
emb|X92557|SEPCCBBCP Genes pccB, bcpA2, y orfX de *S. erythraea* 753 0.0
50 emb|X92556|SEHGATABCP hgtA, bcpA1 y orf122 de *S. erythraea* 753 0.0
gb|U00012|U00012 Cósmido de *Mycobacterium leprae* B1308 746 0.0
55 emb|X63470|MLBCCPG Gen para la carboxibiotina de *M. leprae* 743 0.0
gb|U35023|CGU35023 Tiosulfato de *Corynebacterium glutamicum* 695 0.0
gb|U24659|SVU24659 Glucosa de *Streptomyces venezuelae* 599 e-170
60 gb|AE000742|AE000742 Sección de *Aquifex aeolicus* 74 413 e-113
gb|U67563|U67563 Sección 105 de *Methanococcus jannaschii* 405 e-111
65 gb|L36530|MQSPYRCARB Piruvato carboxilasa de *Aedes aegypti* 400 e-110
gb|AF132152.1|AF132152 Clon de *Drosophila melanogaster* 396 e-108

ES 2 283 323 T3

gb|L09192|MUSMPYR Piruvato carboxilasa de *Mus musculus* 393 e-107
gb|U36585|RNU36585 Piruvato carboxilasa de *Rattus norvegicus* 391 e-107
5 gb|U32314|RNU32314 Piruvato carboxilasa de *Rattus norvegicus* 391 e-107
gb|L14862|ANAACCC Biotina de 49.1 kDa de *Anabaena sp.* (PCC 7120) 388 e-106
gb|U59234|SPU59234 Biotina PCC7942 de *Synechococcus* 387 e-106
10 gb|U0464|HSU04641 mARN de piruvato carboxilasa (PC) humana 387 e-106
ref|NM_000920.1|PC| Piruvato carboxilasa (PC) de *Homo sapiens* 386 e-105
15 gb|AE001090|AE001090 Sección 17 de *Archaeoglobus fulgidus* 383 e-104
dbj|D84432|BACJH642 Región 283 kb del ADN de *Bacillus subtilis* 382 e-104
emb|Z99116|BSUB0013 Genoma completo de *Bacillus subtilis* 382 e-104
20 gb|AE000942|AE000942 *Methanobacterium thermoautotrophicum* 382 e-104
gb|S72370|S72370 Piruvato carboxilasa humana, riñón 380 e-104
25 dbj|D64001|SYCCPNC *Synechocystis sp.* PCC6803 completo 379 e-103
gb|L14612|PSEACCBC Carboxibiotina de *Pseudomonas aeruginosa* 376 e-103
gb|U32778|U32778 Sección 93 de *Haemophilus influenzae* Rd 375 e-102
30 emb|Z36087|SCYBR218C Cromosoma II de *S. cerevisiae* 374 e-102
gb|U35647|SCU35647 Piruvato de *Saccharomyces cerevisiae* 374 e-102
35 gb|J03889|YSCPCB Piruvato carboxilasa de levadura (*S. cerevisiae*) 374 e-102
gb|U90879|ATU90879 Biotina carboxilasa de *Arabidopsis thaliana* 374 e-102
emb|Z72584|SCYGL062W Cromosoma VII de *S. cerevisiae* 374 e-102
40 emb|X59890|SCPYC2G Gen PYC2 para el piruvato de *S. cerivisiae* 373 e-102
gb|AE000749|AE000749 Sección 81 de *Aquifex aeolicus* 371 e-101
45 gb|AE001286|AE001286 Sección 13 de *Chlamydia trachomatis* 370 e-101
gb|AE001604|AE001604 Sección 20 de *Chlamydia pneumoniae* 369 e-100
gb|AF007100|AF007100 Biotina carboxilasa de *Glycine max* 368 e-100
50 emb|Z95556|MTCY07A7 *Mycobacterium tuberculosis* H37Rv 367 e-100
emb|Z19549|MTBCARBCP Gen para la biotina de *M. tuberculosis* 367 e-100
55 gb|AF068249|AF068249 Biotina carboxilasa de *Glycine max* 366 1e-99
gb|L38260|TOBBCSO Acetil CoA de *Nicotiana tabacum* 363 7e-99
gb|U36245|BSU36245 Carboxibiotina de *Bacillus subtilis* 362 2e-98
60 gb|AF097728|AF097728 Piruvato de *Aspergillus terreus* 361 3e-98
emb|AJ235272|RPXX03 Cepa Madrid E de *Rickettsia prowazekii* 360 1e-97
65 dbj|D83706|D83706 ADN de *Bacillus stearothermophilus* 360 1e-97
gb|AE000744|AE000744 Sección 76 de *Aquifex aeolicus* 358 3e-97

ES 2 283 323 T3

emb|AL109846.1|SPBC17G9 Cromosoma II de *S. pombe* 356 1e-96
dbj|D78170|D78170 ADN de levadura para piruvato carboxilasa 353 1e-95
5 gb|M79446|ECOFABG Gen de la biotina carboxilasa de *Escherichia coli* 352 2e-95
gb|M83198|ECOFABEGF Carboxibiotina de *Escherichia coli* 352 2e-95
gb|AE000404|AE000404 *Escherichia coli* K-12 MG1655 352 2e-95
10 gb|U18997.1|ECOOW67 *Escherichia coli* K-12 cromosómico 352 2e-95
gb|M80458|ECOACOAC Biotina carboxilasa y biotina de *E. coli* 352 2e-95
15 gb|U51439|REU51439 Piruvato carboxilasa de *Rhizobium etli* 351 5e-95
emb|Y13917|BSY13917 ppsE, yngL, yngK de *Bacillus subtilis* 348 3e-94
emb|Z99113|BSUB0010 Genoma completo de *Bacillus subtilis* 348 3e-94
20 gb|AE001274.1|AE001274 Cromosoma 1 de *Leishmania major* 347 6e-94
gb|AF042099|AF042099 *Sulfolobus metallicus* supuesto 346 1e-93
25 emb|Z81052.1|CED2023 Cósmido D2023 de *Caenorhabditis elegans* 162 3e-92
emb|Z79700|MTCY10D7 *Mycobacterium tuberculosis* H37Rv 341 4e-92
emb|Z99111|BSUB0008 Genoma completo de *Bacillus subtilis* 340 1e-91
30 gb|U12536|ATU12536 3-metilcrotonilo de *Arabidopsis thaliana* 338 4e-91
emb|Y11106|PPPYC1 Gen PYC1 de *P. pastoris* 338 4e-91
35 gb|AE001529|AE001529 Cepa J99 de *Helicobacter pylori* 334 5e-90
gb|AE000553.1|AE000553 *Helicobacter pylori* 26695333 7e-90
emb|Y09548|CGPYC Gen pyc de *Corynebacterium glutamicum* 333 1e-89
40 gb|AF038548|AF038548 Piruvato de *Corynebacterium glutamicum* 333 1e-89
ref|NM_000282.1|PCCA| Propionil coenzima de *Homo sapiens* 333 1e-89
45 gb|M22631|RATPCOA Propionil CoA carboxilasa alfa de rata 332 2e-89
gb|U08469|GMU08469 3-metilcrotonil CoA de *Glycine max* 328 3e-88
emb|Z83018|MTCY349 *Mycobacterium tuberculosis* H37Rv 318 4e-85
50 emb|AJ243652.1|PFL243652 Gen uahA de *Pseudomonas fluorescens* 316 1e-84
emb|Z36077|SCYBR208C Cromosoma II de *S. cerevisiae* 312 2e-83
55 gb|M64926|YSCUAMD Gen de la urea amidoliasa de levadura (DUR1.2) 311 5e-83
emb|Z97025|BSZ97025 nprE, yla[A,B,C,D,E,F de *Bacillus subtilis* 300 1e-79
emb|Z81074.1|CEF32B6 Cósmido F32B6 de *Caenorhabditis elegans* 131 7e-78
60 gb|U00024|MTU00024 Cósmido tbc2 de *Mycobacterium tuberculosis* 284 7e-75
gb|AD000009|MSGY2 Secuencia de *Mycobacterium tuberculosis* 284 7e-75
65 gb|U34393|GMU34393 Acetil CoA carboxilasa de *Glycine max* 259 2e-67
gb|U49829|CELF27D9 Cósmido F27D9 de *Caenorhabditis elegans* 186 4e-59

ES 2 283 323 T3

emb|AJ010111.1|BCE010111 pycA, ctaA, ctaB de *Bacillus cereus* 208 5e-52
gb|U19183|ZMU19183 Acetil-coenzima A carboxilasa de *Zea mays* 208 5e-52
5 gb|U10187|TAU10187 *Triticum aestivum* Tam 107 206 2e-51
gb|AF029895|AF029895 Acetil-coenzima A de *Triticum aestivum* 205 5e-51
gb|J03808|RATACACA mARN de acetil-coenzima A carboxilasa de rata 204 8e-51
10 emb|X80045|OAACOAC mARN para acetil CoA carboxilasa de *O. aries* 203 1e-50
emb|X68968|HSACOAC mARN para acetil CoA de *H. sapiens* 203 2e-50
15 emb|AJ132890.1|BTA132890 mARN para acetil de *Bos taurus* 202 2e-50
gb|J03541|CHKCOACA mARN de acetil CoA carboxilasa de pollo 202 3e-50
dbj|D34630|ATHACCRNA mARN de *Arabidopsis thaliana* 199 2e-49
20 gb|L25042|ALFACCASE Acetil CoA carboxilasa de *Medicago sativa* 198 5e-49
emb|Z71631|SCYNR016C Cromosoma XIV de *S. cerevisiae* 193 2e-47
25 gb|M92156|YSCFAS3A Acetil coA de *Saccharomyces cerevisiae* 193 2e-47
emb|Z49809|SC8261X Cósmido 8261 del cromosoma XIII de *S. cerevisiae* 192 3e-47
emb|Z22558|SCHFAIGN Gen HFA1 de *S. cerevisiae* 192 3e-47
30 dbj|D78165|D78165 ADN de *Saccharomyces cerevisiae* 192 3e-47
emb|Z46886|UMACCGEN Gen ACC para el acetil CoA de *U. maydis* 190 1e-46
35 ref|NM_001093.1|ACACB| Acetil Coenzima A de *Homo sapiens* 181 5e-44

Propionil-CoA carboxilasa (accA1)

40 gb|AF113603.1|AF113603 *S. coelicolor* supuesto 1101 0.0 (secuencia problema)
gb|AF113604.1|AF113604 *Streptomyces coelicolor* supuesto 1090 0.0
45 gb|AF126429.1|AF126429 *Streptomyces venezuelae* JadJ (jadJ) 967 0.0
emb|Z92771|MTCY71 *Mycobacterium tuberculosis* H37Rv 758 0.0
emb|X92557|SBFCCBSCP Genes pccB, bcpA2, y orfX de *S. erythraea* 753 0.0
50 emb|X92556|SEHGTAACP hgtA, bcpA1 y orf122 de *S. erythraea* 753 0.0
gb|U00012|U00012 Cósmido B1308 de *Mycobacterium leprae* 745 0.0
55 emb|X63470|MLBCCPG Gen para la carboxibiotina de *M. leprae* 742 0.0
gb|U35023|CGU35023 Tiosulfato de *Corynebacterium glutamicum* 694 0.0
gb|U24659|SVU24659 Glucosa de *Streptomyces venezuelae* 596 e-169
60 gb|AE000742|AE000742 Sección 74 de *Aquifex aeolicus* 417 e-115
gb|U67563|U67563 Sección 105 de *Methanococcus jannaschii* 413 e-114
65 gb|L36530|MQSPYRCARB Piruvato carboxilasa de *Aedes aegypti* 404 e-111
gb|AF132152.1|AF132152 Clon de *Drosophila melanogaster* 400 e-110

ES 2 283 323 T3

gb|L09192|MUSMPYR Piruvato carboxilasa de *Mus musculus* 397 e-109
gb|U36585|RNU36585 Piruvato carboxilasa de *Rattus norvegicus* 395 e-108
5 gb|U32314|RNU32314 Piruvato carboxilasa de *Rattus norvegicus* 395 e-108
gb|L14862|ANAACCC Biotina de 49.1 kDa de *Anabaena sp.* (PCC 7120) 394 e-108
gb|U04641|HSU04641 mARN de piruvato carboxilasa (PC) humana 391 e-107
10 gb|U59234|SPU59234 Biotina carboxilasa de *Synechococcus* PCC7942 391 e-107
ref|NM_000920.1|PC| Piruvato carboxilasa (PC) de *Homo sapiens* 390 e-107
15 gb|AE001090|AE001090 Sección 17 de *Archaeoglobus fulgidus* 389 e-106
gb|AE000942|AE000942 *Methanobacterium thermoautotrophicum* 386 e-105
gb|S72370|S72370 Piruvato carboxilasa humana, riñón 384 e-105
20 dbj|D84432|BACJH642 Región 283kb del ADN de *Bacillus subtilis* 383 e-105
emb|Z99116|BSUB0013 Genoma completo de *Bacillus subtilis* 383 e-105
25 dbj|D64001|SYCCPNC *Synechocystis sp.* PCC6803 383 e-104
gb|U35647|SCU35647 Piruvato de *Saccharomyces cerevisiae* 382 e-104
emb|Z36087|SCYBR218C Cromosoma II de *S. cerevisiae* 382 e-104
30 emb|Z72584|SCYGL062W Cromosoma VII de *S. cerevisiae* 381 e-104
gb|J03889|YSCPCB Piruvato carboxilasa de levadura (*S. cerevisiae*) 381 e-104
35 gb|L14612|PSEACCBC Carboxibiotina de *Pseudomonas aeruginosa* 381 e-104
emb|X59890|SCPYC2G Gen PYC2 para el piruvato de *S. cerevisiae* 381 e-104
gb|U32778|U32778 Sección 93 de *Haemophilus influenzae* Rd 380 e-104
40 gb|U90879|ATU90879 Biotina carboxilasa de *Arabidopsis thaliana* 377 e-103
gb|AE000749|AE000749 Sección 81 de 109 de *Aquifex aeolicus* 377 e-103
45 gb|AE001286|AE001286 Sección 13 de *Chlamydia trachomatis* 375 e-102
gb|AE001604|AE001604 Sección 20 de *Chlamydia pneumoniae* 374 e-102
gb|AF007100|AF007100 Biotina carboxilasa de *Glycine max* 372 e-101
50 emb|Z95556|MTCY07A7 *Mycobacterium tuberculosis* H37Rv 369 e-100
emb|Z19549|MTBCARBCP Gen para la biotina de *M. tuberculosis* 369 e-100
55 gb|AF068249|AF068249 Biotina carboxilasa de *Glycine max* 369 e-100
gb|L38260|TOBBCSO Acetil CoA de *Nicotiana tabacum* 367 e-100
gb|AF097728|AF097728 Piruvato de *Aspergillus terreus* 366 1e-99
60 gb|AE000744|AE000744 Sección 76 de 109 de *Aquifex aeolicus* 364 4e-99
dbj|D83706|D83706 ADN de *Bacillus stearothermophilus* 363 7e-99
65 gb|U36245|BSU36245 Carboxibiotina de *Bacillus subtilis* 363 7e-99
emb|AL109846.1|SPBC17G9 Cromosoma II de *S. pombe* 362 2e-98

ES 2 283 323 T3

emb|AJ235272|RPXX03 Cepa Madrid E de *Rickettsia prowazekii* 361 3e-98
dbj|D78170|D78170 ADN de levadura para piruvato carboxilasa 359 2e-97
5 gb|M80458|ECOACOAC Biotina carboxilasa y biotina de *E. coli* 358 3e-97
gb|M79446|ECOFABG Gen de la biotina carboxilasa de *Escherichia coli* 358 3e-97
gb|M83198|ECOFABEGF Carboxibiotina de *Escherichia coli* 358 3e-97
10 gb|AE000404|AE000404 *Escherichia coli* K-12 MG1655 358 3e-97
gb|U18997.1|ECOUW67 *Escherichia coli* K-12 cromosómico 358 3e-97
15 gb|U51439|REU51439 Piruvato carboxilasa de *Rhizobium etli* 355 3e-96
emb|Y13917|BSY13917 ppsE, yngL, yngK de *Bacillus subtilis* 354 4e-96
emb|Z99113|BSUB0010 Genoma completo de *Bacillus subtilis* 354 4e-96
20 gb|AE001274.1|AE001274 Cromosoma 1 de *Leishmania major* 351 3e-95
gb|AF042099|AF042099 *Sulfolobus metallicus putative* 350 9e-95
25 emb|Z79700|MTCY10D7 *Mycobacterium tuberculosis* H37Rv 347 6e-94
emb|Z81052.1|CED2023 Cósmido D2023 de *Caenorhabditis elegans* 168 1e-93
emb|Y11106|PPPYC1 Gen PYC1 de *P. pastoris* 345 2e-93
30 emb|Z99111|BSUB0008 Genoma completo de *Bacillus subtilis* 343 8e-93
ref|NM_000282.1|PCCA| Propionil coenzima de *Homo sapiens* 340 6e-92
35 gb|M22631|RATPCOA Propionil CoA carboxilasa alfa de rata 340 1e-91
gb|U12536|ATU12536 3-metilcrotonilo de *Arabidopsis thaliana* 339 2e-91
emb|Y09548|CGPYC Gen pyc de *Corynebacterium glutamicum* 338 4e-91
40 gb|AF038548|AF038548 Piruvato de *Corynebacterium glutamicum* 338 4e-91
gb|AE001529|AE001529 Cepa J99 de *Helicobacter pylori* 337 5e-91
45 gb|AE000553.1|AE000553 *Helicobacter pylori* 26695 336 1e-90
gb|U08469|GMU08469 3-metilcrotonil CoA de *Glycine max* 329 2e-88
emb|AJ243652.1|PFL243652 Gen uahA de *Pseudomonas fluorescens* 323 1e-86
50 emb|Z83018|MTCY349 *Mycobacterium tuberculosis* H37Rv 321 3e-86
emb|Z36077|SCYBR208C Cromosoma II de *S. cerevisiae* 314 5e-84
55 gb|M64926|YSCUAMD Gen de urea amidoliasa (DUR1.2) de levadura 312 2e-83
emb|Z97025|BSZ97025 nprE, yla[A,B,C,D,E] de *Bacillus subtilis* 303 1e-80
emb|Z81074.1|CEF32B6 Cósmido F32B6 de *Caenorhabditis elegans* 130 1e-78
60 gb|U00024|MTU00024 Cósmido tbc2 de *Mycobacterium tuberculosis* 287 6e-76
gb|AD000009|MSGY2 Secuencia de *Mycobacterium tuberculosis* 287 6e-76
65 gb|U34393|GMU34393 Acetil CoA carboxilasa de *Glycine max* 262 3e-68
gb|U49829|CELF27D9 Cósmido F27D9 de *Caenorhabditis elegans* 190 2e-61

ES 2 283 323 T3

gb|U10187|TAU10187 *Triticum aestivum* Tam 107 213 2e-53

gb|U19183|ZMU19183 Acetil-coenzima A carboxilasa de *Zea mays* 212 3e-53

5 emb|AJ010111.1|BCE010111 *pycA*, *ctaA*, *ctaB* de *Bacillus cereus* 212 4e-53

gb|AF029895|AF029895 Acetil-coenzima de *Triticum aestivum* 209 2e-52

gb|J03808|RATACACA Acetil-coenzima A carboxilasa de rata 205 4e-51

10

emb|X80045|OAACOAC mARN para acetil CoA de *O. aries* 205 5e-51

emb|X68968|HSACOAC mARN para acetil CoA de *H. sapiens* 204 8e-51

15

dbj|D34630|ATHACCRNA mARN de *Arabidopsis thaliana* 203 1e-50

emb|AJ132890.1|BTA132890 mARN para acetil CoA de *Bos taurus* 203 1e-50

gb|J03541|CHKCOACA mARN para acetil CoA carboxilasa de pollo 203 1e-50

20

gb|L25042|ALFACCASE Acetil CoA carboxilasa de *Medicago sativa* 202 2e-50

emb|Z71631|SCYNR016C Cromosoma XIV de *S. cerevisiae* 196 1e-48

25

gb|M92156|YSCFAS3A Acetil CoA de *Saccharomyces cerevisiae* 196 1e-48

emb|Z49809|SC8261X Cósmido 8261 del cromosoma XIII de *S. cerevisiae* 195 4e-48

emb|Z22558|SCHFA1GN gen HFA1 de *S. cerevisiae* 195 4e-48

30

dbj|D78165|D78165 ADN de *Saccharomyces cerevisiae* 195 4e-48

emb|Z46886|UMACCGEN Gen ACC para el acetil coa de *U. maydis* 188 5e-46

35

gb|L20784|CCXACOAC Acetil CoA de *Cyclotella cryptica* 182 2e-44

Los expertos en la técnica reconocerán que, debido a la naturaleza degenerada del código genético, puede usarse una variedad de compuestos de ADN que difieren en sus secuencias de nucleótidos para codificar una secuencia dada de aminoácidos de la invención. En la presente memoria se hace referencia a la secuencia de ADN nativa que codifica las enzimas biosintéticas en las tablas anteriores meramente para ilustrar una realización preferida de la invención, y la invención incluye compuestos de ADN de cualquier secuencia que codifique las secuencias de aminoácidos de los policétidos y proteínas de las enzimas utilizadas en los métodos de la invención. De manera similar, un policétido puede tolerar típicamente una o más sustituciones, supresiones o inserciones de aminoácidos en su secuencia de aminoácidos sin la pérdida o pérdida significativa de una actividad deseada. La presente invención incluye tales polipéptidos con secuencias de aminoácidos alternativas, y las secuencias de aminoácidos codificadas por las secuencias de ADN mostradas en la presente memoria ilustran meramente realizaciones preferidas de la invención.

Por consiguiente, en una realización especialmente preferida, la presente invención proporciona moléculas de ADN en forma de vectores o plásmidos de expresión de ADN recombinante, como se describe con más detalle más abajo, que codifican una o más enzimas biosintéticas precursoras. Generalmente, tales vectores pueden replicarse en el citoplasma de la célula hospedante, o bien integrarse en el ADN cromosómico de la célula hospedante. En cualquier caso, el vector puede ser un vector estable (esto es, el vector sigue presente durante muchas divisiones celulares, aunque sea sólo con presión selectiva) o un vector transitorio (esto es, las células hospedantes pierden el vector gradualmente según aumentan los números de divisiones celulares). La invención proporciona moléculas de ADN en forma aislada (esto es, no pura, peso existente en una preparación en una abundancia y/o concentración no encontrada en la naturaleza) y purificada (esto es, sustancialmente libre de materiales contaminantes o sustancialmente libre de materiales con los que el ADN correspondiente se encontraría en la naturaleza).

En una realización importante, la invención proporciona métodos para la expresión heteróloga de uno o más genes biosintéticos implicados en la biosíntesis de S-metilmalonil-CoA y vectores de expresión de ADN recombinante útiles en el método. Por consiguiente, se describen vectores de expresión recombinantes que incluyen tales ácidos nucleicos. El término vector de expresión se refiere a un ácido nucleico que puede introducirse en una célula hospedante o sistema de transcripción y traducción libre de célula. Un vector de expresión puede mantenerse permanentemente o transitoriamente en una célula, bien como parte del ADN cromosómico u otro o en cualquier compartimento celular, tal como un vector de replicación en el citoplasma. Un vector de expresión comprende también un promotor que dirige la expresión de un ARN, que se traduce típicamente en un polipéptido en la célula o extracto celular. Para la traducción eficiente de un ARN en proteína, el vector de expresión contiene también típicamente una secuencia de sitio de unión a ribosoma colocada antes del extremo 5' del codón de iniciación de la secuencia codificadora del gen a expresar.

Otros elementos, tales como intensificadores, secuencias señales de secreción y uno o más genes marcadores por los cuales las células hospedantes que contienen el vector puedan identificarse y/o seleccionarse, pueden también estar presentes en un vector de expresión. Se prefieren marcadores seleccionables, esto es, genes que confieren resistencia o sensibilidad a antibiótico, y confieren un fenotipo seleccionable a las células transformadas cuando las células se cultivan en un medio selectivo apropiado.

Los diversos componentes de un vector de expresión pueden variar ampliamente, dependiendo del uso que se pretende del vector y de la(s) célula(s) hospedante(s) en la(s) que se pretende que el vector se replique o dirija la expresión. Los componentes de vectores de expresión adecuados para la expresión de genes y mantenimiento de vectores en *E. coli*, levaduras, *Streptomyces*, y otras células comúnmente usadas son ampliamente conocidos y comercialmente disponibles. Por ejemplo, los promotores adecuados para incluir en los vectores de expresión de la invención incluyen aquellos que funcionan en células hospedantes eucariotas y procariotas. Los promotores pueden comprender secuencias reguladoras que permitan la regulación de la expresión con relación al crecimiento de la célula hospedante o que causen que la expresión de un gen se active o desactive en respuesta a un estímulo químico o físico. Para *E. coli* y algunas otras células hospedantes bacterianas, pueden usarse promotores derivados de genes de enzimas biosintéticas, enzimas que confieren resistencia a antibióticos, y proteínas de fagos e incluyen, por ejemplo, los promotores de galactosa, lactosa (*lac*), maltosa, triptófano (*trp*), beta-lactamasa (*bla*), bacteriófago lambda PL, y T5. Además, pueden usarse también promotores sintéticos, tales como el promotor *tac* (Patente de los EE.UU. N° 4.551.433). Para los vectores de expresión en *E. coli*, resulta útil incluir un origen de replicación de *E. coli*, tal como el de pUC, p1P, p11 y pBR.

Por consiguiente, los vectores de expresión recombinantes contienen al menos un sistema de expresión, que, a su vez, está compuesto por al menos una porción de secuencias codificadoras de PKS y/o de otros genes biosintéticos unidos de manera operativa a un promotor y opcionalmente secuencias de terminación que funcionen para producir la expresión de la secuencia codificadora en células hospedantes compatibles. Las células hospedantes se modifican por transformación con los vectores de expresión de ADN recombinante de la invención para contener las secuencias del sistema de expresión como elementos extracromosómicos, o bien integrados en el cromosoma. Las células hospedantes resultantes de la invención son útiles para producir enzimas PKS así como policétidos y antibióticos u otros compuestos útiles derivados de la invención.

Las células hospedantes preferidas a los efectos de seleccionar componentes del vector para los vectores de expresión de la presente invención incluyen células hospedantes fúngicas tales como de levaduras y células hospedantes procariotas tales como *E. coli*, pero pueden usarse también células hospedantes de mamíferos. En hospedantes tales como levaduras, plantas o células de mamífero que no producen policétidos de manera ordinaria, puede ser necesario proporcionar, también típicamente por medios recombinantes, holo-ACP sintetas adecuadas para convertir la PKS producida de manera recombinante en funcional. La provisión de tales enzimas se describe, por ejemplo, en las publicaciones PCT Nos. WO 97/13845 y 98/27203.

Las células hospedantes recombinantes de la invención pueden expresar todos los genes biosintéticos de los policétidos o sólo un subgrupo de los mismos. Por ejemplo, si se expresan sólo los genes para una PKS en una célula hospedante que no produce por lo demás enzimas modificadoras de policétidos (tales como enzimas de hidroxilación, epoxidación o glicosilación) que puedan actuar en el policétido producido, entonces la célula hospedante produce policétidos no modificados, llamados agliconas macrolídicas. Tales agliconas macrolídicas pueden hidroxilarse y glicosilarse añadiéndolas a la fermentación de una cepa tal como, por ejemplo, *Streptomyces antibioticus* o *Saccharopolyspora erythraea*, que contiene las enzimas de modificación requeridas.

Existe una amplia variedad de diversos organismos que pueden modificar agliconas macrolídicas para proporcionar compuestos con, o que pueden modificarse fácilmente para tener, actividades útiles. Por ejemplo, *Saccharopolyspora erythraea* puede convertir 6-dEB en una variedad de compuestos útiles. La eritronolida 6-dEB es convertida por el producto del gen *eryF* en eritronolida B, que es, a su vez, glicosilada por el producto del gen *eryB* para obtener 3-O-micarosileritronolida B, que contiene L-micarosa en C-3. El producto enzimático del gen *eriC* convierte seguidamente este compuesto en eritromicina D por glicosilación con D-desosamina en C-5. La eritromicina D, por lo tanto, difiere de la 6-dEB en la glicosilación y en la adición de un grupo hidroxilo en C-6. La eritromicina D puede convertirse en eritromicina B en una reacción catalizada por el producto del gen *eryG* metilando del resto de L-micarosa en C-3. La eritromicina D se convierte en eritromicina B por la adición de un grupo hidroxilo en C-12 en una reacción catalizada por el producto del gen *eriK*. La eritromicina A se obtiene de la eritromicina C por metilación del resto de micarosa en una reacción catalizada por el producto del gen *eryG*. Los policétidos no modificados proporcionados por la presente invención, tales como, por ejemplo, 6-dEB producido en *E. coli*, pueden proporcionarse a cultivos de *S. erythraea* y convertirse a los derivados correspondientes de eritromicina A, B, C y D, conforme al procedimiento proporcionado en los ejemplos más abajo. Para asegurar que sólo se produce el compuesto deseado, se puede usar un mutante en *eryA* de *S. erythraea* que es incapaz de producir 6-dEB pero que todavía puede llevar a cabo las conversiones deseadas (Weber *et al.*, 1985, J. Bacterial. 164(1): 425-433). También, se pueden emplear otras cepas mutantes, tales como mutantes en *eryB*, *eriC*, *eryG* y/o *eriK*, o cepas mutantes que tienen mutaciones en múltiples genes, para acumular un compuesto preferido. La conversión puede llevarse a cabo también en grandes fermentadores para producción comercial.

Además, hay otros organismos útiles que pueden emplearse para hidroxilar y/o glicosilar los compuestos de la invención. Como se describe anteriormente, los organismos pueden ser mutantes incapaces de producir el policétido producido normalmente en ese organismo, la fermentación puede llevarse a cabo en placas o en grandes fermentado-

res, y los compuestos producidos pueden alterarse químicamente tras la fermentación. De este modo, *Streptomyces venezuelae*, que produce picromicina, contiene enzimas que pueden transferir un grupo desosaminilo al hidroxilo en C-5 y un grupo hidroxilo a la posición C-12. Además, *S. venezuelae* contiene una actividad de glucosilación que glucosila un grupo hidroxilo 2' del azúcar desosamina. Esta última modificación reduce la actividad antibiótica, pero el resto glucosilo se elimina por acción enzimática antes de la liberación del policétido de la célula. Otro organismo, *S. narbonensis*, contiene las mismas enzimas de modificación que *S. venezuelae*, excepto la hidroxilasa en C-12. Por consiguiente, la presente invención proporciona los compuestos producidos por hidroxilación y glucosilación de las agliconas macrolídicas de la invención por acción de las enzimas endógenas de *S. narbonensis* y *S. venezuelae*.

Otros organismos adecuados para producir los compuestos de la invención incluyen *Micromonospora megalomicea*, *Streptomyces antibioticus*, *S. fradiae* y *S. thermotolerans*. *M. megalomicea* glucosila el hidroxilo C-3 con micarosa, el hidroxilo C-5 con desosamina, y el hidroxilo C-6 con megosamina, e hidroxila la posición C-6. *S. antibioticus* produce oleandomicina y contiene enzimas que hidroxilan las posiciones C-6 y C-12, glucosilan el hidroxilo C-3 con oleandrosa y el hidroxilo C-5 con desosamina, y forman un epóxido en C-8-C-8a. *S. fradiae* contiene enzimas que glucosilan el hidroxilo C-5 con micaminosa y seguidamente el hidroxilo 4' de la micaminosa con micarosa, formando un disacárido. *S. thermotolerans* contiene las mismas actividades que *S. fradiae*, así como actividades de acilación. Por consiguiente, la presente invención proporciona los compuestos producidos por hidroxilación y glucosilación de las agliconas macrolídicas de la invención por acción de las enzimas endógenas de *M. megalomicea*, *S. antibioticus*, *S. fradiae* y *S. thermotolerans*.

La presente invención proporciona también métodos y construcciones genéticas para producir los compuestos glucosilados y/o hidroxilados de la invención directamente en la célula hospedante de interés. Por consiguiente, los genes que codifican las enzimas de modificación de policétidos pueden incluirse en las células hospedantes de la invención. La falta de resistencia adecuada a un policétido puede superarse proporcionando a la célula hospedante un gen de resistencia a MLS (*ermE* y *mgf/lrm*, por ejemplo), que confiere resistencia a varios macrólidos de 14 miembros (véanse Cundliffe, 1989, *Annu. Rev. Microbiol.* 43: 207-33; Jenkins y Cundliffe, 1991, *Gene* 108: 55-62; y Cundliffe, 1992, *Gene* 115: 75-84).

Las células hospedantes recombinantes de la invención pueden usarse para producir policétidos (tanto agliconas macrolídicas como sus derivados modificados) que existen en la naturaleza o se producen por tecnología de ADN recombinante. En una realización importante, las células hospedantes recombinantes de la invención se usan para producir enzimas PKS híbridas. A los efectos de la invención, una PKS híbrida es una PKS recombinante que comprende todo o parte de uno o más módulos de elongación, módulo de carga y/o dominio tioesterasa/ciclasa de una primera PKS y todo o parte de uno o más módulos de elongación, módulo de carga y/o dominio tioesterasa/ciclasa de una segunda PKS.

Los expertos en la técnica reconocerán que todo o parte de la primera, o bien de la segunda PKS en una PKS híbrida de la invención no necesita aislarse de una fuente existente en la naturaleza. Por ejemplo, sólo una pequeña porción de un dominio AT determina su especificidad. Véanse la Solicitud de Patente PCT N° WO US99/15047 (publicación PCT N° WO 00/08138), y Lau *et al.*, más abajo. El estado de la técnica en síntesis de ADN permite al técnico construir compuestos de ADN de *novo* de tamaño suficiente para construir una porción útil de un módulo o dominio de una PKS. Por consiguiente, las secuencias codificadoras de derivación deseadas pueden sintetizarse usando métodos estándar de síntesis en fase sólida tales como los descritos por Jaye *et al.*, 1984 *J. Biol. Chem.* 259: 6331, y los instrumentos para la síntesis automatizada están comercialmente disponibles en, por ejemplo, Applied Biosystems, Inc. A los efectos de la invención, se considera que tales compuestos de ADN sintéticos son una porción de una PKS.

Una PKS híbrida a los efectos de la presente invención puede resultar no sólo:

(i) de fusiones de secuencias codificadoras de dominios heterólogos (donde heterólogo significa que los dominios en un módulo se derivan de al menos dos módulos diferentes existentes en la naturaleza) para producir una secuencia codificadora de un módulo híbrido contenida en un gen de PKS cuyo producto se incorpora a una PKS,

sino también:

(ii) de fusiones de secuencias codificadoras de módulos heterólogos (donde módulo heterólogo significa dos módulos adyacentes uno del otro que no están adyacentes uno del otro en enzimas PKS existentes en la naturaleza) para producir una secuencia codificadora híbrida contenida en un gen de PKS cuyo producto se incorpora a una PKS,

(iii) de la expresión de uno o más genes de PKS de un primer "cluster" de genes de PKS con uno o más genes de PKS de un segundo "cluster" de genes de PKS, y

(iv) de combinaciones de los anteriores.

En la presente memoria se describen varias PKSs híbridas de la invención que ilustran estas diversas alternativas.

Los métodos recombinantes para manipular genes de PKS modulares para producir enzimas PKS híbridas se describen en las Patentes de los EE.UU. Nos. 5.672.491; 5.843.718; 5.830.750; y 5.712.146; y en las publicaciones PCT Nos. 98/49315 y 97/02358. Se han usado varias estrategias de ingeniería genética con DEBS para demostrar que

ES 2 283 323 T3

las estructuras de los policétidos pueden manipularse para producir productos naturales novedosos, principalmente análogos de las eritromicinas (véanse las publicaciones de patentes referenciadas *ut supra* y Hutchinson, 1998, *Curr. Opin. Microbiol.* 1: 319-329, y Baltz, 1998, *Trends Microbiol.* 6: 76-83).

5 Estas técnicas incluyen: (i) supresión o inserción de módulos para controlar la longitud de la cadena, (ii) inactivación de los dominios de reducción/deshidratación para evitar pasos de procesamiento de carbonos beta, (iii) sustitución de dominios AT para alterar unidades iniciadoras y de elongación, (iv) adición de dominios de reducción/deshidratación para introducir actividades catalíticas, y (v) sustitución de dominios cetorreductasa KR para controlar la estereoquímica de los hidroxilos. Además, se han usado mutantes bloqueados creados por ingeniería genética
10 de DEBS para la biosíntesis dirigida por precursores de análogos que incorporan unidades iniciadoras derivadas sintéticamente. Por ejemplo, se produjeron más de 100 policétidos novedosos creando por ingeniería genética cambios únicos y combinados en múltiples módulos de DEBS. Se construyeron enzimas PKS híbridas basadas en DEBS por mutagénesis de casetes, en las que varios dominios de DEBS se remplazaron con dominios de la PKS de rapamicina (véase Schweke *et al.*, 1995, *Proc. Nat. Acad. Sci. USA* 92: 7839-7843), o se suprimió un más de los dominios KR de
15 DEBS. Se combinaron remplazamientos o supresiones de dominios únicos funcionales para generar enzimas DEBS con sustituciones de dominios catalíticos dobles y triples (véase McDaniel *et al.*, 1999, *Proc. Nat. Acad. Sci. USA* 96: 1846-1851).

Los métodos para generar bibliotecas de policétidos han mejorado enormemente mediante la clonación de genes
20 de PKS como un juego de tres o más plásmidos mutuamente seleccionables, portando cada uno un gen de PKS salvaje o mutante diferente, e introduciendo seguidamente todas las posibles combinaciones de los plásmidos con secuencias codificadoras de PKS salvajes, mutantes e híbridas en el mismo hospedante (véanse la Solicitud de Patente de los EE.UU. N° de Serie 60/129.731, archivada el 16 de abril 1999, y la Pub. PCT N° WO 98/27203). Este método puede incorporar también el uso de un mutante KS1°, que puede producir por biosíntesis mutacional policétidos producidos
25 a partir de unidades iniciadoras dicetónicas (véase Jacobsen *et al.*, 1997, *Science* 277, 367-369), así como el uso de un gen truncado que conduce a macrólidos de 12 miembros o un gen elongado que conduce a cetólidos de 16 miembros. Además, utilizando adicionalmente uno o más vectores que codifiquen genes de biosíntesis y transferencia de glicosilos, tales como los de la presente invención para megosamina, desosamina, oleandrosa, cladinosina y/o micarosa (en cualquier combinación), puede prepararse una gran colección de policétidos glicosilados.

30 La siguiente tabla enumera referencias que describen genes de PKS ilustrativos y las enzimas correspondientes que pueden utilizarse en la construcción de las PKSs híbridas recombinantes y los compuestos de ADN correspondientes que las codifican. Se presentan también varias referencias que describen enzimas de adaptación y los genes correspondientes que pueden emplearse conforme a los métodos de la invención.

35 Avermectina

Patente de los EE.UU. N° 5.252.474 otorgada a Merck.

40 **MacNeil et al.**, 1993, *Industrial Microorganisms: Basic and Applied Molecular Genetics*, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

45 **MacNeil et al.**, 1992, *Gene* 115: 119-125, Complex Organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase.

Candicidina (FR008)

50 **Hu et al.**, 1994, *Mol. Microbiol.* 14: 163-172.

Epotilona

55 Publicación de Patente PCT N° WO 00/031247 otorgada a Kosan.

Eritromicina

60 Publicación PCT N° 93/13663 otorgada a Abbott.

Patente de los EE.UU. N° 5.824.513 otorgada a Abbott.

65 **Donadio et al.**, 1991, *Science* 252: 675-9.

Cortes et al., 8 Nov. 1990, *Nature* 348: 176-8, An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of *Saccharopolyspora erythraea*.

ES 2 283 323 T3

Enzimas de Glicosilación

Publicación de Solicitud de Patente PCT N° 97/23630 otorgada a Abbott.

5

FK-506

Motamedi et al., 1998, The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506, *Eur. J. Biochem.* 256: 528-534.

10

Motamedi et al., 1997, Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK506, *Eur. J. Biochem.* 244: 74-80.

15 Metiltransferasa

EE.UU. 5.264.355, expedida el 23 Nov. 1993, Methylating enzyme from *Streptomyces* MA6858.31 O-desmethyl-FKSO6 methyltransferase.

20

Motamedi et al., 1996, Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK506 and FK520, *J. Bacteriol.* 178: 5243-5248.

25 FK-520

25

Publicación de Patente PCT N° WO 00/020601 otorgada a Kosan.

Véase también Nielsen *et al.*, 1991, *Biochem.* 30: 5789-96 (enzymology of piperolate incorporation).

30

Lovastatina

Patente de los EE.UU. N° 5.744.350 otorgada a Merck.

35

Narbomicina (y Picromicina)

Publicación de Patente PCT N° WO 99/61599 otorgada a Kosan.

40

Nemadectina

MacNeil et al., 1993, *ut supra*.

45

Nidamicina

Kakavas et al., 1997, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol.* 179: 7515-7522.

50

Oleandomicina

55 **Swan et al.**, 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet.* 242: 358-362.

Publicación de Patente PCT N° WO 00/026349 otorgada a Kosan.

60

Olando et al., 1998, Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet.* 259(3): 299-308.

Platenolida

65

Publicación de Solicitud de Patente EP N° 791.656 otorgada a Lilly.

ES 2 283 323 T3

Rapamicina

Schwecke et al., Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA* 92: 7839-7843.

5

Aparicio et al., 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, *Gene* 169: 9-16.

10 Rifamicina

August et al., 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deduction from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S669, *Chemistry & Biology*, 5(2): 69-79.

15

Sorafeno

Patente de los EE.UU. N° 5.716.849 otorgada a Novartis.

20

Schupp et al., 1995, *J. Bacteriology* 177: 3673-3679. A *Sorangium cellulosum* (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

25 Espiramicina

Patente de los EE.UU. N° 5.098.837 otorgada a Lilly.

30 Gen Activador

Patente de los EE.UU. N° 5.514.544 otorgada a Lilly.

35 Tilosina

Publicación EP N° 791.655 otorgada a Lilly.

40 **Kuhstoss et al.**, 1996, *Gene* 183: 231-6, Production of a novel polyketide through the construction of a hybrid polyketide synthase.

Patente de los EE.UU. N° 5.876.991 otorgada a Lilly.

45 Enzimas de Adaptación

Merson-Davies y Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

50

Como ilustra a Tabla anterior, existe una amplia variedad de genes de PKS que sirven como fuentes fácilmente disponibles de ADN e información para usar en la construcción de los compuestos de ADN que codifican las PKSs híbridas de la invención.

55 En la construcción de las PKSs híbridas, pueden ser de ayuda ciertos métodos generales. Por ejemplo, es con frecuencia beneficioso retener el marco del módulo que se va a alterar para producir la PKS híbrida. Por consiguiente, si se desea añadir funcionalidades DH y ER a un módulo, se prefiere con frecuencia remplazar el dominio KR del módulo original por un segmento que contenga los dominios KR, DH y ER de otro módulo, en vez de meramente insertar los dominios DH y ER. Se puede alterar la especificidad estereoquímica de un módulo por remplazamiento del dominio KS por un dominio KS de un módulo que especifique una estereoquímica diferente, Véase Lau *et al.*, 1999, "Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units" *Biochemistry* 38(5): 1643-1651. Se puede alterar la especificidad de un domino AT cambiando sólo un pequeño segmento del dominio. Véase Lau *et al.*, *ut supra*. Se pueden aprovechar también regiones enlazadoras conocidas en proteínas PKS para enlazar módulos de dos PKSs diferentes para crear una PKS híbrida. Véase Gokhale *et al.*, 16 Abr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases", *Science* 284: 482-485.

65

Los compuestos de ADN que codifican PKSs híbridas pueden ser y son con frecuencia híbridos de más de dos genes de PKS. Incluso cuando sólo se utilizan dos genes, hay con frecuencia dos o más módulos en el gen híbrido en los que todo o parte del módulo deriva de un segundo (o tercer) gen de PKS.

La invención proporciona también bibliotecas de genes de PKS, proteínas PKS, y finalmente, de policétidos, que se construyen generando modificaciones en una PKS para que los complejos proteicos producidos tengan actividades alteradas en uno o más sentidos y produzcan así policétidos distintos del producto natural de la PKS. Pueden prepararse así policétidos novedosos, o prepararse policétidos en general más fácilmente, usando este método. Proporcionando un gran número de genes diferentes o “clusters” de genes derivados de un “cluster” de genes de PKS existente en la naturaleza, cada uno de los cuales se ha modificado de manera diferente a partir del “cluster” nativo, puede producirse una biblioteca de policétidos combinatoria eficazmente como resultado de las múltiples variaciones de estas actividades. Como se describirá adicionalmente más abajo, las imposiciones y límites de esta realización de la invención pueden describirse sobre los niveles de policétido, proteína, y la secuencia de nucleótidos codificadora.

Existen al menos cinco grados de libertad para construir una PKS híbrida en términos del policétido que se producirá. Primero, la longitud de la cadena del policétido está determinada por el número de módulos de elongación en la PKS, y la presente invención incluye PKSs híbridas que contienen 6, así como menos o más de 6, módulos de elongación. Segundo, la naturaleza de la cadena principal carbonada de la PKS está determinada por las especificidades de las aciltransferasas que determinan la naturaleza de las unidades de elongación en cada posición, p.ej., malonilo, metilmalonilo, etilmalonilo, u otro malonilo sustituido. Tercero, la especificidad del módulo de carga tiene también un efecto sobre la cadena principal carbonada del policétido. El módulo de carga puede usar una unidad iniciadora diferente, tal como acetilo, butirilo, y similares. Como se señala anteriormente, otro método para variar la especificidad del módulo de carga implica inactivar la actividad KS en el módulo de elongación 1 (KS1) y proporcionar sustratos alternativos, llamados dicétidos, que son análogos sintetizados químicamente de los productos dicetónicos del módulo de elongación 1, para el módulo de elongación 2. Este abordaje se ilustró en las publicaciones PCT Nos. 97/02358 y 99/03986, en las que la actividad KS1 se inactivó por mutación. Cuarto, el estado de oxidación en varias posiciones del policétido estará determinado por las porciones deshidratasa y reductasa de los módulos. Esto determinará la presencia y localización de los restos cetona y alcohol y los dobles enlaces C-C o enlaces sencillos C-C en el policétido. Finalmente, la estereoquímica del policétido resultante está en función de tres aspectos de la sintasa. El primer aspecto está relacionado con la especificidad AT/KS asociada con los malonilos sustituidos como unidades de elongación, que afecta a la estereoquímica sólo cuando no existe ciclo reductor o cuando contiene sólo una cetoreductasa, puesto que la deshidratasa suprimiría la quiralidad. Segundo, la especificidad de la cetoreductasa puede determinar la quiralidad de cualquier beta-OH. Finalmente, la especificidad enoilreductasa para los malonilos sustituidos como unidades de elongación puede influenciar la estereoquímica cuando está disponible un KR/DH/ER completo.

Por consiguiente, los sistemas de PKS modulares permiten generalmente sintetizar una amplia gama de policétidos. Comparados con los sistemas de PKS aromáticos, los sistemas de PKS modulares aceptan una amplia gama de unidades iniciadoras, incluyendo monómeros alifáticos (acetilo, propionilo, butirilo, isovalerilo, etc.), aromáticos (aminohidroxibenzoilo), alicíclicos (ciclohexanoilo) y heterocíclicos (tiazolilo). Ciertas PKSs modulares tienen una especificidad suavizada hacia sus unidades iniciadoras (Kao *et al.*, 1994, *Science, ut supra*). Las PKSs modulares muestran también una variedad considerable con respecto a la elección de las unidades iniciadoras en cada ciclo de condensación. El grado de beta-cetorreducción después de una reacción de condensación puede alterarse por manipulación genética (Donadio *et al.*, 1991, *Science, ut supra*; Donadio *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90: 7119-7123). Similarmente, el tamaño del producto policétídico puede variarse diseñando mutantes con el número apropiado de módulos (Kao *et al.*, 1994, *J. Am. Chem. Soc.* 116: 11612-11613). Por último, las enzimas PKS modulares son particularmente bien conocidas para la generación de una gama impresionante de centros asimétricos en sus productos de manera muy controlada. Los policétidos, antibióticos y otros compuestos producidos por los métodos de la invención son típicamente formas estereoisómeras únicas. Aunque los compuestos de la invención pueden darse en mezclas de estereoisómeros, puede ser beneficioso en algunos casos generar estereoisómeros individuales. Por consiguiente, el potencial combinatorio en las rutas de PKSs modulares basadas en cualquier “andamio” de PKSs modulares existentes en la naturaleza es virtualmente ilimitado.

Aunque las PKSs híbridas se producen lo más frecuentemente “mezclando y pegando” porciones de secuencias codificadoras de PKS, pueden usarse también mutaciones en el ADN que codifica una PKS para introducir, alterar o suprimir una actividad en el polipéptido codificado. Las mutaciones pueden realizarse en las secuencias nativas usando técnicas convencionales. Los sustratos para la mutación pueden ser un “cluster” entero de genes o sólo uno o dos de ellos; el sustrato para la mutación puede ser también porciones de uno o más de estos genes. Las técnicas para la mutación incluyen preparar oligonucleótidos sintéticos que incluyen las mutaciones e insertar la secuencia mutada en el gen que codifica una subunidad de PKS usando digestión con endonucleasas de restricción. Véanse, p.ej., Kundel, 1985 *Proc. Natl. Acad. Sci. USA* 82: 448; Geisselsoder *et al.*, 1987, *BioTechniques* 5: 786. Alternativamente, las mutaciones pueden efectuarse usando un cebador desapareado (generalmente de 10-20 nucleótidos de longitud) que se hibrida con la secuencia de nucleótidos nativa, a una temperatura inferior a la temperatura de fusión del dúplex desapareado. El cebador puede hacerse específico manteniendo la longitud del cebador y la composición de bases en unos límites relativamente estrechos y manteniendo la base mutada centralmente situada. Véase Zoller y Smith, 1983, *Methods Enzymol.* 100: 468. La elongación del cebador se efectúa usando ADN polimerasa, el producto se clona, y se seleccionan los clones que contienen el ADN mutado, derivados por segregación de la hebra de cebador elongada. La identificación puede llevarse a cabo usando el cebador mutante como sonda de hibridación. La técnica es aplicable también para generar mutaciones en múltiples puntos. Véase, p.ej., Dalbie-McFarland *et al.*, 1982, *Proc. Natl. Acad. Sci. USA* 79: 6409. Puede usarse también mutagénesis por PCR para efectuar las mutaciones deseadas.

Puede llevarse a cabo también la mutagénesis aleatoria de porciones seleccionadas de las secuencias de nucleótidos que codifican las actividades enzimáticas por varias técnicas diferentes conocidas en la técnica, p.ej., insertando

un enlazador nucleotídico aleatoriamente en un plásmido, irradiando con rayos-X o luz ultravioleta, incorporando nucleótidos incorrectos durante la síntesis de ADN *in vitro*, por mutagénesis por PCR con tendencia al error (“error-prone PCR”), preparando mutantes sintéticos, o dañando el ADN plasmídico *in vitro* con productos químicos. Los mutágenos químicos incluyen, por ejemplo, bisulfito de sodio, ácido nitroso, nitrosoguanidina, hidroxilamina, agentes que dañan o eliminan bases previniendo así el emparejamiento normal de bases, tales como hidracina o ácido fórmico, análogos de precursores nucleotídicos tales como 5-bromouracilo, 2-aminopurina, o agentes intercalantes de acridina tales como proflavina, acriflavina, quinacrina, y similares. Generalmente, el ADN plasmídico o los fragmentos de ADN se tratan con mutágenos químicos, se utilizan para transformar *E. coli* y se propagan como un combinado o genoteca de plásmidos mutantes.

Al construir una PKS híbrida de la invención, las regiones que codifican actividad enzimática, esto es, las regiones que codifican las actividades correspondientes de diferentes PKS sintetas o de diferentes localizaciones en la misma PKS, pueden recuperarse, por ejemplo, usando técnicas de PCR con cebadores apropiados. Por regiones que codifican actividades “correspondientes” se entiende aquellas regiones que codifican el mismo tipo general de actividad. Por ejemplo, una actividad KR codificada en una localización de un “cluster” de genes “corresponde” a una actividad codificadora KR en otra localización en el “cluster” de genes o en un “cluster” de genes diferente. Similarmente, un ciclo de reductasa completo podría considerarse correspondiente. Por ejemplo, KR/DH/ER puede corresponder a una KR sola.

Si se va a realizar el remplazamiento de una región diana en particular en una PKS hospedante, este remplazamiento puede llevarse a cabo *in vitro* usando enzimas de restricción adecuadas. El remplazamiento puede efectuarse también *in vivo* usando técnicas recombinantes que implican secuencias homólogas flanqueando el gen de remplazamiento en un plásmido donante y una región receptora en un plásmido receptor. Tales sistemas, que implican ventajosamente plásmidos con diferentes sensibilidades a la temperatura, se describen, por ejemplo, en la publicación PCT N° WO 96/40968. Los vectores usados para realizar las diversas operaciones para remplazar la actividad enzimática en los genes de PKS hospedantes o para mantener mutaciones en estas regiones de los genes de PKS hospedantes pueden elegirse para contener secuencias control unidas funcionalmente a las secuencias codificadoras resultantes de manera tal que pueda efectuarse la expresión de las secuencias codificadoras en un hospedante apropiado.

Sin embargo, pueden usarse asimismo vectores de clonación simples. Si los vectores de clonación empleados para obtener genes de PKS que codifican una PKS derivada carecen de secuencias control para la expresión unidas funcionalmente a las secuencias de nucleótidos codificadoras, las secuencias de nucleótidos pueden insertarse en vectores de expresión apropiados. Esto no necesita hacerse individualmente, sino que un combinado de secuencias de nucleótidos codificadoras aisladas puede insertarse en vectores de expresión, los vectores resultantes usarse para transformar o transfectar células hospedantes, y las células resultantes disponerse en placas para obtener colonias individuales. La invención proporciona una variedad de compuestos de ADN recombinante en los que las diversas secuencias codificadoras de los dominios y módulos de la PKS están flanqueados por sitios de reconocimiento de enzimas de restricción no existentes en la naturaleza.

Las diversas secuencias de nucleótidos de PKS pueden clonarse en uno o más vectores recombinantes como cassetes individuales, con diferentes elementos de control, o bajo el control de, p.ej., un único promotor. Las regiones codificadoras de subunidades de PKS pueden incluir sitios de restricción flanqueantes para permitir la fácil supresión e inserción de otras secuencias codificadoras de subunidades de PKS para poder generar PKSs híbridas. El diseño de tales sitios de restricción únicos es conocido por los expertos en la técnica y puede llevarse a cabo usando las técnicas descritas anteriormente, tales como mutagénesis dirigida y PCR.

Los vectores de expresión que contienen secuencias de nucleótidos que codifican una variedad de enzimas PKS para la producción de diferentes policétidos se usan entonces para transformar células hospedantes apropiadas para construir la genoteca. En un abordaje sencillo, se usa una mezcla de tales vectores para transformar las células hospedantes seleccionadas y las células resultantes se disponen en placas para obtener colonias individuales y se seleccionan para identificar los transformantes positivos. Cada colonia individual tiene la capacidad de producir una PKS sintasa en particular y en última instancia un policétido en particular. Típicamente, habrá duplicaciones en algunas, la mayoría o todas las colonias; el subconjunto de las colonias transformadas que contiene una PKS diferente en cada colonia miembro puede considerarse una genoteca. Alternativamente, los vectores de expresión pueden usarse individualmente para transformar hospedantes, hospedantes transformados que se recopilan en una genoteca. Existe una variedad de estrategias disponibles para obtener gran diversidad colonias que contiene cada una de ellas un “cluster” de genes de PKS derivado del “cluster” de genes del hospedante existente en la naturaleza de tal forma que cada colonia en la genoteca produce una PKS diferente y en última instancia un policétido diferente. El número de policétidos diferentes que produce la genoteca es típicamente al menos cuatro, más típicamente al menos diez, y preferiblemente al menos 20, y más preferiblemente al menos 50, reflejando números similares de “clusters” de genes de PKS alterados diferentes y productos de genes de PKS. El número de miembros en la genoteca se elige arbitrariamente; sin embargo, los grados de libertad resumidos anteriormente con respecto a la variación de iniciador, unidades de elongación, estereoquímica, estado de oxidación y longitud de la cadena posibilita la producción de genotecas bastante grandes.

Los métodos para introducir los vectores recombinantes de la invención en hospedantes adecuados son conocidos por los expertos en la técnica e incluyen típicamente el uso de CaCl₂ o agentes tales como cationes divalentes, lipofeción, DMSO, transformación de protoplastos, infección, transfección y electroporación. Las colonias productoras de policétido pueden identificarse y aislarse usando técnicas conocidas y los policétidos producidos caracterizarse adi-

cionalmente. Los policétidos producidos por estas colonias pueden usarse colectivamente en un panel para representar una genoteca o pueden evaluarse individualmente con respecto a su actividad.

Las bibliotecas de la invención pueden por tanto considerarse a cuatro niveles: (1) una gran diversidad de colonias cada una con una secuencia codificadora de PKS diferente; (2) las proteínas producidas por las secuencias codificadoras; (3) los policétidos producidos a partir de las proteínas recopiladas en una PKS funcional; y (4) los antibióticos o compuestos con otras actividades deseadas derivados de los policétidos.

Las colonias en la genoteca se inducen para producir las sintasas relevantes y por tanto para producir los policétidos relevantes para obtener una biblioteca de policétidos. Los policétidos secretados al medio pueden someterse a escrutinio sobre la base de la unión a dianas deseadas, tales como receptores, proteínas señal, y similares. Pueden usarse los sobrenadantes tal cual para el escrutinio, o puede efectuarse primero la purificación parcial o completa de los policétidos. Típicamente, tales métodos de escrutinio implican detectar la unión de cada miembro de la biblioteca al receptor u otro ligando diana. La unión puede detectarse bien directamente o a través de un ensayo de competición. Los medios para someter a escrutinio tales bibliotecas con respecto a la unión son bien conocidos en la técnica. Alternativamente, pueden ensayarse miembros policetídicos individuales de la biblioteca frente a la diana deseada. En este caso, pueden incluirse más fácilmente escrutinios en los que se mida la respuesta biológica de la diana. La actividad antibiótica puede verificarse usando ensayos típicos tales como los descritos en Lehrer *et al.*, 1991, *J. Immunol. Meth.* 137: 167-173, y en los Ejemplos más abajo.

La invención proporciona métodos para la preparación de un gran número de policétidos. Estos policétidos son compuestos intermedios útiles en la formación de compuestos con actividad antibiótica o de otra clase a través de reacciones de hidroxilación, epoxidación y glicosilación como se describe anteriormente. En general, los productos policetídicos de la PKS deben modificarse adicionalmente, típicamente por hidroxilación y glicosilación, para presentar actividad antibiótica. La hidroxilación da como resultado los policétidos novedosos de la invención que contienen grupos hidroxilo en C-6, que pueden conseguirse usando la hidroxilasa codificada por el gen *eryF*, y/o C-12, que pueden conseguirse usando la hidroxilasa codificada por el gen *picK* o *eriK*. También está disponible el gen *oleP* en forma recombinante, que puede usarse para expresar el producto del gen *oleP* en cualquier célula hospedante. Una célula hospedante, tal como una célula hospedante de *Streptomyces* o una célula hospedante de *Saccharopolyspora erythraea*, modificada para expresar el gen *oleP* puede usarse por tanto para producir policétidos que comprenden el epóxido C-8-C-8a presente en la oleandomicina. Por consiguiente, la invención proporciona tales policétidos modificados. La presencia de grupos hidroxilo en estas posiciones puede intensificar la actividad antibiótica del compuesto resultante en relación con su homólogo no hidroxilado.

Los métodos para glicosilar son generalmente conocidos en la técnica; la glicosilación puede efectuarse intracelularmente proporcionando las enzimas de glicosilación apropiadas o puede efectuarse *in vitro* usando medios químicos de síntesis como se describe en la presente memoria y en la publicación PCT N° WO 98/49315. Preferiblemente, la glicosilación con desosamina, micarosa y/o megosamina se efectúa conforme a los métodos de la invención en células hospedantes recombinantes proporcionadas por la invención. En general, los abordajes para efectuar la glicosilación reflejan los descritos anteriormente con respecto a la hidroxilación. Las enzimas purificadas, aisladas de fuentes naturales o producidas de forma recombinante pueden usarse *in vitro*. Alternativamente y como se señala anteriormente, la glicosilación puede efectuarse intracelularmente usando glicosilasas intracelulares endógenas o producidas de forma recombinante. Además, pueden emplearse métodos químicos de síntesis.

Los policétidos modulares antibióticos pueden contener cualquiera entre un número diferente de azúcares, aunque la D-desosamina, o un análogo cercano de la misma, es el más común. La eritromicina, picromicina, megalomicina, narbomicina y metilmicina contienen desosamina. La eritromicina contiene también L-cladinoso (3-O-metilmicaroso). La tilosina contiene micaminosa (4-hidroxidesosamína), micarosa y 6-desoxi-D-alosa. La 2-acetil-1-bromodesosamina se ha usado como donante para glicosilar policétidos por Masamune *et al.*, 1975, *J. Am. Chem. Soc.* 97: 3512-3513. Otros donantes, aparentemente más estables, incluyen fluoruros de glicosilo, tioglicósidos, y tricloroacetimidatos; véanse Woodward *et al.*, 1981, *J. Am. Chem. Soc.* 103: 3215; Martin *et al.*, 1997, *J. Am. Chem. Soc.* 119: 3193; Toshima *et al.*, 1995, *J. Am. Chem. Soc.* 117: 3717; Matsumoto *et al.*, 1988, *Tetrahedron Lett.* 29: 3575. La glicosilación puede efectuarse también usando las agliconas policetídicas como materiales de partida y usando *Saccharopolyspora erythraea* o *Streptomyces venezuelae* u otra célula hospedante para realizar la conversión, preferiblemente usando mutantes incapaces de sintetizar macrólidos, como se discute anteriormente.

Por consiguiente, puede producirse una amplia variedad de policétidos por las enzimas PKS híbridas de la invención. Estos policétidos son útiles como antibióticos y como compuestos intermedios en la síntesis de otros compuestos útiles. En un aspecto importante, la invención proporciona métodos para producir compuestos antibióticos relacionados en estructura con la eritromicina, un compuesto antibiótico potente. La invención proporciona también compuestos cetólidos novedosos, compuestos policetídicos con actividad antibiótica potente de interés importante debido a la actividad frente a cepas bacterianas resistentes a antibiótico. Véase Griesgraber *et al.*, 1996, *J. Antibiot.* 49: 465-477. La mayoría, si no todos los cetólidos preparados hasta la fecha se sintetizan usando eritromicina A, un derivado de 6-dEB, como compuesto intermedio. Véanse Griesgraber *et al.*, *ut supra*; Agouridas *et al.*, 1998, *J. Med. Chem.* 41: 4080-4100, Patentes de los EE.UU. Nos. 5.770.579; 5.760.233; 5.750.510; 5.747.467; 5.747.466; 5.656.607; 5.635.485; 5.614.614; 5.556.118; 5.543.400; 5.527.780; 5.444.051; 5.439.890; 5.439.889; y las publicaciones PCT Nos. WO 98/09978 y 98/28316.

Como se señala anteriormente, los genes de PKS híbridos de la invención pueden expresarse en una célula hospedante que contiene los genes biosintéticos de desosamina, megosamina y/o micarosa y los genes de la transferasa correspondientes, así como el(los) gen(es) de hidroxilasa requeridos, que pueden ser *picK*, *megK* o *eryK* (para la posición C-12) y/o *megF* o *eryF* (para la posición C-6). Los compuestos resultantes tienen actividad antibiótica pero pueden modificarse adicionalmente, como se describe en las publicaciones de patentes a las que se hace referencia anteriormente, para dar un compuesto deseado con propiedades mejoradas u otra cosa que se desee. Alternativamente, pueden producirse en la célula hospedante recombinante los compuestos de tipo aglicona, y llevarse a cabo los pasos de glicosilación e hidroxilación deseados *in vitro* o *in vivo*, en este último caso proporcionando la aglicona a la célula convertidora, como se describe anteriormente.

Como se describe anteriormente, existe una amplia variedad de organismos diversos que pueden modificar compuestos tales como los descritos en la presente memoria para proporcionar compuestos con, o que pueden modificarse fácilmente para tener actividades útiles. Por ejemplo, *Saccharopolyspora erythraea* puede convertir 6-dEB en una variedad de compuestos útiles. Los compuestos proporcionados por la presente invención pueden proporcionarse a cultivos de *Saccharopolyspora erythraea* y convertirse en los derivados correspondientes de las eritromicinas A, B, C y D conforme al procedimiento proporcionado en los Ejemplos, más abajo. Para asegurarse de que sólo se produce el compuesto deseado, se puede usar un mutante en *eryA* de *S. erythraea* que es incapaz de producir 6-dEB pero que aún puede llevar a cabo las conversiones deseadas (Weber *et al.*, 1985, *J. Bacteriol.* 164(1): 425-433). También, se pueden emplear otras cepas mutantes, tales como mutantes en *eryB*, *eryC*, *eryG*, y/o *eriK*, o cepas mutantes que tienen mutaciones en múltiples genes, para acumular un compuesto preferido. La conversión puede llevarse a cabo también en grandes fermentadores para producción comercial. Cada una de las eritromicinas A, B, C y D tiene actividad antibiótica, aunque la eritromicina A tiene la mayor actividad antibiótica. Además, cada uno de estos compuestos puede formar, por tratamiento con un ácido débil, un hemiacetal de C-6 a C-9 con actividad inductora de la motilidad. Para la formación de hemiacetales con actividad inductora de la motilidad, se prefieren las eritromicinas B, C y D, ya que la presencia de un hidroxilo en C-12 permite la formación de un compuesto inactivo que tiene un hemiacetal formado entre C-9 y C-12.

Por consiguiente, la presente invención proporciona los compuestos producidos por hidroxilación y glicosilación de los compuestos de la invención por acción de las enzimas endógenas de *Saccharopolyspora erythraea* y cepas mutantes de *S. erythraea*. Tales compuestos son útiles como antibióticos o como inductores de la motilidad directamente o tras su modificación química. Para el uso como antibióticos, los compuestos de la invención pueden usarse directamente sin modificación química adicional. Las eritromicinas A, B, C y D tienen todas actividad antibiótica, y los compuestos correspondientes de la invención que resultan de la modificación de los compuestos por *Saccharopolyspora erythraea* tienen también actividad antibiótica. Estos compuestos pueden modificarse químicamente, sin embargo, para proporcionar otros compuestos de la invención con actividad antibiótica potente. Por ejemplo, la alquilación de la eritromicina en el hidroxilo C-6 puede usarse para producir antibióticos potentes (la claritromicina es C-6-O-metilo), y otras modificaciones útiles se describen en, por ejemplo, Griesgraber *et al.*, 1996, *J. Antibiot.* 49: 465-477, Agouridas *et al.*, 1998, *J. Med. Chem.* 41: 4080-4100, Patentes de los EE.UU. Nos. 5.770.579; 5.760.233; 5.750.510; 5.747.467; 5.747.466; 5.656.607; 5.635.485; 5.614.614; 5.556.118; 5.543.400; 5.527.780; 5.444.051; 5.439.890; y 5.439.889; y publicaciones PCT Nos. WO 98/09978 y 98/28316.

Para el uso como inductores de la motilidad, los compuestos de la invención pueden usarse directamente sin modificación química adicional. La eritromicina y ciertos análogos de la eritromicina son agonistas potentes del receptor de la motilina que pueden usarse clínicamente como agentes procinéticos para inducir la fase III de los complejos motores migratorios, para aumentar la peristalsis esofágica y la presión del LES (del inglés "Lower Esophageal Sphincter", esfínter esofágico inferior) en pacientes con GERD (del inglés "Gastroesophageal Reflux Disease", enfermedad del reflujo gastroesofágico), para acelerar el vaciado gástrico en pacientes con paresia gástrica, y para estimular contracciones de la vesícula biliar en pacientes tras la eliminación de piedras vesiculares y en diabéticos con neuropatía autonómica. Véase Omura *et al.*, 1987, Macrolides with gastrointestinal motor stimulating activity, *J. Med. Chem.* 30: 1941-3. Los compuestos correspondientes de la invención que resultan de la modificación de los compuestos por *Saccharopolyspora erythraea* tienen también actividad inductora de la motilidad, particularmente tras la conversión, que también puede ocurrir *in vivo*, al hemiacetal de C-6 a C-9 por tratamiento con ácido débil. Se prefieren especialmente los compuestos que no tienen el hidroxilo en C-12 para usar como agonistas de la motilina. Estos compuestos pueden también modificarse además químicamente, sin embargo, para proporcionar otros compuestos de la invención con actividad inductora de la motilidad potente.

Además, y también como se señala anteriormente, hay otros organismos útiles que pueden emplearse para hidroxilar y/o glicosilar los compuestos de la invención. Como se describe anteriormente, los organismos pueden ser mutantes incapaces de producir el policétido producido normalmente en ese organismo, la fermentación puede llevarse a cabo en placas o en grandes fermentadores, y los compuestos producidos pueden alterarse químicamente tras la fermentación. Además de *Saccharopolyspora erythraea*, pueden usarse también *Streptomyces venezuelae*, *S. narbonensis*, *S. antibioticus*, *Micromonospora megalomicea*, *S. fradiae* y *S. thermotolerans*. Además de la actividad antibiótica, los compuestos de la invención producidos por tratamiento con enzimas de *M. megalomicea* pueden tener también actividad antiparasitaria. Por consiguiente, la presente invención proporciona los compuestos producidos por hidroxilación y glicosilación por acción de las enzimas endógenas de *S. erythraea*, *S. venezuelae*, *S. narbonensis*, *S. antibioticus*, *M. megalomicea*, *S. fradiae* y *S. thermotolerans*.

ES 2 283 323 T3

Los compuestos de la invención pueden aislarse de los caldos de fermentación de estas células cultivadas y purificarse por procedimientos estándar. Los compuestos pueden formularse fácilmente para proporcionar las composiciones farmacéuticas de la invención. Las composiciones farmacéuticas de la invención pueden usarse en forma de una preparación farmacéutica, por ejemplo, en forma sólida, semisólida o líquida. Esta preparación contendrá uno o más de las composiciones de la invención como un ingrediente activo mezclado con un portador o excipiente orgánico o inorgánico adecuado para su aplicación externa, entérica o parenteral. El ingrediente activo puede combinarse, por ejemplo, con los portadores farmacéuticamente aceptables usuales no tóxicos para comprimidos, pelets, cápsulas, supositorios, soluciones, emulsiones, suspensiones, y cualquier otra forma adecuada para el uso.

Los portadores que pueden usarse incluyen agua, glucosa, lactosa, goma de acacia, gelatina, manitol, pasta de almidón, trisilicato de magnesio, talco, almidón de maíz, queratina, sílice coloidal, almidón de patata, urea, y otros portadores adecuados para usar en preparaciones para fabricación, en forma sólida, semisólida o líquida. Además, pueden usarse agentes de estabilización, espesamiento y coloración y perfumes auxiliares. Por ejemplo, los compuestos de la invención pueden utilizarse con hidroxipropilmetilcelulosa esencialmente como se describe en la Patente de los EE.UU. N° 4.916.138, o con un tensioactivo esencialmente como se describe en la publicación de patente EPO N° 428.169.

Las formas de dosificación orales pueden prepararse esencialmente como se describe en Hondo *et al.*, *Transplantation Proceedings XIX*, Sup. 6: 17-22. Las formas de dosificación para aplicación externa pueden prepararse esencialmente como se describe en la publicación de patente EPO N° 423.714. El compuesto activo se incluye en la composición farmacéutica en cantidad suficiente para producir el efecto deseado sobre el proceso de la enfermedad o afección.

Para el tratamiento de afecciones y enfermedades causadas por infección, puede administrarse un compuesto de la invención por vía oral, tópica, parenteral, por pulverización para inhalación, o por vía rectal en formulaciones unitarias de dosificación que contienen portadores, coadyuvantes y vehículos farmacéuticamente aceptables convencionales no tóxicos. El término parenteral, tal como se usa en la presente invención, incluye inyecciones subcutáneas e inyección intravenosa, intramuscular e intrainestinal, o técnicas de infiltración.

Los niveles de dosificación de los compuestos de la invención son del orden de aproximadamente 0,01 mg a aproximadamente 50 mg por kilogramo de peso corporal por día, preferiblemente de aproximadamente 0,1 mg a aproximadamente 10 mg por kilogramo de peso corporal por día. Los niveles de dosificación son útiles en el tratamiento de las afecciones indicadas anteriormente (de aproximadamente 0,7 mg a aproximadamente 3,5 mg por paciente por día, suponiendo un paciente de 70 kg). Además, los compuestos de la invención pueden administrarse intermitentemente, esto es, a intervalos semi-semanales, semanales, semi-mensuales o mensuales.

La cantidad de ingrediente activo que puede combinarse con los materiales portadores para producir una forma de dosificación única variarán dependiendo del receptor tratado y del modo particular de administración. Por ejemplo, una formulación pretendida para administración oral a humanos puede contener de 0,5 mg a 5 gm de agente activo combinado con una cantidad apropiada o conveniente de material portador, que puede variar de aproximadamente un 5 por ciento a aproximadamente un 95 por ciento de la composición total. Las formas unitarias de dosificación contendrán generalmente de aproximadamente 0,5 mg a aproximadamente 500 mg de ingrediente activo. Para administración externa, los compuestos de la invención pueden formularse en el intervalo del, por ejemplo, 0,00001% al 60% en peso, preferiblemente del 0,001% al 10% en peso, y más preferiblemente del aproximadamente 0,005% al 0,8% en peso.

Se entenderá, si embargo, que el nivel de dosis específico para cualquier paciente en particular dependerá de varios factores. Estos factores incluyen la actividad del compuesto empleado; la edad, peso corporal, salud general, sexo y dieta del sujeto; el tiempo y ruta de administración y la tasa de excreción del fármaco; si se emplea en el tratamiento una combinación de fármacos; y la gravedad de la enfermedad o afección en particular para la que se busca la terapia.

Habiéndose proporcionado anteriormente una descripción detallada de la invención, los siguientes ejemplos se ofrecen a los efectos de ilustrar la invención y no deben interpretarse como que son una limitación al alcance de la invención o reivindicaciones.

Ejemplo 1

Producción de Metilmalonil-CoA en E. coli

Este ejemplo describe, en la parte A, la clonación y expresión de metilmalonil-CoA mutasa, y en la parte B, la clonación y expresión de metilmalonil-CoA epimerasa, en *E. coli*.

A. Clonación y expresión de metilmalonil-CoA mutasa

La metilmalonil-CoA mutasa se clonó de *Propionibacterium shermanii* y se expresó en *E. coli*. La holoenzima mm-CoA mutasa se obtuvo cultivando las células en presencia de hidroxocobalamina y demostró ser activa sin adición de vitamina B12. La metilmalonil-CoA se produjo *in vivo*, tal como se vio por análisis de CoA usando una cepa *panD* de BL21 (DE3).

ES 2 283 323 T3

Para respaldar la producción modular de policétidos en *E. coli*, la invención proporciona métodos y reactivos para producir (S)-metilmalonil-CoA, que no está presente de forma natural en *E. coli*, sobreexpresando mm-CoA mutasa y mm-CoA epimerasa en *E. coli*. Se expresó una versión activa, con una cola FLAG, de la mm-CoA mutasa de *S. cinnamomensis* en células XL1Blue, que se cultivaron en presencia de hidroxocobalamina en un medio sintético libre de vitaminas para producir la holoenzima activa. Los niveles de CoA en las células se analizaron alimentando con β -alanina marcada; para este propósito resulta beneficioso tener una cepa *panD*, que es auxótrofa para β -alanina. El ADN de la mutasa sufrió sobrecruzamiento en la cepa *panD* de SJ16, una cepa *recA*⁺, de tal forma que el análisis de CoA tuvo que llevarse a cabo sin el *panD*. Esto dio como resultado una relación de señal a ruido inferior, pero aún pudieron detectarse niveles de mm-CoA elevados. Como alternativa a los genes de *S. cinnamomensis*, la invención proporciona una mm-CoA mutasa de *P. shermanii* clonada en un vector de expresión de *E. coli*, que es activa sin adición de vitamina B12, y que eleva los niveles de mm-CoA en *E. coli* en una cepa *panD* compatible con el ADN de la mutasa.

Propionibacterium freudenreichii subsp. shermanii se obtuvo como una picadura en agar zumo de tomate derivada de un espécimen liofilizado de NCIMB, Escocia (NCIMB # 9885). Se usó una cepa gg3 de *E. coli*, una versión *panD* de BL21 (DE3), para el análisis de CoA. Se usaron también las cepas de *E. coli* gg1 y gg2, versiones *recA*⁻ de la cepa SJ16 *panD*. El vector pKK** es una versión de pKK223-3 en el que la región de clonación se altera para ir de *NdeI* a *EcoRI* y un sitio *NdeI* extra se suprime. El crecimiento de *P. shermanii* y la preparación del ADN genómico se llevó a cabo como se describe en la bibliografía.

La subclonación de metilmalonil-CoA mutasa de *P. shermanii* en *E. coli* se llevó a cabo como sigue. El gen de la metilmalonil-CoA mutasa consiste en dos subunidades, *mutA* y *mutB*, que se amplificaron por PCR a partir del ADN genómico de *P. shermanii* en un total de cuatro fragmentos. Se usaron sitios de restricción existentes en la naturaleza para reconstruir el gen. Se introdujeron sitios de restricción únicos en ambos extremos del gen para su clonación, y el codón de iniciación para el gen *mutB* se cambió de GTG a ATG. Como se ilustra más abajo, estos cuatro fragmentos se clonaron en un vector BluescriptTM (Stratagene), se secuenciaron, y seguidamente se juntaron para formar el gen completo de la mutasa. El gen se clonó a continuación en los vectores de expresión pET22b y pKK** entre los sitios de restricción *NdeI* y *HindIII*, para formar pET-MUT y pKK**-MUT.

El pET-MUT se usó para transformar células competentes de BL21(DE3) y más tarde células gg3, que son una versión *panD* de BL21(DE3). El pKK**-MUT se usó para transformar SJ16 *panD* y XL1Blue. El ADN se ensayó por escrutinio de varias colonias con *NdeI* y *HindIII*, para determinar si el gen de la mutasa estaba todavía presente o si había sufrido sobrecruzamiento.

Para el análisis por SDS-PAGE, las células de la cepa BL21(DE3) que contenían pET-MUT (y pET solo, como control) se cultivaron aeróbicamente a 27°C en medio MUT con 100 μ g/mL de carbenicilina (carb) (el medio MUT es M9 sales, glucosa, tiamina, elementos traza y aminoácidos, como se ha descrito previamente para la expresión de metionina sintasa (Amaratunga M. *et al.*, *A synthetic module for the metH gene permits facile mutagenesis of the cobalamin-binding region of Escherichia coli methionine synthase: initial characterization of seven mutant proteins*. *Biochemistry*, 1996, **35**(7): p. 2453-63)). Se usaron cultivos crecidos durante una noche (250 μ L) para inocular 25 mL de medio MUT (carb), que se cultivaron a 27°C hasta una OD₆₀₀ de aproximadamente 0,5. Los cultivos se indujeron seguidamente con IPTG a una concentración final de 1 mM. Se dejaron dos cultivos a 27°C durante tres horas, mientras que los cultivos duplicados se cultivaron a 37°C durante dos horas. Las células se recogieron por centrifugación y los sedimentos se almacenaron a -80°C antes del análisis. Las células se lisaron por sonicación y se examinaron ambas fases soluble e insoluble por SDS-PAGE. Este procedimiento se repitió para las células de la cepa XL1Blue que contenían pKK**-MUT.

Para la expresión de mm-CoA mutasa activa (con hidroxocobalamina), se cultivaron células de la cepa gg3 que contenía pET-MUT (y pET solo, como control) en medio MUT (carb) y β -alanina 5 μ M durante aproximadamente 20 horas a 27°C. Las siguientes operaciones se realizaron en un cuarto oscuro con una luz roja de seguridad: se inocularon matraces de 125 mL, que contenían cada uno 25 mL de medio MUT con carb y β -alanina 5 μ M y envueltos en papel de aluminio, con hidroxocobalamina 5 μ M y seguidamente con 250 μ L de los cultivos iniciales respectivos. Después de agitar durante una noche a 27°C, los cultivos se indujeron con IPTG a una concentración final de 1 mM y se cultivaron durante 4:45 horas más, momento en el cual se recogieron (en tubos Falcon envueltos en papel de aluminio) por centrifugación a 4000 rpm durante 10 minutos. Los sedimentos se almacenaron a -80°C antes de ensayarse.

El ensayo de la mutasa se realizó como sigue. Todas las operaciones se realizaron en oscuridad o bajo una luz roja de seguridad. El sedimento de 25 mL de cultivo se dejó descongelar, se lavó con tampón C (fosfato de potasio 50 mM, pH 7,4, EDTA 5 mM, glicerol al 10%), y se resuspendió en 0,5 mL de tampón C que contenía inhibidores de proteasa (1 comprimido por 10 mL de tampón). Después de sonicar sobre hielo, el extracto se clarificó por centrifugación a 4°C durante 10 minutos a velocidad máxima en una microcentrífuga Eppendorf; el sobrenadante se sometió a ensayo. Los ensayos de enzima contenían, en un volumen final de 100 μ L, (2R,2S)-metilmalonil-CoA 0,2 mM, el extracto con mutasa, y tampón C que contenía inhibidores de proteasa. Las reacciones para los ensayos con vitamina B12 fueron como los anteriores pero contenían vitamina B12 0,01 mM, en cuyo caso el extracto con mutasa se incubó con la vitamina B12 en un volumen total de 75 μ L durante 5 minutos a 30°C antes del inicio de la reacción con metilmalonil-CoA. Después de la duración deseada de la incubación a 30°C, la reacción se paró por adición de 50 μ L de ácido tricloroacético (TCA) al 10% y se colocó sobre hielo durante aproximadamente 10 minutos. Los desechos celulares y proteína precipitada se retiraron por centrifugación durante 5 minutos en una microcentrífuga Eppendorf a 4°C.

ES 2 283 323 T3

Una alícuota (100 μL) del sobrenadante se inyectó en el HPLC para cuantificar la conversión de metilmalonil-CoA en succinil-CoA. Se tomó un tiempo tras 20 minutos de incubación a 30°C, y la muestra se ensayó con respeto a la conversión de mm-CoA en succinil-CoA. Todas las operaciones se realizaron exclusivamente bajo una luz roja de seguridad hasta que la reacción se paró por adición de TCA.

El análisis de CoA se realizó como se describe en la bibliografía, excepto que se añadió hidroxocobalamina 5 μM en el momento de la inducción con IPTG, y los tubos se envolvieron en papel de aluminio y se cultivaron a 27°C en vez de a 30°C. Los picos de CoA, que eluyeron en aproximadamente un minuto cada uno, se recogieron manualmente, así como aproximadamente un minuto de muestra tanto antes como después de cada pico. En algunos ensayos, las fracciones se recogieron cada 30 segundos. Todas las muestras se sometieron a recuento en un contador de centelleo.

Las dos subunidades del gen que codifica la metilmalonil-CoA mutasa están traduccionalmente acoplados - el codón de iniciación GTG de la subunidad situada hacia el extremo 3' *mutB* solapa con el codón ATG de *mutA*. El inicio mediante valina codificada por GTG se mutó a un inicio mediante metionina codificada por ATG (que no altera ningún otro aminoácido), porque *E. coli* utiliza el inicio mediante metionina más eficientemente. La secuenciación del gen de la mm-CoA mutasa reveló una discrepancia entre la secuencia observada y la secuencia publicada. Un "GC" en vez de un "CG" cambiaba dos aminoácidos de Asp, Val a Glu, Leu. La estructura cristalina de la mm-CoA mutasa de *P. shermanii* mostró que los dos aminoácidos son de hecho Glu, Leu, así que la secuencia publicada tiene un error. El gen de la mm-CoA mutasa se subclonó en dos sistemas de expresión en *E. coli* diferentes: pET, que está bajo el control del fuerte promotor T7, y pKK, que usa el promotor con escapes *tac*. Primero fue necesario encontrar cepas en las que el ADN de la mutasa no sufriera sobrecruzamiento. Se ha observado previamente que una versión con una cola FLAG de la mutasa de *S. cinnamonensis* sufría sobrecruzamiento en SJ16 *panD* y en BL21(DE3), que son cepas *recA*⁺ ambas, pero no en XL1Blue, que es *recA*⁻. Este ADN de la mutasa (*P. shermanii*) sufría también sobrecruzamiento en las células SJ16 pero no en las células BL21(DE3). Por consiguiente, se creó una versión *panD* de BL21(DE3) (gg3) para usar con el vector pET. Se creó también una versión *recA*⁻ de SJ16 (gg1, gg2) para usar con el sistema pKK; sin embargo, el ADN de la mutasa sufría sobrecruzamiento en esta cepa también.

Se ensayaron diferentes condiciones de crecimiento para encontrar las condiciones en las que las dos subunidades de la mutasa se expresaban en la fase soluble en aproximadamente iguales relaciones en moles. En general, parecía que la temperatura más alta de 37°C causaba que la mutasa apareciera predominantemente en forma insoluble. El crecimiento exclusivamente a 27°C dio como resultado proteína soluble con una relación de subunidades aproximadamente igual.

La Figura 3 muestra la comparación de los niveles de acil-CoAs *in vivo* en cepas BL21(DE3) *panD* con y sin mm-CoA mutasa. Para cada CoA, se determinó la relación entre la cantidad en la cepa que contenía la mutasa y la cantidad en la cepa control. Interesantemente, el malonil-CoA aumentó aproximadamente 25 veces y el succinil-CoA aproximadamente 3 veces. El acetil-CoA y la CoA aumentaron sólo ligeramente, y no se detectó propionil-CoA en ningún caso.

Para expresar la mutasa activa *in vivo*, fue necesario cultivar las células en un medio definido (medio MUT) que permite la reabsorción del precursor de vitamina B12 hidroxocobalamina; esto es similar a un protocolo establecido para la expresión de metionina sintasa activa, que requiere también vitamina B12. Se demostró que los extractos celulares que sobreexpresaban la mutasa convertían mm-CoA en succinil-CoA sin la adición de vitamina B12. Se ensayó sólo un tiempo (a los 20 minutos) para confirmar la actividad; la actividad específica de la mutasa no se determinó.

Por consiguiente, la metilmalonil-CoA mutasa se expresó como la holoenzima activa en *E. coli*, y la metilmalonil-CoA se produjo *in vivo*. Debido a que existe una epimerización química lenta espontánea entre (*R*) y (*S*)-mm-CoA (aproximadamente un 3% en 15 minutos), puede resultar de ayuda determinar las cantidades relativas de estos diastereoisómeros en las células que sobreexpresan la mutasa. Puede haber presente (*S*)-mm-CoA suficiente para respaldar la producción de policétidos en algunas células sin la adición de una epimerasa. Para facilitar la producción eventual de policétidos en *E. coli*, el gen de la mutasa puede incorporarse al cromosoma de la célula B21 *panD* u otra célula hospedante.

La Figura 2 muestra la construcción de pSK - MUT, en la que cuatro fragmentos de PCR se secuenciaron y se juntaron para formar el gen de la mutasa completo en pSK-bluescript.

En los experimentos que siguieron, se determinó la actividad específica de la mutasa y se completó un análisis de CoAs en profundidad. Como se muestra en la Figura 3, los niveles de CoAs en las células se analizaron otra vez usando una cepa *panD*, que es auxótrofa para β -alanina. Las células se alimentaron con ³H- β -alanina que se incorporó a las acil-CoAs, que se separaron mediante HPLC y se sometieron a recuento. Se examinaron los combinados de CoAs para los extractos celulares con y sin la mutasa, así como con y sin hidroxocobalamina.

Para ensayar si los acil-CoAs se degradaban en TCA, se llevaron a cabo los siguientes ensayos. La mezcla de CoAs consistía en malonil-, metilmalonil-, succinil-, acetil- y propionil-CoA 1,6 mM cada una, más CoA 0,5 mM. Se añadió una alícuota (10 μL) de esta mezcla a 100 μL de TCA al 10%, se inyectaron inmediatamente 50 μL en el HPLC para el análisis de CoAs, y lo que quedaba se congeló rápidamente en hielo seco. La porción congelada se dejó seguidamente descongelar y se cargó inmediatamente en el HPLC. De nuevo, se añadieron 10 μL de la mezcla de CoAs a 100 μL de TCA al 10%, se dejaron 50 μL en hielo durante 15 minutos y seguidamente se inyectaron en el HPLC, lo que quedaba

se dejó a 4°C durante una noche y se inyectó en el HPLC a la mañana siguiente. Se anotó el área bajo cada pico de CoA. Se siguió este mismo procedimiento sólo que usando una mezcla de TCA y tampón A del ensayo de mutasa.

El análisis de CoAs que se describe aquí se lleva a cabo en células que se lisan en TCA al 10%. Por consiguiente, es importante determinar si las CoAs se degradan significativamente en TCA y en una mezcla de TCA y tampón A del ensayo de mutasa. Los ensayos mostraron que el porcentaje de cada CoA relativo al combinado total de CoAs, así como la cantidad global de CoA, se mantenía constante después de congelar/descongelar, después de dejar en hielo durante 15 minutos, y después de dejar la muestra durante una noche a 4°C. Por consiguiente, las CoAs son estables en TCA y en el tampón de ensayo de la mutasa después de lisar las células o después de completar los ensayos, y antes del análisis por HPLC.

Aunque las CoAs son estables en TCA y tampón a 4°C, se degradaron a 30°C, la temperatura a la que se realizó el ensayo de la mutasa. En cinco minutos en las condiciones de ensayo, aproximadamente un 4% del metilmalonil-CoA se hidrolizó a CoA. El succinil-CoA se hidrolizó a una tasa comparable. Por consiguiente, el ensayo de la mutasa es subóptimo para unos resultados sumamente cuantitativos.

Cuando se incubó metilmalonil-CoA 0,2 mM con un lisado bruto de los extractos celulares que sobreexpresaban la mutasa, se produjo succinil-CoA. No se observó succinil-CoA cuando se incubó metilmalonil-CoA con lisados de la cepa control (que contenía el vector plasmídico pero sin los genes de la mutasa). En estas condiciones de expresión y ensayo, se observó una actividad específica de aproximadamente 0,04 U/mg en los extractos brutos. Cuando las células que sobreexpresaban la mutasa se cultivaron en medio MUT sin hidroxocobalamina, no se observó actividad mutasa; sin embargo, pudo detectarse actividad mutasa añadiendo vitamina B12 *in vitro*. La adición de vitamina B12 a los extractos que se cultivaron en presencia de hidroxocobalamina dio como resultado un aumento de la actividad mutasa, sugiriendo que una cantidad significativa de la mutasa expresada está presente como la apo-enzima. Esto podría haber sucedido porque la enzima se expresaba más rápido de lo que la hidroxocobalamina podía transportarse dentro de la célula, o porque el cofactor vitamina B12 se perdía durante la preparación del extracto.

La Figura 4 muestra la comparación de los niveles de acil-CoAs *in vivo* con y sin la mutasa y con y sin hidroxocobalamina. En las células que sobreexpresaban la mutasa y cultivadas con hidroxocobalamina, el metilmalonil-CoA comprendía un 13% del combinado global de CoAs, mientras que en otras células no fue detectable el metilmalonil-CoA. El nivel de fondo de los recuentos es aproximadamente un 0,25% del número global de recuentos en las CoAs, sugiriendo que cualquier metilmalonil-CoA presente en cepas de *E. coli* que no sobreexpresan la mutasa comprendería como mucho un 0,25% del combinado global de CoAs, o un 2% de la cantidad de metilmalonil-CoA observada en la cepa que sobreexpresaba la mutasa. La composición del combinado de CoAs observado en la cepa *panD* de *E. coli* es consistente con la observada previamente para mutantes *panD* de *E. coli* cultivados en glucosa.

Por consiguiente, la metilmalonil-CoA mutasa de *P. shermanii* se ha sobreexpresado como la holoenzima activa en *E. coli* y se ha demostrado que produce (2*R*)-metilmalonil-CoA *in vivo*. La conversión de (2*R*)- en (2*S*)-metilmalonil-CoA a través de la metilmalonil-CoA epimerasa debería proporcionar un suministro adecuado del isómero correcto de la metilmalonil-CoA para respaldar la producción heteróloga de policétidos complejos en *E. coli*.

La Figura 4 muestra los resultados del análisis de CoAs de *E. coli* que sobreexpresa metilmalonil-CoA mutasa. Se muestran los niveles de ³H detectados en las fracciones recogidas del análisis por HPLC de los extractos libres de células de *E. coli* alimentadas con ³H β-alanina que hospedaban el vector de control pET cultivadas sin hidroxocobalamina (línea sólida), pET cultivadas con hidroxocobalamina (puntos-trazos), pET que sobreexpresa la mutasa y cultivadas sin hidroxocobalamina (puntos), o bien pET que sobreexpresa la mutasa y cultivadas con hidroxocobalamina (trazos).

B. Clonación y expresión de metilmalonil-CoA epimerasa

La mm-CoA epimerasa de *Propionibacterium shermanii* se purificó y se usó para obtener la secuencia proteica del extremo N-terminal así como una secuencia peptídica interna a partir de los péptidos generados con LysC. El gen de la epimerasa se clonó usando sondas de hibridación diseñadas a partir de las secuencias peptídicas.

Se obtuvo y se cultivó *Propionibacterium freudenreichii* subsp. *shermanii* como se describe en la parte A. La purificación de mm-CoA epimerasa de *P. shermanii* se basó en una modificación del procedimiento publicado. El procedimiento utilizó un cultivo de 10 L, que se lisó por sonicación seguido de cromatografía en columna en el orden: DE-52, Hidroxiapatita, Fenil-Sepharose, intercambio iónico MonoQ, y C-8 RP HPLC.

Todas las operaciones se realizaron a 4°C, excepto el C-8 RP HPLC, que se realizó a temperatura ambiente, y todos los tampones contenían PMSF 0,1 mM, a menos que se indique otra cosa. El ensayo de la epimerasa se realizó esencialmente como se describe en la bibliografía. La concentración de proteínas se determinó usando el método de Bradford. No se determinó el rendimiento global de actividad epimerasa.

Más específicamente, se resuspendió pasta celular (75 g) en 50 mL de tampón (Tris-HCl 50 M, pH 7,5, KCl 0,1 M, PMSF 0,2 mM, EDTA 1 mM) y se sonicó usando una macrosonda con un diámetro de 1,2 cm. Con pulsos de 0,5 segundos ON y 0,3 segundos OFF, las células se sonicaron dos veces durante 30 segundos cada vez a una potencia de 4, y a continuación cinco veces durante 30 segundos cada vez a una potencia de 6. Se obtuvo un sobrenadante transparente, color ámbar (53,3 mL), después de centrifugar durante 35 minutos a 12.000 rpm.

ES 2 283 323 T3

El extracto bruto anterior se aplicó a una columna (diámetro 2,5 cm, altura 15 cm) de 73 mL de resina DE-52 equilibrada con Tris-HCl 50 mM, pH 7,5, KCl 0,1 M. La columna se lavó a 1 mL/min con tres volúmenes de columna del tampón anterior, y a continuación se aplicó un gradiente lineal hasta Tris-HCl 50 mM, pH 7,5, KCl 0,5 M a lo largo de siete volúmenes de columna. Se recogieron fracciones de 6 mL y se ensayaron para determinar la actividad epimerasa. La epimerasa se encontró predominantemente en el flujo a través y en varias fracciones tempranas. El flujo a través y las fracciones activas se combinaron (325 mL) y dializaron frente a 4 litros de Tris-HCl 50 mM, pH 7,5, glicerol al 10%, y a continuación 4 litros de fosfato de sodio 10 mM, pH 6,5, glicerol al 5% (volumen final 250 mL).

Se equilibró una columna de biogel de hidroxiapatita HTP de 7,5 mL (diámetro 1,5 cm, altura 16 cm) con fosfato de sodio 10 mM, pH 6,5, glicerol al 5%. Después de cargar la solución enzimática (usando inyecciones repetidas) y lavar con tres volúmenes de columna del tampón anterior, se efectuó un gradiente hasta fosfato de sodio 200 mM, pH 6,5, glicerol al 5%, a lo largo de 20 volúmenes de columna a un flujo de 1 mL/min. Se ensayaron las fracciones de 2 mL para determinar la actividad epimerasa, y las fracciones que contenían actividad epimerasa se combinaron en un total de 99 mL.

A la muestra de 99 mL anterior, se añadió sulfato de amonio sólido hasta una concentración final de 1,5 M, lentamente y con agitación a 4°C durante 30 minutos. Esta suspensión (100 mL) se cargó, por inyección repetida, en una columna de 6,6 mL (1 cm x altura 8,5 cm) de resina de Fenil-Sepharose equilibrada en tampón fosfato de sodio 20 mM, pH 6,5, sulfato de amonio 1,5 M. La columna se lavó a 1 mL/min con tres volúmenes de columna de este tampón, y a continuación se aplicó un gradiente lineal hasta tampón fosfato de sodio 20 mM, pH 6,5, glicerol al 10%, a lo largo de 24 volúmenes de columna. Después de ensayar las fracciones de 3 mL para determinar la actividad epimerasa, las fracciones que contenían actividad epimerasa se combinaron y dializaron frente a Tris-HCl 50 mM, pH 7,5.

Se equilibró una columna mono-Q 5/5 pre-empaquetada con Tris-HCl 25 mM, pH 7,5 a 0,5 mL/min. La muestra del paso anterior se cargó en la columna, que se lavó seguidamente con 5 volúmenes de columna del tampón anterior, y a continuación se aplicó un gradiente lineal hasta Tris-HCl 50 mM, pH 7,5, NaCl 1 M, glicerol al 5%, a lo largo de 50 volúmenes de columna. Las fracciones de 1 mL se ensayaron para determinar la actividad epimerasa. Varias fracciones que contenían actividad epimerasa se almacenaron por separado; la fracción con más actividad se usó para el siguiente paso de purificación.

Se equilibró una columna de fase inversa con agua que contenía ácido trifluoroacético al 0,1%; se inyectaron 120 μ L (concentrados a partir de 0,5 mL de la fracción activa anterior, usando un microconcentrador Amicon) en la columna a un flujo de 0,2 mL/min y se lavó durante 5 minutos con el sistema de disolvente anterior. A continuación se implementó un gradiente lineal a lo largo de 50 minutos hasta acetonitrilo que contenía ácido trifluoroacético al 0,1%. Los picos se recogieron manualmente y el pico correspondiente a la epimerasa (determinado por SDS-PAGE) se secó completamente, se resuspendió en agua y se almacenó a -80°C.

Para la digestión mediada por LysC de la epimerasa purificada por HPLC, la fracción de epimerasa (11751rp2-B, 200 μ L) recogida del análisis por HPLC de fase inversa se secó completamente y se resuspendió en 40 μ L de agua. A 30 μ L de la muestra se añadieron 5 μ L de Tris-HCl 1 M, pH 8, 1,5 μ L de DTT 0,1 M, 2 μ L de proteasa LysC (0,2 μ g). Una reacción de control contenía todos los componentes anteriores menos la epimerasa. Las reacciones se incubaron durante una noche a 37°C. Se diluyó una alícuota de la reacción (5 μ L) hasta 60 μ L con agua y se cargó en el HPLC, usando el mismo programa de HPLC que se usó para purificar la epimerasa. El HPLC analítico mostró que la digestión con LysC no era completa. Se añadió una alícuota adicional (0,2 μ g) de LysC a las reacciones y se continuó la incubación durante una noche a 37°C. Después de la incubación durante una noche, se diluyó una alícuota de la reacción (5 μ L) hasta 60 μ L con agua y se sometió al HPLC. El HPLC mostró que la digestión era completa. Se cargó el resto de la reacción en el HPLC y se recogieron picos individuales manualmente. El análisis por HPLC de la reacción de control no mostró péptidos significativos provenientes de la auto-digestión de la LysC.

Una alícuota de la epimerasa pura, así como un péptido recogido a partir del procedimiento descrito anteriormente, se entregaron para la secuenciación de aminoácidos del extremo N-terminal. Basándose en las secuencias de aminoácidos anteriores, se diseñaron varios cebadores degenerados como se describe más abajo, que introduzcan sitios de restricción únicos en cada extremo del producto de PCR eventual. Estos cebadores se usaron en la PCR con ADN genómico de *P. shermanii* para obtener un producto de 200 pares de bases, que se clonó en un vector BluescriptTM (Stratagene) y se entregó para su secuenciación.

Se preparó una genoteca de cósmidos de *P. shermanii*, esencialmente como se describe en el manual de cósmidos de Stratagene. El título para esta genoteca de cósmidos fue de aproximadamente 11 cfu (unidades de formación de colonias, del inglés "colony forming units") por μ L, para un rendimiento total de 5556 cfu. Se preparó una genoteca de plásmidos de *P. shermanii* digiriendo ADN genómico de *P. shermanii* con *SacI* y ligando la mezcla resultante en un vector BluescriptTM también cortado con *SacI*. Para determinar el tamaño medio del inserto (2 kb), se digirieron diez colonias al azar con *SacI*. La mezcla de ligación se reutilizó para transformar 5 veces, se combinó y se dispuso en una gran placa de LB (carb), dando como resultado un "césped" de colonias que se rasparon juntas y se resuspendieron en LB como la genoteca de plásmidos. El título de esta genoteca de plásmidos fue de aproximadamente 64.000 cfu por μ L.

Se prepararon varios cebadores degenerados basados en las secuencias de aminoácidos y se usaron en la PCR con ADN genómico de *P. shermanii* para obtener un producto de 180 pares de bases, que se clonó en un vector

ES 2 283 323 T3

Bluescript™ y se secuenció. Se prepararon varias sondas diferentes. La primera sonda se preparó usando el método de cebado al azar para incorporar ³²P o bien digoxigenina en el fragmento de la epimerasa. Se preparó una sonda a partir del fragmento clonado por amplificación del fragmento por PCR, usando el método de marcaje con digoxigenina. El producto de PCR se aisló mediante gel, se cuantificó, y se usó como sonda para la genoteca de cósmidos. Las colonias que hibridaban con la sonda se re-estriaron a partir de las placas originales, y se escogieron cinco colonias de las placas re-estriadas, se aislaron los cósmidos, y las secuencias insertadas se sometieron a escrutinio para detectar el gen de la epimerasa por PCR. Varios cósmidos que se consiguieron como positivos por PCR sobre la base de la secuencia de ADN de la epimerasa, se sometieron a secuenciación de ADN usando cebadores específicos para la epimerasa. El cósmido designado 117-167-A7 contenía la secuencia completa de la epimerasa.

La secuencia del supuesto gen de epimerasa contenido en el cósmido 117-167-A7 se alineó con la secuencia del extremo N-terminal de la epimerasa ya conocida. Los varios cientos de pares de bases en dirección 3' desde esta secuencia se tradujeron en los tres marcos de lectura y se encontró un codón de terminación en uno de los marcos que proporcionaba una proteína del tamaño esperado. La secuencia entera se usó para buscar en la base de datos de proteínas a través de un análisis con BLAST, y la secuencia mostró una alta homología con la secuencia de una supuesta epimerasa de *S. coelicolor* identificada conforme a los métodos de la invención. Se diseñaron cebadores para PCR basados en la secuencia de ADN de la epimerasa de *P. shermanii* clonada y el gen se amplificó del ADN genómico de *P. shermanii* con sitios *NdeI* y *BamHI* en el extremo 5', se destruyó un sitio *NdeI* interno cerca del extremo 5', y se introdujeron sitios *NheI* y *AvrII* en el extremo 3'. Después de la PCR, el producto de 447 pb se clonó en un vector Bluescript (143-6-11) y se secuenció. También, se diseñaron cuatro cebadores de secuenciación adicionales para proporcionar un exceso de varias veces de cobertura del gen de la epimerasa. La secuencia completa del gen de la epimerasa proporcionado en forma aislada y recombinante por la presente invención se proporciona más abajo como SEC ID N° 1, y la secuencia proteica de la epimerasa se proporciona más abajo como SEC ID N° 2. El gen de la epimerasa se clonó a continuación en el vector de expresión pET; la construcción se nombró pET-epsperm.

Para la clonación de genes de epimerasa de *B. subtilis* (descrito por Haller *et al.*, *ut supra*) y *S. coelicolor* (del cósmido 8F4 en el proyecto de secuenciación del genoma de *S. coelicolor*), se diseñaron cebadores para amplificar por PCR estos genes a partir de sus ADNs genómicos respectivos y para incorporar un sitio *PacI*, o bien *NdeI* en el extremo 5', y un sitio *NsiI* en el extremo 3'. Los productos de PCR se clonaron en un vector Bluescript™ y se secuenciaron. Se obtuvieron clones libres de mutaciones para la epimerasa de *S. coelicolor*, pero la epimerasa de *B. subtilis* contenía mutaciones en dos puntos en los tres clones ensayados: T por C en el par de bases 37, y A por G en el par de bases 158. Cuando se repitió la PCR para este gen de epimerasa y el producto se clonó y secuenció, las mismas mutaciones estaban presentes, lo que implicaba que la secuencia original era errónea. Las epimerasas clonadas de *B. subtilis* y *S. coelicolor* se clonaron como fragmentos *NdeI/NsiI* en un vector intermedio 116-172a, un plásmido pET Bluescript™ que contenía el promotor T7 y secuencias de terminación. Las epimerasas clonadas de *B. subtilis* y *S. coelicolor* son pET-epsub y pET-epcoel, respectivamente. Los genes de epimerasa se cortaron y separaron también junto con el promotor T7 como fragmentos *PacI/NsiI*, como se muestra esquemáticamente abajo

— *PacI* — promotor T7 — gen de epimerasa — *NsiI* —

y se clonaron en el vector 133-9b digerido con *PacI/NsiI*, para formar un único operón con el gen de la epimerasa situado en dirección 3' respecto a los dos genes de mutasa. El gen de la epimerasa de *P. shermanii* se clonó como arriba excepto que se clonó en 116-172a como un fragmento *NdeI/AvrII*, se cortó junto con el promotor T7 como un fragmento *PacI/NheI*, y se clonó en 133-9b entre los sitios *PacI* y *NheI*. Las construcciones son pET-mutAB-T7-epsperm, pET-mutAB-T7-epsub, y pET-mutAB-T7-epcoel.

Como alternativa a la mutasa de *P. shermanii*, *S. coelicolor* y *B. subtilis*, se puede clonar por PCR a partir de ADN genómico de *E. coli* el gen único de la *Sbm* ("sleeping beauty mutase"). Se preparó ADN genómico de *E. coli* BL21(DE3)/*panD* usando un "kit" (estuche) adquirido de Qiagen. El gen de *Sbm* ("Sleeping beauty mutase", una metilmalonil-CoA mutasa) se amplificó por PCR a partir de ADN genómico de *E. coli* BL21(DE3)/*panD*. El fragmento de PCR se aisló mediante gel, se clonó en PCRscript y se secuenció para proporcionar el clon libre de mutaciones 143-11-54. Cortado y separado como un fragmento *NdeI/SacI*, *sbm* se clonó en pET22b, de ahí como un fragmento *NdeI/XhoI* en pET16b para introducir una cola de histidinas en el extremo N-terminal (143-49-2). *Sbm* se clonó también entre *NdeI* y *SpeI* en 116-95B.43, un vector pET22b que permite la clonación subsiguiente de los genes de epimerasa en dirección 3' respecto al *sbm*. Esa construcción se nombró 143-40-39.

Se cultivaron células de la cepa BL21(DE3) que contenían pET-epsperm, pET-epsub, pET-epcoel o un vector pET de control durante una noche a 37°C en 2 mL de LB que contenía 100 µg/mL de carbenicilina. El cultivo de partida (250 µL) se usó para inocular 25 mL de LB que contenía 100 µg/mL de carbenicilina. Los cultivos se dejaron crecer a 37°C hasta una OD de aproximadamente 0,4, seguidamente se indujeron con IPTG a una concentración final de 1 mM y se cultivaron durante 3 horas más a 30°C. Las células se recogieron por centrifugación a 4000 rpm durante 10 minutos, y los sedimentos se almacenaron a -80°C antes del ensayo. La epimerasa de *P. shermanii* se expresó bien en *E. coli*; el análisis en gel con SDS reveló una proteína sobreexpresada de aproximadamente 22 KDa. La epimerasa de *S. coelicolor* también se expresó bien, con un peso molecular de aproximadamente 19 KDa, y la epimerasa de *B. subtilis* se expresó, pero mayormente en cuerpos de inclusión (una banda apenas visible está presente a aproximadamente 19 KDa), lo cual puede superarse mediante el uso de sistemas de expresión alternativos.

Se midió la actividad epimerasa en los extractos brutos de *E. coli* que hospedaba pET-epsherm, pET-epsub, pET-epcoel, o bien un vector pET de control. El ensayo de la epimerasa acopla la transcarboxilasa, que convierte (S)-metilmalonil-CoA en propionil-CoA, con malato deshidrogenada, que convierte NDAH en NAD⁺, produciendo un descenso en la absorbancia a 340 nm. El ensayo se inicia con una mezcla racémica de (R,S)-metilmalonil-CoA; cuando el isómero (S) se consume como se describe más abajo, se observa una tasa de fondo estacionaria a aproximadamente un décimo de la tasa inicial. Cuando se añade al ensayo un extracto que contiene epimerasa, el isómero (R) se convierte en (S), dando como resultado una disminución adicional de la absorbancia. En los extractos de *E. coli* brutos, sin embargo, se observa una tasa de fondo significativa, debido probablemente a una NADH oxidasa endógena. Por tanto la epimerasa debe expresarse en un nivel suficientemente alto para concluir que es activa. El ensayo se llevó a cabo como sigue.

El sedimento de aproximadamente 20 mL de cultivo se descongeló y resuspendió en 2 mL de tampón de ensayo IX que contenía un comprimido de un cóctel de inhibidores de proteasa. Las células se rompieron por sonicación (dos ciclos de sonicación de 30 segundos cada uno a una potencia establecida de 2 [pulso ON 0,5 s/pulso OFF 0,5 s]). Después de centrifugar durante 10 minutos a 13000 rpm en una centrífuga Eppendorf, los sobrenadantes se guardaron para ensayar. La actividad metilmalonil-CoA epimerasa se ensayó usando una modificación del método de Leadly *et al.*, *Biochem. J.* 197: 413-419, "Purification and characterization of methylmalonyl CoA epimerase from *Propionibacterium shermanii*" (1981). Los ensayos se realizaron a 30°C con una cubeta de plástico de 1 cm de recorrido, en un volumen final de 1,5 mL. Las mezclas de reacción contenían tampón fosfato de potasio 0,2 M, pH 6,9, sulfato de amonio 0,1 M, piruvato de sodio 5 mM, (2R,2S)-metilmalonil-CoA 0,08 mM, 0,05 unidades de transcarboxilasa parcialmente purificada, NADH 0,16 mM, y 2,5 unidades de malato deshidrogenasa. La reacción se inició con (2R,2S)-metilmalonil-CoA y se monitorizó el descenso de absorbancia a 340 nm, que reflejaba la desaparición del isómero 2S. Cuando el descenso de absorbancia a 340 nm alcanzó el nivel basal (usualmente alrededor de un 10% de la tasa inicial de la transcarboxilasa), se añadió un extracto que contenía epimerasa y se observó un descenso adicional en la absorbancia. Los productos químicos y enzimas usados en el ensayo de epimerasa se adquirieron de Sigma, excepto la transcarboxilasa, que se obtuvo como una preparación en bruto de Case Western Reserve.

Los extractos brutos que hospedaban ambas epimerasas de *P. shermanii* y *S. coelicolor* tenían actividades específicas (aproximadamente 30 unidades/mg) al menos 10 veces mayores que la del control. Sin embargo, no se observó actividad por encima del nivel de fondo en el extracto que hospedaba la epimerasa de *B. subtilis*, posiblemente porque no se expresaba en un nivel suficientemente alto o, como se señala anteriormente, se expresaba como cuerpos de inclusión insolubles. La construcción pET-mutAB-T7-epsherm se expresó también en *E. coli*. El extracto bruto resultante contenía actividad epimerasa que estaba significativamente por encima del nivel de fondo; por consiguiente, la epimerasa es funcional en esta construcción. La mutasa no interfería en el ensayo de epimerasa, porque estas células se cultivaron sin adición de hidroxocobalamina, el cofactor para la actividad mutasa. Estos resultados muestran que se puede expresar tanto mutasa activa como epimerasa activa en una célula de *E. coli*. Estos resultados muestran también que la metilmalonil-CoA epimerasa de *P. shermanii* se clonó, se expresó en *E. coli*, y activa, y que la supuesta epimerasa de *S. coelicolor* es una metilmalonil-CoA epimerasa. Estos genes pueden integrarse en el cromosoma de una cepa *panD* de *E. coli* u otra cepa y usarse para la producción de policétidos construidos por completo o en parte a partir de metilmalonil-CoA.

Ejemplo 2

Producción de Metilmalonil-CoA en Levaduras

Este ejemplo describe la construcción de cepas de *Saccharomyces cerevisiae* optimizadas para la sobreproducción de policétidos. En particular, este ejemplo describe la construcción de cepas hospedantes de levadura que (i) producen sustratos y enzimas de modificación post-traduccionales necesarias para expresar policétidos producidos por policétido sintasas modulares; (ii) tienen las deficiencias nutricionales necesarias para permitir la selección de positivos de al menos tres plásmidos compatibles; y/o (iii) son adecuadas para permitir el marcaje radioactivo de combinados de acil-CoA y policétido sintasas, y demuestra que tales cepas pueden expresar una PKS modular y producir un policétido complejo en niveles adecuados para el desarrollo comercial. En este ejemplo las referencias se citan mediante un número correspondiente a la lista numerada de referencias más abajo.

Con modificaciones apropiadas de la cepa, *S. cerevisiae* es un hospedante ideal para la producción de policétidos. *S. cerevisiae* es capaz de producir niveles muy altos de policétidos. La introducción de un gen para la PKS iterativa, 6-MSAS, junto con el gen para Sfp, una P-pant transferasa de *B. subtilis*, condujo a la producción de la impresionante cifra de 2 g/L de 6-MSA en matraces agitados sin optimización Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9. La genética de las levaduras se entiende muy bien. Los genes pueden insertarse fácilmente en el cromosoma, y la secuencia del genoma completo proporciona conocimientos relevantes respecto a las rutas metabólicas y sitios de inserción neutrales. Además, están disponibles varios promotores fuertes y controlables. Las proteínas tienen menos tendencia a formar cuerpos de inclusión en levaduras, comparado con *E. coli*. Las levaduras tienen un tiempo de duplicación relativamente corto en comparación con los organismos nativos productores de policétidos. *S. cerevisiae* tiene un tiempo de duplicación de 1 a 2 h comparado con las 4 a 24 h de un productor típico de policétidos, lo cual tiene beneficios obvios en desarrollo genético, desarrollo de procesos y producción a gran escala.

El hecho de que las levaduras crecen como células individuales proporciona un beneficio adicional frente a los organismos filamentosos (productores típicos de policétidos). Las fermentaciones miceliarias son viscosas y se comportan frecuentemente como fluidos no newtonianos. Esta reología del fluido proporciona un obstáculo importante al científico del proceso en términos tanto del transporte uniforme de nutrientes a las células como del manejo del caldo de fermentación. Emplear una levadura como hospedante, incluso a elevadas densidades celulares, evita tales impedimentos. Debido al extenso historial de las levaduras en la producción de proteínas en células individuales y la expresión de proteínas recombinantes, se han desarrollado protocolos de fermentación escalables. Las levaduras pueden cultivarse en fermentaciones de alimentación discontinua hasta densidades celulares muy altas (>100 g/L de biomasa), comparado con los productores típicos de policétidos (10-20 g/L de biomasa). Por consiguiente, comparando organismos con la misma productividad específica (g policétido/g biomasa/día), las levaduras proporcionarían una productividad volumétrica mayor (g policétido/L/día). Finalmente, *S. cerevisiae* está clasificado por la FDA como un organismo “Generalmente Considerado Como Seguro” (GRAS, del inglés “Generally Regarded As Safe”). La clasificación GRAS facilitará la aprobación de fármacos producidos en levaduras comparadas con otras células hospedantes.

S. cerevisiae tiene también desventajas como hospedante para la biosíntesis de policétidos, la mayoría relacionadas con el hecho de que las levaduras no evolucionaron para producir policétidos. Las levaduras no contienen metilmalonil-CoA, un precursor necesario para la biosíntesis de muchos policétidos. Las levaduras no tienen una P-pant transferasa adecuada capaz de la modificación post-traduccional necesaria de los dominios ACP de una PKS. Los codones de las levaduras tienden a A+T, mientras que la mayoría de los productores de policétidos tienen muchos codones G+C; por consiguiente, las levaduras pueden tener bajas cantidades de algunos tARNs necesarios para la expresión de genes de PKS. En este ejemplo se describe la corrección de estas deficiencias, y la invención proporciona también células hospedantes de levadura modificadas útiles para facilitar el análisis del éxito.

Otras cuestiones potenciales caso por caso con las levaduras incluyen la posibilidad de que algunos productos policétidos pueden ser tóxicos o pueden requerir modificaciones adicionales para su maduración (p.ej., glicosilación, hidroxilación por P450). Pueden tomarse varios métodos proporcionados por la invención para sortear estas cuestiones si aparecieran. Para la toxicidad, la producción puede controlarse para tener lugar en la fase estacionaria del crecimiento (como con la producción de 6-MSA); los factores de resistencia del hospedante salvaje pueden introducirse en la levadura hospedante (p.ej., metilación de ribosomas para algunos antibióticos); puede producirse un precursor no tóxico del policétido y convertirse *ex vivo* (p.ej., producir 6-dEB en una cepa y convertirlo en eritromicina en otra), y otros. Pueden lograrse modificaciones adicionales del policétido por clonación y expresión de enzimas en la cepa hospedante, transformación química o enzimática y/o transformación biosintética en una segunda cepa (p.ej., convertir análogos de 6-dEB en análogos de eritromicina proporcionando 6-dEB a una cepa de *Streptomyces* o *Saccharopolyspora* con capacidad de glicosilación e hidroxilación por P450).

La mayoría de las PKSs modulares requieren uno de ellos o ambos entre malonil-CoA o (2*S*)-metilmalonil-CoA como fuente de unidades de 2 carbonos para la biosíntesis de policétidos. Las reservas de malonil-CoA en las levaduras son moderadamente suficientes para la síntesis de policétidos, como ilustra la producción de grandes cantidades de 6-MSA en levaduras. Sin embargo, *S. cerevisiae* no produce (2*S*)-metilmalonil-CoA y no posee rutas biosintéticas para la biosíntesis de metilmalonil-CoA. Por tanto, debe introducirse una ruta biosintética heteróloga en *S. cerevisiae* para respaldar la biosíntesis de policétidos que usen (2*S*)-metilmalonil-CoA como precursor.

Existen tres vías o rutas biosintéticas para la síntesis de metilmalonil-CoA que pueden introducirse por ingeniería genética en una levadura, como se muestra en la Figura 5. Se ha demostrado que estas rutas producen metilmalonil-CoA en *E. coli* y pueden usarse para producir metilmalonil-CoA en levaduras. Este ejemplo describe la identificación de un sistema para la producción de metilmalonil-CoA en levaduras, y un método para introducirlo en el cromosoma de la levadura.

La ruta de la metilmalonil-CoA mutasa dependiente de vitamina B12 produce (2*R*)-metilmalonil-CoA a partir de succinil-CoA. El (2*R*)-metilmalonil-CoA se convierte en (2*S*)-metilmalonil-CoA a través de la metilmalonil-CoA epimerasa, como se muestra anteriormente. Estas enzimas están presentes en una variedad de organismos, pero no en levaduras; las búsquedas con BLAST de las bases de datos genómicas disponibles revelan al menos 10 metilmalonil-CoA mutasas y 10 metilmalonil-CoA epimerasas en varios organismos. La metilmalonil-CoA mutasa de *Propionibacterium shermanii* se ha expresado en *E. coli* como la apo-enzima, que requiere la adición de vitamina B12 para su actividad *in vitro* (McKie, N., *et al.*, Adenosylcobalamin-dependent methylmalonyl-CoA mutase from *Propionibacterium shermanii*. Active holoenzyme produced from *Escherichia coli*. Biochem. J., 1990. 269(2): p. 293-8). Usando un medio que permita la captación del precursor de vitamina B12 hidroxocobalamina (Amaratunga, M., *et al.*, A synthetic module for the *metH* gene permits facile mutagenesis of the cobalamin-binding region of *Escherichia coli* methionine synthase: initial characterization of seven mutant proteins. Biochemistry, 1996. 35(7): p. 2453-63), y conforme a los métodos de la invención, se puede expresar la holoenzima metilmalonil-CoA mutasa de *P. shermanii* activa en *E. coli* y producir (2*R*)-metilmalonil-CoA en tales células. Además, puede emplearse la metilmalonil-CoA mutasa de una sola unidad de *E. coli*. La presente invención proporciona también los genes que codifican la metilmalonil-CoA epimerasa de *B. subtilis*, *P. shermanii* y *S. coelicolor* y métodos para usarlos en la conversión de (2*R*)-metilmalonil-CoA en el diastereoisómero (2*S*)-metilmalonil-CoA necesario. Un método preferido es expresar en levaduras la metilmalonil-CoA mutasa de *E. coli*, porque es un único ORF, y los codones necesarios son abundantes en levaduras. Alternativamente, puede usarse la enzima de *P. shermanii*.

PCC cataliza la carboxilación dependiente de biotina de propionil-CoA para producir (2S)-metilmalonil-CoA, como se muestra anteriormente; la ruta incluye también un proteína portadora de biotina/biotina carboxilasa. En *S. coelicolor*, Rodríguez y Gramajo identificaron los genes para PCC (*pccB*) y ambas proteína portadora de biotina/biotina carboxilasa (*accA1*) (Rodríguez, E. y H. Gramajo, *Genetic and biochemical characterization of the alpha and beta components of a propionyl-CoA carboxylase complex of Streptomyces coelicolor* A3(2). *Microbiology*, 1999. **145**(Pt 11); p. 3109-19). La introducción en *E. coli* de *pccB* y *accA1* de *S. coelicolor* junto con propionil-CoA ligasa (como suministro de propionil-CoA), da como resultado la producción de metilmalonil-CoA en ese organismo. Una búsqueda de la base de datos genómica revela a *B. subtilis* como fuente adicional de la enzimas implicadas en la ruta de la PCC.

En una realización de la invención, se pueden expresar los genes *pccB* y *accA1* de *S. coelicolor* en levaduras, porque éstos se expresan y las proteínas son funcionales en *E. coli*. Si la utilización de codones resultara subóptima al expresar los genes de *S. coelicolor* en levaduras, pueden emplearse homólogos de *B. subtilis*. Si los niveles de propionil-CoA fueran subóptimos para PCC, puede coexpresarse una propionil-CoA ligasa en la levadura hospedante. La propionil-CoA intracelular puede aumentarse enormemente en *E. coli* expresando la propionil-CoA ligasa de *Salmonella*, PrpE, y suplementando el medio de crecimiento con propionato, como se describe más abajo.

Un método adicional para la producción de (2S)-metilmalonil-CoA proporcionado por la presente invención utiliza los genes *matB* y *matC* de *Rhizobium* (An, J.H. e Y.S. Kim, *A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in Rhizobium trifolii-cloning, sequencing, and expression of the enzymes in Escherichia coli*. *Eur. J. Biochem.*, 1998. **2257**(2): p. 395-402); o *S. coelicolor* (véase esquemáticamente arriba). Los genes *matABC* codifican una ruta biosintética que convierte malonato en acetyl-CoA a través de la formación de malonyl-CoA por medio de MatB y la descarboxilación subsiguiente por MatA. MatB, la malonyl-CoA ligasa, acepta también metilmalonato como sustrato (An, J.H. e Y.S. Kim, *A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in Rhizobium trifolii-cloning, sequencing, and expression of the enzymes in Escherichia coli*. *Eur. J. Biochem.*, 1998. **2257**(2): p. 395-402); y cataliza la formación de metilmalonil-CoA. Los sustratos malonato o metilmalonato entran en la célula a través de un transportador diácido, el producto del gen *matC*. Khosla *et al.*, han demostrado que cuando *E. coli* que contiene el *matBC* de *Rhizobium* se alimenta con (2R,2S)-metilmalonato, se produce (2R,2S)-metilmalonil-CoA. Además, cuando una cepa de *S. coelicolor* que expresa los genes para la síntesis de la aglicona policétidica, 6-desoxieritronolida B (6-dEB), y que contiene *matBC* de *Rhizobium* se alimenta con metilmalonato, se observa un aumento de 3 veces en la producción de 6-dEB. Conforme a los métodos de la invención, se pueden expresar los genes *matB* y *matBC* de *Rhizobium* en levaduras, porque éstos se expresan y las proteínas son funcionales en *E. coli* y *S. coelicolor*, o, alternativamente, los genes *matBC* de *S. coelicolor*.

Las PKSs activas requieren la fosfopanteteinilación post-traducciona en cada ACP de cada módulo, pero las levaduras no contienen una P-pant transferasa con la especificidad necesaria (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide producing prokaryotic and eukaryotic hosts*. *Proc. Natl. Acad. Sci. USA*, 1998. **95**(2): p. 505-9). Trabajos previos (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide producing prokaryotic and eukaryotic hosts*. *Proc. Natl. Acad. Sci. USA*, 1998. **95**(2): p. 505-9) han demostrado que la introducción del gen de la P-pant transferasa de *B. subtilis*, *sfp*, en levaduras da como resultado una Sfp expresada capaz de modificar una PKS iterativa, 6-MSAS. Gokhale *et al.*, demostraron que los dominios ACP en la PKS DEBS son sustratos para la Sfp, así que Sfp es una enzima de modificación general para PKSs (Gokhale, R.S., *et al.*, *Dissecting and exploiting intermodular communication in polyketide synthases*. *Science*, 1999. **284**(5413): p. 482-5). En las células hospedantes de levaduras preferidas de la invención, el gen *sfp* se inserta en un sitio neutral del cromosoma de la levadura.

En el desarrollo de un sistema para producir policétidos y optimizar los procedimientos de fermentación, la capacidad para medir concentraciones intracelulares de sustratos (esto es, acil-CoAs) y de la PKS resulta beneficiosa. En la mayoría de las células, los ésteres de CoA no están presentes en cantidades suficientes para permitir la medición directa por HPLC usando detección ultravioleta u otros métodos simples de detección. En *E. coli*, el método de elección para cuantificar los combinados de CoAs es alimentar con [³H] β -alanina un mutante deficiente en aspartato descarboxilasa (*PanD*), que no puede producir β -alanina endógena (Jackowski, S. y C.O. Rock, *Regulation of coenzyme A biosynthesis*. *J. Bacteriol.*, 1981. **148**(3): p. 926-32). La cepa *PanD* incorpora aproximadamente diez veces más radiactividad en los combinados de CoAs que lo que hace *E. coli* salvaje. Como la β -alanina es un precursor directo de CoA, el marcaje radiactivo entra en el combinado de CoAs sin dilución, y los acil-CoAs pueden separarse en el HPLC y cuantificarse por medición de la radiactividad. Como no hay dilución del radioisótopo, la radiactividad medida refleja las concentraciones intracelulares exactas de acil-CoAs.

Las búsquedas con BLAST no revelaron un homólogo del *PanD* de *E. coli* en el genoma de levaduras; sin embargo, la levadura puede ser auxótrofa para β -alanina o pantotenato. De hecho, para la biosíntesis de CoA, la levadura requiere bien pantotenato exógeno, que entra en la célula a través del transportador Fen2p, o β -alanina exógena, que entra por medio de la aminoácido permeasa general (Gap1p) (Stolz, J. y N. Sauer, *The fenpropimorph resistance gene FEN2 from Saccharomyces cerevisiae encodes a plasma membrane H⁺ pantothenate symporter*. *J. Biol. Chem.*, 1999. **274**(26): p. 18747-52). La [³H] β -alanina se incorpora a los combinados de CoAs de la levadura (véase más abajo), pero por el momento se desconoce si tiene lugar la dilución del isótopo debido a la producción de β -alanina endógena por alguna ruta desconocida. Por consiguiente, para posibilitar la cuantificación, se puede determinar la actividad específica de los combinados de CoAs en la levadura marcada con [³H] β -alanina exógena. Las células que producen los policétidos

expresan generalmente bajos niveles de PKSs de alto peso molecular, que son apenas detectables en SDS-PAGE usando tinción de proteínas. La capacidad para marcar CoA con [³H] β-alanina puede usarse también para cuantificar una PKS expresada en las células hospedantes porque el resto de fosfopanteteína de la CoA que contiene β-alanina se transfiere al dominio APC en cada módulo de una PKS. Por consiguiente, conociendo la actividad específica de las CoAs intracelulares marcadas, puede cuantificarse una PKS simplemente por radiactividad tras SDS-PAGE.

El contenido en G+C de la mayoría de los genes de PKS está en el intervalo del 60 al 70%, mientras que el de los genes de levaduras es el 40%. Por consiguiente, algunos tARNs necesarios para traducir los genes de PKS son escasos (pero no ausentes) en levaduras. Sin embargo, se han expresado muchos genes con alto contenido en G+C en levaduras. Como ejemplos, el gran gen (1560 pb) DHFR-TS de *Leishmania major* (63% G+C) se expresa bien en levaduras, a pesar del hecho de que contiene varios codones raramente usados en levaduras (Grumont, R., W. Sirawaraporn, y D.V. Santi, *Heterologous expression of the difunctional thymidylate synthase-dihydrofolate reductase from Leishmania major*. *Biochemistry*, 1988. **27**(10): p. 3776-84). Además, como se menciona más abajo, la PKS 6-MSAS (G+C = 58%) se expresa también bien en levaduras (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide producing prokaryotic and eukaryotic hosts*. *Proc. Natl. Acad. Sci. USA*, 1998. **95**(2): p. 505-9). Por consiguiente, se puede demostrar la aplicabilidad general de un sistema de expresión en levaduras sin la preocupación inicial de problemas potenciales de utilización de codones. De todas formas, si una PKS no se expresa bien en levaduras, la presente invención proporciona varios métodos para resolver un problema de “utilización de codones” observado con un policétido en particular.

Primero, se pueden cambiar los codones en el extremo 5' del gen para reflejar los encontrados más frecuentemente en genes de levaduras. Batard *et al.*, (Batard, Y., *et al.*, *Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems [In Process Citation]*). *Arch. Biochem. Biophys.*, 2000. **379**(1): p. 161-9; emplearon satisfactoriamente un método similar para expresar en una levadura genes de trigo para un P450 y una P450 reductasa con alto contenido de G+C (56%) y fuerte tendencia a una utilización de codones desfavorable para la levadura. Otro método es introducir genes de tARN de levaduras con anti-codones codificados para representar los codones comunes en las secuencias de PKS. Un método similar se ha usado satisfactoriamente en *E. coli* para intensificar la expresión de genes ricos en G+C (Carstens, C.-P., *et al.*, *New BL21-CodonPlus™ Cells Correct Codon Bias in GC-Rich Genomes*. *Strategies Newsletter from Stratagene Corp.*, 2000. **13**(1): p. 31-33), incluyendo genes de PKS de Actinomicetos. Un tercer método es sintetizar químicamente el gen con codones optimizados para la expresión en levaduras. El coste para contratar la síntesis de un gen de 30.000 pb (p.ej., una PKS de ~6 módulos), incluyendo la verificación de la secuencia, es aproximadamente \$3 por base, o aproximadamente \$100.000. Para un producto valioso (p.ej., eptilona), el coste no es prohibitivo.

En una realización ilustrativa de la invención, se emplea una cepa de levadura deficiente en la biosíntesis de Ura, Trp, His y Leu como hospedante para permitir la selección de plásmidos que contienen estos marcadores. Este hospedante se modifica conforme a los métodos de la invención introduciendo genes que producen el sustrato metilmalonil-CoA necesario y la P-pant transferasa para las modificaciones post-traduccionales de PKSs. Éstos se integran preferiblemente en el cromosoma de la levadura, porque son necesarios para la producción de cualquier policétido. Para validar la expresión funcional de los genes del sustrato, se pueden medir las reservas de metilmalonil-CoA. Para validar la actividad P-pant transferasa, se puede coexpresar 6-MSAS y medir la [³H] fosfopanteteinilación de la enzima así como la producción de 6-MSA. Si alguno de ellos resultara deficiente, se puede aumentar el número de copias del gen.

Para la expresión de genes de PKS, se pueden usar vectores de replicación basados en el replicón del círculo de 2 micras, porque los plásmidos pueden tener que rescatarse para su análisis si surgiera un problema. Un “cluster” de genes de PKS modular típico (p.ej., 3 ORFs, ~10 kb cada uno, como en la eritromicina) puede introducirse en tres o más vectores; tales plásmidos (que contienen marcadores Ura, Trp y Leu) están disponibles y son similares a los usados en los estudios de expresión de 6-MSAS en levaduras. Una PKS que consiste en tres grandes proteínas puede reconstituirse funcionalmente a partir de genes expresados por separado (Xue, Q., *et al.*, *A multiplasmid approach to preparing large libraries of polyketides*. *Proc. Natl. Acad. Sci. USA* 1999. **96**(21): p. 11740-5). Una vez que se establece un sistema para una PKS particular de interés, se pueden integrar los genes de la PKS en sitios neutrales estables del cromosoma.

Los promotores preferidos incluyen el promotor de la alcohol deshidrogenasa 2 (ADH2) represible por glucosa y el promotor inducible por galactosa (GAL1). El primero se ha usado para producir grandes cantidades del policétido 6-MSA en levaduras, y el último es muy controlable por la galactosa en el medio.

Una PKS modular modelo que se puede usar para optimizar la levadura hospedante es la bien estudiada DEBS1. En este sistema modelo, el primer ORF de la PKS modular para la biosíntesis de eritromicina (DEBS1) se ha fusionado con un dominio tioesterasa (TE) y produce una lactona tricetónica fácilmente detectable cuando se expresa en *S. coelicolor*, y más recientemente en *E. coli* (Kao, C.M., *et al.*, *Engineered biosynthesis of a triketide lactone from an incomplete modular polyketide synthase*. *J. Am. Chem. Soc.*, 1994. **116**(25): p. 11612-11613; Cortes, J., *et al.*, *Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage*. *Science*, 1995. **268**(5216): p. 1487-9). El gen contiene 2 módulos de PKS, tiene aproximadamente 12 kb, y produce una proteína que tiene 300 kDa. Este modelo permite optimizar el hospedante modificado por ingeniería genética con respecto a los niveles de acil-CoAs y modificaciones post-traduccionales, la PKS con respecto al contenido en G+C, y desarrollar los métodos analíticos necesarios. Una vez optimizado para DEBS1, se puede expresar cualquier PKS modular dada.

Previamente, se ha demostrado que el gen fúngico que codifica la ácido 6-metilsalicílico sintasa (6-MSAS) de *Penicillium patulum* se expresaba en *S. cerevisiae* y *E. coli* y que se producía el policétido ácido 6-metilsalicílico (Kealey, J.T., et al., *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9). En ambos hospedantes bacteriano y de levadura, la producción del policétido requirió la coexpresión de 6-MSAS y una fosfopanteteinil-transferasa (Sfp) heteróloga, que se requería para convertir la apo-PKS expresada en la holoenzima. La producción de 6-MSA en *E. coli* era dependiente de temperatura y glicerol y los niveles de producción (~60 mg/L) eran inferiores que los del hospedante nativo, *P. patulum*. En la levadura, los genes de 6-MSAS y *sfp* se coexpresaron a partir de plásmidos de replicación separados, y la expresión génica estuvo dirigida por el promotor de la alcohol deshidrogenasa 2 (ADH2) represible por glucosa. En una fermentación no optimizada en matraces con agitación, el sistema en la levadura produjo 6-MSA en niveles de 2.000 mg/mL. Ésta fue la primera publicación de la expresión de un gen de PKS intacto en levaduras o *E. coli*, y demostró que pueden producirse niveles extraordinariamente altos de policétidos en levaduras.

Previamente, se desarrolló un sistema de dos vectores para la expresión heteróloga de tres genes que comprendían el “cluster” de genes de policétidos producidos por DEBS (Ziermann, R., Betlach, M., *A Two-vector System for the Production of Recombinant Polyketides in Streptomyces*. J. Ind. Microbiol. Biotech., 2000, **24**:46-50). Se clonaron los genes individuales de DEBS y combinaciones por parejas de dos de tales genes en dirección 3' respecto al promotor de actinorrodina (*act1*) en dos vectores de *Streptomyces* compatibles: el vector de replicación autónoma, pKAO127'Kan', y el vector de integración, pSET152. Cuando los plásmidos resultantes se usaron para transformar, bien simultáneamente o secuencialmente, el hospedante heterólogo, *Streptomyces lividans* K4-114, se produjo el producto policétidico, 6-dEB. Este trabajo demostró que los genes de DEBS podían dividirse y expresarse en plásmidos separados, y que en el hospedante heterólogo tuvo lugar la *trans*-complementación eficiente de las proteínas de las subunidades de la policétido sintasa modular.

Para facilitar la biosíntesis combinatoria de policétidos producidos por PKSs modulares de tipo I se ha desarrollado un sistema de tres plásmidos para la expresión heteróloga de DEBS (Xue, Q., et al., *A multiplasmid approach to preparing large libraries of polyketides*. Proc. Natl. Acad. Sci. USA 1999. **96**(21): p. 11740-5). Los genes de PKS *eryA* que codifican las tres subunidades de DEBS se clonaron individualmente en tres vectores de *Streptomyces* compatibles que portaban marcadores de resistencia a antibiótico mutuamente seleccionables. Una cepa de *Streptomyces lividans* transformada con los tres plásmidos produjo 6-dEB en un nivel similar al de la cepa transformada con un único plásmido que contenía los tres genes. La utilidad de este sistema en biosíntesis combinatoria se demostró a través de la producción de una gran biblioteca de más de 60 macrolactonas policétídicas modificadas, usando versiones de cada plásmido construidas para contener mutaciones definidas. Se introdujeron combinaciones de estos conjuntos de vectores en *S. lividans*, dando como resultado cepas que producían una amplia gama de análogos de 6-dEB. Este método puede extenderse a cualquier PKS modular y tiene el potencial de producir miles de nuevos productos naturales, incluyendo unos derivados de la modificación adicional de los productos de la PKS por enzimas de adaptación. Además, la capacidad para expresar las PKSs modulares (tales como DEBS) a partir de estos tres plásmidos separados proporciona ventajas en la comercialización de la producción de policétidos por expresión heteróloga de PKSs modulares en levaduras y *E. coli* conforme a los métodos de la presente invención.

Como se describe en el Ejemplo 1, los genes traduccionalmente acoplados *mutA* y *mutB*, que codifican las subunidades α y β de la metilmalonil-CoA mutasa de *Propionibacterium shermanii*, se amplificaron por PCR y se insertaron en un vector de expresión en *E. coli* que contenía un promotor T7. El codón de iniciación GTG existente en la naturaleza para *mutB* se cambió por ATG para facilitar la expresión (Amaratunga M. et al., *A synthetic module for the methH gene permits facile mutagenesis of the cobalamin-binding region of Escherichia coli methionine synthase: initial characterization of seven mutant proteins*. Biochemistry, 1996, **35**(7): p. 2453-63). La expresión heteróloga de los genes de la mutasa en medio que contenía [3 H] β -alanina y el precursor de adenosilcobalamina (coenzima B₁₂), hidroxocobalamina, proporcionó metilmalonil-CoA mutasa activa. El análisis por HPLC de los extractos de *E. coli* BL21(DE3)/*panD* que hospedaba los genes de la mutasa indicó la producción de metilmalonil-CoA, que comprendía el 13% del combinado de CoAs intracelular (mostrado en la Figura 6). Este trabajo demuestra que se puede introducir una ruta de biosíntesis para un sustrato importante de una PKS en un hospedante heterólogo, y que se puede medir la concentración intracelular de acil-CoAs. Conforme a la presente invención, puede emplearse también el gen de la metilmalonil-CoA mutasa (*sbm*) de *E. coli*, que tiene una utilización de codones similar al de levaduras y que codifica un único polipéptido (Haller et al., *Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by Escherichia coli*. Biochemistry, 2000. **39**(16): p. 4622-9).

La Figura 6 muestra el análisis de acil-CoAs de *E. coli* que sobreexpresa metilmalonil-CoA mutasa. Se muestran los niveles de 3 H detectados en las fracciones recogidas del HPLC de los extractos libres de células de *E. coli* alimentadas con [3 H] β -alanina que hospedaban el vector de control pET (línea sólida), o bien pET que sobreexpresa la mutasa (línea de trazos).

Como se describe en el Ejemplo 1, se purificó la metilmalonil-CoA epimerasa de *P. shermanii* y se obtuvo la secuencia proteica del extremo N-terminal e interna. Se diseñaron cebadores para PCR degenerados basados en las secuencias de aminoácidos y se usaron para amplificar un producto de PCR de 180 pb del ADN genómico de *P. shermanii*. Los genes de la metilmalonil-CoA epimerasa de *B. subtilis* (Haller et al., *Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by Escherichia coli*. Biochemistry, 2000. **39**(16): p. 4622-9) y *S. coelicolor* pueden emplearse también en los métodos de la presente invención.

No se detecta propionil-CoA en células de *E. coli* SJ16 cultivadas en presencia de [³H] β-alanina con o sin la adición de propionato en el medio de crecimiento. Cuando las células de *E. coli* SJ16 se transformaron con un plásmido derivado de pACYC que contenía el gen de la propionil-CoA ligasa (*prpE*) de *Salmonella typhimurium* bajo el control del promotor *lac*, se observó una pequeña cantidad de propionil-CoA (~ 0,2% del combinado total de CoAs) en los extractos celulares. Cuando se incluyó propionato de sodio 5 mM en el medio de cultivo, se produjo aproximadamente 14 veces más propionil-CoA (~ 3% del combinado total de CoAs).

La Figura 7 muestra el análisis de acil-CoAs en *S. cerevisiae*. Se muestra el nivel de ³H detectado en las fracciones recogidas del análisis por HPLC de los extractos libres de células de *S. cerevisiae* alimentadas con [³H] β-alanina después de crecer durante 24 horas (línea sólida), 48 horas (línea de trazos) y 66 horas (línea de puntos). La cepa de levadura InvSc1 (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9), cultivada en medio YNB sintético sin pantotenato y β-alanina, se usó para el análisis de acil-CoAs. Los cultivos de levadura privados de β-alanina se alimentaron con [³H] β-alanina y los cultivos se dejaron crecer durante 24, 48 y 66 horas a 30°C. Las células se rompieron con perlas de vidrio en presencia de TCA frío al 10% y los acil-CoAs se separaron por HPLC y se cuantificaron por recuento de centelleo. Los combinados de CoAs de la levadura estaban marcados con [³H], pero el grado de dilución del isótopo queda sin estar claro. Se puede medir la actividad específica de CoA total en estas cepas para determinar el grado de dilución del isótopo.

Para los genes de PKS y estudios iniciales de los genes de rutas metabólicas, se pueden emplear los conjuntos análogos de vectores de clonación “bluescript” y vectores lanzadera de replicación basados en el círculo de 2 micras de levaduras usados en la producción de 6-MSA (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9). Con estos vectores, la expresión en levaduras está dirigida por el promotor de la alcohol deshidrogenasa 2 (ADH2), que es fuertemente reprimido por glucosa y es muy activo después del agotamiento de la glucosa que tiene lugar después de que el cultivo alcanza una alta densidad. Ambos conjuntos de vectores tienen un “casete de clonación común” que contiene, de 5’ a 3’, un sitio de clonación múltiple (L1), el promotor ADH2 (u otro), un sitio de restricción *NdeI*, un sitio de clonación múltiple (L2), un terminador ADH2 (u otro), y un sitio de clonación múltiple (L3). Debido al exceso de sitios de restricción en los vectores lanzadera de levaduras, los genes de interés se introducen primero en vectores de clonación “bluescript” intermedios por medio del sitio *NdeI*, para generar el codón de iniciación ATG, y un sitio de restricción en dirección 3’ en el sitio de clonación múltiple L2 que es común al “bluescript” y a los vectores lanzadera de levaduras (mostrados en la Figura 8). El casete con el promotor-gen se corta y separa a continuación como un fragmento L1-L2 y se transfiere al vector de expresión en levaduras que contiene el terminador transcripcional.

Las cepas hospedantes para sistemas modelo incluyen las cepas de levaduras comúnmente disponibles con deficiencias nutricionales (Ura, Trp, His, Leu) que pueden hospedar al menos tres vectores de replicación (véase más abajo). Si es necesario expresar más de tres genes de PKS simultáneamente, se pueden clonar casetes múltiples promotor-gen de PKS-terminador en el mismo vector o usar un cuarto vector de replicación con un marcador nutricional diferente (esto es, Leu) o un marcador antibiótico (esto es, G418). Se puede construir también un conjunto análogo de vectores de clonación “bluescript” y vectores de expresión/lanzadera de levaduras que contengan un promotor inducible por galactosa. El sistema activador del promotor de galactosa Gal4 está más fuertemente regulado que el promotor ADH2, y puede ser beneficioso o necesario para la expresión de proteínas que son tóxicas para la levadura (Mylin, L.M., *et al.*, *Regulated GAL4 expression cassette providing controllable and high-level output from high-copy galactose promoters in yeast*. Methods Enzymol., 1990. **185**: p. 297-308).

Los genes implicados en la producción de sustratos (p.ej., metilmalonil-CoA y/o propionil-CoA), y el gen *spf* pueden preferiblemente integrarse de manera estable en el cromosoma de la levadura en un número apropiado de copias para producir niveles adecuados de los acil-CoAs deseados y las modificaciones post-traduccionales de la PKS. Los genes pueden introducirse primero en el vector de clonación “bluescript” intermedio como se ha descrito. A continuación, el fragmento que contiene el casete promotor-gen-terminador puede transferirse como un fragmento L1-L3 a un vector de “integración en delta” en levaduras (Lee, F.W. y N.A. Da Silva, *Improved efficiency and stability of multiple cloned gene insertions at the delta sequences of Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol., 1997. **48**(3): p. 339-45; Lee, F.W. y N.A. Da Silva, *Sequential delta-integration for the regulated insertion of cloned genes in Saccharomyces cerevisiae*. Biotechnol. Prog., 1997. **13**(4): p. 368-73) que permita la integración cromosómica de los casetes en uno o más de las cerca de 425 secuencias delta dispersadas a lo largo del cromosoma de la levadura (véase la Figura 9). Estos vectores tienen sitios de clonación compatibles con los presentes en los sitios de clonación L1-L3 para permitir la transferencia directa de los casetes promotor-gen-terminador como fragmentos L1-L3. Contienen también el marcador de selección escindible Ura3 flanqueado por dos repeticiones *hisG* bacterianas (“URA Blaster”), posibilitando la inserción de genes múltiples idénticos o diferentes en el cromosoma de la levadura por integraciones repetitivas. Tras la selección sobre la base de la integración de los genes en medio sin uracilo, el fragmento del gen Ura3 se elimina seleccionando sobre la base de la pérdida del marcador por medio de recombinación de escindidos por selección positiva con ácido 5-fluoroorótico (FOA), que hace que el gen Ura3 resulte tóxico. Esto posibilita la introducción de rutas estables necesarias para los precursores de acil-CoAs y Spf en la levadura, mientras que se conserva el marcador Ura para permitir su uso subsiguiente en plásmidos que contengan otros genes.

La mutasa de un único gen, Sbm (“Sleeping beauty mutase”), de *E. coli* (Haller *et al.*, *Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by Escherichia coli*. Biochemistry, 2000. **39**(16): p. 4622-9), puede clonarse como sigue. Se usaron cebadores diseñados basándose en la secuencia de ADN para amplificar

por PCR el gen *sbm* a partir de ADN genómico de *E. coli* como un fragmento *NdeI*-L2. La estrategia general para clonar los genes en vectores de expresión en levaduras sigue la de Kealey *et al.* (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9) (véase la Figura 9). Se pueden clonar primero los genes como fragmentos *NdeI*-L2 en el vector de clonación intermedio “bluescript”. El casete promotor-gen-terminador puede cortarse y separarse a continuación como un fragmento L1-L3, transferirse al vector de integración en levaduras, digerido con enzimas de restricción de L1/L3, e introducirse en el cromosoma de la levadura como se describe anteriormente. Como alternativa a la *Sbm*, se puede usar la mutasa de dos genes de *P. shermanii*; los genes acoplados traduccionalmente se han amplificado cada uno por PCR como fragmentos *NdeI*-L2 y pueden integrarse en la levadura como se describe anteriormente.

Los genes que codifican *matABC* se han clonado en un vector “bluescript” (An, J.H. e Y.S. Kim, *A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in Rhizobium trifolii-cloning, sequencing, and expression of the enzymes in Escherichia coli*. Eur. J. Biochem., 1998. **225**(2): p. 395-402). Se pueden aislar los genes de *matB* (metilmalonyl-CoA ligasa) y *matC* (transportador de ácidos dicarboxílicos) por PCR, cada uno como un fragmento *NdeI*-L2, e integrarlos en el cromosoma de la levadura como se describe anteriormente y se muestra esquemáticamente más abajo. Las levaduras transformadas con *matBC* se tratarán con ácido metilmalónico, y los extractos celulares pueden analizarse para determinar la metilmalonyl-CoA.

Los genes *pccB* y *accA1* implicados en la ruta de carboxilación de la propionil-CoA en *S. coelicolor* pueden amplificarse por PCR a partir de ADN genómico. Como se muestra en la Figura 9, los genes pueden clonarse en el vector “bluescript” intermedio entre *NdeI* y L2, y a continuación transferirse al vector de integración por medio de L1/L3. Se pueden expresar los genes de *S. coelicolor* que se ha demostrado que son eficaces en *E. coli*; si la utilización de codones fuera subóptima, se pueden emplear los ortólogos de *B. subtilis* (discutido anteriormente).

La Figura 9 muestra un método general para clonar genes en vectores de expresión en levaduras.

En una realización, las células de levadura hospedantes recombinantes de la invención coexpresan la transferasa P-pant de *B. subtilis*, Sfp, con una PKS para convertir la apo-PKS en su holoforma. El gen *spf* está disponible en vectores de clonación Bluescript™ y lanzadera/de expresión de levaduras y es funcional en levaduras (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9), de forma que se pueden simplemente construir cepas estables que expresen este gen. Pueden introducirse de una a varias copias (según se determine óptimo) del gen *spf* en secuencias delta en el cromosoma de la levadura como se describe anteriormente. Se puede ensayar la actividad del gen *spf* integrado coexpresando 6-MSAS en un vector de replicación, midiendo la producción de 6-MSA dependiente de Sfp (Kealey, J.T., *et al.*, *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts*. Proc. Natl. Acad. Sci. USA, 1998. **95**(2): p. 505-9) y cuantificando la incorporación de [³H] β-alanina en el dominio ACP de la PKS (véase más abajo). Esto permite determinar el número óptimo de copias del gen *spf* necesarias para la producción máxima del policétido.

El gen de la PKS modular, DEBS1+TE, está disponible como un fragmento *NdeI*-*EcoRI*, que puede introducirse fácilmente en un vector lanzadera/de expresión de levaduras como se indica en la Figura 9. Las cepas de levadura que expresan DEBS1+TE se analizan para determinar la [³H] fosfopanteteinilación de la PKS, y la producción de lactona tricetónica por cromatografía de líquidos/espectrometría de masas.

El marcaje con ³H de los acil-CoAs intracelulares se lleva a cabo como sigue. Las células se tratan con [³H] β-alanina (disponible a 50 Ci/mmol) en medio definido sin pantotenato, posibilitando al precursor radiactivo del pantotenato entrar en el combinado de CoAs. A continuación las células se rompen, los ésteres de CoA se separan por HPLC, y la radiactividad se cuantifica por recuento de centelleo en líquido, como se describe anteriormente.

Las células hospedantes de *Saccharomyces cerevisiae* se cultivan, y los extractos se preparan como sigue. Se inocula medio YNB mínimo definido (1 mL) sin pantotenato pero que contiene β-alanina 1 μM con una única colonia de *S. cerevisiae* (InvSc1, o cepa con supresión en Fen2b) de una placa con YPD. El cultivo se deja crecer hasta la fase estacionaria y se usan 10 μL del cultivo estacionario para inocular el medio anterior sin β-alanina y pantotenato. El cultivo se incuba durante 4 horas y se usan 10 μL del cultivo “privado de alimento” para inocular medio (1 mL) que contiene 10 μCi de [³H] β-alanina (50 Ci/mmol; 0,2 μM de β-alanina). Después del crecimiento del cultivo durante los tiempos apropiados, las células de 1 mL de cultivo se recogen por centrifugación y se lavan con agua. Las células se resuspenden en 200 μL de ácido tricloroacético (TCA) al 10%, que contiene acil-CoAs estándar sin marcar como marcadores cromatográficos (malonyl-, metilmalonyl-, succinil-, acetil-, propionil-CoA, y CoA). Las células se rompen mediante vórtex con perlas de vidrio, y el sobrenadante se analiza por HPLC.

El análisis por HPLC se realiza usando una columna 5 μ ODS-3 INERTSIL HPLC de 150 x 4,6 mm adquirida de Matachem Technologies. El tampón A para el HPLC es fosfato de sodio monobásico 100 mM, acetato de sodio 75 mM, pH 4,6 y el tampón B es tampón A al 70%, metanol al 30%. La columna de HPLC se equilibra con tampón B al 10% a un flujo de 1 mL/min. Después de la inyección, se implementa un gradiente lineal hasta un 40% de tampón B durante 35 minutos, y a continuación un gradiente lineal hasta un 90% de tampón B durante 20 minutos. El gradiente proporciona la separación en la línea base de los acil-CoAs estándar. El eluyente se monitoriza a 260 nm y las fracciones se recogen y se cuantifican en un contador de centelleo.

ES 2 283 323 T3

La determinación de la actividad específica del combinado total de CoAs se lleva a cabo como sigue. Los cultivos de *S. cerevisiae* se marcan con 100 μCi de [^3H] β -alanina como se describe anteriormente. Las células de levadura se rompen y el extracto se trata con hidroxilamina 100 μM , pH 8,5, para convertir todos los acil-CoAs en CoA. La CoA marcada se aísla por HPLC como se describe anteriormente y se convierte en acetil-CoA con acetil-CoA sintasa de *E. coli* (Sigma), usando [^{14}C]-acetato como sustrato. La [^3H , ^{14}C]-acetil-CoA se separa por HPLC y los marcajes duales se cuantifican por recuento de centelleo. Los mmol de CoA se determinan por el ^{14}C , y la actividad específica de CoA se determina a partir de las dpm de ^3H por mmol de CoA. La dilución de isótopos, que refleja la producción endógena de β -alanina, se calcula mediante la actividad específica de [^3H] CoA/actividad específica de [^3H] β -alanina usada en el ensayo.

El análisis de los niveles de expresión de PKS se lleva a cabo como sigue. Cada dominio ACP de cada módulo de una PKS activa se modifica post-traduccionalmente con fosfopanteteína derivada de CoA. Usando células de levadura tratadas con [^3H] β -alanina (descritas anteriormente), se puede marcar la PKS con tritio de alta actividad específica. La proteína se separará por SDS-PAGE, se eluirá y se determinará la radiactividad por recuento de centelleo en líquido.

Ejemplo 3

Conversión de Eritronolidas en Eritromicinas

De disuelve una muestra de un policétido (~50 a 100 g) en 0,6 mL de etanol y se diluye hasta 3 mL con agua estéril. Esta solución se usa para cubrir un cultivo de tres días de *Saccharomyces cerevisiae* WHM34 (un mutante en *eryA*) crecido sobre una placa de R2YE agar de 100 mm a 30°C. Después de secarla, la placa se incuba a 30°C durante 4 días. El agar se corta y a continuación se somete a extracción tres veces con porciones de 100 mL de trietilamina al 1% en acetato de etilo. Los extractos se combinan y se evaporan. El producto bruto se purifica por HPLC preparativa (C-18 fase inversa, gradiente de agua-acetonitrilo que contiene ácido acético al 1%). Las fracciones se analizan por espectrometría de masas, y las que contienen el compuesto puro se combinan, se neutralizan con trietilamina y se evaporan hasta conseguir un jarabe. El jarabe se disuelve en agua y se somete a extracción tres veces con volúmenes iguales de acetato de etilo. Los extractos orgánicos se combinan, se lavan una vez con NaHCO_3 acuoso saturado, se secan sobre Na_2SO_4 , se filtran, y se evaporan para proporcionar ~0,15 mg de producto. El producto es un compuesto glicosilado e hidroxilado correspondiente a eritromicina A, B, C y D, pero que difiere de ellas como el compuesto proporcionado difería de 6-dEB.

Ejemplo 4

Medición de la Actividad Antibacteriana

La actividad antibacteriana se determina usando ensayos de difusión en disco con *Bacillus cereus* como organismo de ensayo, o bien por medición de las concentraciones mínimas inhibitoras (MIC, del inglés “minimum inhibitory concentrations”) en cultivo líquido frente a cepas sensibles y resistentes de *Staphylococcus pneumoniae*.

Ejemplo 5

Evaluación de la Actividad Antiparasitaria

Los compuestos pueden someterse a escrutinio inicialmente *in vitro* usando cultivos de las cepas de *P. falciparum* FCR-3 y K1, y a continuación *in vivo* usando ratones infectados con *P. berghei*. La toxicidad sobre las células de mamíferos puede determinarse en células FM3A o KB. Los compuestos pueden someterse a escrutinio también sobre la base de la actividad frente a *P. berhei*. Los compuestos se ensayan también en estudios animales y pruebas clínicas para ensayar la actividad antiparasitaria ampliamente (antimalaria, tripanosomiasis y Leishmaniasis).

ES 2 283 323 T3

REIVINDICACIONES

5 1. Una célula hospedante de *E. coli* recombinante para la producción eficaz de 2S-metilmalonil-CoA que comprende uno o más vectores de expresión, comprendiendo dichos vectores de expresión en combinación:

(a) los genes *mutA* y *mutB* de metilmalonil-CoA mutasa; y

(b) un gen de metilmalonil-CoA epimerasa.

10 2. La célula hospedante de la reivindicación 1, en la que los genes de metilmalonil-CoA mutasa son de *Propionibacterium shermanii* o bien de *Streptomyces cinnamomensis*.

15 3. La célula hospedante de la reivindicación 1 ó 2, en la que el gen de la epimerasa deriva de *Propionibacterium shermanii* que tiene la secuencia de nucleótidos

```
20      atg agt aat gag gat ctt ttc atc tgt atc gat cac gtg gca tat gcg      48
      tgc ccc gac gcc gac gag gct tcc aag tac tac cag gag acc ttc ggc      96
      tgg cat gag ctc cac cgc gag gag aac ccg gag cag gga gtc gtc gag      144
      atc atg atg gcc ccg gct gcg aag ctg acc gag cac atg acc cag gtt      192
25      cag gtc atg gcc ccg ctc aac gac gag tgc acc gtt gcc aag tgg ctt      240
      gcc aag cac aat ggt cgc gcc gga ctg cac cac atg gca tgg cgt gtc      288
30      gat gac atc gac gcc gtc agc gcc acc ctg cgc gag cgc ggc gtg cag      336
      ctg ctg tat gac gag ccc aag ctc ggc acc ggc ggc aac cgc atc aac      384
35      ttc atg cat ccc aag tgc ggc aag ggc gtg ctc atc gag ctc acc cag      432
      tac ccg aag aac tga      447
```

40 4. La célula hospedante de cualquiera de las reivindicaciones 1 a 3, que comprende además al menos un gen de policétido sintasa (PKS) modular.

45 5. La célula hospedante de la reivindicación 4, en la que el policétido es producido por una policétido sintasa modular (PKS) producida por el gen de PKS en dicha célula hospedante, y

dicho gen de PKS está contenido en un vector que se replica de manera extra-cromosómica o se integra en el ADN cromosómico de tal célula hospedante.

50 6. La célula hospedante de cualquiera de las reivindicaciones 1 a 5, en medio que contiene hidroxocobalamina.

7. La célula hospedante de cualquiera de las reivindicaciones 1 a 6, que comprende dos vectores de expresión, uno de los cuales se integra en el ADN cromosómico de dicha célula.

55 8. La célula hospedante de cualquiera de las reivindicaciones 1 a 7, que comprende dos vectores de expresión, uno de los cuales es un plásmido.

9. La célula hospedante de cualquiera de las reivindicaciones 1 a 8, en la que dichos genes de metilmalonil-CoA mutasa son los genes *mutA* y *mutB* de la metilmalonil-CoA mutasa de *Propionibacterium shermanii*.

60 10. La célula hospedante de cualquiera de las reivindicaciones 1 a 8, en la que dichos genes de metilmalonil-CoA mutasa son los genes *mutA* y *mutB* de la metilmalonil-CoA mutasa de *Streptomyces cinnamomensis*.

65 11. La célula hospedante de cualquiera de las reivindicaciones 1 a 10, en la que uno o más de dichos genes está bajo el control de un promotor de un gen de *E. coli*.

12. La célula hospedante de cualquiera de las reivindicaciones 4 a 11, en la que dicha PKS es una 6-desoxieritrolida B sintasa.

ES 2 283 323 T3

13. Una célula hospedante de *E. coli* recombinante que comprende un sistema de expresión de propionil-CoA carboxilasa (pcc) que comprende los genes pccB y accA1 de *S. coelicolor* en el que dicho sistema de expresión de pcc produce una enzima capaz de sintetizar 2S-metilmalonil-CoA y que comprende además al menos un sistema de expresión para una policétido sintasa (PKS) modular.

5

14. La célula hospedante de la reivindicación 13, que comprende además un sistema de expresión de fosfopanteteinil-transferasa.

10

15. La célula hospedante de la reivindicación 14, en la que la fosfopanteteinil-transferasa es el gen *sfp* de *Bacillus subtilis*.

16. La célula hospedante de cualquiera de las reivindicaciones 13 a 15, en la que la célula hospedante comprende además un sistema de expresión para biotina transferasa.

15

17. La célula hospedante de la reivindicación 16, en la que el sistema de expresión para biotina transferasa es el gen *birA* de *E. coli*.

20

18. La célula hospedante de cualquiera de las reivindicaciones 14 a 17, en la que la PKS es desoxieritronolida B sintasa (DEBS).

19. La célula hospedante de cualquiera de las reivindicaciones 14 a 17, en la que el policétido es 6-desoxieritronolida B (6-dEB).

25

20. Un método para producir un policétido, que comprende cultivar la célula hospedante de cualquiera de las reivindicaciones 4 a 19 en condiciones tales que dicho gen de PKS se expresa para producir una PKS funcional, dicho 2S-metilmalonil-CoA se produce, y dicha PKS funcional sintetiza un policétido que incorpora dicho 2S-metilmalonil-CoA.

30

35

40

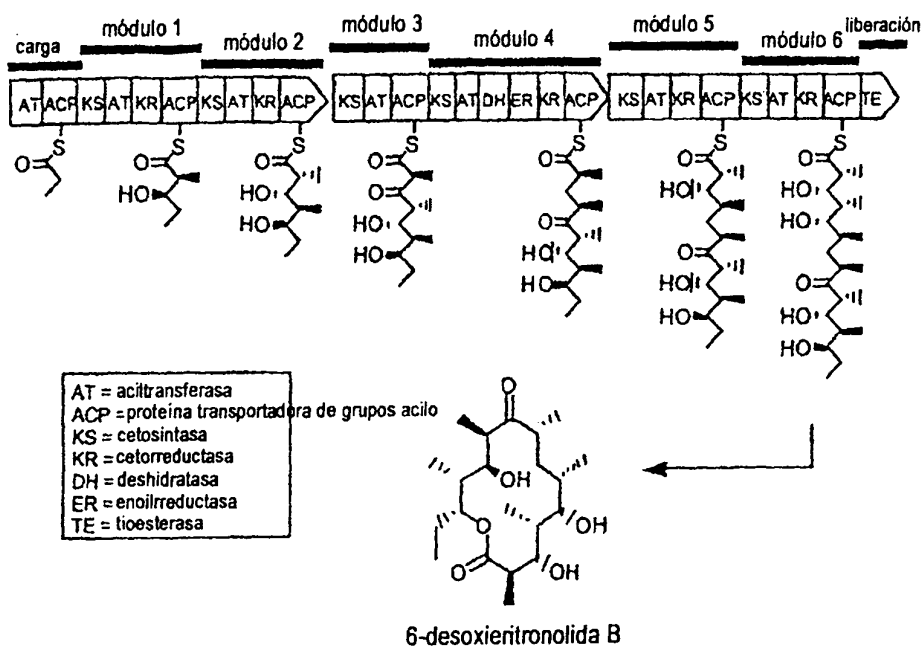
45

50

55

60

65



Biosíntesis de 6-Desoxieritronolida B (6-dEB), la Aglicona de Eritromicina, por una PKS Modular

Figura 1

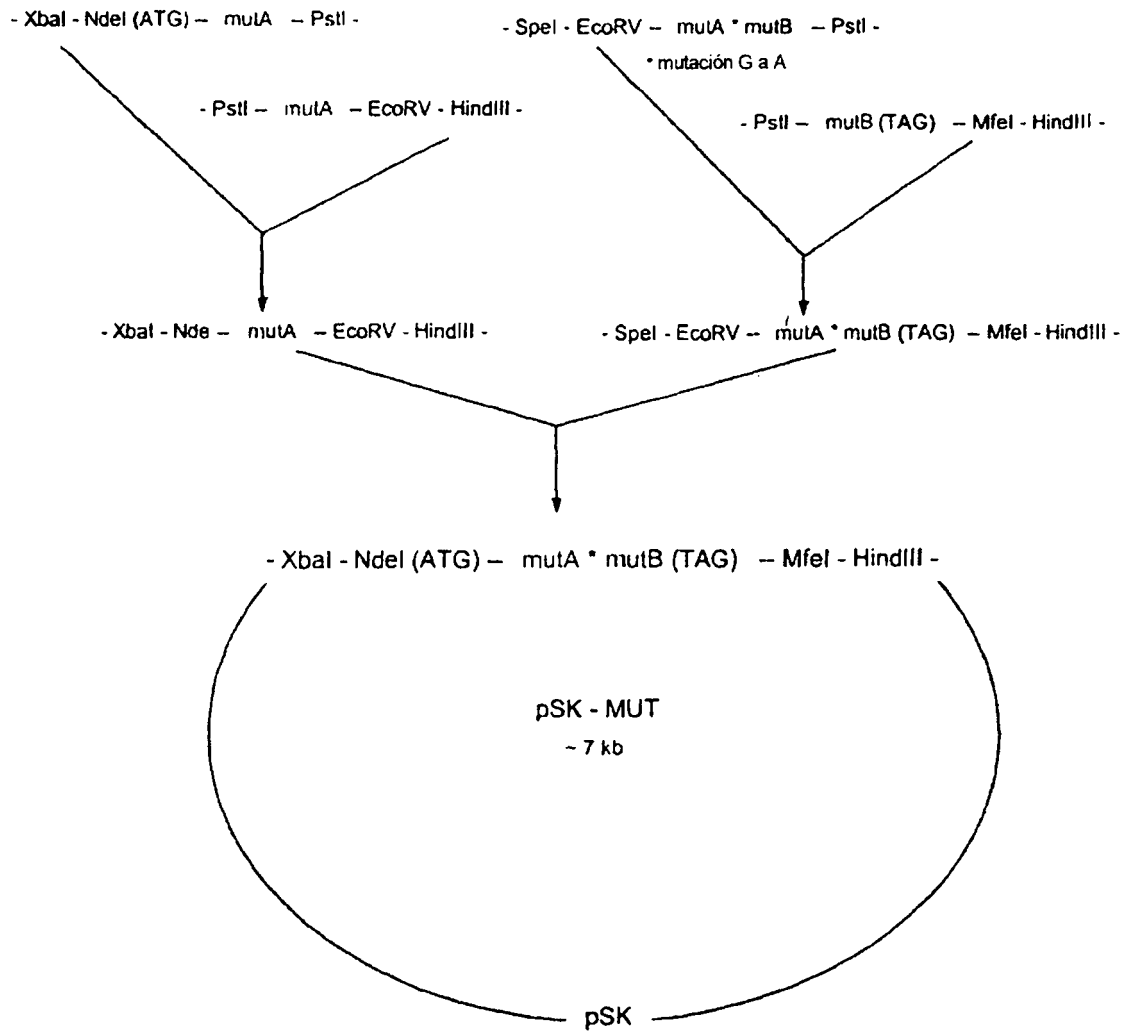


Figura 2

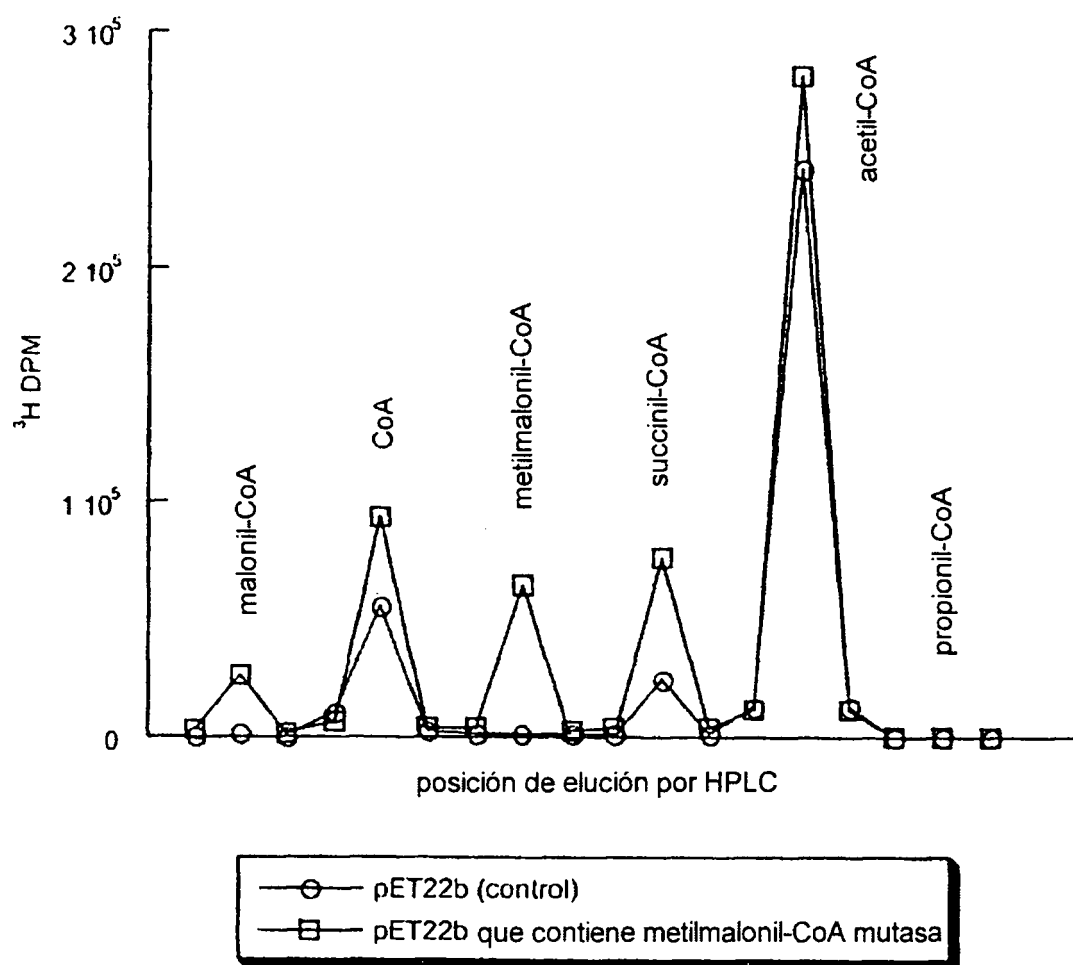


Figura 3

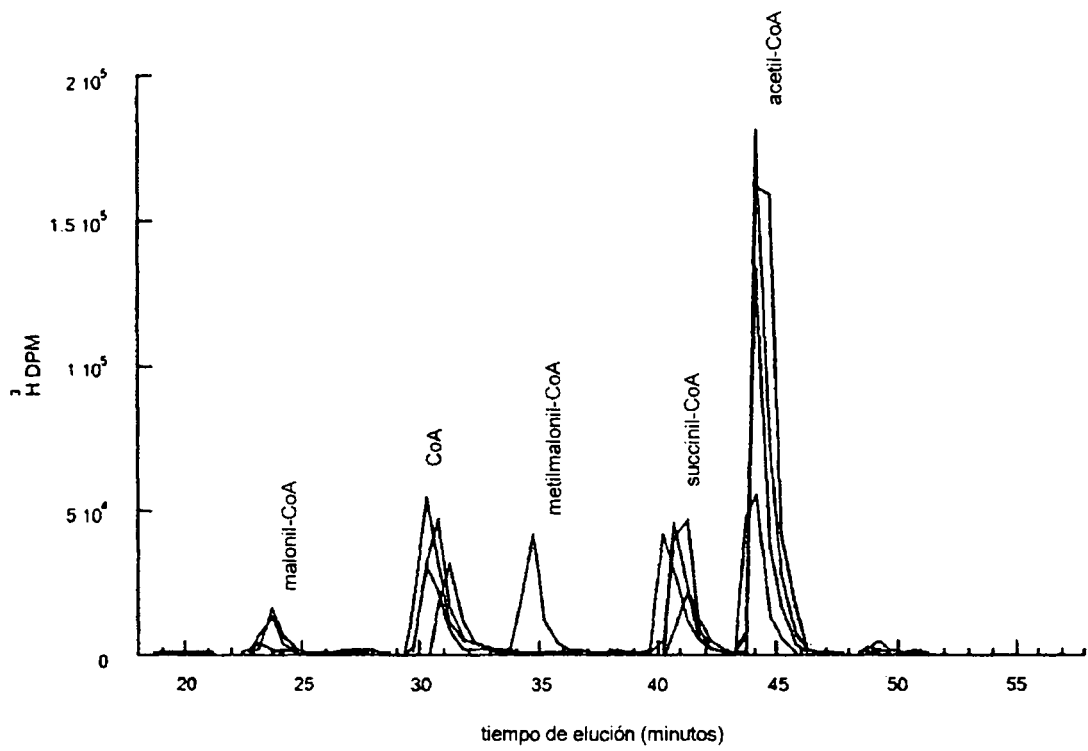
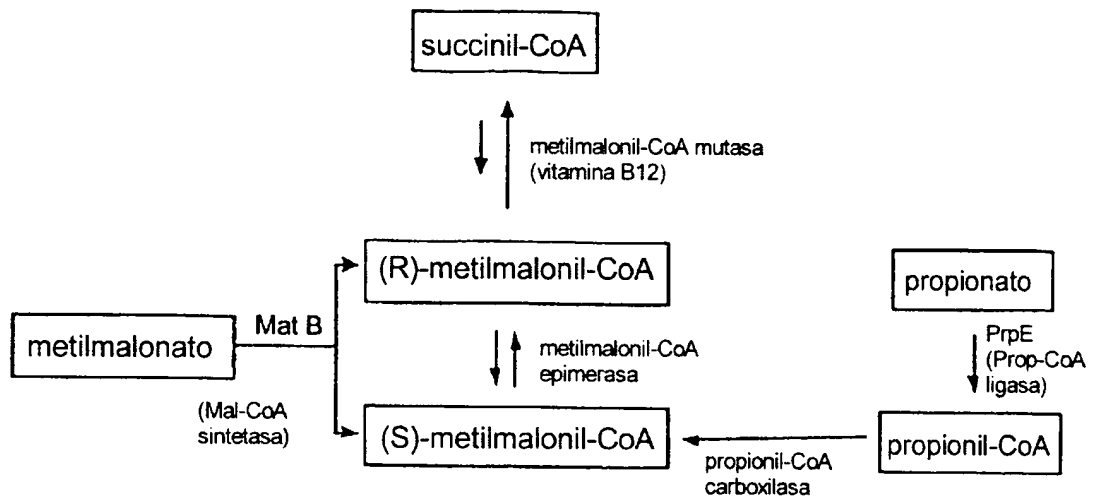


Figura 4



Vías biosintéticas para (S)-metilmalonil-CoA

Figura 5

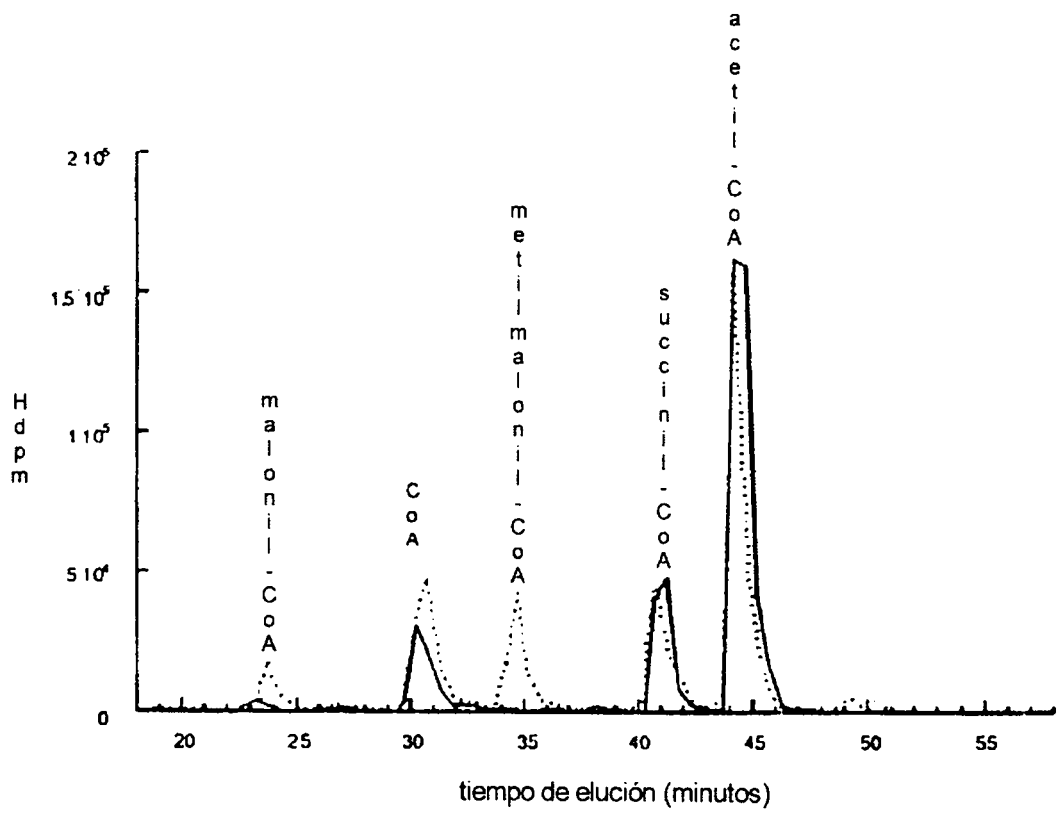


Figura 6

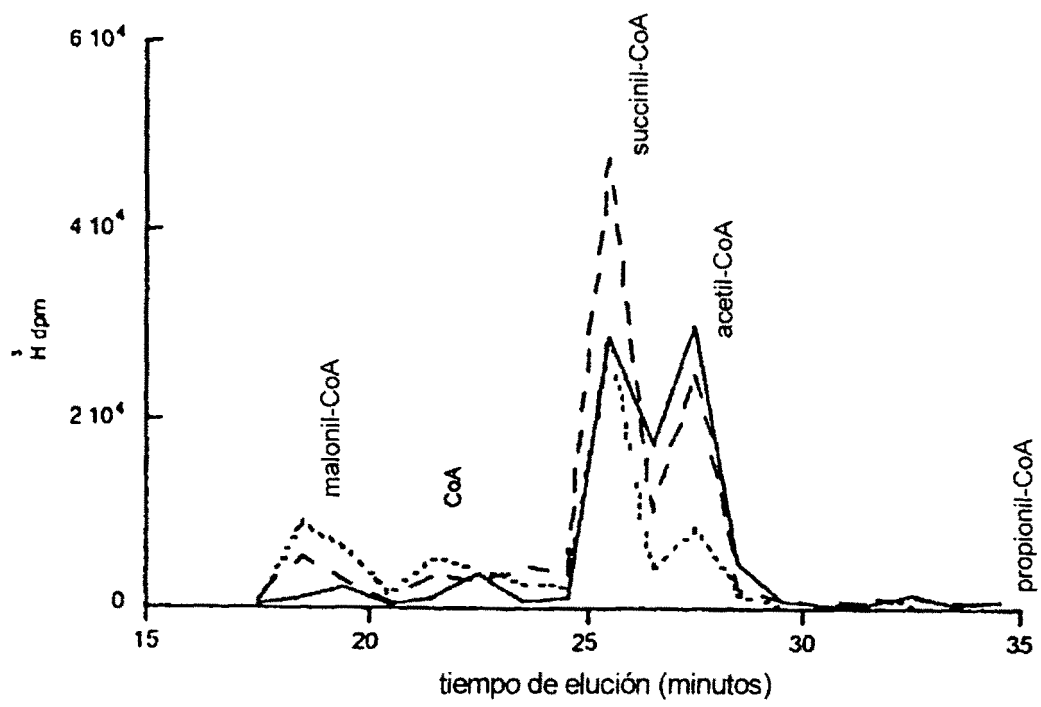


Figura 7

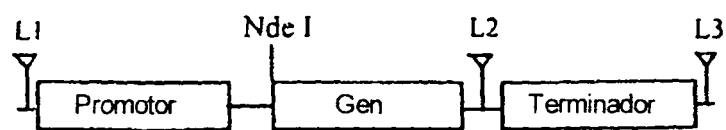
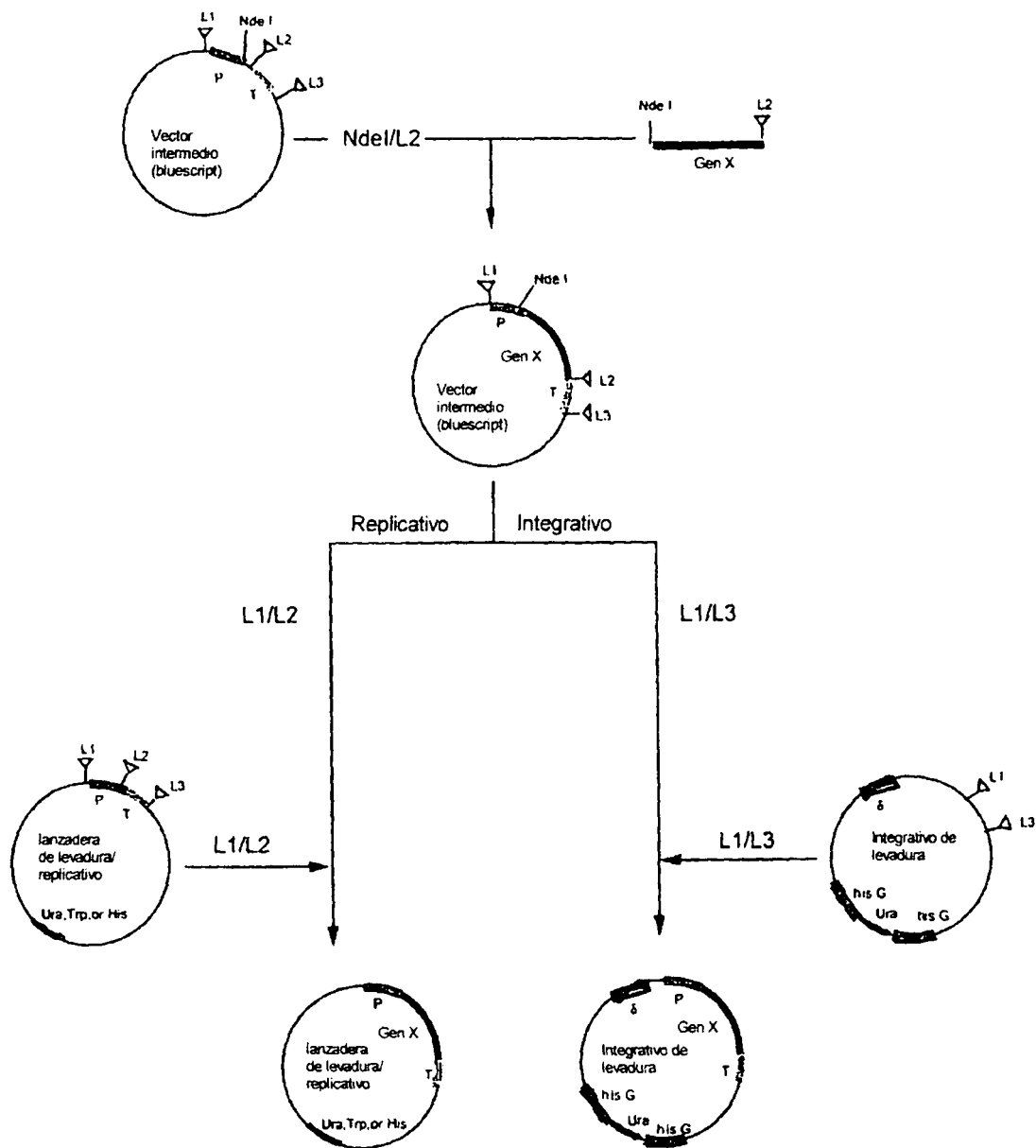


Figura 8



P = promotor de levadura
 T = terminador de levadura
 L1 = BamHI, Not I
 L2 = XbaI, EcoRI, Sall, RsrII, AvrII, NsiI, SpeI
 L3 = Xho, KpnI

Figura 9

ES 2 283 323 T3

LISTA DE SECUENCIAS

<110> KOSAN Biosciences, Inc.

5 <120> PRODUCCIÓN HETERÓLOGA DE POLICÉPTIDOS

<130> 30062-20049.40

10 <140> PCT/US 00/29775

<141> 2000-10-27

15 <150> 60/161, 703

<151> 1999-10-27

<160> 2

20 <170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 447

25 <212> ADN

<213> *Propionibacterium shermanii*

30 <220>

<221> CDS

<222> (1)...(447)

35 <400> 1

	atg agt aat gag gat ctt ttc atc tgt atc gat cac gtg gca tat gcg	48
	Met Ser Asn Glu Asp Leu Phe Ile Cys Ile Asp His Val Ala Tyr Ala	
40	1 5 10 15	
	tgc ccc gac gcc gac gag gct tcc aag tac tac cag gag acc ttc ggc	96
	Cys Pro Asp Ala Asp Glu Ala Ser Lys Tyr Tyr Gln Glu Thr Phe Gly	
	20 25 30	
45	tgg cat gag ctc cac cgc gag gag aac ccg gag cag gga gtc gtc gag	144
	Trp His Glu Leu His Arg Glu Glu Asn Pro Glu Gln Gly Val Val Glu	
	35 40 45	
50	atc atg atg gcc ccg gct gcg aag ctg acc gag cac atg acc cag gtt	192
	Ile Met Met Ala Pro Ala Ala Lys Leu Thr Glu His Met Thr Gln Val	
	50 55 60	
55	cag gtc atg gcc ccg ctc aac gac gag tcg acc gtt gcc aag tgg ctt	240
	Gln Val Met Ala Pro Leu Asn Asp Glu Ser Thr Val Ala Lys Trp Leu	
	65 70 75 80	
60	gcc aag cac aat ggt cgc gcc gga ctg cac cac atg gca tgg cgt gtc	288
	Ala Lys His Asn Gly Arg Ala Gly Leu His His Met Ala Trp Arg Val	
	85 90 95	
65	gat gac atc gac gcc gtc agc gcc acc ctg cgc gag cgc ggc gtg cag	336

