

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 724 373**

51 Int. Cl.:

C12N 9/02 (2006.01)

C12P 33/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.01.2015 PCT/EP2015/050866**

87 Fecha y número de publicación internacional: **23.07.2015 WO15107185**

96 Fecha de presentación y número de la solicitud europea: **19.01.2015 E 15702160 (1)**

97 Fecha y número de publicación de la concesión europea: **02.01.2019 EP 3097113**

54 Título: **Polipéptido citocromo P450 novedoso con actividad enzimática aumentada**

30 Prioridad:

20.01.2014 EP 14305071

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

10.09.2019

73 Titular/es:

**SANOFI (100.0%)
54, rue La Boétie
75008 Paris, FR**

72 Inventor/es:

**BERTIN, MARINE y
DUMAS, BRUNO**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 724 373 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

Polipéptido citocromo P450 novedoso con actividad enzimática aumentada

5 La presente invención se refiere a monooxigenasas del citocromo P450 con una actividad enzimática aumentada.

Los P450 mitocondriales de mamíferos constituyen una familia pequeña de citocromos que llevan a cabo reacciones específicas dentro de las mitocondrias y desempeñan un papel importante en el metabolismo de una diversidad de compuestos hidrófobos. Por ejemplo, P450c11beta y P450c11AS (codificados por los genes CYP11B1 y CYP11B2) llevan a cabo la etapa final de la biosíntesis de glucocorticoides y mineralocorticoides, respectivamente. El P450c11beta es un esteroide 11beta-hidroxilasa clásica que convierte 11-desoxicortisol en hidrocortisona, mientras que el P450c11AS cataliza la transformación de desoxicorticosterona en aldosterona en tres etapas consecutivas. El CYP27A1 es otro ejemplo de un P450 mitocondrial interno que es un esteroide 27-hidroxilasa y vitamina D25 hidroxilasa. Finalmente, el P450 mitocondrial más emblemático es el P450scc (enzima de escisión de cadena lateral (*Side Chain Cleaving*)), codificado por el gen CYP11A1, que escinde la cadena lateral de colesterol, transformando así el colesterol en pregnenolona por medio de dos hidroxilaciones consecutivas y una escisión final. El P450scc lleva a cabo la primera etapa clave de la biosíntesis de esteroides transformando un esteroide en un esteroide.

La capacidad del P450 para catalizar la oxidación regioespecífica, quimioespecífica y estereoquímica de una amplia serie de sustratos refleja sus funciones biológicas y los convierte en candidatos importantes para aplicaciones biotecnológicas. En particular, las hormonas esteroideas se utilizan ampliamente como antiinflamatorios, anticonceptivos y antiproliferativos. En los mamíferos, la síntesis de estos esteroides comienza con la reacción de escisión de la cadena lateral del colesterol para dar pregnenolona. La pregnenolona sirve como base para la producción de otras hormonas esteroideas tales como hidrocortisona y están asociados grandes intereses con su conversión industrial a gran escala a partir de sustratos económicos tales como el colesterol y sus análogos derivados de plantas. No obstante, la reacción de escisión de la cadena lateral del colesterol para dar pregnenolona es una etapa limitante en el proceso general de esteroides. En los mamíferos, está catalizada por la enzima CYP11A1 unida a membrana.

Sin embargo, por múltiples motivos, el uso biotecnológico e industrial de los P450 es difícil de implementar y no es satisfactorio. El P450 mitocondrial utiliza una cadena específica de transportadores de electrones constituida por dos proteínas, a saber, la ferredoxina reductasa (FdxR) y la ferredoxina (Fdx1), también denominadas adrenodoxina reductasa (AdR) y adrenodoxina (Adx). En la situación natural *in vivo*, la FdxR gana un electrón procedente del NADPH y la Fdx1 transporta un electrón desde la FdxR hasta el P450 mitocondrial. Se ha desarrollado un sistema *in vitro* en el que la FdxR se encuentra en una concentración catalítica (0,5 µM) y la Fdx1 en una concentración de saturación (10 µM). Para que este sistema *in vitro* sea eficaz, los electrones deben fluir de forma apropiada al P450.

Para tratar de superar estas dificultades, algunos autores han fusionado los tres péptidos P450scc (o P450c11 beta), Fdx1 y FdxR utilizando secuencias bisagra o enlazadoras variables. Pero aunque esta triple fusión da como resultado una proteína funcional, su eficacia es baja en comparación con el polipéptido "auténtico" y no se puede utilizar a escala industrial. Además, el entorno mitocondrial es difícil de imitar en un sistema recombinante microbiano. Por ejemplo, el polipéptido P450scc puede direccionarse adecuadamente a las mitocondrias de levadura. Sin embargo, el polipéptido direccionado no puede convertir esteroide en pregnenolona debido a la ausencia de sustrato en las mitocondrias y/o al plegamiento o al direccionamiento incorrectos de P450scc.

Se utilizan diferentes sistemas recombinantes para producir pregnenolona biosintética a partir de esteroide vegetal. Por ejemplo, se ha desarrollado un sistema de biosíntesis en la levadura *Saccharomyces cerevisiae*. En este sistema, el P450scc se direcciona al exterior de las mitocondrias en la membrana plasmática junto con FdxR y Fdx1, y la ruta del esteroide se dirige a producir esteroide en la membrana. Además, el P450c11beta puede direccionarse a las mitocondrias junto con Fdx y FdxR1, en las que puede convertir 11-desoxicortisol en cortisol. Se ha desarrollado un sistema de bioconversión en *Bacillus megaterium*. En este sistema, el colesterol se puede convertir en pregnenolona por medio de formas maduras de P450scc, Fdx1 y FdxR1. La biosíntesis continúa siendo la tecnología más atractiva a escala industrial, dado que el sustrato se produce como una molécula soluble a partir de una fuente de carbono sencilla por parte del huésped, evitando por lo tanto la carga de utilizar detergentes.

Recientemente, se han realizado intentos para producir nuevos polipéptidos P450scc con propiedades mejoradas. Por ejemplo, se ha creado un mutante de P450scc recombinante que porta una mutación K193E. Este mutante es más soluble y, por lo tanto, menos propenso a la agregación que el polipéptido recombinante de tipo silvestre. Por lo tanto, se obtiene un nivel de expresión más elevado con este mutante sin cambiar sus características enzimáticas. No obstante, todavía existe la necesidad de obtener polipéptidos P450scc mejorados con una mayor actividad enzimática que permitan una producción optimizada de hormonas esteroideas que sean particularmente adecuados para procesos industriales.

Los inventores han desarrollado un nuevo polipéptido P450scc que porta mutaciones específicas. Este mutante presenta inesperadamente una actividad enzimática mejorada en términos de conversión de sustrato en hormonas esteroideas.

La presente invención se refiere así a una nueva monooxigenasa del citocromo P450 aislada que comprende o consiste en una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1, en la que dicha secuencia comprende una treonina en la posición correspondiente a la posición 225 y una mutación de ácido aspártico en la posición correspondiente a la posición 289. En algunas formas de realización, la nueva enzima P450 es al menos el 85%, 90%, 95%, 98% o 99% idéntica a la SEQ ID NO: 1.

La invención también se refiere a un ácido nucleico aislado que comprende o consiste en una secuencia de nucleótidos que codifica dicha enzima, a un vector que comprende dicho ácido nucleico que está asociado operativamente con secuencias de control de la expresión y a una célula huésped que contiene dicho ácido nucleico o dicho vector, siendo dicha célula huésped, por ejemplo, un microorganismo.

En una forma de realización, la invención proporciona un microorganismo manipulado genéticamente capaz de convertir un sustrato en un precursor de hormonas esteroideas. El sustrato puede ser, por ejemplo, un monoalcohol policíclico insaturado que tiene una cadena lateral alifática, tal como colesterol, un análogo de colesterol o un derivado de colesterol. El microorganismo manipulado genéticamente comprende, por ejemplo, un ácido nucleico que codifica una nueva enzima P450 presentada en la invención, y opcionalmente una secuencia de ácidos nucleicos que codifica una adrenodoxina (Adx) y/o una secuencia de ácidos nucleicos que codifica una adrenodoxina reductasa (AdR).

En otro aspecto, la invención proporciona un procedimiento *in vitro* para preparar una enzima P450 presentada en la invención, comprendiendo dicho procedimiento las etapas siguientes:

- a) cultivar una célula huésped en condiciones adecuadas para obtener la expresión de la enzima P450 y
- b) recuperar la enzima expresada.

En otra forma de realización, una enzima P450 presentada en la invención se utiliza para producir un precursor de hormonas esteroideas.

La invención se refiere además a un procedimiento para producir un precursor de hormonas esteroideas, que comprende las etapas siguientes:

- a) proporcionar un microorganismo que expresa una enzima P450 presentada en la invención,
- b) cultivar dicho microorganismo en condiciones que permitan la expresión de la enzima P450,
- c) poner en contacto el cultivo de microorganismos obtenido en la etapa b) con un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo de colesterol y un derivado de colesterol, en condiciones que permiten la producción, por el microorganismo, de un precursor de hormonas esteroideas a partir de dicho sustrato y
- d) recuperar el precursor de hormonas esteroideas producido.

La invención también se refiere a un procedimiento, tal como un procedimiento *in vitro*, para producir un precursor de hormonas esteroideas, que comprende las etapas siguientes:

- a) poner en contacto una enzima P450 presentada en la invención con un polipéptido adrenodoxina (Adx) aislado, un polipéptido adrenodoxina reductasa (AdR) aislado y un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, análogos y derivados del colesterol en condiciones que permitan la transformación de dicho sustrato en un precursor de hormonas esteroideas y
- b) recuperar el precursor de hormonas esteroideas obtenido.

Las hormonas esteroideas se utilizan ampliamente como antiinflamatorios, anticonceptivos y antiproliferativos. La invención es de utilidad para producir un precursor de hormonas esteroideas y ayudar a la síntesis de hormonas esteroideas para su uso en seres humanos. El precursor de hormonas esteroideas, tales como la pregnelona, sirve como base para la producción de otras hormonas esteroideas, tales como la hidrocortisona, y están asociados grandes intereses con su conversión industrial a gran escala.

Descripcion de la invencion

Enzimas

Por una enzima "aislada" o un polipéptido "aislado" se quiere decir que la enzima o el polipéptido ya no está en su entorno natural dentro del organismo en el que se expresa originariamente. Cuando se hace referencia a una enzima o un polipéptido, "purificado" significa que la molécula indicada está presente en ausencia sustancial de otras macromoléculas biológicas del mismo tipo. El término "purificado", tal como se usa en el presente documento, significa que está presente al menos el 75% en peso, al menos el 85% en peso, al menos el 95% en peso o al menos el 98% en peso de macromoléculas biológicas del mismo tipo.

La expresión "enzima citocromo P450" o "enzima P450" se refiere a una monooxigenasa que es capaz de catalizar determinadas reacciones tales como las que se examinan por Van Bogaert *et al.*, 2011, FEBS J. 278 (2): 206-221 o por Urlacher y Girhard, 2011, Trends in Biotechnology 30 (1): 26-36, o en el sitio de Internet dmelson.uthsc.edu/CytochromeP450.html. Como ejemplo de dicha enzima P450 de tipo silvestre, la secuencia SEQ ID NO: 4 representa la forma madura de CYP11A1 de *Bos taurus*. La secuencia completa de CYP11A1 de *Bos taurus* que incluye un péptido de tránsito se referencia como P00189 en la base de datos UniProtKB/Swiss-Prot (última actualización de la secuencia: 21 de julio de 1986).

Las enzimas citocromo P450 presentadas en la invención pueden estar unidas a membrana (insolubles) o ser citoplásmicas (solubles) en sus respectivos huéspedes originales. Por ejemplo, el P450scc (SEQ ID NO: 4) cataliza la reacción de escisión de la cadena lateral de monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, análogos de colesterol y derivados de los mismos para dar pregnenolona u otros precursores y derivados de hormonas esteroideas, como ejemplo no limitante.

Los inventores han desarrollado una nueva enzima P450 de secuencia SEQ ID NO: 1 con actividad enzimática mejorada. En comparación con la SEQ ID NO: 4 (enzima scc de tipo silvestre con R en la posición 225 y N en la posición 289), la secuencia de aminoácidos SEQ ID NO: 1 comprende dos mutaciones: una mutación de arginina a treonina en la posición correspondiente a la posición 225 de la secuencia SEQ ID NO: 4 y una mutación de asparagina a ácido aspártico en la posición correspondiente a la posición 289 de la secuencia SEQ ID NO: 4.

En una forma de realización específica, la nueva enzima P450 presentada en la invención muestra una actividad de monooxigenasa que cataliza la transformación de colesterol en pregnenolona. En una forma de realización particular, la nueva enzima P450 muestra al menos el 80% o más, en particular al menos el 85%, 90%, 95%, 100%, 125%, 150%, 175%, 200%, 225%, 250%, 275%, más particularmente al menos el 300% o más de la actividad de monooxigenasa de la enzima P450 de secuencia SEQ ID NO: 4.

La enzima P450 aislada presentada en la invención comprende o consiste en una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1, en la que dicha secuencia comprende una treonina en la posición correspondiente a la posición 225 y un ácido aspártico en la posición correspondiente a la posición 289. En algunas formas de realización, la nueva enzima P450 es al menos el 85%, 90%, 95%, 98% o 99% idéntica a la SEQ ID NO: 1.

Se pueden introducir variaciones en la secuencia de aminoácidos mediante la sustitución, la eliminación o la inserción de uno o más codones en la secuencia de ácidos nucleicos que codifica la enzima, lo que da como resultado un cambio en la secuencia de aminoácidos de la enzima. Las sustituciones de aminoácidos pueden ser conservativas o no conservativas. En algunas formas de realización, las sustituciones son sustituciones conservativas, en las que un aminoácido se sustituye por otro aminoácido con propiedades químicas y/o estructurales similares.

Una "sustitución conservativa de aminoácidos" es una en la que un residuo de aminoácido se sustituye por otro residuo de aminoácido que tiene un grupo R de cadena lateral con propiedades químicas similares (por ejemplo, carga o hidrofobicidad). En general, una sustitución de aminoácidos conservativa no cambiará sustancialmente las propiedades funcionales de una proteína. Ejemplos de grupos de aminoácidos que tienen cadenas laterales con propiedades químicas similares incluyen 1) cadenas laterales alifáticas: glicina, alanina, valina, leucina e isoleucina; 2) cadenas laterales de hidroxilo alifáticas: serina y treonina; 3) cadenas laterales que contienen amida: asparagina y glutamina; 4) cadenas laterales aromáticas: fenilalanina, tirosina y triptófano; 5) cadenas laterales básicas: lisina, arginina e histidina; 6) cadenas laterales ácidas: ácido aspártico y ácido glutámico; y 7) cadenas laterales que contienen azufre: cisteína y metionina. Los grupos de sustitución conservativa de aminoácidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina, glutamato-aspartato y asparagina-glutamina.

Por ejemplo, una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1 puede ser un polipéptido que tiene al menos una sustitución en un residuo de aminoácido particular. En algunas formas de realización, una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1 tiene al menos aproximadamente el 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% de identidad de secuencia de aminoácidos con la secuencia SEQ ID NO: 1. La identidad de secuencia de aminoácidos se define como el porcentaje de residuos de aminoácidos de la secuencia al menos el 80% idéntica a la SEQ ID NO: 1 que son idénticos a los residuos de aminoácidos de la secuencia de referencia SEQ ID NO: 1, después de alinear las secuencias e introducir huecos, si es necesario, para lograr el porcentaje máximo de identidad de secuencia, y sin considerar las sustituciones conservativas como parte de la identidad de secuencia. La identidad de secuencia

se puede determinar a lo largo de toda la longitud de la secuencia al menos el 80% idéntica, la longitud total de la secuencia de referencia, o ambas. El porcentaje de identidad para secuencias de aminoácidos se puede calcular realizando un alineamiento general por pares basado en el algoritmo de alineamiento de Needleman-Wunsch para encontrar el alineamiento óptimo (incluidos los huecos) de dos secuencias a lo largo de toda su longitud, por ejemplo, utilizando Needle, y utilizando la matriz BLOSUM62 con una penalización de apertura de hueco de 10 y una penalización de extensión de hueco de 0,5.

En la enzima P450 que comprende o consiste en una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1, las modificaciones de aminoácidos en comparación con la SEQ ID NO: 1 se ubican normalmente en unas posiciones que no socavan significativamente la actividad biológica de la enzima. De hecho, la secuencia de aminoácidos del citocromo P450 que se presenta al menos el 80% idéntica a la SEQ ID NO: 1 muestra al menos la misma actividad biológica que el polipéptido de secuencia SEQ ID NO: 1.

Una "misma actividad biológica" puede significar una misma función biológica. Por lo tanto, un polipéptido que tiene una misma actividad biológica que el polipéptido de secuencia SEQ ID NO: 1 puede ser, por ejemplo, un polipéptido que tiene actividad de monooxigenasa. Las técnicas para determinar la actividad de monooxigenasa de una enzima son bien conocidas por los expertos. Por ejemplo, un polipéptido que tiene la misma actividad biológica que el polipéptido de secuencia SEQ ID NO: 1 puede ser un polipéptido capaz de catalizar la reacción de escisión de la cadena lateral de monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, análogos y derivados de colesterol para dar pregnenolona u otros precursores y derivados de hormonas esteroideas, como ejemplo no limitante. En este caso, la actividad catalizadora de un compuesto se puede evaluar fácilmente *in vitro* (tal como se muestra por Woods, S.T., J. Sadleir, *et al.* (1998) "Expression of catalytically active human cytochrome p450scc in Escherichia coli and mutagenesis of isoleucine-462."; *Arch Biochem Biophys* 353(1): 109-15) o *in vivo* (como se muestra por Dupont, C., R. Spagnoli, *et al.* (1998) "Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast."; *Nat Biotechnol* 16(2): 186-9) por parte de un experto en la técnica, por ejemplo, por medio del protocolo descrito en el Ejemplo 1.

Normalmente, la actividad catalizadora se puede evaluar en un ensayo de conversión *in vitro* en el que se puede aplicar 150 mM de tampón HEPES, ajustado a pH 7,4, que contiene Tween-20 al 0,05% y MgCl₂ 1 mM, como tampón de reacción. 1 unidad de glucosa-6-fosfato deshidrogenasa, con 5 mM de glucosa-6-fosfato como sustrato, puede servir como sistema de regeneración de NADPH. Normalmente, las concentraciones de CYP11A1, Adx y AdR pueden ser de 1 μM, 20 μM y 0,5 μM, o de 0,25 μM, 5 μM y 0,125 μM, respectivamente. 20 μM de colesterol, disuelto en 2-hidroxipropil-β-ciclodextrina al 45%, puede servir como sustrato para CYP11A1. Normalmente, las muestras están precalentadas a 37 °C y la reacción se inició mediante la adición de NADPH a una concentración final de 100 μM. Las mezclas se pueden incubar a 37 °C con agitación durante 30 s. La reacción se puede detener hirviendo las muestras en agua durante 30 s. Para permitir la detección fotométrica a 240 nm, los esteroides se pueden convertir en sus derivados 3-ceto-Δ⁴ utilizando un colesterol oxidasa de *Nocardia spec.* Normalmente, se añaden 20 μl de una solución de colesterol oxidasa (5 mg de colesterol oxidasa y 5 mg de colato de Na disueltos en 5 ml de tampón HEPES 50 mM, pH 7, que contiene Tween-20 al 0,05%) a las muestras. Después de la incubación a 37 °C durante 1 h, se puede añadir 11-desoxicorticosterona (DOC) a las mezclas de reacción como patrón interno, operación seguida de una extracción de 2 veces con volúmenes iguales de acetato de etilo. Después de la evaporación, los extractos pueden disolverse en acetonitrilo/agua.

La actividad catalizadora también se puede analizar en un ensayo de conversión *in vivo* en el que la conversión de 300 μM de colesterol en progesterona se evalúa normalmente mediante HPLC después de 24 h. Se puede cultivar *Bacillus megaterium* en medio TB que contiene 10 μg/ml de tetraciclina a 37 °C con agitación de 180 rpm. La expresión de la proteína se puede inducir añadiendo 0,25 g de xilosa disueltos en 1 ml de agua, operación seguida de la posterior adición del sustrato, disuelto en 2-hidroxipropil-β-ciclodextrina.

Se puede realizar el alineamiento de la secuencia SEQ ID NO: 1 con una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1. Como estas dos secuencias tienen un alto porcentaje de identidad, la mayor parte de sus aminoácidos son idénticos. Por lo tanto, la numeración de aminoácidos se puede definir para ambas secuencias de forma que al menos el 80% de los aminoácidos presentes en las posiciones que portan el mismo número en ambas secuencias sean idénticos (introduciendo huecos, si es necesario, para lograr el porcentaje máximo de identidad de secuencia). En este contexto, "una posición correspondiente a la posición X" de la secuencia SEQ ID NO: 1 significa la posición que porta dicho número X en la secuencia al menos el 80% idéntica a la SEQ ID NO: 1. La secuencia SEQ ID NO: 1, por lo tanto, puede considerarse una "secuencia de referencia" para la numeración de las posiciones de aminoácidos en un alineamiento con una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1, de forma que al menos el 80% de los aminoácidos presentes en una "posición correspondiente" en ambas secuencias alineadas sean idénticos.

En una forma de realización, la enzima P450 aislada presentada en la invención comprende o consiste en una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1, comprendiendo dicha secuencia tanto una treonina en la posición correspondiente a la posición 225 como un ácido aspártico en la posición correspondiente a la posición 289.

En otra forma de realización, la enzima P450 divulgada comprende o consiste en una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 2 y la SEQ ID NO: 3.

5 En otra forma de realización más, la enzima P450 presentada en la invención consiste en la secuencia SEQ ID NO: 1.

La enzima presentada en la invención puede incluir una o más etiquetas, que pueden facilitar su purificación. Por ejemplo, la etiqueta puede ser una etiqueta de poli-histidina (poli-His).

10 *Ácidos nucleicos, vectores, células huésped y procedimiento de producción de la enzima*

En una forma de realización, la invención presenta un ácido nucleico aislado que comprende o consiste en una secuencia que codifica una enzima P450 presentada en la invención.

15 Los ácidos nucleicos aislados presentados en la invención, también denominados polinucleótidos, pueden ser moléculas de ADN o ARN, que codifican la enzima P450 definida en la sección "Enzimas", aunque teniendo en cuenta la degeneración del código genético. Los ácidos nucleicos aislados pueden obtenerse mediante técnicas estándar bien conocidas por los expertos en la técnica, tales como amplificación o polimerización de ADN *in vitro*, síntesis génica *in vitro*, ligadura de oligonucleótidos, o mediante una combinación de estas técnicas.

20 Los ácidos nucleicos presentados en la invención se encuentran ventajosamente en forma aislada o purificada. Los términos "purificado" y "aislado" tienen el mismo significado que se definió anteriormente.

25 Como entenderán los expertos en la técnica, puede ser ventajoso en algunos casos producir moléculas de nucleótidos que codifican polipéptidos que posean codones de origen no natural en la molécula de nucleótidos naturales. Por ejemplo, los codones preferidos por un huésped procarionta o eucariota particular pueden seleccionarse para aumentar la tasa de expresión de polipéptidos recombinantes.

30 Un ácido nucleico presentado en la presente invención también puede incluir secuencias que codifican etiquetas, proteínas transportadoras, péptidos señal o secuencias no transcritas o traducidas que aumentan la expresión o la estabilidad de la molécula.

35 Los ácidos nucleicos presentados en la invención pueden utilizarse para producir una enzima P450 recombinante en un sistema de expresión adecuado. La expresión "sistema de expresión" significa una célula huésped y un vector compatible en condiciones adecuadas, por ejemplo para la expresión de una proteína codificada por un ácido nucleico extraño transportado por el vector e introducido en la célula huésped.

Normalmente, el ácido nucleico puede estar incluido en cualquier vector adecuado.

40 Por lo tanto, en algunas formas de realización, la invención presenta un vector que comprende un ácido nucleico tal como se ha definido anteriormente asociado operativamente con elementos de control de la expresión y una célula huésped que contiene dicho ácido nucleico o dicho vector.

45 Los términos y las expresiones "vector", "vector de clonación" y "vector de la expresión" significan el vehículo mediante el que se puede introducir una secuencia de ADN o ARN (por ejemplo, un gen extraño) en una célula huésped, para transformar el huésped y promover la expresión (por ejemplo, la transcripción y la traducción) de la secuencia introducida.

50 Se puede utilizar cualquier vector de expresión, tal como un plásmido, cósmido, episoma, cromosoma artificial, fago o un vector vírico. Los ejemplos de plásmidos incluyen plásmidos de replicación que comprenden un origen de replicación, o plásmidos integrativos, tales como, por ejemplo, pUC, pcDNA, pBR y similares.

55 Otros ejemplos de vectores incluyen vectores para células animales tales como pAGE107 (Miyaji H *et al.* 1990), pAGE103 (Mizukami T *et al.* 1987), pHSG274 (Brady G *et al.* 1984), pKCR (O'Hare K *et al.* 1981), pSG1 beta d2-4- (Miyaji H *et al.* 1990) y similares.

60 Los ejemplos de vectores víricos incluyen vectores de adenovirus, retrovirus, herpesvirus y AAV. Pueden producirse virus recombinantes mediante técnicas conocidas en la técnica, tales como mediante la transfección de células de empaquetamiento o mediante la transfección transitoria con plásmidos o virus auxiliares. Los ejemplos típicos de células de empaquetamiento de virus incluyen células PA317, células PsiCRIP, células GPenv+, células 293, etc. Se pueden encontrar protocolos detallados para producir dichos virus recombinantes de replicación defectuosa, por ejemplo, en los documentos WO 95/14785, WO 96/22378, US 5.882.877, US 6.013.516, US 4.861.719, US 5.278.056 y WO 94/19478.

El vector de expresión puede comprender un casete de expresión funcional, tal como un casete de expresión que comprende una secuencia de ácidos nucleicos que codifica un polipéptido presentado en la invención, que está unido operativamente a secuencias de control de la expresión.

5 "Secuencia(s) de control de la expresión" se refiere al o a los elementos necesarios para la expresión de un polipéptido y, opcionalmente, para su regulación. La o las secuencias de control de la expresión pueden incluir, por ejemplo, una secuencia promotora, señales para el inicio y la terminación de la traducción, así como regiones apropiadas para la regulación de la traducción, tales como un promotor, un potenciador, un terminador y similares, para provocar o dirigir la expresión de dicho polipéptido.

10 En una forma de realización, un vector presentado en la invención también puede comprender un gen marcador, por ejemplo, un gen que posibilita seleccionar entre un organismo transformado y un organismo que no contiene el ADN extraño transfectedo. Un gen marcador puede ser un gen que confiere resistencia a un antibiótico.

15 Según una forma de realización específica, un vector presentado en la invención puede comprender además una secuencia de ácidos nucleicos que codifica una adrenodoxina (Adx) y/o una secuencia de ácidos nucleicos que codifica una adrenodoxina reductasa (AdR). Adx y AdR son asociados de óxido-reducción de enzimas P450.

20 "AdR" se refiere a adrenodoxina reductasa (EC: 1.18.1.6) o adrenodoxina-NADP+ reductasa, la enzima que se conoce como el primer componente en el sistema de transferencia de electrones mitocondrial citocromo P450 y que está implicada en la biosíntesis de todas las hormonas esteroideas.

25 En una forma de realización específica, dicha enzima AdR se selecciona del grupo que consiste en AR (NADPH:adrenodoxina oxidoreductasa (EC = 1.18.1.6) codificada por el gen arh1) de *Schizosaccharomyces pombe* o de *Saccharomyces cerevisiae* y FNR (ferredoxina-NADP reductasa (EC = 1.18.1.2) codificada por el gen fpr de *Escherichia coli*.

30 En una forma de realización específica, la enzima AdR es AdR de *Bos taurus* (referenciada como P08165 en UniProtKB/Swiss-Prot, última actualización de secuencia: 15 de julio de 1998). En otra forma de realización específica, la secuencia de proteínas de dicha AdR es la SEQ ID NO: 7. En otra forma de realización específica, la secuencia de proteínas de dicha AdR es una variante de la SEQ ID NO: 7, siempre que conserve su actividad biológica.

35 Por "Adx" se entiende adrenodoxina o ferredoxina 1, la proteína que es conocida por su actividad de transferencia de electrones desde la adrenodoxina reductasa hasta el CYP11A1.

En una forma de realización específica, dicha proteína Adx se selecciona del grupo que consiste en Fdx de origen mamífero, Etp1fd de *Schizosaccharomyces pombe* y Yah1 de *Saccharomyces cerevisiae*.

40 En una forma de realización específica, la proteína Adx es Adx de *Bos taurus* (referenciada como P00257 en UniProtKB/Swiss-Prot, última actualización de secuencia: 1 de julio de 1989). En otra forma de realización específica, la secuencia de proteínas de dicha Adx es la SEQ ID NO: 8. En otra forma de realización específica, la secuencia de proteínas de dicha Adx es una variante de la SEQ ID NO: 8, siempre que conserve su actividad biológica.

45 Se puede utilizar un vector o un ácido nucleico presentado en la invención para transformar células huésped según técnicas conocidas comúnmente por los expertos en la técnica. La inserción de dicho vector en la célula huésped puede ser transitoria o estable.

50 El vector también puede contener secuencias que codifican señales específicas que desencadenan la secreción de la proteína traducida o su direccionamiento a los compartimentos u orgánulos celulares. Estas diversas señales de control se seleccionan según la célula huésped y pueden insertarse en vectores que se autorreplican en la célula huésped, o en vectores que integran el genoma de dicho huésped.

55 Las células huésped pueden ser procariotas o eucariotas, incluidas, entre otras, bacterias, levaduras, células vegetales, células animales, células de insectos y células de mamíferos, incluidas líneas celulares que están disponibles comercialmente. En algunas formas de realización, las células huésped para la expresión son *Escherichia coli*, *Lactobacilli*, *Bacillus*, en particular *Bacillus megaterium*, bacterias probióticas, *Pichia pastoris*, *Saccharomyces cerevisiae*, células de insecto, células vegetales, células COS y células CHO. En una forma de realización, la célula huésped es una célula procariota, en particular una célula bacteriana. En otra forma de realización particular, la célula huésped es una célula eucariota, tal como una célula de levadura tal como una célula de *Saccharomyces cerevisiae*.

60 Los sistemas de expresión comunes incluyen células huésped bacterianas y vectores plasmídicos, células huésped de insectos y vectores de Baculovirus, y células huésped de mamíferos y vectores. Los ejemplos específicos incluyen *E. coli*, *B. megaterium*, levaduras *Kluyveromyces* o *Saccharomyces*, líneas celulares de mamíferos (por

ejemplo, células Vero, células CHO, células 3T3, células COS, etc.), así como cultivos de células de mamíferos primarios o establecidos (por ejemplo, producidos a partir de linfoblastos, fibroblastos, células embrionarias, células epiteliales, células nerviosas, adipocitos, etc.). Los ejemplos también incluyen células SP2/0-Ag14 de ratón (ATCC CRL1581), células P3X63-Ag8.653 de ratón (ATCC CRL1580), células CHO en las que un gen de dihidrofolato reductasa (también denominado "gen DHFR") es defectuoso (Urlaub G *et al*; 1980), células YB2/3HL.P2.G11.16Ag.20 de rata (ATCC CRL1662, también denominadas "células YB2/0") y similares.

En otra forma de realización, la invención presenta una célula que ha sido transfectada, infectada o transformada por un ácido nucleico y/o un vector proporcionado en el presente documento. En consecuencia, la presente invención se refiere también a una célula huésped que contiene un ácido nucleico y/o un vector descrito en el presente documento, así como a la progenie y/o derivados de dichas células huésped. En una forma de realización, la célula huésped es un microorganismo manipulado genéticamente.

Tal como se utiliza en el presente documento, el "microorganismo" puede ser un *Escherichia coli*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus subtilis*, *Kluyveromyces lactis*, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*.

Tal como se muestra por Woods, S.T., J. Sadleir, *et al.* (1998) ((Expression of catalytically active human cytochrome p450_{sc} in *Escherichia coli* and mutagenesis of isoleucine-462. *Arch Biochem Biophys* 353(1): 109-15), la forma madura de P450_{sc} humano se expresó en *Escherichia coli* con el fin de proporcionar una fuente más conveniente de la enzima humana y permitir que se realizaran estudios de estructura-función utilizando mutagénesis dirigida al sitio. Este sistema de expresión permitió producir mayores cantidades de citocromo activo que las que se han aislado previamente a partir de mitocondria placentaria. El P450_{sc} expresado se purificó hasta casi homogeneidad y se demostró que tenía propiedades catalizadoras comparables a la enzima purificada a partir de la placenta humana. La forma madura de adrenodoxina humana también se expresó en *E. coli* y apoyó la actividad de escisión de la cadena lateral de colesterol con la misma V_{max} que se observó con adrenodoxina bovina pero con una Km superior. La mutación de Ile-462 a Leu en P450_{sc} humano provocó una reducción de la constante de velocidad catalítica (k_{cat}) con colesterol como sustrato, aumentó la Km para 22R-hidroxicolesterol, pero no afectó a las constantes cinéticas para 20 alfa-hidroxicolesterol.

Tal como se muestra por Duport, C., R. Spagnoli, *et al.* (1998) (Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nat Biotechnol* 16(2): 186-9), las dos primeras etapas de la ruta esteroideogénica se reprodujeron en *Saccharomyces cerevisiae*. La manipulación genética de la biosíntesis de esteroides mediante la alteración del gen delta 22-desaturasa y la introducción de la actividad de delta 7-reductasa de *Arabidopsis thaliana* y la coexpresión de citocromo P450 de escisión de la cadena lateral bovino, adrenodoxina y adrenodoxina reductasa, conducen a la biosíntesis de pregnenolona a partir de una fuente de carbono sencilla. Después de la coexpresión adicional de 3-beta-hidroxiesteroide deshidrogenasa/isomerasa humana, la pregnenolona se metabolizó adicionalmente para dar progesterona. La formación de esteroides parece estar acoplada a la biosíntesis de esteroides de levadura.

En Szczebara, F. M., C. Chandelier, *et al.* (2003) (Total biosynthesis of hydrocortisone from a simple carbon source in yeast; *Nat Biotechnol* 21(2): 143-9), la producción de hidrocortisona, el principal glucocorticoide suprarrenal de mamíferos y un intermediario importante de la síntesis de fármacos esteroideos, se comunicó a partir de una fuente de carbono sencilla por cepas de *Saccharomyces cerevisiae* recombinantes. Una ruta biosintética artificial y completamente autosuficiente que involucra 13 genes manipulados genéticamente se ensambla y se expresa en una única cepa de levadura. La biosíntesis de esteroides endógenos se redireccionó para producir esteroides compatibles para que sirvieran como sustratos para la parte heteróloga de la ruta. La biosíntesis involucró ocho proteínas de mamíferos (formas maduras de CYP11A1, adrenodoxina (ADX) y adrenodoxina reductasa (ADR); formas mitocondriales de ADX y CYP11B1; 3beta-HSD, CYP17A1 y CYP21A1). La optimización implicó la modulación de los dos sistemas mitocondriales y la interrupción de reacciones secundarias no deseadas asociadas con los productos de los genes ATF2, GCY1 e YPR1. La hidrocortisona fue el principal esteroide producido. Este trabajo demostró la viabilidad de transferir una ruta biosintética compleja de eucariotas superiores a microorganismos.

En una forma de realización específica, el microorganismo es *Bacillus megaterium*, en particular, la cepa denominada MS941 de *Bacillus megaterium*. La expresión "cepa MS941 de *Bacillus megaterium*" se refiere a la cepa referenciada por Wittchen y Meinhardt, 1995, *Appl Microbiol Biotechnol*. 42: 871-877, y derivada de la cepa DSM319 (Deutsche Stammsammlung von Mikroorganismen und Zellkulturen).

El sistema descrito anteriormente que utiliza *E. coli* y el sistema que utiliza *B. megaterio* son dos sistemas que son capaces de realizar la conversión de esteroides y derivados de esteroides, ya sea *in vitro* con *E. coli* o *in cellulo* con *B. megaterio*. Estas tecnologías pueden evaluar las capacidades de P450_{sc} y los sustratos para funcionar conjuntamente. Ambas tecnologías requieren un procedimiento eficaz para solubilizar los sustratos de forma que sean accesibles a la enzima. Esta solubilización podría ser exigente a nivel industrial (gran volumen) tanto en términos financieros como científicos.

5 Un tercer organismo *S.cerevisiae* es un organismo particularmente ventajoso en este sentido; su amplia fermentación está bien dominada a niveles de metros cúbicos o más elevados y tiene una potente capacidad para producir esteroides tales como la pregnenolona a partir de una fuente de carbono sencilla tal como glucosa y/o etanol, evitando así detergentes o productos químicos solubilizantes tales como betaciclodextrina (tal como se ha mencionado anteriormente: Duport, C., R. Spagnoli, *et al.* (1998); Szczebara, FM, C. Chandelier, *et al.* (2003)). Para permitir la síntesis de esteroides *in vivo*, la biosíntesis de esteroides se direccionó a la producción de un esteroide compatible con P450scc.

10 Como la fermentación de *S. cerevisiae* se domina bien a gran escala, la introducción del ADNc de P450scc descrita recientemente en una *S. cerevisiae* direccionada para campesterol, u otros esteroides descritos como sustratos, en presencia de un portador de electrones apropiado es un recurso para la producción de esteroides. Este organismo recombinante se puede utilizar para evaluar un nuevo ADNc de P450scc tal como se describe en el presente documento.

15 Por lo tanto, en otra forma de realización específica, el microorganismo es *Saccharomyces cerevisiae*.

20 La expresión microorganismo "manipulado genéticamente" se refiere a cualquier microorganismo que haya sido modificado mediante técnicas de ingeniería genética conocidas en el campo por los expertos en la técnica. Para procedimientos relacionados con microorganismos particulares, el experto en la técnica puede recurrir a manuales de referencia tales como para levadura: "Methods in yeast genetics" - A laboratory course manual by M Rose, F Winston y P Hieter. p 198. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York. 1990. ISBN 0-87969-354-1.

25 Estas técnicas son técnicas convencionales, a menos que se indique lo contrario, en los campos de la bioinformática, biología celular, cultivo celular, biología molecular, biología transgénica, microbiología, ADN recombinante e inmunología, que se encuentran dentro de los conocimientos de la técnica. Dichas técnicas se explican completamente en la literatura.

30 Tal como se utiliza en el presente documento, "técnicas de ingeniería genética" se refiere a tecnologías que permiten la expresión o la sobreexpresión de la expresión génica conocidas en el campo por los expertos en la técnica. La expresión o la sobreexpresión de un gen de interés se puede lograr introduciendo en una célula o un microorganismo una secuencia de ácidos nucleicos exógena que comprende dicho gen de interés, mediante cualquier tecnología de transformación conocida por los expertos en la técnica.

35 El término "transformación" o "transfección" significa la introducción de un gen o una secuencia de ADN o ARN "extraños" (es decir, heterólogos) en una célula huésped, de tal forma que la célula huésped expresará el gen o la secuencia introducidos para producir una sustancia deseada, normalmente una proteína o enzima codificada por el gen introducido o la secuencia introducida. La transfección de la célula huésped se puede realizar utilizando cualquier técnica estándar, tal como transformación química, electroporación, precipitación con fosfato cálcico o lipofección. Las técnicas de transformación también incluyen, por ejemplo, la técnica de transformación de protoplastos mediada por PEG (Barg et al, 2005, Microbial Processes and Products 165-184).

45 "Secuencia(s) exógena(s) de ácidos nucleicos" se refiere a la o las secuencias de ácidos nucleicos que no se expresan originariamente en el microorganismo considerado y/o en la forma que lo hacen en la cepa natural del microorganismo (en términos de nivel de expresión, por ejemplo), y que se han utilizado para transformar dicho microorganismo con el fin de obtener un microorganismo manipulado genéticamente tal como se ha indicado anteriormente. En una forma de realización específica, la secuencia de ácidos nucleicos exógena se origina a partir de otra especie diferente del microorganismo considerado (por ejemplo, otra especie de microorganismo u organismo). En otra forma de realización específica, la secuencia exógena se origina a partir del mismo microorganismo.

50 Dicha o dichas secuencias de ácidos nucleicos exógenas pueden codificar proteínas de interés tales como el citocromo P450 y los asociados de óxido-reducción AdR y Adx, y la expresión "ADN exógeno" puede designar cada secuencia individual o abarcar una secuencia completa que comprende cada secuencia individual. Como ejemplo no limitante, dichas secuencias de ácidos nucleicos exógenas se integran en el genoma de dicho microorganismo mediante técnicas conocidas en el campo, tales como mediante recombinación homóloga.

60 En un aspecto, la invención proporciona un microorganismo manipulado genéticamente capaz de convertir un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, análogos y derivados de colesterol en un precursor de hormonas esteroideas, en el que dicho microorganismo comprende un ácido nucleico presentado en la invención, y opcionalmente una secuencia de ácidos nucleicos exógena que codifica una Adx y/o una secuencia de ácidos nucleicos exógena que codifica una AdR.

65 En una forma de realización específica, el microorganismo manipulado genéticamente es *Saccharomyces cerevisiae*.

En otra forma de realización específica, el microorganismo manipulado genéticamente es *Bacillus megaterium*, en particular la cepa MS941 de *Bacillus megaterium*.

5 La presente invención también se refiere a un procedimiento *in vitro* para preparar una enzima P450 presentada en la invención, comprendiendo dicho procedimiento las etapas siguientes:

a) cultivar una célula huésped en condiciones adecuadas para obtener la expresión de la enzima P450 y

10 b) recuperar la enzima expresada.

Las "condiciones adecuadas para obtener la expresión de la enzima P450" son bien conocidas por los expertos en la técnica. Normalmente se utilizan medios LB (25 g/l), TB (24 g/l de extracto de levadura, 12 g/l de triptona, 0,4% de glicerol, tampón de fosfato de potasio 10 mM) o EnPresso™ Tablet para el cultivo de la célula huésped. Se puede realizar un precultivo inoculando 50 ml de medio que contiene 10 µg/ml de tetraciclina con células de una placa o solución madre de glicerol. Después se puede realizar un cultivo principal inoculando 50 ml de medio que contiene 10 µg/ml de tetraciclina con 500 µl de muestra del precultivo. El cultivo principal se cultiva normalmente hasta que se logra una densidad óptica de ~0,4. La expresión de proteínas se puede inducir después de la adición de 0,25 g de xilosa disueltos en 1 ml de agua destilada.

La enzima P450 se puede purificar por medio de procedimientos bien conocidos de purificación: se puede purificar a partir de lisados o extractos celulares, cuerpos de inclusión o a partir del sobrenadante del cultivo mediante procedimientos tales como cromatografía HPLC, técnicas de inmunoafinidad con anticuerpos específicos y similares.

25 Alternativamente, la enzima P450 puede expresarse *in vitro* con un sistema de transcripción y traducción exento de células a partir de una matriz de ADN o ARN que contiene los elementos necesarios para su expresión en un lisado celular o un sistema reconstituido (por ejemplo, Rapid Translation System®, Roche Diagnostics o Retic Lysate IVT™, Ambion).

30 Procedimientos de producción de un precursor de hormonas esteroideas

Tal como se muestra en los Ejemplos 2-4, los inventores han desarrollado una nueva enzima P450_{scc} que porta mutaciones específicas que muestran una actividad enzimática mejorada en términos de conversión de sustrato en precursor de hormonas esteroideas.

35 Por lo tanto, en otro aspecto, la invención se refiere al uso de una enzima P450 presentada en la invención para producir un precursor de hormonas esteroideas.

La invención proporciona también procedimientos para producir un precursor de hormonas esteroideas tal como se describe a continuación.

Un primer procedimiento para producir un precursor de hormonas esteroideas comprende las etapas siguientes:

45 a) proporcionar un microorganismo tal como se ha descrito anteriormente,

b) cultivar dicho microorganismo en condiciones que permitan la expresión de una enzima P450 presentada en la invención,

50 c) poner en contacto el cultivo de microorganismos obtenido en la etapa b) con un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo de colesterol y un derivado de colesterol, en condiciones que permitan la producción, por microorganismo, de un precursor de hormonas esteroideas de dicho sustrato y

d) recuperar el precursor de hormonas esteroideas producido.

55 Según una forma de realización específica, las etapas b) y c) se realizan simultáneamente. En este caso, el procedimiento para producir un precursor de hormonas esteroideas comprende las etapas siguientes:

60 a) proporcionar un microorganismo tal como se ha descrito anteriormente,

b) cultivar dicho microorganismo en presencia de un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo de colesterol y un derivado de colesterol, en condiciones que permitan la expresión de una enzima P450 presentada en la invención y permitir la producción, por el microorganismo, del precursor de la hormona esteroidea a partir de dicho sustrato y

65

c) recuperar el precursor de hormonas esteroideas producido.

Un segundo procedimiento para producir un precursor de hormonas esteroideas comprende las etapas siguientes:

5 a) poner en contacto una enzima P450 presentada en la invención con un polipéptido Adx aislado, un polipéptido AdR aislado y un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo de colesterol y un derivado de colesterol en condiciones que permitan la transformación de dicho sustrato en un precursor de hormonas esteroideas y

10 b) recuperar el precursor de hormonas esteroideas obtenido.

15 Tal como se utiliza en el presente documento, el término "sustrato" abarca monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como fitosterol derivado de cicloartenol y lanosterol. Entre estos sustratos, "colesterol, análogos de colesterol y derivados de los mismos" se refiere a una lista de sustratos seleccionados del grupo que consiste en colesterol, brassicasterol, campesterol, ergostadienol tal como ergosta-5,22-dienol, ergosta-5,24(28)-dienol, ergosta-5,24(28)-dien-3-beta-ol, ergosta-5,24(25)-dienol, ergostatrienol tal como ergosta-5,22,24(25)-trienol, ergosta-5,22,24(28)-trienol, ergosta-5,7,22-trienol, ergostatetraenol tal como ergosta-5,7,22,24(25) o ergosta-5,7,22,24(28), desmosterol, beta-sitosterol, generol, una mezcla de oxisteroles, estigmasterol, vitamina D, 7-dehidrocolesterol y ergosterol. Las mezclas de esteroides que se utilizan actualmente en procesos industriales también están abarcadas por esta definición de sustratos, tales como generol 100 y ADM90 (que comprende brassicasterol + campesterol + estigmasterol + beta-sitosterol en diferentes proporciones).

20 En una forma de realización particular, el sustrato se selecciona del grupo que consiste en colesterol, campesterol, desmosterol y ergosta-5,24(28)-dien-3beta-ol. En una forma de realización específica, el sustrato es ergosterol.

25 En una forma de realización específica, el "precursor de hormonas esteroideas" se selecciona del grupo que consiste en pregnenolona, 7-deshidropregnenolona, hidroxiergosterol e hidroxiestigmasterol. En otra forma de realización específica, dicho precursor de hormonas esteroideas se selecciona del grupo de análogos de colesterol hidroxilados y secoesteroides (tales como vitaminas D2 y D3 como derivados de los análogos de colesterol 7-dehidrocolesterol y ergosterol). En una forma de realización particular, el precursor de hormonas esteroideas es pregnenolona.

30 "Cultivar el microorganismo en presencia de un sustrato" significa hacer interactuar físicamente dicho microorganismo con un sustrato tal como se ha definido anteriormente. Esta interacción se puede lograr dentro de un medio de cultivo o no. En una forma de realización específica, dicho medio de cultivo contiene agentes para la permeabilización de un microorganismo según la presente invención y/o la solubilización de un sustrato según la presente invención. La disolución del sustrato en estos agentes puede ser anterior a la adición al cultivo de microorganismos.

35 "Poner en contacto una enzima P450 con un sustrato" significa hacer interactuar físicamente dicha enzima P450 con un sustrato tal como se ha definido anteriormente. Esta interacción se puede lograr dentro de un medio o no. En una forma de realización específica, dicho medio contiene agentes para la solubilización de un sustrato según la presente invención. La disolución del sustrato en estos agentes puede ser anterior a la adición a dicho medio.

40 Los agentes utilizados para la disolución del sustrato pueden seleccionarse del grupo que consiste en etanol, Tween-80, tergitol, polivinilpirrolidona (PVP), saponinas (tales como saponina de *Quillaja* que está contenida en extractos brutos del árbol de corteza de jabón *Quillaja saponaria*, por ejemplo), ciclodextrinas y derivados de las mismas (por ejemplo, 2-hidroxiopropil- β -ciclodextrina). En otra forma de realización específica, se pueden utilizar mezclas de estos agentes, tales como una mezcla de etanol y Tween-80, una mezcla de tergitol y etanol, una mezcla de saponinas (por ejemplo, saponina de *Quillaja*) y ciclodextrinas como ejemplos no limitantes. Se pueden utilizar derivados de ciclodextrina, tales como 2-hidroxiopropil- β -ciclodextrina. En otra forma de realización específica, los sustratos se cocrystalizan con polivinilpirrolidona (PVP).

45 En otra forma de realización, el sustrato se disuelve en primer lugar en 2-hidroxiopropil- β -ciclodextrina antes de la adición al cultivo de microorganismos, en el que dicho cultivo contiene saponina de *Quillaja*. En otra forma de realización, los sustratos se disuelven en una solución que comprende un porcentaje de 2-hidroxiopropil- β -ciclodextrina que varía del 10 al 60%, por ejemplo, del 20 al 50%, por ejemplo, del 40 al 50%, y un porcentaje de saponina de *Quillaja* que varía del 1 al 10%, por ejemplo, del 2 al 8%, por ejemplo, del 3 al 6%. En otra forma de realización, los sustratos se disuelven en una solución que comprende el 45% de 2-hidroxiopropil- β -ciclodextrina y el 4% de saponina de *Quillaja*.

50 En otra forma de realización, la concentración final de 2-hidroxiopropil- β -ciclodextrina en el cultivo de microorganismos se encuentra entre el 1 y el 4%, por ejemplo, entre el 2 y el 3%, tal como el 2,25% tal como se ilustra en el ejemplo 2. En otra forma de realización, la concentración final de saponina de *Quillaja* en el cultivo de microorganismos se encuentra entre el 0,05 y el 0,25%, por ejemplo, entre el 0,075 y el 0,225%, por ejemplo, entre el 0,1 y el 0,2%.

El "cultivo de dicho microorganismo en condiciones que permiten la expresión de dichas secuencias de ácidos nucleicos exógenas" se puede realizar según cualquier procedimiento de cultivo e inducción bien conocido en el campo de la biotecnología, tal como se describe por Bleif *et al.* (Appl Microbiol Biotechnol (2012) 93: 1135-1146) o por Korneli *et al.* (Journal of Biotechnology 163 (2013) 87-96).

Las "condiciones que permiten la transformación de dicho sustrato en un precursor de hormonas esteroideas" incluyen condiciones *in vitro*, tales como los reactivos utilizados, las concentraciones de reactivos, la temperatura, el uso de agitación, la duración de la reacción o la incubación, etc., que favorecen la transformación de dicho sustrato en un precursor de hormonas esteroideas. Estas condiciones pueden determinarse y ajustarse mediante procedimientos conocidos por los expertos en la técnica.

Como ejemplo no limitante, la reacción se puede realizar en presencia de glucosa-6-fosfato deshidrogenasa y glucosa-6-fosfato, normalmente en 150 mM de tampón HEPES, por ejemplo, ajustado a pH 7,4, que contiene típicamente Tween-20 al 0,05%. y MgCl₂ 1 mM. Las concentraciones de la enzima P450, Adx y AdR pueden ser, por ejemplo, de 1 μM, 20 μM y 0,5 μM, o de 0,25 μM, 5 μM y 0,125 μM, respectivamente. Los monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, análogo de colesterol o derivado de colesterol pueden disolverse en una 2-hidroxipropil-β-ciclodextrina al 45%. Las muestras pueden precalentarse a 37 °C y se puede añadir NADPH a una concentración final de 100 μM. Las mezclas pueden incubarse a 37 °C con agitación.

En una forma de realización específica, el precursor de hormonas esteroideas utilizado en cualquier procedimiento para producir un precursor de hormonas esteroideas es pregnenolona. Tal como se utiliza en el presente documento, "pregnenolona" se refiere a una hormona esteroidea también denominada 3β-hidroxi pregn-5-en-20-ona.

Breve descripción de las secuencias

SEQ ID NO: 1 muestra la secuencia de aminoácidos de CYP11A1 que comprende una treonina en la posición 225 y un ácido aspártico en la posición 289 (polipéptido SA1).

SEQ ID NO: 2 muestra la secuencia de aminoácidos de CYP11A1 que comprende una treonina en la posición 225.

SEQ ID NO: 3 muestra la secuencia de aminoácidos de CYP11A1 que comprende un ácido aspártico en la posición 289.

SEQ ID NO: 4 muestra la secuencia de aminoácidos de CYP11A1 de Bos taurus de WT (polipéptido SA4; P450scc de WT).

SEQ ID NO: 5 muestra la secuencia del plásmido pSMF2.1_CYP11A1BYM que codifica SA1 de SEQ ID NO: 1.

SEQ ID NO: 6 muestra la secuencia de aminoácidos del polipéptido SA6 (P450scc-I1A-K193E).

SEQ ID NO: 7 Muestra la secuencia de aminoácidos de AdR.

SEQ ID NO: 8 Muestra la secuencia de aminoácidos de Adx.

Breve descripción de las figuras

Figura 1: Cromatograma de HPLC que muestra la conversión *in vitro* de colesterol 20 μM por los polipéptidos citocromo P450 SA1 (SEQ ID NO: 1) y SA4 (control que tiene la secuencia SEQ ID NO: 4). Leyenda: S: sustrato, P: producto principal, DOC: 11-desoxicorticosterona.

Figura 2: Cuantificación de pregnenolona a partir de cromatografía HPLC que muestra la conversión *in vivo* de diferentes sustratos en pregnenolona por el polipéptido citocromo P450 SA1 (SEQ ID NO: 1).

Figura 3: Cromatograma de HPLC que muestra la conversión *in vivo* de colesterol 300 μM por los polipéptidos citocromo P450 SA1 (SEQ ID NO: 1) y SA6 (control que tiene la secuencia SEQ ID NO: 6) después de 24 h. Leyenda: Col.: colesterol, Prog.: pregnenolona.

Figura 4: Transcurso del tiempo para la conversión *in vivo* de colesterol 300 μM por los polipéptidos citocromo P450 SA1 (SEQ ID NO: 1) y SA6 (control que tiene la secuencia SEQ ID NO: 6).

Figura 5: Mapa vectorial de pSMF2.1_CYP11A1BYM que codifica SA1 (SEQ ID NO: 1).

Ejemplos

Ejemplo 1: Materiales y procedimientos

5 Síntesis de proteínas

Se obtuvieron variantes de CYP11A1 mediante síntesis génica y se clonaron en pTRC99A, fusionadas con una etiqueta de poli-His. Se cotransformaron células de *E. coli* C43DE3 con cada vector y el pGro12 que codifica la chaperona. La purificación se realizó tal como se describe por Janocha *et al.* (Biochim Biophys Acta. (2011), enero; 1814 (1): 126-31). La pureza se evaluó mediante SDS-PAGE. Las concentraciones de las proteínas purificadas se determinaron mediante espectros de diferencia de CO, después de tratamiento con ditionito de sodio y exposición a CO.

Se purificaron Adx y AdR según Uhlmann *et al.* (Biochem. Biophys. Res. Commun., 188 (1992), p. 1131-1138) y Sagara *et al.* (Biol. Pharm. Bull., 16 (1993), p. 627-630), respectivamente.

Ensayo de conversión *in vitro*

Para ensayos enzimáticos *in vitro* se aplicaron como tampón de reacción 150 mM de tampón HEPES, ajustado a pH 7,4, que contenía Tween-20 al 0,05% y MgCl₂ 1 mM. 1 unidad de glucosa-6-fosfato deshidrogenasa, con glucosa-6-fosfato 5 mM como sustrato, sirvió como un sistema de regeneración de NADPH.

Para P450scc WT (anotado como "SA4", SEQ ID NO: 4) y P450scc-11A-K193E (anotado como "SA6", SEQ ID NO: 6) las concentraciones de CYP11A1, Adx y AdR fueron de 1 μM, 20 μM y 0,5 μM, y para P450scc-R225T-N289D (anotado como "SA1", SEQ ID NO: 1) las concentraciones de CYP11A1, Adx y AdR fueron de 0,25 μM, 5 μM y 0,125 μM, respectivamente.

20 μM de colesterol, disuelto en 2-hidroxipropil-β-ciclodextrina al 45%, sirvió como sustrato para CYP11A1.

30 Las muestras se precalentaron a 37 °C y la reacción se inició mediante la adición de NADPH a una concentración final de 100 μM. Las mezclas se incubaron a 37 °C con agitación durante 30 s. La reacción se detuvo hirviendo las muestras en agua durante 30 s.

35 Para permitir la detección fotométrica a 240 nm, los esteroides se convirtieron en sus derivados 3-ceto-Δ₄ utilizando una colesterol oxidasa de *Nocardia spec.*

Se añadieron 20 μl de una solución de colesterol oxidasa (5 mg de colesterol oxidasa y 5 mg de colato de Na disueltos en 5 ml de tampón HEPES 50 mM, pH 7, que contenía Tween-20 al 0,05%) a las muestras. Después de la incubación a 37 °C durante 1 h, se añadió 11-desoxicorticosterona (DOC) a las mezclas de reacción como patrón interno, operación seguida de una extracción de 2 veces con volúmenes iguales de acetato de etilo. Después de la evaporación, los extractos se disolvieron en acetonitrilo/agua.

Ensayo de conversión *in vivo*

45 La conversión *in vivo* de colesterol 300 μM en pregnenolona se evaluó mediante HPLC después de 24 h. Se cultivó *Bacillus megaterium* en medio TB que contenía 10 μg/ml de tetraciclina a 37 °C con agitación de 180 rpm. La expresión de la proteína se indujo mediante la adición de 0,25 g de xilosa disuelta en 1 ml de agua, seguida de la adición subsiguiente del sustrato, disuelto en 2-hidroxipropil-β-ciclodextrina.

50 **Ejemplo 2: Conversión *in vitro* de colesterol por los polipéptidos citocromo P450 SA1 (SEQ ID NO: 1) y SA4 (control WT, SEQ ID NO: 4)**

Después de la evaluación profunda de la concentración de cada preparación de P450scc, se utilizaron SA1, SA4 y SA6 en los ensayos *in vitro* siguientes. Cada una de estas isoformas se reconstituyó a una concentración de 1 μM en presencia de colesterol como sustrato (20 μM) en cantidades saturadas de Adx y AdR.

Mientras que SA4 y SA6 mostraron una actividad idéntica, SA1 mostró una actividad 2-3 veces mayor en comparación con SA4 y SA6 (**Figura 1**). Dado que el sustrato ya está agotado después del tiempo de incubación, es difícil evaluar las diferencias entre los polipéptidos SA1 y SA1/SA4 en un tiempo de incubación más prolongado.

60 **Ejemplo 3: Conversión de colesterol *in vivo* por los polipéptidos citocromo P450 SA1 (SEQ ID NO: 1) y SA6 (control, SEQ ID NO: 6)**

65 Como no se observaron diferencias medibles entre SA4 y SA6, los inventores enfocaron su interés en comparar los polipéptidos SA1 y SA6. Recientemente se ha desarrollado un nuevo sistema en el que el colesterol se metaboliza dando pregnenolona por medio de *Bacillus megaterium* recombinante que expresa P450scc, Fdx1 y FdxR en

presencia de colesterol solubilizado. Los ADNc correspondientes a SA1 y SA6 se optimizaron para codón BIAS de *Bacillus megaterium* y se transfirieron al plásmido pSFM2.1 utilizando los sitios de restricción apropiados. Ambos plásmidos PSFM2.1 que portan respectivamente el codón optimizado SA1 y SA6 se transfirieron a MS941 de *Bacillus megaterium* utilizando un protocolo clásico de transformación de protoplastos.

5 La actividad de conversión de colesterol *in vivo* se evaluó después de 24 h y 48 h para los polipéptidos SA1 y SA6 expresados en *Bacillus megaterium*.

10 Según experimentos *in vitro*, el SA1 mostró ahora una actividad 2 veces mayor en comparación con SA6 (Figuras 3 y 4).

Ejemplo 4: Conversión *in vivo* de diferentes sustratos en pregnenolona por medio del polipéptido citocromo P450 SA1 (SEQ ID NO: 1)

15 Finalmente, se utilizó el sistema mejorado para analizar la capacidad de conversión del polipéptido SA1 con varios sustratos de interés biotecnológico: campesterol, desmosterol, ergosta-5,24(28)-dien-3 β -ol y una mezcla de varios 20, 22-OH-oxiesteroles.

20 Cada sustrato se convirtió en un producto principal con el mismo tiempo de retención que la progesterona, lo que indica que CYP11A1 fue capaz de escindir la cadena lateral de cada uno de estos sustratos, produciendo pregnenolona.

25 La formación de pregnenolona se cuantificó para todos los sustratos (Figura 2). Los oxiesteroles polares se convirtieron a una tasa comparable al colesterol (~35 mg/l después de 48 horas). La conversión de los esteroides más hidrófobos campesterol, desmosterol y ergostadienol solo se produjo a una tasa de aproximadamente el 19% después de 48 horas, en comparación con el colesterol.

30 Tomados en conjunto, los resultados muestran que todos los esteroides analizados fueron capaces de penetrar a través de la membrana celular de *Bacillus megaterium* y se convirtieron en pregnenolona por medio del mutante SAP de CYP11A1.

Los inventores informan así en el presente documento una nueva proteína P450scc, denominada SA1, que muestra una mayor actividad de conversión hacia varios esteroides, incluido el colesterol.

35 Listado de secuencias

<110> SANOFI S.A.

<120> POLIPÉPTIDO CITOCROMO P450 NOVEDOSO CON ACTIVIDAD ENZIMÁTICA AUMENTADA

40 <130> BET 13P3205

<150> EP 14305071.4

45 <151> 20/01/2014

<160> 8

<170> PatentIn versión 3.5

50 <210> 1

<211> 481

<212> PRT

<213> Secuencia artificial

55 <220>

<223> Secuencia de aminoácidos de CYP11A1 mutada

<400> 1

ES 2 724 373 T3

Ile Ser Thr Lys Thr Pro Arg Pro Tyr Ser Glu Ile Pro Ser Pro Gly
 1 5 10 15

Asp Asn Gly Trp Leu Asn Leu Tyr His Phe Trp Arg Glu Lys Gly Ser
 20 25 30

Gln Arg Ile His Phe Arg His Ile Glu Asn Phe Gln Lys Tyr Gly Pro
 35 40 45

Ile Tyr Arg Glu Lys Leu Gly Asn Leu Glu Ser Val Tyr Ile Ile His
 50 55 60

Pro Glu Asp Val Ala His Leu Phe Lys Phe Glu Gly Ser Tyr Pro Glu
 65 70 75 80

Arg Tyr Asp Ile Pro Pro Trp Leu Ala Tyr His Arg Tyr Tyr Gln Lys
 85 90 95

Pro Ile Gly Val Leu Phe Lys Lys Ser Gly Thr Trp Lys Lys Asp Arg
 100 105 110

Val Val Leu Asn Thr Glu Val Met Ala Pro Glu Ala Ile Lys Asn Phe
 115 120 125

Ile Pro Leu Leu Asn Pro Val Ser Gln Asp Phe Val Ser Leu Leu His
 130 135 140

Lys Arg Ile Lys Gln Gln Gly Ser Gly Lys Phe Val Gly Asp Ile Lys

ES 2 724 373 T3

Leu Ser Lys Asp Lys Asp Leu Ile His Phe Arg Asn Leu Gly Phe Gly
 405 410 415

Trp Gly Val Arg Gln Cys Val Gly Arg Arg Ile Ala Glu Leu Glu Met
 420 425 430

Thr Leu Phe Leu Ile His Ile Leu Glu Asn Phe Lys Val Glu Met Gln
 435 440 445

His Ile Gly Asp Val Asp Thr Ile Phe Asn Leu Ile Leu Thr Pro Asp
 450 455 460

Lys Pro Ile Phe Leu Val Phe Arg Pro Phe Asn Gln Asp Pro Pro Gln
 465 470 475 480

Ala

<210> 2

<211> 481

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de CYP11A1 mutada

<400> 2

Ile Ser Thr Lys Thr Pro Arg Pro Tyr Ser Glu Ile Pro Ser Pro Gly
 1 5 10 15

Asp Asn Gly Trp Leu Asn Leu Tyr His Phe Trp Arg Glu Lys Gly Ser
 20 25 30

Gln Arg Ile His Phe Arg His Ile Glu Asn Phe Gln Lys Tyr Gly Pro
 35 40 45

Ile Tyr Arg Glu Lys Leu Gly Asn Leu Glu Ser Val Tyr Ile Ile His
 50 55 60

Pro Glu Asp Val Ala His Leu Phe Lys Phe Glu Gly Ser Tyr Pro Glu
 65 70 75 80

Arg Tyr Asp Ile Pro Pro Trp Leu Ala Tyr His Arg Tyr Tyr Gln Lys
 85 90 95

Pro Ile Gly Val Leu Phe Lys Lys Ser Gly Thr Trp Lys Lys Asp Arg
 100 105 110

Val Val Leu Asn Thr Glu Val Met Ala Pro Glu Ala Ile Lys Asn Phe

ES 2 724 373 T3

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

ES 2 724 373 T3

Ile Pro Ala Lys Thr Leu Val Gln Val Ala Ile Tyr Ala Met Gly Arg
 370 375 380

Asp Pro Ala Phe Phe Ser Ser Pro Asp Lys Phe Asp Pro Thr Arg Trp
 385 390 395 400

Leu Ser Lys Asp Lys Asp Leu Ile His Phe Arg Asn Leu Gly Phe Gly
 405 410 415

Trp Gly Val Arg Gln Cys Val Gly Arg Arg Ile Ala Glu Leu Glu Met
 420 425 430

Thr Leu Phe Leu Ile His Ile Leu Glu Asn Phe Lys Val Glu Met Gln
 435 440 445

His Ile Gly Asp Val Asp Thr Ile Phe Asn Leu Ile Leu Thr Pro Asp
 450 455 460

Lys Pro Ile Phe Leu Val Phe Arg Pro Phe Asn Gln Asp Pro Pro Gln
 465 470 475 480

Ala

<210> 3

<211> 481

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia de aminoácidos de CYP11A1 mutada

10

<400> 3

Ile Ser Thr Lys Thr Pro Arg Pro Tyr Ser Glu Ile Pro Ser Pro Gly
 1 5 10 15

Asp Asn Gly Trp Leu Asn Leu Tyr His Phe Trp Arg Glu Lys Gly Ser
 20 25 30

Gln Arg Ile His Phe Arg His Ile Glu Asn Phe Gln Lys Tyr Gly Pro
 35 40 45

Ile Tyr Arg Glu Lys Leu Gly Asn Leu Glu Ser Val Tyr Ile Ile His
 50 55 60

Pro Glu Asp Val Ala His Leu Phe Lys Phe Glu Gly Ser Tyr Pro Glu
 65 70 75 80

Arg Tyr Asp Ile Pro Pro Trp Leu Ala Tyr His Arg Tyr Tyr Gln Lys

ES 2 724 373 T3

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

ES 2 724 373 T3

Leu Lys Ala Ser Ile Lys Glu Thr Leu Arg Leu His Pro Ile Ser Val
 340 345 350

Thr Leu Gln Arg Tyr Pro Glu Ser Asp Leu Val Leu Gln Asp Tyr Leu
 355 360 365

Ile Pro Ala Lys Thr Leu Val Gln Val Ala Ile Tyr Ala Met Gly Arg
 370 375 380

Asp Pro Ala Phe Phe Ser Ser Pro Asp Lys Phe Asp Pro Thr Arg Trp
 385 390 395 400

Leu Ser Lys Asp Lys Asp Leu Ile His Phe Arg Asn Leu Gly Phe Gly
 405 410 415

Trp Gly Val Arg Gln Cys Val Gly Arg Arg Ile Ala Glu Leu Glu Met
 420 425 430

Thr Leu Phe Leu Ile His Ile Leu Glu Asn Phe Lys Val Glu Met Gln
 435 440 445

His Ile Gly Asp Val Asp Thr Ile Phe Asn Leu Ile Leu Thr Pro Asp
 450 455 460

Lys Pro Ile Phe Leu Val Phe Arg Pro Phe Asn Gln Asp Pro Pro Gln
 465 470 475 480

Ala

<210> 4
 <211> 481

5

<212> PRT
 <213> Bos taurus

<400> 4

10

Ile Ser Thr Lys Thr Pro Arg Pro Tyr Ser Glu Ile Pro Ser Pro Gly
 1 5 10 15

Asp Asn Gly Trp Leu Asn Leu Tyr His Phe Trp Arg Glu Lys Gly Ser
 20 25 30

Gln Arg Ile His Phe Arg His Ile Glu Asn Phe Gln Lys Tyr Gly Pro
 35 40 45

Ile Tyr Arg Glu Lys Leu Gly Asn Leu Glu Ser Val Tyr Ile Ile His
 50 55 60

ES 2 724 373 T3

Pro Glu Asp Val Ala His Leu Phe Lys Phe Glu Gly Ser Tyr Pro Glu
65 70 75 80

Arg Tyr Asp Ile Pro Pro Trp Leu Ala Tyr His Arg Tyr Tyr Gln Lys
85 90 95

Pro Ile Gly Val Leu Phe Lys Lys Ser Gly Thr Trp Lys Lys Asp Arg
100 105 110

Val Val Leu Asn Thr Glu Val Met Ala Pro Glu Ala Ile Lys Asn Phe
115 120 125

Ile Pro Leu Leu Asn Pro Val Ser Gln Asp Phe Val Ser Leu Leu His
130 135 140

Lys Arg Ile Lys Gln Gln Gly Ser Gly Lys Phe Val Gly Asp Ile Lys
145 150 155 160

Glu Asp Leu Phe His Phe Ala Phe Glu Ser Ile Thr Asn Val Met Phe
165 170 175

Gly Glu Arg Leu Gly Met Leu Glu Glu Thr Val Asn Pro Glu Ala Gln
180 185 190

Lys Phe Ile Asp Ala Val Tyr Lys Met Phe His Thr Ser Val Pro Leu
195 200 205

Leu Asn Val Pro Pro Glu Leu Tyr Arg Leu Phe Arg Thr Lys Thr Trp
210 215 220

Arg Asp His Val Ala Ala Trp Asp Thr Ile Phe Asn Lys Ala Glu Lys
225 230 235 240

Tyr Thr Glu Ile Phe Tyr Gln Asp Leu Arg Arg Lys Thr Glu Phe Arg
245 250 255

Asn Tyr Pro Gly Ile Leu Tyr Cys Leu Leu Lys Ser Glu Lys Met Leu
260 265 270

Leu Glu Asp Val Lys Ala Asn Ile Thr Glu Met Leu Ala Gly Gly Val
275 280 285

Asn Thr Thr Ser Met Thr Leu Gln Trp His Leu Tyr Glu Met Ala Arg
290 295 300

Ser Leu Asn Val Gln Glu Met Leu Arg Glu Glu Val Leu Asn Ala Arg
305 310 315 320

ES 2 724 373 T3

Arg Gln Ala Glu Gly Asp Ile Ser Lys Met Leu Gln Met Val Pro Leu
 325 330 335

Leu Lys Ala Ser Ile Lys Glu Thr Leu Arg Leu His Pro Ile Ser Val
 340 345 350

Thr Leu Gln Arg Tyr Pro Glu Ser Asp Leu Val Leu Gln Asp Tyr Leu
 355 360 365

Ile Pro Ala Lys Thr Leu Val Gln Val Ala Ile Tyr Ala Met Gly Arg
 370 375 380

Asp Pro Ala Phe Phe Ser Ser Pro Asp Lys Phe Asp Pro Thr Arg Trp
 385 390 395 400

Leu Ser Lys Asp Lys Asp Leu Ile His Phe Arg Asn Leu Gly Phe Gly
 405 410 415

Trp Gly Val Arg Gln Cys Val Gly Arg Arg Ile Ala Glu Leu Glu Met
 420 425 430

Thr Leu Phe Leu Ile His Ile Leu Glu Asn Phe Lys Val Glu Met Gln
 435 440 445

His Ile Gly Asp Val Asp Thr Ile Phe Asn Leu Ile Leu Thr Pro Asp
 450 455 460

Lys Pro Ile Phe Leu Val Phe Arg Pro Phe Asn Gln Asp Pro Pro Gln
 465 470 475 480

Ala

<210> 5
 <211> 9745
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> plásmido

10

<400> 5

catggtctca cttttccact ttttgtcttg tccactaaaa cccttgattt ttcactgaa 60
 taaatgctac tattaggaca cataatatta aaagaaacct ccatctattht agttattht 120
 ttggtcactt ataactttaa cagatggggt ttttctgtgc aaccaattht aagggttht 180
 aatactthta aacacataca taccaacact tcaacgcacc ttcagcaac taaaataaaa 240
 atgacgttat ttctatatgt atcaagataa gaaagaacaa gttcaaaacc atcaaaaaa 300

ES 2 724 373 T3

gacacctttt caggtgcttt ttttatttta taaactcatt ccctgatctc gacttcgttc	360
tttttttacc tctcggttat gagttagttc aaattcgttc tttttagggtt ctaaactcgtg	420
tttttcttgg aattgtgctg ttttatcctt taccttgtct acaaaccctt taaaaacgtt	480
tttaaaggct ttttaagccgt ctgtacgttc ctttaagatca acgtgatata ggtttgctaa	540
cctttgcggt cacttaacta acttataggg gtaacactta aaaaagaatc aataacgata	600
gaaaccgctc ctaaagcagc tgcatttttt cctaacgaag aaggcaatag ttcacattta	660
ttgtctaaat gagaatggac tctagaagaa acttcgtttt taatcgtatt taaaacaatg	720
ggatgagatt caattatatg atttctcaag ataacagctt ctatatcaaa tgtattaagg	780
atattggtta atccaattcc gatataaaag ccaaagtttt gaagtgcatt taacatttct	840
acatcatttt tatttgccgcg ttccacaatc tcttttcgag aaatattctt ttcttcttta	900
gagagcgaag ccagtaacgc tttttcagaa gcatataatt cccaacagcc tcgatttcca	960
cagctgcatt tgggtccatt aaaatctatc gtcatatgac ccatttcccc agaaaaacc	1020
tgaacacctt tatacaattc gttgttaata acaagtcag ttccaattcc gatattaata	1080
ctgatgtaaa cgatgttttc atagtttttt gtcataccaa atactttttc accgtatgct	1140
cctgcattag cttcattttc acaaaaaacc ggaacattaa actcactctc aattaanaac	1200
tgcaaatctt tgatattcca atttaagtta ggcatgaaaa taatttgctg atgacgatct	1260
acaaggcctg gaacacaaat tcctattccg actagaccat aaggggactc aggcataatg	1320
gttacaaaac catgaataag tgcaataaaa atctctttta cttcactagc ggaagaacta	1380
gacaagtcag aagtcttctc gagaataata tttccttcta agtcggttag aattccgtta	1440
agatagtcga ctccctatac aataccaatc gagtagcctg cattcttatt aaaaacaagc	1500
attacaggtc ttctgcccgc tctagattgc cctgccccaa tttcaaaaat aaaatctttt	1560
tcaagcagtg tatttacttg agaggagaca gtagacttgt ttaatcctgt aatctcagag	1620
agagttgcc tggagacagc ggagttcttc aaaatttcat ctaatattaa tttttgattc	1680
atTTTTTTta ctaaagcttg atctgcaatt tgaataataa ccaactcctt gtttatccac	1740
cgaactaagt tgggtgttttt tgaagcttga attagatatt taaaagtatc atatctaata	1800
ttataactaa attttctaaa aaaaacattg aaataaacat taaattaata tatgatggaa	1860
ttgtagttag tttacaattc caacaaacta actcaattaa gctagctgat ggataaactt	1920
gttcacttaa ttaactagt aatcaagga ggtgaatata caatgattag cacaaaaaca	1980
cctcgccctt attctgaaat tcctagccct ggtgataatg gatggttaaa tttatatcat	2040
ttttggcgtg aaaaaggtag ccaacgcatt ctttttcgtc atattgaaaa ttttcaaaaa	2100
tatggacctt tttatcgcga aaaattagga aatttagaaa gcgtatatat tattcatcct	2160

ES 2 724 373 T3

gaagatgtag ctcatttatt taaatttgaa ggatcttata ctgaacgcta tgatattcct 2220
 ccttggttg cttatcatcg ttattatcaa aaacctattg gcgtattatt taaaaaatct 2280
 ggaacatgga aaaaagatcg tgtagtatta aatacagaag taatggctcc tgaagctatt 2340
 aaaaatttta ttccgttatt aaatcctgta tctcaagatt ttgtatctct ttacataaa 2400
 cgtattaaac aacaaggatc tggaaaaatt gttggagaca ttaaagaaga tttatttcat 2460
 tttgcgtttg aatctattac aaatgttatg tttggagaac gccttggaat gttagaagaa 2520
 acggtaaacc ctgaagctca aaaatttatt gatgctgtat ataaaatgtt tcatacatct 2580
 gtacctttat taaacgtacc toctgaactt tatcgtcttt ttogaacgaa aacgtggaca 2640
 gatcatgtag ctgcttggga tacaattttt aataaaagcg aaaaatacac ggaaattttt 2700
 tatcaagatt tacgtcgtaa aacagaattt cgtaattatc cgggaattct ttattgttta 2760
 cttaaaagcg aaaagatgtt attagaagat gtaaaagcta ataccacaga aatgtagca 2820
 ggtggagtag atacaacaag catgacatta caatggcacc tttatgaaat ggctcgcagc 2880
 ttaaaccgtac aagaaatgtt acgtgaagaa gtattaaacg ctgcgtcgtca agctgaagggt 2940
 gatatttcta aaatgttaca aatggttcca ttattaaaag cttctattaa agaaacgtta 3000
 cgtttacatc caattagcgt aacgcttcaa cgttatcctg aatctgattt agtattacaa 3060
 gattatctta ttctgctaa aacattagta caagtagcta tttatgctat gggacgtgat 3120
 cctgcttttt tttcttctcc tgataaattt gatcctacac gttggttata taaagataaa 3180
 gatcttattc attttcgcaa tcttggtatt ggatggggag tacgtcaatg tgtaggacgt 3240
 cgtattgctg aattagaaat gacgcttttt cttattcaca ttcttgaaaa ctttaaagtg 3300
 gaaatgcaac atattggaga tgtagacacg atttttaact taattcttac gcctgataaa 3360
 cctatttttt tagtatttctg tccttttaat caagatcctc ctcaagctta ataaacgcgt 3420
 ggtaccaa atcaaggagt aatatacaat gtctacacaa gaacaaacac ctcaaatttg 3480
 tgtagtagga tctggacctg ctggatttta tacagctcaa catcttttaa aacatcatc 3540
 tcgcgctcat gtagatattt atgaaaaaca acttgtaacct tttggattag ttcgttttgg 3600
 agtagctcct gatcatcctg aagtaaaaaa cgtaattaac acatttacac agacagctcg 3660
 ttctgatcgt tgtgcttttt atggaaatgt agaagtagga cgtgatgtaa cagtacaaga 3720
 acttcaagat gcttatcatg ctgtagtatt atcttatggt gctgaagatc atcaagcttt 3780
 agatattcca ggtgaagaat tacctgggtg attttctgct cgtgcttttg taggatggta 3840
 taatggatta cctgaaaatc gtgaattagc tctgattta tcttgtgata cagctgtaat 3900
 tttaggacaa ggcaacgtag ctttagatgt agctcgtatt ttattaacac ctccggatca 3960
 tttagaaaaa acggatatta cagaagctgc tcttgagct ttacgtcaat ctctgtgtaa 4020
 aacagtatgg attgtaggac gtcgtggacc tttacaagta gcttttacga ttaaagaact 4080

ES 2 724 373 T3

tcgCGAAAtg attcaattac ctggaacacg tcctatgtta gatcctgctg attttttagg	4140
acttcaggat cgtattaaag aagctgcacg tcctcgtaaa cgtttaatgg aattattatt	4200
acgtacagct acagaaaaac ctggtgtaga agaagctgct cgtagagcat ctgcttctcg	4260
tgcttgggga ttacgttttt ttcgtagccc tcaacaagta ttaccttctc ctgatggacg	4320
tcgtgctgct ggaattcgtt tagctgtaac acgtttagaa ggtattggag aagctacacg	4380
tgctgtacct acaggtgatg tagaagattt accttgtgga cttgtattaa gctctattgg	4440
atataaatct cgtcctattg atccttctgt accttttgat cctaaattag gtgtagtacc	4500
taatattgaa ggaagttag tagatgtacc tggattatat tgttctggat gggtaaaacg	4560
tggacctaca ggtgtaatta caacaacaat gacagatagc tttttaacag gccaaattct	4620
tttacaagat cttaaagctg gacatttacc ttcaggacct cgtcctggat ctgcttttat	4680
taaagcttta cttgattctc gtggagtatg gcctgtatct ttttctgatt gggaaaaatt	4740
agatgctgaa gaagtatcta gaggacaagc ttctggaaaa cctcgtgaaa aattattaga	4800
tcctcaagaa atgcttcgtt tacttggcca ctaataagag ctctgtacaa atcaaggagg	4860
tgaatataca atgtcttctt ctgaagataa aataacagtc cactttataa accgtgatgg	4920
tgaaacatta acaaccaaag gaaaaattgg tgactctctg ctagatggtg tggttcaaaa	4980
taatctagat attgatggtt ttggtgcatg tgagggaacc ttggcttgtt ctacctgtca	5040
cctcatcttt gaacagcaca tatttgagaa attggaagca atcactgatg aggagaatga	5100
catgcttgat ctggcatatg gactaacaga tagatcgcgg ttgggctgcc agatctgttt	5160
gacaaaggct atggacaata tgactgttcg agtaccatag catgcgcggc cgccatgccg	5220
gctaaacctc gcgaacggat tcaccgggtcc aagaattgga gctaattaat tcttgccggag	5280
aactgtgaat gcgcaaacca accttggca gaacatatcc atcgcgtccg ccatctccag	5340
cagccgcacg cggcgcactc cgggccgcgt tgctggcgtt tttccatagg ctccgcccc	5400
ctgacgagca tcacaaaaat cgacgtcaa gtcagaggtg gcgaaaccg acaggactat	5460
aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc	5520
cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct	5580
cacgctgtag gtatctcagt tcggtgtagg tcgttcgtc caagctgggc tgtgtgcacg	5640
aacccccgt tcagcccagc cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc	5700
cggtaagaca cgacttatcg cactggcag cagccactgg taacaggatt agcagagcga	5760
ggtatgtagg cggtgctaca gagttcttga agtggggcc taactacggc tacactagaa	5820
ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta	5880
gctcttgatc cggcaaaaa accaccgctg gtagcgggtg ttttttgtt tgcaagcagc	5940

ES 2 724 373 T3

agattacgcg cagaaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 6000
 acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 6060
 tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 6120
 agtaaaactg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 6180
 gtctatttcg ttcacccata gttgcctgac tccccgctgt gtagataact acgatacggg 6240
 agggcttacc atctggcccc agtgctgcaa tgataccgcg agaccacgc tcaccggctc 6300
 cagatttata agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 6360
 ctttatccgc ctccatccag totattaatt gttgccggga agctagagta agtagttcgc 6420
 cagttaatag tttgcgcaac gttgttgcca ttgctgcagg catcgtggtg tcacgctcgt 6480
 cgtttggtat ggcttcattc agctccgggt cccaacgatc aaggcgagtt acatgatccc 6540
 ccattgtgtg caaaaaagcg gtttagctct tcggctctcc gatcgttgtc agaagtaagt 6600
 tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 6660
 catccgtaag atgcttttct gtgactggtg agtactcaac caagtattc tgagaatagt 6720
 gtatgcggcg accgagttgc tcttgcccg cgtcaacacg ggataatacc gcgccacata 6780
 gcagaacttt aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 6840
 tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcacccaac tgatcttcag 6900
 catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 6960
 aaaaaggaat aaggcgaca cggaaatggt gaatactcat actcttcctt tttcaatatt 7020
 attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 7080
 aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct gacgtctaag 7140
 aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 7200
 ttcaagaatt cctgttataa aaaaaggatc aattttgaac tctctccca agttgatccc 7260
 ttaacgattt agaaatccct ttgagaatgt ttatatacat tcaaggtaac cagccaacta 7320
 atgacaatga ttcctgaaaa aagtaataac aaattactat acagataagt tgactgatca 7380
 acttccatag gtaacaacct ttgatcaagt aagggtatgg ataataaacc acctacaatt 7440
 gcaatacctg ttccctctga taaaaagctg gtaaagttaa gcaaactcat tccagcacca 7500
 gcttctgct gtttcaagct acttgaaca attggtgata taactgtttt ggtgaacgaa 7560
 agcccaccta aaacaaatac gattataatt gtcatgaacc atgatgttgt ttctaaaaga 7620
 aaggaagcag ttaaaaagct aacagaaaga aatgtaactc cgatgtttaa cacgtataaa 7680
 ggacctcttc tatcaacaag tatcccacca atgtagccga aaataatgac actcattggt 7740
 ccagggaaaa taattacact tccgatttcg gcagtaacta gctggtgaac atctttcatc 7800
 atataaggaa ccatagagac aaaccctgct actgttccaa atataattcc cccacaaaaga 7860

ES 2 724 373 T3

actccaatca taaaaggtat atttttccct aatccgggat caacaaaag atctgttact 7920
 ttcctgatat gttttacaaa tatcaggaat gacagcacgc taacgataag aaaagaaatg 7980
 ctatatgatg ttgtaaacaa cataaaaaat acaatgccta cagacattag tataattcct 8040
 ttgatataca aatgaccttt tatccttact tctttcttta ataatttcat aagaaacgga 8100
 acagtgataa ttgttatcat aggaatgagt agaagatagg accaatgaat ataatgggct 8160
 atcatccac caatcgctgg accgactcct tctccatgg ctactatcga tccaataaga 8220
 ccaaatgctt taccctatt ttcctttgga atatagcgcg caactacaac cattacgagt 8280
 gctgaaatg cagctgcacc agccccttga ataaaacgag ccataataag taaggaaaag 8340
 aaagaatggc caacaaacc aattaccgac ccgaaacaat ttattataat tccaaatagg 8400
 agtaaccttt tgatgcctaa ttgatcagat agctttccat atacagctgt tccaatggaa 8460
 aaggtaaca taaaggctgt gttcaccag tttgtactcg caggtggttt attaaaatca 8520
 tttgcaatat caggtaatga gacgttcaaa accatttcat ttaatacgt aaaaaagat 8580
 aaaatgcaaa gccaaattaa aatttggtt tgctgtaa atcgattgtga ataggatgta 8640
 ttcacatttc accctccaat aatgagggca gacgtagttt atagggttaa tgatagcgtt 8700
 ccctctttta attgaaccct gttacattca ttacacttca taattaattc ctccctaaact 8760
 tgattaaaac attttaccac atataaacta agttttaaat tcagtatttc atcacttata 8820
 caacaatatg gcccgtttgt tgaactactc ttaataaaa taatttttcc gttcccaatt 8880
 ccacattgca ataatagaaa atccatcttc atcggctttt tcgtcatcat ctgtatgaat 8940
 caaatcgctt tcttctgtgt catcaagggt taatttttta tgtatttctt ttaacaaacc 9000
 accataggag attaaccttt tacggtgtaa accttcctcc aatcagaca aacgtttcaa 9060
 attcttttct tcatcatcgg tcataaaatc cgtatccttt acaggatatt ttgcagtttc 9120
 gtcaattgcc gattgtatat ccgatttata tttatttttc ggtcgaatca tttgaacttt 9180
 tacatttggg tcatagtcta atttcattgc ctttttccaa aattgaatcc attgtttttg 9240
 attcacgtag ttttctgtat tcttaaaata agttggttcc acacatacca atacatgcat 9300
 gtgctgatta taagaattat ctttattatt tattgtcact tccggtgcac gcataaaacc 9360
 aacaagattt ttattaattt ttttatattg catcattcgg cgaatcctt gagccatatac 9420
 tgacaaactc ttatttaatt cttcgccatc ataaacattt ttaactgtta atgtgagaaa 9480
 caaccaacga actggtggct tttgtttaat aacttcagca acaacctttt gtgactgaat 9540
 gccatgtttc attgctctcc tccagttgca cattggacaa agcctggatt taaaaacca 9600
 cactcgatac aactttcttt cgcctgtttc acgattttgt ttatactcta atatttcagc 9660
 acaatctttt actctttcag cttttttaa ttcaagaata tgcagaagtt caaagtaatc 9720
 aacattagcg attttctttt ctctc 9745

5 <210> 6
 <211> 481
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> Secuencia de aminoácidos de CYP11A1 mutada

ES 2 724 373 T3

<400> 6

Ala Ser Thr Lys Thr Pro Arg Pro Tyr Ser Glu Ile Pro Ser Pro Gly
1 5 10 15

Asp Asn Gly Trp Leu Asn Leu Tyr His Phe Trp Arg Glu Lys Gly Ser
20 25 30

Gln Arg Ile His Phe Arg His Ile Glu Asn Phe Gln Lys Tyr Gly Pro
35 40 45

Ile Tyr Arg Glu Lys Leu Gly Asn Leu Glu Ser Val Tyr Ile Ile His
50 55 60

Pro Glu Asp Val Ala His Leu Phe Lys Phe Glu Gly Ser Tyr Pro Glu
65 70 75 80

Arg Tyr Asp Ile Pro Pro Trp Leu Ala Tyr His Arg Tyr Tyr Gln Lys
85 90 95

Pro Ile Gly Val Leu Phe Lys Lys Ser Gly Thr Trp Lys Lys Asp Arg
100 105 110

Val Val Leu Asn Thr Glu Val Met Ala Pro Glu Ala Ile Lys Asn Phe
115 120 125

Ile Pro Leu Leu Asn Pro Val Ser Gln Asp Phe Val Ser Leu Leu His
130 135 140

Lys Arg Ile Lys Gln Gln Gly Ser Gly Glu Phe Val Gly Asp Ile Lys
145 150 155 160

Glu Asp Leu Phe His Phe Ala Phe Glu Ser Ile Thr Asn Val Met Phe
165 170 175

Gly Glu Arg Leu Gly Met Leu Glu Glu Thr Val Asn Pro Glu Ala Gln
180 185 190

Lys Phe Ile Asp Ala Val Tyr Lys Met Phe His Thr Ser Val Pro Leu
195 200 205

ES 2 724 373 T3

Leu Asn Val Pro Pro Glu Leu Tyr Arg Leu Phe Arg Thr Lys Thr Trp
 210 215 220
 Arg Asp His Val Ala Ala Trp Asp Thr Ile Phe Asn Lys Ala Glu Lys
 225 230 235 240
 Tyr Thr Glu Ile Phe Tyr Gln Asp Leu Arg Arg Lys Thr Glu Phe Arg
 245 250 255
 Asn Tyr Pro Gly Ile Leu Tyr Cys Leu Leu Lys Ser Glu Lys Met Leu
 260 265 270
 Leu Glu Asp Val Lys Ala Asn Ile Thr Glu Met Leu Ala Gly Gly Val
 275 280 285
 Asn Thr Thr Ser Met Thr Leu Gln Trp His Leu Tyr Glu Met Ala Arg
 290 295 300
 Ser Leu Asn Val Gln Glu Met Leu Arg Glu Glu Val Leu Asn Ala Arg
 305 310 315 320
 Arg Gln Ala Glu Gly Asp Ile Ser Lys Met Leu Gln Met Val Pro Leu
 325 330 335
 Leu Lys Ala Ser Ile Lys Glu Thr Leu Arg Leu His Pro Ile Ser Val
 340 345 350
 Thr Leu Gln Arg Tyr Pro Glu Ser Asp Leu Val Leu Gln Asp Tyr Leu
 355 360 365
 Ile Pro Ala Lys Thr Leu Val Gln Val Ala Ile Tyr Ala Met Gly Arg
 370 375 380
 Asp Pro Ala Phe Phe Ser Ser Pro Asp Lys Phe Asp Pro Thr Arg Trp
 385 390 395 400
 Leu Ser Lys Asp Lys Asp Leu Ile His Phe Arg Asn Leu Gly Phe Gly
 405 410 415
 Trp Gly Val Arg Gln Cys Val Gly Arg Arg Ile Ala Glu Leu Glu Met
 420 425 430
 Thr Leu Phe Leu Ile His Ile Leu Glu Asn Phe Lys Val Glu Met Gln
 435 440 445
 His Ile Gly Asp Val Asp Thr Ile Phe Asn Leu Ile Leu Thr Pro Asp
 450 455 460
 Lys Pro Ile Phe Leu Val Phe Arg Pro Phe Asn Gln Asp Pro Pro Gln
 465 470 475 480

Ala

5 <210> 7
 <211> 461

ES 2 724 373 T3

<212> PRT
 <213> Bos taurus

5 <400> 7

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

ES 2 724 373 T3

180 185 190

Val Trp Ile Val Gly Arg Arg Gly Pro Leu Gln Val Ala Phe Thr Ile
195 200 205

Lys Glu Leu Arg Glu Met Ile Gln Leu Pro Gly Thr Arg Pro Met Leu
210 215 220

Asp Pro Ala Asp Phe Leu Gly Leu Gln Asp Arg Ile Lys Glu Ala Ala
225 230 235 240

Arg Pro Arg Lys Arg Leu Met Glu Leu Leu Leu Arg Thr Ala Thr Glu
245 250 255

Lys Pro Gly Val Glu Glu Ala Ala Arg Arg Ala Ser Ala Ser Arg Ala
260 265 270

Trp Gly Leu Arg Phe Phe Arg Ser Pro Gln Gln Val Leu Pro Ser Pro
275 280 285

Asp Gly Arg Arg Ala Ala Gly Ile Arg Leu Ala Val Thr Arg Leu Glu
290 295 300

Gly Ile Gly Glu Ala Thr Arg Ala Val Pro Thr Gly Asp Val Glu Asp
305 310 315 320

Leu Pro Cys Gly Leu Val Leu Ser Ser Ile Gly Tyr Lys Ser Arg Pro
325 330 335

Ile Asp Pro Ser Val Pro Phe Asp Pro Lys Leu Gly Val Val Pro Asn
340 345 350

Met Glu Gly Arg Val Val Asp Val Pro Gly Leu Tyr Cys Ser Gly Trp
355 360 365

Val Lys Arg Gly Pro Thr Gly Val Ile Thr Thr Thr Met Thr Asp Ser
370 375 380

Phe Leu Thr Gly Gln Ile Leu Leu Gln Asp Leu Lys Ala Gly His Leu
385 390 395 400

Pro Ser Gly Pro Arg Pro Gly Ser Ala Phe Ile Lys Ala Leu Leu Asp
405 410 415

Ser Arg Gly Val Trp Pro Val Ser Phe Ser Asp Trp Glu Lys Leu Asp
420 425 430

Ala Glu Glu Val Ser Arg Gly Gln Ala Ser Gly Lys Pro Arg Glu Lys
435 440 445

Leu Leu Asp Pro Gln Glu Met Leu Arg Leu Leu Gly His
450 455 460

5 <210> 8
<211> 109
<212> PRT

ES 2 724 373 T3

<213> Bos taurus

<400> 8

Met Ser Ser Ser Glu Asp Lys Ile Thr Val His Phe Ile Asn Arg Asp
 1 5 10 15

Gly Glu Thr Leu Thr Thr Lys Gly Lys Ile Gly Asp Ser Leu Leu Asp
 20 25 30

Val Val Val Gln Asn Asn Leu Asp Ile Asp Gly Phe Gly Ala Cys Glu
 35 40 45

Gly Thr Leu Ala Cys Ser Thr Cys His Leu Ile Phe Glu Gln His Ile
 50 55 60

Phe Glu Lys Leu Glu Ala Ile Thr Asp Glu Glu Asn Asp Met Leu Asp
 65 70 75 80

Leu Ala Tyr Gly Leu Thr Asp Arg Ser Arg Leu Gly Cys Gln Ile Cys
 85 90 95

Leu Thr Lys Ala Met Asp Asn Met Thr Val Arg Val Pro
 100 105

5

REIVINDICACIONES

- 5 1. Una monooxigenasa del citocromo P450 aislada que comprende o consiste en una secuencia de aminoácidos al menos el 80% idéntica a la SEQ ID NO: 1, en la que dicha secuencia comprende una treonina en la posición correspondiente a la posición 225 de la SEQ ID NO: 1 y un ácido aspártico en la posición correspondiente a la posición 289 de SEQ ID NO: 1.
2. La enzima según la reivindicación 1, que consiste en la secuencia SEQ ID NO: 1.
- 10 3. Un ácido nucleico aislado que comprende o consiste en una secuencia de nucleótidos que codifica una enzima tal como se define en las reivindicaciones 1 o 2.
4. Un vector que comprende un ácido nucleico tal como se define en la reivindicación 3 que está asociado operativamente con secuencias de control de la expresión.
- 15 5. Vector según la reivindicación 4, que además comprende una secuencia de ácidos nucleicos que codifica adrenodoxina (Adx) y/o una secuencia de ácidos nucleicos que codifica adrenodoxina reductasa (AdR).
- 20 6. Una célula huésped que contiene un ácido nucleico tal como se define en la reivindicación 3 o un vector tal como se define en las reivindicaciones 4 o 5.
7. La célula huésped según la reivindicación 6, que es un microorganismo manipulado genéticamente.
- 25 8. La célula huésped según la reivindicación 6 o 7, que es *Saccharomyces cerevisiae*.
9. Un microorganismo manipulado genéticamente capaz de convertir un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo de colesterol y un derivado de colesterol en un precursor de hormonas esteroideas, comprendiendo dicho microorganismo un ácido nucleico tal como se define en la reivindicación 3, y opcionalmente una secuencia de ácidos nucleicos que codifica adrenodoxina (Adx) y/o una secuencia de ácidos nucleicos que codifica adrenodoxina reductasa (AdR).
- 30 10. Un procedimiento *in vitro* para preparar una enzima tal como se define en las reivindicaciones 1 o 2, comprendiendo dicho procedimiento las etapas siguientes:
- 35 a) cultivar una célula huésped tal como se define en una cualquiera de las reivindicaciones 6 a 8 en condiciones adecuadas para obtener la expresión de la enzima tal como se define en las reivindicaciones 1 o 2 y
- 40 b) recuperar la enzima expresada.
11. El uso de una enzima tal como se define en las reivindicaciones 1 o 2 para producir un precursor de hormonas esteroideas.
- 45 12. Un procedimiento para producir un precursor de hormonas esteroideas, que comprende las etapas siguientes:
- a) proporcionar un microorganismo tal como se define en una cualquiera de las reivindicaciones 7 a 9,
- b) cultivar dicho microorganismo en condiciones que permitan la expresión de una enzima tal como se define en las reivindicaciones 1 o 2,
- 50 c) poner en contacto el cultivo de microorganismos obtenido en la etapa b) con un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo de colesterol y un derivado de colesterol en condiciones que permitan la producción por el microorganismo de un precursor de hormonas esteroideas a partir de dicho sustrato y
- 55 d) recuperar el precursor de hormonas esteroideas producido.
13. Un procedimiento para producir un precursor de hormonas esteroideas, que comprende las etapas siguientes:
- 60 a) poner en contacto una enzima tal como se define en las reivindicaciones 1 o 2 con un polipéptido adrenodoxina (Adx) aislado, un polipéptido adrenodoxina reductasa (AdR) y un sustrato seleccionado del grupo que consiste en monoalcoholes policíclicos e insaturados que tienen una cadena lateral alifática tales como colesterol, un análogo del colesterol y un derivado del colesterol en condiciones que permiten la transformación de dicho sustrato en un precursor de hormonas esteroideas y
- 65 b) recuperar el precursor de hormonas esteroideas obtenido.

14. El procedimiento según la reivindicación 12 o 13, en el que el precursor de hormonas esteroideas es pregnenolona.

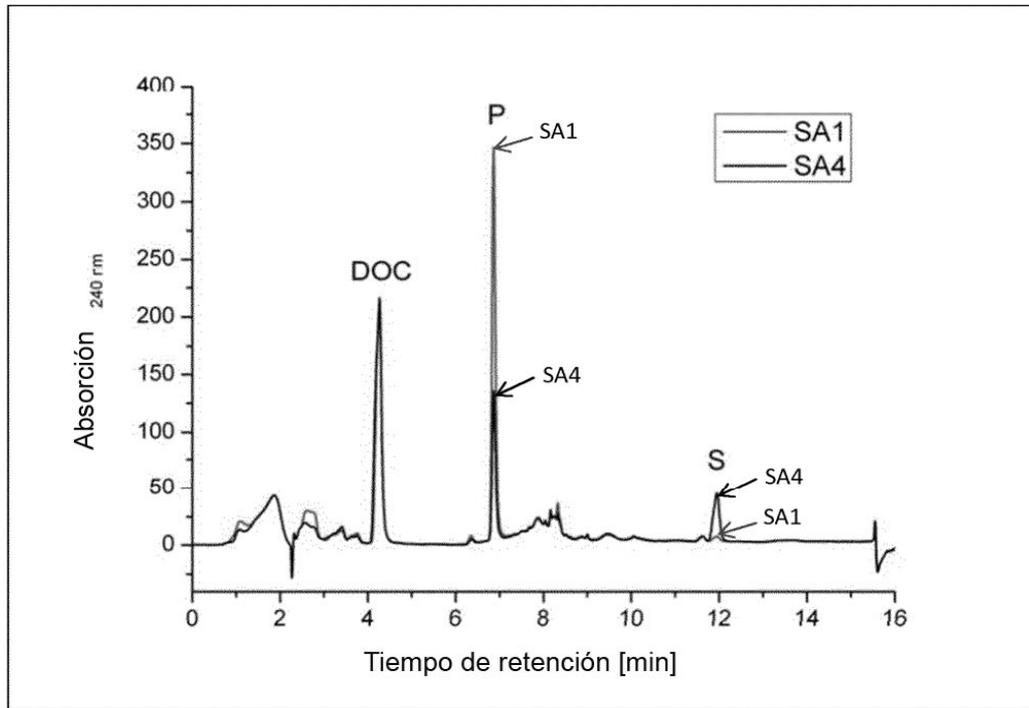


FIG.1

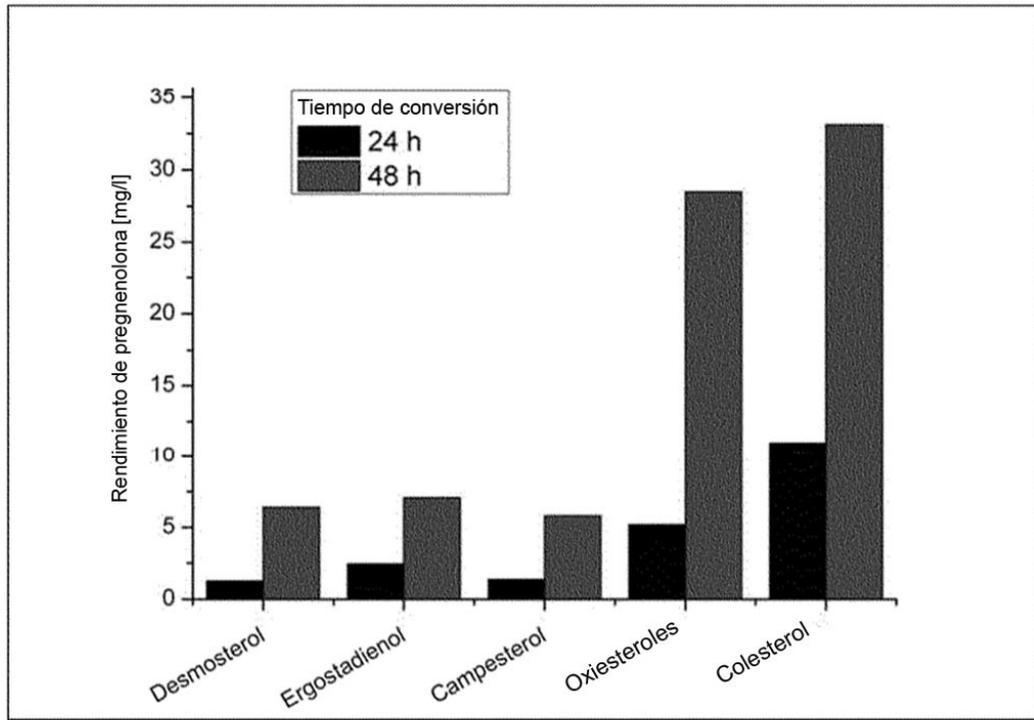


FIG.2

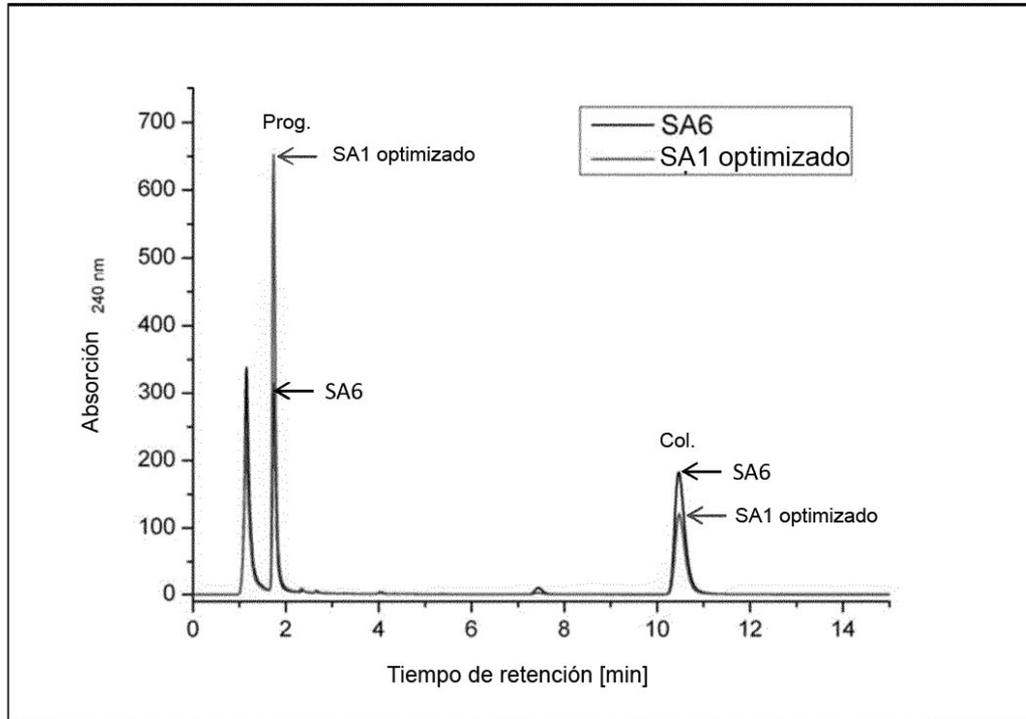


FIG.3

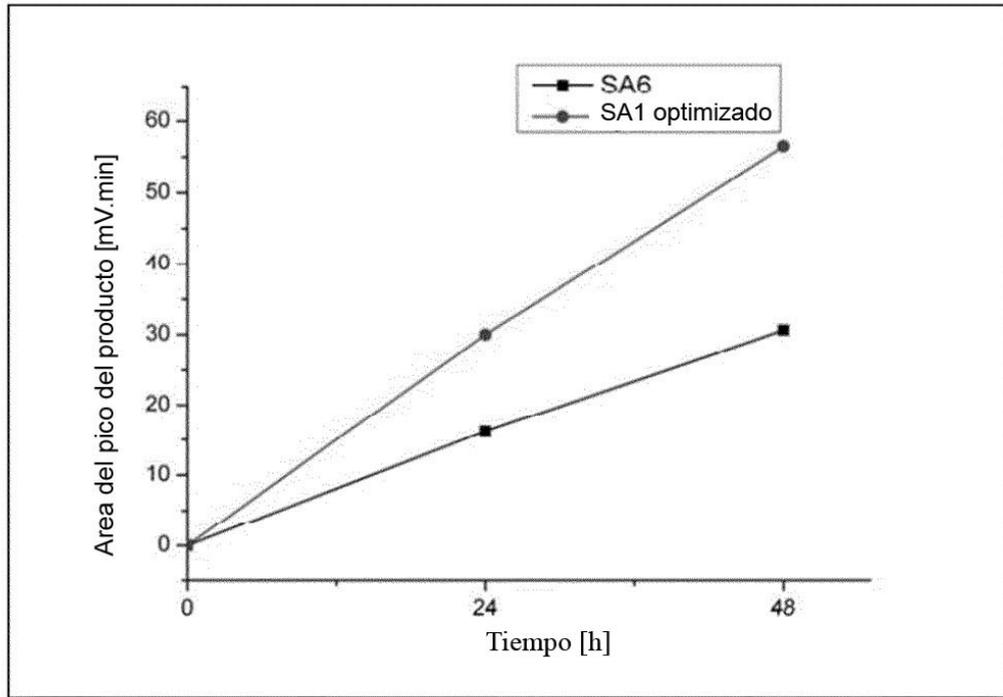


FIG.4

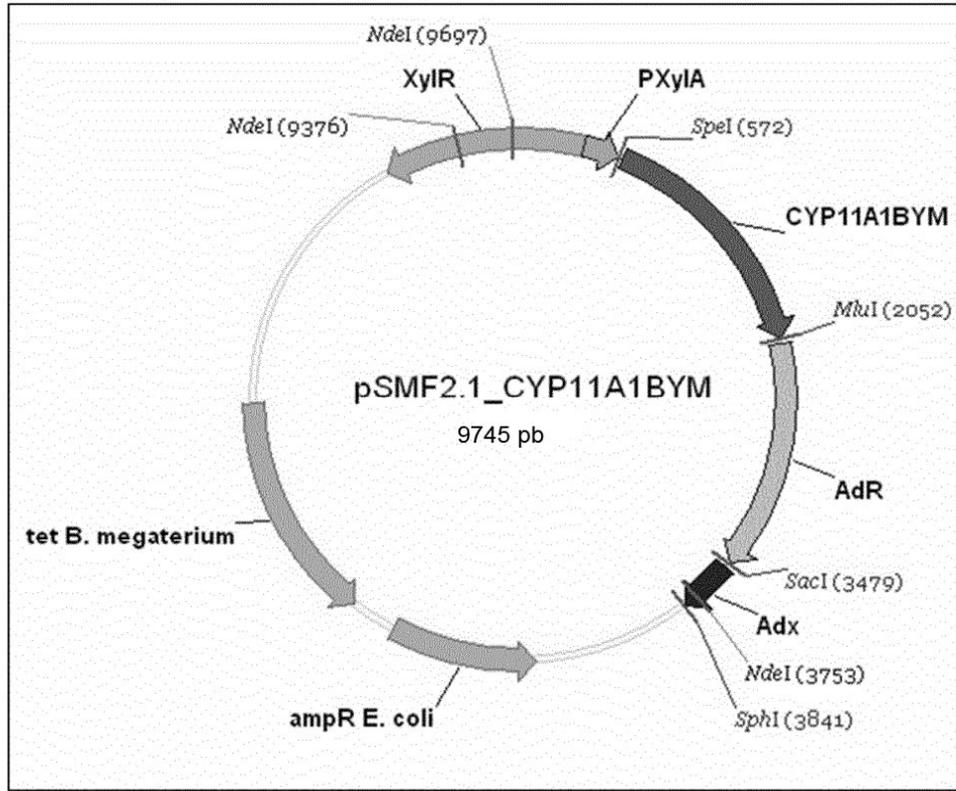


FIG.5