

NOVEL CYTOCHROME P450 POLYPEPTIDE WITH INCREASED ENZYMATIC ACTIVITY

The present invention concerns cytochrome P450 monooxygenases with increased enzymatic activity.

Mammalian mitochondrial P450s constitute a small family of cytochromes that perform specific reactions inside mitochondria and play an important role in the metabolism of a variety of hydrophobic compounds. For example, P450c11beta and P450c11AS (encoded by the CYP11B1 and CYP11B2 genes) perform the final step of glucocorticoid and mineralocorticoid biosynthesis, respectively. P450c11beta is a classical steroid 11beta-hydroxylase that converts 11deoxycortisol into hydrocortisone, while P450c11AS catalyses the transformation of deoxycorticosterone into aldosterone in three consecutive steps. CYP27A1 is another example of an inner mitochondrial P450 which is a sterol 27-hydroxylase and vitamin D25 hydroxylase. Finally, the most emblematic mitochondrial P450 is P450scc (Side Chain Cleaving enzyme), encoded by the CYP11A1 gene, which cleaves the cholesterol side chain thus transforming cholesterol into pregnenolone by two consecutive hydroxylations and a final cleaving. P450scc performs the first key step of steroid biosynthesis by transforming a sterol into steroid.

The P450 ability to catalyze the regio-, chemo- and stereospecific oxidation of a vast number of substrates reflects their biological roles and makes them important candidates for biotechnological applications. Particularly, steroid hormones are widely used as anti-inflammatory, contraceptive and antiproliferative drugs. In mammals, the synthesis of these steroids starts with the side-chain cleaving reaction of cholesterol to pregnenolone. Pregnenolone serves as a basis for the production of further steroid hormones such as hydrocortisone and great interests are associated with its industrial large scale conversion from low-priced substrates such as cholesterol and its plant-derived analogues. However, side-chain cleaving reaction of cholesterol to pregnenolone is a limiting step in steroids overall process. In mammals, it is catalyzed by the membrane-bound CYP11A1 enzyme.

However, for multiple reasons, biotechnological and industrial use of the P450s is difficult to implement and is not satisfactory. Mitochondrial P450 use a specific chain of electron transporters made of two proteins, namely ferredoxin reductase (FdxR) and ferredoxin (Fdx1), also named adrenodoxin reductase (AdR) and adrenodoxin (Adx). In the natural situation *in vivo*, FdxR gains electron from NADPH and Fdx1 shuttles electron from FdxR to the mitochondrial P450. An *in vitro* system has been developed where FdxR is at a catalytic concentration (0.5 μ M) and Fdx1 at a saturation concentration (10 μ M). In

order for this *in vitro* system to be efficient, electrons must be appropriately fluxed to the P450.

To try to overcome these difficulties, some authors fused the three peptides P450scc (or P450c11beta), Fdx1 and FdxR using variable hinge or linker sequences. Even if this triple fusion results in a functional protein, its efficacy is low compared to the “bona fide” polypeptide and it cannot be used at industrial scale. Moreover, the mitochondrial environment is difficult to mimic in a microbial recombinant system. For example, the P450scc polypeptide can be properly targeted to yeast mitochondria. However, the targeted polypeptide cannot convert sterol into pregnenolone due to absence of substrate in the mitochondria and/or to improper folding or targeting of P450scc.

Different recombinant systems are used to produce biosynthetic pregnenolone from plant sterol. For instance, a biosynthesis system was developed in the yeast *Saccharomyces cerevisiae*. In this system, P450scc is targeted outside the mitochondria at the plasma membrane together with FdxR and Fdx1, and the sterol pathway is routed to produce sterol at the membrane. Also, P450c11beta may be targeted to mitochondria together with Fdx and FdxR1 where it may convert 11-deoxycortisol into cortisol. A bioconversion system was developed in *Bacillus megaterium*. In this system, cholesterol can be converted into pregnenolone by mature forms of P450scc, Fdx1 and FdxR1. Biosynthesis remains the most attractive technology at the industrial scale since the substrate is made as a soluble molecule from a simple carbon source by the host therefore avoiding the burden of using detergents.

Recently, efforts have been made to produce new P450scc polypeptides with improved properties. For instance, a recombinant P450scc mutant bearing a K193E mutation was created. This mutant is more soluble and thus less prone to aggregation than the recombinant wild-type polypeptide. Therefore, a higher expression level is obtained with this mutant without changing its enzymatic characteristics. However, there is still a need for improved P450scc polypeptides with increased enzymatic activity allowing optimized production of steroid hormones, particularly suitable for industrial processes.

The inventors have developed a new P450scc polypeptide bearing specific mutations. This mutant shows unexpectedly an improved enzymatic activity in terms of substrate conversion into steroid hormones.

The present invention thus concerns a new isolated P450 enzyme comprising or consisting of an amino acid sequence at least 80% identical to SEQ ID NO: 1, wherein said sequence comprises a threonine at position corresponding to position 225 and/or an aspartic acid mutation at position corresponding to position 289. In some embodiments, the new P450 enzyme is at least 85%, 90%, 95%, 98% or 99% identical to SEQ ID NO:1.

The invention also concerns an isolated nucleic acid which comprises or consists of a nucleotide sequence encoding said enzyme, a vector comprising said nucleic acid which is operatively associated with expression control sequences, and a host cell containing said nucleic acid or said vector, said host cell being, for example, a microorganism.

5 In one embodiment, the invention provides a genetically engineered microorganism capable of converting a substrate into a steroid hormone precursor. The substrate can be, for example, a polycyclic, unsaturated mono alcohol having an aliphatic side chain, such as cholesterol, a cholesterol analogue or a cholesterol derivative. The genetically engineered microorganism comprises, for example, a nucleic acid encoding a new P450
10 enzyme featured in the invention, and optionally a nucleic acid sequence encoding an adrenodoxin (Adx) and/or a nucleic acid sequence encoding an adrenodoxin reductase (AdR).

In another aspect, the invention provides an *in vitro* method for preparing a P450 enzyme featured in the invention, said method comprising the steps of:

- 15 a) culturing a host cell under conditions suitable to obtain expression of the P450 enzyme; and
b) recovering the expressed enzyme.

20 In another embodiment, a P450 enzyme featured in the invention is used for producing a steroid hormone precursor.

The invention further relates to a method for producing a steroid hormone precursor, comprising the steps of:

- 25 a) providing a microorganism expressing a P450 enzyme featured in the invention,
b) culturing said microorganism under conditions allowing the expression of the P450 enzyme,
c) contacting the microorganism culture obtained at step b) with a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol
30 derivative, in conditions allowing the production, by the microorganism, of a steroid hormone precursor from said substrate, and
d) recovering the steroid hormone precursor produced.

35 The invention also pertains to a method, such as an *in vitro* method, for producing a steroid hormone precursor, comprising the steps of:

- a) contacting a P450 enzyme featured in the invention with an isolated adrenodoxin (Adx) polypeptide, an isolated adrenodoxin reductase (AdR) polypeptide and a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, cholesterol analogues and derivatives in conditions allowing the transformation of said substrate into a steroid hormone precursor, and
- b) recovering the steroid hormone precursor obtained.

Steroid hormones are widely used as anti-inflammatory, contraceptive and antiproliferative drugs. The invention is of utility to produce a steroid hormone precursor and aid the synthesis of steroid hormones for use in humans. Steroid hormone precursor, such as pregnenolone, serves as a basis for the production of further steroid hormones such as hydrocortisone and great interests are associated with its industrial large scale conversion.

DESCRIPTION OF THE INVENTION

Enzymes

By an "isolated" enzyme or polypeptide, it is intended that the enzyme or polypeptide is no longer in its natural environment within the organism in which it is originally expressed. When referring to an enzyme or polypeptide, "purified" means that the indicated molecule is present in the substantial absence of other biological macromolecules of the same type. The term "purified" as used herein means at least 75% by weight, at least 85% by weight, at least 95% by weight, or at least 98% by weight, of biological macromolecules of the same type are present.

The term "cytochrome P450 enzyme" or "P450 enzyme" refers to a monooxygenase which is capable of catalyzing certain reactions such as those reviewed in Van Bogaert *et al*, 2011, FEBS J. 278(2): 206-221 or in Urlacher and Girhard, 2011, Trends in Biotechnology 30(1): 26-36, or at the website dmelson.uthsc.edu/CytochromeP450.html. As an example of such a wild-type P450 enzyme, the sequence SEQ ID NO: 4 represents the mature form of *Bos taurus* CYP11A1. The complete sequence of *Bos taurus* CYP11A1 including a transit peptide is referenced as P00189 in UniProtKB/Swiss-Prot database (last sequence update: July 21, 1986).

The cytochrome P450 enzymes featured in the invention may be membrane-bound (insoluble) or cytoplasmic (soluble) in their respective original hosts. For example, P450scc (SEQ ID NO:4) catalyzes the side-chain cleavage reaction of polycyclic and

unsaturated mono alcohols having an aliphatic side chain such as cholesterol, cholesterol analogues and derivatives thereof into pregnenolone or other steroid hormone precursors and derivatives, as a non-limiting example.

The inventors have developed a new P450 enzyme of sequence SEQ ID NO: 1 with improved enzymatic activity. In comparison with SEQ ID NO: 4 (wild-type scc enzyme with R at position 225 and N at position 289), the amino acid sequence SEQ ID NO: 1 comprises two mutations: an arginine to threonine mutation at position corresponding to position 225 of the sequence SEQ ID NO: 4 and an asparagine to aspartic acid mutation at position corresponding to position 289 of the sequence SEQ ID NO: 4.

In a specific embodiment, the new P450 enzyme featured in the invention displays a monooxygenase activity catalyzing the transformation of cholesterol into pregnenolone. In a particular embodiment, the new P450 enzyme displays at least 80% or more, in particular at least 85%, 90%, 95%, 100%, 125%, 150%, 175%, 200%, 225%, 250%, 275%, more particularly at least 300% or more of the monooxygenase activity of the P450 enzyme of sequence SEQ ID NO: 4.

The isolated P450 enzyme featured in the invention comprises or consists of an amino acid sequence at least 80% identical to SEQ ID NO: 1, wherein said sequence comprises a threonine at position corresponding to position 225 and/or an aspartic acid at position corresponding to position 289. In some embodiments, the new P450 enzyme is at least 85%, 90%, 95%, 98% or 99% identical to SEQ ID NO:1.

Variations in amino acid sequence may be introduced by substitution, deletion or insertion of one or more codons into the nucleic acid sequence encoding the enzyme that results in a change in the amino acid sequence of the enzyme. Amino acid substitutions may be conservative or non-conservative. In some embodiments, substitutions are conservative substitutions, in which one amino acid is substituted for another amino acid with similar structural and/or chemical properties.

A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain R group with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. Examples of groups of amino acids that have side chains with similar chemical properties include 1) aliphatic side chains: glycine, alanine, valine, leucine, and isoleucine; 2) aliphatic-hydroxyl side chains: serine and threonine; 3) amide-containing side chains: asparagine and glutamine; 4) aromatic side chains: phenylalanine, tyrosine, and tryptophan; 5) basic side chains: lysine, arginine, and histidine; 6) acidic side chains: aspartic acid and glutamic acid; and 7) sulfur-containing side chains: cysteine and methionine. Conservative amino

acid substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, glutamate-aspartate, and asparagine-glutamine.

For example, an amino acid sequence at least 80% identical to SEQ ID NO: 1 may be a polypeptide having at least one substitution at a particular amino acid residue. In some embodiments, an amino acid sequence at least 80% identical to SEQ ID NO: 1 has at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% amino acid sequence identity to the sequence SEQ ID NO: 1. Amino acid sequence identity is defined as the percentage of amino acid residues in the sequence at least 80% identical to SEQ ID NO: 1 that are identical with the amino acid residues in the reference sequence SEQ ID NO: 1, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percentage of sequence identity, and not considering any conservative substitutions as part of the sequence identity. Sequence identity may be determined over the full length of the at least 80% identical sequence, the full length of the reference sequence, or both. The percentage of identity for amino acid sequences may be calculated by performing a pairwise global alignment based on the Needleman-Wunsch alignment algorithm to find the optimum alignment (including gaps) of two sequences along their entire length, for instance using Needle, and using the BLOSUM62 matrix with a gap opening penalty of 10 and a gap extension penalty of 0.5.

In the P450 enzyme comprising or consisting of an amino acid sequence at least 80% identical to SEQ ID NO: 1, the amino acid modifications as compared with SEQ ID NO: 1 are typically located at positions such that they do not significantly undermine the biological activity of the enzyme. Indeed, the cytochrome P450 amino acid sequence presenting at least 80% identical to SEQ ID NO: 1 exhibits at least the same biological activity as the polypeptide of sequence SEQ ID NO: 1.

A "same biological activity" may denote a same biological function. Therefore, a polypeptide having a same biological activity as the polypeptide of sequence SEQ ID NO: 1 may for instance be a polypeptide having monooxygenase activity. Techniques to determine the monooxygenase activity of an enzyme are well known from the skilled person. For instance, a polypeptide having a same biological activity as the polypeptide of sequence SEQ ID NO: 1 may be a polypeptide able to catalyse the side-chain cleavage reaction of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, cholesterol analogues and derivatives to pregnenolone or other steroid hormone precursors and derivatives, as a non-limiting example. In this case, the catalysing activity of a compound can easily be evaluated *in vitro* (as shown in 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) or *in vivo* (as shown in Duport, C., R. Spagnoli, et al. (1998) "Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast.”; Nat Biotechnol **16**(2): 186-9) by the person skilled in the art, for example by

5 means of the protocol described in Example 1.

Typically, catalysing activity may be tested in an *in vitro* conversion assay wherein 150 mM of HEPES buffer, adjusted to pH 7.4, containing 0.05 % Tween-20 and 1 mM MgCl₂, may be applied as reaction buffer. 1 unit of glucose-6-phosphate dehydrogenase, with 5 mM glucose-6-phosphate as substrate, may serve as a NADPH-regenerating

10 system. Typically, the concentrations of CYP11A1, Adx and AdR may be of 1 µM, 20 µM and 0.5 µM, or of 0.25 µM, 5 µM and 0.125 µM, respectively. 20 µM of cholesterol, dissolved in 45% 2-hydroxypropyl-β-cyclodextrin, may serve as substrate for CYP11A1. Typically, the samples are pre-warmed to 37°C and the reaction was started by addition of NADPH to a final concentration of 100 µM. The mixtures may be incubated at 37°C with

15 agitation for 30 s. The reaction may be stopped by boiling the samples in water for 30 s. To allow photometric detection at 240 nm, the steroids may be converted into their 3-keto-Δ⁴ derivatives using a cholesterol oxidase from *Nocardia spec.* Typically, 20 µl of a cholesterol oxidase solution (5 mg cholesterol oxidase and 5 mg Na-cholate dissolved in 5 ml of 50 mM HEPES buffer pH 7, containing 0.05 % Tween-20) is added to the samples.

20 After incubation at 37°C for 1 h, 11-deoxycorticosterone (DOC) may be added to the reaction mixtures as an internal standard, followed by a 2-times extraction with equal volumes of ethylacetate. After evaporation, the extracts may be dissolved in acetonitrile/water.

Catalysing activity may also be tested in an *in vivo* conversion assay wherein the

25 conversion of 300 µM cholesterol into progesterone is typically evaluated by HPLC after 24 h. *Bacillus megaterium* may be cultivated in TB-medium containing 10 µg/ml tetracycline at 37°C with 180 rpm shaking. Protein expression may be induced by adding 0.25 g of xylose dissolved in 1 mL water, followed by the subsequent addition of the substrate, dissolved in 2-hydroxypropyl-β-cyclodextrin.

Alignment of the sequence SEQ ID NO: 1 with an amino acid sequence at least

30 80% identical to SEQ ID NO: 1 may be performed. As these two sequences have a high percentage of identity, most of their amino acids are identical. Therefore, amino acid numbering may be defined for both sequences in such a way that at least 80% of the amino acids present at positions bearing the same number in both sequences are

35 identical (by introducing gaps, if necessary, to achieve the maximum percentage of sequence identity). In this context, “a position corresponding to position X” of the

sequence SEQ ID NO: 1 denotes the position bearing said number X in the sequence at least 80% identical to SEQ ID NO: 1. The sequence SEQ ID NO: 1 may therefore be taken as a "reference sequence" for the numbering of amino acid positions in an alignment with an amino acid sequence at least 80% identical to SEQ ID NO: 1, in such a way that at least 80% of the amino acids present at a "corresponding position" in both aligned sequences are identical.

In one embodiment, the isolated P450 enzyme featured in the invention comprises or consists of an amino acid sequence at least 80% identical to SEQ ID NO: 1, wherein said sequence comprises both a threonine at position corresponding to position 225 and an aspartic acid at position corresponding to position 289.

In another embodiment, the P450 enzyme featured in the invention comprises or consists of a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3.

In still another embodiment, the P450 enzyme featured in the invention consists of the sequence SEQ ID NO: 1.

The enzyme featured in the invention may include one or more tag(s), which may facilitate its purification. For example, the tag can be a poly-histidine (poly-His) tag.

Nucleic acids, vectors, host cells and method of producing the enzyme

In one embodiment, the invention features an isolated nucleic acid which comprises or consists of a sequence encoding a P450 enzyme featured in the invention.

Isolated nucleic acids featured in the invention, also called polynucleotides, may be DNA or RNA molecules, that encode the P450 enzyme defined in the section "Enzymes", while taking into account the degeneracy of the genetic code. The isolated nucleic acids can be obtained by standard techniques well known by those skilled in the art, such as by *in vitro* DNA amplification or polymerisation, *in vitro* gene synthesis, oligonucleotide ligation, or by a combination of these techniques.

The nucleic acids featured in the invention are advantageously in isolated or purified form. The terms "purified" and "isolated" have the same meaning as defined above.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide molecules possessing codons non-naturally occurring in the natural nucleotide molecule. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of recombinant polypeptide expression.

A nucleic acid featured in this invention can also include sequences encoding tags, carrier proteins, signal peptides, or non-transcribed or translated sequences increasing expression or stability of the molecule.

The nucleic acids featured in the invention may be used to produce a recombinant P450 enzyme in a suitable expression system. The term "expression system" means a host cell and compatible vector under suitable conditions, e.g. for the expression of a protein encoded by a foreign nucleic acid carried by the vector and introduced to the host cell.

Typically, the nucleic acid may be included in any suitable vector.

Thus, in some embodiments, the invention features a vector comprising a nucleic acid as defined above operatively associated with expression control elements, and a host cell containing said nucleic acid or said vector.

The terms "vector", "cloning vector" and "expression vector" mean the vehicle by which a DNA or RNA sequence (e.g. a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (e.g. transcription and translation) of the introduced sequence.

Any expression vector can be used, such as a plasmid, cosmid, episome, artificial chromosome, phage or a viral vector. Examples of plasmids include replicating plasmids comprising an origin of replication, or integrative plasmids, such as for instance pUC, pcDNA, pBR, and the like.

Other examples of vectors include vectors for animal cells such as 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) and the like.

Examples of viral vectors include adenoviral, retroviral, herpes virus and AAV vectors. Recombinant viruses may be produced by techniques known in the art, such as by transfecting packaging cells or by transient transfection with helper plasmids or viruses. Typical examples of virus packaging cells include PA317 cells, PsiCRIP cells, GPenv+ cells, 293 cells, etc. Detailed protocols for producing such replication-defective recombinant viruses may be found for instance in WO 95/14785, WO 96/22378, US 5,882,877, US 6,013,516, US 4,861,719, US 5,278,056 and WO 94/19478.

The expression vector may comprise a functional expression cassette, such as an expression cassette comprising a nucleic acid sequence encoding a polypeptide featured in the invention, which is operably linked to expression control sequences.

"Expression control sequence(s)" refers to element(s) necessary for expression of a polypeptide and, optionally, for its regulation. Expression control sequence(s) may for instance include a promoter sequence, signals for initiation and termination of translation,

as well as appropriate regions for regulation of translation, such as a promoter, enhancer, terminator and the like, to cause or direct expression of said polypeptide.

In one embodiment, a vector featured in the invention may also comprise a marker gene, for example a gene making it possible to select between a transformed organism and an organism which does not contain the transfected foreign DNA. A marker gene may be a gene that confers resistance to an antibiotic.

According to a specific embodiment, a vector featured in the invention may further comprise a nucleic acid sequence encoding an adrenodoxin (Adx) and/or a nucleic acid sequence encoding an adrenodoxin reductase (AdR). Adx and AdR are oxydo-reduction partners of P450 enzymes.

“AdR” refers to Adrenodoxin reductase (EC: 1.18.1.6) or Adrenodoxin-NADP+ reductase, the enzyme which is known as the first component in the mitochondrial Cytochrome P450 electron transfer system and which is involved in the biosynthesis of all steroid hormones.

In a specific embodiment, said AdR enzyme is selected from the group consisting of AR (NADPH:adrenodoxin oxidoreductase (EC=1.18.1.6) encoded by *arh1* gene) from *Schizosaccharomyces pombe* or from *Saccharomyces cerevisiae* and FNR (Ferredoxin-NADP reductase (EC=1.18.1.2) encoded by *fpr* gene) from *Escherichia coli*.

In a specific embodiment, the AdR enzyme is AdR from *Bos taurus* (referenced as P08165 in UniProtKB/Swiss-Prot, last sequence update: July 15, 1998). In another specific embodiment, the protein sequence of said AdR is SEQ ID NO: 7. In another specific embodiment, the protein sequence of said AdR is a variant of SEQ ID NO: 7, provided it retains its biological activity.

By “Adx” is meant Adrenodoxin or Ferredoxin 1, the protein which is known for its activity of transferring electrons from Adrenodoxin reductase to CYP11A1.

In a specific embodiment, said Adx protein is selected from the group consisting of Fdx from mammalian origin, Etp1fd from *Schizosaccharomyces pombe* and Yah1 from *Saccharomyces cerevisiae*.

In a specific embodiment, the Adx protein is Adx from *Bos taurus* (referenced as P00257 in UniProtKB/Swiss-Prot, last sequence update: July 1, 1989). In another specific embodiment, the protein sequence of said Adx is SEQ ID NO: 8. In another specific embodiment, the protein sequence of said Adx is a variant of SEQ ID NO: 8, provided it retains its biological activity.

A vector or nucleic acid featured in the invention can be used to transform host cells according to techniques commonly known to those skilled in the art. Insertion of said vector into the host cell may be transient or stable.

The vector may also contain sequences encoding specific signals which trigger the secretion of the translated protein or its targeting to cellular compartments or organelles. These various control signals are selected according to the host cell and may be inserted into vectors which self-replicate in the host cell, or into vectors which integrate the genome of said host.

Host cells may be prokaryotic or eukaryotic, including but not limited to bacteria, yeasts, plant cells, animal cells, insect cells, and mammalian cells, including cell lines which are commercially available. In some embodiments, the expression host cells are *Escherichia coli*, *Lactobacilli*, *Bacillus*, in particular *Bacillus megaterium*, probiotic bacteria, *Pichia pastoris*, *Saccharomyces cerevisiae*, insect cells, plant cells, COS cells and CHO cells. In one embodiment, the host cell is a prokaryotic cell, in particular a bacterial cell. In another particular embodiment, the host cell is a eukaryotic cell, such as a yeast cell such as a cell of *Saccharomyces cerevisiae*.

Common expression systems include bacterial host cells and plasmid vectors, insect host cells and Baculovirus vectors, and mammalian host cells and vectors. Specific examples include *E. coli*, *B. megaterium*, *Kluyveromyces* or *Saccharomyces* yeasts, mammalian cell lines (e.g., Vero cells, CHO cells, 3T3 cells, COS cells, etc.) as well as primary or established mammalian cell cultures (e.g., produced from lymphoblasts, fibroblasts, embryonic cells, epithelial cells, nervous cells, adipocytes, etc.). Examples also include mouse SP2/0-Ag14 cells (ATCC CRL1581), mouse P3X63-Ag8.653 cells (ATCC CRL1580), CHO cells in which a dihydrofolate reductase gene (also referred to as "DHFR gene") is defective (Urlaub G *et al*; 1980), rat YB2/3HL.P2.G11.16Ag.20 cells (ATCC CRL1662, also referred to as "YB2/0 cell"), and the like.

In another embodiment, the invention features a cell which has been transfected, infected or transformed by a nucleic acid and/or a vector provided herein. Consequently, the present invention further concerns a host cell containing a nucleic acid and/or a vector described herein, as well as progeny and/or derivatives of such host cells. In one embodiment, the host cell is a genetically engineered microorganism.

As used herein, the "microorganism" can be an *Escherichia coli*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus subtilis*, *Kluyveromyces lactis*, *Saccharomyces cerevisiae* or *Schizosaccharomyces pombe*.

As shown in 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), the mature form of human P450scc was expressed in *Escherichia coli* in order to provide a more convenient source of the human enzyme and to enable structure-function studies to be done using site-directed mutagenesis. This

expression system enabled to produce larger quantities of active cytochrome than have previously been isolated from placental mitochondria. The expressed P450scc was purified to near homogeneity and shown to have catalytic properties comparable to the enzyme purified from the human placenta. The mature form of human adrenodoxin was also expressed in *E. coli* and supported cholesterol side chain cleavage activity with the same Vmax as that observed using bovine adrenodoxin but with a higher Km. Mutation of Ile-462 to Leu in human P450scc caused a decrease in the catalytic rate constant (kcat) with cholesterol as substrate, increased the Km for 22R-hydroxycholesterol, but did not affect the kinetic constants for 20 alpha-hydroxycholesterol.

As shown in Duport, C., R. Spagnoli, et al. (1998) (Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. Nat Biotechnol **16**(2): 186-9), the first two steps of the steroidogenic pathway were reproduced in *Saccharomyces cerevisiae*. Engineering of sterol biosynthesis by disruption of the delta 22-desaturase gene and introduction of the Arabidopsis thaliana delta 7-reductase activity and coexpression of bovine side chain cleavage cytochrome P450, adrenodoxin, and adrenodoxin reductase, lead to pregnenolone biosynthesis from a simple carbon source. Following additional coexpression of human 3 beta-hydroxysteroid dehydrogenase/isomerase, pregnenolone was further metabolized to progesterone. Steroid formation appeared to be coupled to yeast sterol biosynthesis.

In 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), the production of hydrocortisone, the major adrenal glucocorticoid of mammals and an important intermediate of steroidal drug synthesis, was reported from a simple carbon source by recombinant *Saccharomyces cerevisiae* strains. An artificial and fully self-sufficient biosynthetic pathway involving 13 engineered genes was assembled and expressed in a single yeast strain. Endogenous sterol biosynthesis was rerouted to produce compatible sterols to serve as substrates for the heterologous part of the pathway. Biosynthesis involved eight mammalian proteins (mature forms of CYP11A1, adrenodoxin (ADX), and adrenodoxin reductase (ADR); mitochondrial forms of ADX and CYP11B1; 3beta-HSD, CYP17A1, and CYP21A1). Optimization involved modulating the two mitochondrial systems and disrupting of unwanted side reactions associated with ATF2, GCY1, and YPR1 gene products. Hydrocortisone was the major steroid produced. This work demonstrated the feasibility of transferring a complex biosynthetic pathway from higher eukaryotes into microorganisms.

In a specific embodiment, the microorganism is *Bacillus megaterium*, in particular the strain referred to as *Bacillus megaterium* MS941. The expression "*Bacillus*

megaterium MS941 strain” refers to the strain referenced in Wittchen and Meinhardt, 1995, Appl Microbiol Biotechnol. 42: 871-877, and derived from the DSM319 strain (Deutsche Stammsammlung von Mikroorganismen und Zellkulturen).

The system described previously using *E Coli* and the system using *B. megaterium* are two systems that are capable of realizing sterol and sterol derivatives conversion either *in vitro* with *E.coli* or *in cellulo* with *B. megaterium*. These technologies can evaluate the capacities of P450scc and substrates to function together. Both technologies require an efficient method to solubilize the substrates so that they are accessible to the enzyme. This solubilization could be challenging at the industrial level (large volume) both in financial and scientific terms.

A third organism *S.cerevisiae* is a particularly advantageous organism in this regard; its large fermentation is well mastered at cubic meter or higher levels and it has the potent capacity to produce steroids such as pregnenolone from a simple carbon source such as glucose and/or ethanol thus avoiding of detergent or solubilizing chemical such as beta-cyclodextrin (as referenced above: Duport, C., R. Spagnoli, et al. (1998); Szczebara, F. M., C. Chandelier, et al. (2003)). To permit *in vivo* steroid synthesis, the sterol biosynthesis was routed to produce a P450scc compatible sterol.

As *S. cerevisiae* fermentation is well mastered at large scale, introduction of the newly described P450scc cDNA into a suitable *S.cerevisiae* routed for campesterol, or other sterols described as substrates, in the presence of an appropriate electron carrier is an asset for production of steroid. This recombinant organism can be used for evaluating a new P450scc cDNA as described herein.

Therefore, in another specific embodiment, the microorganism is *Saccharomyces cerevisiae*.

The term “genetically engineered” microorganism refers to any microorganism which has been modified by genetic engineering techniques known in the field by the skilled in the art. For methods related to particular microorganisms, the person skilled in the art can refer to reference manuals such as for yeast: “Methods in yeast genetics” — A laboratory course manual by M Rose, F Winston and P Hieter. pp 198. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 1990. ISBN 0-87969-354-1.

These techniques are conventional techniques, unless otherwise indicated, in the fields of bioinformatics, cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature.

As used herein, “genetic engineering techniques” relate to technologies allowing the expression or overexpression of gene expression known in the field by the skilled in the

art. The expression or overexpression of a gene of interest can be achieved by introducing into a cell or a microorganism an exogenous nucleic acid sequence comprising such gene of interest, by any transformation technologies known by the skilled in the art.

5 The term "transformation" or "transfection" means the introduction of a "foreign" (i.e. heterologous) gene, DNA or RNA sequence to a host cell, so that the host cell will express the introduced gene or sequence to produce a desired substance, typically a protein or enzyme coded by the introduced gene or sequence. The transfection of the host cell may be performed using any standard technique, such as chemical transformation, electroporation, phosphate calcium precipitation or lipofection. The transformation techniques also include for instance the PEG-mediated protoplast transformation technique (Barg *et al*, 2005, Microbial Processes and Products 165-184).

10 "Exogenous nucleic acid sequence(s)" relates to nucleic acid sequence(s) non originally and/or non naturally expressed in the considered microorganism or the way it is in the natural strain of microorganism (in term of expression level for example), and which have been used to transform said microorganism in order to obtain a genetically engineered microorganism as referred above. In a specific embodiment, the exogenous nucleic acid sequence originates from another species than the considered microorganism (e.g. another species of microorganism or organism). In another specific embodiment, exogenous sequence originates from the same microorganism.

20 Said exogenous nucleic acid sequence(s) may encode proteins of interest such as cytochrome P450 and oxydo-reduction partners AdR and Adx and the expression "exogenous DNA" can designate each individual sequence or encompass a whole sequence comprising each individual sequence. As a non-limiting example, said exogenous nucleic acid sequences are integrated into the genome of said microorganism by techniques known in the field, such as by homologous recombination.

30 In one aspect, the invention provides a genetically engineered microorganism capable of converting a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, cholesterol analogues and derivatives into a steroid hormone precursor, wherein said microorganism comprises a nucleic acid featured in the invention, and optionally an exogenous nucleic acid sequence encoding an Adx and/or an exogenous nucleic acid sequence encoding an AdR.

35 In a specific embodiment, the genetically engineered microorganism is *Saccharomyces cerevisiae*.

In another specific embodiment, the genetically engineered microorganism is *Bacillus megaterium*, in particular the *Bacillus megaterium* MS941 strain.

The present invention also relates to an *in vitro* method for preparing a P450 enzyme featured in the invention, said method comprising the steps of:

- a) culturing a host cell under conditions suitable to obtain expression of the P450 enzyme; and
- b) recovering the expressed enzyme.

"Conditions suitable to obtain expression of the P450 enzyme" are well known by the skilled in the art. Typically, either LB- (25 g/L), TB- (24 g/L yeast extract, 12 g/L tryptone, 0.4 % glycerol, 10 mM potassium phosphate buffer) or EnPresso™ Tablet medium are used for the cultivation of the host cell. A pre-culture may be performed by inoculating 50 mL medium containing 10 µg/mL tetracycline with cells from a plate or glycerol stock. A main culture may then be performed by inoculating 50 mL medium containing 10 µg/mL tetracycline with 500 µL sample of the pre-culture. The main culture is typically grown until an optical density of ~0.4 is reached. Protein expression may be induced after addition of 0.25 g xylose dissolved in 1 mL distilled water.

The P450 enzyme can then be purified by means of well-known procedures for purification: it may be purified from lysates or cell extracts, inclusion bodies or from the culture supernatant by methods such as HPLC chromatography, immunoaffinity techniques with specific antibodies, and the like.

Alternatively, the P450 enzyme may be expressed *in vitro* with a cell-free transcription and translation system from a DNA or RNA matrix containing required elements for its expression in a cell lysate or reconstituted system (for example, Rapid Translation System®, Roche Diagnostics or Retic Lysate IVT™, Ambion).

Methods of producing steroid hormone precursor

As shown in Examples 2-4, the inventors have developed a new P450scc enzyme bearing specific mutations which shows an improved enzymatic activity in terms of substrate conversion into steroid hormone precursor.

Therefore, in another aspect, the invention pertains to the use of a P450 enzyme featured in the invention for producing a steroid hormone precursor.

The invention further provides methods for producing a steroid hormone precursor as described below.

A first method for producing a steroid hormone precursor comprises the steps of:

- a) providing a microorganism as described above,

b) culturing said microorganism under conditions allowing the expression of a P450 enzyme featured in the invention,

c) contacting the microorganism culture obtained at step b) with a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol derivative, in conditions allowing the production, by the microorganism, of a steroid hormone precursor from said substrate, and

d) recovering the steroid hormone precursor produced.

According to a specific embodiment, steps b) and c) are performed simultaneously.

In this case, the method for producing a steroid hormone precursor comprises the steps of:

a) providing a microorganism as described above,

b) culturing said microorganism in the presence of a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol derivative, in conditions allowing the expression of a P450 enzyme featured in the invention and allowing the production, by the microorganism, of the steroid hormone precursor from said substrate, and

c) recovering the steroid hormone precursor produced.

A second method for producing a steroid hormone precursor comprises the steps of:

a) contacting a P450 enzyme featured in the invention with an isolated Adx polypeptide, an isolated AdR polypeptide and a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol derivative in conditions allowing the transformation of said substrate into a steroid hormone precursor, and

b) recovering the steroid hormone precursor obtained.

As used herein, the term "substrate" encompasses polycyclic and unsaturated mono alcohols having an aliphatic side chain such as phytosterol derivated from cycloartenol and lanosterol. Among these substrates, "cholesterol, cholesterol analogues and derivatives thereof" refers to a list of substrates selected from the group consisting of cholesterol, brassicasterol, campesterol, ergostadienol such as ergosta 5, 22 dienol, ergosta 5, 24 (28) dienol, ergosta 5, 24(28) diene 3beta ol, ergosta 5, 24 (25) dienol, ergostatrienol such as ergosta 5, 22, 24 (25) trienol, ergosta 5, 22, 24 (28) trienol, ergosta 5, 7 , 22 trienol, ergostatetrenol such as ergosta 5,7, 22, 24 (25) ou ergosta 5,7, 22, 24

(28), desmosterol, beta-sitosterol, generol, a mixture of oxysterols, stigmasterol, vitamin D, 7-Dehydrocholesterol and ergosterol. Sterol mixes currently used in industrial processes are also encompassed in this definition of substrates, such as generol 100 and ADM90 (comprising brassicasterol + campesterol + stigmasterol + beta-sitosterol at different ratios).

In a particular embodiment, the substrate is selected from the group consisting of cholesterol, campesterol, desmosterol and ergosta 5, 24(28) diene 3beta ol. In a specific embodiment, the substrate is ergosterol.

In a specific embodiment, the "steroid hormone precursor" is selected from the group consisting of pregnenolone, 7-Dehydropregnenolone, Hydroxyergosterol, and Hydroxystigmasterol. In another specific embodiment, said steroid hormone precursor is selected from the group of hydroxylated cholesterol analogues and secosteroids (such as vitamins D2 and D3 as derivatives of the cholesterol analogues 7-dehydrocholesterol and ergosterol). In a particular embodiment, the steroid hormone precursor is pregnenolone.

"Culturing the microorganism in the presence of a substrate" means physically interacting said microorganism with a substrate as defined above. This interaction can be achieved within a culture medium or not. In a specific embodiment, said culture medium contains agents for the permeabilization of a microorganism according to the present invention and/or the solubilization of a substrate according to the present invention. Dissolution of the substrate in these agents can be prior to the addition to microorganism culture.

"Contacting a P450 enzyme with a substrate" means physically interacting said P450 enzyme with a substrate as defined above. This interaction can be achieved within a medium or not. In a specific embodiment, said medium contains agents for the solubilization of a substrate according to the present invention. Dissolution of the substrate in these agents can be prior to the addition to said medium.

The agents used for the dissolution of the substrate can be selected from the group consisting of ethanol, Tween-80, tergitol, polyvinylpyrrolidone (PVP), saponins (such as *Quillaja* saponin which is contained within crude extracts of the soap bark tree *Quillaja saponaria* for example), cyclodextrins and derivatives thereof (e.g. 2-hydroxypropyl- β -cyclodextrin). In another specific embodiment, mixtures of these agents can be used, such as a mixture of ethanol and Tween-80, a mixture of tergitol and ethanol, a mixture of saponins (e.g. *Quillaja* saponin) and cyclodextrins as non-limiting examples. Cyclodextrin derivatives can be used, such as 2-hydroxypropyl- β -cyclodextrin. In another specific embodiment, substrates are co-crystallized with polyvinylpyrrolidone (PVP).

In another embodiment, substrate is first dissolved in 2-hydroxypropyl- β -cyclodextrin prior to addition to the microorganism culture wherein said culture contains *Quillaja* saponin. In another embodiment, substrates are dissolved in a solution comprising a percentage of 2-hydroxypropyl- β -cyclodextrin ranging from 10 to 60%, e.g., 20 to 50%,
5 e.g., 40 to 50%, and a percentage of *Quillaja* saponin ranging from 1 to 10%, e.g., 2 to 8%, e.g., 3 to 6%. In another embodiment, substrates are dissolved in a solution comprising 45% of 2-hydroxypropyl- β -cyclodextrin and 4% of *Quillaja* saponin.

In another embodiment, the final concentration of 2-hydroxypropyl- β -cyclodextrin in the microorganism culture is between 1 and 4%, e.g., between 2 and 3%, such as 2.25%
10 as illustrated in example 2. In another embodiment, the final concentration of *Quillaja* saponin in the microorganism culture is between 0.05 and 0.25%, e.g., between 0.075 and 0.225%, e.g., between 0.1 and 0.2%.

The "culturing of said microorganism under conditions allowing the expression of said exogenous nucleic acid sequences" may be performed according to any well-known
15 culturing and inducing methods in the biotechnology field such as described in Bleif et al. (Appl Microbiol Biotechnol (2012) 93:1135–1146) or in Korneli et al. (Journal of Biotechnology 163 (2013) 87–96).

"Conditions allowing the transformation of said substrate into a steroid hormone precursor" include *in vitro* conditions, such as the reagents used, the reagent
20 concentrations, the temperature, the use of agitation, the duration of the reaction or incubation, etc., which favors the transformation of said substrate into a steroid hormone precursor. These conditions may be determined and adjusted by methods known to those skilled in the art.

As a non-limiting example, the reaction may be performed in presence of glucose-6-phosphate dehydrogenase and glucose-6-phosphate, typically in 150 mM of HEPES
25 buffer, for example adjusted to pH 7.4, containing typically 0.05 % Tween-20 and 1 mM MgCl₂. The concentrations of the P450 enzyme, Adx and AdR may for instance be of 1 μ M, 20 μ M and 0.5 μ M, or of 0.25 μ M, 5 μ M and 0.125 μ M, respectively. The polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol,
30 cholesterol analogue or cholesterol derivative may be dissolved in 45% 2-hydroxypropyl- β -cyclodextrin. The samples may be pre-warmed to 37°C and NADPH may be added to a final concentration of 100 μ M. The mixtures may be incubated at 37°C with agitation.

In a specific embodiment, the steroid hormone precursor used in any method for producing a steroid hormone precursor is pregnenolone. As used herein, "pregnenolone"
35 refers to a steroid hormone also referred to as 3 β -hydroxypregn-5-en-20-one.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO: 1 shows the amino acid sequence of CYP11A1 comprising a threonine at position 225 and an aspartic acid at position 289 (polypeptide SA1).

SEQ ID NO: 2 shows the amino acid sequence of CYP11A1 comprising a threonine at position 225.

SEQ ID NO: 3 shows the amino acid sequence of CYP11A1 comprising an aspartic acid, at position 289.

SEQ ID NO: 4 shows the amino acid sequence of WT Bos taurus CYP11A1 (polypeptide SA4; WT P450scc).

SEQ ID NO: 5 shows the sequence of the plasmid pSMF2.1_CYP11A1BYM encoding for SA1 of SEQ ID NO: 1.

SEQ ID NO: 6 shows the amino acid sequence of the polypeptide SA6 (P450scc-l1A-K193E).

SEQ ID NO: 7 shows the amino acid sequence of AdR.

SEQ ID NO: 8 shows the amino acid sequence of Adx.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: HPLC-chromatogram showing the *in vitro* conversion of 20 μ M cholesterol by the cytochrome P450 polypeptides SA1 (SEQ ID NO:1) and SA4 (control having sequence SEQ ID NO: 4). Legend: S: substrate, P: main product, DOC: 11-deoxycorticosterone.

Figure 2: Pregnenolone quantification from HPLC-chromatogram showing the *in vivo* conversion of different substrates into pregnenolone by the cytochrome P450 polypeptide SA1 (SEQ ID NO:1).

Figure 3: HPLC-chromatogram showing the *in vivo* conversion of 300 μ M cholesterol by the cytochrome P450 polypeptides SA1 (SEQ ID NO:1) and SA6 (control having sequence SEQ ID NO: 6) after 24 h. Legend: Chol.: cholesterol, Prog.: pregnenolone.

Figure 4: Time course for the *in vivo* conversion of 300 μ M cholesterol by the cytochrome P450 polypeptides SA1 (SEQ ID NO:1) and SA6 (control having sequence SEQ ID NO: 6).

Figure 5: Vector map of pSMF2.1_CYP11A1BYM encoding for SA1 (SEQ ID NO: 1).

EXAMPLES**Example 1: Materials and Methods***Protein synthesis*

CYP11A1 variants were obtained by gene synthesis and cloned into pTRC99A, fused to a poly-His tag. C43DE3- *E. coli* cells were co-transformed with each vector and the chaperone-encoding pGro12. The purification was performed as described in Janocha *et al.* (*Biochim Biophys Acta*, (2011) Jan;1814(1):126-31). The purity was assessed by SDS-PAGE. Concentrations of the purified proteins were determined by CO-difference spectra, after treatment with sodium dithionite and exposure to CO.

Adx and AdR were purified according to Uhlmann *et al.* (Biochem. Biophys. Res. Commun., 188 (1992), pp. 1131–1138) and Sagara *et al.* (Biol. Pharm. Bull., 16 (1993), pp. 627–630), respectively.

In vitro conversion assay

For *in vitro* enzyme assays, 150 mM of HEPES buffer, adjusted to pH 7.4, containing 0.05 % Tween-20 and 1 mM MgCl₂, were applied as reaction buffer. 1 unit of glucose-6-phosphate dehydrogenase, with 5 mM glucose-6-phosphate as substrate, served as a NADPH-regenerating system.

For WT P450scc (noted “SA4”, SEQ ID NO: 4) and P450scc-I1A-K193E (noted “SA6”, SEQ ID NO: 6) the concentrations of CYP11A1, Adx and AdR were of 1 μ M, 20 μ M and 0.5 μ M, and for P450scc-R225T-N289D (noted “SA1”, SEQ ID NO: 1) the concentrations of CYP11A1, Adx and AdR were of 0.25 μ M, 5 μ M and 0.125 μ M, respectively.

20 μ M of cholesterol, dissolved in 45 % 2-hydroxypropyl- β -cyclodextrin, served as substrate for CYP11A1.

The samples were pre-warmed to 37°C and the reaction was started by addition of NADPH to a final concentration of 100 μ M. The mixtures were incubated at 37°C with agitation for 30 s. The reaction was stopped by boiling the samples in water for 30 s.

To allow photometric detection at 240 nm, the steroids were converted into their 3-keto- Δ^4 derivatives using a cholesterol oxidase from *Nocardia spec.*

20 μ l of a cholesterol oxidase solution (5 mg cholesterol oxidase and 5 mg Na-cholate dissolved in 5 ml of 50 mM HEPES buffer pH 7, containing 0.05 % Tween-20) were added to the samples. After incubation at 37°C for 1 h, 11-deoxycorticosterone (DOC) was added to the reaction mixtures as an internal standard, followed by a 2-times

extraction with equal volumes of ethylacetate. After evaporation, the extracts were dissolved in acetonitrile/water.

In vivo conversion assay

The *in vivo* conversion of 300 μ M cholesterol into pregnenolone was evaluated by HPLC after 24 h. *Bacillus megaterium* was cultivated in TB-medium containing 10 μ g/ml tetracycline at 37°C with 180 rpm shaking. Protein expression was induced by adding 0.25 g of xylose dissolved in 1 mL water, followed by the subsequent addition of the substrate, dissolved in 2-hydroxypropyl- β -cyclodextrin.

Example 2: *in vitro* conversion of cholesterol by the cytochrome P450 polypeptides SA1 (SEQ ID NO:1) and SA4 (WT control, SEQ ID NO:4)

After acute evaluation of the concentration of each P450scc preparation, SA1, SA4 and SA6 were used in the following *in vitro* assays. Each of these isoforms was reconstituted at a concentration of 1 μ M in the presence of cholesterol as a substrate (20 μ M) in saturating amounts of Adx and AdR.

While SA4 and SA6 showed identical activity, SA1 showed a 2-3-fold higher activity compared to SA4 and SA6 (**Fig. 1**). Since the substrate is already depleted after the incubation time, it is difficult to assess the differences between the SA1 and SA1/SA4 polypeptides at longer incubation time.

Example 3: *in vivo* conversion of cholesterol by the cytochrome P450 polypeptides SA1 (SEQ ID NO:1) and SA6 (control, SEQ ID NO:6)

As no measurable differences were observed between SA4 and SA6, the inventors focused their interest on comparing the SA1 and SA6 polypeptides. A new system was recently developed in which cholesterol is metabolized into pregnenolone by recombinant *Bacillus megaterium* expressing P450scc, Fdx1 and FdxR in the presence of solubilized cholesterol. SA1 and SA6 corresponding cDNAs were optimized for *Bacillus megaterium* codon BIAS and transferred into the plasmid pSFM2.1 using the appropriate restriction sites. Both pSFM2.1 plasmids bearing respectively the codon optimized SA1 and SA6 were transferred into *Bacillus megaterium* MS941 using a classical protoplast transformation protocol.

The *in vivo* cholesterol conversion activity was assessed after 24 h and 48 h for the SA1 and SA6 polypeptides expressed in *Bacillus megaterium*.

In accordance with the *in vitro* experiments, SA1 exhibited now a 2-fold higher activity compared to SA6 (**Fig. 3 and 4**).

Example 4: *in vivo* conversion of different substrates into pregnenolone by the cytochrome P450 polypeptide SA1 (SEQ ID NO:1)

Finally the improved system was used to test the SA1 polypeptide conversion capacity with various substrates of biotechnological interest: campesterol, desmosterol, ergosta-5,24(28)-dien-3 β -ol and a mixture of various 20, 22-OH oxysterols.

Each substrate was converted to one main product with the same retention time as progesterone, indicating that CYP11A1 was able to cleave the side-chain of each of these substrates, yielding pregnenolone.

Pregnenolone formation was quantified for every substrate (**Fig. 2**). The polar oxysterols were converted at a rate comparable to cholesterol (~ 35 mg/L after 48 hours). Conversion of the more hydrophobic sterols campesterol, desmosterol and ergostadienol only occurred at a rate of approximately 19 % after 48 hours, compared to cholesterol.

Taken together, the results show that all of the tested steroids were able to permeate through the cell membrane of *Bacillus megaterium* and were converted to pregnenolone by the CYP11A1 SA1 mutant.

The inventors thus report here a new P450scc protein, named SA1, which shows an increased conversion activity towards various sterols including cholesterol.

CLAIMS

1. An isolated P450 enzyme comprising or consisting of an amino acid sequence at least 80% identical to SEQ ID NO: 1, wherein said sequence comprises a threonine at position corresponding to position 225 and/or an aspartic acid at position corresponding to position 289.
2. The enzyme according to claim 1 comprising or consisting of a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3.
3. The enzyme according to claim 1 or 2, consisting of the sequence SEQ ID NO: 1.
4. An isolated nucleic acid which comprises or consists of a nucleotide sequence encoding an enzyme as defined in any one of claims 1 to 3.
5. A vector comprising a nucleic acid as defined in claim 4 which is operatively associated with expression control sequences.
6. The vector according to claim 5, further comprising a nucleic acid sequence encoding adrenodoxin (Adx) and/or a nucleic acid sequence encoding adrenodoxin reductase (AdR).
7. A host cell containing a nucleic acid as defined in claim 4 or a vector as defined in claims 5 or 6.
8. The host cell according to claim 7, which is a genetically engineered microorganism.
9. The host cell according to claim 7 or 8 which is *Saccharomyces cerevisiae*.
10. A genetically engineered microorganism capable of converting a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol derivative, into a steroid hormone precursor, wherein said microorganism comprises a nucleic acid as defined in claim 4, and optionally a nucleic acid sequence encoding adrenodoxin (Adx) and/or a nucleic acid sequence encoding adrenodoxin reductase (AdR).

11. An *in vitro* method for preparing an enzyme as defined in any one of claims 1 to 3, said method comprising the steps of:

- a) culturing a host cell as defined in any one of claims 7 to 9 under conditions suitable to obtain expression of the enzyme as defined in any one of claims 1 to 3; and
- b) recovering the expressed enzyme.

12. Use of an enzyme as defined in one of claims 1 to 3 for producing a steroid hormone precursor.

13. A method for producing a steroid hormone precursor, comprising the steps of:

- a) providing a microorganism as defined in any one of claim 8 to 10,
- b) culturing said microorganism under conditions allowing the expression of an enzyme as defined in any one of claims 1 to 3,
- c) contacting the microorganism culture obtained at step b) with a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol derivative, in conditions allowing the production, by the microorganism, of a steroid hormone precursor from said substrate, and
- d) recovering the steroid hormone precursor produced.

14. A method for producing a steroid hormone precursor, comprising the steps of:

- a) contacting an enzyme as defined in any one of claims 1 to 3 with an isolated adrenodoxin (Adx) polypeptide, an isolated adrenodoxin reductase (AdR) polypeptide and a substrate selected from the group consisting of polycyclic and unsaturated mono alcohols having an aliphatic side chain such as cholesterol, a cholesterol analogue and a cholesterol derivative in conditions allowing the transformation of said substrate into a steroid hormone precursor, and
- b) recovering the steroid hormone precursor obtained.

15. The method according to claim 13 or 14, wherein the steroid hormone precursor is pregnenolone.