(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 July 2012 (05.07.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/089613 A1

(51) International Patent Classification:
C12N 9/00 (2006.01) *C12N 5/10* (2006.01)

(21) International Application Number:
PCT/EP2011/073765

(22) International Filing Date:
22 December 2011 (22.12.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

10197201.6	28 December 2010 (28.12.2010)	EP
61/427,609	28 December 2010 (28.12.2010)	US
11160000.3	28 March 2011 (28.03.2011)	EP
61/468,159	28 March 2011 (28.03.2011)	US

(71) Applicant *(for all designated States except US)*: **DSM IP ASSETS B.V.** [NL/NL]; Het Overloon 1, NL-6411 TE Heerlen (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HANS, Marcus** [DE/NL]; Van Slingelandtstraat 28, NL-2582 XR Den Haag (NL). **DULK, DEN, Ben** [NL/NL]; Prinses Wilhelminastraat 10, NL-2313 AW Leiden (NL). **WU, Liang** [NL/NL]; Van Barenstraat 25, NL-2628 LC Delft (NL).

(74) Agents: **KLEIBORN, Paul Erik** et al.; DSM Intellectual Property, P.O. Box 130, NL-6100 AC Echt (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2012/089613 A1

(54) Title: PROCESS TO INCREASE THE PRODUCTION OF A SUCCINYL-COA DERIVED COMPOUND

(57) Abstract: This invention relates to a process to improve the production of a succinyl-CoA derived compound by a eukaryotic cell. This method may be advantageously used in the biological production of e.g. adipic acid.

PROCESS TO INCREASE THE PRODUCTION OF A SUCCINYL-COA DERIVED COMPOUND

Field of the invention

This invention relates to a process to improve the production of a succinyl-CoA derived compound by a eukaryotic cell. The invention furthermore relates to an engineered eukaryotic cell transformed with a polynucleotide encoding a succinyl-CoA ligase obtainable by said process and to a method to produce adipic acid or adipic acid ester or adipic acid thioester comprising culturing said engineered eukaryotic cell. The invention also relates to an engineered eukaryotic cell comprising an enzyme capable of converting an adipic ester, adipic acid ester or adipic acid thioester into 5-formylpentanoate and to a method to produce 5-formylpentanoate by culturing said cell. The invention also relates to an engineered eukaryotic cell comprising an enzyme capable of converting 5-formylpentanoate into 6-amino caproic acid and to a method to produce 6-amino caproic acid by culturing said engineered cell and to a method to produce caprolactam from said 6-amino caproic acid.

Background of the invention

Succinyl-coenzyme A (succinyl-CoA) is an essential intermediate in numerous metabolic pathways and a key precursor in the synthesis of many (industrially relevant) valuable compounds, such as fatty acids, carotenoids, isoprenoids, vitamins, amino acids, lipids, wax esters, (poly)saccharides, polyhydroxyalkanoates, statins, polyketides In particular, succinyl-CoA is also the precursor of the industrially important bulk chemical hexamethylene diamine and adipic acid.

A biosynthetic production pathway of such succinyl-CoA-derived compounds may be naturally present in a microorganism, but often the microorganism must be engineered in order to create a functional biosynthetic pathway, for example by introducing genes encoding specific enzymes.

When natural (wild type) cells are not capable of producing the succinyl-CoA-derived compound of interest, the use of metabolic engineering can provide for

eukaryotic cells expressing heterologous genes that could support such a process. In such cases, the heterologous gene products are usually targeted to the cytosolic compartment of the cell. As the biosynthesis of a succinyl-CoA-derived compound will take place completely or partially in the cytosol of the cell, the supply of sufficient amounts of the precursor succinyl-CoA in the cytosolic compartment is crucial. In eukaryotes, biosynthesis of succinyl-CoA takes place mainly in the mitochondrion. Succinyl-CoA is an intermediate of the tricarboxylic acid or Krebs cycle. Two enzymes in the mitochondria generate and use succinyl-CoA. The alpha-ketoglutarate dehydrogenase complex produces succinyl-CoA, and in a subsequent reaction succinyl-CoA gets converted to succinic acid by a reversible succinyl-CoA ligase enzyme, giving rise to ATP (or GTP) generation. In industrially relevant eukaryotic microorganisms there have been no examples of succinyl-CoA formation outside the mitochondria. In haem biosynthesis, for example, the first step involves the reaction of succinyl-CoA and glycine to aminolevulinate (ALA). Due to the location of the succinyl-CoA, this step takes place in the mitochondrion, and the ALA is then transported into the cytosol where subsequent enzymatic reaction will occur.

It is an aim of the invention to increase the production of a succinyl-CoA derived compound in a eukaryotic cell.

Detailed description of the invention

In a first aspect the present invention provides a method to improve the production of a succinyl-CoA derived compound by a eukaryotic cell, which is capable of producing a succinyl-CoA derived compound, comprising transforming said cell with a polynucleotide encoding a succinyl-CoA ligase.

The term “a” or “an” as used herein is defined as “at least one” unless specified otherwise.

When referring to a noun (e.g. a compound, a cell *etc.*) in the singular, the plural is meant to be included.

When referring to an acid (e.g. adipic acid) the conjugated base or salt is also included.

“A succinyl-CoA ligase” is defined as a polypeptide, which may comprise one, two, or more subunits which may each encoded by a different gene. For example, the succinyl-CoA ligase from Escherichia coli consists of two subunits a and b which are

encoded by the sucC and sucD genes, respectively. Likewise, the succinyl-CoA ligase from Saccharomyces cerevisiae also consists of two subunits a and b which are encoded by lsc1 and lsc 2 genes, respectively. Therefore, “a” polynucleotide encoding a succinyl-CoA ligase” is understood to also encompass instances wherein a plurality of polynucleotides encode for a succinyl-CoA ligase, such as two, three, or more polynucleotides.

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The essential nature of such analogues of naturally occurring amino acids is that, when incorporated into a protein, that protein is specifically reactive to antibodies elicited to the same protein but consisting entirely of naturally occurring amino acids. The terms "polypeptide", "peptide" and "protein" are also inclusive of modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation.

The term “enzyme” as used herein is defined as a protein which catalyses a (bio)chemical reaction in a cell.

The term "polynucleotide" as used herein, includes reference to a deoxyribonucleotide or ribonucleotide polymer, in either single-or double-stranded form, and unless otherwise limited, encompasses known analogues having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides (e.g., peptide nucleic acids). A polynucleotide can be full-length or a subsequence of a native or heterologous structural or regulatory gene. Unless otherwise indicated, the term includes reference to the specified sequence as well as the complementary sequence thereof.

The term "homologous" or “endogenous” when used to indicate the relation between a given (recombinant) nucleic acid or polypeptide molecule and a given host organism or host cell, is understood to mean that in nature the nucleic acid or polypeptide molecule is produced by a host cell or organisms of the same species, preferably of the same variety or strain.

The term "heterologous" when used with respect to a nucleic acid (DNA or RNA) or protein refers to a nucleic acid or protein that does not occur naturally as part of the

organism, cell, genome or DNA or RNA sequence in which it is present, or that is found in a cell or location or locations in the genome or DNA or RNA sequence that differ from that in which it is found in nature. Heterologous nucleic acids or proteins are not endogenous to the cell into which it is introduced, but have been obtained from another cell or synthetically or recombinantly produced.

The term “gene”, as used herein, refers to a nucleic acid sequence containing a template for a nucleic acid polymerase, in yeasts, RNA polymerase II. Genes are transcribed into mRNAs that are then translated into protein.

To increase the likelihood that the introduced protein is expressed in active form in a eukaryotic yeast cell in the method according to the invention, the corresponding encoding nucleotide sequence may be adapted to optimise its codon usage to that of the chosen yeast host cell. Several methods for codon optimisation are known in the art. A preferred method to optimise codon usage of the nucleotide sequences to the eukaryotic cell according to the present invention is codon pair optimization technology as disclosed in WO2008/000632. Codon-pair optimization is a method for producing a polypeptide in a host cell, wherein the nucleotide sequences encoding the polypeptide have been modified with respect to their codon-usage, in particular the codon-pairs that are used, to obtain improved expression of the nucleotide sequence encoding the polypeptide and/or improved production of the polypeptide. Codon pairs are defined as a set of two subsequent triplets (codons) in a coding sequence.

Usually, a nucleotide sequence encoding an protein, is operably linked to a promoter that causes sufficient expression of the corresponding nucleotide sequence in the yeast cell according to the present invention to confer to the cell the ability to produce a dicarboxylic acid.

As used herein, the term "operably linked" refers to a linkage of polynucleotide elements (or coding sequences or nucleic acid sequence) in a functional relationship. A nucleic acid sequence is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the coding sequence.

As used herein, the term "promoter" refers to a nucleic acid fragment that functions to control the transcription of one or more genes, located upstream with respect to the direction of transcription of the transcription initiation site of the gene, and

is structurally identified by the presence of a binding site for DNA-dependent RNA polymerase, transcription initiation sites and any other DNA sequences known to one of skilled in the art. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation.

A promoter that could be used to achieve expression of a nucleotide sequence coding an protein, may be not native to the nucleotide sequence coding for the protein to be expressed, i.e. a promoter that is heterologous to the nucleotide sequence (coding sequence) to which it is operably linked. Preferably, the promoter is homologous, i.e. endogenous to the host cell.

Suitable promoters in yeast cells are known to the skilled man in the art. Suitable promotors may be, but are not limited to TDH1, TDH3, GAL7, GAL10, GAL1, CYC1, HIS3, ADH1, PH05, ADC1, ACT1, TRP1, URA3, LEU2, ENO1, TPI1,. Other suitable promoters include PDC1, GPD1, PGK1, and TEF1.

Usually a nucleotide sequence encoding a protein comprises a terminator. Any terminator, which is functional in the cell, may be used in the present invention. Preferred terminators are obtained from natural genes of the host cell. Suitable terminator sequences are well known in the art. Preferably, such terminators are combined with mutations that prevent nonsense mediated mRNA decay in the host cell of the invention (see for example: Shirley et al., 2002, Genetics 161:1465-1482).

The applicants have surprisingly found that the production of a succinyl-CoA derived compound by a eukaryotic cell, which is capable of producing a succinyl-CoA derived compound, is higher when said cell has been transformed with a polynucleotide encoding a succinyl-CoA ligase as compared to a cell which is not transformed with a polynucleotide encoding a succinyl-CoA ligase. Within the context of the invention “the production of a succinyl-CoA derived compound” is defined as the amount of succinyl-CoA derived compound per volume of fermentation media. In order to establish whether the production of a succinyl-CoA derived compound of a eukaryotic cell is improved upon transformation with a polynucleotide encoding a succinyl-CoA ligase as compared to the production of a succinyl-CoA derived compound of a eukaryotic cell which has not been transformed with a polynucleotide encoding a succinyl-CoA ligase, the fermentation conditions, particularly the length of the fermentation for both cells are preferably similar of even more preferably substantially identical or even advantageously exactly the same.

Preferably the fermentation conditions are selected such that the amount of succinyl-CoA derived compound increases up to completion of the fermentation.

Within the context of the specification "succinyl-CoA ligase" is defined as an enzyme capable of reversible coupling of coenzyme A with succinic acid.

In an embodiment the succinyl-CoA ligase belongs to enzyme class of E.C.6.2.1.5 or E.C. 6.2.1.4. The EC enzyme class is a class wherein the enzyme is classified or may be classified, on the basis of the Enzyme Nomenclature provided by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), which nomenclature may be found at http://www.chem.qmul.ac.uk/iubmb/enzyme/. Other suitable enzymes that have not (yet) been classified in a specified class but may be classified as such, are meant to be included.

In a preferred embodiment the succinyl-CoA ligase is active in the cytosol or the cell. The biosynthesis of succinyl-CoA-derived compound may take place completely or partially in the cytosol. Amino acid sequences of succinyl-CoA ligase may comprise a targeting signal, for instance a peroxisomal or mitochondrial targeting signal. A person skilled in the art may know methods to determine the localization of proteins, for example, as described in Emanuelsson et al. 2007. Nature Protocols 2: 953-971. In the event the succinyl-CoA ligase comprises a targeting signal, it may be preferred that the eukaryotic cell in the method according to the invention comprises a truncated form of the enzyme, wherein the targeting signal is removed. Deletion of targeting signal may localize the protein in the cytosol, which may may result in increased production of said succinyl-CoA-derived compound.

The succinyl-CoA ligase may comprise any suitable ligase and can derived from any organism. The organism may be a eukaryote, bacterium or an archeon.

In an embodiment the succinyl-CoA ligase is derived from a eukaryote, more preferably from Saccharomyces, e.g. S. cerevisae. A suitable eukaryotic succinyl-CoA ligase may be encoded by S. cerevisiae genes LSC1 and/or LSC2 and/or homologues thereof.

In another embodiment the succinyl-CoA ligase is prokaryotic, i.e. derived from a bacterium or archeon, more preferably from Escherichia coli. Transformation of a eukaryotic cell which is capable of producing a succinyl-CoA derived compound with a polynucleotide encoding a yeast, such as Saccharomyces, succinyl-CoA ligase instead

of a prokaryotic succinyl-CoA ligase, in some cases may not result in improving the production of a succinyl-CoA derived compound. A suitable prokaryotic succinyl-CoA ligase may be encoded by E. coli genes sucC and/or sucD and homologues thereof.

In a specific embodiment, the succinyl-CoA ligase may comprise an amino acid sequence comprising SEQ ID's 15 and/or 16 and/or comprising SEQ ID's 17 and/or 18, and/or homologues thereof. Such succinyl-CoA ligases may for instance be encoded by polynucleotide sequences comprising SEQ ID's 1 and/or 2 or comprising SEQ ID's 3 and 4, respectively, and/or homologues thereof. The skilled person also will be able to construe functional analogues of these sequences, which may be used as an alternative, based on common general knowledge.

In a preferred embodiment, the succinyl-CoA ligase comprises an amino acid sequence comprising SEQ ID's 15 and/or 16 and/or homologues thereof. Such succinyl-CoA ligase may for instance be encoded by polynucleotide sequences comprising SEQ ID's 1 and/or 2, respectively, and/or homologues thereof. The skilled person also will be able to construe functional analogues of these sequences, which may be used as an alternative, based on common general knowledge.

The term "homologue" is used herein in particular for polynucleotides or polypeptides having a sequence identity of at least 30%, preferably at least 40%, more preferably at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, in particular at least 85%, more in particular at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99%. The term homologue is also meant to include nucleic acid sequences (polynucleotide sequences) which differ from another nucleic acid sequence due to the degeneracy of the genetic code and encode the same polypeptide sequence.

Sequence identity is herein defined as a relationship between two or more amino acid (polypeptide or protein) sequences or two or more nucleic acid (polynucleotide) sequences, as determined by comparing the sequences. Usually, sequence identities or similarities are compared over the whole length of the sequences compared. In the art, "identity" also means the degree of sequence relatedness between amino acid or nucleic acid sequences, as the case may be, as determined by the match between strings of such sequences.

Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. The percent identity between two amino acid sequences or between two nucleotide sequences may be determined using the Needleman and Wunsch algorithm (Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453) (incorporated herein by reference). Both amino acid sequences and nucleotide sequences can be aligned by the algorithm. The Needleman-Wunsch algorithm has been implemented in the computer program NEEDLE. For the purpose of this invention the NEEDLE program from the EMBOSS package was used (version 2.8.0 or higher, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice, P. LongdenJ. and Bleasby,A. Trends in Genetics 16, (6) pp 276 - 277, http://emboss.bioinformatics.nl/), incorporated herein by reference in their entireties). For protein sequences EBLOSUM62 is used for the substitution matrix. For nucleotide sequence, EDNAFULL is used. The optional parameters used are a gap-open penalty of 10 and a gap extension penalty of 0.5. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

Within the context of the invention “a succinyl-CoA derived product” is understood to include any product which may be produced by a biosynthetic pathway sequence which includes succinyl-CoA, or by a production method including one or more biochemical step. Said pathway is not necessarily limited to a biological function. For example, WO2009/113853, (incorporated herein by reference) describes a microorganisms which is capable of producing a succinyl-CoA derived product, which microorganisms has been transformed with polynucleotides encoding enzymes that are part of a biosynthetic pathway to produce, *inter alia*, adipic acid, which in the context of the invention is a succinyl-CoA derived compound. In this biosynthetic pathway adipic acid is formed by coupling of acetyl-CoA with succinyl-CoA followed by a number of subsequent enzymatic reactions. The production of adipic acid may however not be a biological function of said biosynthetic pathway. Examples of such succinyl-CoA derived product include, in addition to adipic acid, fatty acids, carotenoids, isoprenoids, vitamins, amino acids, lipids, wax esters, (poly)saccharides, polyhydroxyalkanoates,

statins, polyketides, and haem. In a preferred embodiment the succinyl-CoA derived compound is adipic acid.

Suitable eukaryotic cells can be selected in particular from the group of fungi; metazoan; *Viridiplantae* (in particular *Arabidopsis* and *Chlamydomonadales)*; *Diplomonads* (in particular *Giardiinae*); *Entamoebidae* (in particular *Entaboeba*); *Euglenozoa* (in particular *Euglena*); *Pelobiontida* (in particular *Mastigamoeba*); and *Alveolata* (in particular *Cryptosporidium*).

Suitable fungi in particular include fungi and yeasts selected amongst the group of *Rhizopus, Neurospora, Penicillium, Aspergillus*, *Piromyces*, *Trichosporon, Candida, Hansenula, Kluyveromyces, Saccharomyces, Rhodotorula, Schizosaccharomyces, Yarrowia (such as Yarrowia lypolytica*).

In a preferred embodiment the eukaryotic cell is yeast, more preferably Saccharomyces cerevisiae. Compared to bacteria, such as *E. coli*, yeast provides a very suitable alternative to produce the above-mentioned succinyl-CoA derived products, in that yeast is not susceptible to phage or other infection since yeast-based processes may be run at low pH. Therefore, the use of yeast does not require a sterile process, thereby lowering the cost price of the product of interest.

Transformation of eukaryotic cells may be done by methods known in the art, e.g. as described by R. D. Gietz and R. H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method (2007), *Nature Protocols* 2, 31 - 34.

In a second aspect the present invention provides an engineered eukaryotic cell transformed with a polynucleotide encoding a succinyl-CoA ligase obtainable by a process according to the first aspect of the invention. The engineered eukaryotic cell according to the second aspect of the invention is preferably suitable to produce one or more enzymes such that a succinyl-CoA derived compound can be produced in a biosynthetic pathway. One such a biosynthetic pathway is an adipic acid biosynthetic pathway. Therefore, in another embodiment the engineered eukaryotic cell of the second aspect invention further comprises an adipic acid biosynthetic pathway, preferably a CoA dependent adipic acid biosynthetic pathway.

An engineered eukaryotic cell according to the second aspect of the invention is defined herein as a cell which contains, or is transformed or genetically modified with a nucleotide sequence that does not naturally occur in the eukaryotic cell, or it contains additional copy or copies of an endogenous nucleic acid sequence, or it comprises a

deletion or disruption of an endogenous nucleic acid sequence. A wild-type eukaryotic cell is herein defined as the parental cell of the recombinant eukaryotic cell.

A genetic modification with a nucleotide sequence is used herein to indicate that a gene or a nucleotide sequence is introduced into a (eukaryotic) cell by any available means. A nucleotide sequence or gene may be prepared according to any method in the art, for instance extracted from an organism or synthesized by chemical means.

The adipic acid biosynthetic pathway preferably comprises an enzyme selected from the group consisting of an enzyme capable of acyl group transfer such as for example a thiolase, an enzyme capable of catalysing the reduction of a carbon-carbon double bond of a 2,3-enoate moiety or a 2-enoyl moiety, as for example an enoyl reductase, an enzyme capable of catalysing the dehydration of a 3-hydroxyacyl ester or 3-hydroxyacyl thioester to a 2-enoyl ester or thioester such as for example a dehydratase, an enzyme capable of catalysing the reduction of a carbonyl group to an alcohol group or capable of catalysing the reduction of a 3-oxoacyl ester or 3-oxoacyl thioester to the corresponding 3-hydroxyacyl ester or thioester, such as a ketoreductase or a dehydrogenase, and an enzyme capable of converting an adipic acid ester or adipic acid thioester into adipic acid such as an acyl-CoA transferase or an acyl-CoA hydrolase.

In a specific embodiment, the thiolase is a beta-ketoadipyl CoA thiolase, e.g. from Acinetobacter. Such thiolase may comprise a sequence comprising SEQ ID 20 and/or a homologue thereof. Such thiolase is for instance encoded by a gene comprising the sequence comprising SEQ ID 9.

In a specific embodiment, the enoyl reductase is an enoyl-CoA reductase, e.g. from Candida tropicalis. It may comprise a sequence comprising SEQ ID 23 and/or a homologue thereof.

In a specific embodiment, the dehydratase is an enoyl-CoA dehydratase, e.g. from Acinetobacter. Such dehydratase may comprise a sequence comprising SEQ ID 22 and/or a homologue thereof. Such dehydratase is for instance encoded by a gene comprising the sequence comprising SEQ ID 11.

In a specific embodiment, the ketoreductase or dehydrogenase comprises a sequence comprising SEQ ID 21 and/or a homologue thereof. Such ketoreductase or a dehydrogenase is for instance encoded by a gene comprising the sequence comprising SEQ ID 10.

In a specific embodiment, the acyl-CoA transferase comprises a sequence comprising SEQ ID's 24 and/or 25 and/or a homologue thereof. Such acyl-CoA transferase is for instance encoded by a gene comprising the sequence comprising SEQ ID's 13 and/or 14, respectively

In a third aspect the present invention provides a method to produce adipic acid or adipic acid ester or adipic acid thioester comprising culturing the engineered eukaryotic cell according to the second aspect of the invention.

In a fourth aspect the engineered eukaryotic cell according to the second aspect of the invention further comprises an enzyme capable of converting an adipic ester, adipic acid ester or adipic acid thioester into 5-formylpentanoate.

In a fifth aspect the present invention provides a method to produce 5-formylpentanoate comprising culturing the engineered eukaryotic cell according to the fourth aspect of the invention.

In a sixth aspect the present invention provides an engineered eukaryotic cell according to the fourth aspect of the invention further comprising an enzyme capable of converting 5-formylpentanoate into 6-amino caproic acid.

In a seventh aspect the present invention provides a method to produce 6-amino caproic acid comprising culturing the engineered eukaryotic cell according to the sixth aspect of the invention.

In an eighth aspect the present invention provides a method to produce caprolactam comprising preparing 6-amino caproic acid according to the method according to the seventh aspect of the invention and cyclising the 6-amino caproic acid, thereby forming caprolactam.

EXAMPLES

1. Material and methods

Oligonucleotides were synthesized by Invitrogen (Carlsbad CA, US). DNA sequencing was performed at SEQLAB (Göttingen, Germany) or by Baseclear (Leiden, The Netherlands). DNA synthesis was carried out at GeneArt (Regensburg, Germany) or DNA2.0 (Menlo Park, CA, USA). Restriction enzymes were supplied by Invitrogen or New England Biolabs. Strains used for transformation are Escherichia coli DH10B

electromax competent cells (Invitrogen) according to the protocol supplied by manufacturer. All nucleotides were codon optimized for transformation in yeast.

2. Generation of gene expression cassettes for adipic acid production in S. cerevisiae integration vectors

All DNA fragments used in this example were ordered as synthetic DNA from DNA2.0 (Menlo Park, CA, USA) or Geneart. The following DNA molecules were used: DNA encoding an Acinetobacter thiolase (SEQ ID 9), an Acinetobacter ketoreductase (SEQ ID 10), an Acinetobacter dehydratase (SEQ ID 11), a Candida tropicalis enoylreductase (SEQ ID 12) and an Acinetobacter acyl-CoA transferase (SEQ ID's 13 and 14). All DNA molecules (10 μg) were restricted using the restruction enzymes SapI. The desired DNA fragments yielded the promoter-gene-terminator cassette and was purified out of the agarose gel using QiaQuick Gel Extraction Kit (Qiagen, Hilden, Germany). Furthermore, the vector backbones pRS414 (SEQ ID 6, Trp1 selection marker), pRS415 (SEQ ID 7, Leu2 selection marker) and pRS416 (SEQ ID 8, Ura3 selection marker) were synthesized by DNA2.0. The three so obtained vectors were digested with the restriction enzymes XhoI and NotI. The linearized vectors were in analogy to the pathway expression cassettes purified from agarose gels.

3. Generation of S. cerevisiae expression plasmids and transformation of S. cerevisiae with adipic acid pathway genes

In vivo homologous recombination in *S. cerevisiae* was done essentially as described in Kazuko Lida, Tomoko Tada, Hidetoshi Iida, Molecular cloning in yeast by in vivo homologous recombination of the yeast putative a1 subunit of the voltage-gated calcium channel (2004), FEBS Letters, vol. 576, 291-296, except that more than one fragments were inserted in the vector. The final expression vectors were assembled by transformation of the linearized DNA fragments. All transformation were done with strain S. cerevisiae CENPK2-1C (genotype MATa; ura3-52; trp1-289; leu2-3,112; his3Δ 1; MAL2-8^{C}; SUC2). For the construction of plasmid pADI154, the linearized vector backbone SEQ ID 6 and the expression cassettes for the genes adi21 (SEQ ID 9), adi 22 (SEQ ID 10) and adi23 (SEQ ID 11) were cotransformed. In analogy, to yield plasmid pADI155, linearized vector backbone SEQ ID 7 was used in combination with

the expression cassettes of genes adi8 (SEQ ID 12), adi24 (SEQ ID 13), and adi25 (SEQ ID 14).

Transformants were plated on Yeast Nitrogen Base (YNB) w/o AA (Difco) + 2% glucose + addition of compounds to overcome the remaining auxotrophies (addition of leucin and histidine for generation of pADI154, and addition of Tryptophane and histidine for the assembly of pADI155). Cultivation in liquid medium was carried out using Verduyn medium (C. Verduyn, E. Postma, "Effect of Benzoic Acid on Metabolic Fluxes in Yeasts: A Continuous-Culture Study on the Regulation of Respiration and Alcoholic Fermentation", (1992) *Yeast*, vol. 8, 501-517). When transformants were identified based on selective growth, the plasmids in S. cerevisiae were isolated using the ZymoResearch Yeast Plasmid Isolation Kit (ZymoResearch Corporation, USA). The so obtained plasmids were subsequently transformed in E. Coli Top10 or XL1-blue cells and subsequently purified out of E. coli. This step ensured a high quality of the plasmid DNA. Finally, the plasmids were sequenced to ensure that only correct DNA sequences are transformed and used for further studies. The sequenced and verified plasmids pAdDI154 and pADI155 were retransformed in S. cerevisiae strains. The strains were restreaked and purified by isolating single colonies. The strains were resteaked on the same type of plates as the ones on which the transformants were initially plated. The single colonies were cultivated in the following medium (Verduyn, 1992) in 24 microtiter plate: (NH4)2SO4, 5 g; KH2PO4, 3 g; MgS04·7H2O, 0.5 g; EDTA, 15 mg; ZnS04·7H2O, 4.5 mg; CoCl2·6H2O, 0.3 mg; MnCl2·4H20, 1 mg; CuSO4·5H2O, 0.3 mg; CaCl2·2H20, 4.5 mg; H3BO3, 1 mg; KI, 0.1 mg; and 0.025 ml silicone antifoam (BDH). Filter-sterilized vitamins were addedafter heat sterilization (20°C) of this medium. Final vitamin concentrations per litre were: biotin, 0.05 mg; calcium pantothenate, 1 mg; nicotinic acid, 1 mg; inositol, 25 mg; thiamine HCl, 1 mg; pyridoxine HCl, 1 mg; and para-aminobenzoic acid, 0.2 mg. For the first biomass phase, (72 hours), 4% galactose was added, for the production phase (96 hours) 8% galactose and 1% calcium carbonate was added. Supernatants were separated and analyzed on adipic acid content using LC-MS.

The following LCMS conditions were applied: column, Waters Acquity UPLC HSS T3, 1.8 µm, 30 mm*2.1 mm; flow, 1 mL/min; temperature, 60ºC. Mass spectrometry ESI in negative mode. Mobile phase A: 0.1 % formic acid in water. Injection volume 5 µl, full loop. Mobile Phase B: 0.1 % formic acid in acetonitrile.

Gradient:

Time (min.)	0,0	0,5	1,0	1,2	1,2	1,6	1,6	2,0
%A	99,9	98,5	90,0	90,0	20,0	20,0	99,9	99,9
%B	0,1	1,5	10,0	10,0	80,0	80,0	0,1	0,1

Adipic acid eluted from the column at 0.98 min elution time. Adipic acid production is presented in Table 1 (duplicate colonies).

Table 1. Adipic acid production

Strain	Adipic acid (mg/L)
S. cerevisiae CENPk2-1c (untransformed, negative control), colony #1	ND
S. cerevisiae CENPk2-1c (untransformed, negative control), colony #2	ND
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pRS416, colony #1)	1,3
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pRS416, colony #2)	1,3

ND, not detected

4. Transformation of S. cerevisiae with adipic acid pathway genes including transformation with polynucleotides encoding succinyl-CoA ligase

Plasmid pADI156 was constructed by recombination between the linearized vector pRS416 backbone SEQ ID 8 and the expression cassettes for the genes sucC (SEQ ID 1), sucD (SEQ ID 2) (sucC and sucD are endogenous E. coli genes), and Lin1129 (SEQ ID 5; Lin1129 corresponds to a gene from *Listeria innocua* encoding acetaldehyde dehydrogenase). SEQ ID's 1, 2, 5, and 8 were cotransformed in S. cerevisiae CENPk2-1c (genotype MATa; ura3-52; trp1-289; leu2-3,112; his3Δ 1; MAL2-8^{C}; SUC2). Plasmid pADI156 was isolated from S. cerevisiae and sequence verified. Plasmid pADI164 was constructed by removing the Lin1129 expression cassette from pADI156 using standard restriction site cloning and was re-transformed in S. cerevisiae strain CENPk2-1c. The resulting transformants contained all three plasmids pADI154, pADI155 and pADI164. As medium for transformation, Verduyn medium was used, containing tryptophane, uracil and leucine, depending on the auxotrophies present. When all three plasmids pADI154, pADI155 and pADI164 were transformed, none of

the three compounds were added. The strain was cultivated and the supernatant of the broth was analyzed for adipic acid production as described in Example 3.

Plasmid pADI157 was constructed by the linearized vector pRS416 backbone (SEQ ID 8) and the expression cassettes for the genes Lsc1 (SEQ ID 3), Lsc2 (SEQ ID 4) and Lin1129 (SEQ ID 5). Lin1129 corresponds to a gene from *Listeria innocua* encoding acetaldehyde dehydrogenase. Lsc1 and Lsc2 are endogenous Saccharomyces genes. The S. cerevisiae succinyl-CoA ligase is localized in the mitochondria. In order to localize the S. cerevisiae succinyl-CoA ligase in the cytosol, the mitochondrial targeting sequence (MTS) of the S. cerevisiae succinyl-CoA ligase subunits was removed. The length of the MTS of proteins encoded by the LSC1 and LSC2 genes were predicted by TargetP 1.1 server (http://www.cbs.dtu.dk/services/TargetP/) according to Emanuelsson et al. 2007. Nature Protocols 2: 953-971, using Non-plant organism group. The MTS of the Lsc1 protein is 24 amino acid at the N-terminus. The MTS of the Lsc2 protein is 30 amino acids at the N-terminus. Subsequently, the MTS of both proteins, with the exception of the N-terminus methionine, was removed, resulting in the protein sequences represented by seq ID 17 and 18, respectively. Vector pRS416 and SEQ ID's 3, 4, and 5 were cotransformed in S. cerevisiae CENPk2-1c (genotype MATa; ura3-52; trp1-289; leu2-3,112; his3Δ 1; MAL2-8^C; SUC2). Plasmid pADI157 was isolated from S. cerevisiae and the sequence was verified. Plasmid pADI165 was constructed by removing the Lin1129 expression cassette from pADI157 using standard restriction site cloning and re-transformed into S. cerevisiae CENPk2-1c. The resulting transformants contained all three plasmids pADI154, pADI155 and pADIS165. As a medium for transformation, Verduyn medium was used, containing tryptophane, uracil and leucine, depending on the auxotrophies. When all three plasmids pADI154, pADI155 and pADI165 were transformed, none of the three compounds were added. The strain was cultivated and the supernatant of the broth was analyzed for adipic acid production as described in Example 3. Adipic acid production is presented in Table 2.

Table 2. Adipic acid production

Strain	Adipic acid (mg/L)
S. cerevisiae CENPk2-1c (transformed with pADI154 and pADI155 and pRS416, colony #1)	0,13
S. cerevisiae CENPk2-1c (transformed with pADI154 and pADI155	0,13

and pRS416, colony #2)	
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pADI164, colony #1)	0,8
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pADI164, colony #2)	0,8
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pADI164, colony #3)	0,8
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pADI165, colony #1)	0,2
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pADI165, colony #2)	0,2
S. cerevisiae CENPk2-1c (transformed with pADI154, pADI155 and pADI165, colony #3)	0,2

Surprisingly, strains harbouring adipic acid pathway genes and which were transformed with yeast succinyl-CoA Lsc1 and Lsc2 did not produce more adipic acid than strains which were not transformed with Lsc1 and Lsc2. However, when the yeast cells were transformed with the prokaryotic succinyl-CoA ligase genes sucC and sucD, the production of adipic acid was improved.

CLAIMS

1. Method to improve the production of a succinyl-CoA derived compound by a eukaryotic cell, which is capable of producing a succinyl-CoA derived compound, comprising transforming said cell with a polynucleotide encoding a succinyl-CoA ligase.
2. Method according to claim 1 wherein the succinyl-CoA ligase is active in the cytosol of the cell.
3. Method according to claim 1 or 2 wherein the succinyl-CoA ligase is prokaryotic.
4. Method according to any one of claim 1-3 wherein the succinyl-CoA ligase comprises an amino acid sequence comprising SEQ ID's 15 and/or 16 and/or homologues thereof.
5. Method according to any one of claim 1-4 wherein the succinyl-CoA derived compound is adipic acid.
6. Method according to any one of claim 1-5 wherein the eukaryotic cell is yeast.
7. An engineered eukaryotic cell transformed with a polynucleotide encoding a succinyl-CoA ligase obtainable by a process according to any of claims 1-6.
8. An engineered eukaryotic cell according to claim 7 further comprising an adipic acid biosynthetic pathway.
9. Engineered eukaryotic cell according to claim 8 wherein the adipic acid biosynthetic pathway comprises an enzyme selected from the group consisting of an enzyme capable of acyl group transfer, an enzyme capable of catalysing the reduction of a carbon-carbon double bond of a 2,3-enoate moiety or a 2-enoyl moiety, an enzyme capable of catalysing the dehydration of a 3-hydroxyacyl ester or 3-hydroxyacyl thioester to a 2-enoyl ester or thioester, an enzyme capable of catalysing the reduction of a carbonyl group to an alcohol group or capable of catalysing the reduction of a 3-oxoacyl ester or 3-oxoacyl thioester to the corresponding 3-hydroxyacyl ester or thioester, and an enzyme capable of converting an adipic acid ester or adipic acid thioester into adipic acid.
10. Method to produce adipic acid or adipic acid ester or adipic acid thioester comprising culturing the engineered eukaryotic cell according to any one of claims 7-9.

11. Engineered eukaryotic cell according to any one of claims 7-9 further comprising an enzyme capable of converting an adipic ester, adipic acid ester or adipic acid thioester into 5-formylpentanoate.
12. Method to produce 5-formylpentanoate by culturing the cell according to claim 11.
13. Engineered eukaryotic cell according to claim 11 further comprising an enzyme capable of converting 5-formylpentanoate into 6-amino caproic acid.
14. Method to produce 6-amino caproic acid comprising culturing the engineered cell according to claim 11.
15. Method to produce caprolactam comprising preparing 6-amino caproic acid according to the method according to claim 12 and cyclising the 6-amino caproic acid, thereby forming caprolactam.

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/073765

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/00 C12N5/10
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/115840 A2 (GENOMATICA INC [US]; BURK MARK J [US]; VAN DIEN STEPHEN J [US]; BURGAR) 25 September 2008 (2008-09-25)	1-4,6-13
A	page 2, lines 7-12 page 8, paragraph 2 page 10, lines 9-12 page 13, paragraph 2 page 62, lines 4-6 page 75, lines 15-18 figure 2	5

A	US 5 487 987 A (FROST JOHN W [US] ET AL) 30 January 1996 (1996-01-30) claim 1	1-15

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 February 2012	27/02/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Barnas, Christoph

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2011/073765

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☒ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purpose of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2011/073765

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008115840 A2	25-09-2008	AU 2008229076 A1	25-09-2008
		BR PI0808988 A2	08-11-2011
		CA 2678946 A1	25-09-2008
		CN 101668861 A	10-03-2010
		EP 2137315 A2	30-12-2009
		JP 2010521182 A	24-06-2010
		TW 200911985 A	16-03-2009
		US 2009075351 A1	19-03-2009
		WO 2008115840 A2	25-09-2008
US 5487987 A	30-01-1996	AU 687874 B2	05-03-1998
		CA 2171327 A1	23-03-1995
		EP 0722504 A1	24-07-1996
		FI 961234 A	15-03-1996
		JP H09505463 A	03-06-1997
		NZ 273231 A	24-11-1997
		US 5487987 A	30-01-1996
		US 5616496 A	01-04-1997
		WO 9507996 A1	23-03-1995