MUT- METHYLOTROPHIC YEAST

TECHNICAL FIELD

The invention refers to production of a protein of interest (POI) in a recombinant methylotrophic yeast which is deficient in alcohol oxidase 1 (AOX1) and alcohol oxidase 2 (AOX2).

BACKGROUND

Proteins produced in recombinant host cell culture have become increasingly important as diagnostic and therapeutic agents. For this purpose, cells are engineered and/or selected to produce unusually high levels of a recombinant or heterologous protein of interest. Optimization of cell culture conditions is important for successful commercial production of recombinant or heterologous proteins.

Successful production of proteins of interest (POI) has been accomplished both with prokaryotic and eukaryotic host cells in cell culture. Eukaryotic host cells, in particular mammalian host cells, yeasts or filamentous fungi, or bacteria are commonly used as production hosts for biopharmaceutical proteins as well as for bulk chemicals. The most prominent examples are methylotrophic yeasts like such as *Pichia pastoris,* which is well reputed for efficient secretion of heterologous proteins. *P. pastoris* has been reclassified into a new genus, *Komagataella*, and split into three species, *K. pastoris*, *K. phaffii*, and *K. pseudopastoris*. Strains commonly used for biotechnological applications belong to two proposed species, *K. pastoris* and *K. phaffii*. The strains GS115, X-33, CBS2612, and CBS7435 are *K. phaffii*, while the strain DSMZ70382 is classified into the type species, *K. pastoris*, which is the reference strain for all the available *P. pastoris* strains (Kurtzman 2009, J Ind Microbiol Biotechnol. 36(11):1435-8). Mattanovich et al. (Microbial Cell Factories 2009, 8:29 doi:10.1186/1475-2859-8-29) describe the genome sequencing of the type strain DSMZ 70382 of *K. pastoris*, and analyzed its secretome and sugar transporters.

P. pastoris strains have been used which are deficient in both AOX genes, AOX1 and AOX2 (referred to as Mut-), or deficient in only AOX2 (referred to as Mut^S), or not deficient in any of the AOX genes (referred to as Mut^+).

Promoters used for protein production in recombinant host cells are either regulated (*e.g.*, induced upon addition of methanol to the medium, methanol-controlled), or constantly active (constitutive). The methanol inducible promoter pAOX1 has been described to control protein expression in Mut-, Mut^+ or Mut^S strains.

Chiruvolu et al. (Enzyme Microb. Technol. 1997, 21:277-283) describe the construction of a Mut- strain used for recombinant protein production. It was determined that the Mut- strain did not grow on methanol, which necessitated the use of another carbon source to provide for growth, maintenance and protein production. A POI has been expressed under the control of pAOX1, which has been induced by injection of methanol into the fermentor to maintain a concentration of about 0.5% (v/v).

Chauhan et al. (Process Biochemistry 1999, 34:139-145) describe utilization of a AOX1 deleted host (designated as Mut-, however, understood to be Mut^S) carrying a gene for expressing HBsAg under the pAOX1 promoter. Protein expression was inducted by methanol, but a high methanol concentration in the broth was found to be toxic to the cells, because the Mut^S cells were found to be sensitive to the methanol concentration.

Karaoglan et al. (Biotechnol. Lett. 1995, DOI 10.1007/s10529-015-1993-z) describe the functional analysis of alcohol dehydrogenase (ADH) genes in P. pastoris. ADH3 (XM_002491337) and ADH (FN392323) genes were disrupted. The double knockout strain also produced ethanol. It is concluded that the ADH gene does not play a role in ethanol metabolism; and *PpADH3* was the only gene responsible for consumption of ethanol in *P. pastoris*.

Singh and Narang (bioRxiv, DOI:10.1101/573519, preprint) describe β-galactosidase expression in Mut^+, a Mut^s (AOX1-) and Mut^- (AOX1- AOX2-) strains of *Komagataella phaffii* (*Pichia pastoris*). It was concluded that formate or/and formaldehyde are probably true inducers since both induce P_{AOX1}expression in Mut^- which cannot synthesize intracellular methanol from formate or formaldehyde, and propose formate as a promising substitute for methanol since it does not appear to suffer from the deficiencies that afflict methanol.

Wei Shen et al. (Microbial Cell Factories 2016, 15(1):1-11) describe a methanol-free *Pichia pastoris* protein expression system. Two kinase mutants, Δgut1 and Δdak, showed strong alcohol oxidase activity under non-methanol carbon sources and were used to construct methanol-free expression systems.

Ia Pla et al. (Biotechnol Prog, 2006, pp 881-888) describe Mut^S and Mut+ *P. pastoris* strains for expressing scFv using AOX promoters and induction by methanol.

EP1905836A1 discloses P. pastoris strains for producing recombinant human interferon alpha, and suggests using an AOX1-disturbed clone which comprises a mutated AOX1 gene, yet no complete deletion of the AOX1 locus.

Ching-Hsiang Chang et al. (BMC Biotechnology 2018, 18(1):81) suggest a flexible pAOX1 induction system in *P. pastoris* using methanol expression regulator 1 (Mxr1)-reprogrammed cells.

Russmayer H. et al. (BMC Biology 2015, 13(1):80) describe the importance of AOX in regulating *P. pastoris* expression systems.

Tomas-Gamisans M. et al. (Microbial Biotechnology 2018, 11(1):224-237) describe a *P. pastoris* genome-scale metabolic model for improved prediction on methanol or glycerol as sole carbon sources.

Moser et al. (Microbial Cell Factories 2017, 16(1):49) describe adaptive laboratory evolution to improve growth and recombinant protein production in *P. pastoris*.

Recombinant protein production in *P. pastoris* requires an intense process scheme, leading to high oxygen demand, and heat production, demanding a high biomass concentration and methanol consumption. Oxygen transfer, cooling and biomass separation in downstream processing is expensive. It is thus desirable to develop processes of low oxygen demand and heat production. It is further desirable to increase the yield of protein production, in particular by efficient use of a carbon source.

SUMMARY OF THE INVENTION

It is the objective of the invention to improve recombinant protein production in methylotrophic yeast.

The objective is solved by the subject of the claims and as further described herein.

The invention provides for a recombinant methanol utilization pathway deficient methylotrophic yeast (Mut-) host cell which is engineered:

a) by one or more genetic modifications to render the host cell methanol utilization pathway deficient, by reducing expression of a first and a second endogenous gene compared to the host cell prior to said one or more genetic modifications, wherein

i. the first endogenous gene encodes alcohol oxidase 1 (AOX1) comprising the amino acid sequence identified as SEQ ID NO:1 or a homologue thereof that is endogenous to the respective wild-type host cell, and

ii. the second endogenous gene encodes alcohol oxidase 2 (AOX2) comprising the amino acid sequence identified as SEQ ID NO:3 or a homologue thereof that is endogenous to the respective wild-type host cell, and

b) by one or more genetic modifications to increase expression of an alcohol dehydrogenase (*ADH2*) gene compared to the host cell prior to said one or more genetic modifications, wherein the *ADH2* gene encodes an alcohol dehydrogenase (ADH2).

Specifically, the ADH2 protein is an alcohol dehydrogenase classified as EC 1.1.1.1.

As described herein, the term “ADH2” shall refer to either a native alcohol dehydrogenase, such as *P. pastoris* alcohol dehydrogenase comprising or consisting of the amino acid sequence identified as SEQ ID NO:50 (UniProtKB - F2QSX6_KOMPC; FR839629 Genomic DNA Translation: CCA38504.1; gene: PP7435_Chr2-0821), or a sequence which has a certain homology (or sequence identity) to SEQ ID NO:50.

Specifically, the ADH2 may originate from a *P. pastoris* strain or may be a homologue or ortholog thereof which is naturally-occurring originating from or endogenous to a wild-type cell of an organism, such as eukaryotes, including *e.g*. yeast, in particular of a methylotrophic yeast strain, species, or genus, or which is a mutant of such naturally-occurring ADH2.

Specifically, the ADH2 protein is naturally-occurring in or endogenous to the species of the host cell, or a mutant thereof.

Specifically, the gene encoding the *P. pastoris* alcohol dehydrogenase, herein referred to as *ADH2* gene, comprises or consists of the nucleotide sequence identified as SEQ ID NO:51, or a homologous polynucleotide (gene) encoding ADH2, which ADH2 has a certain homology (or sequence identity) to SEQ ID NO:50.

Specifically, the *ADH2* gene is endogenous or heterologous to the Mut- host cell. Specifically, the Mut- host cell comprises one or more copies of said *ADH2* gene.

Specifically, the ADH2 is any one of:

a) an ADH2, which is *P. pastoris* ADH2 comprising the amino acid sequence identified as SEQ ID NO:50, or a homologue thereof that is endogenous to a yeast species, in particular methylotrophic yeast; or

b) a mutant of the ADH2 of a), which is at least 60% identical to SEQ ID NO:50.

The homologous sequences are also referred to as ADH2 homologue or *ADH2* homologue.

Exemplary homologues are described in Figure 1:

SEQ ID NO:52: ADH2 amino acid sequence of *Komagataella pastoris*, ATCC 28485
SEQ ID NO:53: *ADH2* gene sequence of *Komagataella pastoris*, ATCC 28485
SEQ ID NO:54: ADH2 amino acid sequence of *Ogataea parapolymorpha,* DL-1
SEQ ID NO:55: *ADH2* gene sequence of *Ogataea parapolymorpha,* DL-1
SEQ ID NO:56: ADH2 amino acid sequence of *Ogataea parapolymorpha,* DL-1
SEQ ID NO:57: *ADH2* gene sequence of *Ogataea parapolymorpha,* DL-1
SEQ ID NO:58: ADH amino acid sequence of *Ogataea polymorpha,* NCYC 495 leu1.1
SEQ ID NO:59: *ADH* gene sequence of *Ogataea polymorpha,* NCYC 495 leu1.1
SEQ ID NO:60: ADH amino acid sequence *Ogataea polymorpha,* NCYC 495 leu1.1
SEQ ID NO:61: *ADH* gene sequence of *Ogataea polymorpha,* NCYC 495 leu1.1
SEQ ID NO:62: ADH2 amino acid sequence of *Saccharomyces cerevisiae,* YJM627
SEQ ID NO:63: *ADH2* gene sequence of *Saccharomyces cerevisiae,* YJM627
SEQ ID NO:64: ADH2 amino acid sequence of *Candida maltosa,* Xu316
SEQ ID NO:65: *ADH2* gene sequence of *Candida maltosa,* Xu316
SEQ ID NO:66: ADH4 amino acid sequence of *Kluyveromyces marxianus*, DMKU3-1042
SEQ ID NO:67: *ADH4* gene sequence of *Kluyveromyces marxianus*, DMKU3-1042
SEQ ID NO:68: ADH1 amino acid sequence of *Escherichia coli*, 7.1982
SEQ ID NO:69: *ADH1* gene sequence of *Escherichia coli*, 7.1982
SEQ ID NO:70: ADH1 amino acid sequence of *Fusarium graminearum*, PH-1
SEQ ID NO:71: *ADH1* gene sequence of *Fusarium graminearum*, PH-1

Specifically, the ADH2 homologue has at least any one of 60%, 70%, 80%, 85%, 90 %, or 95% sequence identity to SEQ ID NO:50. An *ADH2* homologue is herein understood to encode an ADH2 homologue. Specifically, sequence identity is determined as further disclosed herein, for example when comparing the full-length sequence.

Specifically, the homologue or homologous sequence is characterized by the same qualitative function of the ADH2 protein in a wild-type host cell such as in *P. pastoris*, in particular *K. pastoris* or *K. phaffii e.g*., as alcohol dehydrogenase (EC 1.1.1.1.).

Specifically, the homologous sequences of SEQ ID NO:50 is of a species other than *P. pastoris, in particular K. pastoris or K. phaffii e.g*., another yeast of the Komagataella or Pichia genus, and expression of the respective endogenous coding sequences increased (*e.g*., knocked in) as described herein.

If the host cell is of *P. pastoris, in particular K. pastoris or K. phaffii,* the ADH2 protein may comprise or consist of the endogenous sequence *e.g*., SEQ ID NO:50, or a homologue to SEQ ID NO:50 of a different strain or species, or an artificial sequence of a mutant ADH2 protein, which is not naturally-occurring in a wild-type strain or organism, in particular which is not naturally-occurring in a methylotrophic yeast.

According to a specific example, the homologous sequence has at least any one of 0.2-fold, 0.3-fold, 0.4-fold, 0.5.fold, 0.6-fold, 0.7-fold, 0.8-fold, 0.9-fold, 1-fold, 1.1-fold, 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2-fold, or even higher alcohol dehydrogenase activity as compared to the activity of the endogenous naturally-occurring *P. pastoris* alcohol dehydrogenase in a wild-type *P. pastoris*. The alcohol dehydrogenase activity can be measured in a suitable assay *e.g*., by an alcohol dehydrogenase assay using cell free extracts. Cell free extracts can be obtained by mechanical disruption of the cell culture by zirconia/silica/glass beads as described by Karaoglan et al. (Biotechnol Lett. 2016; 38(3):463-9). The alcohol dehydrogenase activity can be measured following the formation of NADH by measuring the absorption increase at a wavelength of 340 nm as described by Walker (Biochemical Education. 1992 21(1):42-43). NAD^+ and alcohol can be used as substrates and are consumed in equimolar concentrations. NADH production is inversely correlated with NAD^+ consumption. Alternatively, a commercial colorimetric alcohol dehydrogenase activity assay kit can be used (MAK053, Sigma-Aldrich). An exemplary assay is herein described in the examples section.

As described herein, the term “AOX1” shall refer to either a native alcohol oxidase 1, such as *P. pastoris* alcohol oxidase 1 comprising or consisting of the amino acid sequence identified as SEQ ID NO:1 (UniProtKB - F2QY27), or a sequence which has a certain homology (or sequence identity) to SEQ ID NO:1, which may be a homologue of the *P. pastoris* alcohol oxidase 1 that is endogenous to a methylotrophic yeast species, in particular which is endogenous to the methylotrophic yeast herein used as a host cell, prior to said one or more genetic modifications to reduce expression of said endogenous alcohol oxidase 1. In particular, the AOX1 protein is an ortholog that is endogenous to the species of the host cell species.

Specifically, the gene encoding the *P. pastoris* alcohol oxidase 1, herein referred to as *AOX1* gene, comprises or consists of the nucleotide sequence identified as SEQ ID NO:2.

The homologous sequences are also referred to as AOX1 homologue or *AOX1* homologue. Specifically, the AOX1 protein or AOX1 homologue is an alcohol oxidase enzyme classified as EC 1.1.3.13.

Specifically, the AOX1 homologue has at least any one of 60%, 70%, 80%, 85%, 90%, or 95% sequence identity to SEQ ID NO:1. An *AOX1* homologue is herein understood to encode an AOX1 homologue. Specifically, sequence identity is determined as further disclosed herein, for example when comparing the full-length sequence.

As described herein, the term “AOX2” shall refer to either a native alcohol oxidase 2, such as *P. pastoris* alcohol oxidase 2 comprising or consisting of the amino acid sequence identified as SEQ ID NO:3 (UniProtKB - F2R038), or a sequence which has a certain homology (or sequence identity) to SEQ ID NO:3, which may be a homologue of the *P. pastoris* alcohol oxidase 2 that is endogenous to a methylotrophic yeast species, in particular which is endogenous to the methylotrophic yeast herein used as a host cell, prior to said one or more genetic modifications to reduce expression of said endogenous alcohol oxidase 2. In particular, the AOX2 protein is an ortholog that is endogenous to the species of the host cell species.

Specifically, the gene encoding the *P. pastoris* alcohol oxidase 2, herein referred to as *AOX2* gene, comprises or consists of the nucleotide sequence identified as SEQ ID NO:4.

The homologous sequences are also referred to as AOX2 homologue or *AOX2* homologue. Specifically, the AOX2 protein or AOX2 homologue is an alcohol oxidase enzyme classified as EC 1.1.3.13.

Specifically, the AOX2 homologue has at least any one of 60%, 70%, 80%, 85%, 90%, or 95% sequence identity to SEQ ID NO:3. An *AOX2* homologue is herein understood to encode an AOX2 homologue. Specifically, sequence identity is determined as further disclosed herein, for example when comparing the full-length sequence.

Specifically, the AOX1 and/or AOX2 proteins are of *P. pastoris* origin, in particular *K. pastoris* or *K. phaffii* origin, if the host cell is *P. pastoris*, in particular *K. pastoris* and *K. phaffii*, respectively. Alternatively, each of the AOX1 and AOX2 proteins comprises a

homologous (or orthologous) sequence of the respective protein of in *P. pastoris*, in particular *K. pastoris* or *K. phaffii,* origin, which homologous (orthologous) sequence is endogenous to a wild-type host cell, if of another origin or species. For example, if the host cell is *K. phaffii*, an endogenous AOX1 or AOX2 protein comprises or consists of the amino acid sequence identified as SEQ ID NO:1 and SEQ ID NO:3, respectively.

According to another example, if the host cell is *K. pastoris*, the endogenous AOX1 protein comprises or consists of the amino acid sequence identified as SEQ ID NO:9 (Komagataella pastoris, ATCC 28485), which is 99.85% identical to SEQ ID NO:1.

According to another example, if the host cell is *K. pastoris*, the endogenous AOX2 protein comprises or consists of the amino acid sequence identified as SEQ ID NO:11, which is 99.40% identical to SEQ ID NO:3.

Exemplary homologues are described in Figure 1:

SEQ ID NO:9: AOX1 amino acid sequence of Komagataella pastoris, ATCC 28485
SEQ ID NO:10: *AOX1* nucleotide sequence of Komagataella pastoris, ATCC 28485
SEQ ID NO:11: AOX2 amino acid sequence of Komagataella pastoris, ATCC 28485
SEQ ID NO:12: *AOX2* nucleotide sequence of Komagataella pastoris, ATCC 28485
SEQ ID NO:13: MOD1 amino acid sequence of Ogataea methanolica JCM 10240
SEQ ID NO:14: *MOD1* nucleotide sequence of Ogataea methanolica JCM 10240
SEQ ID NO:15: MOD2 amino acid sequence of Ogataea methanolica JCM 10240
SEQ ID NO:16: *MOD2* nucleotide sequence of Ogataea methanolica JCM 10240
SEQ ID NO:17: pMOD1 promoter sequence of Ogataea methanolica JCM 10240
SEQ ID NO:18: pMOD2 promoter sequence of Ogataea methanolica JCM 10240
SEQ ID NO:19: MOX amino acid sequence of Ogataea polymorpha NCYC 495 leu 1.1
SEQ ID NO:20: *MOX* nucleotide sequence of Ogataea polymorpha NCYC 495 leu 1.1
SEQ ID NO:21: pMOX promoter sequence of Ogataea polymorpha NCYC 495 leu 1.1

Yet, if the host cell is of a different species (other than *K. pastoris* and/or *K. phaffii*), the AOX1 or AOX2 protein sequence which is endogenous to the host cell is a homologue to SEQ ID NO:1 and SEQ ID NO:3, respectively, and expression of such homologue in the host cell (the orthologous sequence of SEQ ID NO:1 and SEQ ID NO:3, respectively) is reduced for the purpose described herein.

Specifically, any or each of the homologous sequences is characterized by the same qualitative function of the respective AOX1 and AOX2 protein in a wild-type host cell such as in *P. pastoris*, in particular *K. pastoris* or *K. phaffii e.g.*, as alcohol oxidase (EC 1.1.3.13).

Specifically, the respective homologous sequences of SEQ ID NO:1 and SEQ ID NO:3 are of a species other than *P. pastoris, in particular K. pastoris or K. phaffii e.g*., another yeast of the Komagataella or Pichia genus, and expression of the respective endogenous coding sequences reduced or abolished (knocked out) as described herein.

Specifically, both, the AOX1 and AOX2 proteins, are endogenous to the host cell, and the expression of the genes encoding AOX1 and AOX2, respectively, is reduced or deleted.

Specifically, both, the AOX1 and AOX2 proteins, are of the same origin, originating from or endogenous to the same host cell (or host cell species) prior to its engineering for reducing expression of said first and second endogenous genes.

Specifically, both, the AOX1 and AOX2 proteins, are of *P. pastoris* origin, in particular proteins encoded by a respective gene that is endogenous to the host cell, wherein the host cell is *P. pastoris*.

Specifically, both, the AOX1 and AOX2 proteins, are of *Komagataella phaffii* origin, in particular proteins encoded by a respective gene that is endogenous to the host cell, wherein the host cell is *Komagataella phaffii.*

Specifically, both, the AOX1 and AOX2 proteins, are of *Komagataella pastoris* origin, in particular proteins encoded by a respective gene that is endogenous to the host cell, wherein the host cell is *Komagataella pastoris.*

According to a specific aspect, said one or more genetic modifications comprises a disruption, substitution, deletion, knockin or knockout of (i) one or more polynucleotides, or a part thereof; or (ii) an expression control sequence.

According to a specific aspect, said one or more genetic modifications are of one or more endogenous polynucleotides of the host cell described herein, such as coding polynucleotides, including *e.g*., said polynucleotide (or gene) encoding the respective AOX1, AOX2, or ADH2 protein, in particular the wild-type (unmodified or native) protein, which is naturally-occurring in the host cell species, type or strain, or a nucleotide sequence controlling expression of said polynucleotide (or gene).

According to a specific aspect, said one or more genetic modifications are of an expression control sequence, including *e.g*., a promoter, ribosomal binding site, transcriptional or translational start and stop sequences, or of an enhancer or activator sequence.

A variety of methods of engineering a host cell can be employed to modulate (reduce or increase) expression of an endogenous polynucleotide, such as

a) to reduce expression of a gene encoding the respective AOX1 or AOX2 protein, including *e.g*., disrupting the polynucleotide encoding the respective AOX1 or AOX2 protein, disrupting the promoter which is operably linked to such polynucleotide, replacing such promoter with another promoter which has lower promoter activity; or

b) to increase expression of a gene encoding the ADH2 protein, including *e.g*., introducing a polynucleotide encoding the ADH2 protein into the host cell genome, disrupting the promoter which is operably linked to such polynucleotide, replacing such promoter with another promoter which has higher promoter activity.

Specific methods of modifying gene expression employ modulating (*e.g*., activating, up-regulating, inactivating, inhibiting, or down-regulating) regulatory sequences which modulate the expression of a polynucleotide (a gene), such as using respective transcription regulators targeted to the relevant sequences by an RNA guided ribonuclease used in a CRISPR based method of modifying a host cell, *e.g*., regulatory sequences selected from the group consisting of promoter, ribosomal binding sites, transcriptional start or stop sequences, translational start or stop sequences, enhancer or activator sequences, repressor or inhibitor sequences, signal or leader sequences, in particular those which control the expression and/or secretion of a protein.

According to a specific aspect, said one or more genetic modifications include a gain-of-function alteration in the *ADH2* gene resulting in increasing the level or activity of ADH2.

Specifically, said gain-of-function alteration includes a knockin of the *ADH2* gene.

Specifically, said gain-of-function alteration up-regulates the *ADH2* gene expression in said cell.

Specifically, said gain-of-function alteration includes an insertion of a heterologous expression cassette to overexpress the *ADH2* gene in said cell.

Specifically, said heterologous expression cassette comprises a heterologous polynucleotide comprising an *ADH2* gene under the control of a promoter sequence. Such promoter can be any of a constitutive, repressible or inducible promoter.

Specifically, said one or more genetic modifications to increase expression of a gene include one or more genomic mutations including insertion or activation of a gene or genomic sequence which increases expression of a gene or part of a gene by at least 50%, 60%, 70%, 80%, 90%, or 95%, or even more *e.g*., by a knockin of a heterologous gene, or increasing the copy number of the endogenous gene, as compared to the respective host without such genetic modification.

Specifically, the one or more genetic modifications increasing expression comprise genomic mutations which constitutively improve or otherwise increase the expression of one or more endogenous polynucleotides.

Specifically, the one or more genetic modifications increasing expression comprise genomic mutations introducing one or more inducible or repressible regulatory sequences which conditionally improve or otherwise increase the expression of one or more endogenous polynucleotides. Such conditionally active modifications are particularly targeting those regulatory elements and genes which are active and/or expressed dependent on cell culture conditions.

Specifically, the expression of the polynucleotide encoding the ADH2 protein is increased when using the host cell in a method of producing a protein of interest (POI). Specifically, upon genetic modification, expression of the ADH2 protein is increased under conditions of the host cell culture during which the POI is produced.

Specifically, the host cell is genetically modified to increase the amount (*e.g.*, the level, activity or concentration) of the ADH2 protein, by at least any one of 50%, 60%, 70%, 80%, 90%, or 95%, (mol/mol), or even more, compared to the host cell without said modification, *e.g.*, by a knockin of one or more respective *ADH2* genes. According to a specific embodiment, the host cell is genetically modified to comprise one or more insertions of (one or more) genomic sequences, in particular genomic sequences encoding the respective ADH2 protein, which are integrated in the host cell genome. Such host cell is typically provided as a knockin strain.

According to a specific embodiment, once the host cell described herein is cultured in a cell culture, the total amount of the ADH2 protein in the host cell or host cell culture is increased by at least any one of 50%, 60%, 70%, 80%, 90%, or 95%, (activity% or mol/mol), or even by 100% or more, compared to a reference amount expressed or produced by the host cell prior to or without such genetic modification, or compared to a reference amount produced in a respective host cell culture, or compared to the host cell prior to or without said modification.

Specifically, said one or more genetic modifications to reduce expression of a gene, such as the *AOX1* and/or *AOX2* genes, include one or more genomic mutations including deletion or inactivation of a gene or genomic sequence which reduces expression of a gene or part of a gene by at least 50%, 60%, 70%, 80%, 90%, or 95%, or even completely abolishes its expression, *e.g.*, by a knockout of the gene, as compared to the respective host without such genetic modification.

Specifically, the one or more genetic modifications reducing expression comprise genomic mutations which constitutively impair or otherwise reduce the expression of one or more endogenous polynucleotides.

Specifically, the one or more genetic modifications reducing expression comprise genomic mutations introducing one or more inducible or repressible regulatory sequences which conditionally impair or otherwise reduce the expression of one or more endogenous polynucleotides. Such conditionally active modifications are particularly targeting those regulatory elements and genes which are active and/or expressed dependent on cell culture conditions.

Specifically, the expression of said one or more endogenous polynucleotides is reduced thereby reducing expression of the polynucleotide encoding the respective AOX1 or AOX2 protein when using the host cell in a method of producing a protein of interest (POI). Specifically, upon genetic modification, expression of both, the AOX1 and AOX2 proteins, is reduced under conditions of the host cell culture during which the POI is produced.

Specifically, the host cell is genetically modified to reduce the amount (*e.g.*, the level, activity or concentration) of both, the AOX1 and AOX2 proteins, by at least any one of 50%, 60%, 70%, 80%, 90%, or 95%, (activity%, or mol/mol) compared to the host cell without said modification, or even by 100%, *e.g.* to a non-detectable amount, thereby completely abolishing production of both, the AOX1 and AOX2 proteins, *e.g.*, by a knockout of the respective *AOX1* and *AOX2* genes. According to a specific embodiment, the host cell is genetically modified to comprise one or more deletions of (one or more) genomic sequences, in particular genomic sequences encoding the respective AOX1 and/or AOX2 protein. Such host cell is typically provided as a deletion or knockout strain.

According to a specific aspect, said first and/or second endogenous gene is knocked out by said one or more genetic modifications. Specifically, the Mut- host cell is a Δ*AOX1*/Δ*AOX2* knockout strain.

According to a specific aspect, said first and/or second endogenous gene is knocked out by said one or more genetic modifications; and said *ADH2* gene is knocked in by said one or more genetic modifications. Specifically, the Mut- host cell is a Δ*AOX1*/Δ*AOX2* + ADH2-OE strain.

According to a specific embodiment, once the host cell described herein is cultured in a cell culture, the total amount of the respective AOX1 and/or AOX2 protein in the host cell or host cell culture is reduced by at least any one of 50%, 60%, 70%,

80%, 90%, or 95%, (mol/mol), or even by 100%, *e.g*. to a non-detectable amount, compared to a reference amount expressed or produced by the host cell prior to or without such genetic modification, or compared to a reference amount produced in a respective host cell culture, or compared to the host cell prior to or without said modification.

When comparing the host cell described herein for the effect of said genetic modification to increase or reduce production of the respective ADH2, AOX1 or AOX2 protein, it is typically compared to the comparable host cell prior to or without such genetic modification. Comparison is typically made with the same host cell species or type without (or prior to) such genetic modification, which is engineered to produce the recombinant or heterologous POI, in particular when cultured under conditions to produce said POI. However, a comparison can also be made with the same host cell species or type which is not further engineered to produce the recombinant or heterologous POI.

According to a specific aspect, the increase or reduction of the respective ADH2, AOX1 or AOX2 protein is determined by the increase or reduction of the amount (*e.g*., the level, activity or concentration) of the respective protein in the cell. Specifically, the amount of said protein is determined by a suitable method, such as employing a Western Blot, immunofluorescence imaging, flow cytometry or mass spectrometry, in particular wherein mass spectrometry is liquid chromatography–mass spectrometry (LC-MS), or liquid chromatography tandem-mass spectrometry (LC-MS/MS) *e.g*., as described by Doneanu et al. (MAbs. 2012; 4(1): 24–44). According to a specific example, alcohol dehydrogenase activity can be measured by an activity assay using cell free extracts. Cell free extracts can be obtained by mechanical disruption of the cell culture by zirconia/silica/glass beads as described by Karaoglan et al. (Biotechnol Lett. 2016; 38(3): 463-9). The alcohol dehydrogenase activity is measured by directly following the formation of NADH by measuring the absorption increase at a wavelength of 340 nm as described by Walker (Biochemical Education. 1992 21(1):42-43). NAD^+ and an alcohol are used as substrates and are consumed in equimolar concentrations. NADH production is inversely correlated with NAD^+ consumption. Alternatively a commercial colorimetric alcohol dehydrogenase activity assay kit can be used (MAK053, Sigma-Aldrich). Alcohol oxidase activity can be measured calorimetrically with the 2,2′-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid that reacts with hydrogen peroxide as described by Verduyn et al. (Journal of Microbiological Methods. 1984 (2)1: 15-25) or by

measuring the amount of formaldehyde formed as described by Couder and Baratti (Agric. Bioi. Chern. 1980; 44(10):2279-2289). A detailed assay is described herein in the examples section.

According to a specific aspect, the Mut- host cell comprises a heterologous gene of interest expression cassette (GOIEC) comprising an expression cassette promoter (ECP) operably linked to a gene of interest (GOI) encoding a protein of interest (POI).

According to a specific aspect, the Mut- host cell is a recombinant host cell comprising at least one heterologous GOIEC, wherein at least one component or combination of components comprised in the GOIEC is heterologous to the host cell. Specifically, an artificial expression cassette is used, in particular wherein the promoter and GOI are heterologous to each other, not occurring in such combination in nature *e.g*., wherein either one (or only one) of the promoter and GOI is artificial or heterologous to the other and/or to the host cell described herein; the promoter is an endogenous promoter and the GOI is a heterologous GOI; or the promoter is an artificial or heterologous promoter and the GOI is an endogenous GOI; wherein both, the promoter and GOI, are artificial, heterologous or from different origin, such as from a different species or type (strain) of cells compared to the host cell described herein. Specifically, the ECP is not naturally associated with and/or not operably linked to said GOI in the cell which is used as a host cell described herein.

Specifically, the GOIEC comprises a constitutive, inducible or repressible promoter.

Specific examples of constitutive promoter include *e.g*., the pGAP and functional variants thereof, any of the constitutive promoter such as pCS1, published in WO2014139608.

Specific examples of inducible or repressible promoter include *e.g*., the native pAOX1 or pAOX2 and functional variants thereof, any of the regulatory promoter, such as pG1-pG8, and fragments thereof, published in WO2013050551; any of the regulatory promoter, such as pG1 and pG1-x, published in WO2017021541 A1.

Suitable promoter sequences for use with yeast host cells are described in Mattanovich et al. (Methods Mol. Biol. (2012) 824:329-58) and include glycolytic enzymes like triosephosphate isomerase (TPI), phosphoglycerate kinase (PGK), glyceraldehyde-3- phosphate dehydrogenase (GAPDH or GAP) and variants thereof, lactase (LAC) and galactosidase (GAL), *P. pastoris* glucose-6-phosphate isomerase promoter (PPGI), the 3- phosphoglycerate kinase promoter (PPGK), the glycerol

aldehyde phosphate dehydrogenase promoter (pGAP), translation elongation factor promoter (PTEF), and the promoters of *P. pastoris* enolase 1 (PEN01), triose phosphate isomerase (PTPI), ribosomal subunit proteins (PRPS2, PRPS7, PRPS31, PRPL1), alcohol oxidase promoter (PAOX1, PAOX2) or variants thereof with modified characteristics, the formaldehyde dehydrogenase promoter (PFLD), isocitrate lyase promoter (PICL), alpha-ketoisocaproate decarboxylase promoter (PTHI), the promoters of heat shock protein family members (PSSA1, PHSP90, PKAR2), 6-Phosphogluconate dehydrogenase (PGND1), phosphoglycerate mutase (PGPM1), transketolase (PTKL1), phosphatidylinositol synthase (PPIS1), ferro-02-oxidoreductase (PFET3), high affinity iron permease (PFTR1), repressible alkaline phosphatase (PPH08), N-myristoyl transferase (PNMT1), pheromone response transcription factor (PMCM1), ubiquitin (PUBI4), single- stranded DNA endonuclease (PRAD2), the promoter of the major ADP/ATP carrier of the mitochondrial inner membrane (PPET9) (WO2008/128701) and the formate dehydrogenase (FMD) promoter.

Further examples of suitable promoters include *Saccharomyces cerevisiae* enolase (ENO1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase (PGK), and the maltase gene promoter (MAL).

The GAP promoter (pGAP), AOX1 (pAOX1) or AOX2 (pAOX2) promoter or a promoter which is a functional variant thereof and derived from any one of pGAP or pAOX1 or pAOX2 is particularly preferred. pAOX promoters can be induced by methanol and are repressed by glucose. Specifically, the functional variant has at least at least any one of 80%, 85%, 90%, 95%, or 100% sequence identity to the promoter from which it is derived, and has about the same promoter activity (e.g. +/- any one of 50%, 40%, 30%, 20%, or 10%; although the promoter activity may be improved) as compared to the promoter from which it is derived.

According to a specific embodiment, the ECP is methanol-inducible. In particular, the ECP is methanol-controlled. Specifically, the ECP can be fully induced in the methanol containing cell culture. In such case, the methanol may be used not only as a source of energy supplied to the cell culture, but also to induce the POI expression upon inducing the ECP.

According to a specific aspect, the ECP is methanol-inducible by the amount of methanol present in the cell culture used as a carbon source to produce the POI. Specifically, the GOI expression by the heterologous expression cassette is inducible by the methanol-inducible ECP.

Specifically, the ECP is methanol-inducible, and *e.g.*, repressed in the absence of a methanol amount which is less than any one of 0.1%, 0.05%, or 0.01% (v/v) in the cell culture medium or supernatant (herein referred to as a promoter-repressing amount).

Specifically, the ECP is methanol-inducible, and *e.g.*, induced in the presence of a methanol amount which is higher than the promoter-repressing amount *e.g.*, by at least any one of 0.1%, 0.5%, 1%, 1.5%, 2.0%, 2.5%, or 3% (v/v) in the cell culture medium or supernatant (herein referred to as a promoter-inducing amount).

Specifically, the ECP is fully induced by the methanol amount as used in the cell culture method described herein. The ECP promoter is considered to be fully induced, if the culture conditions provide for about maximum induction.

Such amounts in the cell culture medium or supernatant are particularly understood as the amount which upon feeding of the host cell and consumption by the host cell may be detectable. Typically, when producing a POI during the production phase of a cell culture, the cell culture is fed by adding a supplemental carbon source, yet in an amount that is immediately consumed by the cells during POI production, thus, leaving no or only a low remaining amount in the cell culture medium or supernatant, *e.g.* an amount up to 1.0 g/L.

Specifically, the ECP is endogenous or heterologous to the host cell.

Specifically, the ECP is any one of the following:

a) a pAOX1 promoter comprising or consisting of at least any one of 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% sequence identity to SEQ ID NO:5; or

b) a pAOX2 promoter comprising or consisting of at least any one of 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% sequence identity to SEQ ID NO:6; or

c) a promoter comprising or consisting of a nucleotide sequence selected from the group consisting of SEQ ID NO:36-49.

Specifically, any of the methanol-inducible promoters may be used which are listed in Table 38, in particular those comprising or consisting of a nucleotide sequence selected from the group consisting of SEQ ID NO:36-49.

Functional pAOX1 and pAOX2 promoter variants characterized by a sequence identity of at least 60% are exemplified by the exemplary methanol-inducible promoters further described herein. For example, SEQ ID NO:17 (pMOD1 promoter sequence of *Ogataea methanolica* JCM 10240) has a sequence identity of 54.0% compared to SEQ ID NO:5; and SEQ ID NO:18 (pMOD2 promoter sequence of *Ogataea methanolica* JCM 10240) has a sequence identity of 53.7% compared to SEQ ID NO:6. Sequence identity of the pMOD1 and pMOD2 promoter compared to the respective pAOX1 and pAOX2 promoter has been determined by alignment using LALIGN version 36.3.8g Dec, 2017; results refer to sequences aligned with the same sequence orientation and highest overlap (Parameters: Matrix: +5/-4; GAP OPEN: -5; Gap Extend: -4; E() Threshold 10.0; Output format: MARKX 0; Graphics: Yes).

Further exemplary methanol inducible promoter are listed in Table 38, or pSHB17, pALD4, pFDH1, pDAS1, pDAS2, pPMP20, pFBA1-2 pPMP47, pFLD, pFGH1, pTAL1-2, pDAS2, pCAM1, pPP7435_Chr1-0336 as described by Gasser *et al*. (Gasser, Steiger, & Mattanovich, 2015, Microb Cell Fact. 14: 196).

As described herein, the term “pAOX1” shall refer to both, a promoter comprising the sequence identified as SEQ ID NO:5, or a sequence which has a certain homology (or sequence identity) to SEQ ID NO:5. The homologous sequence is also referred to as pAOX1 homologue. The pAOX1 homologue may be a native, naturally-occurring sequence or a mutant thereof *e.g*., produced by any suitable method of mutagenesis.

As described herein, the term “pAOX2” shall refer to both, a promoter comprising the sequence identified as SEQ ID NO:6, or a sequence which has a certain homology (or sequence identity) to SEQ ID NO:6. The homologous sequence is also referred to as pAOX2 homologue. The pAOX2 homologue may be a native, naturally-occurring sequence or a mutant thereof *e.g*., produced by any suitable method of mutagenesis.

A pAOX1 or pAOX2 mutant described herein is specifically characterized by a promoter strength which is about 0.5-fold to at least any one of 1.1-fold, 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2-fold, 2.1-fold, 2.2-fold, 2.3-fold, 2.4-fold, 2.5-fold, 2.6-fold, 2.7-fold, 2.8-fold, 2.9-fold, 3-fold, 3.3-fold, 3.5-fold, 3.8-fold, 4-fold, 4.5-fold, 5-fold, 5.5-fold, or at least 6-fold increased compared to the respective native pAOX1 or pAOX2 promoter when in the induced state, as determined in a comparable expression system or production host cell line.

Specifically, the promoter strength is determined by the expression level of a POI, such as a model protein (e.g., Green Fluorescence Protein, GFP, including *e.g*.,

enhanced GFP, eGFP, Gene Bank Accession no. U57607), and/or the transcription strength, as compared to the reference promoter. Preferably, the transcription analysis is quantitative or semi-quantitative, preferably employing qRT-PCR, DNA microarrays, RNA sequencing and transcriptome analysis.

Specifically, the recombinant host cell described herein comprises only one or multiple heterologous GOIEC, *e.g*. multiple copies of said expression cassettes, such as at least 2, 3, 4, or 5 copies (gene copy number, GCN). For example, the recombinant host cell comprises up to 2, 3, 4, or five copies. Each of the copies may comprise or consist of the same or different sequences, yet includes the ECP operably linked to the GOI.

According to a specific aspect, the heterologous expression cassette is comprised in an autonomously replicating vector or plasmid, or integrated within a chromosome of said host cell.

The expression cassette may be introduced into the host cell and integrated into the host cell genome (or any of its chromosomes) as intrachromosomal element *e.g*., at a specific site of integration or randomly integrated, whereupon a high producer host cell line is selected. Alternatively, the expression cassette may be integrated within an extrachromosomal genetic element, such as a plasmid or an artificial chromosome *e.g*., a yeast artificial chromosome (YAC). According to a specific example, the expression cassette is introduced into the host cell by a vector, in particular an expression vector, which is introduced into the host cell by a suitable transformation technique. For this purpose, the GOI may be ligated into an expression vector.

A preferred yeast expression vector (which is preferably used for expression in yeast) is selected from the group consisting of plasmids derived from pPICZ, pGAPZ, pPIC9, pPICZalfa, pGAPZalfa, pPIC9K, pGAPHis, pPUZZLE or Golden*Pi*CS.

Techniques for transfecting or transforming host cells for introducing a vector or plasmid are well known in the art. These can include electroporation, spheroplasting, lipid vesicle mediated uptake, heat shock mediated uptake, calcium phosphate mediated transfection (calcium phosphate/DNA co-precipitation), viral infection, and particularly using modified viruses such as, for example, modified adenoviruses, microinjection and electroporation.

Transformants as described herein can be obtained by introducing the expression cassette, vector or plasmid DNA into a host and selecting transformants which express the relevant protein or selection marker. Host cells can be treated to introduce

heterologous or foreign DNA by methods conventionally used for transformation of host cells, such as the electric pulse method, the protoplast method, the lithium acetate method, and modified methods thereof. *P. pastoris* is preferably transformed by electroporation. Preferred methods of transformation for the uptake of the recombinant DNA fragment by the microorganism include chemical transformation, electroporation or transformation by protoplastation.

According to a specific aspect, the heterologous GOIEC described herein comprises or consists of an artificial fusion of polynucleotides, including the ECP operably linked to the GOI, and optionally further sequences, such as a signal, leader, or a terminator sequence.

Specifically, the expression cassette comprises the ECP operably linked to the GOI, and optionally further comprises signal and leader sequences, as necessary to express and produce the POI as a secreted protein.

According to a specific aspect, the GOIEC comprises a nucleotide sequence encoding a signal peptide enabling the secretion of the POI.

Specifically, the nucleotide sequence encoding the signal peptide is fused adjacent to the 5'-end of the GOI.

Specifically, the signal peptide is selected from the group consisting of signal sequences from *S. cerevisiae* alpha-mating factor prepro-peptide, the signal peptides from the *P. pastoris* acid phosphatase gene (PHO1) and the extracellular protein X (EPX1) (Heiss, S., V. Puxbaum, C. Gruber, F. Altmann, D. Mattanovich & B. Gasser, Microbiology 2015;161(7):1356-68).

Specifically, any of the signal and/or leader sequences as described in WO2014067926 A1 can be used, in particular SEQ ID NO:22 or SEQ ID NO:23.

Specifically, signal sequences as described in WO2012152823 A1 can be used, in particular the signal sequence of native alpha mating factor of *S. cerevisiae* identified as SEQ ID NO:24, or mutants thereof.

According to a specific aspect, the host cell described herein may undergo one or more further genetic modifications *e.g.*, for improving protein production.

Specifically, the host cell is further engineered to modify one or more genes influencing proteolytic activity used to generate protease deficient strains, in particular a strain deficient in carboxypeptidase Y activity. Particular examples are described in WO1992017595A1. Further examples of a protease deficient Pichia strain with a functional deficiency in a vacuolar protease, such as proteinase A or proteinase B, are

described in US6153424A. Further examples are Pichia strains which have an *ade2* deletion, and/or deletions of one or both of the protease genes, *PEP4* and *PRB1*, are provided by *e.g*., ThermoFisher Scientific.

Specifically, the host cell is engineered to modify at least one nucleic acid sequence encoding a functional gene product, in particular a protease, selected from the group consisting of PEP4, PRB1, YPS1, YPS2, YMP1, YMP2, YMP1, DAP2, GRHl, PRD1, YSP3, and PRB3, as disclosed in WO2010099195A1.

Overexpression or underexpression of genes encoding helper factors is specifically applied to enhance expression of a GOI, e.g. as described in WO2015158800A1.

The POI can be any one of eukaryotic, prokaryotic or synthetic peptides, polypeptides, proteins, or metabolites of a host cell.

According to a specific aspect, the POI is heterologous to the Mut- host cell or the ECP.

Specifically, the POI is heterologous to the host cell species.

Specifically, the POI is a secreted peptide, polypeptide, or protein, *i.e*. secreted from the host cell into the cell culture supernatant.

Specifically, the POI is a eukaryotic protein, preferably a mammalian derived or related protein such as a human protein or a protein comprising a human protein sequence, or a bacterial protein or bacterial derived protein

Preferably, the POI is a therapeutic protein functioning in mammals.

In specific cases, the POI is a multimeric protein, specifically a dimer or tetramer.

Specifically, the POI is a peptide or protein selected from the group consisting of an antigen-binding protein, a therapeutic protein, an enzyme, a peptide, a protein antibiotic, a toxin fusion protein, a carbohydrate - protein conjugate, a structural protein, a regulatory protein, a vaccine antigen, a growth factor, a hormone, a cytokine, a process enzyme.

Specifically, the antigen-binding protein is selected from the group consisting of

a) antibodies or antibody fragments, such as any of chimeric antibodies, humanized antibodies, bi-specific antibodies, Fab, Fd, scFv, diabodies, triabodies, Fv tetramers, minibodies, single-domain antibodies like VH, VHH, IgNARs, or V-NAR;

b) antibody mimetics, such as Adnectins, Affibodies, Affilins, Affimers, Affitins, Alphabodies, Anticalins, Avimers, DARPins, Fynomers, Kunitz domain peptides, Monobodies, or NanoCLAMPS; or

c) fusion proteins comprising one or more immunoglobulin-fold domains, antibody domains or antibody mimetics.

A specific POI is an antigen-binding molecule such as an antibody, or a fragment thereof, in particular an antibody fragment comprising an antigen-binding domain. Among specific POIs are antibodies such as monoclonal antibodies (mAbs), immunoglobulin (Ig) or immunoglobulin class G (IgG), heavy-chain antibodies (HcAb's), or fragments thereof such as fragment-antigen binding (Fab), Fd, single-chain variable fragment (scFv), or engineered variants thereof such as for example Fv dimers (diabodies), Fv trimers (triabodies), Fv tetramers, or minibodies and single-domain antibodies like VH, VHH, IgNARs, or V-NAR, or any protein comprising an immunoglobulin-fold domain. Further antigen-binding molecules may be selected from antibody mimetics, or (alternative) scaffold proteins such as *e.g.,* engineered Kunitz domains, Adnectins, Affibodies, Affiline, Anticalins, or DARPins.

According to a specific aspect, the POI is *e.g.*, BOTOX, Myobloc, Neurobloc, Dysport (or other serotypes of botulinum neurotoxins), alglucosidase alpha, daptomycin, YH-16, choriogonadotropin alpha, filgrastim, cetrorelix, interleukin-2, aldesleukin, teceleulin, denileukin diftitox, interferon alpha-n3 (injection), interferon alpha-nl, DL-8234, interferon, Suntory (gamma-1a), interferon gamma, thymosin alpha 1, tasonermin, DigiFab, ViperaTAb, EchiTAb, CroFab, nesiritide, abatacept, alefacept, Rebif, eptoterminalfa, teriparatide (osteoporosis), calcitonin injectable (bone disease), calcitonin (nasal, osteoporosis), etanercept, hemoglobin glutamer 250 (bovine), drotrecogin alpha, collagenase, carperitide, recombinant human epidermal growth factor (topical gel, wound healing), DWP401, darbepoetin alpha, epoetin omega, epoetin beta, epoetin alpha, desirudin, lepirudin, bivalirudin, nonacog alpha, Mononine, eptacog alpha (activated), recombinant Factor VIII+VWF, Recombinate, recombinant Factor VIII, Factor VIII (recombinant), Alphnmate, octocog alpha, Factor VIII, palifermin, indikinase, tenecteplase, alteplase, pamiteplase, reteplase, nateplase, monteplase, follitropin alpha, rFSH, hpFSH, micafungin, pegfilgrastim, lenograstim, nartograstim, sermorelin, glucagon, exenatide, pramlintide, iniglucerase, galsulfase, Leucotropin, molgramostirn, triptorelin acetate, histrelin (subcutaneous implant, Hydron), deslorelin, histrelin, nafarelin, leuprolide sustained release depot (ATRIGEL), leuprolide implant (DUROS), goserelin, Eutropin, KP-102 program, somatropin, mecasermin (growth failure), enlfavirtide, Org-33408, insulin glargine, insulin glulisine, insulin (inhaled), insulin lispro, insulin deternir, insulin (buccal, RapidMist), mecasermin rinfabate, anakinra,

celmoleukin, 99 mTc-apcitide injection, myelopid, Betaseron, glatiramer acetate, Gepon, sargramostim, oprelvekin, human leukocyte-derived alpha interferons, Bilive, insulin (recombinant), recombinant human insulin, insulin aspart, mecasenin, Roferon-A, interferon-alpha 2, Alfaferone, interferon alfacon-1, interferon alpha, Avonex' recombinant human luteinizing hormone, dornase alpha, trafermin, ziconotide, taltirelin, diboterminalfa, atosiban, becaplermin, eptifibatide, Zemaira, CTC-111, Shanvac-B, HPV vaccine (quadrivalent), octreotide, lanreotide, ancestirn, agalsidase beta, agalsidase alpha, laronidase, prezatide copper acetate (topical gel), rasburicase, ranibizumab, Actimmune, PEG-Intron, Tricomin, recombinant house dust mite allergy desensitization injection, recombinant human parathyroid hormone (PTH) 1-84 (sc, osteoporosis), epoetin delta, transgenic antithrombin III, Granditropin, Vitrase, recombinant insulin, interferon-alpha (oral lozenge), GEM-21S, vapreotide, idursulfase, omnapatrilat, recombinant serum albumin, certolizumab pegol, glucarpidase, human recombinant C1 esterase inhibitor (angioedema), lanoteplase, recombinant human growth hormone, enfuvirtide (needle-free injection, Biojector 2000), VGV-1, interferon (alpha), lucinactant, aviptadil (inhaled, pulmonary disease), icatibant, ecallantide, omiganan, Aurograb, pexigananacetate, ADI-PEG-20, LDI-200, degarelix, cintredelinbesudotox, Favld, MDX-1379, ISAtx-247, liraglutide, teriparatide (osteoporosis), tifacogin, AA4500, T4N5 liposome lotion, catumaxomab, DWP413, ART-123, Chrysalin, desmoteplase, amediplase, corifollitropinalpha, TH-9507, teduglutide, Diamyd, DWP-412, growth hormone (sustained release injection), recombinant G-CSF, insulin (inhaled, AIR), insulin (inhaled, Technosphere), insulin (inhaled, AERx), RGN-303, DiaPep277, interferon beta (hepatitis C viral infection (HCV)), interferon alpha-n3 (oral), belatacept, transdermal insulin patches, AMG-531, MBP-8298, Xerecept, opebacan, AIDSVAX, GV-1001, LymphoScan, ranpirnase, Lipoxysan, lusupultide, MP52 (beta-tricalciumphosphate carrier, bone regeneration), melanoma vaccine, sipuleucel-T, CTP-37, Insegia, vitespen, human thrombin (frozen, surgical bleeding), thrombin, TransMID, alfimeprase, Puricase, terlipressin (intravenous, hepatorenal syndrome), EUR-1008M, recombinant FGF-I (injectable, vascular disease), BDM-E, rotigaptide, ETC-216, P-113, MBI-594AN, duramycin (inhaled, cystic fibrosis), SCV-07, OPI-45, Endostatin, Angiostatin, ABT-510, Bowman Birk Inhibitor Concentrate, XMP-629, 99 mTc-Hynic-Annexin V, kahalalide F, CTCE-9908, teverelix (extended release), ozarelix, rornidepsin, BAY-504798, interleukin4, PRX-321, Pepscan, iboctadekin, rhlactoferrin, TRU-015, IL-21, ATN-161, cilengitide, Albuferon, Biphasix, IRX-2, omega interferon, PCK-3145,

CAP-232, pasireotide, huN901-DMI, ovarian cancer immunotherapeutic vaccine, SB-249553, Oncovax-CL, OncoVax-P, BLP-25, CerVax-16, multi-epitope peptide melanoma vaccine (MART-1, gp100, tyrosinase), nemifitide, rAAT (inhaled), rAAT (dermatological), CGRP (inhaled, asthma), pegsunercept, thymosinbeta4, plitidepsin, GTP-200, ramoplanin, GRASPA, OBI-1, AC-100, salmon calcitonin (oral, eligen), calcitonin (oral, osteoporosis), examorelin, capromorelin, Cardeva, velafermin, 131I-TM-601, KK-220, T-10, ularitide, depelestat, hematide, Chrysalin (topical), rNAPc2, recombinant Factor V111 (PEGylated liposomal), bFGF, PEGylated recombinant staphylokinase variant, V-10153, SonoLysis Prolyse, NeuroVax, CZEN-002, islet cell neogenesis therapy, rGLP-1, BIM-51077, LY-548806, exenatide (controlled release, Medisorb), AVE-0010, GA-GCB, avorelin, ACM-9604, linaclotid eacetate, CETi-1, Hemospan, VAL (injectable), fast-acting insulin (injectable, Viadel), intranasal insulin, insulin (inhaled), insulin (oral, eligen), recombinant methionyl human leptin, pitrakinra subcutancous injection, eczema), pitrakinra (inhaled dry powder, asthma), Multikine, RG-1068, MM-093, NBI-6024, AT-001, PI-0824, Org-39141, Cpn10 (autoimmune diseases/inflammation), talactoferrin (topical), rEV-131 (ophthalmic), rEV-131 (respiratory disease), oral recombinant human insulin (diabetes), RPI-78M, oprelvekin (oral), CYT-99007 CTLA4-Ig, DTY-001, valategrast, interferon alpha-n3 (topical), IRX-3, RDP-58, Tauferon, bile salt stimulated lipase, Merispase, alaline phosphatase, EP-2104R, Melanotan-II, bremelanotide, ATL-104, recombinant human microplasmin, AX-200, SEMAX, ACV-1, Xen-2174, CJC-1008, dynorphin A, SI-6603, LAB GHRH, AER-002, BGC-728, malaria vaccine (virosomes, PeviPRO), ALTU-135, parvovirus B19 vaccine, influenza vaccine (recombinant neuraminidase), malaria/HBV vaccine, anthrax vaccine, Vacc-5q, Vacc-4x, HIV vaccine (oral), HPV vaccine, Tat Toxoid, YSPSL, CHS-13340, PTH(1-34) liposomal cream (Novasome), Ostabolin-C, PTH analog (topical, psoriasis), MBRI-93.02, MTB72F vaccine (tuberculosis), MVA-Ag85A vaccine (tuberculosis), FARA04, BA-210, recombinant plague FIV vaccine, AG-702, OxSODrol, rBetV1, Der-p1/Der-p2/Der-p7 allergen-targeting vaccine (dust mite allergy), PR1 peptide antigen (leukemia), mutant ras vaccine, HPV-16 E7 lipopeptide vaccine, labyrinthin vaccine (adenocarcinoma), CML vaccine, WT1-peptide vaccine (cancer), IDD-5, CDX-110, Pentrys, Norelin, CytoFab, P-9808, VT-111, icrocaptide, telbermin (dermatological, diabetic foot ulcer), rupintrivir, reticulose, rGRF, HA, alpha-galactosidase A, ACE-011, ALTU-140, CGX-1160, angiotensin therapeutic vaccine, D-4F, ETC-642, APP-018, rhMBL, SCV-07 (oral, tuberculosis), DRF-7295, ABT-828,

ErbB2-specific immunotoxin (anticancer), DT3SSIL-3, TST-10088, PRO-1762, Combotox, cholecystokinin-B/gastrin-receptor binding peptides, 111In-hEGF, AE-37, trasnizumab-DM1, Antagonist G, IL-12 (recombinant), PM-02734, IMP-321, rhIGF-BP3, BLX-883, CUV-1647 (topical), L-19 based radioimmunotherapeutics (cancer), Re-188-P-2045, AMG-386, DC/1540/KLH vaccine (cancer), VX-001, AVE-9633, AC-9301, NY-ESO-1 vaccine (peptides), NA17.A2 peptides, melanoma vaccine (pulsed antigen therapeutic), prostate cancer vaccine, CBP-501, recombinant human lactoferrin (dry eye), FX-06, AP-214, WAP-8294A (injectable), ACP-HIP, SUN-11031, peptide YY [3-36] (obesity, intranasal), FGLL, atacicept, BR3-Fc, BN-003, BA-058, human parathyroid hormone 1-34 (nasal, osteoporosis), F-18-CCR1, AT-1100 (celiac disease/diabetes), JPD-003, PTH(7-34) liposomal cream (Novasome), duramycin (ophthalmic, dry eye), CAB-2, CTCE-0214, GlycoPEGylated erythropoietin, EPO-Fc, CNTO-528, AMG-114, JR-013, Factor XIII, aminocandin, PN-951, 716155, SUN-E7001, TH-0318, BAY-73-7977, teverelix (immediate release), EP-51216, hGH (controlled release, Biosphere), OGP-I, sifuvirtide, TV4710, ALG-889, Org-41259, rhCC10, F-991, thymopentin (pulmonary diseases), r(m)CRP, hepatoselective insulin, subalin, L19-IL-2 fusion protein, elafin, NMK-150, ALTU-139, EN-122004, rhTPO, thrombopoietin receptor agonist (thrombocytopenic disorders), AL-108, AL-208, nerve growth factor antagonists (pain), SLV-317, CGX-1007, INNO-105, oral teriparatide (eligen), GEM-OS1, AC-162352, PRX-302, LFn-p24 fusion vaccine (Therapore), EP-1043, S pneumoniae pediatric vaccine, malaria vaccine, Neisseria meningitidis Group B vaccine, neonatal group B streptococcal vaccine, anthrax vaccine, HCV vaccine (gpE1+gpE2+MF-59), otitis media therapy, HCV vaccine (core antigen+ISCOMATRIX), hPTH(1-34) (transdermal, ViaDerm), 768974, SYN-101, PGN-0052, aviscumnine, BIM-23190, tuberculosis vaccine, multi-epitope tyrosinase peptide, cancer vaccine, enkastim, APC-8024, GI-5005, ACC-001, TTS-CD3, vascular-targeted TNF (solid tumors), desmopressin (buccal controlled-release), onercept, or TP-9201, adalimumab (HUMIRA), infliximab (REMICADE™), rituximab (RITUXAN™/MAB THERA™), etanercept (ENBREL™), bevacizumab (AVASTIN™), trastuzumab (HERCEPTIN™), pegrilgrastim (NEULASTA™), or any other suitable POI including biosimilars and biobetters.

According to a specific aspect, a fermentation product is isolated from the cell culture, which fermentation product comprises the POI or a host cell metabolite obtained from the Mut- host cell.

According to a specific aspect, the Mut- host cell is a yeast cell of the genus Pichia, Komagataella, Hansenula, Ogataea or Candida.

Specifically, the Mut- host cell is originating from a strain which is of a yeast selected from the group consisting of a Pichia species, such as *Pichia pastoris*, *Pichia methanolica*, *Pichia kluyveri*, and *Pichia angusta*; Komagataella species, such as *Komagataella pastoris*, *Komagataella pseudopastoris* or *Komagataella phaffii,* Hansenula species, such as *Hansenula polymorpha,* Ogataea species, such as *Ogataea polymorpha,* or *Ogataea parapolymorpha,* and Candida species, such as *Candida utilis*, *Candida cacaoi*, and *Candida boidinii.*

Preferred is the species *Pichia pastoris*. Specifically, the host cell is a *Pichia pastoris* strain selected from the group consisting of CBS 704, CBS 2612, CBS 7435, CBS 9173-9189, DSMZ 70877, X-33, GS115, KM71, KM71H and SMD1168.

Sources: CBS 704 (=NRRL Y-1603 = DSMZ 70382), CBS 2612 (=NRRL Y-7556), CBS 7435 (=NRRL Y-11430), CBS 9173-9189 (CBS strains: CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelculturen, Utrecht, The Netherlands), and DSMZ 70877 (German Collection of Microorganisms and Cell Cultures); strains from Invitrogen, such as X-33, GS115, KM71, KM71H and SMD1168.

According to a specific aspect, the invention provides for a method of producing a protein of interest (POI) comprising culturing a Mut- host cell described herein using methanol as a carbon source to produce the POI, in particular such methanol amounts for use as a source of energy, not (only) for methanol-induction of an ECP.

According to a specific aspect, the method comprises culturing the Mut- host cell using methanol as a carbon source to produce the POI, which Mut- host cell comprises a heterologous gene of interest expression cassette (GOIEC) comprising an expression cassette promoter (ECP) operably linked to a gene of interest (GOI) encoding a protein of interest (POI),

wherein the Mut- host cell is engineered by one or more genetic modifications to reduce expression of a first and a second endogenous gene compared to the host cell prior to said one or more genetic modifications, wherein

a) the first endogenous gene encodes alcohol oxidase 1 (AOX1) comprising the amino acid sequence identified as SEQ ID NO:1 or a homologue thereof, and

b) the second endogenous gene encodes alcohol oxidase 2 (AOX2) comprising the amino acid sequence identified as SEQ ID NO:3 or a homologue thereof.

Specifically, the Mut- host cell is cultured using methanol as a sole carbon source or in a mixture with other carbon sources (or carbohydrates), in particular as a source of energy, such as for growth and/or POI production (synthesis).

Specifically, such other carbon source (herein also referred to as “non-methanol carbon source” is a carbohydrate.

Specifically, the non-methanol carbon source is selected from saccharides, polyols, alcohols, or mixtures of any one or more of the foregoing.

Specifically, the saccharides may be any one or more of monosaccharides, such as a hexose, *e.g*. glucose, fructose, galactose or mannose, or a disaccharides, such as saccharose; or an alcohol or polyol *e.g*., ethanol, or any diol, or triol, *e.g*., glycerol, or a mixture of any of the foregoing. In addition to the methanol amount used as a carbon source as described herein, any such non-methanol carbon source may be additionally used in the cell culture in an amount to produce said POI.

According to a specific embodiment, a cell line of the Mut- host cell is cultured.

Specifically, the cell line is cultured under batch, fed-batch or continuous culture conditions. The culture may be performed in microtiter plates, shake-flasks, or a bioreactor, and optionally starting with a batch phase as the first step, followed by a fed-batch phase or a continuous culture phase as the second step.

According to a specific aspect, the method described herein comprises a growing phase and a production phase.

Specifically, the method comprises the steps:

a) culturing the host cell under growing conditions (growing phase, or “growth phase”); and a further step

b) culturing the host cell under growth-limiting conditions in the presence of methanol as a carbon source (production phase), during which the GOI is expressed to produce said POI.

Specifically, the second step b) follows the first step a).

Specifically,

a) a growing phase, during which the Mut- host cell is cultured using a basal carbon source as a source of energy; is followed by

b) a production phase, during which the Mut- host cell is cultured using a methanol feed thereby producing the POI.

Specifically, the host cell is cultured in the first step under growing conditions in a cell culture medium comprising the first carbon source, *e.g*. in an amount sufficient to

enable growth of the host cell in cell culture, optionally until the amount of the carbon source is consumed, and further culturing can be under growth-limiting conditions.

Specifically, the second carbon source is methanol and optionally one or more further carbon sources (other than methanol), said second carbon source being referred to as supplemental carbon source.

Specifically, said basal carbon source and/or supplemental carbon source (in addition to methanol) can be selected from saccharides, polyols, alcohols, or mixtures of any one or more of the foregoing.

According to a specific embodiment, the basal carbon source is different from the supplemental carbon source, *e.g*. quantitatively and/or qualitatively different. The quantitative difference typically provides for the different conditions to repress or induce the ECP promoter activity.

According to a further specific embodiment the basal and the supplemental carbon sources comprise the same type of molecules or carbohydrates, preferably in different concentrations. According to a further specific embodiment, the carbon source is a mixture of two or more different carbon sources.

Any type of organic carbon source may be used, in particular those typically used for host cell culture, in particular for eukaryotic host cell culture. According to a specific embodiment, the carbon source is a hexose, such as glucose, fructose, galactose or mannose, a disaccharide, such as saccharose, an alcohol, such as glycerol or ethanol, or a mixture thereof.

According to a specifically preferred embodiment, the basal carbon source is selected from the group consisting of glucose, glycerol, ethanol, or mixtures thereof. According to a preferred embodiment, the basal carbon source is glycerol.

According to a further specific embodiment, the supplemental carbon source comprises (in addition to methanol) a hexose such as glucose, fructose, galactose and mannose, a disaccharide, such as saccharose, an alcohol, such as glycerol or ethanol, or a mixture thereof. According to a preferred embodiment, the supplemental carbon source comprises glucose in addition to methanol.

Both of said culturing steps specifically comprise cultivating the cell line in the presence of said carbon sources.

Specifically, said growth phase (step a)) culturing is performed in a batch phase; and said production phase (step b)) culturing is performed in fed-batch or a continuous cultivation phase.

Specifically, the host cells are grown in a carbon source rich medium comprising a basal carbon source during the phase of high growth rate (under growing conditions), step a) (e.g. at least 50%, or at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or up to the maximum growth rate) and producing the POI during a phase of low growth rate (under growth-limiting conditions), step b) (*e.g*. less than 90%, preferably less than 80%, less than 70%, less than 60%, less than 50%, or less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.4%, less than 0.3%, or less than 0.2% of the maximum growth rate) while limiting the carbon source, in particular by feeding a defined minimal medium comprising only the amount of carbon source which is completely consumed when maintaining the cell culture in the production phase.

Specifically, an average methanol concentration of at least any one of 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, or 1.0% (v/v) e.g., up to any one of 3%, 2.5%, 2%, 1.5%, or 1% (v/v) is used in the host cell culture, specifically in the cell culture medium or supernatant, in particular during a production phase.

Specifically, the average methanol concentration is maintained during a production phase of at least 24 hours, preferably, at least any one of 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5% or 3% (v/v).

According to a specific embodiment, the average methanol concentration is 0.5-2% (v/v), during the production phase of at least 24 hours

The average amount or concentration can be calculated as the sum of methanol concentrations as measured in the cell culture, in particular in the cell culture medium or supernatant, at least at the start and at the end of an observation period, and during the observation period *e.g*., at least every 24 h, or per continuous measurement, divided by the number of measurements.

The methanol concentration in the cell culture can be measured using a suitable standard assay like HPLC, *e.g*. determined as a residual concentration in the culture medium upon consumption by the cell culture.

Specifically, the methanol feed is at an average feed rate of at least any one of 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 mg methanol/(g dry biomass*h), or higher *e.g*., 2-20 or 2-15 mg methanol/(g dry biomass*h) during a production phase.

Methanol may be added to the cell culture in one or more instalments *e.g*., by one or more injections, or may be continuously added during a certain period of time while

producing the POI. The average amount can be calculated as the sum of all methanol additions during an observation period divided by the average total dry biomass and by the duration of the observation period. The average total dry biomass is calculated by measuring the dry biomass concentration at least at the start and at the end of an observation period, and optional during the observation period. The dry biomass concentration is then interpolated between start and the end of the observation period. The interpolated dry biomass concentration is multiplied by the reactor volume at each interval, the calculated values for all intervals are summed and divided by the number of intervals. An interval duration is less than or equal to 1 h.

Specifically, the average feed rate is maintained during a production phase of at least 24 hours, preferably, at least any one of 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 mg methanol/(g dry biomass*h), or higher *e.g*., 2-20 or 2-15 mg methanol/(g dry biomass*h).

The observation period is herein understood as a certain period of time during which the cell culture is producing the POI, and particularly understood as a production phase, in particular the production phase of a fed-batch or continuous cell cultivation method. Though the actual POI production process or production phase may be longer than the observation period, the average amount is calculated during a defined observation period.

Specifically, the duration of the POI production process is 10 to 500h.

Specifically, a batch phase is performed for around 10 to 36h.

The term “around” with respect to cultivation time shall mean +/-5% or +/-10%.

For example, the specific batch performance time of around 10 to 36h may be 18 to 39.6h, specifically 19 to 37.8h.

According to a specific embodiment, a batch phase is performed using 10 to 50 g/L glycerol, specifically 20 to 40 g/L glycerol as a basal carbon source in batch media, and cultivation is performed at 25°C for around 27 to 30h, or at 30°C for around 23 to 36h, or at any temperature between 25°C and 30°C during a cultivation time of 23 to 36h. Lowering the glycerol concentration in the batch medium would decrease the length of the batch phase, while increasing the glycerol in the batch medium would even prolong the batch phase. As an alternative to glycerol, glucose can be used, *e.g*. in about the same amounts.

In a typical system of cell culture and POI expression, wherein a batch phase is followed by a fed-batch phase, specifically, the cultivation in the fed-batch phase is

performed for any one of around 15 to 100h, around 15 to 80h, around 15 to 70h, around 15 to 60h, around 15 to 50h, around 15 to 45h, around 15 to 40h, around 15 to 35h, around 15 to 30h, around 15 to 35h, around 15 to 25h, or around 15 to 20h; preferably around 20 to 40h. Specifically, the cultivation in the fed-batch phase is performed for any one of around 100h, around 80h, around 70h, around 60h, around 55h, around 50h, around 45h, around 40h, around 35h, around 33h, around 30h, around 25h, around 20h, or around 15h.

Specifically, the volume specific product formation rate (rP) is the amount of product (mg) formed per Unit Volume (L) and Unit time (h) (mg (L h)$^{-1}$). Volume specific product formation rate is also called space time yield (STY) or volumetric productivity.

Specifically, a fed-batch cultivation of the method described herein is performed such that a space time yield of around 30 mg (L h)$^{-1}$ (meaning 30 mg (L h)$^{-1}$ +/-5% or +/-10%). Specifically a space time yield of around 30 mg (L h)$^{-1}$ is achieved within around 30h fed batch, specifically at least any of 27, 28, 29, 30, 31, 32, or 33 mg (L h)$^{-1}$ within less than any one of 33h, 32h, 31h, 30h, 29h, 28h, 27h, 26h, or 25h fed batch time can be achieved.

Specifically, the POI is expressed in the production phase under growth-limiting conditions, *e.g*. by cultivating the cell line at a growth rate of less than the maximal growth rate, typically less than 90%, preferably less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.4%, less than 0.3%, or less than 0.2% of the maximum growth rate of the cells. Typically the maximum growth rate is individually determined for each type of host cell.

According to a specific aspect, the Mut- host cell is cultured during a production phase under conditions limiting the host cell growth to less than any one of 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% (w/w biomass).

Specifically, the production phase employs a feed medium that provides for a supplemental carbon source in a growth limiting amount to keep the specific growth rate within the range of 0.0001 h^{-1} to 0.2 h^{-1}, preferably 0.005 h^{-1} to 0.15 h^{-1}.

According to a specific aspect, the invention provides for the use of a recombinant methanol utilization pathway deficient methylotrophic yeast (Mut-) host cell in a method of producing a fermentation product which method comprises culturing said Mut- host

cell under conditions that permit the Mut- host cell to use methanol as a substrate for alcohol dehydrogenase (ADH2), and to produce the fermentation product.

Specifically, said Mut- host cell is deficient of alcohol oxidase 1 (AOX1) and alcohol oxidase 2 (AOX2).

Specifically, in said Mut- host cell, the genes encoding alcohol oxidase 1 (AOX1) and alcohol oxidase 2 (AOX2) are knocked out or deleted.

According to a specific aspect, the invention provides for the use of a recombinant methanol utilization pathway deficient methylotrophic yeast (Mut-) host cell in a method of producing a fermentation product which method comprises culturing said Mut- host cell under conditions that permit the Mut- host cell to produce the fermentation product using methanol as a carbon source, which Mut- host cell is engineered by one or more genetic modifications

a) to render the host cell methanol utilization pathway deficient, by reducing expression of a first and a second endogenous gene compared to the host cell prior to said one or more genetic modifications, wherein

i. the first endogenous gene encodes alcohol oxidase 1 (AOX1) comprising the amino acid sequence identified as SEQ ID NO:1 or a homologue thereof that is endogenous to the respective wild-type host cell, and

ii. the second endogenous gene encodes alcohol oxidase 2 (AOX2) comprising the amino acid sequence identified as SEQ ID NO:3 or a homologue thereof that is endogenous to the respective wild-type host cell, and

b) to increase expression of an alcohol dehydrogenase (*ADH2*) gene, wherein the *ADH2* gene encodes an alcohol dehydrogenase (ADH2).

According to a specific aspect, the invention provides for a method for producing a protein of interest (POI) in a host cell, comprising the steps:

a) genetically engineering the host cell to reduce expression (underexpress) of said first and second genes encoding the AOX1 and AOX2, respectively;

b) genetically engineering the host cell to increase expression (overexpress) of a gene encoding ADH2;

c) introducing into the host cell a heterologous expression cassette comprising a gene of interest (GOI) encoding said POI under the control of an expression cassette promoter (ECP);

d) culturing said host cell under conditions to produce said POI using methanol as a carbon source, thereby particularly providing energy for growth and/or POI production;
e) optionally isolating said POI from the cell culture; and
f) optionally purifying said POI.

Specifically, step a) of the method described herein is carried out before, or after, or concomitantly with step b).

Specifically, steps a) and b) of the method described herein is carried out before, or after, or concomitantly with step c).

According to a specific aspect, the host cell is first genetically modified to reduce expression of said first and second genes encoding the AOX1 and AOX2, respectively, and to increase expression of a gene encoding ADH2, before being engineered for producing the POI. According to a specific example, a wild-type host cell is genetically modified according to steps a) and b) of the method described herein. Specifically, the host cell is provided upon introducing said one or more genetic modifications into a wild-type host cell strain for reduction of said first and second genes encoding the AOX1 and AOX2, respectively, and for increasing expression of the gene encoding ADH2.

According to a further aspect, the host cell is first engineered for producing the heterologous or recombinant POI, before being further genetically modified to reduce said first and second genes encoding the AOX1 and AOX2, respectively, and to increase expression of the gene encoding ADH2. According to a specific example, a wild-type host cell may first be engineered to comprise the expression cassette for POI production. Such engineered host cell may then be further modified to reduce said first and second genes encoding the AOX1 and AOX2, respectively, and to increase expression of the gene encoding ADH2.

According to a further aspect, the host cell is undergoing the engineering steps, including the engineering for POI production and genetically modifying for reduction of said first and second genes encoding the AOX1 and AOX2, respectively, and for increasing expression of the gene encoding ADH2, concomitantly *i.*e. in one method step, *e.g.*, employing the respective expression cassette, reagents and tools in one or more reaction mixtures.

Specifically, the method employs method steps to produce the recombinant Mut- host cell as further described herein.

Specifically, the heterologous expression cassette comprises the ECP as further described herein.

Specifically, the POI can be produced by culturing the Mut- host cell in an appropriate medium, producing the POI during a culturing step using a cell culture production medium comprising methanol, and isolating the expressed POI from the cell culture, in particular from the cell culture supernatant or medium upon separating the cells, and optionally purifying it by a method appropriate for the expressed product. Thereby, a purified POI preparation can be produced.

It has surprisingly turned out that the Mut- host cell was insensitive to methanol and could effectively uptake and use significant amounts of methanol as necessary to provide energy for POI production. This was surprising because in the prior art, methanol was found to be toxic to a Mut^S strain.

According to a specific example of the cell culture as described herein, the growth of the Mut- host cells was advantageously limited during the production phase, which reduced the necessity of oxygen supply and cooling.

It was even more surprising that the yield of POI production was increased by a heretofore underestimated mechanism of alcohol dehydrogenase and the activity of ADH2 in methylotrophic yeast, which turned out to result in methanol uptake and effective methanol consumption despite of knocking out AOX1 and AOX2 genes in the methanol utilization pathway deficient methylotrophic yeast.

According to a specific example, a methanol-inducible ECP has been advantageously used in a GOIEC. The methanol amounts as used in the cell culture as a carbon source were sufficient to induce expression of the GOI.

FIGURES

Figure 1: Sequences referred to herein

DETAILED DESCRIPTION OF THE INVENTION

Specific terms as used throughout the specification have the following meaning.

The term “carbon source” as used herein shall mean a fermentable carbon substrate, typically a source carbohydrate, suitable as an energy source for microorganisms, such as those capable of being metabolized by host organisms or

production cell lines, in particular sources selected from the group consisting of monosaccharides, oligosaccharides, polysaccharides, alcohols including glycerol, in the purified form, in minimal media or provided in raw materials, such as a complex nutrient material. The carbon source may be used as described herein as a single carbon source or as a mixture of different carbon sources.

As described herein, methanol is used as a carbon source, *e.g.*, as a sole carbon source during a production phase, or in a mixture with a non-methanol carbon source. Specifically, methanol is co-fed to the cell culture with any non-methanol carbon source.

A non-methanol carbon source is herein understood as a carbon source which is any other than methanol, in particular a methanol-free carbon source.

A “basal carbon source” such as described herein typically is a carbon source suitable for cell growth, such as a nutrient for host cells, in particular for eukaryotic cells. The basal carbon source may be provided in a medium, such as a basal medium or complex medium, but also in a chemically defined medium containing a purified carbon source. The basal carbon source typically is provided in an amount to provide for cell growth, in particular during the growth phase in a cultivation process, for example to obtain cell densities of at least 5 g/L cell dry mass, preferably at least 10 g/L cell dry mass, or at least 15 g/L cell dry mass, e.g. exhibiting viabilities of more than 90% during standard sub-culture steps, preferably more than 95%.

The basal carbon source is typically used in an excess or surplus amount, which is understood as an excess providing energy to increase the biomass, *e.g.* during the cultivation of a cell line with a high specific growth rate, such as during the growth phase of a cell line in a batch or fed-batch cultivation process. This surplus amount is particularly in excess of the limited amount of a supplemental carbon source (as used under growth-limited conditions) to achieve a residual concentration in the fermentation broth that is measurable and typically at least 10 fold higher, preferably at least 50 fold or at least 100 fold higher than during feeding the limited amount of the supplemental carbon source.

A “supplemental carbon source” such as described herein typically is a supplemental substrate facilitating the production of fermentation products by production cell lines, in particular in the production phase of a cultivation process. The production phase specifically follows a growth phase, e.g. in batch, fed-batch and continuous cultivation process. The supplemental carbon source specifically may be contained in the feed of a fed-batch process. The supplemental carbon source is typically employed

in a cell culture under carbon substrate limited conditions, i.e. using the carbon source in a limited amount.

Specifically, in a method described herein methanol is used as a supplemental carbon source.

A “limited amount” of a carbon source or a “limited carbon source” is herein understood to specifically refer to the type and amount of a carbon substrate facilitating the production of fermentation products by production cell lines, in particular in a cultivation process with controlled growth rates of less than the maximum growth rate. The production phase specifically follows a growth phase, e.g. in batch, fed-batch and continuous cultivation process. Cell culture processes may employ batch culture, continuous culture, and fed-batch culture. Batch culture is a culture process by which a small amount of a seed culture solution is added to a medium and cells are grown without adding an additional medium or discharging a culture solution during culture.

Continuous culture is a culture process by which a medium is continuously added and discharged during culture. The continuous culture also includes perfusion culture. Fed-batch culture, which is an intermediate between the batch culture and the continuous culture and also referred to as semi-batch culture, is a culture process by which a medium is continuously or sequentially added during culture but, unlike the continuous culture, a culture solution is not continuously discharged.

Specifically preferred is a fed-batch process which is based on feeding of a growth limiting nutrient substrate to a culture. The fed-batch strategy, including single fed-batch or repeated fed-batch fermentation, is typically used in bio-industrial processes to reach a high cell density in the bioreactor. The controlled addition of the carbon substrate directly affects the growth rate of the culture and helps to avoid overflow metabolism or the formation of unwanted metabolic byproducts. Under carbon source limited conditions, the carbon source specifically may be contained in the feed of a fed-batch process. Thereby, the carbon substrate is provided in a limited amount.

Also in chemostat or continuous culture as described herein, the growth rate can be tightly controlled.

The limited amount of a carbon source is herein particularly understood as the amount of a carbon source necessary to keep a production cell line under growth-limited conditions, e.g. in a production phase or production mode. Such a limited amount may be employed in a fed-batch process, where the carbon source is contained in a feed medium and supplied to the culture at low feed rates for sustained energy delivery, e.g.

to produce a POI, while keeping the biomass at low specific growth rates. A feed medium is typically added to a fermentation broth during the production phase of a cell culture.

The limited amount of a carbon source may, for example, be determined by the residual amount of the carbon source in the cell culture broth, which is below a predetermined threshold or even below the detection limit as measured in a standard (carbohydrate) assay. The residual amount typically would be determined in the fermentation broth upon harvesting a fermentation product.

The limited amount of a carbon source may as well be determined by defining the average feed rate of the carbon source to the fermenter, e.g. as determined by the amount added over the full cultivation process, e.g. the fed-batch phase, per cultivation time, to determine a calculated average amount per time. This average feed rate is kept low to ensure complete usage of the supplemental carbon source by the cell culture, e.g. between 0.6 g L^{-1} h^{-1} (g carbon source per L initial fermentation volume and h time) and 25 g L^{-1} h^{-1}, preferably between 1.6 g L^{-1} h^{-1} and 20 g L^{-1} h^{-1}.

The limited amount of a carbon source may also be determined by measuring the specific growth rate, which specific growth rate is kept low, e.g. lower than the maximum specific growth rate, during the production phase, e.g. within a predetermined range, such as in the range of 0.001 h^{-1} to 0.20 h^{-1}, or 0.005 h^{-1} to 0.20 h^{-1}, preferably between 0.01 h^{-1} and 0.15 h^{-1}.

Specifically, a feed medium is used which is chemically defined and comprising methanol.

The term “chemically defined” with respect to cell culture medium, such as a minimal medium or feed medium in a fed-batch process, shall mean a cultivation medium suitable for the *in vitro* cell culture of a production cell line, in which all of the chemical components and (poly)peptides are known. Typically, a chemically defined medium is entirely free of animal-derived components and represents a pure and consistent cell culture environment.

The term “cell” or “host cell” as used herein shall refer to a single cell, a single cell clone, or a cell line of a host cell. The term “host cell” shall particularly apply to a cell of methylotrophic yeast, which is suitably used for recombination purposes to produce a POI or a host cell metabolite. It is well understood that the term “host cell” does not include human beings. Specifically, host cells as described herein are artificial organisms and derivatives of native (wild-type) host cells. It is well understood that the host cells, methods and uses described herein, *e.g.,* specifically referring to those

comprising one or more genetic modifications, said heterologous expression cassettes or constructs, said transfected or transformed host cells and recombinant proteins, are non-naturally-occurring, “man-made” or synthetic, and are therefore not considered as a result of “law of nature”.

The term “cell line” as used herein refers to an established clone of a particular cell type that has acquired the ability to proliferate over a prolonged period of time. A cell line is typically used for expressing an endogenous or recombinant gene, or products of a metabolic pathway to produce polypeptides or cell metabolites mediated by such polypeptides.

The host cell producing the POI as described herein is also referred to as “production host cell”, and a respective cell line a “production cell line”. A “production cell line” is commonly understood to be a cell line ready-to-use for cell culture in a bioreactor to obtain the product of a production process, such as a POI.

Specific embodiments described herein refer to a Mut- production host cell, which can effectively use ADH2 to enzymatically process methanol thereby providing energy to the cell.

Specific embodiments described herein refer to a production cell line which is engineered to underexpress endogenous genes encoding the AOX1 and AOX2 proteins, and to overexpress a gene encoding ADH2, and is characterized by a high yield of POI production under the control of an ECP described herein, using methanol as a carbon source.

The term “cell culture” or “culturing” or “cultivation” as used herein with respect to a host cell refers to the maintenance of cells in an artificial, *e.g.*, an *in vitro* environment, under conditions favoring growth, differentiation or continued viability, in an active or quiescent state, of the cells, specifically in a controlled bioreactor according to methods known in the industry.

When culturing a cell culture using appropriate culture media, the cells are brought into contact with the media in a culture vessel or with substrate under conditions suitable to support culturing the cells in the cell culture. As described herein, a culture medium is provided that can be used for the growth of host cells *e.g.*, methylotrophic yeast. Standard cell culture techniques are well-known in the art.

The cell cultures as described herein particularly employ techniques which provide for the production of a secreted POI, such as to obtain the POI in the cell culture medium, which is separable from the cellular biomass, herein referred to as “cell culture

supernatant", and may be purified to obtain the POI at a higher degree of purity. When a protein (such as *e.g*., a POI) is produced and secreted by the host cell in a cell culture, it is herein understood that such proteins are secreted into the cell culture supernatant, and can be obtained by separating the cell culture supernatant from the host cell biomass, and optionally further purifying the protein to produce a purified protein preparation.

Cell culture media provide the nutrients necessary to maintain and grow cells in a controlled, artificial and *in vitro* environment. Characteristics and compositions of the cell culture media vary depending on the particular cellular requirements. Important parameters include osmolality, pH, and nutrient formulations. Feeding of nutrients may be done in a continuous or discontinuous mode according to methods known in the art.

Whereas a batch process is a cell culture mode in which all the nutrients necessary for culturing the cells are contained in the initial culture medium, without additional supply of further nutrients during fermentation, in a fed-batch process, after a batch phase, a feeding phase takes place in which one or more nutrients are supplied to the culture by feeding. Although in most processes the mode of feeding is critical and important, the host cell and methods described herein are not restricted with regard to a certain mode of cell culture.

A recombinant POI can be produced using the host cell and the respective cell line described herein, by culturing in an appropriate medium, isolating the expressed product or metabolite from the culture, and optionally purifying it by a suitable method.

Several different approaches for the production of the POI as described herein are preferred. A POI may be expressed, processed and optionally secreted by transfecting or transforming a host cell with an expression vector harboring recombinant DNA encoding the relevant protein, preparing a culture of the transfected or transformed cell, growing the culture, inducing transcription and POI production, and recovering the POI.

In certain embodiments, the cell culture process is a fed-batch process. Specifically, a host cell transformed with a nucleic acid construct encoding a desired recombinant POI, is cultured in a growth phase and transitioned to a production phase in order to produce a desired recombinant POI.

In another embodiment, host cells described herein are cultured in a continuous mode, *e.g.,* employing a chemostat. A continuous fermentation process is characterized by a defined, constant and continuous rate of feeding of fresh culture medium into a

bioreactor, whereby culture broth is at the same time removed from the bioreactor at the same defined, constant and continuous removal rate. By keeping culture medium, feeding rate and removal rate at the same constant level, the cell culture parameters and conditions in the bioreactor remain constant.

A stable cell culture as described herein is specifically understood to refer to a cell culture maintaining the genetic properties, specifically keeping the POI production level high, e.g. at least at a µg level, even after about 20 generations of cultivation, preferably at least 30 generations, more preferably at least 40 generations, most preferred of at least 50 generations. Specifically, a stable recombinant host cell line is provided which is considered a great advantage when used for industrial scale production.

The cell culture described herein is particularly advantageous for methods on an industrial manufacturing scale, *e.g*. with respect to both the volume and the technical system, in combination with a cultivation mode that is based on feeding of nutrients, in particular a fed-batch or batch process, or a continuous or semi-continuous process (*e.g*. chemostat).

The host cell described herein is typically tested for its capacity to express the GOI for POI production, tested for the POI yield by any of the following tests: ELISA, activity assay, HPLC, or other suitable tests, such as SDS-PAGE and Western Blotting techniques, or mass spectrometry.

To determine the effect of one or more genetic modifications on the underexpression or reduction of expression of the genes encoding the AOX1 and/or AOX2 protein(s) in the respective cell culture and *e.g.,* on their effect on POI production, the host cell line may be cultured in microtiter plates, shake flask, or bioreactor using fed-batch or chemostat fermentations in comparison with strains without such genetic modification(s) in the respective cell.

The production method described herein specifically allows for the fermentation on a pilot or industrial scale. The industrial process scale would preferably employ volumes of at least 10 L, specifically at least 50 L, preferably at least 1 m^3, preferably at least 10 m^3, most preferably at least 100 m^3.

Production conditions in industrial scale are preferred, which refer to *e.g.,* fed batch culture in reactor volumes of 100 L to 10 m^3 or larger, employing typical process times of several days, or continuous processes in fermenter volumes of approximately 50 – 1000 L or larger, with dilution rates of approximately 0.02 – 0.15 h^{-1}.

The devices, facilities and methods used for the purpose described herein are specifically suitable for use in and with culturing any desired cell line including prokaryotic and/or eukaryotic cell lines. Further, in embodiments, the devices, facilities and methods are suitable for culturing any cell type including suspension cells or anchorage-dependent (adherent) cells and are suitable for production operations configured for production of pharmaceutical and biopharmaceutical products—such as polypeptide products (POI), nucleic acid products (for example DNA or RNA), or cells and/or viruses such as those used in cellular and/or viral therapies.

In certain embodiments, the cells express or produce a product, such as a recombinant therapeutic or diagnostic product. As described in more detail herein, examples of products produced by cells include, but are not limited to, POIs such as exemplified herein including antibody molecules (*e.g*., monoclonal antibodies, bispecific antibodies), antibody mimetics (polypeptide molecules that bind specifically to antigens but that are not structurally related to antibodies such as *e.g*. DARPins, affibodies, adnectins, or IgNARs), fusion proteins (*e.g*., Fc fusion proteins, chimeric cytokines), other recombinant proteins (*e.g*., glycosylated proteins, enzymes, hormones), or viral therapeutics (*e.g*., anti-cancer oncolytic viruses, viral vectors for gene therapy and viral immunotherapy), cell therapeutics (*e.g*., pluripotent stem cells, mesenchymal stem cells and adult stem cells), vaccines or lipid-encapsulated particles (*e.g*., exosomes, virus-like particles), RNA (such as *e.g*. siRNA) or DNA (such as *e.g*. plasmid DNA), antibiotics or amino acids. In embodiments, the devices, facilities and methods can be used for producing biosimilars.

As mentioned, in certain embodiments, devices, facilities and methods allow for the production of eukaryotic cells, such as for example yeast cells, *e.g*., POIs including proteins, peptides, or antibiotics, amino acids, nucleic acids (such as DNA or RNA), synthesized by said cells in a large-scale manner. Unless stated otherwise herein, the devices, facilities, and methods can include any desired volume or production capacity including but not limited to bench-scale, pilot-scale, and full production scale capacities.

Moreover, and unless stated otherwise herein, the devices, facilities, and methods can include any suitable reactor(s) including but not limited to stirred tank, airlift, fiber, microfiber, hollow fiber, ceramic matrix, fluidized bed, fixed bed, and/or spouted bed bioreactors. As used herein, “reactor” can include a fermentor or fermentation unit, or any other reaction vessel and the term “reactor” is used interchangeably with “fermentor.” For example, in some aspects, an example bioreactor unit can perform one

or more, or all, of the following: feeding of nutrients and/or carbon sources, injection of suitable gas (*e.g.*, oxygen), inlet and outlet flow of fermentation or cell culture medium, separation of gas and liquid phases, maintenance of temperature, maintenance of oxygen and CO_2 levels, maintenance of pH level, agitation (*e.g.*, stirring), and/or cleaning/sterilizing. Example reactor units, such as a fermentation unit, may contain multiple reactors within the unit, for example the unit can have 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, or 100, or more bioreactors in each unit and/or a facility may contain multiple units having a single or multiple reactors within the facility. In various embodiments, the bioreactor can be suitable for batch, semi fed-batch, fed-batch, perfusion, and/or a continuous fermentation processes. Any suitable reactor diameter can be used. In embodiments, the bioreactor can have a volume between about 100 mL and about 50,000 L. Non-limiting examples include a volume of 100 mL, 250 mL, 500 mL, 750 mL, 1 liter, 2 liters, 3 liters, 4 liters, 5 liters, 6 liters, 7 liters, 8 liters, 9 liters, 10 liters, 15 liters, 20 liters, 25 liters, 30 liters, 40 liters, 50 liters, 60 liters, 70 liters, 80 liters, 90 liters, 100 liters, 150 liters, 200 liters, 250 liters, 300 liters, 350 liters, 400 liters, 450 liters, 500 liters, 550 liters, 600 liters, 650 liters, 700 liters, 750 liters, 800 liters, 850 liters, 900 liters, 950 liters, 1000 liters, 1500 liters, 2000 liters, 2500 liters, 3000 liters, 3500 liters, 4000 liters, 4500 liters, 5000 liters, 6000 liters, 7000 liters, 8000 liters, 9000 liters, 10,000 liters, 15,000 liters, 20,000 liters, and/or 50,000 liters. Additionally, suitable reactors can be multi-use, single-use, disposable, or non-disposable and can be formed of any suitable material including metal alloys such as stainless steel (*e.g.*, 316L or any other suitable stainless steel) and Inconel, plastics, and/or glass.

In embodiments and unless stated otherwise herein, the devices, facilities, and methods described herein can also include any suitable unit operation and/or equipment not otherwise mentioned, such as operations and/or equipment for separation, purification, and isolation of such products. Any suitable facility and environment can be used, such as traditional stick-built facilities, modular, mobile and temporary facilities, or any other suitable construction, facility, and/or layout. For example, in some embodiments modular clean-rooms can be used. Additionally, and unless otherwise stated, the devices, systems, and methods described herein can be housed and/or performed in a single location or facility or alternatively be housed and/or performed at separate or multiple locations and/or facilities.

Suitable techniques may encompass culturing in a bioreactor starting with a batch phase, followed by a short exponential fed batch phase at high specific growth rate, further followed by a fed batch phase at a low specific growth rate. Another suitable culture technique may encompass a batch phase followed by a fed-batch phase at any suitable specific growth rate or combinations of specific growth rate such as going from high to low growth rate over POI production time, or from low to high growth rate over POI production time. Another suitable culture technique may encompass a batch phase followed by a continuous culturing phase at a low dilution rate.

A preferred embodiment includes a batch culture to provide biomass followed by a fed-batch culture for high yields POI production.

It is preferred to culture a host cell as described herein in a bioreactor under growth conditions to obtain a cell density of at least 1 g/L cell dry weight, more preferably at least 10 g/L cell dry weight, preferably at least 20 g/L cell dry weight, preferably at least any one of 30, 40, 50, 60, 70, or 80 g/L cell dry weight. It is advantageous to provide for such yields of biomass production on a pilot or industrial scale.

A growth medium allowing the accumulation of biomass, specifically a basal growth medium, typically comprises a carbon source, a nitrogen source, a source for sulphur and a source for phosphate. Typically, such a medium comprises furthermore trace elements and vitamins, and may further comprise amino acids, peptone or yeast extract.

Preferred nitrogen sources include $NH_4H_2PO_4$, or NH_3 or $(NH_4)_2SO_4$;

Preferred sulphur sources include $MgSO_4$, or $(NH_4)_2SO_4$ or K_2SO_4;

Preferred phosphate sources include $NH_4H_2PO_4$, or H_3PO_4, or NaH_2PO_4, KH_2PO_4, Na_2HPO_4 or K_2HPO_4;

Further typical medium components include KCl, $CaCl_2$, and Trace elements such as: Fe, Co, Cu, Ni, Zn, Mo, Mn, I, B;

Preferably the medium is supplemented with vitamin B_7;

A typical growth medium for *P. pastoris* comprises glycerol, sorbitol or glucose, $NH_4H_2PO_4$, $MgSO_4$, KCl, $CaCl_2$, biotin, and trace elements.

In the production phase a production medium is specifically used with only a limited amount of a supplemental carbon source.

Preferably the host cell line is cultured in a mineral medium with a suitable carbon source, thereby further simplifying the isolation process significantly. An example of a preferred mineral medium is one containing an utilizable carbon source (in particular

methanol as described herein optionally in combination with *e.g.,* glucose, glycerol, or sorbitol), salts containing the macro elements (potassium, magnesium, calcium, ammonium, chloride, sulphate, phosphate) and trace elements (copper, iodide, manganese, molybdate, cobalt, zinc, and iron salts, and boric acid), and optionally vitamins or amino acids, *e.g.,* to complement auxotrophies.

Specifically, the cells are cultured under conditions suitable to effect expression of the desired POI, which can be purified from the cells or culture medium, depending on the nature of the expression system and the expressed protein, *e.g.,* whether the protein is fused to a signal peptide and whether the protein is soluble or membrane-bound. As will be understood by the skilled artisan, culture conditions will vary according to factors that include the type of host cell and particular expression vector employed.

A typical production medium comprises a supplemental carbon source, and further $NH_4H_2PO_4$, $MgSO_4$, KCl, $CaCl_2$, biotin, and trace elements.

For example the feed of the supplemental carbon source added to the fermentation may comprise a carbon source with up to 50 wt % utilizable sugars.

The fermentation preferably is carried out at a pH ranging from 3 to 8.

Typical fermentation times are about 24 to 120 hours with temperatures in the range of 20 °C to 35°C, preferably 22-30°C.

The POI is preferably expressed employing conditions to produce titers of at least 1 mg/L, preferably at least 10 mg/L, preferably at least 100 mg/L, most preferred at least 1 g/L.

The term "expression" or “expression cassette” as used herein refers to nucleic acid molecules containing a desired coding sequence and control sequences in operable linkage, so that hosts transformed or transfected with these sequences are capable of producing the encoded proteins or host cell metabolites. In order to effect transformation, the expression system may be included in a vector; however, the relevant DNA may also be integrated into a host cell chromosome. Expression may refer to secreted or non-secreted expression products, including polypeptides or metabolites.

Expression cassettes are conveniently provided as expression constructs *e.g*., in the form of “vectors” or “plasmids”, which are typically DNA sequences that are required for the transcription of cloned recombinant nucleotide sequences, *i.e*. of recombinant genes and the translation of their mRNA in a suitable host organism. Expression vectors or plasmids usually comprise an origin for autonomous replication or a locus for genome integration in the host cells, selectable markers (*e.g.,* an amino acid synthesis gene or

a gene conferring resistance to antibiotics such as zeocin, kanamycin, G418 or hygromycin, nourseothricin), a number of restriction enzyme cleavage sites, a suitable promoter sequence and a transcription terminator, which components are operably linked together. The terms “plasmid” and “vector” as used herein include autonomously replicating nucleotide sequences as well as genome integrating nucleotide sequences, such as artificial chromosomes *e.g*., a yeast artificial chromosome (YAC).

Expression vectors may include but are not limited to cloning vectors, modified cloning vectors and specifically designed plasmids. Preferred expression vectors described herein are expression vectors suitable for expressing of a recombinant gene in a eukaryotic host cell and are selected depending on the host organism. Appropriate expression vectors typically comprise regulatory sequences suitable for expressing DNA encoding a POI in a eukaryotic host cell. Examples of regulatory sequences include promoter, operators, enhancers, ribosomal binding sites, and sequences that control transcription and translation initiation and termination. The regulatory sequences are typically operably linked to the DNA sequence to be expressed.

Specific expression constructs described herein comprise a promoter operably linked to a nucleotide sequence encoding a POI under the transcriptional control of said promoter. Specifically, the promoter is not natively associated with the coding sequence of the POI.

To allow expression of a recombinant nucleotide sequence in a host cell, the expression cassette or vector described herein as GOIEC comprises an ECP, typically a promoter nucleotide sequence which is adjacent to the 5’ end of the coding sequence, *e.g.,* upstream from and adjacent to a gene of interest (GOI), or if a signal or leader sequence is used, upstream from and adjacent to said signal and leader sequence, respectively, to facilitate expression and secretion of the POI. The promoter sequence is typically regulating and initiating transcription of the downstream nucleotide sequence, with which it is operably linked, including in particular the GOI.

Specific expression constructs described herein comprise a polynucleotide encoding the POI linked with a leader sequence which causes secretion of the POI from the host cell. The presence of such a secretion leader sequence in the expression vector is typically required when the POI intended for recombinant expression and secretion is a protein which is not naturally secreted and therefore lacks a natural secretion leader sequence, or its nucleotide sequence has been cloned without its natural secretion leader sequence. In general, any secretion leader sequence effective to cause secretion

of the POI from the host cell may be used. The secretion leader sequence may originate from yeast source, *e.g*. from yeast α-factor such as MFa of *Saccharomyces cerevisiae*, or yeast phosphatase, from mammalian or plant source, or others.

In specific embodiments, multicloning vectors may be used, which are vectors having a multicloning site. Specifically, a desired heterologous gene can be integrated or incorporated at a multicloning site to prepare an expression vector. In the case of multicloning vectors, a promoter is typically placed upstream of the multicloning site.

The term "gene expression", or “expressing a polynucleotide” as used herein, is meant to encompass at least one step selected from the group consisting of DNA transcription into mRNA, mRNA processing, mRNA maturation, mRNA export, translation, protein folding and/or protein transport.

The term “increase expression” herein also referred to as "overexpression" refers to any amount higher than an expression level exhibited by a reference standard, which may be the host cell prior to the genetic alteration to increase expression of a certain polynucleotide, or which is otherwise expressed in a host cell of the same type or species which is not engineered to increase expression of said polynucleotide.

If a host cell does not comprise a given gene product, it is possible to introduce the gene product into the host cell for expression; in this case, any detectable expression is encompassed by the term “overexpression.”

Overexpression of a gene encoding a protein (such as ADH2), is also referred to as overexpression of a protein (such as ADH2). Overexpression can be achieved in any ways known to a skilled person in the art. In general, it can be achieved by increasing transcription/translation of the gene, e.g. by increasing the copy number of the gene or altering or modifying regulatory sequences or sites associated with expression of a gene. For example, the gene can be operably linked to a strong promoter in order to reach high expression levels. Such promoters can be endogenous promoters or heterologous, in particular recombinant promoters. One can substitute a promoter with a heterologous promoter which increases expression of the gene. Using inducible promoters additionally makes it possible to increase the expression in the course of cultivation. Furthermore, overexpression can also be achieved by, for example, modifying the chromosomal location of a particular gene, altering nucleic acid sequences adjacent to a particular gene such as a ribosome binding site or transcription terminator, introducing a frame-shift in the open reading frame, modifying proteins (*e.g*., regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription

of the gene and/or translation of the gene product, or any other conventional means of deregulating expression of a particular gene routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins or deleting or mutating the gene for a transcriptional factor which normally represses expression of the gene desired to be overexpressed. Prolonging the life of the mRNA may also improve the level of expression. For example, certain terminator regions may be used to extend the half-lives of mRNA. If multiple copies of genes are included, the genes can either be located in plasmids of variable copy number or integrated and amplified in the chromosome. It is possible to introduce one or more genes or genomic sequences into the host cell for expression.

According to a specific embodiment, a polynucleotide encoding the ADH2 protein can be presented in a single copy or in multiple copies per cell. The copies may be adjacent to or distant from each other. According to another specific embodiment, overexpression of the ADH2 protein employs recombinant nucleotide sequences encoding the ADH2 protein provided on one or more plasmids suitable for integration into the genome (*i.e*., knockin) of the host cell, in a single copy or in multiple copies per cell. The copies may be adjacent to or distant from each other. Overexpression can be achieved by expressing multiple copies of the polynucleotide, such as 2, 3, 4, 5, 6 or more copies of said polynucleotide per host cell.

A recombinant nucleotide sequence comprising a GOI and a polynucleotide (gene) encoding the ADH2 protein may be provided on one or more autonomously replicating plasmids, and introduced in a single copy or in multiple copies per cell.

Alternatively, the recombinant nucleotide sequence comprising a GOI and a polynucleotide (gene) encoding the ADH2 protein may be present on the same plasmid, and introduced in a single copy or multiple copies per cell.

A heterologous polynucleotide (gene) encoding the ADH2 protein or a heterologous recombinant expression construct comprising the polynucleotide (gene) encoding the ADH2 protein is preferably integrated into the genome of the host cell.

The term "genome" generally refers to the whole hereditary information of an organism that is encoded in the DNA (or RNA). It may be present in the chromosome, on a plasmid or vector, or both. Preferably, polynucleotide (gene) encoding the ADH2 protein is integrated into the chromosome of said cell.

The polynucleotide (gene) encoding the ADH2 protein may be integrated in its natural locus. “Natural locus” means the location on a specific chromosome, where the

polynucleotide (gene) encoding the ADH2 protein is located in a naturally-occurring wild-type cell. However, in another embodiment, the polynucleotide (gene) encoding the ADH2 protein is present in the genome of the host cell not at their natural locus, but integrated ectopically. The term "ectopic integration" means the insertion of a nucleic acid into the genome of a microorganism at a site other than its usual chromosomal locus, *i.e*., predetermined or random integration. In another embodiment, the polynucleotide (gene) encoding the ADH2 protein is integrated into the natural locus and ectopically. Heterologous recombination can be used to achieve random or non-targeted integration. Heterologous recombination refers to recombination between DNA molecules with significantly different sequences.

For yeast cells, the polynucleotide (gene) encoding the ADH2 protein and/or the GOI may be inserted into a desired locus, such as AOX1, GAP, ENO1, TEF, HIS4 (Zamir et al., Proc. NatL Acad. Sci. USA (1981) 78(6):3496-3500), HO (Voth et al. Nucleic Acids Res. 2001 June 15; 29(12): e59), TYR1 (Mirisola et al., Yeast 2007; 24: 761–766), His3, Leu2, Ura3 (Taxis et al., BioTechniques (2006) 40:73-78), Lys2, ADE2, TRP1, GAL1, ADH1 or on the integration of 5S ribosomal RNA gene.

In other embodiments, the polynucleotide (gene) encoding the ADH2 protein and/or the GOI can be integrated in a plasmid or vector. Preferably, the plasmid is a eukaryotic expression vector, preferably a yeast expression vector. Suitable plasmids or vectors are further described herein.

Overexpression of an endogenous or heterologous polynucleotide in a recombinant host cell can be achieved by modifying expression control sequences. Expression control sequences are known in the art and include, for example, promoters, enhancers, polyadenylation signals, transcription terminators, internal ribosome entry sites (IRES), and the like, that provide for the expression of the polynucleotide sequence in a host cell. Expression control sequences interact specifically with cellular proteins involved in transcription. Exemplary expression control sequences are described in, for example, Goeddel, Gene Expression Technology: Methods in Enzymology, Vol. 185, Academic Press, San Diego, Calif. (1990).

In a preferred embodiment, the overexpression is achieved by using an enhancer to express the polynucleotide. Transcriptional enhancers are relatively orientation and position independent, having been found 5' and 3' to the transcription unit, within an intron, as well as within the coding sequence itself. The enhancer may be spliced into the expression vector at a position 5' or 3' to the coding sequence, but is preferably

located at a site 5' from the promoter. Most yeast genes contain only one UAS, which generally lies within a few hundred base pairs of the cap site and most yeast enhancers (UASs) cannot function when located 3´ of the promoter, but enhancers in higher eukaryotes can function both 5´ and 3´ of the promoter.

Many enhancer sequences are known from mammalian genes (globin, RSV, SV40, EMC, elastase, albumin, a-fetoprotein and insulin). One may also use an enhancer from a eukaryotic cell virus, such as the SV40 late enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Specifically, the GOI and/or the ADH2 encoding polynucleotide (gene) as described herein, are operably linked to transcriptional and translational regulatory sequences that provide for expression in the host cells. The term "translational regulatory sequences" as used herein refers to nucleotide sequences that are associated with a gene nucleic acid sequence and which regulate the translation of the gene. Transcriptional and/or translational regulatory sequences can either be located in plasmids or vectors or integrated in the chromosome of the host cell. Transcriptional and/or translational regulatory sequences are located in the same nucleic acid molecule of the gene which it regulates.

Specifically, the overexpression of the ADH2 protein can be achieved by methods known in the art, for example by genetically modifying their endogenous regulatory regions, as described by Marx et al., 2008 (Marx, H., Mattanovich, D. and Sauer, M. *Microb Cell Fact* 7 (2008): 23), such methods include, for example, integration of a recombinant promoter that increases expression of a gene.

For example, overexpression of an endogenous or heterologous polynucleotide in a recombinant host cell can be achieved by modifying the promoters controlling such expression, for example, by replacing a promoter (*e.g*., an endogenous promoter or a promoter which is natively linked to said polynucleotide in a wild-type organism) which is operably linked to said polynucleotide with another, stronger promoter in order to reach high expression levels. Such promoter may be inductive or constitutive. Modification of a promoter may also be performed by mutagenesis methods known in the art.

In a preferred embodiment, expression of both, the polynucleotide encoding the ADH2 protein and the polynucleotide encoding the POI, is driven by an inducible promoter. In another preferred embodiment, expression of both, the polynucleotide

encoding the ADH2 protein and the polynucleotide encoding the POI, is driven by a constitutive promoter. In yet another preferred embodiment, expression of the polynucleotide encoding the ADH2 protein is driven by a constitutive promoter and expression of the polynucleotide encoding the POI is driven by an inducible promoter. In yet another preferred embodiment, expression of the polynucleotide encoding the ADH2 protein is driven by an inducible promoter and expression of the polynucleotide encoding the POI is driven by a constitutive promoter.

Specifically, a methanol-inducible promoter may be employed in expression constructs used to overexpress the gene encoding ADH2 and/or to express a GOI, as further described herein.

As an example, expression of the polynucleotide encoding the ADH2 protein may be driven by a constitutive GAP promoter and expression of the polynucleotide encoding the POI may be driven by the methanol-inducible AOX1 or AOX2 promoter.

In one embodiment, expression of the polynucleotides encoding the ADH2 protein and the POI is driven by the same promoter or same type of promoters in terms of promoter activity (*e.g.*, the promoter strength) and/or expression behaviour (*e.g.*, inducible or constitutive).

The term “reduce expression” herein also referred to as "underexpression" refers to any amount or level (*e.g.*, the activity or concentration) less than an expressed amount or level (*e.g.*, the activity or concentration) exhibited by a reference standard, which may be the host cell prior to the genetic alteration to reduce expression of a certain polynucleotide, or which is otherwise expressed in a host cell of the same type or species which is not engineered to lower expression of said polynucleotide. Reduction of expression as described herein specifically refers to a polynucleotide or gene encoding a defined AOX1 protein or AOX2 protein, in particular a gene that is endogenous to the host cell prior to engineering. In particular, the respective gene product is the defined AOX1 protein or AOX2 protein as described herein. Upon engineering the host cell by genetic modification to reduce expression of said gene the expression of said gene product or polypeptide is at a level which is less than the expression of the same gene product or polypeptide prior to a genetic modification of the host cell or in a comparable host which has not been genetically modified. “Less than” includes, e.g., 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80, 90% or more. No expression of the gene product or a polypeptide is also encompassed by the term “reduction of expression” or “underexpression.”

According to specific embodiments described herein, the host cell is engineered to knock-down or knockout (for inactivation or deletion of a gene or a part thereof) the endogenous host cell gene encoding the AOX1 protein or AOX2 protein (as defined herein, including *e.g*. homologues or orthologues of the sequences naturally-occurring in wild-type *P. pastoris*), or other (coding or non-coding) nucleotide sequences which confer the host cell's ability to express or produce said AOX1 protein or AOX2 protein.

Specifically, a deletion strain is provided, wherein a nucleotide sequence is disrupted.

The term "disrupt" as used herein refers to the significant reduction to complete removal of the expression or activity of one or more endogenous proteins in a host cell, such as by knock-down or knockout. This may be measured as presence of this one or more endogenous proteins in a cell culture or culture medium of the host cell, such as by mass spectrometry wherein the total content of an endogenous protein may be less than a threshold or non-detectable. Alternatively it may be measured as the enzymatic activity of the endogenous protein.

The term "disrupted" specifically refers to a result of genetic engineering by at least one step selected from the group consisting of gene silencing, gene knock-down, gene knockout, delivery of a dominant negative construct, conditional gene knockout, and/or by gene alteration with respect to a specific gene.

The term "knock-down", "reduction" or "deletion" in the context of gene expression as used herein refers to experimental approaches leading to reduced expression of a given gene compared to expression in a control cell. Knock-down of a gene can be achieved by various experimental means such as introducing nucleic acid molecules into the cell which hybridize with parts of the gene's mRNA leading to its degradation (*e.g.,* shRNAs, RNAi, miRNAs) or altering the sequence of the gene in a way that leads to reduced transcription, reduced mRNA stability, diminished mRNA translation, or reduced activity of the encoded protein.

A complete inhibition of expression of a given gene is referred to as "knockout". Knockout of a gene means that no functional transcripts are synthesized from said gene leading to a loss of function normally provided by this gene. Gene knockout is achieved by altering the DNA sequence leading to disruption or deletion of the gene or its regulatory sequences, or part of such gene or regulatory sequences. Knockout technologies include the use of homologous recombination techniques to replace, interrupt or delete crucial parts or the entire gene sequence or the use of DNA- modifying

enzymes such as zinc-finger or mega-nucleases to introduce double strand breaks into DNA of the target gene e.g., described by Gaj et al. (Trends Biotechnol. 2013;31(7):397-405).

Specific embodiments employ one or more knockout plasmids or cassettes which are transformed or transfected into the host cells. By homologous recombination the target gene in the host cells can be disrupted. This procedure is typically repeated until all alleles of the target gene are stably removed.

One specific method for knocking out a specific gene as described herein is the CRISPR-Cas9 methods as described in *e.g.,* Weninger et al. (J. Biotechnol. 2016, 235:139-49). Another method includes the split marker approach as described by e.g. Heiss et al. 2013 (Appl Microbiol Biotechnol. 97(3):1241-9.)

Another embodiment refers to target mRNA degradation by using small interfering RNA (siRNA) to transfect the host cell and targeting a mRNA encoding the target protein expressed endogenously by said host cell.

Expression of a gene may be inhibited or reduced by methods which directly interfere with gene expression, encompassing, but not restricted to, inhibition or reduction of DNA transcription, *e.g*., by use of specific promoter-related repressors, by site specific mutagenesis of a given promoter, by promoter exchange, or inhibition or reduction of translation, *e.g*., by RNAi or non-coding RNA induced post-transcriptional gene silencing. The expression of a dysfunctional, or inactive gene product with reduced activity, can, for example, be achieved by site specific or random mutagenesis, insertions or deletions within the coding gene.

The inhibition or reduction of the activity of gene product can, for example, be achieved by administration of, or incubation with, an inhibitor to the respective enzyme, prior to or simultaneously with protein expression. Examples for such inhibitors include, but are not limited to, an inhibitory peptide, an antibody, an aptamer, a fusion protein or an antibody mimetic against said enzyme, or a ligand or receptor thereof, or an inhibitory peptide or nucleic acid, or a small molecule with similar binding activity.

Gene silencing, gene knock-down and gene knockout refers to techniques by which the expression of a gene is reduced, either through genetic modification or by treatment with an oligonucleotide with a sequence complementary to either an mRNA transcript or a gene. If genetic modification of DNA is done, the result is a knock-down or knockout organism. If the change in gene expression is caused by an oligonucleotide binding to an mRNA or temporarily binding to a gene, this results in a temporary change

in gene expression without modification of the chromosomal DNA and is referred to as a transient knock-down.

In a transient knock-down, which is also encompassed by the above term, the binding of this oligonucleotide to the active gene or its transcripts causes decreased expression through blocking of transcription (in the case of gene-binding), degradation of the mRNA transcript (*e.g.,* by small interfering RNA (siRNA) or antisense RNA) or blocking mRNA translation.

Other approaches to carry out gene silencing, knock-down or knockout are known to the skilled person from the respective literature, and their application in the context of the present invention is considered as routine. Gene knockout refers to techniques by which the expression of a gene is fully blocked, *i.e*. the respective gene is inoperative, or even removed. Methodological approaches to achieve this goal are manifold and known to the skilled person. Examples are the production of a mutant which is dominantly negative for the given gene. Such mutant can be produced by site directed mutagenesis (*e.g*., deletion, partial deletion, insertion or nucleic acid substitution), by use of suitable transposons, or by other approaches which are known to the skilled person from the respective literature, the application of which in the context of the present invention is thus considered as routine. One example is knockout by use of targeted Zinc Finger Nucleases. A respective Kit is provided by Sigma Aldrich as "CompoZR knockout ZFN". Another approach encompasses the use of Transcription activator-like effector nucleases (TALENs).

The delivery of a dominant negative construct involves the introduction of a sequence coding for a dysfunctional gene expression product, *e.g*., by transfection. Said coding sequence is functionally coupled to a strong promoter, in such way that the gene expression of the dysfunctional enzyme overrules the natural expression of the gene expression product, which, in turn, leads to an effective physiological defect of the respective activity of said gene expression product.

A conditional gene knockout allows blocking gene expression in a tissue- or time-specific manner. This is done, for example, by introducing short sequences called loxP sites around the gene of interest. Again, other approaches are known to the skilled person from the respective literature, and their application in the context of the present invention is considered as routine.

One other approach is gene alteration which may lead to a dysfunctional gene product or to a gene product with reduced activity. This approach involves the

introduction of frame shift mutations, nonsense mutations (*i.e*., introduction of a premature stop codon) or mutations which lead to an amino acid substitution which renders the whole gene product dysfunctional, or causing a reduced activity. Such gene alteration can for example be produced by mutagenesis (*e.g*., deletion, partial deletion, insertion or nucleic acid substitution), either unspecific (random) mutagenesis or site directed mutagenesis. Protocols describing the practical application of gene silencing, gene knock-down, gene knockout, delivery of a dominant negative construct, conditional gene knockout, and/or gene alteration are commonly available to the skilled artisan, and are within his routine. The technical teaching provided herein is thus entirely enabled with respect to all conceivable methods leading to an inhibition or reduction of gene expression of a gene product, or to the expression of a dysfunctional, or inactive gene product, or with reduced activity.

Genetic modifications described herein may employ tools, methods and techniques known in the art, such as described by J. Sambrook et al., Molecular Cloning: A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, New York (2001).

The term “endogenous” as used herein is meant to include those molecules and sequences, in particular endogenous genes or proteins, which are present in the wild-type (native) host cell, prior to its modification to reduce expression of the respective endogenous genes and/or reduce the production of the endogenous proteins. In particular, an endogenous nucleic acid molecule (*e.g*., a gene) or protein that does occur in (and can be obtained from) a particular host cell as it is found in nature, is understood to be “host cell endogenous” or “endogenous to the host cell”. Moreover, a cell “endogenously expressing” a nucleic acid or protein expresses that nucleic acid or protein as does a host of the same particular type as it is found in nature. Moreover, a host cell “endogenously producing” or that “endogenously produces” a nucleic acid, protein, or other compound produces that nucleic acid, protein, or compound as does a host cell of the same particular type as it is found in nature.

Thus, even if an endogenous protein is no more produced by a host cell, such as in a knockout mutant of the host cell, where the protein encoding gene is inactivated or deleted, the protein is herein still referred to as “endogenous”.

The term “heterologous” as used herein with respect to a nucleotide sequence, construct such as an expression cassette, amino acid sequence or protein, refers to a compound which is either foreign to a given host cell, *i.e*. “exogenous”, such as not found

in nature in said host cell; or that is naturally found in a given host cell, *e.g*., is “endogenous”, however, in the context of a heterologous construct or integrated in such heterologous construct, *e.g.,* employing a heterologous nucleic acid fused or in conjunction with an endogenous nucleic acid, thereby rendering the construct heterologous. The heterologous nucleotide sequence as found endogenously may also be produced in an unnatural, *e.g.,* greater than expected or greater than naturally found, amount in the cell. The heterologous nucleotide sequence, or a nucleic acid comprising the heterologous nucleotide sequence, possibly differs in sequence from the endogenous nucleotide sequence but encodes the same protein as found endogenously. Specifically, heterologous nucleotide sequences are those not found in the same relationship to a host cell in nature. Any recombinant or artificial nucleotide sequence is understood to be heterologous. An example of a heterologous polynucleotide is a nucleotide sequence not natively associated with a promoter, *e.g.,* to obtain a hybrid promoter, or operably linked to a coding sequence, as described herein. As a result, a hybrid or chimeric polynucleotide may be obtained. A further example of a heterologous compound is a POI encoding polynucleotide operably linked to a transcriptional control element, *e.g*., a promoter, to which an endogenous, naturally-occurring POI coding sequence is not normally operably linked.

The term "operably linked" as used herein refers to the association of nucleotide sequences on a single nucleic acid molecule, *e.g.,* a vector, or an expression cassette, in a way such that the function of one or more nucleotide sequences is affected by at least one other nucleotide sequence present on said nucleic acid molecule. By operably linking, a nucleic acid sequence is placed into a functional relationship with another nucleic acid sequence on the same nucleic acid molecule. For example, a promoter is operably linked with a coding sequence of a recombinant gene, when it is capable of effecting the expression of that coding sequence. As a further example, a nucleic acid encoding a signal peptide is operably linked to a nucleic acid sequence encoding a POI, when it is capable of expressing a protein in the secreted form, such as a preform of a mature protein or the mature protein. Specifically, such nucleic acids operably linked to each other may be immediately linked, *i.e*. without further elements or nucleic acid sequences in between the nucleic acid encoding the signal peptide and the nucleic acid sequence encoding a POI.

The term “methylotrophic yeast” as used herein refers to of yeast genera and species which share a common metabolic pathway that enables them to use methanol

as a sole carbon source for their growth. In a transcriptionally regulated response to methanol induction, several of the enzymes are rapidly synthesized at high levels. Since the promoters controlling the expression of these genes are among the strongest and most strictly regulated yeast promoters, methylotrophic yeast are attractive as hosts for the large scale production of recombinant proteins.

The methylotrophic yeast as described herein is mutated by one or more genetic modifications to render it deficient in the methanol utilization pathway, in particular by underexpressing one or both of the genes encoding the endogenous AOX1 and AOX2 proteins, respectively. A methylotrophic yeast which is underexpressing or otherwise deficient in expressing both, the gene encoding the AOX1 protein and the gene encoding the AOX2 protein is herein understood as “Mut-“. For the purpose describe herein, such Mut- yeast is still referred to as “methylotrophic yeast”, because comprising a functional methanol utilization pathway prior to such genetic modification(s).

A “promoter” sequence is typically understood to be operably linked to a coding sequence, if the promoter controls the transcription of the coding sequence. If a promoter sequence is not natively associated with the coding sequence, its transcription is either not controlled by the promoter in native (wild-type) cells or the sequences are recombined with different contiguous sequences.

The promoter which is used for the purpose described herein, is herein referred to as “ECP”. The ECP may be a constitutive, inducible or repressible promoter. In a specific embodiment, the ECP is a promoter which is inducible by methanol and a methanol carbon source, respectively.

The ECP as described herein in particular initiates, regulates, or otherwise mediates or controls the expression of a POI coding DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms.

An inducible ECP as described herein is specifically understood as being a regulatable promoter, which has different promoter strength in the repressed and induced state. The term “regulatable” with respect to an inducible or repressible regulatory element, such as a promoter described herein shall refer to an element that is repressed in a host cell in the presence of an excess amount of a substance (such as a nutrient in the cell culture medium) *e.g.*, in the growth phase of a batch culture, and de-repressed to induce strong activity *e.g.*, in the production phase (such as upon upon feeding of a supplemental substrate, or adding methanol for methanol-induction),

according to a fed-batch strategy. A regulatory element can as well be designed to be regulatable, such that the element is inactive without addition of a cell culture additive, and active in the presence of such additive. Thus, expression of a POI under the control of such regulatory element can be induced upon addition of such additive.

The strength of the ECP specifically refers to its transcription strength, represented by the efficiency of initiation of transcription occurring at that promoter with high or low frequency. The higher transcription strength, the more frequently transcription will occur at that promoter. Promoter strength is a typical feature of a promoter, because it determines how often a given mRNA sequence is transcribed, effectively giving higher priority for transcription to some genes over others, leading to a higher concentration of the transcript. A gene that codes for a protein that is required in large quantities, for example, typically has a relatively strong promoter. The RNA polymerase can only perform one transcription task at a time and so must prioritize its work to be efficient. Differences in promoter strength are selected to allow for this prioritization.

A strong ECP is herein preferred, in particular an ECP which is relatively strong in the fully induced state, which is typically understood as the state of about maximal activity. The relative strength is commonly determined with respect to a comparable promoter, herein referred to as a reference promoter, which can be a standard promoter, such as the respective pGAP promoter of the cell as used as the host cell.

The frequency of transcription is commonly understood as the transcription rate, e.g. as determined by the amount of a transcript in a suitable assay, e.g. RT-PCR or Northern blotting. For example, the transcription strength of a promoter according to the invention is determined in the host cell which is *P. pastoris* and compared to the native pGAP promoter of *P. pastoris*.

The strength of a promoter to express a gene of interest is commonly understood as the expression strength or the capability of support a high expression level/rate. For example, the expression and/or transcription strength of a promoter of the invention is determined in the host cell which is *P. pastoris* and compared to the native pGAP promoter of *P. pastoris*.

The comparative transcription strength compared to a reference promoter may be determined by standard methods, such as by measuring the quantity of transcripts, e.g. employing a microarray, or else in a cell culture, such as by measuring the quantity of respective gene expression products in recombinant cells. In particular, the

transcription rate may be determined by the transcription strength on a microarray, Northern blot or with quantitative real time PCR (qRT-PCR) or with RNA sequencing (RNA-seq) where the data show the difference of expression level between conditions with high growth rate and conditions with low growth rate, or conditions employing different media composition, and a high signal intensity as compared to the reference promoter.

The expression rate may, for example, be determined by the amount of expression of a reporter gene, such as eGFP.

ECP as described herein exerts a relatively high transcription strength, *e.g.*, reflected by a transcription rate or transcription strength of at least 15% as compared to the native pGAP promoter in the host cell, also called “homologous pGAP promoter”. Preferably the transcription rate or strength is at least any one of 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 100%, or even higher, such as at least any one of 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, or 200%, or even higher, as compared to the native pGAP promoter, such as determined in the (*e.g*. eukaryotic) host cell selected as a host cell for recombination purpose to produce the POI.

The native pGAP promoter typically initiates expression of the *gap* gene encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which is a constitutive promoter present in most living organisms. GAPDH (EC 1.2.1.12), a key enzyme of glycolysis and gluconeogenesis, plays a crucial role in catabolic and anabolic carbohydrate metabolism.

The native pGAP promoter specifically is active in a recombinant eukaryotic cell in a similar way as in a native eukaryotic cell of the same species or strain, including the unmodified (non-recombinant) or recombinant eukaryotic cell. Such native pGAP promoter is commonly understood to be an endogenous promoter, thus, homologous to the host cell, and may serve as a standard or reference promoter for comparison purposes. The relative expression or transcription strength of a promoter as described herein is usually compared to the native pGAP promoter of a cell of the same species or strain that is used as a host for producing a POI.

The term “mutagenesis” as used herein shall refer to a method of providing mutants of a nucleotide sequence, e.g. through insertion, deletion and/or substitution of one or more nucleotides, so to obtain variants thereof with at least one change in the non-coding or coding region. Mutagenesis may be through random, semi-random or site directed mutation. Variants can be produced by a suitable mutagenesis method using a

parent sequence as a reference. Certain mutagenesis methods encompass those methods of engineering the nucleic acid or *de novo* synthesizing a nucleotide sequence using the respective parent sequence information as a template. Specific mutagenesis methods apply rational engineering of a mutant.

The term "nucleotide sequence" or “nucleic acid sequence” used herein refers to either DNA or RNA. "Nucleic acid sequence" or "polynucleotide sequence" or simply “polynucleotide” refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. It includes expression cassettes, self-replicating plasmids, infectious polymers of DNA or RNA, and non-functional DNA or RNA.

The term "protein of interest (POI)" as used herein refers to a polypeptide or a protein that is produced by means of recombinant technology in a host cell. More specifically, the protein may either be a polypeptide not naturally-occurring in the host cell, *i.e*. a heterologous protein, or else may be native to the host cell, *i.e*. a homologous protein to the host cell, but is produced, for example, by transformation with a self-replicating vector containing the nucleic acid sequence encoding the POI, or upon integration by recombinant techniques of one or more copies of the nucleic acid sequence encoding the POI into the genome of the host cell, or by recombinant modification of one or more regulatory sequences controlling the expression of the gene encoding the POI, *e.g.,* of a promoter sequence. In some cases the term POI as used herein also refers to any metabolite product by the host cell as mediated by the recombinantly expressed protein.

The term “sequence identity” of a variant, homologue or orthologue as compared to a parent nucleotide or amino acid sequence indicates the degree of identity of two or more sequences. Two or more amino acid sequences may have the same or conserved amino acid residues at a corresponding position, to a certain degree, up to 100%. Two or more nucleotide sequences may have the same or conserved base pairs at a corresponding position, to a certain degree, up to 100%.

Sequence similarity searching is an effective and reliable strategy for identifying homologs with excess (*e.g*., at least 50%) sequence identity. Sequence similarity search tools frequently used are *e.g*., BLAST, FASTA, and HMMER.

Sequence similarity searches can identify such homologous proteins or genes by detecting excess similarity, and statistically significant similarity that reflects common ancestry. Homologues may encompass orthologues, which are herein understood as

the same protein in different organisms, *e.g*., variants of such protein in different different organisms or species.

An orthologous sequence of the same protein in different organisms or species is typically homologous to the protein sequence, specifically of orthologs originating from the same genus. Typically, orthologs have at least about any one of 25%, 30%, 35%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, or 95% identity, up to 100% sequence identity.

Specifically, orthologs of a protein can be determined upon replacement of said protein or the gene encoding said protein by the orthologous sequences in a knock-out host cell, which host cell has been modified to knockout the respective gene or protein prior to such replacement.

For example, if a putative ADH2, AOX1 or AOX2 protein is functional in a *P. pastoris* host cell replacing the respective endogenous protein in a *P. pastoris* host cell in which the gene encoding such endogenous protein has been knocked out, such putative ADH2, AOX1 or AOX2 protein can be considered a respective homologue for the purpose described herein.

The AOX1 protein comprising or consisting of the amino acid sequence identified as SEQ ID NO:1 is of *K. phaffii* origin. It is well understood that there are homologous sequences present in other methylotrophic yeast host cells. For example, yeast of *Pichia pastoris* comprise the respective homologous sequences. *Pichia pastoris* has been reclassified into a new genus, *Komagataella*, and split into three species, *K. pastoris*, *K. phaffii*, and *K. pseudopastoris*.

Specific homologous sequences of SEQ ID NO:1 are *e.g*., found in *K. pastoris* (*e.g.*, SEQ ID NO:9, such as encoded by the nucleotide sequence comprising or consisting of SEQ ID NO:10), *Ogataea polymorpha* (*e.g.*, SEQ ID NO:19 such as encoded by the nucleotide sequence comprising or consisting of SEQ ID NO:20) or *Ogataea methanolica* (*e.g.*, SEQ ID NO:13 such as encoded by the nucleotide sequence comprising or consisting of SEQ ID NO:14).

The AOX2 protein comprising or consisting of the amino acid sequence identified as SEQ ID NO:3 is of *K. phaffii* origin. It is well understood that there are homologous sequences present in other methylotrophic yeast host cells. For example, yeast of *Pichia pastoris* comprise the respective homologous sequences. *Pichia pastoris* has been reclassified into a new genus, *Komagataella*, and split into three species, *K. pastoris*, *K. phaffii*, and *K. pseudopastoris*.

Specific homologous sequences of SEQ ID NO:3 are *e.g.*, found in *K. pastoris* (*e.g.*, SEQ ID NO:11, such as encoded by the nucleotide sequence comprising or consisting of SEQ ID NO:12), *Ogataea polymorpha* (*e.g.*, SEQ ID NO:19, such as encoded by the nucleotide sequence comprising or consisting of SEQ ID NO:20), or *Ogataea methanolica* (*e.g.*, SEQ ID NO:15, such as encoded by the nucleotide sequence comprising or consisting of SEQ ID NO:16).

Ogataea polymorpha has only one alcohol oxidase, herein exemplified by SEQ ID NO:19. Thus, reducing expression of AOX1 and AOX2 in *Ogataea polymorpha* as described herein is effectively carried out by reducing expression of the endogenous alcohol oxidase of *Ogataea polymorpha.*

Any homologous sequence of an AOX1 or AOX2 protein with a certain sequence identity described herein, in particular any such protein which is an ortholog of the *P. pastoris* AOX1 or AOX2 protein, is included in the definition of the respective AOX1 protein or AOX2 protein, as described herein.

The ADH2 protein comprising or consisting of the amino acid sequence identified as SEQ ID NO:50 is of *K. phaffii* origin. It is well understood that there are homologous sequences present in other methylotrophic yeast host cells. For example, yeast of *Pichia pastoris* comprise the respective homologous sequences*. Pichia pastoris* has been reclassified into a new genus, *Komagataella*, and split into three species, *K. pastoris*, *K. phaffii*, and *K. pseudopastoris*.

Specific homologous sequences of SEQ ID NO:50 are *e.g.*, any one of SEQ ID NO:52, 54, 56, 58, 60, 62, 64, 66, 68, or 70.

Any homologous sequence of an ADH2 protein with a certain sequence identity described herein, in particular any such protein which is an ortholog of the *P. pastoris* ADH2 protein, as described herein.

“Percent (%) amino acid sequence identity” with respect to an amino acid sequence, homologs and orthologues described herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific polypeptide sequence, after aligning the sequence and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

For purposes described herein, the sequence identity between two amino acid sequences is determined using the NCBI BLAST program version BLASTP 2.8.1 with the following exemplary parameters: Program: blastp, Word size: 6, Expect value: 10, Hitlist size: 100, Gapcosts: 11.1, Matrix: BLOSUM62, Filter string: F, Compositional adjustment: Conditional compositional score matrix adjustment.

"Percent (%) identity" with respect to a nucleotide sequence *e.g*., of a promoter or a gene, is defined as the percentage of nucleotides in a candidate DNA sequence that is identical with the nucleotides in the DNA sequence, after aligning the sequence and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent nucleotide sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

For purposes described herein (unless indicated otherwise), the sequence identity between two amino acid sequences is determined using the NCBI BLAST program version BLASTN 2.8.1 with the following exemplary parameters: Program: blastn, Word size: 11, Expect threshold: 10, Hitlist size: 100, Gap Costs: 5.2, Match/Mismatch Scores: 2,-3, Filter string: Low complexity regions, Mark for lookup table only.

The term “isolated” or “isolation” as used herein with respect to a POI shall refer to such compound that has been sufficiently separated from the environment with which it would naturally be associated, in particular a cell culture supernatant, so as to exist in “purified” or “substantially pure” form. Yet, “isolated” does not necessarily mean the exclusion of artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification. Isolated compounds can be further formulated to produce preparations thereof, and still for practical purposes be isolated - for example, a POI can be mixed with pharmaceutically acceptable carriers or excipients when used in diagnosis or therapy.

The term “purified” as used herein shall refer to a preparation comprising at least 50% (mol/mol), preferably at least 60%, 70%, 80%, 90% or 95% of a compound (*e.g*., a POI). Purity is measured by methods appropriate for the compound (*e.g.,*

chromatographic methods, polyacrylamide gel electrophoresis, HPLC analysis, and the like). An isolated, purified POI as described herein may be obtained by purifying the cell culture supernatants to reduce impurities.

As isolation and purification methods for obtaining a recombinant polypeptide or protein product, methods, such as methods utilizing difference in solubility, such as salting out and solvent precipitation, methods utilizing difference in molecular weight, such as ultrafiltration and gel electrophoresis, methods utilizing difference in electric charge, such as ion-exchange chromatography, methods utilizing specific affinity, such as affinity chromatography, methods utilizing difference in hydrophobicity, such as reverse phase high performance liquid chromatography, and methods utilizing difference in isoelectric point, such as isoelectric focusing may be used.

The following standard methods are preferred: cell (debris) separation and wash by Microfiltration or Tangential Flow Filter (TFF) or centrifugation, POI purification by precipitation or heat treatment, POI activation by enzymatic digest, POI purification by chromatography, such as ion exchange (IEX), hydrophobic interaction chromatography (HIC), Affinity chromatography, size exclusion (SEC) or HPLC Chromatography, POI precipitation of concentration and washing by ultrafiltration steps.

A highly purified product is essentially free from contaminating proteins, and preferably has a purity of at least 90%, more preferred at least 95%, or even at least 98%, up to 100%. The purified products may be obtained by purification of the cell culture supernatant or else from cellular debris.

An isolated and purified POI can be identified by conventional methods such as Western blot, HPLC, activity assay, or ELISA.

The term “recombinant” as used herein shall mean “being prepared by or the result of genetic engineering. A recombinant host may be engineered to delete and/or inactivate one or more nucleotides or nucleotide sequences, and may specifically comprise an expression vector or cloning vector containing a recombinant nucleic acid sequence, in particular employing nucleotide sequence foreign to the host. A recombinant protein is produced by expressing a respective recombinant nucleic acid in a host. The term “recombinant” with respect to a POI as used herein, includes a POI that is prepared, expressed, created or isolated by recombinant means, such as a POI isolated from a host cell transformed to express the POI. In accordance with the present invention conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art may be employed. Such techniques are explained

fully in the literature. See, *e.g.,* Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, (1982).

Certain recombinant host cells are “engineered” host cells which are understood as host cells which have been manipulated using genetic engineering, *i.e*. by human intervention. When a host cell is engineered to reduce expression or to underexpress a given gene or the respective protein, the host cell is manipulated such that the host cell has no longer the capability to express such gene and protein, respectively, compared to the host cell under the same condition prior to manipulation, or compared to the host cells which are not engineered such that said gene or protein is underexpressed.

The foregoing description will be more fully understood with reference to the following examples. Such examples are, however, merely representative of methods of practicing one or more embodiments of the present invention and should not be read as limiting the scope of invention.

EXAMPLES

Example 1: Generation of Methanol utilization negative strains of *Pichia pastoris*.

In order to generate the methanol utilization negative strains (Mut^-) two genes responsible for the methanol utilization named *AOX1* and *AOX2* were deleted from the genome of *Pichia pastoris* (syn. *Komagataella phaffii*).

a) For this purpose the *P. pastoris* strain (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands) was made electrocompetent. The strain was inoculated into 100 mL YPD media (main culture) for 16-20 hours (25°C; 180 rpm) and harvested at an optical density (OD_{600}) from 1.8 – 3 by centrifugation (5min; 1500g; 4°C) in two 50 mL falcon tubes. The cell pellet was resuspended in 10 mL YPD + 20mM HEPES + 25mM DTT and incubated (30 min; 25°C; 180 rpm). After the incubation period the falcon tubes were filled with 40 mL ice cold sterile distilled water and centrifuged (5min; 1500g; 4°C) (Eppendorf AG, Germany). The cell pellet was resuspended in ice cold sterile 1mM HEPES buffer, pH 8 and centrifuged (5min; 1500g; 4°C). The cell pellet was resuspended in 45 mL ice cold 1 M sorbitol and centrifuged (5min; 1500g; 4°C). The pellet was

resuspended in 500 µL ice cold 1M sorbitol and 80 µL aliquoted into ice cold 1.5 mL Eppendorf tubes. The aliquoted electro competent cells were kept at -80°C till used.

b) Cultivation of yeast strains was done in YPD media (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose) or YPD agar plates (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose, 20 g/L agar-agar) containing 500 µg/mL Geneticin or 200 µg/mL Hygromycin when necessary for selection.

c) For generating the deletion strains of *AOX1* and *AOX2* the split marker cassette approach was used (Gasser et al., 2013). The DNA fragments for generating the gene deletions are found in Table 1. The split marker cassette was carrying an antibiotic resistance cassette for Geneticin flanked by LoxP sites.

d) The deletions were done by adding 0.5 µg AOX1 split marker cassette 1 and 0.5 µg AOX1 split marker cassette 2 to the aliquoted electro competent cells and incubated for 5 min on ice. The electroporation was performed at 2 kV for 4 milliseconds (Gene Pulser, Bio-Rad Laboratories, Inc, USA). After transformation the electroporated cells were suspended in 1 mL YPD media and regenerated for 1.5 h to 3 h on 30°C shaking at 700 rpm on a thermoshaker (Eppendorf AG, Germany). Later 20 µL and 150 µL of the cell suspension was plated on YPD plates containing 500 µg/mL Geneticin for selection and incubated on 28°C for 48 hours. The colonies that appeared were re-streaked onto fresh YPD plates containing 500 µg/mL Geneticin. The correct disruption of *AOX1* locus was verified by PCR on genomic DNA using the primers AOX1_ctrl_Fwd and AOX1_ctrl_Rev (Table 2) binding outside of the deletion cassette. One clone was selected based on PCR amplification and sequencing of the PCR amplicon confirming correct deletion of the desired gene creating a *P. pastoris aox1*Δ::KanMX strain. A liquid culture was incubated from a single positive colony and made electrocompetent as explained in Example 1a), except for the addition of 500 µg/mL Geneticin to the liquid medium for generating the main culture. The strains were electroporated with 500 µg pTAC_Cre_Hph_Mx4 plasmid carrying a Cre recombinase needed for deletion of the Geneticin antibiotic resistance cassette between the LoxP regions and a Hygromycin resistance cassette for selection (Marx, Mattanovich, & Sauer, 2008). The electroporation was done

as described and selected for loss of Geneticin resistance by restreaking the transformants in parallel on YPD with 500 µg/mL Geneticin and 200 µg/mL Hygromycin and YPD plates with only 200 µg/mL Hygromycin. Clones were selected which were unable to grow on YPD plates with 500 µg/mL Geneticin and 200 µg/mL Hygromycin, but could grow on YPD plates with only 200 µg/mL Hygromycin. The successful deletion of the *AOX1* coding region and antibiotic marker was verified by PCR amplification with the primers AOX1_ctrl_Fwd and AOX1_ctrl_Rev (Table 2) and sequencing of the PCR amplicon (Microsynth AG, Swiss). The generated strain is called *P. pastoris* CBS2612 Δ*aox1* and has a Mut^S phenotype. It was selected for further genetic manipulation.

e) The *P. pastoris* CBS2612 Δ*aox1* was used to generate electro competent cells with the protocol described in a) and was electroporated with 0.5 µg AOX2 split marker cassette 1 and 0.5 µg AOX2 split marker cassette 2 (Table 1) with the procedure described in d). The antibiotic marker was removed with the same procedure as described in d). The successful deletion of the *AOX2* coding region and antibiotic marker was verified by PCR amplification with the primers AOX2_ctrl_fwd & AOX2_ctrl_rev (Table 2) and sequencing of the PCR amplicon (Microsynth AG, Swiss). The generated strain is called *P. pastoris* CBS2612 Δ*aox1*Δ*aox2* and has a methanol utilization negative (Mut^-) phenotype.

f) Genomic DNA for PCR amplifications was isolated with the Wizard® Genomic DNA Purification Kit (Promega Corporation, USA) as per the manufacturer´s recommendation. The PCR amplification reactions were done with the Q5 polymerase (New England Biolabs, Inc., USA) as per the manufacturer´s recommendations.

Table 1: Split marker cassette DNA sequence used for generating the AOX1 and AOX2 deletion strains.

DNA fragment	DNA sequence 5' to 3'
AOX1 split marker cassette 1	(SEQ ID NO:25) AGGGGTCCAAGTAAGAAGCTTCTTGCTGTAGAATTTGGGCATATGTGCTGGTGACAAAG GCATCTCTGCCTTGAGTTTCTGACGGCGGGACAGCATTCTTACCGGATATATAACACCA ATTGCCAGCACCACCAATCTCAGAGGTACCCCTAACAAACTTAATAAAATCTTGGGTAT CAACTTCATTAAGCTTTGTAGTTTGCAAGTACTTATAAACAAAATTCCGTAAGGTGTCG TCTTGAGGCTGGGACTTGACAAACTGCCAAAATGGCAACAAATCTACTGGCTTGGCCAT AATTTTGACATTCGAGTCATCAAAGGTAAATTCAACCGGAGACTTGTATTCTTTATTGA TAACTTTCTCATATAGGACATTGTCAGGAACACGATGAAACCAGGATGCCCCCAAATCC AATGAGACTGAGGTTTCATGAGTCGCAACCAACCTACCTCCAATACGGTCCCTACCCTC TAAAATCAACGCATTCACGCCATTGCTTTTGAGATCGACTGCAGCTTTGATGCCTGAAA TCCCAGCGCCTACAATGATGACATTTGGATTTGGTTGACTCATGTTGGTATTGTGAAAT AGACGCAGATCGGGAACACTGAAAAATAACAGTTATTATTCGAGATCTAACATCCAAAG ACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACACAT AAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACC TCCACTCCTCTTCTCCTCAACACCCACTTTTGCCATCGAAAAACCAGCCCAGTTATTGG GCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCTACTAACACCATGACTTT ATTAGCCTGTCTATCCTGGCCCCCCTGGCGAGGTTCATGTTTGTTTATTTCCGAATGCA ACAAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTC AGTACGCTGCAGGTCGACAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGT TATTAGGTCTAGATCGGTACCGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGA CGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCC ATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTAC GGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTG TCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCT TCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAA CATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAAC ATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGC GACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCA AAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAA TTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACT CACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAG GTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTT TGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAAT GAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTG AACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCG
AOX1 split marker cassette 2	(SEQ ID NO:26) AAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGT TACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCA AGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAA ACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCT GGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCG ATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCG AGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCA TAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATA ACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATC GCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTC ATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGC AGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTG TTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCA AATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTT AAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGG TACCATTCGAGAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATTAGG TGATATCAGATCCACTGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTT TATTTGTAACCTATATAGTATAGGATTTTTTTTGTCATTTTGTTTCTTCTCGTACGAGC TTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAGGGGTTTGGGAAA ATCATTCGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGAT

	TAAGTGAGACGTTCGTTTGTGCAAGCTTCAACGATGCCAAAAGGGTATAATAAGCGTCA TTTGCAGCATTGTGAAGAAAACTATGTGGCAAGCCAAGCCTGCGAAGAATGTATTTTAA GTTTGACTTTGATGTATTCACTTGATTAAGCCATAATTCTCGAGTATCTATGATTGGAA GTATGGGAATGGTGATACCCGCATTCTTCAGTGTCTTGAGGTCTCCTATCAGATTATGC CCAACTAAAGCAACCGGAGGAGGAGATTTCATGGTAAATTTCTCTGACTTTTGGTCATC AGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGA CAGTAAATGAAGTCCCACCAATAAAGAAATCCTTGTTATCAGGAACAAACTTCTTGTTT CGAACTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGG AGCGGGCAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGA TGCCAACTTCAGTGACAACGTTGCTATTTCGTTCAAACCATTCCGAATCCAGAGAAATC AAAGTTGTTTGTCTACTATTGATCCAAGCCAGTGCGGTCTTGAAACTGACAATAGTGTG CTCGTGTTTTGAGGTCATCTTTGT
AOX2 split marker cassette 1	(SEQ ID NO:27) GTACGGGTTTACTGATTTGACATATCTTGGTACTAACGTTACCAATGGTGTTCCAAATA ACGCAGATGATGAGCGTGGTTGCATTGCTGGATTTGACAACACTGGTTTCGTGCTGGGA ACTTCATCCTCGTTGTTTAATCAGTTTATTCTGCAACTGAATACGAGTGATCTTTCAGG AGCAATTTACCAAATCATTGAGCATTTTCTGACTGGACTTAGCGAAGACGAAGACGACA TTGCTATCTATTCCCCCAACCCTTTCTACAAAAGTACGTATGCAGGAGTAGGTGCCATT GCGGAAAATGACACCCTTTACTTGGTTGATGGTGGAGAGGATAACCAAAACGTCCCTCT GCAGCCTCTACTTCAAAAGGAGCGTGACGTTGATATCATCTTTGCGTTTGACAACAGTG CAGACACTGACCTCTCTTGGCCAAACGGTTCATCATTAGTCAACACCTACATGAGACAG TTTTCTTCTCAAGCAAATGGAACAACGTTCCCTTATGTACCTGATACCACCACTTTCCT AAACTTGAATCTTTCGAGTAAGCCAACCTTCTTTGGTTGTGATGCTAGAAATTTGACAG ACATTGTTGAAGGCACGGATCACATTCCTCCCCTGGTTGTTTATCTGGCCAATAGACCT TTCTCGTATTGGAGTAACACTTCAACTTTCAAGTTAGACTACTCTGAATCCGAGAAGAG AGGAATGATCCAAAACGGTTTTGAAGTGTCGTCTCGTTTGAACATGACTATTGATGAAG AATGGCGTACTTGTGTTGGATGTGCAATCATTCGTAGACAGCAGGAGAGATCCAATGCA ACACAAACAGAGCAATGTAGAAGATGTTTTGAGAATTATTGTTGGAACGGTGATATTGA CACTTCCACCGAAGATATCCCCGTTAATTTTACCACTACTGGAGCAACCAATGAGGAGA ATGACAACTCCACTTCAATATCATCGGCCAATTCGGTAGCACCTTCCAAACTTTGGTAC CAAGCACCATTGCTGTTGGTCGGCCTTGTCGCATTCTTCATCTAGTACGTACGCTGCAG GTCGACAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATTAGGTCTAG ATCGGTACCGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTC AGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCAT TTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTG CAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGC GCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATAC TTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCA TGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGAT TTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCG ATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTG CCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTT CCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGAT CCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTG TTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCT TTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTT GGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGA AAGAAATGCATAAGCTTTTGCCATTCTCACCG
AOX2 split marker cassette 2	(SEQ ID NO:28) AAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGT TACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCA AGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAA ACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCT GGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCG ATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCG AGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCA TAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATA ACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATC GCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTC ATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGC AGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTG

	TTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCA AATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTT AAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGG TACCATTCGAGAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATTAGG TGATATCAGATCCACTTCCTCTTACGGCTTTCTTTCCCAAAAAATCATTGGGGAAATGT GCCCCTCATCAGAGTCCAATGACCCATGAATAAAGTTTCTTGTACTGTTTAAGACGATG AATTGCAACGATAATCCGAGCAGTTTACGGGGTACATCACGTGCTTTGCATATGATCTC GGAGTCGGATCAGTTCCGGATGTGATGTATTACCCCATAGTTTCAAACTCTAATGCAGC CGCCAAGTGCCATACACCCTCCATCAATCTATGCTTAAAGTTTTTCACCATCGTTGGGT GGTGATGATGACTCGCTTAGTCTCTGCTGTTCGATATTAACTTTGTAAGGATCGCCCTT GGATGGAAAATTGAGGGGTTGTAACCTGAATTTGCAGGCTACTTACATTGGACTTTTGA GAAGGCTGGACGGTTGATGAAGAGGGCTGGGTGCAGAGGAATGGAAAAAAATTTAGTTG AGAGGACTGCTTGAAATTTTAGGAAATGGAGTCCTTTAAGCTGACAAAACTTCAAGGAT GGGGATTTTCATGTAGCTTTTTCATGCCTTCGACAAGCTAAAGGAAGGTAATTGATTCT GGATAAATGGATATTTGATCTGCTTTAGCAGATGTCAAAGTTCTACTAGTGATAGTCTG GTATCTCGTAGCCTTCAATTGGGCGTATCTTACTCGAAGTGTTATATTTTTAGCTGACG AGACGAAGAACGAGAGAGTATTGACACATTCAGAGGTAAGACAATATGTCGTATTATCA AAATAAGTATCGAACCTCTATTAGGAGCCTACTGGCTCAAATGTGCAACCTTAGTGGTG ATTGTCTCTGCTTCTTGATCACAATCTGTCGTGTTTGAGAGTGCCGATGTATGATTTTT AGTAAATGTTTTTCAGAAAAGGCGCTAAGTAAATAACCAGTAAGTAATAAATAACGTAA AAGTGATTTGAATCATAAAAGAATCAAGATAGAGGTCAAAGCATAGATAATCCCCC

Table 2: Polymerase chain reaction primers

Primer Name	DNA sequence 5' to 3'
AOX1_ctrl_Fwd	GGCTGGAAATAGATGTAGGGAG (SEQ ID NO:29)
AOX1_ctrl_Rev	TCGCATCTCCGCAAATTTCTC (SEQ ID NO:30)
AOX2_ctrl_fwd	GATCCCATTCCCTATCCATGT (SEQ ID NO:31)
AOX2_ctrl_rev	CTCTCCCCCTCGTAATCTT (SEQ ID NO:32)

<u>Example 2</u>: Production of intracellular eGFP with *P. pastoris* Δ*aox1* and *P. pastoris* Δ*aox1*Δ*aox2* under the methanol inducible *AOX1* promoter.

To test the protein producing ability and promoter activity the *P. pastoris* Δ*aox1*Δ*aox2* and *P. pastoris* Δ*aox1* strains were transformed with an expression constructs for enhanced green fluorescent protein (eGFP) (Table 4). The eGFP coding sequence was under the expression control of the P_{AOX1}: PP7435_chr4 (237941…238898).

a) The expression construct BB3aN_pAOX1_GFP_ScCYCtt was assembled using the Golden Gate assembly procedure as described (Prielhofer et al., 2017) from the plasmids BB1_12_pAOX1, BB1_23_eGFP, BB1_34_ScCYC1tt and BB3aN_14*. The plasmids and sequences are available in the Golden PiCS kit # 1000000133 (Addgene, Inc., USA). Before electroporation the plasmids were linearized with the restriction enzyme AscI (New England Biolabs, Inc., USA) as per the manufacturer´s protocol and purified with the Hi Yield® Gel/PCR DNA Fragment Extraction Kits (Süd-Laborbedarf GmbH, Germany). 500 ng of the

linearized plasmid was transformed into electrocompetent *P. pastoris* ∆*aox1*∆*aox2* and *P. pastoris* ∆*aox1* as previously described in Example **1a)** and **1d)**. Positive transformants were selected on YPD with 100 µg/L Nourseothricin and used for later screening experiments.

b) Small scale screening experiments of intracellular eGFP expression in the *P. pastoris* ∆*aox1*∆*aox2* and *P. pastoris* ∆*aox1*. Ten transformants from Example **2a)** were picked and inoculated into an overnight culture in 24 deep well plates containing 2 mL YPD and 100 µg/L Nourseothricin per well. All transformants were tested in duplicates. The 24 well plates were sealed with an air permissible membrane and incubated at 25°C and 280 rpm. For screening of the intracellular expression level of eGFP the overnight cultures were transferred to 24 deep well plates with 2.5 mL minimal ASMv6 media with 25 g/L polysaccharide and 0.3% amylase (m2p-labs GmbH, Germany) for slow glucose release and incubated for two hours followed by an addition of either 0.2% (v/v) or 1% (v/v) methanol for induction of eGFP production. eGFP measurements were done 4 and 20 hours after induction using a Gallios flow cytometer (Beckman Coulter , Inc., USA). For this purpose the cells were diluted to an OD_{600} of 0.5 in phosphate buffered saline containing 0.24 g/L KH_2PO_4, 1.8 g/L Na_2HPO_4*$2H_2O$, 0.2 g/L KCl, 8 g/L NaCl. 20,000 events were measured. FX values were calculated with the software FACS Express version 3 (De Novo Software, USA) using the equation:

$$FX = \left(\frac{FL1}{FSC^{1.5}}\right) * 8000$$

$$FX = Dimensionless\ value$$

$$FL1 = Fluorescence\ measured\ with\ a\ 505 - 545\ nm\ filter$$

$$FSC = Forward\ scatter$$

The method was already described (Ata, Prielhofer, Gasser, Mattanovich, & Çalık, 2017; Hohenblum, Borth, & Mattanovich, 2003; Prielhofer et al., 2013).

c) Minimal media ASMv6: 6.3 g/L $(NH_4)_2HPO_4$, 0.8 g/L $(NH_4)_2SO_4$, 0.49 g/L $MgSO_4$*$7H_2O$, 2.64 g/L KCl, 0.054 g/L $CaCl_2$*$2H_2O$, 22 g/L citric acid monohydrate, 1.47 mL/L PTM0 trace metals, 0.8 mg/L biotin 20 mL/L NH_4OH (25%); pH set to 6.5 with KOH.

d) The results are displayed in Table 3. Fluorescence was a proxy to determine the intracellular eGFP levels and the intracellular eGFP levels were a proxy for determining the activity of the P_{AOX1}. The results show that at 20 hours under 1%

methanol induction the promoter is active in the *P. pastoris* Δ*aox1*Δ*aox2* strain and eGFP is produced. The induction of the *P. pastoris* Δ*aox1*Δ*aox2* BB3aN_pAOX1_GFP_ScCYCtt strain is better at 1% than at 0.2% methanol after 20 h, no difference between methanol concentrations is observed in the *P. pastoris* Δ*aox1* strain.

Table 3: Results (FX values) with standard deviation from experiment described in Example 2b)

			negative control	
	P. pastoris Δ*aox1*Δ*aox2* BB3aN_pAOX1_GFP_ScCYCtt	*P. pastoris* Δ*aox1* BB3aN_pAOX1_GFP_ScCYCtt	*P. pastoris* Δ*aox1*Δ*aox2*	*P. pastoris* Δ*aox1*
0.2% MeOH at 4h	4.39 ± 0.90	6.26 ± 1.74	2.94	2.92
1% MeOH at 4h	5.59 ±1.09	7.39 ± 1.57	2.95	2.89
0.2% MeOH at 20h	37.81 ± 12.71	84.18 ± 4.57	2.37	2.70
1% MeOH at 20h	62.89 ± 8.81	89.85 ± 3.90	2.35	2.42

Table 4: Coding sequences of the Genes of interest expressed in *P. pastoris* Δ*aox1*Δ*aox2* and *P. pastoris* Δ*aox1*.

Gene of interest name	DNA sequence 5′ to 3′
enhanced green fluorescent protein (eGFP)	(SEQ ID NO:33) ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCG AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGA GGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGC AAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGC AGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTC CGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGA ACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG GCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGAC AAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGA CGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGA CCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA
Human serum albumin (HSA) with its native secretion leader	(SEQ ID NO:34) ATGAAGTGGGTTACTTTCATCTCCTTGTTGTTCTTGTTCTCCTCAGCTTACT CCAGAGGTGTTTTCAGAAGAGATGCTCACAAGTCCGAGGTTGCTCACAGATT CAAGGACTTGGGTGAAGAGAACTTCAAGGCTTTGGTTTTGATCGCTTTCGCT CAGTACTTGCAGCAGTGTCCATTCGAGGACCACGTTAAGTTGGTTAACGAGG TTACTGAGTTCGCTAAGACTTGTGTTGCTGACGAATCCGCTGAGAACTGTGA TAAGTCCTTGCACACTTTGTTCGGTGACAAGTTGTGTACTGTTGCTACTTTG AGAGAAACTTACGGTGAGATGGCTGACTGTTGTGCTAAGCAAGAGCCTGAGA GAAACGAGTGTTTCTTGCAACACAAGGACGACAACCCAAACTTGCCAAGATT GGTTAGACCAGAGGTTGACGTTATGTGTACTGCTTTCCACGACAACGAAGAG ACTTTCTTGAAGAAGTACTTGTACGAGATCGCTAGAAGACACCCATACTTCT ACGCTCCAGAGTTGTTGTTCTTCGCTAAGAGATACAAGGCTGCTTTCACTGA GTGTTGTCAGGCTGCTGATAAGGCTGCTTGTTTGTTGCCAAAGTTGGACGAG TTGAGAGATGAGGGTAAGGCTTCTTCCGCTAAGCAGAGATTGAAGTGTGCTT

	CCTTGCAGAAGTTCGGAGAGAGAGCTTTTAAGGCTTGGGCTGTTGCTAGATT GTCCCAGAGATTCCCAAAGGCTGAGTTCGCTGAGGTTTCCAAGTTGGTTACT GACTTGACTAAGGTTCACACAGAGTGTTGTCACGGTGACTTGTTGGAATGTG CTGATGACAGAGCTGACTTGGCTAAGTACATCTGTGAGAACCAGGATTCCAT CTCCTCCAAGTTGAAAGAATGTTGTGAGAAGCCTTTGTTGGAGAAGTCCCAC TGTATCGCTGAGGTTGAAAACGACGAAATGCCAGCTGACTTGCCATCTTTGG CTGCTGACTTCGTTGAATCCAAGGACGTCTGCAAGAACTACGCTGAGGCTAA GGACGTTTTCTTGGGTATGTTCTTGTATGAGTACGCTAGAAGACATCCAGAC TACTCCGTTGTTTTGTTGTTGAGATTGGCTAAGACTTACGAGACTACTTTGG AGAAGTGTTGTGCTGCTGCTGACCCACATGAGTGTTACGCTAAGGTTTTCGA CGAGTTCAAGCCATTGGTTGAGGAACCACAGAACTTGATCAAGCAGAACTGT GAGTTGTTCGAGCAGTTGGGTGAGTACAAGTTCCAGAACGCTTTGTTGGTTA GATACACTAAGAAGGTTCCACAGGTTTCCACTCCAACTTTGGTTGAGGTTTC CAGAAACTTGGGTAAGGTTGGTTCCAAGTGTTGTAAGCACCCAGAGGCTAAG AGAATGCCATGTGCTGAGGACTACTTGTCTGTTGTTTTGAACCAGTTGTGTG TCTTGCACGAAAAGACACCAGTTTCCGACAGAGTTACTAAGTGTTGTACTGA ATCCTTGGTTAACAGAAGACCTTGTTTCTCCGCTTTGGAGGTTGACGAGACT TACGTTCCAAAAGAGTTCAACGCTGAGACTTTCACTTTCCACGCTGACATCT GTACTTTGTCCGAGAAAGAGAGACAGATCAAGAAGCAGACTGCTTTGGTTGA GTTGGTTAAGCACAAGCCAAAGGCTACAAAAGAGCAGTTGAAGGCTGTTATG GACGACTTCGCTGCTTTCGTTGAGAAATGTTGTAAGGCTGACGACAAAGAGA CTTGTTTCGCTGAAGAGGGTAAGAAGTTGGTTGCTGCTTCCCAAGCTGCTTT GGGTCTGTAA
variable region of a camelid antibody (VHH) with the *S. cerevisiae* α-mating factor leader	(SEQ ID NO:35) ATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGCCTCCTCTGCAT TGGCTGCCCCTGTTAACACTACCACTGAAGACGAGACTGCTCAAATTCCAGC TGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTT TTGCCTTTCTCTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTA TCGCTTCCATTGCTGCTAAGGAAGAGGGTGTCTCTCTCGAGAAGAGACAAGC CGGTGGTTCATTAAGATTGTCCTGTGCTGCCTCTGGTAGAACTTTCACTTCT TTCGCAATGGGTTGGTTTAGACAAGCACCTGGAAAAGAGAGAGAGTTTGTTG CTTCTATCTCCAGATCCGGTACTTTAACTAGATACGCTGACTCTGCCAAGGG TAGATTCACTATTTCTGTTGACAACGCCAAGAACACTGTTTCTTTGCAAATG GACAACCTTAACCCAGATGACACCGCAGTCTATTACTGTGCCGCTGACTTGC ACAGACCATACGGTCCAGGAACCCAAAGATCCGATGAGTACGATTCTTGGGG TCAGGGAACTCAAGTCACTGTCTCTTCAGGTGGTGGATCTGGTGGTGGAGGT TCAGGTGGTGGAGGATCCGGTGGTGGTGGTTCTGGTGGTGGTGGATCTGGTG GAGGTGAAGTTCAACTTGTCGAATCCGGTGGTGCACTTGTCCAACCTGGTGG ATCTCTTAGACTTTCTTGTGCCGCCTCCGGTTTTCCTGTTAACCGTTACTCT ATGCGTTGGTACAGACAAGCCCCTGGAAAAGAACGTGAATGGGTTGCCGGAA TGTCCTCAGCTGGTGACAGATCCTCCTACGAAGATTCTGTGAAGGGACGTTT CACCATCTCCAGAGATGACGCCCGTAACACCGTTTACCTTCAAATGAACTCC CTTAAGCCTGAGGATACTGCCGTCTACTATTGTAACGTGAATGTCGGATTTG AATACTGGGGACAGGGAACCCAAGTTACTGTCTCTTCCGGTGGACATCACCA CCACCATCACTAATAG

<u>Example 3:</u> Generation of *P. pastoris* Δ*aox1* and *P. pastoris* Δ*aox1*Δ*aox2* strains producing secreted HSA and VHH under the methanol inducible *AOX1* promoter.

To test the ability to produce secreted recombinant proteins in the *P. pastoris* Δ*aox1*Δ*aox2* strain and compare it with the *P. pastoris* Δ*aox1* strains, the strains were transformed with expression constructs for two secreted model proteins: (1) Human serum albumin with its native secretion leader (HSA) or (2) variable region of a camelid antibody with the *S. cerevisiae* α-mating factor secretion leader (VHH). The coding sequence of these genes of interest (codon-optimized and synthesized by external providers) can be found in Table 4.

a) The pPM2pN21_pAOX1_HSAopt_CycTT and pPM2pZ30_pAOX1_αMF-vHH_CycTT expression constructs used for HSA and VHH production are derivatives of the pPuzzle ZeoR vector described in WO2008128701A2, consisting of the *E.coli* pUC19 ori and the Zeocin antibiotic resistance cassette. In this case of pPM2pN21_pAOX1_HSAopt_CycTT the Zeocin resistance is exchanged for Nourseothricin resistance via restriction and ligation. Additionally the vectors are carrying an integration sequence that is homologous to the PGI locus PP7435_Chr3 (1366329…1367193) for efficient integration. The expression vector is described in more details elsewhere (Gasser et al., 2013; Stadlmayr et al., 2010). Expression of the gene of interest (GOI) was mediated by the P_{AOX1} PP7435_chr4 (237941…238898) and the *Saccharomyces cerevisiae* CYC1 transcription terminator. The gene for human serum albumin (HSA) (GenBank NP_000468) was codon optimized for *P. pastoris* and synthesized. It has a native secretion leader and is therefore secreted into the supernatant. The gene for VHH is codon optimized for *P. pastoris* and synthetized (Table 4), it has an N-terminal *S. cerevisiae* α-mating type leader for secretion into the supernatant. For the purpose of transformation of the expression constructs the circular vectors were linearized by restriction in the PGI1 homologous sequence with XmnI (New England Biolabs, Inc., USA) and purified with the Hi Yield® Gel/PCR DNA Fragment Extraction Kits (Süd-Laborbedarf GmbH, Germany).

b) Electroporation of electrocompetent *P. pastoris* ∆*aox1*∆*aox2* and *P. pastoris* ∆*aox1* with 500 ng linearized pPM2pN21_pAOX1_HSAopt_CycTT and pPM2pZ30_pAOX1_αMF-vHH_CycTT plasmid and selection were carried out as previously described in Example **1a)** and Example **1d)**. The selection was carried out on YPD plates with 100 µg/mL Nourseothricin or 25 µg/mL Zeocin, respectively.

<u>Example 4</u>: Small scale screening of the HSA and VHH producing *P. pastoris* ∆*aox1*∆*aox2* and *P. pastoris* ∆*aox1*.

a) For the pre-culture the transformants were inoculated in 2 mL YPD with 100 µg/mL Nourseothricin or 25 µg/mL Zeocin based on the antibiotic resistance used for selection. For each expression construct twelve transformants were picked for screening. Pre-culture and screening cultures were cultivated in 24 well plates sealed with an air permeable membrane and incubated on 25°C on 280 rpm. The screening culture was inoculated with a start optical density (OD_{600}) of 8 into 2 mL of minimal media (ASMv6) with a slow glucose release system based on 6 mm feedbeads (Kuhner Shaker GmbH, Germany) to keep the cultures in glucose limit. The strains were compared with different methanol feed procedures differing in total methanol received and duration (Table 5).

b) After the incubation period 1 mL of the each culture was removed and centrifuged in a pre-weighted Eppendorf tube. The supernatant was removed and the protein concentration was measured with the Caliper LabChip GXII Touch (PerkinElmer, inc., USA) as per the manufacturer´s instructions. The wet cell weight was determined by weighting the Eppendorf tube with the cell pellet and calculated as follows: Weight (full) – Weight (empty) = Wet cell weight (WCW) (g/L). Out of this data the yield was calculated: Yield (µg/g) = Protein concentration / Wet cell weight. Data of transformants that had double the concentration or had no detectible protein in the supernatant were removed from analysis as outliers. The outliers are considered as transformants that have either two copies of the expression construct or no copy at all (Aw & Polizzi, 2013; Schwarzhans et al., 2016).

Table 5: Overview of the screening strategies used for testing the secreted protein production yield of the transformed strains. * The first shot was 0.5% methanol.

Protocol	Incubation period	Feedbeads	Start OD_{600}	Methanol pulse	Total methanol	Methanol shot time points
Standard	48 h	12 mm	8	4 x	3.5% (v/v)	4 h*, 19 h, 27 h, 43 h
One shot	48 h	3x6 mm	8	1 x	1% (v/v)	3 h
Two shot	48 h	3x6 mm	8	2 x	2% (v/v)	3 h, 23 h
One shot - extended	72 h	3x6 mm	8	1 x	1% (v/v)	3 h
Two shot - extended	72 h	3x6 mm	8	2 x	2% (v/v)	3 h, 43 h

c) The results show that the *P. pastoris* Δ*aox1*Δ*aox2* can produce secreted proteins under the induction of the P_{AOX1} and that the yield is comparable to the *P. pastoris* Δ*aox1* used as industry standard (Table 6). In the "Two shot – extended" strategy the *P. pastoris* Δ*aox1*Δ*aox2* shows a better yield indicating that under longer cultivation times with less methanol the *P. pastoris* Δ*aox1*Δ*aox2* has an yield advantage. Furthermore this shows that it is possible to use limited glucose conditions to screen *P. pastoris* Δ*aox1*Δ*aox2* strains producing secreted proteins controlled by the P_{AOX1} and methanol induction.

Table 6: Average secreted product yield with standard deviation in µg product / g WCW of the tested strains in different screening conditions.

	Standard	One shot	Two shot	One shot extended	Two shot extended
P. pastoris Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT	733 ± 59	219 ± 32	379 ± 38	0	410 ± 16
P. pastoris Δ*aox1* pPM2pZ30_pAOX1_αMF-vHH_CycTT	1465 ± 239	195 ± 66	413 ± 114	0	239 ± 77
P. pastoris Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT	443 ± 111	138 ± 24	251 ± 110	228 ± 33	363 ± 55
P. pastoris Δ*aox1* pPM2pN21_pAOX1_HSAopt_CycTT	840 ± 36	79 ± 9	388 ± 61	83 ± 18	277 ± 23

Example 5: Bioreactor cultivations

To determine the behavior and process parameters of *P. pastoris* Δ*aox1*Δ*aox2* in fed-batch mode in a recombinant protein production setting, bioreactor cultivations were performed. The cultivations were performed as follows.

a) DASGIP bioreactors were used with a working volume of 0.7 L (Eppendorf AG, Germany). One Bioreactor system consists of four reactors that are arranged in one bio-block for controlling the temperature. Each reactor was connected to 4 peristaltic pumps that were software controlled. Additionally each reactor had 2 balances available that were connected to the DASGIP control software (Eppendorf AG, Germany) for adjusting the pump speed gravimetrically. Each reactor was connected with a controllable gas supply (pressured air, N_2, O_2 could be mixed in any desired amount) and a gas analyzer for O_2 and CO_2 concentration measurement in the reactor off gas. The reactors had a pH probe and Dissolved Oxygen (DO) probe connected to the DASGIP control software. The DASGIP control software was recording all parameters in one minute intervals.
b) The bioreactor cultivation media consisted of BSM medium (Mellitzer et al., 2014): 11.48 g/L H_3PO_4, 0.5 g/L $CaCl_2*2H_2O$, 7.5 g/L $MgSO_4*7H_2O$, 9 g/L K_2SO_4, 2 g/L KOH, 40g/L Glycerol, 0.25 g/L NaCl, 4.35 mL/L PTM0, 0.87 mg/L Biotin, 0.1 mL/L Glanapon 2000, pH set to 5.5 with 25% NH_3.
c) PTM0 consisted of: 6.0 g $CuSO_4*5H_2O$, 0.08 g NaI, 3.36 g $MnSO_4*H_2O$, 0.2 g $Na_2MoO_4*2H_2O$, 0.02 g H_3BO_3, 0.82 g $CoCl_2$, 20.0 g $ZnCl_2$, 65.0 g $FeSO_4*7H_2O$, 5 mL/L H_2SO_4 (95 %-98 %).
d) The Glucose feed media consisted of: 50% (w/w) glucose, 2.08 mg/kg Biotin, 10.4 mL/kg PTM0. The methanol feed media was: 50% (v/v) or 100% (v/v) methanol. The glycerol feed media consisted of: 60% (w/w) glycerol, 2.08 mg/kg Biotin, 10.4 mL/kg PTM0.
e) The Dissolved Oxygen (DO) set point was 20%. In certain cases the DO control was deactivated and the agitation and aeration were manually set to a constant 750 rpm and 9.5 sL/h. The pH was set to 5.0 or 5.5 with either 12.5% or 25% NH_3 controlled by the DASGIP control software. Acid control was achieved with 10% H_3PO_4 by manual addition when necessary. The temperature was set at 25°C. The start OD_{600} was 2 and the start volume was 300 mL plus 15 mL of inoculation culture.

f) Sampling was done on a daily basis (approximately every 24 hours). First a 3 mL aspirate was taken from the reactor to remove the dead volume of the sampling port. Then 9 mL of sample were taken. 3 x 2 mL were pipetted into reweighted 2 mL Eppendorf tubes and 1 x 1.5 mL into one 1.5 mL Eppendorf tube. The samples were centrifuged (16,000 g, 10 min, 4°C). The supernatant was collected for protein and HPLC analysis and stored at -20°C. The pellet was washed by resuspension in 1 mL 0.1 M HCl to remove trace salts and centrifuged again (16,000 g, 10 min, 4°C). The pellet was then dried for 24 hours at 105°C to determine the dry cell weight. The dry cell weight was calculated as follows: (Weight (full) – Weight (empty)) / 2 = Dry cell weight (g/L) and calculated as the average of three replicates. If only a HPLC sample was need only 2 mL of sample were taken.

g) Cell viability was measured by staining the cell suspension with propidium iodide. For this the cell suspension from the reactor sample was diluted with phosphate buffered saline to on OD_{600} of 0.5 and mixed with a stock solution of propidium iodide to a final concentration of 10 µM prior to measurement with the Gallios flow cytometer (Beckman coulter , Inc., USA) with a filter of 590 - 650 nm. 50,000 events were measured per sample.

<u>Example 6</u>: Determining the evaporation rate of methanol from the Bioreactors without cells.

To assess the evaporation rate of methanol from the reactors by aeration and agitation the reactors were filled with sterile media and pulsed with methanol, samples were taken to determine the methanol concentration.

a) For this example two reactors were filled with 310 mL of BSM media without glycerol and two reactors were filled with 500 mL BSM media without glycerol to simulate the media at the end of the batch phase where the glycerol is consumed by the growing culture.

b) The reactor stirrer speed was set to 760 rpm and gassing to 9.5 sL/h as would be the case in a cultivation of the P*. pastoris* Δ*aox1*Δ*aox2*. The parameters can be found in Table 7.

c) A 50% (v/v) methanol pulse was added manually to increase the methanol concentration to 1% (v/v) and a sample was taken to determine the actual achieved concentration. Samples were taken at 3.4, 6.5, 22.4, 31.0, 47.9 hours by first removing 3 mL of dead volume from the sample port and discarding the aspirate. Immediately after that a 4 mL sample was taken.

d) HPLC measurement of methanol concentration were done as described previously (Blumhoff, Steiger, Marx, Mattanovich, & Sauer, 2013). For identification and quantification pure standards were used. The column was an Aminex HPX-87H (Bio-Rad Laboratories, Inc, USA) run at 60°C with a 4 mM H_2SO_4 mobile phase at 0.6 mL/min. The detector was a refraction index detector RID-10 A (Shimadzu, Corp., Japan) and the calculations were done with the LabSolutions v5.85 software (Shimadzu, Corp., Japan).

e) The evaporation rate was calculated only from the first and last sample with the biggest time and concentration difference. The changes in concentration between the adjacent samples were marginal and measurement error could have a significant impact on the calculation. The data can be found in Table 8**.** R1 and R2 filled with 500 mL media had a mean value of 0.063 $g*L^{-1}*h^{-1}$.

Table 7: Reactor parameters and methanol pulse volume.

Reactor	Volume (mL)	Agitation (rpm)	Gassing (sL/h)	50% (v/v) methanol (ml)
R1	310	760	9.5	6.2
R2	310	760	9.5	6.2
R3	500	760	9.5	10
R4	500	760	9.5	10

Table 8: Methanol concentration at the sampling timepoints and the calculated evaporation rate.

	Time (h)						
	0	3.4	6.5	22.4	31.0	46.9	
Reactor	Methanol (g/L)						dc/dt ($g\ L^{-1}\ h^{-1}$)
R1	7.91	6.13	7.53	6.76	6.10	5.49	0.052
R2	7.80	7.14	7.49	6.72	6.28	4.29	0.075
R3	7.47	7.73	5.07*	7.05	6.93	6.43	0.022
R4	7.54	7.80	6.33*	6.97	4.39*	6.49	0.022

*are too low and are considered as outliers.

<u>Example 7</u>: Determining the methanol uptake rate of *P. pastoris* Δ*aox1*Δ*aox2*

To determine the methanol uptake rate the *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT strain was cultivated in a bioreactor. The culture was grown till a certain biomass concentration. Then a methanol pulse was applied and samples were taken immediately after the pulse and approximately 20 hours later. The goal was to determine the methanol uptake rate of the Mut^- strain and compare it to the methanol evaporation rate measured in Example 6.

a) Pre-culture: 24 hours prior to reactor inoculation 50 mL YPD containing 100 µg/L Nourseothricin were inoculated with *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT. 3 hours prior to inoculation of the reactors the pre-culture was diluted by another 50 mL YPD containing 100 µg/L Nourseothricin. Before inoculation the appropriate amount of culture was centrifuged (1500g, 5 min, 20°C) and resuspended in 15 mL of BSM media with an OD_{600} of 42.
b) The reactors filled with 300 mL BSM media were inoculated with 15 mL of *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT culture. The target inoculation OD_{600} in the reactor was 2. At the end of the batch phase as indicated by a dissolved oxygen spike, a 50% (w/w) glucose feed was started at 2.4 mL/h for 24 hours to increase the biomass. Two hours after the glucose feed start a 9.5 mL 50% (v/v) methanol shot was given to increase the methanol concentration to 1.5% (measured concentration was R1 = 1.47% and R2= 1.48%). This was done to induce methanol consumption. At the end of the glucose feed phase samples for cell dry weight and HPLC were taken.
c) After the glucose phase the agitation and gassing was set to a constant 750 rpm and 9.5 sL/h. An additional 50% methanol pulse was added to increase the concentration to 1.5% and immediately a sample was taken (measured concentration was R1 = 1.36% and R2= 1.36%). The concentration was measured again after 19.5 hours and used to determine the specific methanol uptake rate ($q_{methanol}$).
d) The methanol concentration decrease (dc/dt) for this experiment was substantially higher at 0.37 g L^{-1} h^{-1} on average compared to the values obtained for the evaporation rate in Example 6e), Table 8 that ranges from 0.022 to 0.063 g L^{-1} h^{-1}.

The specific methanol uptake rate was calculated based on the data represented in Table 9 as follows.

$$q_{methanol} = ((C^{t0}_{methanol} - C^{t19.5}_{methanol}) / C_{biomass}) / (t_0 - t_{19.5})$$

The volume was constant over the measured time period. The average specific methanol uptake rate ($q_{methanol}$) without subtracted evaporation was 5.07 mg g^{-1} h^{-1}. For the calculations of the specific methanol uptake rate with subtracted evaporation an evaporation rate of 22 mg L^{-1} h^{-1} was estimated based on the results in Example 6e), Table 8. This outcome was completely new and unexpected. Till now it was reported and accepted in published literature that the Mut^- is unable to metabolise methanol and that the decrease of methanol is due to evaporation loss (Looser et al., 2015).

Table 9: Data overview specific methanol uptake rate ($q_{methanol}$) and apparent methanol loss (dc/dt).

Reactor	Volume (mL)	CDW (g/L)	Methanol at 0 h (g/L)	Methanol at 19.5 h (g/L)	dc/dt (g L^{-1} h^{-1})	$q_{methanol}$ (mg g^{-1} h^{-1})	$q_{methanol\text{-}evaporated}$ (subtracted evaporation) (mg g^{-1} h^{-1})
R1	378	73.4	10.71	3.67	0.361	4.92	4.61
R2	372	72.6	10.74	3.35	0.379	5.22	4.90

<u>Example 8</u>: Cultivation strategy 1 – Applying a constant glucose/methanol co-feed to the *P. pastoris* ∆*aox1*∆*aox2*

The *P. pastoris* ∆*aox1*∆*aox2* pPM2pN21_pAOX1_HSAopt_CycTT strain was cultivated in a recombinant protein production scenario. The strain was fed with a constant limited glucose feed and induced with methanol for protein production.

a) The inoculation was done as described in Example 7a) b). The cultivation was separated into two phases. (1) Phase one: The batch was started at OD_{600} of 2 in BSM medium. The batch phase end was indicated by a dissolved oxygen spike at 22.27 h for reactor R1 and 21.52 h for reactor R2.
b) (2) Phase two: A fed-batch phase with a 50% (w/w) glucose feed was started after phase one at a feed rate of 2.4 mL/h for 97 hours. At the same time a 50% (v/v) methanol pulse was added with the aim to increase the methanol concertation to 1.5% (v/v). A HPLC sample was taken as described in Example 6d) to measure

the exact concentration and an additional pulse was added if necessary. A methanol feed calculated based on the predicted biomass concentration and the specific methanol uptake rate of 5 mg g^{-1} h^{-1} as measured in Example 7d) was applied. The methanol concentration was measured daily at line by HPLC.

c) The methanol feed was calculated in hourly intervals as follows:

$$R_{methanol} = q_{methanol} * X_{predicted} * t_{interval}$$

$$T_{methanol} = T_{methanol-\,previous\,interval} - R_{methanol} + A_{methanol-\,previous\,interval} - \left(\frac{T_{methanol-previous\,interval}}{V_{reactor-previous\,interval}} * V_{sample-previous\,interval}\right)$$

$$A_{methanol} = V_{reactor} * C_{methanol,target} - T_{methanol}$$

$$F_{methanol} = \left(\frac{A_{methanol}}{\rho_{methanol}}\right) * 0.002$$

$$V_{reactor} = V_{reactor-previous\,interval} + F_{Glucose} + F_{methanol}$$

$$X_{predicted} = X_{predicted-previous} + \left(Y_{\left(\frac{x}{s}\right)} * \left(F_{glucose} * \rho_{50\%\,glucose} * 50\%\right)\right) - \left(V_{sample} * \left(\frac{X_{predicted-previous\,interval}}{V_{reactor-previous\,interval}}\right)\right)$$

$q_{methanol} = specific\ methanol\ uptake\ rate\ (mg\ g^{-1}\ h^{-1})$

$X_{predicted} = predicted\ total\ biomass$ in cell dry weight (g)

$t_{interval} = time\ interval\ (h)$

$R_{methanol} = methanol\ consumption\ at\ t_{interval}\ (mg)$

$T_{methanol} = total\ methanol\ (mg)$

$A_{methanol} = methanol\ addition\ (mg)$

$F_{methanol} = 50\%\ (v/v)\ methanol\ feed\ (mL)$

$C_{methanol\,target} = target\ methanol\ concentration\ (mg/mL)$

$V_{reactor} = reactor\ volumen\ (mL)$

$F_{glucose} = 50\%\ (v/v)\ \ glucose\ feed\ (mL)$

$V_{Sample} = volume\ of\ sample\ if\ applicable\ in\ the\ interval, else\ it\ is\ 0$

$Y_{\left(\frac{x}{s}\right)} = biomass\ yield\ on\ glucose\ (g/g)$

d) Because of the predicted specific methanol uptake rate based on example **7d)** it was possible to keep the methanol concentration at excess during the bioreactor cultivation from 1.19% to 1.5% (v/v) of methanol with only once per day at line methanol concentration measurements and feed adjustment.

e) The process and productivity data can be found in Table 10. The overall average specific productivity was 29.4 µg g^{-1} h^{-1}. The methanol concentration at the end of the cultivation was 10.4 and 10.0 g/L (1% (v/v) methanol corresponds to 7.92 g/L) for reactor R1 and R2. The total amount of consumed methanol in phase two by reactor R1 and R2 was 25.03 g and 24.07 g. This was calculated by the following equation:

$$T_{consumed\ methanol} = \left(\frac{(m_{start} - m_{end})}{\rho_{50\%\ methanol}} * 50\% * \rho_{methanol}\right) - (C_{methanol-end} * V_{reactor-end})$$

$T_{consumed\ methanol} = total\ consumed\ methanol\ (g)$

$m_{start} = 50\%\ methanol\ container\ weight\ at\ phase\ start\ (g)$

$m_{end} = 50\%\ methanol\ container\ weight\ at\ feedend\ (g)$

$\rho_{50\%\ methanol} = 50\%\ methanol\ density\ (g/ml)$

$\rho_{methanol} = 100\%\ methanol\ density\ (g/ml)$

$C_{methanol-end} = methanol\ concentration\ at\ feedend\ (g/L)$

$V_{reactor-end} = reactor\ volume\ at\ feedend\ (L)$

Table 10: Bioreactor cultivation process data and specific productivity (q_P) for Example 8. The methanol concentration was adjusted 2.22 hours after the sample was taken by an additional 50% (v/v) methanol pulse for R1 = 5.6 mL, R2 = 2.3 mL.

Time (h)	Volume (mL)		YDM (g/L)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) (µg g^{-1} h^{-1})		Methanol concentration (g/L)	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
22.30	318.2	318.0	25.2	25.0	0.0	0.0				
24.60									4.9*	9.0*
43.22	382.9	378.3	66.4	67.0	75.4	75.2	64.50	63.63	10.3	11.9
68.20	458.5	453.1	99.0	101.5	164.0	142.5	31.74	23.35	9.4	10.0
92.45	528.9	521.5	119.8	122.0	235.9	211.6	18.83	17.28	11.0	10.7
116.37	621.6	613.5	135.1	136.9	363.6	338.4	28.30	27.17	10.3	9.9
120.12	625.3	617.6	136.0	137.3	390.9	378.6	28.90	30.17	10.4	10.0

* Represents a control sample after the methanol pulse.

<u>Example 9:</u> Cultivation strategy 2 - A feed strategy with a separated glucose feed phase and a methanol only feed phase.

The *P. pastoris* ∆*aox1*∆*aox2* pPM2pN21_pAOX1_HSAopt_CycTT strain was tested in a recombinant protein production scenario where first a limited glucose feed was applied to increase the biomass followed by a separated phase with a methanol pulse and feed to induce protein production.

a) The bioreactor cultivation was separated into three phases. (1) Phase one consisted of the batch phase on BSM media with a start OD_{600} of 2. The inoculation was done as described in Example 7a) b). The batch phase lasted for 19.68 and 19.50 hours for reactor R3 and R4, respectively. (2) Phase two was a 50% (w/w) glucose feed at 4.8 mL/h for 25 hours to increase the biomass concentration. (3) Phase three was started with a 50% (v/v) methanol pulse to reach a target concentration of 1.5% (v/v) and a methanol feed profile calculated based on the predicted cell dry weigh and specific methanol uptake rate as described in Example 8c) for 72.7 hours. Methanol concentration was measured at line with HPLC every day as described in Example 6d) to measure the exact concentration and an additional compensation pules was added if necessary. In this phase the reactor stirrer speed was set to a constant 760 rpm and gassing to 9.5 sL/h.

b) The process and productivity data can be found in Table 11. The maximal and minimal methanol concentration throughout the cultivations ranged from 4.3 g/L to 12.55 g/L. The overall average specific productivity was 32.9 µg g^{-1} h^{-1}. The methanol concentration at the end of the cultivation was 7.10 and 7.47 g/L for reactor R3 and R4. The amount of consumed methanol in Phase three by reactor R3 and R4 was 12.0 g and 12.6 g. This was calculated as in Example 8e). Because the biomass was constant in phase three the methanol uptake rate ($q_{methanol}$) for phase three was calculated as in the following equation.

$$q_{methanol} = \frac{T_{consumed\ methanol}}{X_{biomass-average} * t_{phase\ 3}}$$

$$T_{consumed\ methanol} = total\ consumed\ methanol\ in\ phase\ 3\ (mg)$$

$$X_{biomass-average} = average\ biomass\ in\ phase\ 3\ (g)$$

$$t_{phase\ 3} = duration\ of\ phase\ 3\ (h)$$

In phase three the $q_{methanol}$ for reactor R3 is 3.79 mg g^{-1} h^{-1} and for reactor R4 it is 3.92 mg g^{-1} h^{-1}.

c) The total biomass in Table 11 was corrected for the sample withdraw of 12 mL and shows that the biomass was not increasing. Overall the total biomass in phase three decreased by 4.5% and 3.4% for reactor R3 and R4. Astonishingly, the culture was nonetheless producing secreted recombinant proteins. This shows that the *P. pastoris* ∆*aox1*∆*aox2* pPM2pN21_pAOX1_HSAopt_CycTT strain can efficiently produce recombinant secreted proteins even when fed only with methanol at no apparent growth. The total average amount of proteins produced in phase three with methanol as the only carbon source is 105 mg.

$$T_{biomass} = V_{reactor} * CDW + \sum \left(CDW_{previous} * V_{sample} \right)$$

$$T_{biomass} = total\ corrected\ biomass$$

$$CDW = cell\ dry\ weight$$

$$V_{sample} = volume\ of\ sample$$

Table 11: Bioreactor cultivation process data and productivity (q_P) for Example 9. The total biomass was corrected for 12 mL sampling as in Example 9 b).

	Time (h)	Volume (mL)		YDM (g/L)		Total Biomass (g)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) (µg g^{-1} h^{-1})		Methanol concentration (g/L)	
Phase		R3	R4	R3	R4	R3	R4	R3	R4	R3	R4	R3	R4
1	20.13	317.9	318.5	25.5	25.5	8.1	8.1						
2	29.92	366.5	368.0	67.2	67.0	24.9	24.9	0.0	0.0	0.0	0.0		
	45.02	445.8	447.8	108.0	107.7	49.2	49.3	0.0	0.0	0.0	0.0		
3	47.08											12.1*	12.5*
	53.00	444.2	446.1	104.4	103.7	48.8	48.6	46.1	45.7	35.56	35.6		
	69.58	432.2	435.7	103.3	102.9	48.3	48.5	106.7	107.8	22.51	23.3	4.3	5.9
	93.93	434.7	437.0	98.5	99.1	47.7	48.2	244.0	251.0	38.67	39.9	6.6	7.4
	118.12	434.4	435.9	94.2	95.3	47.0	47.6	350.9	350.2	33.07	30.0	7.1	7.5

* Represents a control sample after the methanol pulse.

<u>Example 10</u>: Cultivation strategy 3 - A feed strategy with a glucose/methanol co-feed phase and a separated methanol only feed phase.

The *P. pastoris* ∆*aox1*∆*aox2* pPM2pN21_pAOX1_HSAopt_CycTT strain was tested in a recombinant protein production scenario where after the batch phase a limited glucose feed and an additional methanol pulse and feed was applied to achieve a biomass increase and recombinant protein production simultaneously. After the desired biomass was reached the glucose feed was stopped but the methanol feed continued for the rest of the cultivation.

a) This bioreactor cultivation was separated into three phases. (1) Phase one was the batch phase. For this the reactors were inoculated with the production strain *P. pastoris* ∆*aox1*∆*aox2* pPM2pN21_pAOX1_HSAopt_CycTT with a start OD_{600} of 2. The inoculation was done as described in Example 7a) b). The batch phase lasted for 19.35 and 19.37 hours for reactor R1 and R2, respectively. The end of the batch phase was indicated by a dissolved oxygen peak. (2) At this point Phase two was started. Phase two consisted of a 50% (w/w) glucose feed at 4.8 mL/h for 25 hours. At the start of Phase two a 50% (v/v) methanol pulse was applied to increase methanol concentration to the target of 1.5% (v/v) and a subsequent methanol feed was started to counteract methanol consumption, evaporation and dilution by the glucose feed. (3) Phase three consisted of a methanol only feed for 72.9 hours. In this phase the reactor stirrer speed was set to a constant 760 rpm and gassing to 9.5 sL/h. Methanol concentration was measured at line with HPLC every day as described in Example 6d). An additional compensation pulse was added if necessary.

b) The methanol feed was calculated in hourly intervals as in Example 9b):

c) The process and productivity data can be found in Table 12. The maximal and minimal methanol concentration throughout the cultivation of the two repeats ranged from 6.9 g/L to 11.4 g/L. The overall average specific productivity was 45.8 µg g^{-1} h^{-1}, the average specific productivity in phase three was 34.0 µg g^{-1} h^{-1}. The methanol concentration at the end of the cultivation was 8.0 g/L for reactor R1 and R2. The amount of consumed methanol in phase three by reactor R1 and R2 was 14.4 g and 14.1 g. This was calculated by the equation as shown in Example 9b). Because the biomass was constant in phase three the methanol uptake rate was calculated ($q_{methanol}$) for phase three as shown in Example 9b). In phase three the $q_{methanol}$ for reactor R1 was 4.61 mg g^{-1} h^{-1} and for reactor R2 it was 4.54 mg g^{-1} h^{-1}. Overall the biomass decreased by 5.7% and 5.5% for reactor R1 and R2 in phase three. Again, this shows that with the *P. pastoris* ∆*aox1*∆*aox2* strain production at no apparent growth is possible. The average total amount of recombinant protein produced in phase three with methanol as the only carbon source is 106 mg which is similar to the 105 mg of recombinant protein produced in Example 9, phase three. This is also illustrated by the similar specific productivity in phase three from Example 9 at 32.9 µg g^{-1} h^{-1} and at 34.0 µg g^{-1} h^{-1} in this example. In conclusion the productivity in phase three (methanol

only feed phase) did not depend whether the cultures were induced in phase two (glucose feed phase) or not. Because recombinant protein production is independent of growth the methanol only feed strategy has several advantages when used with the *P. pastoris* ∆*aox1*∆*aox2* strain. In a bioreactor cultivation without a methanol only feed phase as in Example 8 the process is constrained by the maximal reactor volume and the yeast dry mass concentration. After a certain time this constraints stop the cultivation process either by reaching the maximal volume or maximal desired biomass concentration. By using a methanol feed phase in combination with the *P. pastoris* ∆*aox1*∆*aox2* as in Example 9 the cultivation time is no longer limited by the biomass concentration or volume as the biomass is not increasing and the volume increase is negligible. As a consequence cultures at high biomass concentrations can be kept in the bioreactor for longer periods of time without reaching these constraints and allow for longer production phases that increase the concentration of the protein of interest. A methanol only feed phase is also applied when using the methanol utilization slow *P. pastoris* ∆*aox1* strain as shown in the following Example 12 but these advantages are not present there because *P. pastoris* ∆*aox1* is continuously growing on a methanol only feed and therefore exhibits the same constraints as discussed. Restricting the *P. pastoris* ∆*aox1* to the same methanol feed rate as the *P. pastoris* ∆*aox1*∆*aox2* results in productivity loss. Further process related improvements of the *P. pastoris* ∆*aox1*∆*aox2* strain are discussed in Example 12.

Table 12: Bioreactor cultivation process data and specific productivity (q_P) for Example 10. The total biomass was corrected for 12 mL sampling as in Example 9 b).

	Time (h)	Volume (mL)		YDM (g/L)		Total Biomass (g)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) ($\mu g\ g^{-1}\ h^{-1}$)		Methanol concentration (g/L)	
Phase		R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1	20.13	318.0	317.7	25.0	24.9	7.9	7.9	0.0	0.0	0.0	0.0		
2	21.87											11.0*	11.4*
	29.92	374.5	374.2	65.3	64.6	24.7	24.5	47.7	51.9	88.6	97.3		
	45.02	464.2	461.8	104.0	103.1	49.4	48.7	203.0	243.4	87.1	108.6	7.2	6.9
3	53.00	462.7	460.9	99.7	99.5	48.4	48.2	246.1	265.5	38.6	21.5		
3	69.58	457.4	455.6	95.9	96.3	47.4	47.4	357.3	371.8	47.3	44.9	7.9	7.5
	93.93	459.8	458.3	91.9	91.2	46.9	46.5	459.5	515.0	33.9	47.4	8.2	7.1
	118.12	462.0	460.8	88.3	87.3	46.6	46.0	494.8	568.2	14.8	21.6	8.0	8.0

* Represents a control sample after the methanol pulse.

<u>Example 11</u>: Cultivation strategy 3 - A feed strategy with a glucose/methanol co-feed phase and a separated methanol only feed phase applied to *P. pastoris* ∆*aox1*∆*aox2* secreting VHH.

To check secreted recombinant protein production with another secreted protein the bioreactor cultivation described in Example 10a) was repeated with the strain *P. pastoris* ∆*aox1*∆*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT.

a) As in Example 10a), this bioreactor cultivation was separated into three phases. (1) Phase one was the batch phase. For this the reactors were inoculated with the production strain *P. pastoris* ∆*aox1*∆*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT with a start OD_{600} of 2 as described in Example 7a) b). The batch phase lasted for 18.79 and 19.33 hours for reactor R1 and R2, respectively. The end of the batch phase was indicated by a dissolved oxygen peak. (2) At this point Phase two was started. Phase two consisted of a 50% (w/w) glucose feed at 4.8 mL/h for 33.9 hours to increase the biomass even higher than in Example **10**. At the start of Phase two a 50% (v/v) methanol pulse was added to increase the methanol concentration to the target value of 1.5% (v/v) and a subsequent methanol feed was started to counteract methanol consumption, evaporation and dilution by the glucose feed. (3) Phase three consisted of a methanol only feed for 63.9 hours. The stirrer speed was set to a constant 760 rpm and gassing to

9.5 sL/h. Methanol concentration was measured at line with HPLC every day as described in Example 6d). An additional compensation pulse was added if necessary.

b) The methanol feed was calculated as described in Example 9b)

The process and productivity data can be found in Table 13. The maximal and minimal methanol concentration throughout the cultivation of the two repeats ranged from 8.7 g/L (R1) to 11.3 g/L (R1). The overall average specific productivity was 118.0 µg g^{-1} h^{-1} and in phase three 88.2 µg g^{-1} h^{-1}. The methanol concentration at the end of the cultivation was 10.5 and 10.8 g/L for reactor R1 and R2. The amount of consumed methanol by reactor R1 and R2 was 26.6 g and 26.0 g. This was calculated by the following equation as shown in Example 8e). This amount is higher as in Example 10 due to the higher biomass concentration, but still significantly (5-times) lower than in a Mut^S strain (as described in Example 12). Because the biomass was constant in phase three the methanol uptake rate ($q_{methanol}$) was calculated for phase three as shown in Example 9b). In phase three the $q_{methanol}$ for reactor R1 was 4.75 mg g^{-1} h^{-1} and for reactor R2 it was 4.68 mg g^{-1} h^{-1}.

c) Overall the biomass decreased by 4.4 % and 4.9 % for reactor R1 and R2 in phase three as in previous Examples. The data in Table 13 clearly show that the *P. pastoris* Δ*aox1*Δ*aox2* strains can produce secreted recombinant proteins even in gram per liter amounts. In the methanol only feed phase the vHH concentration increased by 815.5 mg/L, meaning that an average total of 323.1 mg of antibody fragment was produced using methanol as the only carbon source.

Table 13: Bioreactor cultivation process data and productivity (qP) for Example 11. The total biomass was corrected for 12 mL sampling as in Example 9 b).

	Time (h)	Volume (mL)		YDM (g/L)		Total biomass (g)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) (µg g^{-1} h^{-1})		Methanol concentration (g/L)	
Phase		R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1	20.00	316.7	316.7	25.1	25.2	7.9	8.0	0.0	0.0	0.0	0.0		
2	21.42											10.9*	10.9*
	28.22	358.4	358.7	60.2	60.1	21.9	21.9	58.2	59.7	137.8	141.5		
	44.83	449.9	449.8	105.8	105.2	48.6	48.3	494.4	464.5	222.8	208.3	7.8	7.9
	53.58	499.8	499.4	124.1	122.9	64.3	63.7	776.8	854.1	176.0	246.2		
3	68.83	502.5	503.3	115.5	116.3	61.8	62.3	1156.0	1030.0	142.5	72.2	8.7	9.5
	92.00	511.3	511.1	110.8	109.4	61.8	61.1	1337.2	1374.0	55.8	97.5	11.3	9.8
	117.92	525.0	527.3	104.7	102.7	61.5	60.6	1623.2	1638.6	85.7	85.8	10.5	10.8

* Represents a control sample after the methanol pulse.

<u>Example 11.1</u>: Process parameters obtained with *P. pastoris* ∆*aox1*∆*aox2* (Mut$^-$) strains compared to a methanol utilization slow *P. pastoris* ∆*aox1* (Mut^S) cultivated with an established bioreactor cultivation protocol.

For comparison the *P. pastoris* ∆*aox1* pPM2pN21_pAOX1_HSAopt_CycTT was cultivated with an established cultivation protocol for the Mut^S phenotype (Potvin, Ahmad, & Zhang, 2012).

a) This bioreactor cultivation was separated into four phases. (1) Phase one was the batch phase. For this the reactors were inoculated with the production strain *P. pastoris* ∆*aox1* pPM2pN21_pAOX1_HSAopt_CycTT with a start OD_{600} of 2 as described in Example 7a) b). The batch phase lasted for 20.17 and 20.30 hours for reactors R1 and R2, respectively. The end of the batch phase was indicated by a dissolved oxygen peak. (2) Phase two is a linearly increasing (y = 0.225x + 1.95) 60% glycerol feed for 8 hours. (3) Phase three was a co-feed phase with a linearly decreasing (y = 3.75 – 0.111x) 60% glycerol feed and a linearly increasing (y = 0.028x + 0.6) 100% methanol feed for 18 hours (4) Phase four is a methanol only feed phase with a linearly increasing 100% methanol feed (y = 0.028x + 1.10) for 72 hours. The total run time was 119.25 hours.

b) The glycerol and methanol feed was gravimetrically controlled based on the equations in a) by the DASGIP control software (Eppendorf AG, Germany)

c) The process and productivity data can be found in Table 13.1. The overall average specific productivity from phase three to four (the production phases) was 61.7 µg g^{-1} h^{-1}. The average total amount of consumed methanol was 165.8 g over the whole cultivation period and 150.6 g in phase four (methanol only feed phase). The residual methanol concentration in the culture broth was considered

to be zero as this was a methanol limited cultivation. Phase four in this example corresponds to phase three in Examples 9, 10 and 11. Based on the average biomass in phase four the methanol uptake rate ($q_{methanol}$) as shown in Example 9b) was calculated. The $q_{methanol}$ in phase four for reactor R1 is 37.1 mg g^{-1} h^{-1} and 37.6 mg g^{-1} h^{-1} for reactor R2. The $q_{methanol}$ in phase four facilitates an unwanted biomass increase. 53.2% (R1) and 52.4% (R2) of the total biomass at cultivation finish are generated during phase four growth on methanol.

d) The comparison of strain related process parameters are depicted in table Table 13.2 and an overview of the specific methanol uptake rates and feed rates from the presented examples can be found in Table 13.3. By using the *P. pastoris* Δ*aox1*Δ*aox2* Mut$^-$ for recombinant protein production several of the key processes parameters improved considerably compared to a process with the *P. pastoris* Δ*aox1*. The heat of reaction is reduced substantially by more than 80% leading to a reduced need for cooling. The specific oxygen uptake rate and oxygen transfer rate is reduced by more than 80% leading to a reduced need for mixing and aeration, reducing the flow rate of aeration as well as the need to supply pure oxygen to the bioreactor vessels. The lower specific methanol uptake rate reduces the amount of methanol needed in a cultivation. Methanol is toxic and flammable.

Use of the Mut$^-$ strains represents a technical and safety improvement as it reduces the quantities of methanol that need to be handled and stored in a production facility. Another advantage is the lower sensitivity of the *P. pastoris* Δ*aox1*Δ*aox2* to high methanol concentrations. This is confirmed by cell viability data of the strain *P. pastoris* Δ*aox1*Δ*aox2* in Example 10 and *P. pastoris* Δ*aox1* in this example. In Example 10 the viability of the Mut$^-$ cells in reactors R1 and R2 at the end of the process was 99.8% and 99.7%. In contrast the cell viability of the Mut^S cells in reactors R1 and R2 in this example was 95.9% and 96.5%. The lower sensitivity and higher viability of the *P. pastoris* Δ*aox1*Δ*aox2* strain has an effect on the purity of the recombinant produced secreted protein. Lysed cells release proteases that degrade the protein of interest and add unwanted soluble protein in the supernatant, both effects lead to lower purity and loss of the protein of interest in the supernatant. The purity of the *P. pastoris* Δ*aox1* in this example was 72% and 77% for reactors R1 and R2, in comparison the purity of the *P. pastoris* Δ*aox1*Δ*aox2* in Example 10 was 85% for both reactors R1 and R2.

Table 13.1: Bioreactor cultivation process data and specific productivity ($q_{product}$) for Example 11.1.

	Time (h)	Volume (mL)		YDM (g/L)		Total biomass (g)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) (µg g^{-1} h^{-1})	
Phase		R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1	20.3	318.1	316.5	26.6	25.9	8.5	8.2				
2	28.5	330.4	328.6	51.6	51.4	17.0	16.9	0.0	0.0	0.00	0.0
3	47.1	392.5	390.3	99.6	100.4	39.1	39.2	129.6	112.9	65.57	56.6
4	69.8	428.4	426.5	111.5	111.4	47.7	47.5	260.9	306.7	36.89	53.9
	92.9	487.8	487.3	126.9	124.6	61.9	60.7	548.1	534.7	67.14	56.7
	119.3	581.9	584.9	143.7	140.8	83.6	82.3	893.9	927.9	61.55	72.7

Table 13.2: Comparison of key bioreactor cultivation parameters overall and on the methanol only feed phases of *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT in Example 10 and *P. pastoris* Δ*aox1* pPM2pN21_pAOX1_HSAopt_CycTT in Example 11.1.

		P. pastoris Δ*aox1*Δ*aox2* pPM2pN21_ pAOX1_HSAopt_CycTT Example 10	*P. pastoris* Δ*aox1* pPM2pN21_ pAOX1_HSAopt_CycTT Example 11.1	Change
Overall	HSA concentration (mg/L)	531	911	-42%
	$q_{product}$ (µg g^{-1} h^{-1})	46	62	-26%
Methanol only feed phase: Example 10 - Phase three; Example 11.1 - Phase four	Duration (h)	73.1	72.2	
	Heat production rate (W/L)	3	19.2	-84%
	OTR (mM/h)	23	149	-85%
	Heat of reaction (kJ) Integrated OTR (MA 120 min)	359	2286	-84%
	Heat of combustion (kJ)	324	2468	-87%
	qO_2 (mmol g^{-1} h^{-1})			-81%
	$q_{methanol}$ (mg g^{-1} h^{-1})	4.6	37.4	-88%
	Protein (mg)	106	250	-58%
	Protein/methanol (mg g^{-1})	7	1.7	+311%
	Protein/0_2 (mg mol^{-1})	135	51	+164%
	Protein/Heat of reaction (mg kJ^{-1})	0.30	0.11	+172%

Table 13.3: Specific methanol uptake rates and methanol feed rates based on average cell dry weight in the methanol only feed phase. *As Example 11.1 has a limited methanol feed the $q_{methanol}$ and feed rate are considered equal.

	Example 7		Example 9		Example 10		Example 11		Example 11.1	
Reactor	R1	R2	R3	R4	R1	R2	R1	R2	R1	R2
$q_{methanol}$ (mg g^{-1} h^{-1})	4.92	5.22	3.79	3.92	4.61	4.54	4.75	4.68	37.1*	37.1*
Feed rate (mg g^{-1} h^{-1})	NA	NA	4.79	4.89	5.80	5.73	5.78	7.71	37.1	37.6

Example 12: Generation of methanol utilization negative and alcohol dehydrogenase defective strains.

The methanol consumption of the *P. pastoris* Mut$^-$ strain observed in Example **7** was unexpected and new. Based on this knowledge the hypothesis was formed that alcohol dehydrogenases might be responsible for this characteristic. To test the effect of alcohol dehydrogenases on methanol consumption in the *P. pastoris* Δ*aox1*Δ*aox2* two potential alcohol dehydrogenases were selected *ADH2:* PP7435_Chr2-0821 and *ADH900*: PP7435_Chr2-0990 for deletion. Three strains were created, (1) an *ADH2* defective strain, (2) an *ADH900* defective strain and (3) a double deletion *ADH2* & *ADH900* strain by exchanging the coding region of the gene with an antibiotic resistance. Effectively the strains (1) *P. pastoris* Δ*aox1*Δ*aox2 adh2*Δ::HphR, (2) *P. pastoris* Δ*aox1*Δ*aox2 Adh900*Δ::KanMX and (3) *P. pastoris* Δ*aox1*Δ*aox2 adh2*Δ::HphR a*dh900*Δ::KanMX were created.

a) *P. pastoris* Δ*aox1*Δ*aox2* was made electrocompetent as described in Example **1a)**. For generating the *ADH2* deletions the spilt marker approach was used described in example **1b)**. The electrocompetent cells were transformed with 500 ng of Adh2 split marker cassette 1 and 500 ng Adh2 split marker cassette 2 as described in Example **1d)**. The cassette sequences can be found in Table 14**.** The transformants were selected on YPD plates with 200 µg/mL Hygromycin. One clone was selected based on PCR amplification and sequencing of the PCR amplicon. The successful substitution of the *ADH2* coding region with the antibiotic marker was verified by PCR amplification with the primers Adh2_KO_ctrl_fwd & Adh2_KO_ctrl_rev (Table 15) and sequencing of the PCR amplicon (Microsynth AG, Swiss). The generated strain is called (1) *P. pastoris* Δ*aox1*Δ*aox2 adh2*Δ::HphR.

b) The *P. pastoris* Δ*aox1*Δ*aox2* strain was made electrocompetent as described in Example **1a)**. The electrocompetent cells were transformed with 500 ng of Adh900 split marker cassette 1 and 500 ng Adh900 split marker cassette 2 as described in Example **1d)**. The cassette sequences can be found in Table 14. The transformants were selected on YPD plates with 500 µg/mL Geneticin. One clone was selected based on PCR amplification and sequencing of the PCR amplicon. The successful substitution of the *ADH900* coding region with antibiotic marker was verified by PCR amplification with the primers AdhII_KO_Ctrl_fwd & AdhII_KO_Ctrl_rev (Table 15) and sequencing of the PCR amplicon (Microsynth AG, Swiss). The generated strain is called (2) *P. pastoris* Δ*aox1*Δ*aox2* Adh900Δ::KanMX.

c) The *P. pastoris* Δ*aox1*Δ*aox2 adh2*Δ::HphR strain was made electrocompetent as described in Example **1a)** apart from that 200 µg/mL Hygromycin were added to the main culture medium. The electrocompetent cells were transformed with 500 ng of Adh900 split marker cassette 1 and 500 ng Adh900 split marker cassette 2 as described in Example **1d)**. The cassette sequences can be found in Table 14. The transformants were selected on YPD plates with 200 µg/mL Hygromycin and 500 µg/mL Geneticin. One clone was selected based on PCR amplification and sequencing of the PCR amplicon. The successful substitution of the *ADH900* coding region with the antibiotic marker was verified by PCR amplification with the primers AdhII_KO_Ctrl_fwd & AdhII_KO_Ctrl_rev (Table 15) and sequencing of the PCR amplicon (Microsynth AG, Swiss). The generated strain is called (3) *P. pastoris* Δ*aox1*Δ*aox2* Adh2Δ::HphR Adh900Δ::KanMX.

d) Genomic DNA for PCR amplifications was isolated with the Wizard® Genomic DNA Purification Kit (Promega Corporation, USA) as per manufacturer's recommendations. The PCR amplification reactions were done with the Q5 polymerase (New England Biolabs, Inc., USA) as per manufacturer's recommendations.

Table 14: Split marker cassette DNA sequence used for generating the Adh2 and Adh900 deletion strains.

DNA fragment	DNA sequence 5′ to 3′
Adh2 split marker cassette 1	(SEQ ID NO:72) CGTATCTACCGATGATGGCACCAGCCTCCATCTGTTCGTAGACCTTAGCAAGTTCAGACA GACCGATAATCTTGATAGGAGCCTTGACCAAACCTCTGGTGAACAAGTCGATGGCCTCGG CACTGTCCTCTCTGTTTCCAACGTAAGATCCCTTGATCTCGATGGACTTCAGAACGTGCC AGAAAACGTCAGAGTTGACAACGGCACCAGATGGCAGACCAACCAAAACAACCTTACCCA AAGTTCTAACGTATTGGACAGATTGGTTGATAGCATGTGGGGAAACGGAGACGTTAATAA CACCGTGTGGACCACCGTTGGTGAGCTTTTGGACTTCAGCAACGACGTCCTTAGTCTTAG TGAAGTCGACGAAGACCTCAGCACCCAAGGACTTGACAAATTCACCCTTGTCGGCACCAC CATCAATACCCAAAACTCTCAAACCCAGAGCCTTGGCGTATTGAACGGCAAGAGAACCCA GTCCTCCACCAGCACCAGAAATGGCAACCCATTGGCCAATACGCAAGTCAGCGGTCTTAA GAGCCTTGTAAACGGTGATACCAGCACACAGAATTGGGGCAACTTCAGCCAAGTCAGCCT CCTTTGGAATTCTGGCGGCTTGGGTGGCATCAGCAGTAGCATACTGCTGGAAAGATCCGT CGTGGGTGAAACCAGACAGGTCAGCCTTGGCACAACTGGATTCAGCACCTTGGATACAGT ACTCACAGTTCAAACAAGAACCGTTCAACCATTTGATACCAGCGTAGTCACCGATAGTGG ATCTGATATCACCTAATAACTTCGTATAGCATACATTATACGAAGTTATATTAAGGGTTC TCGAATGGTACCTTGCTCACATGTTGATCTCCAGCTTGCAAATTAAAGCCTTCGAGCGTC CCAAAACCTTCTCAAGCAAGGTTTTCAGTATAATGTTACATGCGTACACGCGTCTGTACA GAAAAAAAAGAAAAATTTGAAATATAAATAACGTTCTTAATACTAACATAACTATAAAAA AATAAATAGGGACCTAGACTTCAGGTTGTCTAACTCCTTCCTTTTCGGTTAGAGCGGATG TGGGGGGAGGGCGTGAATGTAAGCGTGACATAACTAATTACATGATATCGACAAAGGAAA AGGGGGACGGATCTCCGAGGTAAAATAGAACAACTACAATATAAAAAAACTATACAAATG ACAAGTTCTTGAAAACAAGAATCTTTTTATTGTCAGTACTGATTATTCCTTTGCCCTCGG ACGAGTGCTGGGGCGTCGGTTTCCACTATCGGCGAGTACTTCTACACAGCCATCGGTCCA GACGGCCGCGCTTCTGCGGGCGATTTGTGTACGCCCGACAGTCCCGGCTCCGGATCGGAC GATTGCGTCGCATCGACCCTGCGCCCAAGCTGCATCATCGAAATTGCCGTCAACCAAGCT CTGATAGAGTTGGTCAAGACCAATGCGGAGCATATACGCCCGGAGCCGCGGCGATCCTGC AAGCTCCGGATGCCTCCGCTCGAAGTAGCGCGTCTGCTGCTCCATACAAGCCAACCACGG CCTCCAGAAGAAGATGTTGGCGACCTCGTATTGGGAATCCCCGAACATCGCCTCGCTCCA GTCAATGACCGCTGTTATGCGGCCATTGTCCGTCAGGACATTGTTGGAGCCGAAATCCGC GTGCACGAGGTGCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAGAGC CTGCGCGACGGACGCACTGACGGTGTCGTCCATCACAGTTTGCCAGTGATACACATGGGG ATCAGCAATCGCGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCCTTGCGGTCC GAATGGGCCGAACCCGCTCGTCTGGCTAAGATCGGCCGCAGCGATCGCATCCATGGCCTC CGCGACCGGCTGCAGAACAGCGGGCAGTTCGGTTTCAGGCAGGTCT
Adh2 split marker cassette 2	(SEQ ID NO:73) AGATGTTGGCGACCTCGTATTGGGAATCCCCGAACATCGCCTCGCTCCAGTCAATGACCG CTGTTATGCGGCCATTGTCCGTCAGGACATTGTTGGAGCCGAAATCCGCGTGCACGAGGT GCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAGAGCCTGCGCGACGG ACGCACTGACGGTGTCGTCCATCACAGTTTGCCAGTGATACACATGGGGATCAGCAATCG CGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCCTTGCGGTCCGAATGGGCCGA ACCCGCTCGTCTGGCTAAGATCGGCCGCAGCGATCGCATCCATGGCCTCCGCGACCGGCT GCAGAACAGCGGGCAGTTCGGTTTCAGGCAGGTCTTGCAACGTGACACCCTGTGCACGGC GGGAGATGCAATAGGTCAGGCTCTCGCTGAATTCCCCAATGTCAAGCACTTCCGGAATCG GGAGCGCGGCCGATGCAAAGTGCCGATAAACATAACGATCTTTGTAGAAACCATCGGCGC AGCTATTTACCCGCAGGACATATCCACGCCCTCCTACATCGAAGCTGAAAGCACGAGATT CTTCGCCCTCCGAGAGCTGCATCAGGTCGGAGACGCTGTCGAACTTTTCGATCAGAAACT TCTCGACAGACGTCGCGGTGAGTTCAGGCTTTTTACCCATGGTTTAGTTCCTCACCTTGT CGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAACACCGCCCTTAGATT AGATTGCTATGCTTTCTTTCTAATGAGCAAGAAGTAAAAAAAGTTGTAATAGAACAAGAA AAATGAAACTGAAACTTGAGAAATTGAAGACCGTTTATTAACTTAAATATCAATGGGAGG TCATCGAAAGAGAAAAAAATCAAAAAAAAAAAATTTTCAAGAAAAAGAAACGTGATAAAA ATTTTTATTGCCTTTTTAGACGAAGAAAAAGAAACGAGGCGGTCTCTTTTTTCTTTTCCA AACCTTTAGTACGGGTAATTAACGACACCCTAGAGGAAGAAAGAGGGGAAATTTAGTATG CTGTGCTTGGGGGTTTTGNAAATGGTACGGCGATGCGCGGAATCCGAGAAAATCTGGAAG AGTAAAAAAGGAGTAGAAACATTTTGAAGCTATGGTGTGTGGTACCGATCTAGACCTAAT AACTTCGTATAGCATACATTATACGAAGTTATATTAAGGGTTGTCGACCTGCAGCGTACG

	GCACGAATTCGCACCCCGGAGAGCGCTCACCCCCGTTTTCAAACAGCGGGGGGAGCACAA AATGTTGAAAACTACACAGATCTTTTCGGACACCGGTCGCTTTATGTAGTCGACATGCAG ATTCTCCCAAATGGAAAACGAGATTGGACAATTTGTGGAGTTGGAAAGGGGGGTGGGAAT CAACGAAATTAGCAGATTCATGGGCAATTGGCAGGACTGGGCAGAAGGGGTGAGAATTGC AATCGAATGGAACAGGCACTCCCGTTGCGAAATCAAAAAAGTCTCGCTATCTGAACTGAT TTTTTTTAAGCAGCAACTTACGGTCAATACATCTCCGATGGAGGAATTTTTCACCCCTCG CTAACTAGATGGGCCCCTTCTAAGAAATTTGGGTTTAAGGTTGGGCAGTCAGTCAGTGCA CCAATGCTAACTGCCATTTGTCCAAAGAGGGGTGCAAGGATGAGGGACCGTTGAGAATAA GATTTGGGGTGTTAATCGGTGATACTGATTTGTCAAAGAGTGGGGAGGACTGCTGGGCAT TGTTCACCCCCCTAGTTGTTAGAGTTCGATAGCCGGCCGAATCACCCCCCTCTTCTTACA TAATCATTGTCACTATGTGGGGTCTCTACAGTCTCACCCTGCGATCCGGGACGACGCCGC GAAATTAGGGGGCAAGTCTCCTCCGGGCATGCAATATTGGTAACAGGATCAATTGATGCG AGAAAAGTTGGAGGGGGTGTAAAATTCAAGCCCACAAAGTCACACCCTTATGCCTGTAGA GGGGCAATCGGAGAGCAGCCATGGGGTGT
Adh900 split marker cassette 1	(SEQ ID NO:74) CACTCCAGTTGGGCCATTACCGAACATTTTGCCATTGTAGGCGATTAGTAAGTATTAACA AGACAGCTGACTATACGTTTATTCTCAAACAATATTTCCCTTTTTGGTTTTGACCTCGCT TTAATCAATTTTTCAGACCTGATCCCACCTACTTTTCTTCGGCCTCAACTTCAATCTGAC TCTTCTCTCTCAATTGGTACCAACCAGCCAGAAAATGTCCTTCCGTTACTTGAAACGGCA TTTCTCTACAGCTACAAACGCAATTGCTCTCCTTAGCAGACCTGAATTCAAAATAGGTCG AATTGTGGACGTCGTGAAACATCCAAATGCAGACAAACTTTATGTCTCGTCGATTTCTGT GGGAAACAATTATGCCTCGGGTACATCCAACACCCTAACCGTTTGCAGCGGCTTGGTGGA CTACTTTTCAGTTCCCGAATTGCTTCAGCGACGGGTCGTTGTGGTCACAAACCTCAAGCC ATCGAAGATGAGAGGTGTAACATCGGAGGCAATGCTTTTGGCAGGGGAAAAGTCGGGGAA AGTGGAATTGGTCGAGCCGCCAATGTCCGGGAGAGAGGGCGAATCACTCCACTTCGAAGG TGTAGAAATTACATCAGAGGAGAGCGCCAATCAATTGCATTTGCCTGCTAAGCGATTGAA GAAGTCAGAGTGGAGTCAACTGGCGGAAGGTCTACAGACAAATGACCAGCGTGAAGTGGT CTTCCACAGCCAAATTGGCTCCAAACGAATTTACGCTTTAGTAGGAGCGAGTACTGAAAA ATGCACGTTAGCGACTCTTGCGCAGGCCGTCGTACGATAAGGGCAATATGGTTGAGAACG TTCCTCACCCAAATAAAATCATCGTACGCTGCAGGTCGACAACCCTTAATATAACTTCGT ATAATGTATGCTATACGAAGTTATTAGGTCTAGATCGGTACCGACATGGAGGCCCAGAAT ACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTAC ATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACG GCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTC ACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGAT TTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCT CACATCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGC GATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCG GGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTC TGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACT GGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATG CATGGTTACTCACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGA TTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAAT CACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGC CTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCG
Adh900 split marker cassette 2	(SEQ ID NO:75) AAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTT ACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAG CATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAACA GCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCA GTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGC GTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGAT TTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTT TTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATT TTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGA TACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAA CGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTG ATGCTCGATGAGTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACT TGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGA TTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAG

	TAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGGTACCATTCGAGAACC CTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATTAGGTGATATCAGATCCACT CTGTAGTGAGGGTTGGTGGTCTGACGAACATCCAGCAAGGTGTTCCACCTGAAATTTTTC ACCTTGGAGGGTAATGTGATGACGCCATTTCCTGTGCAAATGCTTTTCGTTTTGAACAGT GCAACTTTTGTATCAGATCTTCATCTACTTGATGCCATCTCAACAAATCCCTCATTTACT AGCGTGTGAAGGAATCTAGATTTTCCACTGATAAGCCAATTTGTCGGAAATCCCCCGCGC GGGAGTTGGCGTTCAGTACGAGCCACACACGTTTCTTTTGGACAACCAAAGCATCCGCCT GAAGGGACAACTTGCATTCAACGGCTTCAGTTGGAAACGTCAGAGCTGACCTATAGTTTG CTAGAACCGTTTTCTCTGTTTACGTTTACGTCTCCTCAAATTTGCGCTCGGTATGTCCTT CCTAATTAGCGGGAAAAGCTGTTCTTAGTTAATACGGAGAAAGTTTCGGGGTTACCGTTC CGGGAAGAGGAGGGGTCATCTCTCTCATCTCATCCAACCATTAAGTTTCTTCCAAAACTT CAGGATAATCAGTTTAACCACCGACAGGAGTCAGATTTGAGATTGACAGAAAGTTTTTCC GTCCATTTCCTCATCTTGTCGCCGTTATCAGTCAATCTCTATGGTTATCTGGAATTTCTT TTTTCTTTTAATTCATCTTCTTTTTATCCCGCGCCTTTGGCGTTCTAGCTCATCTCATGA AAACAAAACCCTCTCATGTTCGGATAATTCCAGCGGCTTTCACTTTCAGATGACACATAG ATTGGACTCAACCATGGCTATCTGGGGTATACGGACGTTGGCAAGGGCGTTAATTTTTCA GGACAAACGGAAATGCCATGGCTCCAGGGAAAGGCATTCCTATTGCAAACCTAGACCGTC GAACCTCTCCTATCGCCTACCAGTCACCCAGCTATCCCTAGGCAACTCATCTCCTTCAAG CGGATTGCAACCTGCTAAGCCAAATTAGATCTGGCCACAGAAATGCCGCAATATTTCTTG GCTCTCCCCTCCC

Table 15: Polymerase chain reaction primers.

Primer Name	DNA sequence 5′ to 3′
Adh2_KO_ctrl_fwd	GAATTGAGCCAAAAAAGGAGAGG (SEQ ID NO:76)
Adh2_KO_ctrl_rev	GATGGAATAGGAGACTAGGTGTG (SEQ ID NO:77)
AdhII_KO_Ctrl_fwd	TGGTTGAGACGTTTGTATTG (SEQ ID NO:78)
AdhII_KO_Ctrl_rev	TGGGTTGGGAGTTTAGTG (SEQ ID NO:79)

<u>Example 13</u>: Generation of Adh2 and Adh900 overexpression methanol utilization negative strains.

For the purpose of checking the effect of *ADH2* and *ADH900* overexpression. An expression construct was created being composed of a constitutive promoter P_{GAP}: PP7425_Chr1 (596296…596790) and the *ADH2* coding sequence or the *ADH900* coding sequence, respectively. The *ADH2* and *ADH900* coding sequence (Table 16) were modified to eliminate BbsI and BsaI restriction sites in the coding sequence without affecting the amino acid sequence of the gene product. The generated strains were designated *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pGAP_Adh2_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pGAP_Adh900_CycTT.

a) For the purpose of using the Golden Gate assembly method the restriction sites of restriction enzymes BbsI and BsaI (New England Biolabs, Inc., USA) needed to be removed from the coding sequence without affecting the amino acid sequence of the gene product. This process is called Curing. The coding sequence of *ADH2* PP7435_Chr2-0821 was modified at c.45G>A and c.660C>G. The coding sequence of *ADH900* PP7435_Chr2-0990 was modified at c.42C>G. The cured coding sequence used for Golden Gate assembly can be found in

Table 15. Note that the first 12 base pairs and the last 15 base pairs are not part of the coding sequence and are needed for Golden Gate assembly.

b) Golden Gate assembly as used here was already described (Prielhofer et al., 2017). (1) The expression construct BB3aZ_pGAP_Adh2_CycTT was assembled as follows. The Adh2_GG_cured DNA fragment (Table 15) was cloned into the BB1_23 backbone, creating the BB1_23_Adh2. The expression construct was generated by Golden Gate assembly of BB3aZ_14* (backbone), BB1_23_Adh2 (coding sequence) BB1_12_pGAP (promoter), BB1_34_ScCYC1tt (terminator). (2) The expression construct BB3aZ_pGAP_Adh900_CycTT was assembled as follows. The Adh900_GG_cured DNA fragment (Table 15) was cloned into the BB1_23 backbone, creating the BB1_23_Adh900. The expression construct was generated by Golden Gate assembly of BB3aZ_14* (backbone), BB1_23_Adh900 (coding sequence) BB1_12_pGAP (promoter), BB1_34_ScCYC1tt (terminator). The plasmids and sequences are available in the Golden PiCS kit # 1000000133 (Addgene, Inc., USA).

Table 16: *ADH2* and *ADH900* native coding sequence and *ADH2* cured coding sequence with mutations in c.45G>A and c.660C>G and *ADH900* cured coding sequence with mutations in c.42C>G used for Golden Gate assembly. The first 12 base pairs and the last 15 base pairs are not part of the coding sequence and are used for Golden Gate assembly.

DNA fragment	DNA sequence 5’ to 3’
ADH2 coding sequence	(SEQ ID NO:80) ATGTCTCCAACTATCCCAACTACACAAAAGGCTGTTATCTTCGAGACCAACGGCG GTCCCCTAGAGTACAAGGACATTCCAGTCCCAAAGCCAAAGTCAAACGAACTTTT GATCAACGTTAAGTACTCCGGTGTCTGTCACACTGATTTGCACGCCTGGAAGGGT GACTGGCCATTGGACAACAAGCTTCCTTTGGTTGGTGGTCACGAAGGTGCTGGTG TCGTTGTCGCTTACGGTGAGAACGTCACTGGATGGGAGATCGGTGACTACGCTGG TATCAAATGGTTGAACGGTTCTTGTTTGAACTGTGAGTACTGTATCCAAGGTGCT GAATCCAGTTGTGCCAAGGCTGACCTGTCTGGTTTCACCCACGACGGATCTTTCC AGCAGTATGCTACTGCTGATGCCACCCAAGCCGCCAGAATTCCAAAGGAGGCTGA CTTGGCTGAAGTTGCCCCAATTCTGTGTGCTGGTATCACCGTTTACAAGGCTCTT AAGACCGCTGACTTGCGTATTGGCCAATGGGTTGCCATTTCTGGTGCTGGTGGAG GACTGGGTTCTCTTGCCGTTCAATACGCCAAGGCTCTGGGTTTGAGAGTTTTGGG TATTGATGGTGGTGCCGACAAGGGTGAATTTGTCAAGTCCTTGGGTGCTGAGGTC TTCGTCGACTTCACTAAGACTAAGGACGTCGTTGCTGAAGTCCAAAAGCTCACCA ACGGTGGTCCACACGGTGTTATTAACGTCTCCGTTTCCCCACATGCTATCAACCA ATCTGTCCAATACGTTAGAACTTTGGGTAAGGTTGTTTTGGTTGGTCTGCCATCT GGTGCCGTTGTCAACTCTGACGTTTTCTGGCACGTTCTGAAGTCCATCGAGATCA AGGGATCTTACGTTGGAAACAGAGAGGACAGTGCCGAGGCCATCGACTTGTTCAC CAGAGGTTTGGTCAAGGCTCCTATCAAGATTATCGGTCTGTCTGAACTTGCTAAG GTCTACGAACAGATGGAGGCTGGTGCCATCATCGGTAGATACGTTGTGGACACTT CCAAATAA

ADH900 coding sequence	(SEQ ID NO:81) ATGTCTGTGATGAAAGCCCTCGTGTACGGTGGTAAGAACGTCTTCGCCTGGAAAA ACTTCCCTAAACCAACTATCTTGCACCCAACAGATGTCATCGTTAAGACGGTGGC TACTACCATCTGCGGAACAGACTTGCACATCTTGAAAGGTGATGTTCCAGAGGTC AAACCTGAAACCGTCTTGGGTCATGAAGCAATTGGAGTCGTCGAATCTATCGGTG ATAACGTCAAAAACTTCAGCATTGGTGATAAGGTGCTGGTTTCATGCATCACCAG TTGTGGAAGCTGTTACTACTGTAAGAGAAACTTGCAGAGTCATTGCAAGACCGGT GGATGGAAATTAGGTCACGATTTGAACGGTACGCAGGCTGAGTTTGTCCGTATCC CATATGGAGACTTCTCATTGCACCGTATTCCTCATGAAGCAGATGAAAAGGCAGT TCTGATGCTGTCTGACATCTTACCTACTGCTTACGAAGTTGGTGTTCTTGCCGGA AATGTCCAAAAGGGAGACTCAGTTGCCATTGTCGGCGCCGGTCCAGTTGGTCTTG CCGCTCTGCTGACTGTCAAAGCCTTTGAGCCTTCTGAAATTATTATGATTGACAC TAACGATGAAAGACTGAGTGCCTCCTTGAAATTGGGAGCCACCAAGGCAGTCAAC CCAACCAAGGTCAGCAGTGTCAAAGATGCTGTTTATGATATTGTCAATGCCACTG TCCGCGTCAAGGAGAACGACCTGGAGCCAGGTGTCGATGTTGCCATTGAGTGTGT TGGTGTTCCTGACACGTTTGCAACTTGTGAAGAGATTATCGCCCCAGGTGGCCGT ATTGCCAATGTTGGTGTTCACGGCACTAAAGTGGATTTACAACTGCAAGACCTAT GGATCAAGAACATTGCTATCACCACCGGTTTGGTAGCCACATACTCCACTAAAGA CCTGTTGAAGCGAGTCTCTGACAAGTCTCTAGACCCTACACCACTGGTTACACAT GAGTTCAAGTTCAGTGAATTTGAGAAGGCCTATGAGACTTCTCAAAATGCTGCCA CCACCAAAGCCATCAAGATTTTCTTATCTGCCGATTAA
Adh2_GG_cured	(SEQ ID NO:82) GATAGGTCTCACATGTCTCCAACTATCCCAACTACACAAAAGGCTGTTATCTTCG AAACCAACGGCGGTCCCCTAGAGTACAAGGACATTCCAGTCCCAAAGCCAAAGTC AAACGAACTTTTGATCAACGTTAAGTACTCCGGTGTCTGTCACACTGATTTGCAC GCCTGGAAGGGTGACTGGCCATTGGACAACAAGCTTCCTTTGGTTGGTGGTCACG AAGGTGCTGGTGTCGTTGTCGCTTACGGTGAGAACGTCACTGGATGGGAGATCGG TGACTACGCTGGTATCAAATGGTTGAACGGTTCTTGTTTGAACTGTGAGTACTGT ATCCAAGGTGCTGAATCCAGTTGTGCCAAGGCTGACCTGTCTGGTTTCACCCACG ACGGATCTTTCCAGCAGTATGCTACTGCTGATGCCACCCAAGCCGCCAGAATTCC AAAGGAGGCTGACTTGGCTGAAGTTGCCCCAATTCTGTGTGCTGGTATCACCGTT TACAAGGCTCTTAAGACCGCTGACTTGCGTATTGGCCAATGGGTTGCCATTTCTG GTGCTGGTGGAGGACTGGGTTCTCTTGCCGTTCAATACGCCAAGGCTCTGGGTTT GAGAGTTTTGGGTATTGATGGTGGTGCCGACAAGGGTGAATTTGTCAAGTCCTTG GGTGCTGAGGTGTTCGTCGACTTCACTAAGACTAAGGACGTCGTTGCTGAAGTCC AAAAGCTCACCAACGGTGGTCCACACGGTGTTATTAACGTCTCCGTTTCCCCACA TGCTATCAACCAATCTGTCCAATACGTTAGAACTTTGGGTAAGGTTGTTTTGGTT GGTCTGCCATCTGGTGCCGTTGTCAACTCTGACGTTTTCTGGCACGTTCTGAAGT CCATCGAGATCAAGGGATCTTACGTTGGAAACAGAGAGGACAGTGCCGAGGCCAT CGACTTGTTCACCAGAGGTTTGGTCAAGGCTCCTATCAAGATTATCGGTCTGTCT GAACTTGCTAAGGTCTACGAACAGATGGAGGCTGGTGCCATCATCGGTAGATACG TTGTGGACACTTCCAAATAAGCTTAGAGACCGATC
Adh900_GG_cured	(SEQ ID NO:83) GATAGGTCTCACATGTCTGTGATGAAAGCCCTCGTGTACGGTGGTAAGAACGTGT TCGCCTGGAAAAACTTCCCTAAACCAACTATCTTGCACCCAACAGATGTCATCGT TAAGACGGTGGCTACTACCATCTGCGGAACAGACTTGCACATCTTGAAAGGTGAT GTTCCAGAGGTCAAACCTGAAACCGTCTTGGGTCATGAAGCAATTGGAGTCGTCG AATCTATCGGTGATAACGTCAAAAACTTCAGCATTGGTGATAAGGTGCTGGTTTC ATGCATCACCAGTTGTGGAAGCTGTTACTACTGTAAGAGAAACTTGCAGAGTCAT TGCAAGACCGGTGGATGGAAATTAGGTCACGATTTGAACGGTACGCAGGCTGAGT TTGTCCGTATCCCATATGGAGACTTCTCATTGCACCGTATTCCTCATGAAGCAGA TGAAAAGGCAGTTCTGATGCTGTCTGACATCTTACCTACTGCTTACGAAGTTGGT GTTCTTGCCGGAAATGTCCAAAAGGGAGACTCAGTTGCCATTGTCGGCGCCGGTC CAGTTGGTCTTGCCGCTCTGCTGACTGTCAAAGCCTTTGAGCCTTCTGAAATTAT TATGATTGACACTAACGATGAAAGACTGAGTGCCTCCTTGAAATTGGGAGCCACC AAGGCAGTCAACCCAACCAAGGTCAGCAGTGTCAAAGATGCTGTTTATGATATTG TCAATGCCACTGTCCGCGTCAAGGAGAACGACCTGGAGCCAGGTGTCGATGTTGC CATTGAGTGTGTTGGTGTTCCTGACACGTTTGCAACTTGTGAAGAGATTATCGCC CCAGGTGGCCGTATTGCCAATGTTGGTGTTCACGGCACTAAAGTGGATTTACAAC TGCAAGACCTATGGATCAAGAACATTGCTATCACCACCGGTTTGGTAGCCACATA CTCCACTAAAGACCTGTTGAAGCGAGTCTCTGACAAGTCTCTAGACCCTACACCA

	CTGGTTACACATGAGTTCAAGTTCAGTGAATTTGAGAAGGCCTATGAGACTTCTC AAAATGCTGCCACCACCAAAGCCATCAAGATTTTCTTATCTGCCGATTAAGCTTA GAGACCGATC

c) The *P. pastoris* ∆*aox1*∆*aox2* strain was made electrocompetent as described in Example **1a)**. The BB3aZ_pGAP_Adh2_CycTT expression construct and the BB3aZ_pGAP_Adh900_CycTT expression construct was linearized with AscI (New England Biolabs, Inc., USA) as per the manufacturer´s protocol and purified with the Hi Yield® Gel/PCR DNA Fragment Extraction Kits (Süd-Laborbedarf GmbH, Germany). 500 ng of the linearized plasmid was transformed into electrocompetent *P. pastoris* ∆*aox1*∆*aox2* as previously described in Example **1a)** and **1d)**. Positive transformants were selected on YPD plates with 25 µg/mL Zeocin. The successful integration of the expression construct was verified by PCR amplification with primers 109_BB3aN_ctrl_fwd and pGAP_goi_rev_v2 (Table 17) with genomic DNA as template. The created strains are called *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh2_CycTT and *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh900_CycTT.

d) Genomic DNA for PCR amplifications was isolated with the Wizard® Genomic DNA Purification Kit (Promega Corporation, USA) as per manufacturer's recommendations. The PCR amplification reactions were done with the Q5 polymerase (New England Biolabs, Inc., USA) as per manufacturer's recommendations.

Table 17: Polymerase chain reaction primers.

Primer Name	DNA sequence 5′ to 3′
109_BB3aN_ctrl_fwd	TTGATCTTTTCTACGGGGTGG (SEQ ID NO:84)
pGAP_goi_rev_v2	GGTGTTTTGAAGTGGTACGG (SEQ ID NO:85)

<u>Example 14</u>: Measurement of alcohol dehydrogenase activity in cell free extract of methanol utilization negative alcohol dehydrogenase defective strains.

To check for the successful deletion of the alcohol dehydrogenases on the phenotype level the alcohol dehydrogenase activity in cell free extracts with ethanol as a substrate was measured. Ethanol is generally regarded as the primary substrate for Adh2.

a) An overnight culture was done in 2 mL of YPD media in 24 well plates sealed by an air permeable membrane at 25°C and 280 rpm. The strains used, were from Example **12a)** *P. pastoris* Δ*aox1*Δ*aox2 Adh2*Δ::HphR, Example **12b)** *P. pastoris* Δ*aox1*Δ*aox2 Adh2*Δ::HphR *Adh900*Δ::KanMX and Example **1e)** *P. pastoris* Δ*aox1*Δ*aox2*. As an additional control the *P. pastoris* X33 (Thermo Fisher Scientific Inc., USA) and the *P. pastoris* X33 Δ*Adh2* (Nocon et al., 2014) was used.

b) The cell free extracts were prepared by centrifuging (16.000 g, 5 min, 4°C) the overnight culture and resuspending it in 1 mL phosphate buffered saline. After a second centrifuge step (16.000 g, 5 min, 4°C) the cells were resuspended in 0.5 mL of cell lysis buffer with glass beats. The cultures were lysed in a ribolyser (MP Biomedicals, Inc., USA) by bead beating for 3 x 20 seconds at 6 m/s with 1 minute cooling on ice in-between steps. After the lysis step the cultures were centrifuged (16.000 g, 5 min, 4°C) and the supernatant was transferred to a fresh Eppendorf tube and centrifuged again (16.000 g, 30 min, 4°C) to remove any carried over cell debris. After the second centrifugation step the supernatant was stored at -20°C till use.

c) The cell lysis buffer consisted of 20 mM HEPES, 420 mM NaCl, 1.5 mM $MgCl_2$, 10% Glycerol, 1 SIGMAFAST™ Protease Inhibitor Cocktail Tablets (Sigma-Aldrich GmbH) per 50 mL. The assay buffer consisted of 100 mM MOPS, 5 mM $MgSO_4$, 2 mM NAD^+ at pH 8.9.

d) The protein concertation of the cell free extracts was measured by Pierce™ BCA Protein Assay (Thermo Scientific, Inc., USA) as per manufacturer's recommendations and uniformly adjusted to a common concentration of 3.8 mg/mL for all samples.

e) The alcohol dehydrogenase activity assays were done in a 96 well plate. The measurements were done in a microplate reader (Tecan Group Ltd., Swiss) by

measuring the absorbance of NADH at 340 nm. Temperature was set at 42°C. To start the assay 20 µL cell free extracts were added to the assay buffer and equilibrated for 10 to 15 minutes before the addition of 1 M of ethanol as a substrate. The total end volume was 300 µL. The activity in mU/mg was calculated from the maximal linear absorption increase after addition of the substrate ethanol. One activity unit corresponds to 1 µmol substrate (NAD^+) consumed per minute. This was calculated from the absorption data using the Lambert-Beer law and the coefficient $\varepsilon_{NADH} = 6220\ M^{-1}\ cm^{-1}$.

f) The results show clearly the effect of AHD gene deletion on the Alcohol dehydrogenase activity of the cell free extracts (Table 18). By deleting the *ADH2* gene an activity reduction of 94% is achieved. This is additionally confirmed by the P. pastoris X33 strains. By deleting also the second alcohol dehydrogenase gene *ADH900* a combined activity reduction by 99% is observed.

Table 18: Alcohol dehydrogenase activity of cell free extracts on ethanol as a substrate.

	Alcohol dehydrogenase activity (mU/mg)		
	Mean	Standard Error	Clones tested
P. pastoris CBS2612 Δ*aox1*Δ*aox2*	1293.8	244.9	3
P. pastoris CBS2612 Δ*aox1*Δ*aox2* *Adh2*Δ::HphR	80.8	7.9	6
P. pastoris CBS2612 Δ*aox1*Δ*aox2* *Adh2*Δ::HphR *Adh900*Δ::KanMX	8.0	0.4	6
P. pastoris X33	1196.5	28.3	3
P. pastoris X33 *Adh2*Δ::HphR	88.5	2.1	7

<u>Example 15:</u> Measurement of methanol uptake rates of methanol utilization negative and alcohol dehydrogenase deficient strains.

To determine the methanol uptake rate the Mut^- and alcohol dehydrogenase deficient strains were cultivated in a bioreactor. The strains tested were the *P. pastoris* Δ*aox1*Δ*aox2* *Adh2*Δ::HphR, *P. pastoris* Δ*aox1*Δ*aox2* *Adh900*Δ::KanMX and *P. pastoris* Δ*aox1*Δ*aox2* *Adh2*Δ::HphR *Adh900*Δ::KanMX. The cultures was grown till a certain biomass concentration. Then a methanol pulse was applied and the methanol concentration was measured immediately after the pulse and approximately 20 hours later. The experimental setup is already described in detail in Example **7**. The goal was

to determine the specific methanol uptake rate of the alcohol dehydrogenase deficient strains and compare it to the methanol uptake rate measured in Example **7**.

a) The reactors filled with 300 mL BSM media were inoculated with 15 mL of *P. pastoris* Δ*aox1*Δ*aox2 adh2*Δ::HphR (reactor aR2 and aR4) and *P. pastoris* Δ*aox1*Δ*aox2 adh900*Δ::KanMX (reactor aR1 and aR3). The target start OD_{600} was 2. At the end of the batch phase as indicated by a dissolved oxygen spike, a 50% (w/w) glucose feed was started at 2.8 mL/h for 24 hours to increase the biomass. Two hours after the glucose feed start a 50%(v/v) methanol shot was given to increase the methanol concentration to 1.5% (measured concentration was aR1=1.64%, aR2=1.66%, aR3=1.59% and aR4=1.67%). This was done to induce methanol consumption. At the end of the glucose feed phase samples for cell dry weight and HPLC were taken.

b) After the glucose feed phase the agitation and gassing was set to a constant 750 rpm and 9.5 sL/h. An additional 50% methanol pulse was added to increase the concentration to 1.5% and immediately a HPLC sample was taken (measured concentration was aR1=1.36%, aR2=1.44%, aR3=1.34% and aR4=1.45%). The concentration was measured again after 18.4 hours and used to determine the specific methanol uptake rate ($q_{methanol}$).

c) A separate bioreactor cultivation was started to measure the uptake rate of the double *ADH* deletion strain *P. pastoris* Δ*aox1*Δ*aox2 adh2*Δ::HphR *adh900*Δ::KanMX. The cultivation was carried as explained in this example and Example **7**. The *P. pastoris* Δ*aox1*Δ*aox2 Adh2*Δ::HphR *Adh900*Δ::KanMX was inoculated into reactor cR3 and cR4. The target start OD_{600} was 2. At the end of the batch phase as indicated by a dissolved oxygen spike, a 50% (w/w) glucose feed was started at 2.4 mL/h for 24 hours to increase the biomass. Two hours after the glucose feed start a 50% (v/v) methanol shot was given to increase the methanol concentration to 1.5% (measured concentration was cR3=1.50% and cR4=1.47%). This was done to induce methanol consumption. At the end of the glucose feed phase samples for cell dry weight and HPLC were taken.

d) After the glucose feed phase the agitation and gassing was set to a constant 750 rpm and 9.5 sL/h. An additional 50% methanol pulse was added to increase the concentration to 1.5% and immediately a HPLC sample was taken (measured concentration was cR3=1.37% and cR4=1.49%). The concentration was

measured again after 19.5 hours and used to determine the specific methanol uptake rate ($q_{methanol}$).

e) The specific methanol uptake rate was calculated as in Example **7d)**. By deleting the *ADH2* a surprising and substantial reduction in methanol uptake rate was achieved (Table 19). The dc/dt of methanol is at 0.07 to 0.06 g L^{-1} h^{-1} for aR2 and aR4 which is only slightly higher that the evaporation observed in Example **6**. In contrast the methanol uptake rate of the *ADH900* deletion strain is not reduced and was in fact slightly higher than the measured uptake rate of the *P. pastoris* Δ*aox1*Δ*aox2* in Example **7**. This difference can be attributed to slightly different conditions between Example **7** and this example, the difference being a slightly higher reactor volume and methanol concentration after the methanol pulse. The double *ADH* deletion strain does not show an observable reduction of the methanol uptake rate compared to the already low uptake rate of the *ADH2* deletion strain. These results unexpectedly confirmed that the *ADH2* gene and its product, the enzyme Adh2, are to the biggest extend responsible for the characteristics observed for *P. pastoris* Δ*aox1*Δ*aox2* and that the observations in Example **7** are not the consequence of spontaneous methanol oxidation or the promiscuous activity of any other enzyme.

Therefore, it was concluded that the main responsible gene and enzyme for methanol uptake in the *P. pastoris* Δ*aox1*Δ*aox2* is the *ADH2* gene and its product the enzyme Adh2.

Table 19: Overview specific methanol uptake rates ($q_{methanol}$) and apparent methanol loss (dc/dt) for the *ADH* deletion strains. * The methanol concentration is measured at 19.5 h.

Reactor	*ADH* gene	Volume (mL)	CDW (g/L)	Methanol at 0 h (g/L)	Methanol at 18.4 h (g/L)	dc/dt (g L^{-1} h^{-1})	$q_{methanol}$ (mg g^{-1} h^{-1})
aR1	*Adh900*Δ::KanMX	392	73.4	10.75	3.49	0.40	5.38
aR2	*Adh2*Δ::HphR	385	74.1	11.40	10.20	0.07	0.89
aR3	*Adh900*Δ::KanMX	395	71.2	10.61	3.24	0.40	5.63
aR4	*Adh2*Δ::HphR	382	74.7	11.47	10.34	0.06	0.82
cR3	*Adh2*Δ::HphR *Adh900*Δ::KanMX	373	73.9	10.8	10.0*	0.04	0.55
cR4	*Adh2*Δ::HphR *Adh900*Δ::KanMX	369	73.9	11.8	10.6*	0.06	0.85

<u>Example 16</u>: Measurement of methanol uptake rates of methanol utilization negative and alcohol dehydrogenase overexpression strains.

To confirm that Adh2 is the responsible enzyme for the consumption of methanol in the *P. pastoris* ∆*aox1*∆*aox2* and to investigate if it is possible to increase the methanol uptake rate with overexpression of the *ADH2* or *ADH900* genes the specific methanol uptake rate of *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh2_CycTT and *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh900_CycTT strain was measured in a bioreactor cultivation. The experiment was done as described in Example **7** and **15**.

a) The reactors filled with 300 mL BSM media were inoculated with 15 mL of *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh2_CycTT (reactor R1 and R3) and *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh900_CycTT (R2 and R4). The target start OD_{600} was 2. At the end of the batch phase as indicated by a dissolved oxygen spike, a 50% (w/w) glucose feed was started at 2.8 mL/h for 24 hours to increase the biomass. Two hours after the glucose feed start a 50% (v/v) methanol shot was given to increase the methanol concentration to 1.5% (measured concentration was R1=1.56%, R2=1.53%, R3=1.52% and R4=1.54%). This was done to induce methanol consumption. At the end of the glucose feed phase samples for cell dry weight and HPLC were taken.
b) After the glucose feed phase the agitation and gassing was set to a constant 750 rpm and 9.5 sL/h. An additional 50% methanol pulse was added to increase the concentration to 1.5% and immediately a HPLC sample was taken (measured concentration was R1=1.30%, R2=1.30%, R3=1.35% and R4=1.30%). The concentration was measured again after 4.1, 20.1 hours and used to determine the specific methanol uptake rate ($q_{methanol}$).
c) An additional sampling time point at 4.1 hours was chosen because a higher methanol uptake rate was expected. The average methanol uptake rate at 4.1 hours was 7.72 mg g^{-1} h^{-1} for the *ADH2* overexpressing strain *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh2_CycTT and 5.61 mg g^{-1} h^{-1} for the *ADH900* overexpressing strain *P. pastoris* ∆*aox1*∆*aox2* BB3aZ_pGAP_Adh900_CycTT. After 20.1 hours the average uptake rate decreases to 6.35 mg g^{-1} h^{-1} for the *ADH2* overexpressing strain and to 5.16 mg g^{-1} h^{-1} for the *ADH900* overexpressing strain (Table 20). As previously discussed in Example **15e)** no significant difference to the parent *P. pastoris* ∆*aox1*∆*aox2* can be observed when deleting the *ADH900* gene and the same is true when overexpressing the

ADH900 gene. On the other hand, the *ADH2* overexpressing strain had a 37% and 23% higher uptake rate at 4.1 and 21.7 hours compared to the *ADH900* overexpressing strain. Further underlining the unexpected finding that Adh2 is the responsible enzyme for the consumption of methanol and that it is possible to increase the methanol consumption with overexpression of the *ADH2* gene.

Table 20: Overview specific methanol uptake rates ($q_{methanol}$) and apparent methanol loss (dc/dt) for the ADH overexpressing strains.

Reactor	*ADH* gene	Volume (mL)	CDW (g/L)	Methanol at 0 h (g/L)	Methanol at 4.1 h (g/L)	Methanol at 20.1 h (g/L)	$q_{methanol}$ (mg g^{-1} h^{-1}) At 4.1 h	$q_{methanol}$ (mg g^{-1} h^{-1}) At 20.1 h
R1	P_{GAP}*ADH2*	398	72.9	10.3	8.0	1.2	7.65	6.18
R2	P_{GAP}*ADH900*	400	71.7	10.3	8.7	2.9	5.30	5.09
R3	P_{GAP}*ADH2*	406	70.7	10.7	8.4	1.3	7.80	6.51
R4	P_{GAP}*ADH900*	399	71.3	10.3	8.6	2.8	5.91	5.23

<u>Example 17:</u> Strain generation with methanol-inducible promoters for *ADH2* overexpression.

For the purpose of investigating the effect of methanol inducible *ADH2* overexpression, two overexpression constructs were created being composed of methanol inducible promoters P_{AOX1} PP7435_chr4 (237941…238898) and P_{FLD1} PP7435_Chr3 (262922…263518) controlling the expression of the *ADH2* coding sequence. The *ADH2* coding sequence was modified to eliminate BbsI and BsaI restriction sites in the coding sequence without affecting the amino acid sequence of the gene product (**Table 16**). The generated strains were designated *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pAOX1_Adh2_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pFLD1_Adh2_CycTT.

a) The expression constructs were created using Golden Gate assembly as already described (Prielhofer et al., 2017). (1) The expression construct BB3aZ_pAOX1_Adh2_CycTT was assembled as follows. The Adh2_GG_cured DNA fragment (**Table 16**) was cloned into the BB1_23 backbone, creating the BB1_23_Adh2. The expression construct was generated by Golden Gate assembly of BB3aZ_14* (backbone), BB1_23_Adh2 (coding sequence) BB1_12_pAOX1 (promoter), BB1_34_ScCYC1tt (terminator). (2) The expression construct BB3aZ_pFLD1_Adh2_CycTT was generated by Golden Gate assembly of BB3aZ_14* (backbone), BB1_23_Adh2 (coding sequence) BB1_12_pFLD1

(promoter), BB1_34_ScCYC1tt (terminator). The plasmids and sequences are available in the Golden PiCS kit # 1000000133 (Addgene, Inc., USA).

b) The *P. pastoris* Δ*aox1*Δ*aox2* strain was made electrocompetent as described in Example **1a)**. The BB3aZ_pAOX1_Adh2_CycTT expression construct and the BB3aZ_pFLD1_Adh2_CycTT expression construct was linearized with AscI (New England Biolabs, Inc., USA) as per the manufacturer´s protocol and purified with the Hi Yield® Gel/PCR DNA Fragment Extraction Kits (Süd-Laborbedarf GmbH, Germany). 500 ng of the linearized plasmid was transformed into electrocompetent *P. pastoris* Δ*aox1*Δ*aox2* as previously described in Example **1a)** and **1d)**. Positive transformants were selected on YPD plates with 25 µg/mL Zeocin. The successful integration of the expression construct was verified by PCR amplification with primers 109_BB3aN_ctrl_fwd and pGAP_goi_rev_v2 (**Table 17**) with genomic DNA as template. The created strains are called *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pAOX1_Adh2_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pFLD1_Adh2_CycTT.

c) Genomic DNA for PCR amplifications was isolated with the Wizard® Genomic DNA Purification Kit (Promega Corporation, USA) as per manufacturer's recommendations. The PCR amplification reactions were done with the Q5 polymerase (New England Biolabs, Inc., USA) as per manufacturer's recommendations.

<u>Example 18</u>: Measurement of specific methanol uptake rates of methanol utilization negative strains with methanol inducible promoters for *AHD2* overexpression.

To investigate the effect of *ADH2* overexpression with methanol inducible promotors on the specific methanol uptake rate, a bioreactor cultivation was set up as described in Example **7** and Example **16**. For this purpose, the strains *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pAOX1_Adh2_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pFLD1_Adh2_CycTT generated in Example **17** were used.

a) The reactors filled with 300 mL BSM media were inoculated with 15 mL of *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pAOX1_Adh2_CycTT (reactor R1 and R2) and *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pFLD1_Adh2_CycTT (R3 and R4). The target start OD_{600} was 2. At the end of the batch phase as indicated by a dissolved oxygen spike, a 50% (w/w) glucose feed was started at 2.8 mL/h for 24 hours to increase

the biomass. Two hours after the glucose feed start a 50% (v/v) methanol shot was given to increase the methanol concentration to 1.5% (measured concentration was R1=1.57%, R2=1.57%, R3=1.57% and R4=1.70%). This was done to induce methanol consumption and the methanol inducible promotors. At the end of the glucose feed phase samples for cell dry weight and HPLC were taken.

b) After the glucose feed phase the agitation and gassing was set to a constant 750 rpm and 9.5 sL/h. An additional 50% methanol pulse was added to increase the concentration to 1.5% and immediately a HPLC sample was taken (measured concentration was R1=1.42%, R2=1.39%, R3=1.39% and R4=1.42%). The concentration was measured again after 4.2 hours and used to determine the specific methanol uptake rate ($q_{methanol}$). After the initial pulse was nearly consumed a second methanol pulse was applied and immediately a HPLC sample was taken (measured concentration was R1=1.61%, R2=1.65%, R3=1.50% and R4=1.47%). The concentration was measured again after 6.2 hours and used to determine the specific methanol uptake rate ($q_{methanol}$) a second time.

c) The specific methanol uptake rate ($q_{methanol}$) was determined using two methanol pulses. The time between the first pulse and the sampling time point was 4.2 hours. The time between the second methanol pulse and the sampling point was 6.2 hours. The average methanol uptake rate after 4.2 hours after the first pulse was 8.3 mg g^{-1} h^{-1} for the P_{FLD1}*ADH2* overexpressing strain *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pFLD1_Adh2_CycTT and 11.6 mg g^{-1} h^{-1} for the P_{AOX1}*ADH2* overexpressing strain *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pAOX1_Adh2_CycTT (**Table 21**). 6.2 hours after the second methanol pulse the average uptake rate increased to 10.1 mg g^{-1} h^{-1} for the P_{FLD1}*ADH2* overexpressing strain and to 13.8 mg g^{-1} h^{-1} for the P_{AOX1}*ADH2* overexpressing strain (**Table 22**). This data shows that longer methanol induction times lead to an increased expression of Adh2 and specific methanol uptake rate when methanol inducible promotors are used. The strains can therefore sustain and even increase the specific methanol uptake rate over time in a medium with methanol as the only energy and carbon source. Compared to the *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT strain described in Example **7** the specific methanol uptake rate was increased on average by 1.6 fold for the *P. pastoris* Δ*aox1*Δ*aox2*

BB3aZ_pFLD1_Adh2_CycTT and by 2.3 fold for the *P. pastoris* Δ*aox1*Δ*aox2* BB3aZ_pAOX1_Adh2_CycTT.

Table 21: Overview of the specific methanol uptake rates ($q_{methanol}$) with the methanol inducible *ADH2* overexpression after the first methanol pulse.

Reactor	*ADH* gene	Volume (mL)	CDW (g/L)	Methanol 1st pulse (g/L)	Methanol 4.2 h after 1st methanol pulse (g/L)	$q_{methanol}$ (mg g^{-1} h^{-1})
R1	P_{AOX1}*ADH2*	400	77.7	11.2	7.6	11.2
R2	P_{AOX1}*ADH2*	401	76.4	11.0	7.2	11.9
R3	P_{FLD1}*ADH2*	400	75.1	11.0	8.5	8.0
R4	P_{FLD1}*ADH2*	401	75.8	11.3	8.6	8.5

Table 22: Overview of the specific methanol uptake rates ($q_{methanol}$) with the methanol inducible *ADH2* overexpression after the second methanol pulse.

Reactor	*ADH* gene	Volume (mL)	CDW (g/L)	Methanol 2nd pulse (g/L)	Methanol 6.2 h after 2st methanol pulse (g/L)	$q_{methanol}$ (mg g^{-1} h^{-1})
R1	P_{AOX1}*ADH2*	394	72.7	12.7	6.3	14.4
R2	P_{AOX1}*ADH2*	393	72.6	13.1	7.1	13.3
R3	P_{FLD1}*ADH2*	393	71.1	11.9	7.4	10.2
R4	P_{FLD1}*ADH2*	394	73.0	11.7	7.1	10.1

<u>Example 19</u>: Generation of *ADH2* overexpressing strains producing a secreted recombinant protein.

To investigate if the *ADH2* overexpression and the consequential increase in specific methanol uptake rate have an impact on the recombinant protein production the *P. pastoris* Δ*aox1*Δ*aox2* producing HSA and vHH from Example **3** were transformed with the P_{AOX1}*ADH2* and P_{FLD1}*ADH2* overexpression constructs and screened in small scale with an adapted protocol described in Example **4.**

a) The overexpression constructs were done using Golden Gate assembly as already described (Prielhofer et al., 2017). (1) The expression construct BB3aK_pAOX1_Adh2_CycTT was generated by Golden Gate assembly of BB3aK_14* (backbone), BB1_23_Adh2 (coding sequence) from Example **17**, BB1_12_pAOX1 (promoter), BB1_34_ScCYC1tt (terminator). (2) The expression construct BB3aK_pFLD1_Adh2_CycTT was generated by Golden Gate assembly of BB3aK_14* (backbone), BB1_23_Adh2 (coding sequence) from Example **17**, BB1_12_pFLD1 (promoter), BB1_34_ScCYC1tt (terminator). The plasmids and sequences are available in the Golden PiCS kit # 1000000133 (Addgene, Inc., USA).

b) The *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT strain were made electrocompetent as described in Example **1a)**. The BB3aK_pAOX1_Adh2_CycTT expression construct and the BB3aK_pFLD1_Adh2_CycTT expression construct was linearized with AscI (New England Biolabs, Inc., USA) as per the manufacturer´s protocol and purified with the Hi Yield® Gel/PCR DNA Fragment Extraction Kits (Süd-Laborbedarf GmbH, Germany). 500 ng of the linearized plasmid was transformed into electrocompetent *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT as previously described in Example **1a)** and **1d)**. Positive transformants were selected on YPD plates with 500 µg/mL geneticin and 25 µg/mL zeocin for the *P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT_BB3aK_pAOX1_Adh2_CycTT and the *P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT_BB3aK_pFLD1_Adh2_CycTT transformants. The *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT_BB3aK_pAOX1_Adh2_CycTT and *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT_BB3aK_pFLD1_Adh2_CycTT transformants were selected on YPD plates with 500 µg/mL geneticin and 100 µg/L nourseothricin. Multiple clones per transformation were selected for further screening (Example **20**). The created strains are named: *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{AOX1}*ADH2*, *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{FLD1}*ADH2*, *P.*

pastoris Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{AOX1}*ADH2* and *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{FLD1}*ADH2*.

<u>Example 20</u>: Small scale screening of *ADH2* overexpressing strains producing a secreted recombinant protein.

Multiple clones from the transformants described in Example **19** were tested in small scale screening to investigate the impact of the *ADH2* overexpression on recombinant protein production. The screening procedure was adapted from the two shot–extended protocol and the standard protocol described in Example **4.**

a) For the pre-culture of the *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{AOX1}*ADH2* and *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{FLD1}*ADH2* clones were inoculated in 2 mL YPD with 500 µg/mL geneticin and 100 µg/mL nourseothricin, the *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{AOX1}*ADH2* and *P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{FLD1}*ADH2* clones were inoculated on 500 µg/mL Geneticin and 25 µg/mL Zeocin. The parental strains were inoculated in two replicates on YPD with 100 µg/mL Nourseothricin or 25 µg/mL Zeocin based on the antibiotic resistance used for selection. For each expression construct eleven clones were picked for screening. Pre-culture and screening cultures were cultivated in 24 well plates sealed with an air permeable membrane and incubated on 25°C on 280 rpm. The screening culture was inoculated with a start optical density (OD_{600}) of 8 into 2 mL of minimal media (ASMv6) with a slow glucose release system EnPump200 (Enpresso GmbH, Germany) based on a polysaccharide solution and an enzyme to keep the cultures in glucose limit. The strains were compared with two different methanol feed procedures differing in total methanol received and incubation time (**Table 23**).
b) After the incubation period 1 mL of each culture was removed and centrifuged in a pre-weighted Eppendorf tube. The supernatant was removed and the protein concentration was measured with the Caliper LabChip GXII Touch (PerkinElmer, inc., USA) as per the manufacturer´s instructions. The wet cell weight was determined by weighting the Eppendorf tubes with the cell pellet and calculated as follows: Weight (full) – weight (empty) = wet cell weight (WCW) (g/L). Out of this data the yield was calculated: Yield (µg/g) = protein concentration / wet cell weight.

Table 23: Overview of the screening strategies used for testing the secreted protein production yield of the transformed strains in Example 20. *The first shot was 0.5% (v/v) methanol.

Protocol	Incubation period	Polysaccharide (g/L)	Enzyme (%)	Methanol shot	Total methanol (v/v)	Methanol shot time points (h)
Standard	48 h	25	0.35	4 x	3.5%	4* ,19, 27,43
Two shot - extended	72 h	25	0.20	2 x	2%	3, 43

c) The results are summed up in **Table 24**. Surprisingly, the overexpression increased the protein yield (µg/g) by up to 1.7 fold for vHH and 2.3 fold for HSA when compared to the parental Mut$^-$ strains. Indeed, proving that the overexpression of the *ADH2* improves recombinant protein production of the *P. pastoris* Δ*aox1*Δ*aox2*.

d) Average performing strains were selected for bioreactor cultivation. The successful integration of the expression construct was verified by PCR amplification with primers 109_BB3aN_ctrl_fwd and pGAP_goi_rev_v2 (**Table 17**) with genomic DNA as template. Genomic DNA for PCR amplifications was isolated with the Wizard® Genomic DNA Purification Kit (Promega Corporation, USA) as per manufacturer's recommendations. The PCR amplification reactions were done with the Q5 polymerase (New England Biolabs, Inc., USA) as per manufacturer's recommendations.

Table 24: Average secreted product yield in µg product / g WCW with standard deviation in different screening conditions. *t-test statistically significant difference ($p<0.05$) from the parent strain.

		Screening protocol	
Descriptive name	Name	Two shot - extended	Standard Mut^S
Parent strain: Δ*aox1*Δ*aox2* P_{AOX1}vHH	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT	1766	1428
Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{AOX1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT BB3aK_pAOX1_Adh2_CycTT	2394* ±640	2441* ±211
Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{FLD1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT BB3aK_pFLD1_Adh2_CycTT	1907 ±381	1810* ±175
Parent strain: Δ*aox1*Δ*aox2* P_{AOX1}HSA	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT	635	459
Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{AOX1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pAOX1_Adh2_CycTT	1485* ±183	711* ±135
Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{FLD1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pFLD1_Adh2_CycTT	1198 ±96	621* ±85

<u>Example 21</u>: The methanol utilization negative strain with *ADH2* overexpression producing HSA as a model protein. Cultivated with strategy 3 - A feed strategy with a glucose/methanol co-feed phase and a separated methanol only feed phase.

A bioreactor cultivation was performed to evaluate the recombinant protein producing ability of the methanol utilization negative *ADH2* overexpressing strain generated in Example **19** and selected in Example **20**. For this purpose, *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pAOX1_Adh2_CycTT (Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{AOX1}*ADH2*) and *P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pFLD1_Adh2_CycTT (Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{FLD1}*ADH2*) strains were cultivated with the strategy 3 as described in Example **10** and Example **11**.

a) This bioreactor cultivation was separated into three phases. (1) Phase one was the batch phase. The reactors were inoculated with the production strains with a start OD_{600} of 2. The inoculation was done as described in Example **7a) b)**. The end of the batch phase was indicated by a dissolved oxygen spike. (2) At this point Phase two was started. Phase two consisted of a 50% (w/w) glucose feed

at 4.8 mL/h for 25 hours. At the start of Phase two a 50% (v/v) methanol pulse was applied to increase methanol concentration to the target of 1.5% (v/v) and a subsequent methanol feed was started to counteract methanol consumption, evaporation and dilution by the glucose feed. (3) Phase three consisted of a methanol only feed for 19.6 (R1, R2) and 21.6 (R5, R6) hours. Methanol concentration was measured at line with HPLC as described in Example **6d)**. An additional compensation pulse was added if necessary. The methanol feed was calculated in hourly intervals as in Example **8** and **9b)**. The strains used in each of the reactors R1, R2, R5 and R6 are identified in **Table 25**.

b) The process and productivity data can be found in **Table 26** and **Table 27**. The maximal and minimal methanol concentration throughout the cultivation of reactors R1, R2, R5 and R6 ranged from 8.0 g/L to 13.6 g/L. Reactors R1, R2 and R5, R6 were producing HSA as a model protein. The specific productivity (q_P) at 68.7 hours in phase 3 shows a positive impact of the *ADH2* overexpression with either the P_{AOX1} or P_{FLD1} compared to Example **10** (**Table 28**, **Table 29**). A weighted average of the q_P was calculated for Example **10,** timepoints 45.02 to 69.58 hours for easier comparison. The q_P in Example **10** from timepoint 45.02 to 69.58 h is on average 40.9 µg g^{-1} h^{-1}. The q_P in the present example for reactor the P_{AOX1}*ADH2* overexpression (R1, R2) from timepoint 49.1 to 68.7 hours is on average 84.3 µg g^{-1} h^{-1}. This is a **2 fold** increase compared to Example **10** (**Table 28**). The P_{FLD1}*ADH2* overexpression (R5, R6) in the similar timeframe (47.1 to 68.7 hours) show an average q_P of 71 µg g^{-1} h^{-1}. This represents a **1.7 fold** increase in q_P (**Table 29**). Volumetric productivity at timepoint 68.7 hours is increased by **1.21 fold** for P_{AOX1}*ADH2* overexpression (**Table 30**) and **1.13** for the P_{FLD1}*ADH2* overexpression (**Table 31**). The increased q_P and volumetric productivity in this example is demonstrating the benefits of the *ADH2* overexpression in the *P. pastoris* ∆*aox1*∆*aox2* strain for recombinant protein production.

Table 25: Overview of the strains used in Example 21 and Example 22.

Reactor	Descriptive name	Name
R1	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{AOX1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pAOX1_Adh2_CycTT
R2	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{AOX1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pAOX1_Adh2_CycTT
R3	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{AOX1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT BB3aK_pAOX1_Adh2_CycTT
R4	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{AOX1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT BB3aK_pAOX1_Adh2_CycTT
R5	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{FLD1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pFLD1_Adh2_CycTT
R6	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}HSA P_{FLD1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pN21_pAOX1_HSAopt_CycTT BB3aK_pFLD1_Adh2_CycTT
R7	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{FLD1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT BB3aK_pFLD1_Adh2_CycTT
R8	*P. pastoris* Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{FLD1}*ADH2*	*P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_αMF-vHH_CycTT BB3aK_pFLD1_Adh2_CycTT

Table 26: Bioreactor cultivation process data and specific productivity (q_P) for HSA with the **P_{AOX1}*ADH2*** overexpression from Example 21. *Represents a control sample after the methanol pulse.

P_{AOX1} HSA P_{AOX1} *ADH2*	Time (h)	Volume (mL)		YDM (g/L)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) ($\mu g\ g^{-1}\ h^{-1}$)		Methanol concentration (g/L)	
Phase		R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1	24.0	318	317	24.8	24.0					0.0	0.0
2	26.2									*9.4	*9.9
	49.1	502	501	100.4	100.2	208.9	218.6	96.0	101.1	13.8	13.6
3	68.7	517	517	95.2	95.8	445.6	423.8	90.2	78.3	8.0	8.7

Table 27: Bioreactor cultivation process data and specific productivity (q_P) for HSA with the **P_{FLD1}*ADH2*** overexpression from Example 21. *Represents a control sample after the methanol pulse.

P_{AOX1} HSA P_{FLD1} *ADH2*	Time (h)	Volume (mL)		YDM (g/L)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) ($\mu g\ g^{-1}\ h^{-1}$)		Methanol concentration (g/L)	
Phase		R5	R6	R5	R6	R5	R6	R5	R6	R5	R6
1	21.9	308	306	25.2	25.2						
2	23.8									*10.6	*11.0
	47.1	488	486	101.2	100.5	208.5	203.9	94.2	93.1	11.6	11.0
3	68.7	515	513	95.4	95.2	412.5	395.6	73.0	69.0	10.8	10.4

Table 28: Comparison of specific productivity (q_P) of methanol utilization negative strain producing HSA from Example 10 to the **P_{AOX1}*ADH2*** overexpressing strain from Example 21 in phase 3.

Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA			Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{AOX1}*ADH2***			Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA		Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{AOX1}*ADH2***		**Fold increase**
	q_P ($\mu g\ g^{-1}\ h^{-1}$)			q_P ($\mu g\ g^{-1}\ h^{-1}$)			q_P ($\mu g\ g^{-1}\ h^{-1}$)			
Time (h)	R1	R2	Time (h)	R1	R2	Time (h)	Average	Time (h)	Average	
45.02			49.1			45.02		49.1		
53.00	38.6	21.5				53.00	**Weighted average**			
69.58	47.3	44.9	68.7	90.2	78.3	**69.58**	**40.9**	**68.7**	**84.3**	**2.06**

Table 29: Comparison of specific productivity (q_P) of methanol utilization negative strain producing HSA from Example 10 to the **P_{FLD1}*ADH2*** overexpressing strain from Example 21 in phase 3.

Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA			Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{FLD1}*ADH2***			Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA		Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{FLD1}*ADH2***		**Fold increase**
	q_P ($\mu g\ g^{-1}\ h^{-1}$)			q_P ($\mu g\ g^{-1}\ h^{-1}$)			q_P ($\mu g\ g^{-1}\ h^{-1}$)			
Time (h)	R1	R2	Time (h)	R5	R6	Time (h)	Average	Time (h)	Average	
45.02			47.1			45.02		47.1		
53.00	38.6	21.5				53.00	**Weighted average**			
69.58	47.3	44.9	68.7	73.0	69.0	**69.58**	**40.9**	**68.7**	**71.0**	**1.74**

Table 30: Comparison of volumetric productivity of methanol utilization negative strain producing HSA from Example 10 to the **P_{AOX1}*ADH2*** overexpressing strain from Example 21. * Corrected recombinant concentration (C_{cP}) for the biomass volume, $C_{cP}= C_p*(1-C_x*F_c)$, $F_c=0.0033$.

	Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA		Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{AOX1}*ADH2***		Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA	Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{AOX1}*ADH2***	**Fold increase**
	R1	R2	R1	R2	Average	Average	
Recombinant protein (mg/L)	357.3	371.8	445.6	423.8	364.6	434.7	
Biomass (g/L)	95.9	96.3	95.2	95.8	96.1	95.5	
*Corrected recombinant protein (mg/L)	244.2	253.6	305.6	289.8	248.9	297.7	
Time (h)					69.58	68.7	
Volumetric productivity (mg L^{-1} h^{-1})					**3.58**	**4.33**	**1.21**

Table 31: Comparison of volumetric productivity of methanol utilization negative strain producing HSA from Example 10 to the **P_{FLD1}*ADH2*** overexpressing strain from Example 21. * Corrected recombinant concentration (C_{cP}) for the biomass volume, $C_{cP}= C_p*(1-C_x*F_c)$, $F_c=0.0033$.

	Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA		Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{FLD1}*ADH2***		Example 10 Δ*aox1*Δ*aox2* P_{AOX1}HSA	Example 21 Δ*aox1*Δ*aox2* P_{AOX1}HSA **P_{FLD1}*ADH2***	Fold increase
	R1	R2	R5	R6	Average	Average	
Recombinant protein (mg/L)	357.3	371.8	412.5	395.6	364.6	404.1	
Biomass (g/L)	95.9	96.3	95.4	95.2	96.1	95.3	
*Corrected recombinant protein (mg/L)	244.2	253.6	282.6	271.3	248.9	277.0	
Time (h)					69.58	68.7	
Volumetric productivity (mg L^{-1} h^{-1})					**3.58**	**4.03**	**1.13**

Example 22: The methanol utilization negative strain with *ADH2* overexpression producing vHH as a model protein. Cultivated with strategy 3 - A feed strategy with a glucose/methanol co-feed phase and a separated methanol only feed phase.

A bioreactor cultivation was performed to evaluate the recombinant protein producing ability of the methanol utilization negative *ADH2* overexpressing strain generated in Example **19** and selected in Example **20**. For this purpose, *P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_vHH_CycTT BB3aK_pAOX1_Adh2_CycTT (Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{AOX1}*ADH2*) and *P. pastoris* Δ*aox1*Δ*aox2* pPM2pZ30_pAOX1_vHH_CycTT BB3aK_pFLD1_Adh2_CycTT (Δ*aox1*Δ*aox2* P_{AOX1}vHH P_{FLD1}*ADH2*) strains were cultivated with the strategy 3 as described in Example **10** and Example **11**.

a) This bioreactor cultivation was separated into three phases. (1) Phase one was the batch phase. The reactors were inoculated with the production strains with a start OD_{600} of 2. The inoculation was done as described in Example **7a) b)**. The end of the batch phase was indicated by a dissolved oxygen spike. (2) At this point Phase two was started. Phase two consisted of a 50% (w/w) glucose feed at 4.8 mL/h for 25 hours. At the start of Phase two a 50% (v/v) methanol pulse was applied to increase methanol concentration to the target of 1.5% (v/v) and a subsequent methanol feed was started to counteract methanol consumption, evaporation and dilution by the glucose feed. (3) Phase three consisted of a methanol only feed for 43.6 (R3, R4) and 44.6 (R7, R8) hours. Methanol concentration was measured at line with HPLC as described in Example **6d)**. An additional compensation pulse was added if necessary. The methanol feed was calculated in hourly intervals as in Example **8** and **9b)**. The strains used in each reactors R3, R4, R7 and R8 can be found in **Table 25**.
b) The process and productivity data can be found in **Table 32** and **Table 33**. The maximal and minimal methanol concentration throughout the cultivation of reactors R3, R4, R7 and R8 ranged from 8.6 g/L to 14.4 g/L. Reactors R3, R4 and R7, R8 were producing vHH as a model protein. The *ADH2* overexpression had a positive impact on specific productivity (q_P) compared to Example **11**. For easier comparison a weighted average of the q_P was calculated for phase two (timepoints 20.0 to 53.6 hours) of Example **11**. The comparison shows that P_{AOX1}*ADH2* overexpression (R3, R4) has a **1.71 fold** and the P_{FLD1}*ADH2*

overexpression (R7, R8) a **1.78** fold increase on q_P in phase two (**Table 34**, **Table 35**). At later timepoints in phase three the improvements are even larger. At timepoints 92.7 and 91.7 the q_P is the increase by **3.76 fold** (P_{AOX1}*ADH2* overexpression) and **3.86 fold** (P_{FLD1}*ADH2* overexpression) (**Table 34**, **Table 35**). Additionally, volumetric productivity was improved at least **1.9 fold** compared to the parental strain in Example **11** in both cases (**Table 35**, **Table 36**). The increased q_P and volumetric productivity in this example is demonstrating the benefits of the *ADH2* overexpression in the *P. pastoris* ∆*aox1*∆*aox2* strain for recombinant protein production.

Table 32: Bioreactor cultivation process data and specific productivity (q_P) for vHH with the **P_{AOX1}*ADH2*** overexpression from Example 22. *Represents a control sample after the methanol pulse.

P_{AOX1} vHH P_{AOX1} *ADH2*	Time (h)	Volume (mL)		YDM (g/L)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) ($\mu g\ g^{-1}\ h^{-1}$)		Methanol concentration (g/L)	
Phase		R3	R4	R3	R4	R3	R4	R3	R4	R3	R4
1	24.0	318	318	24.9	24.3						
2	26.2									*9.8	*10.3
	49.1	501	502	102.6	100.6	696.1	783.0	310.3	359.4	13.9	14.4
3	68.7	518	518	94.9	94.6	1688.5	1724.9	376.1	362.0	9.4	9.7
	92.7	544	543	86.7	85.2	2466.5	2364.8	307.4	268.3	8.6	8.9

Table 33: Bioreactor cultivation process data and specific productivity (q_P) for vHH with the **P_{FLD1}*ADH2*** overexpression from Example 22. *Represents a control sample after the methanol pulse.

P_{AOX1} vHH P_{FLD1} *ADH2*	Time (h)	Volume (mL)		YDM (g/L)		Recombinant protein concentration (mg/L)		Specific productivity (q_P) ($\mu g\ g^{-1}\ h^{-1}$)		Methanol concentration (g/L)	
Phase		R7	R8	R7	R8	R7	R8	R7	R8	R7	R8
1	21.9	306	307	24.5	24.3						
2	23.8									*10.8	*10.7
	47.1	485	484	100.3	98.6	738.1	761.1	339.0	358.0	10.8	11.5
3	68.7	514	513	93.3	92.3	1511.7	1555.3	284.7	297.8	11.2	11.7
	91.7	516	514	85.8	85.2	2270.9	2364.4	285.7	305.6	10.5	10.5

Table 34: Comparison of specific productivity (q_P) of methanol utilization negative strain producing vHH from Example 11 to the **P_{AOX1}*ADH2*** overexpressing strain from Example 22.

Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH			Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{AOX1}*ADH2***			Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH		Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{AOX1}*ADH2***		Fold increase
	q_P(µg g^{-1} h^{-1})			q_P(µg g^{-1} h^{-1})			q_P(µg g^{-1} h^{-1})		q_P(µg g^{-1} h^{-1})	
Time (h)	R1	R2	Time (h)	R3	R4	Time (h)	Average	Time (h)	Average	
20.00			24.0			20.00	**Weighted average 195.8**	24.0		
28.22	137.8	141.5				28.22				
44.83	222.8	208.3				44.83				
53.58	176.0	246.2	49.1	310.3	359.4	**53.58**		**49.1**	**335**	**1.71**
68.83	142.5	72.2	68.7	376.1	362	**68.83**	**107.4**	**68.7**	**369**	**3.44**
92.00	55.8	97.5	92.7	307.4	268.3	**92.00**	**76.7**	**92.7**	**288**	**3.76**

Table 35: Comparison of specific productivity (q_P) of methanol utilization negative strain producing vHH from Example 11 to the **P_{FLD1}*ADH2*** overexpressing strain from Example 22.

Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH			Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{FLD1}*ADH2***			Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH		Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{FLD1}*ADH2***		Fold increase
	q_P(µg g^{-1} h^{-1})			q_P(µg g^{-1} h^{-1})			q_P(µg g^{-1} h^{-1})		q_P(µg g^{-1} h^{-1})	
Time (h)	R1	R2	Time (h)	R7	R8	Time (h)	Average	Time (h)	Average	
20.00			21.9			20.00	**Weighted average 195.8**	21.9		
28.22	137.8	141.5				28.22				
44.83	222.8	208.3				44.83				
53.58	176.0	246.2	47.1	339.0	358.0	**53.58**		**47.1**	**349**	**1.78**
68.83	142.5	72.2	68.7	284.7	297.8	**68.83**	**107.4**	**68.7**	**291**	**2.71**
92.00	55.8	97.5	91.7	285.7	305.6	**92.00**	**76.7**	**91.7**	**296**	**3.86**

Table 36: Comparison of volumetric productivity of methanol utilization negative strain producing vHH from Example 11 to the **P_{AOX1}*ADH2*** overexpressing strain from Example 22. * Corrected recombinant concentration (C_{cP}) for the biomass volume, $C_{cP}= C_p*(1-C_x*F_c)$, $F_c=0.0033$.

	Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH		Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{AOX1}*ADH2***		Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH	Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{AOX1}*ADH2***	**Fold increase**
	R1	R2	R3	R4	Average	Average	
Recombinant protein (mg/L)	1337.2	1374.0	2466.5	2364.8	1355.6	2415.7	
Biomass (g/L)	110.8	109.4	86.7	85.2	110.1	86.0	
*Corrected recombinant protein (mg/L)	848.3	878.0	1760.8	1699.9	863.1	1730.4	
Time (h)	92.0	92.0	92.7	92.7	92.0	92.7	
Volumetric productivity (mg L^{-1} h^{-1})					**9.38**	**18.67**	**1.99**

Table 37: Comparison of volumetric productivity of methanol utilization negative strain producing vHH from Example 11 to the **P_{FLD1}*ADH2*** overexpressing strain from Example 22. * Corrected recombinant concentration (C_{cP}) for the biomass volume, $C_{cP}= C_p*(1-C_x*F_c)$, $F_c=0.0033$.

	Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH		Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{FLD1}*ADH2***		Example 11 Δ*aox1*Δ*aox2* P_{AOX1}vHH	Example 22 Δ*aox1*Δ*aox2* P_{AOX1}vHH **P_{FLD1}*ADH2***	**Fold increase**
	R1	R2	R7	R8	Average	Average	
Recombinant protein (mg/L)	1337.2	1374.0	2270.9	2364.4	1355.6	2317.7	
Biomass (g/L)	110.8	109.4	85.8	85.2	110.1	85.5	
*Corrected recombinant protein (mg/L)	848.3	878.0	1627.9	1699.6	863.1	1663.8	
Time (h)	92.0	92.0	91.7	91.7	92.00	91.7	
Volumetric productivity (mg L^{-1} h^{-1})					**9.38**	**18.14**	**1.93**

Table 38: Methanol inducible promoters and their respective chromosomal positions in the strain *P. pastoris* CBS7435 (Gasser, Steiger, & Mattanovich, 2015)

P_{SHB17}	PP7435_chr2 (340616…341606)	(SEQ ID NO:36) GCAAGGCAACTGAGAAATTGAATAGTGGTTTCAAGCCCGCTGACTTTTT GTATTATCTCAATGTCGGTGTTTCACAGTCCCCAGAAGGGGGCTTTGCC TTCAAGGGAGACGGAAGAGACATCGTCAACCCTGGGGAGAAGTATTTCA AATGGCGCAAGTTCGCTAATTTTTACGATTAAGCAGTGCTGTATGGGGT AGTTAATAAATCGGGAATATCCTTCTGACGTGACTGTAACAAATCTCTT TTTACGTGGTGCGCATACTGGACAGAGGCAGAGTCTCAATTTCTTCTTT TGAGACAGGCTACTACAGCCTGTGATTCCTCTTGGTACTTGGATTTGCT TTTATCTGGCTCCGTTGGGAACTGTGCCTGGGTTTTGAAGTATCTTGTG GATGTGTTTCTAACACTTTTTCAATCTTCTTGGAGTGAGAATGCAGGAC TTTGAACATCGTCTAGCTCGTTGGTAGGTGAACCGTTTTACCTTGCATG TGGTTAGGAGTTTTCTGGAGTAACCAAGACCGTCTTATCATCGCCGTAA AATCGCTCTTACTGTCGCTAATAATCCCGCTGGAAGAGAAGTTCGAACA GAAGTAGCACGCAAAGCTCTTGTCAAATGAGAATTGTTAATCGTTTGAC AGGTCACACTCGTGGGCTATGTACGATCAACTTGCCGGCTGTTGCTGGA GAGATGACACCAGTTGTGGCATGGCCAATTGGTATTCAGCCGTACCACT GTATGGAAAATGAGATTATCTTGTTCTTGATCTAGTTTCTTGCCATTTT AGAGTTGCCACATTCGTAGGTTTCAGTACCAATAATGGTAACTTCCAAA CTTCCAACGCAGATACCAGAGATCTGCCGATCCTTCCCCAACAATAGGA GCTTACTACGCCATACATATAGCCTATCTATTTTCACTTTCGCGTGGGT GCTTCTATATAAACGGTTCCCCATCTTCCGTTTCATACTACTTGAATTT TAAGCACTAAA
P_{ALD4}	PP7435_chr2 (1466285…1467148)	(SEQ ID NO:37) CTTTTCTTTGGGCAAGGAAAAATCAAGAAAAAGCAGAGGTTAAAGTTTT CAGGGGAATGGCAATTGCTTTATATATGGGAGAAAGTTAACTACGTCGG TGCTGTAGGCGTAGAGAGCGACTGGAGAATGCGTGATGAGGTCGTCTCT TTTCGCCCCCCCTTGGCGGGGTAAAAATTGCACTACTGCAGAATTACTA CACCCCTATTCCGAGGAGACGGAGTGCGACAAAAATGGTAAAGTTCACC CTAGTCTGCGACTTTTAATTGACGGACACCGGCGTTTACATGCGAAAAA AACTAAAGTGCGCGCATTTCACGGCCGAGGGGGGTCCCACTTGGGACTG AGAGGGGGTGGGATCTGAAATCGAGGAGGTATCAAGACCCCCCGTTTCT CAACTCCCTAATCAAAAATTACGAAGTCCTCGTTGGAAAGGAGTTAAAA TAATTAAGCGGGGTCGGACGCCATACCGAGGTTATCTTGCAGGCATTTT ACTAATATTGGAATTCGGAGCTCAACTTGCAACCAGGCAGGGTTTAGCT ATGTAATCAATGTAATCAATATAATAAAGCACTACCACATCGAAGGTTT GGGAGGGAGGCCAATAGTGTCCCCCACAGGGTGCTGATATCGCGATTCT TGGGTGAGGAGACACATATTTCACTCCTCTCACCAACCAACCAAGCGGC TCCTCGCAAGATGATTTATCCGATTATCCGGACACTATACTCCCATCCA GTTTGATGCCGATTTCATCGATTGTCCTAAATAATCCTTAAATATGTAT AGAACGGTACCCTGGGGTTACATAATCCTTATTTAATAATCCCTCCCCC ACCGCTTTTCTTTTTTTTTCTTCTTATTGTC
P_{FDH1}	PP7435_chr3 (423504…424503)	(SEQ ID NO:38) AAATGGCAGAAGGATCAGCCTGGACGAAGCAACCAGTTCCAACTGCTAA GTAAAGAAGATGCTAGACGAAGGAGACTTCAGAGGTGAAAAGTTTGCAA GAAGAGAGCTGCGGGAAATAAATTTTCAATTTAAGGACTTGAGTGCGTC CATATTCGTGTACGTGTCCAACTGTTTTCCATTACCTAAGAAAAACATA AAGATTAAAAAGATAAACCCAATCGGGAAACTTTAGCGTGCCGTTTCGG ATTCCGAAAAACTTTTGGAGCGCCAGATGACTATGGAAAGAGGAGTGTA CCAAAATGGCAAGTCGGGGGCTACTCACCGGATAGCCAATACATTCTCT AGGAACCAGGGATGAATCCAGGTTTTTGTTGTCACGGTAGGTCAAGCAT TCACTTCTTAGGAATATCTCGTTGAAAGCTACTTGAAATCCCATTGGGT GCGGAACCAGCTTCTAATTAAATAGTTCGATGATGTTCTCTAAGTGGGA CTCTACGGCTCAAACTTCTACACAGCATCATCTTAGTAGTCCCTTCCCA AAACACCATTCTAGGTTTCGGAACGTAACGAAACAATGTTCCTCTCTTC ACATTGGGCCGTTACTCTAGCCTTCCGAAGAACCAATAAAAGGGACCGG CTGAAACGGGTGTGGAAACTCCTGTCCAGTTTATGGCAAAGGCTACAGA AATCCCAATCTTGTCGGGATGTTGCTCCTCCCAAACGCCATATTGTACT

		GCAGTTGGTGCGCATTTTAGGGAAAATTTACCCCAGATGTCCTGATTTT CGAGGGCTACCCCCAACTCCCTGTGCTTATACTTAGTCTAATTCTATTC AGTGTGCTGACCTACACGTAATGATGTCGTAACCCAGTTAAATGGCCGA AAAACTATTTAAGTAAGTTTATTTCTCCTCCAGATGAGACTCTCCTTCT TTTCTCCGCTAGTTATCAAACTATAAACCTATTTTACCTCAAATACCTC CAACATCACCCACTTAAACA
P_{DAS1}	PP7435_chr3 (634140…634688)	(SEQ ID NO:39) AATGATATAAACAACAATTGAGTGACAGGTCTACTTTGTTCTCAAAAGG CCATAACCATCTGTTTGCATCTCTTATCACCACACCATCCTCCTCATCT GGCCTTCAATTGTGGGGAACAACTAGCATCCCAACACCAGACTAACTCC ACCCAGATGAAACCAGTTGTCGCTTACCAGTCAATGAATGTTGAGCTAA CGTTCCTTGAAACTCGAATGATCCCAGCCTTGCTGCGTATCATCCCTCC GCTATTCCGCCGCTTGCTCCAACCATGTTTCCGCCTTTTTCGAACAAGT TCAAATACCTATCTTTGGCAGGACTTTTCCTCCTGCCTTTTTTAGCCTC AGGTCTCGGTTAGCCTCTAGGCAAATTCTGGTCTTCATACCTATATCAA CTTTTCATCAGATAGCCTTTGGGTTCAAAAAAGAACTAAAGCAGGATGC CTGATATATAAATCCCAGATGATCTGCTTTTGAAACTATTTTCAGTATC TTGATTCGTTTACTTACAAACAACTATTGTTGATTTTATCTGGAGAATA ATCGAACAAA
P_{DAS2}	PP7435_chr3 (632201…633200)	(SEQ ID NO:40) ATTACTGTTTTGGGCAATCCTGTTGATAAGACGCATTCTAGAGTTGTTT CATGAAAGGGTTACGGGTGTTGATTGGTTTGAGATATGCCAGAGGACAG ATCAATCTGTGGTTTGCTAAACTGGAAGTCTGGTAAGGACTCTAGCAAG TCCGTTACTCAAAAAGTCATACCAAGTAAGATTACGTAACACCTGGGCA TGACTTTCTAAGTTAGCAAGTCACCAAGAGGGTCCTATTTAACGTTTGG CGGTATCTGAAACACAAGACTTGCCTATCCCATAGTACATCATATTACC TGTCAAGCTATGCTACCCCACAGAAATACCCCAAAAGTTGAAGTGAAAA AATGAAAATTACTGGTAACTTCACCCCATAACAAACTTAATAATTTCTG TAGCCAATGAAAGTAAACCCCATTCAATGTTCCGAGATTTAGTATACTT GCCCCTATAAGAAACGAAGGATTTCAGCTTCCTTACCCCATGAACAGAA ATCTTCCATTTACCCCCCACTGGAGAGATCCGCCCAAACGAACAGATAA TAGAAAAAAGAAATTCGGACAAATAGAACACTTTCTCAGCCAATTAAAG TCATTCCATGCACTCCCTTTAGCTGCCGTTCCATCCCTTTGTTGAGCAA CACCATCGTTAGCCAGTACGAAAGAGGAAACTTAACCGATACCTTGGAG AAATCTAAGGCGCGAATGAGTTTAGCCTAGATATCCTTAGTGAAGGGTT GTTCCGATACTTCTCCACATTCAGTCATAGATGGGCAGCTTTGTTATCA TGAAGAGACGGAAACGGGCATTAAGGGTTAACCGCCAAATTATATAAAG ACAACATGTCCCCAGTTTAAAGTTTTTCTTTCCTATTCTTGTATCCTGA GTGACCGTTGTGTTTAATATAACAAGTTCGTTTTAACTTAAGACCAAAA CCAGTTACAACAAATTATAACCCCTCTAAACACTAAAGTTCACTCTTAT CAAACTATCAAACATCAAAA
P_{PMP20}	PP7435_ Chr1 (2418090…2419089)	(SEQ ID NO:41) GTCAACTGCGTACTCTTTTGTCGAATGGACTACTGAATCTGCCTCGATA GCCACTATAGGAAGGTCCATAGAGGCCAGTTTTTCAACTAGTCTTGGTG GAAAGAAACCGACAAAGCCTTTCATGGAGTCACCGATACTGAAAGGTTC AAACAAAGAATGCTTGGGTAGTCTCTTAATACCCATGGCAACGAAAAAG GGGTCTTCATTGTTCAACATGAATTCGTATCCACCTTTAATGTAGTCAT AAAGCTGCTGAAGTTCCGAATCAGTGATGGAACTGTCTACAGTGACAAT ATAGGAGTTCTCAATCACCTTATATCCAGTCGAATATATCTGGATAGGG TCGGGTCTCACTGTGGAAGATTCAAATGGGTTAGATCCCTGTAATTTCA GCGATGGAGACTCAGTATGATGGGGCAAGGAAAACGGCAATTGGATATT CAATTGGTCAAGAGATGGTATCAAAAGCGAGTGTGCCAGGGTAGCCACG GTAGCCACTGATGCTAATCTGATAATTTTCATTTCTGGAGTGTCAAAAC AGTAGTGATAAAAGGCTATGAAGGAGGTTGTCTAGGGGCTCGCGGAGGA AAGTGATTCAAACAGACCTGCCAAAAAGAGAAAAAAGAGGGAATCCCTG TTCTTTCCAATGGAAATGACGTAACTTTAACTTGAAAAATACCCCAACC AGAAGGGTTCAAACTCAACAAGGATTGCGTAATTCCTACAAGTAGCTTA GAGCTGGGGGAGAGACAACTGAAGGCAGCTTAACGATAACGCGGGGGGA TTGGTGCACGACTCGAAAGGAGGTATCTTAGTCTTGTAACCTCTTTTTT CCAGAGGCTATTCAAGATTCATAGGCGATATCGATGTGGAGAAGGGTGA ACAATATAAAAGGCTGGAGAGATGTCAATGAAGCAGCTGGATAGATTTC

		AAATTTTCTAGATTTCAGAGTAATCGCACAAAACGAAGGAATCCCACCA AGCAAAAAAAAAAATCTAAG
P_{FBA1-2}	PP7435_Chr1 (1162918…1163621)	(SEQ ID NO:42) AAATTAATCCATAAGATAAGGCAAATGTGCTTAAGTAATTGAAAACAGT GTTGTGATTATATAAGCATGGTATTTGAATAGAACTACTGGGGTTAACT TATCTAGTAGGATGGAAGTTGAGGGAGATCAAGATGCTTAAAGAAAAGG ATTGGCCAATATGAAAGCCATAATTAGCAATACTTATTTAATCAGATAA TTGTGGGGCATTGTGACTTGACTTTTACCAGGACTTCAAACCTCAACCA TTTAAACAGTTATAGAAGACGTACCGTCACTTTTGCTTTTAATGTGATC TAAATGTGATCACATGAACTCAAACTAAAATGATATCTTTTACTGGACA AAAATGTTATCCTGCAAACAGAAAGCTTTCTTCTATTCTAAGAAGAACA TTTACATTGGTGGGAAACCTGAAAACAGAAAATAAATACTCCCCAGTGA CCCTATGAGCAGGATTTTTGCATCCCTATTGTAGGCCTTTCAAACTCAC ACCTAATATTTCCCGCCACTCACACTATCAATGATCACTTCCCAGTTCT CTTCTTCCCCTATTCGTACCATGCAACCCTTACACGCCTTTTCCATTTC GGTTCGGATGCGACTTCCAGTCTGTGGGGTACGTAGCCTATTCTCTTAG CCGGTATTTAAACATACAAATTCACCCAAATTCTACCTTGATAAGGTAA TTGATTAATTTCATAAAT
P_{PMP47}	PP7435_Chr3 (2033196…2034195)	(SEQ ID NO:43) AGCTCAGATTGGAAATGATTTTTGATCCTACCAAGAAGCCTTTGATTTC CAGAATCTCCGCTAAGTAAGTAACCCCCGCAAACGCATGCATCCATGCA AACAAAATACTAACAATTTTAGCCCCGTTGTTGAGAAACCCAGAAAATT GAATGTTCAACCAATCCAGACGATCAATAAGAAAAAAGGCCCAAAGGCT ACTTCCAAACCTGCTGCCGCCAAACCTGCTCCTTCAAAAGCCGGTCCCA AGGGAGGTAAGAAGGTGAGAAAGCCAAAGAAGACAGTTGAAGAATTGGA TCAGGAAATGGCTGACTACTTTGAAAATAAGAATTAGCCCAACAAAATA TGTACAAGTATTATATAAATGAATCTACATGGTGTGTTTTATTTAGATC CTCCAAACCAAGGAAAGAAACTAAACTTATCTCCGGACTTACGAGTCAA ATAACTATCCGCAGTTCCTTGGAACTCAGACTTTCTTCCATAAGCGGTC ATATCATCTTTGGACTGTGGGAATCCTGGACGAATCTTTGAAATGTCAT AATCTTGCTCTCTATCTCCAAGCACAGCGTCCGGTAAATGCTGGTTCTT CTTTCTCAGATGAATCTTGGATTTAACAAATAAAGCCGTGCCTATGGCT AATGTACTCAAAAACAAAGTCTGCTTCCAGAATTTCGCAAACGATGGAA TGCCATTTCCTGTAAATGTACTCATTGAACCTATGTTTGATTAAAGTTG GTGTGAAGTCATCAAACGAGAGTAAAATCAGATACTCGTGCACCGGCCA AAATTGACTGAGCTAATCTCTGCAGGCTTGACATCCGAACACAACAAAT AGGCGACAAATCTTAACTATCTAATCGTAGGCTATGGTAGAACTTTGTG GGGGTAGAGGAAGACTACAACAGCAAGACAAAACAAAAGAGTCATAGTT TGACTCTCTGCTTTTTTCTTCTTTCTCTTCTTTTTCTTCCTCCATATTC GTTATTTATTTCGAACTGGA
P_{FLD}	PP7435_Chr3 (262519…263518)	(SEQ ID NO:44) CAGCCATTAATCTCACCTCAGTTTTTGAATCAGTAGAATTTTTAATGAA ACAAACGGTTGGTATATTATTTGATAGAGTTGCCAAATTTCCAAAGATA AATTTTTCATCAGGTAATATCCTGAATACCGTAACATAGTGACTATTGG AAGACACTGCTATCATATTATATTTCGGATAAAAATCCAAACCCCAGAC CGACCTCTTGAGTCTCAACTCCAAGTCAGCCGCAACTTTAATTATCCGT GGATTGGGAGCTAGTTTGGACAACGCATCAGTATAATATAACTTTACGG TTCCATTATCAGACGCTATTGCAAGAACTTCCTTTCCATTGATCTCGCC AATGCGGCAGTAATTGATATCGTAGGGTAGGTCTGGAAAGACGCTGGCG CTTGTGTCCCATTCTGCAGGAATCTCTGGCACGGTGCTAATGGTAGTTA TCCAACGGAGCTGAGGTAGTCGATATATCTGGATATGCCGCCTATAGGA TAAAAACAGGAGAGGGTGAACCTTGCTTATGGCTACTAGATTGTTCTTG TACTCTGAATTCTCATTATGGGAAACTAAACTAATCTCATCTGTGTGTT GCAGTACTATTGAATCGTTGTAGTATCTACCTGGAGGGCATTCCATGAA TTAGTGAGATAACAGAGTTGGGTAACTAGAGAGAATAATAGACGTATGC ATGATTACTACACAACGGATGTCGCACTCTTTCCTTAGTTAAAACTATC ATCCAATCACAAGATGCGGGCTGGAAAGACTTGCTCCCGAAGGATAATC TTCTGCTTCTATCTCCCTTCCTCATATGGTTTCGCAGGGCTCATGCCCC TTCTTCCTTCGAACTGCCCGATGAGGAAGTCCTTAGCCTATCAAAGAAT TCGGGACCATCATCGATTTTTAGAGCCTTACCTGATCGCAATCAGGATT

		TCACTACTCATATAAATACATCGCTCAAAGCTCCAACTTTGCTTGTTCA TACAATTCTTGATATTCACA
P_{FGH1}	PP7435_Chr3 (555587…556586)	(SEQ ID NO:45) TGGTTCCCTCTCGGTCCAATACCAAAAATATTATCACCATACAGGTCTC CCTTCGATACCAGTGCAAAGTTGAACCGTGGGATTACCTTGGAATCTAC AAAAATAGTGTCACTCACAAGTTTGTCATCAACCACGCTGCCGCTTGCA AAGGAGAACTGAACATGAAGGTTGTTAGGGTTTGTTATATTGGAATAAG TGGTGGATTTGTTGAAGGCGAACGCACCAAAGCTACATCCGTCCTGAGC ACACTGTGAATTTGTCACGGAATTGACCAAGAGGTCAGACGATCCTGTA TCCCATTGAGCCGTTATGCTTTGTGGGGGAAACCCTATTTCTATCGTAC TAAGAAAACCAATGGTGAACTCATATTCGGTATCAATGGCGACGATTCC AGCATAGCCTGTAGACAGTAACAACACTAGGGCAACAGCAACTAACATA TCTTCATTGATGAAACGTTGTGATCGGTGTGACTTTTATAGTAAAAGCT ACAACTGTTTGAAATACCAAGATATCATTGTGAATGGCTCAAAAGGGTA ATACATCTGAAAAACCTGAAGTGTGGAAAATTCCGATGGAGCCAACTCA TGATAACGCAGAAGTCCCATTTTGCCATCTTCTCTTGGTATGAAACGGT AGAAAATGATCCGAGTATGCCAATTGATACTCTTGATTCATGCCCTATA GTTTGCGTAGGGTTTAATTGATCTCCTGGTCTATCGATCTGGGACGCAA TGTAGACCCCATTAGTGGAAACACTGAAAGGGATCCAACACTCTAGGCG GACCCGCTCACAGTCATTTCAGGACAATCACCACAGGAATCAACTACTT CTCCCAGTCTTCCTTGCGTGAAGCTTCAAGCCTACAACATAACACTTCT TACTTAATCTTTGATTCTCGAATTGTTTACCCAATCTTGACAACTTAGC CTAAGCAATACTCTGGGGTTATATATAGCAATTGCTCTTCCTCGCTGTA GCGTTCATTCCATCTTTCTA
P_{TAL1-2}	PP7435_Chr2 (644082…645082)	(SEQ ID NO:46) GATATCGATCTACACTTAATAGTAGATGACGAGGCATCTCTCCAATAGG TACCATATCTGGTGTTTCTTGTAATTTAAGAATCTGTTGGTCTATGAAT GTAGATTTGTCATGAACAATGATATATGGGTCAGGAGGACAAGATGGTT TCTCTGAGTTGGGTTGTTGAGGTGCCTGGCAAGACTTCGGAGCGTTGAT ATCCCCAAGACTTGTAGTGACCGATAGTTGAAGCGTGTGTTTGCAGGAA CGGCACATCAATGCAACTTTCGTAACTTTGGAATTGAGAGTTGATGCAC TGATGACGATACCCGAAATTTTGACGATTTTACCAATATGACTTGAAGA CAAGTCTCTCATTGAAACCTTATTATCGTTACTAAGCAAAACGAGCTGA CAAGAAGGGAAGGTGGTCGGTATTTCCTCGTTGTTCAAATATATGATTC TCCTGGCAATATCTGTGATGGCCTGTTCAAAAAGTGGAATCATTTCTGC AGGATCATCTACCAACTTTTTATTGAGCTCCTCATTGAATACGATTAAG TGGTCATTTTGAATCGTCAGTAAGTACTTGTTTACAAGTAAATTCTGTC TGAGTTGTTCTCTGTAGATGTACTGATTTTCCATACGAAACTCCAAAAT GAACGAACGGAATGCCTTAATGACCTCACTGAACTGGTCATCGTTCTGT TCTCCGGGAAGGACACTTGTGTTAAAGACTGATGCTCTATCAAAGGACA TTGCAACAAAGTATAAACGGTTGTGAGCGGGAAAAAGATGTGTAGGTAA TTGTCGTAGATGAGACTGATTCAGTAGAAAACGCGTCCTGCACTATTTT TTTCTTTCTTCATTACATTTCCTAATCGGGACAAAATGAATCTAAAGAC GTGGTTATGTAGTACACGCATCGATAGGCTATCCCCATACCAAAACACT ATTTTACCCCATCCTTGACAGGTTATAAATATGCGATAGTATGAGTATC TTCAAATTCAGCTGAAATATC
P_{DAS2}	PP7435_Chr3 (633689-634688)	SEQ ID NO:47) AATAAAAAAAACGTTATAGAAAGAAATTGGACTACGATATGCTCCAATCC AAATTGTCAAAATTGACCACCGAAAAAGAACAATTGGAATTTGACAAGA GGAACAACTCACTAGATTCTCAAACGGAGCGTCACCTAGAGTCAGTTTC CAAGTCAATTACAGAAAGTTTGGAAACAGAAGAGGAGTATCTACAATTG AATTCCAAACTTAAAGTCGAGCTGTCCGAATTCATGTCGCTAAGGCTTT CTTACTTGGACCCCATTTTTGAAAGTTTCATTAAAGTTCAGTCAAAAAT TTTCATGGACATTTATGACACATTAAAGAGCGGACTACCTTATGTTGAT TCTCTATCCAAAGAGGATTATCAGTCCAAGATCTTGGACTCTAGAATAG ATAACATTCTGTCGAAAATGGAAGCGCTGAACCTTCAAGCTTACATTGA TGATTAGAGCAATGATATAAACAACAATTGAGTGACAGGTCTACTTTGT TCTCAAAAGGCCATAACCATCTGTTTGCATCTCTTATCACCACACCATC CTCCTCATCTGGCCTTCAATTGTGGGGAACAACTAGCATCCCAACACCA GACTAACTCCACCCAGATGAAACCAGTTGTCGCTTACCAGTCAATGAAT GTTGAGCTAACGTTCCTTGAAACTCGAATGATCCCAGCCTTGCTGCGTA

		TCATCCCTCCGCTATTCCGCCGCTTGCTCCAACCATGTTTCCGCCTTTT TCGAACAAGTTCAAATACCTATCTTTGGCAGGACTTTTCCTCCTGCCTT TTTTAGCCTCAGGTCTCGGTTAGCCTCTAGGCAAATTCTGGTCTTCATA CCTATATCAACTTTTCATCAGATAGCCTTTGGGTTCAAAAAAGAACTAA AGCAGGATGCCTGATATATAAATCCCAGATGATCTGCTTTTGAAACTAT TTTCAGTATCTTGATTCGTTTACTTACAAACAACTATTGTTGATTTTAT CTGGAGAATAATCGAACAAA
P_{CAM1}	PP7435_Chr3 (178828-179827)	SEQ ID NO:48) ATTGTTGTGAATACTCTCCTTCATTTGGATTTCTTGGACTTCGGACTCT CTTGATCTCTCTTCGAAAGTTTTAACTCTGTTCATGTATAATTTTACCC GCTGTAGGTCGCTCATAATACCATGAGTATGCACATCTTTTACTCCATT AACTTTCAGGTATGCAAAATACAATGAAGATAGTATATAGCTCAAAGAA TTTAGCATTTTGCATTGATCTAATTGTGACATTTTCTCTATGATATCAT CTAGCTTCTTAAACTCGAGAATCTCGTCCAACGAGGCAGAAACATTGTC CAGTCTTACGTCAAGATTATTCACGAGTTTCTGGACCGTATCAACGTTT TCCATCTTAAGATTACAGTAAGTATCGTCCTTTTGAACTGCAAAGGTAG AAAAGTTAATTTTTGATTTGGTAGTACACTATGAAACTTGCTCACCCCA ATCTTTCCTCCTGACAGGTTGATCTTTATCCCTCTACTAAATTGCCCCA AGTGTATCAAGTAGACTAGATCTCGCGAAAGAACAGCCTAATAAACTCC GAAGCATGATGGCCTCTATCCGGAAAACGTTAAGAGATGTGGCAACAGG AGGGCACATAGAATTTTTAAAGACGCTGAAGAATGCTATCATAGTCCGT AAAAATGTGATAGTACTTTGTTTAGTGCGTACGCCACTTATTCGGGGCC AATAGCTAAACCCAGGTTTGCTGGCAGCAAATTCAACTGTAGATTGAAT CTCTCTAACAATAATGGTGTTCAATCCCCTGGCTGGTCACGGGGAGGAC TATCTTGCGTGATCCGCTTGGAAAATGTTGTGTATCCCTTTCTCAATTG CGGAAAGCATCTGCTACTTCCCATAGGCACCAGTTACCCAATTGATATT TCCAAAAAAGATTACCATATGTTCATCTAGAAGTATAAATACAAGTGGA CATTCAATGAATATTTCATTCAATTAGTCATTGACACTTTCATCAACTT ACTACGTCTTATTCAACAAT
P PP7435_Chr1-0336	PP7435_Chr1 (615194-615193)	SEQ ID NO:49) TATACGGTCTATCCACTTTGGAAACGATGTAGTTGAAACGGGGAAGTAA TAGTGGTTCCCAAACGACATGAAGAGGTTATATAAGTTTGCAAGAGGGT GACACCATTTTAGTTGTGGTTCCCGGGTATTTTTTTAATCTTTTTAGTC TAAGATAGCCTCCCCAGATATTACCGAGTTGGGCCATTTGGGGCGGTAT CGGTGGTATCTGATGGTAGCGCGTTTTTACATGCCTGTGCATTGAACTG GCAAAGAGTATACTATCGTGGGGCCCTGAAGGAGGCAGCAAATGGACCG TCAATTGGTTGATCAGGGACTCAAGACAGGTATTGAGCTTTTCAAACAA AAAGAGTATAGGCGCTGCTACAAGGCATTTACTTCTACTATCAATTTCA TTGAGAATGATCCCGAGTTGGCCGCCAGCTGTGTATCTCAACTGATATC TCTGTTAGATTGTAGGGCAGCCTGTTTGGAAAAGCTAGATCAATTGAAT ATGGCCTTGAAAGATGGTCTTAAAATGATCAAGAGAGAGTGCCACAACT GCAAGGGTTATTTGAGAACTTGCAAAATTTTAGACCTACAAGGGAAGAT CAGTGAGGCTTTGTCTACAGCAAGAGAAGGGATCTCCATAATAGAAACT AGAAGAGATCAGGATAATCAATTTAGATATTCCAAGGTTCTTTTGGAAC AATTAAAGGAACTGAAAAATGCACTGAAAATCAAATTGGACAAGAAAAA TCAGCTACACTTCAAAGTTTTAAAGTTTGACGCACCAGTGCCTTGTACA AAGAAACTAAGATTAGTCACTCCAAGAACAATAGATCCTTCCATTTTTT TGCCGATAGAGCTAGTGAAGCTGATCTTTCGCCTGTTGAATTTCTCAGA CATGTATGCCTGTTTATTGGTCTCAACAAAATGGAACTCAATTATATCC TCATCACCGGAACTGTTTCGAAAACTTCAGTTGAAATCCCAACTGTCCA ACAAGGCGTTAAACAATTGT

References:

Ata, Ö., Prielhofer, R., Gasser, B., Mattanovich, D., & Çalık, P. (2017). Transcriptional engineering of the glyceraldehyde-3-phosphate dehydrogenase promoter for improved heterologous protein production in Pichia pastoris: Transcriptional Engineering of the P GAP for Gene Overexpression in P. pastoris. Biotechnology and Bioengineering, 114(10), 2319–2327. https://doi.org/10.1002/bit.26363

Aw, R., & Polizzi, K. M. (2013). Can too many copies spoil the broth? Microbial Cell Factories, 12, 128. https://doi.org/10.1186/1475-2859-12-128

Blumhoff, M., Steiger, M. G., Marx, H., Mattanovich, D., & Sauer, M. (2013). Six novel constitutive promoters for metabolic engineering of Aspergillus niger. Applied Microbiology and Biotechnology, 97(1), 259–267. https://doi.org/10.1007/s00253-012-4207-9

Gasser, B., Prielhofer, R., Marx, H., Maurer, M., Nocon, J., Steiger, M., … Mattanovich, D. (2013). Pichia pastoris : protein production host and model organism for biomedical research. Future Microbiology, 8(2), 191–208. https://doi.org/10.2217/fmb.12.133

Gasser, B., Steiger, M. G., & Mattanovich, D. (2015). Methanol regulated yeast promoters: production vehicles and toolbox for synthetic biology. Microbial Cell Factories, 14(1). https://doi.org/10.1186/s12934-015-0387-1

Hohenblum, H., Borth, N., & Mattanovich, D. (2003). Assessing viability and cell-associated product of recombinant protein producing Pichia pastoris with flow cytometry. Journal of Biotechnology, 102(3), 281–290. https://doi.org/10.1016/S0168-1656(03)00049-X

Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., … Kovar, K. (2015). Cultivation strategies to enhance productivity of Pichia pastoris: A review. Biotechnology Advances, 33(6), 1177–1193. https://doi.org/10.1016/j.biotechadv.2015.05.008

Marx, H., Mattanovich, D., & Sauer, M. (2008). Overexpression of the riboflavin biosynthetic pathway in Pichia pastoris. Microbial Cell Factories, 7(1), 23. https://doi.org/10.1186/1475-2859-7-23

Mellitzer, A., Ruth, C., Gustafsson, C., Welch, M., Birner-Grünberger, R., Weis, R., … Glieder, A. (2014). Synergistic modular promoter and gene optimization to push cellulase secretion by Pichia pastoris beyond existing benchmarks. Journal of Biotechnology, 191, 187–195. https://doi.org/10.1016/j.jbiotec.2014.08.035

Nocon, J., Steiger, M. G., Pfeffer, M., Sohn, S. B., Kim, T. Y., Maurer, M., … Mattanovich, D. (2014). Model based engineering of Pichia pastoris central metabolism enhances recombinant protein production. Metabolic Engineering, 24, 129–138. https://doi.org/10.1016/j.ymben.2014.05.011

Prielhofer, R., Barrero, J. J., Steuer, S., Gassler, T., Zahrl, R., Baumann, K., … Marx, H. (2017). GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast Pichia pastoris. BMC Systems Biology, 11(1). https://doi.org/10.1186/s12918-017-0492-3

Prielhofer, R., Maurer, M., Klein, J., Wenger, J., Kiziak, C., Gasser, B., & Mattanovich, D. (2013). Induction without methanol: novel regulated promoters enable high-level expression in Pichia pastoris. Microbial Cell Factories, 12(1), 5. https://doi.org/10.1186/1475-2859-12-5

Schwarzhans, J.-P., Wibberg, D., Winkler, A., Luttermann, T., Kalinowski, J., & Friehs, K. (2016). Integration event induced changes in recombinant protein productivity in Pichia pastoris discovered by whole genome sequencing and derived vector optimization. Microbial Cell Factories, 15, 84. https://doi.org/10.1186/s12934-016-0486-7

Stadlmayr, G., Mecklenbräuker, A., Rothmüller, M., Maurer, M., Sauer, M., Mattanovich, D., & Gasser, B. (2010). Identification and characterisation of novel Pichia pastoris promoters for heterologous protein production. Journal of Biotechnology, 150(4), 519–529. https://doi.org/10.1016/j.jbiotec.2010.09.957

CLAIMS

1. A recombinant methanol utilization pathway deficient methylotrophic yeast (Mut-) host cell which is engineered:

a) by one or more genetic modifications to render the host cell methanol utilization pathway deficient, by reducing expression of a first and a second endogenous gene compared to the host cell prior to said one or more genetic modifications, wherein

i. the first endogenous gene encodes alcohol oxidase 1 (AOX1) comprising the amino acid sequence identified as SEQ ID NO:1 or a homologue thereof that is endogenous to the respective wild-type host cell, and

ii. the second endogenous gene encodes alcohol oxidase 2 (AOX2) comprising the amino acid sequence identified as SEQ ID NO:3 or a homologue thereof that is endogenous to the respective wild-type host cell, and

b) by one or more genetic modifications to increase expression of an alcohol dehydrogenase (*ADH2*) gene compared to the host cell prior to said one or more genetic modifications, wherein the *ADH2* gene encodes an alcohol dehydrogenase (ADH2).

2. The Mut- host cell of claim 1, wherein said one or more genetic modifications comprise a disruption, substitution, deletion, knockin or knockout of (i) one or more polynucleotides, or a part thereof; or (ii) an expression control sequence.

3. The Mut- host cell of claim 2, wherein said expression control sequence is selected from the group consisting of a promoter, a ribosomal binding site, transcriptional or translational start and stop sequences, an enhancer and activator sequence.

4. The Mut- host cell of any one of claims 1 to 3, wherein said first and/or second endogenous gene is knocked out by said one or more genetic modifications.

5. The Mut- host cell of any one of claims 1 to 4, wherein the *ADH2* gene is endogenous or heterologous to the Mut- host cell.

6. The Mut- host cell of any one of claims 1 to 5, wherein the ADH2 is any one of:

a) an ADH2, which is *P. pastoris* ADH2 comprising the amino acid sequence identified as SEQ ID NO:50, or a homologue thereof that is endogenous to a yeast species;

b) a mutant of the ADH2 of a), which is at least 60% identical to SEQ ID NO:50.

7. The Mut- host cell of any one of claims 1 to 6, wherein said one or more genetic modifications include a gain-of-function alteration in the *ADH2* gene resulting in increasing the level or activity of ADH2.

8. The Mut- host cell of claim 7, wherein said gain-of-function alteration includes a knockin of the *ADH2* gene.

9. The Mut- host cell of claim 7 or 8, wherein said gain-of-function alteration up-regulates the *ADH2* gene expression in said cell.

10. The Mut- host cell of any one of claims 7 to 9, wherein said gain-of-function alteration includes an insertion of a heterologous expression cassette to overexpress the *ADH2* gene in said cell.

11. The Mut- host cell of claim 10, wherein said heterologous expression cassette comprises a heterologous polynucleotide comprising an *ADH2* gene under the control of a promoter sequence.

12. The Mut- host cell of any one of claims 1 to 11, which comprises a heterologous gene of interest expression cassette (GOIEC) comprising an expression cassette promoter (ECP) operably linked to a gene of interest (GOI) encoding a protein of interest (POI).

13. The Mut- host cell of claim 12, wherein the ECP is a methanol-inducible promoter.

14. The Mut- host cell of claim 13, wherein the ECP is any one of the following:

a) a pAOX1 promoter comprising at least 90% sequence identity to SEQ ID NO:5;

b) a pAOX2 promoter comprising at least 90% sequence identity to SEQ ID NO:6; or

c) a promoter comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:36-49.

15. The Mut- host cell of any one of claims 12 to 14, wherein the GOIEC further comprises a nucleotide sequence encoding a signal peptide enabling the secretion of the POI.

16. The Mut- host cell of claim 15, wherein the nucleotide sequence encoding the signal peptide is fused adjacent to the 5'-end of the GOI.

17. The Mut- host cell of any one of claims 12 to 16, wherein the POI is heterologous to the Mut- host cell or the ECP.

18. The Mut- host cell of any one of claims 12 to 17, wherein the POI is a peptide or protein selected from the group consisting of an antigen-binding protein, a therapeutic protein, an enzyme, a peptide, a protein antibiotic, a toxin fusion protein, a carbohydrate - protein conjugate, a structural protein, a regulatory protein, a vaccine antigen, a growth factor, a hormone, a cytokine, a process enzyme.

19. The Mut- host cell of any one of claims 1 to 18, wherein the Mut- host cell is a yeast cell of the genus Pichia, Komagataella, Hansenula, Ogataea or Candida.

20. A method of producing a protein of interest (POI) comprising culturing a Mut- host cell of any one of claims 1 to 19 using methanol as a carbon source to produce the POI.

21. The method of claim 20, wherein a fermentation product is isolated from the cell culture, which fermentation product comprises the POI or a host cell metabolite obtained from the Mut- host cell.

22. The method of claim 20 or 21, wherein

a) a growing phase, during which the Mut- host cell is cultured using a basal carbon source as a source of energy; is followed by

b) a production phase, during which the Mut- host cell is cultured using a methanol feed thereby producing the POI.

23. The method of claim 22, wherein an average methanol concentration of 0.5-2.0% (v/v) is used in the host cell culture during the production phase of at least 24 hours.

24. The method of claim 22 or 23, wherein the methanol feed is at an average feed rate of at least 2 mg methanol/ (g dry biomass*h) during the production phase of at least 24 hours.

25. The method of any one of claims 22 to 24, wherein the Mut- host cell is cultured during the production phase under conditions limiting the host cell growth to less than 10% (w/w biomass).

26. Use of a recombinant methanol utilization pathway deficient methylotrophic yeast (Mut-) host cell in a method of producing a fermentation product which method comprises culturing said Mut- host cell under conditions that permit the Mut- host cell to use methanol as a substrate for alcohol dehydrogenase (ADH2), and to produce the fermentation product, wherein the method is of any one of claims 20 to 25.

27. Use of a recombinant methanol utilization pathway deficient methylotrophic yeast (Mut-) host cell in a method of producing a fermentation product which method comprises culturing said Mut- host cell under conditions that permit the Mut- host cell to produce the fermentation product using methanol as a carbon source, which Mut- host cell is engineered by one or more genetic modifications

a) to render the host cell methanol utilization pathway deficient, by reducing expression of a first and a second endogenous gene compared to the host cell prior to said one or more genetic modifications, wherein

i. the first endogenous gene encodes alcohol oxidase 1 (AOX1) comprising the amino acid sequence identified as SEQ ID NO:1 or a homologue thereof that is endogenous to the respective wild-type host cell, and

ii. the second endogenous gene encodes alcohol oxidase 2 (AOX2) comprising the amino acid sequence identified as SEQ ID NO:3 or a homologue thereof that is endogenous to the respective wild-type host cell, and

b) to increase expression of an alcohol dehydrogenase (*ADH2*) gene, wherein the *ADH2* gene encodes an alcohol dehydrogenase (ADH2).