

Method for reducing misincorporation of non-canonical branched-chain amino acids

Description

The present invention relates to a method for producing a recombinant polypeptide of interest in a microbial host cell, comprising (a) introducing a polynucleotide encoding the polypeptide of interest into a microbial host cell which has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19) is modulated (such as increased) in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell, and (b) expressing said polypeptide of interest in said microbial host cell. Moreover, the present invention relates to a method for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest expressed in a microbial host cell. Further envisaged by the present invention is a microbial host cell comprising (a) a recombinant polynucleotide encoding a polypeptide of interest, and (b) a recombinant polynucleotide encoding a polypeptide having an enzymatic activity.

BACKGROUND OF THE INVENTION

Expression of recombinant proteins in microbial hosts such as *Escherichia coli* (*E. coli*) has become a standard technique for the manufacturing of recombinant therapeutic proteins such as insulin. Today, a variety of biopharmaceutical products produced by microbial cells such as *E. coli* are on the market.

However, overexpression of recombinant proteins is known to cause stress to the microbial host, often resulting in misfolding of proteins or incorporation of non-wanted amino acids. Altered proteins may display unwanted properties such as modified biological activity, increased sensitivity to proteolysis or immunogenicity. In particular, the misincorporation of amino acids into recombinant proteins is a problem, since pharmaceutical products such as recombinant insulin have to be homogenous and pure. As a result, recombinant proteins comprising misincorporated amino acids have to be removed from the product.

Branched-chain amino acids (BCAAs) are amino acids with a non-linear molecular structure. BCAAs have aliphatic side-chains with a branch, i.e. a central carbon atom is bound to three

or more carbon atoms. Leucine, isoleucine and valine are so called proteinogenic or canonical BCAA. Besides the canonical BCAAs, non-canonical or non-proteinogenic BCAAs also exist. Such non-canonical BCAAs (ncBCAAs) are norleucine, norvaline, homoisoleucine and β -methylnorleucine.

Non-canonical branched chain amino acids (ncBCAA) such as norleucine and norvaline have been reported to be respectively misincorporated in place of leucine and methionine in recombinant proteins expressed in *E. coli* (Apostol I. *et al.*, 1997, Incorporation of norvaline at leucine positions in recombinant human hemoglobin expressed in *Escherichia coli*. Journal of Biological Chemistry 272.46: 28980-28988; Tsai *et al.*, (1988), Control of misincorporation of de novo synthesized norleucine into recombinant interleukin-2 in *E. coli*. Biochemical and biophysical research communications 156(2):733-739). Synthesis and accumulation of ncBCAA is a result of the low specificity of the *leu* and *ilv*-operon-coded enzymes involved in the BCAA biosynthetic pathway for α -ketoacids. Several studies have shown that the enzymes of the leucine biosynthetic pathway, encoded by the *leuABCD* operon, are crucial for the production of non-canonical branched chain amino acids such as norleucine. In their canonical mode, the enzymes of the leucine pathway convert α -ketoisovalerate to α -ketoisocaproate; meaning that they cause a one carbon addition to a five carbon α -ketoacid. Moreover, the leucine biosynthetic pathway was also shown to display quite broad substrate specificity and its enzymes can act on a variety of α -ketoacids. For example, enzymes of the leucine pathway can also convert α -ketovalerate to α -ketocaproate, which is the precursor of norleucine.

Non-canonical branched chain amino acids (ncBCAA) misincorporation into nascent recombinant proteins happens due to promiscuity of amino-acyl tRNA synthetases (aaRS). The fidelity of protein synthesis counts on the aptitude of aaRSs to charge the appropriate canonical amino acid onto its corresponding tRNA (e.g. reviewed by Reitz *et al.*, 2018, Synthesis of non-canonical branched-chain amino acids in *Escherichia coli* and approaches to avoid their incorporation into recombinant proteins, Curr Opin Biotechnol. Oct;53:248-253). Such fidelity can be jeopardized by a number of non-canonical amino acids, particularly ncBCAA, which are structurally similar to their canonical equivalents (Martinis SA, Fox GE. Non-standard amino acid recognition by *Escherichia coli* leucyl-tRNA synthetase. Nucleic Acids Symp Ser. 1997;36:125–128). For instance, leucyl-tRNA synthetase (leuRS) must distinguish between leucine and the non-canonical counterpart norvaline, which only differ by a single methyl group (Apostol, I., *et al.*, 1997, Incorporation of norvaline at leucine positions in recombinant human hemoglobin expressed in *Escherichia coli*. Journal of Biological Chemistry, 272(46), 28980-28988.). The same applies for methionyl-tRNA synthetase (metRS), which must discriminate between methionine and norleucine (Kiick, K. L., *et al.*, 2001, Identification of an expanded set of translationally active methionine analogues in *Escherichia coli*. FEBS Letters, 502(1-2), 25-30.), and isoleucyl-tRNA synthetase (ileRS),

which must differentiate between isoleucine and β -methylnorleucine (Muramatsu, R., *et al.*, 2003, Finding of an isoleucine derivative of a recombinant protein for pharmaceutical use. Journal of pharmaceutical and biomedical analysis, 31(5), 979-987.). For example, when *E. coli* is grown on a mineral salt medium and methionine is limiting, norleucine can undergo acylation onto a methionyl transfer RNA and subsequently become incorporated into the recombinant protein whenever methionine codons are translated.

Although different expression systems for recombinant protein production are available and adaption of culture conditions for improved recombinant protein production have been reported, the problem of a potential overflow metabolism and misincorporation of unwanted amino acids into the recombinant protein of interest is still unsolved.

The conditions under which misincorporation of non-canonical branched chain amino acids into heterologous recombinant proteins occur and how to efficiently prevent such misincorporations are not completely understood. A method for the production of recombinant polypeptides without misincorporation of non-canonical branched-chain amino acids would therefore be highly desired.

The technical problem underlying the present invention can be seen as the provision of means and methods for complying with the aforementioned needs. The technical problem is solved by the embodiments characterized in the claims and herein.

SUMMARY OF THE INVENTION

The present invention relates to a method for producing a recombinant polypeptide of interest in a microbial host cell, comprising the steps of

- (a) introducing a polynucleotide encoding the polypeptide of interest into a microbial host cell which has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19) is modulated (such as increased) in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell, and
- (b) expressing said polypeptide of interest in said microbial host cell.

The present invention further relates to a method for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest expressed in a microbial host cell, said method comprising

- (a) modulating (such as increasing) an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19) in the microbial host cell,
- (b) introducing a polynucleotide encoding the polypeptide of interest into said microbial host cell, and
- (c) expressing said polypeptide of interest in said microbial host cell.

Further envisaged by the present invention is a microbial host cell comprising:

- (a) a recombinant polynucleotide encoding a polypeptide of interest, and
- (b) a recombinant polynucleotide encoding a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19).

Further envisaged by the present invention is a bioreactor comprising the microbial host cell of the present invention. In some embodiments, the bioreactor has a volume of at least 10 L.

Moreover, the present invention relates to the use of a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19), or of a polynucleotide encoding said polypeptide for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest produced in a microbial host cell.

Finally, the present invention concerns the use of a microbial host cell for producing a recombinant polypeptide of interest, wherein the microbial host cell has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19) is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell.

In an embodiment of the method, the use, or the microbial host cell of the present invention, the microbial host cell does not express an endogenous polypeptide having said enzymatic activity. Accordingly, the endogenous, i.e. naturally occurring polynucleotide, encoding for the polypeptide having said enzymatic activity has been deleted, i.e. knocked out in the microbial host cell.

In an embodiment of the method, the use, or the microbial host cell of the present invention, the polypeptide of interest is a therapeutic polypeptide such as a proinsulin, an insulin or an insulin analogue.

In an embodiment of the method, the use, or the microbial host cell of the present invention, the polynucleotide encoding the polypeptide of interest and/or the polynucleotide encoding the polypeptide having said enzymatic activity is operably linked to an inducible promoter.

In an embodiment of the method, the use, or the microbial host cell of the present invention, the microbial host cell is an *Escherichia coli* cell.

DETAILED DESCRIPTION OF THE INVENTION

As set forth above, the present invention relates to a method for producing a recombinant polypeptide of interest in a microbial host cell. The method comprises (a) the introduction of a polynucleotide encoding the polypeptide of interest into a microbial host cell which has been modified such that an enzymatic activity as set forth elsewhere herein is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell, and (b) expressing said polypeptide of interest in said microbial host cell.

In accordance with the present invention, a recombinant polypeptide of interest shall be produced in a microbial host cell. The term "recombinant polypeptide" as used herein refers to a genetically engineered polypeptide. Accordingly, the polypeptide to be produced shall be heterologous with respect to the microbial host cell which means that the host cell does not naturally express the polypeptide of interest. The term "heterologous", thus, means that the polynucleotide/polypeptide does not occur naturally in the microbial host cell.

In a preferred embodiment of the present invention, the recombinant polypeptide of interest to be produced is a therapeutic polypeptide.

In particular, it is envisaged that the recombinant polypeptide of interest is an antibody or antigen-binding fragment thereof, an enzyme, a receptor, a secreted protein, a fusion protein,

or a hormone, in particular a peptide hormone (such as insulin, or a precursor thereof such as a proinsulin).

In a preferred embodiment, the recombinant polypeptide of interest is an antibody or antigen-binding fragment thereof. The antibody is, preferably, selected from a multispecific antibody, a human antibody, a humanized antibody, a chimeric antibody, and a single-chain antibody. Preferably, the antigen-binding fragment of the antibody is selected from the group consisting of a Fab fragment, a Fab' fragment, a F(ab')₂ fragment, a scFv fragment, and a Fv fragment. For example, the antigen-binding fragment is a F(ab')₂ fragment.

In another preferred embodiment, the recombinant polypeptide of interest is proinsulin or insulin. Insulin is a peptide hormone which is naturally secreted by the islets of Langerhans and functions in the regulation of the metabolism of carbohydrates and fats, particularly the conversion of glucose to glycogen. The insulin may be a naturally occurring insulin, in particular human insulin, or an analog of a naturally occurring insulin, in particular an analog of human insulin. Accordingly, the term "insulin" encompasses naturally occurring insulins and analogs thereof. An insulin analog is an altered form of insulin, different from any occurring in nature, but still available to the human body for performing the same action as human insulin in terms of glycemic control. Preferably, the insulin analog is selected from insulin lispro, insulin aspart, insulin glulisine, insulin detemir, and insulin glargine.

Other recombinant polypeptides of interest are hirudin, somatotropin, an interleukin such as interleukin-2, or hemoglobin.

It is to be understood that the recombinant polypeptide of interest shall comprise at least one isoleucine residue, at least one leucine residue, and/or at least one methionine residue.

In an embodiment, the recombinant polypeptide of interest is a leucine-rich polypeptide. For example, at least 5%, at least 10%, or at least 15% or all amino acids of the polypeptide are leucine residues. For example, the recombinant polypeptide of interest may have a length of 96 amino acids and may comprise 14 leucine residues. Alternatively, the recombinant polypeptide of interest may have a length of 44 amino acids and may comprise 8 leucine residues. Alternatively, the recombinant polypeptide of interest may have a length of 52 amino acids and may comprise 6 leucine residues.

In another embodiment, the recombinant polypeptide of interest is an isoleucine-rich polypeptide. For example, at least 5%, at least 10%, or at least 15% or all amino acids of the polypeptide are isoleucine residues.

In another embodiment, the recombinant polypeptide of interest is a methionine-rich polypeptide. For example, at least 5%, at least 10%, or at least 15% or all amino acids of the polypeptide are methionine residues.

In particular, the analog of insulin is insulin glargine. Insulin glargine is a long-acting basal insulin analogue. Insulin glargine is produced by recombinant DNA technology using a non-pathogenic laboratory strain of *Escherichia coli* (K12) as the production organism. It is an analogue of human insulin made by replacing the asparagine residue at position A21 of the A-chain with glycine and adding two arginine residues to the C-terminus (positions B31 and 32) of the B-chain. The resulting protein is soluble at pH 4 and forms microprecipitates at physiological pH 7.4. Small amounts of insulin glargine are slowly released from microprecipitates giving the drug a long duration of action (up to 24 hours) and no pronounced peak concentration. It is marketed under the name Lantus®. The CAS Registry Number for insulin glargine is 160337-95-1.

The recombinant polypeptide of interest is preferably not a polypeptide that is expressed in the organism in the untransformed state. Preferably, it is also not a polypeptide which is expressed as a selection marker (i.e. the polypeptide shall not confer a resistance such as a resistance to an antibiotic to the cell). Further, it is preferably not a polypeptide which is expressed as reporter polypeptide (such as a fluorescent polypeptide or the GUS-polypeptide).

In an embodiment of the method of the present invention, the recombinant polypeptide of interest and the polypeptide having an enzymatic activity as referred to herein are co-expressed. Thus, the recombinant polypeptide of interest is not the polypeptide having an enzymatic activity as referred to herein in connection with the method of the present invention. Preferably, the recombinant polynucleotide encoding the polypeptide of interest and the polynucleotide encoding the polypeptide having an enzymatic activity as referred to herein are not present on the same plasmid. Accordingly, it is envisaged that they are expressed from different plasmids. Thus, the polypeptides are comprised by different DNA molecules.

In accordance with the present invention, the recombinant polypeptide of interest is produced by introduction of a recombinant polynucleotide encoding the polypeptide of interest into the microbial host cell as defined herein and by expressing said recombinant polypeptide of interest in said microbial host cell.

Preferably, the polynucleotide encoding the polypeptide of interest is comprised by an expression plasmid, i.e. a plasmid which allows for the expression of said polynucleotide. The same applies to the polynucleotide encoding for a polypeptide having enzymatic activity as defined elsewhere herein. In order to express the polynucleotide, said polynucleotide is

operably linked to a promoter. The term “operably linked” as used herein refers to a functional linkage between the promoter and the polynucleotide to be expressed, such that the promoter is able to initiate transcription of said polynucleotide. Preferred promoters are described elsewhere herein.

The expression plasmid may comprise further elements. Preferably, the plasmid additionally comprises a selectable marker gene which allows for the selection of microbial host cells. For example, this can be a gene which confers resistance to an antibiotic such as ampicillin, chloramphenicol, kanamycin, or tetracycline.

Furthermore, it is envisaged that the plasmid comprises a replicon which allows for the replication of the plasmid in the microbial host cell. In a preferred embodiment, only one copy of the plasmid comprising the polynucleotide encoding the polypeptide having an enzymatic activity as set forth herein in connection with the method of the present invention is present in the microbial host cell. Thus, the copy number of said plasmid in the microbial host cell shall be one. This may be achieved by the presence of *ori2* and its elements *repE*, *sopA*, *sopB*, and *sopC* which ensures 1 copy of the plasmid per cell. In contrast, the polynucleotide encoding the polypeptide of interest may be present on a medium or high copy plasmid.

Methods of introducing a polynucleotide into a microbial host cell are well known in the art. Preferably, said polynucleotide is introduced into the microbial host cell by transformation. Transformation is the process by which an organism acquires a heterogeneous or recombinant polynucleotide. In an embodiment, the transformation of the microbial host cell involves use of divalent cations such as calcium chloride to increase the permeability of the membrane of the host cell, making the host cell chemically competent, thereby increasing the likelihood of uptake of the recombinant polynucleotide. In another embodiment, the polynucleotide is transformed into the host cell by electroporation. Further, the polynucleotide may be stably introduced into the host cell, i.e. into the chromosome of the host cell.

The microbial host cell is preferably a bacterium. More preferably, the microbial host cell is an *Escherichia coli* (*E. coli*) cell. *E. coli* is a Gram-negative gammaproteobacterium. The descendants of two isolates, K-12 and B strain, are used routinely in molecular biology as both a tool and a model organism. Preferably, the *E. coli* cell is an *E. coli* cell of the strain K12. Preferably, the *E. coli* cell strain is *E. coli* K12 BW25113 (see Grenier, 2014, Genome Announc. Sep-Oct; 2(5): e01038-14).

The present invention is not limited to *E. coli*, as the misincorporation of ncBCAAs also occurs in other bacteria such as *B. subtilis* or *S. marcescens*.

According to the present invention, the polynucleotide encoding the polypeptide of interest shall be introduced into a microbial host cell. Said microbial host cell shall have been modified such that an enzymatic activity as set forth elsewhere herein has been modulated (such as increased) in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell. Thus, the method of the present invention may further comprise (prior to step a), i.e. the introduction step) the step of providing or obtaining a microbial host cell having a modulated enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19). In this particular method said enzymatic activity shall be modulated as compared to an unmodified microbial host cell. Alternatively, an exemplary method of the present invention may comprise prior to step a), the step of modulating an enzymatic activity as set forth above in the microbial host cell.

In an embodiment of the present invention, the term “modulating an enzymatic activity” refers to increasing the enzymatic activity (as compared to the control or unmodified enzymatic activity). In another embodiment, the term refers to decreasing the enzymatic activity (as compared to the control or unmodified enzymatic activity).

Methods of increasing enzymatic activity of a polypeptide are well known in the art. For example, a polypeptide having enzymatic activity can be mutated to increase its activity (e.g. by enzyme design or enzyme engineering).

In an embodiment of the present invention, the enzymatic activity as referred to herein is increased in the microbial host cell into which the polynucleotide encoding the polypeptide of interest is introduced. Preferably, the enzymatic activity is increased by at least 10% as compared to the corresponding enzymatic activity in the unmodified microbial host cell. More preferably, the enzymatic activity is increased by at least 20%, even more preferably by at least 30%, and even more preferably by at least 50% as compared to the enzymatic activity in the unmodified microbial host cell. In an embodiment the enzymatic activity is increased by at least 100% or in another embodiment by at least 150%. Further, it is envisaged that the enzymatic activity is increased by 20% to 300%, such as by 50% to 200%. In an embodiment, the enzymatic activity is increased by overexpressing a polynucleotide encoding a polypeptide having said enzymatic activity.

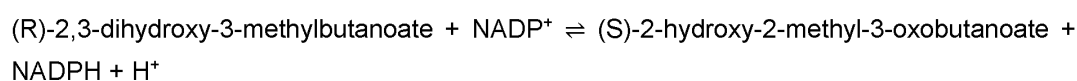
In an embodiment of the present invention, the enzymatic activity as referred to herein is decreased in the microbial host cell into which the polynucleotide encoding the polypeptide of interest is introduced. Preferably, the enzymatic activity is decreased by at least 10% as compared to the corresponding enzymatic activity in the unmodified microbial host cell. More

preferably, the enzymatic activity is decreased by at least 20%, even more preferably by at least 30%, and even more preferably by at least 50% as compared to the enzymatic activity in the unmodified microbial host cell. In an embodiment the enzymatic activity is decreased by at least 80% as compared to the enzymatic activity in the unmodified microbial host cell. However, it is envisaged that the enzymatic activity is not completely knocked-out. For example, the host cell might retain at least 5%, at least 10% or at least 20% of the enzymatic activity of the unmodified microbial host cell. Thus, the enzymatic activity might be decreased by 20% to 80%. Alternatively, the enzymatic activity might be decreased by 30 to 70 %. In an embodiment, the enzymatic activity is decreased by antisense RNAs which inhibit expression of the polynucleotide encoding a polypeptide having said enzymatic activity. Thus, the antisense RNAs should be complementary to the target gene.

In certain embodiments, the unmodified microbial host cell may be a wild-type cell. In a particular embodiment, an unmodified microbial host cell is of the same strain which has been modified. For example, the unmodified microbial host cell may be the *E. coli* K 12 strain, which may be subsequently modified as described herein.

The enzymatic activity to be modulated such as increased in the microbial host cell is, preferably, selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase activity (EC 4.3.1.19). The aforementioned enzymatic activities can be determined by assays well known in the art. The number in brackets is the so-called Enzyme Commission number (abbreviated "EC number"). As known in the art, every EC number consists of the letters EC followed by four numbers separated by periods. Those numbers represent a progressively finer classification of the enzyme. Based on the EC number, the reaction catalyzed by an enzyme is specified.

In an embodiment of the present invention, the enzymatic activity to be modulated is ketol-acid reductoisomerase(NADP(+)) activity. Accordingly, the activity of a ketol-acid reductoisomerase shall be modulated, preferably increased. The systematic name for this enzyme is *(R)-2,3-dihydroxy-3-methylbutanoate:NADP⁺ oxidoreductase (isomerizing)*. The EC number for this enzyme is EC 1.1.1.86. A ketol-acid reductoisomerase shall be capable of catalyzing the chemical reaction of (R)-2,3-dihydroxy-3-methylbutanoate + NADP⁺ into (S)-2-hydroxy-2-methyl-3-oxobutanoate, NADPH, and H⁺. Accordingly, it shall be capable of catalyzing the following chemical reaction:



In another embodiment of the present invention, the enzymatic activity to be modulated is acetohydroxyacid synthase activity. Accordingly, the activity of an acetohydroxyacid synthase shall be modulated (e.g. increased or decreased). This enzyme is also known as acetolactate synthase. The EC number for this enzyme is EC 2.2.1.6. An acetohydroxyacid synthase shall be capable of catalyzing the conversion of two pyruvate molecules to an acetolactate molecule and carbon dioxide. The reaction uses thiamine pyrophosphate to link the two pyruvate molecules. Accordingly, it shall be capable of catalyzing the following chemical reaction:



In another embodiment of the present invention, the enzymatic activity to be modulated is L-threonine dehydratase activity. Accordingly, the activity of a threonine synthase shall be modulated. The systematic name for this enzyme is *L-threonine ammonia-lyase (2-oxobutanoate-forming)*. The EC number for this enzyme is EC 4.3.1.19.

In another embodiment of the present invention, the enzymatic activity to be modulated is aspartate kinase activity. Accordingly, the activity of an aspartate kinase activity shall be modulated. The systematic name for this enzyme is ATP:L-aspartate 4-phosphotransferase. The EC number for this enzyme is EC 2.7.2.4. An aspartate kinase shall be capable of catalyzing the conversion of beta-aspartyl phosphate from aspartic acid and ATP. Threonine serves as an allosteric regulator of this enzyme to control the biosynthetic pathway from aspartic acid to threonine. Accordingly, it shall be capable of catalyzing the following chemical reaction:



In another embodiment of the present invention, the enzymatic activity to be modulated is homoserine dehydrogenase activity. Accordingly, the activity of a homoserine dehydrogenase shall be modulated. The systematic name for this enzyme is *L-homoserine:NAD(P)⁺ oxidoreductase*. The EC number for this enzyme is EC 1.1.1.3. A homoserine dehydrogenase shall be capable of catalyzing the chemical reaction of L-homoserine and NAD⁺ (or NADP⁺) into L-aspartate 4-semialdehyde, NADH (or NADPH), and H⁺. Accordingly, it shall be capable of catalyzing the following chemical reaction:



The enzyme from *Escherichia coli* which has aspartate kinase activity also catalyses the reaction of EC 1.1.1.3 homoserine dehydrogenase. Accordingly, said enzyme is a bifunctional enzyme which has both aspartate kinase activity and homoserine dehydrogenase activity. Accordingly, it is envisaged that both enzymatic activities, i.e. the activity of an aspartate kinase and the activity of a homoserine dehydrogenase are modulated in the microbial host cell as compared to the aspartate kinase activity and the homoserine dehydrogenase activity in an unmodified microbial host cell.

The enzymatic activity as referred to herein is preferably increased by introducing a polynucleotide encoding a polypeptide having said enzymatic activity into the microbial host cell and expressing said polynucleotide. For example, the polynucleotide shall be overexpressed. The polypeptide encoded by said polynucleotide shall be an enzyme having the enzymatic activity as referred to herein. Preferred polypeptides conferring said enzymatic activities are disclosed elsewhere in this specification in detail. It is to be understood that the polypeptides referred to herein may also exhibit further biological activities.

It is to be understood that said polynucleotide encoding a polypeptide having said enzymatic activity is not the polynucleotide encoding the polypeptide of interest. Accordingly, the microbial host cell as referred to herein comprises a) a polynucleotide encoding a polypeptide having an enzymatic activity as referred to herein and b) polynucleotide encoding the polypeptide of interest. Accordingly, the polynucleotides under a) and b) are co-expressed in the microbial host cell. Further, it is to be understood that both polynucleotides are recombinant polynucleotides which have been introduced into the host cell artificially.

Preferred sequences for the polynucleotide encoding a polypeptide having an enzymatic activity as referred to herein are provided below.

In an embodiment of the present invention, the polynucleotide encodes a polypeptide having ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86). Preferably, said polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 3, and/or
- ii) encodes a polypeptide comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 4.

In another embodiment of the present invention, the polynucleotide encodes a polypeptide having L-threonine dehydratase activity. Preferably, said polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 11, and/or

- ii) encodes a polypeptide comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 12.

In an embodiment, the enzymatic activity to be modulated is acetohydroxyacid synthase (AHAS) activity, in particular, the activity of AHAS isoform I, or variants thereof, AHAS isoform II, or variants thereof, or AHAS isoform III, or variants thereof. The activity of AHAS isoform I, or variants thereof, is preferably decreased in order to reduce misincorporation of ncBCAAs. The activity of AHAS isoforms II and III, or variants thereof, is preferably increased in order to reduce misincorporation of ncBCAAs.

As known in the art, a functional acetohydroxyacid synthase (e.g. AHAS I, AHAS II or AHAS III) comprises two large subunits and two small subunits which form a tetramer (which has acetohydroxyacid synthase activity). The two large subunits as well the two small subunits are identical. Accordingly, acetohydroxyacid synthase activity is preferably increased in the microbial host cell by introducing a first polynucleotide encoding for the large subunit of the acetohydroxyacid synthase and a second polynucleotide encoding for the small subunit of the acetohydroxyacid synthase into the microbial host cell and by expressing the polynucleotides. In the microbial cell, two large subunits and two small units preferably form a tetramer. It is to be understood that said tetramer shall have acetohydroxyacid synthase activity. The first and the second polynucleotide are preferably present in the same construct and are preferably expressed in a bicistronic manner under control of a single promoter.

AHAS I or variants thereof

In an embodiment of the present invention, said first polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 23, and/or
- ii) encodes large subunit of an acetohydroxyacid synthase, said large subunit comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 24,

and said second polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 25, and/or
- ii) encodes a small subunit of an acetohydroxyacid synthase, said small subunit comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 26.

SEQ ID NO: 24 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform I from *E. coli*, whereas SEQ ID NO: 26 is the amino acid sequence of the small subunit of said isoform.

AHAS II or variants thereof

In another embodiment of the present invention, said first polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 5, 27 or 31,
- ii) encodes large subunit of an acetohydroxyacid synthase, said large subunit comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 6, 28 or 32

and said second polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 7, 29, or 33 and/or
- ii) encodes the small subunit of an acetohydroxyacid synthase, said small subunit comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 8, 30 or 34.

SEQ ID NO: 6 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform II from *E. coli*, whereas SEQ ID NO: 8 is the amino acid sequence of the small subunit of said isoform. In a preferred embodiment, the AHAS isoform of *E. coli* is used (or a variant thereof).

SEQ ID NO: 28 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform III from *Shigella boydii*, whereas SEQ ID NO: 30 is the amino acid sequence of the small subunit of said isoform.

SEQ ID NO: 32 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform III from *Serratia marcescens*, whereas SEQ ID NO: 34 is the amino acid sequence of the small subunit of said isoform.

AHAS III of variants thereof

In another embodiment of the present invention, said first polynucleotide

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 13, 35, 39, or 43 and/or

- ii) encodes large subunit of an acetohydroxyacid synthase, said large subunit comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 14, 36, 40, or 44 and said second polynucleotide
 - i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 15, 37, 41, or 45 and/or
 - ii) encodes a small subunit of an acetohydroxyacid synthase, said small subunit comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 16, 38, 42, or 46.

SEQ ID NO: 14 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform III from *E. coli*, whereas SEQ ID NO: 16 is the amino acid sequence of the small subunit of said isoform. In a preferred embodiment, the AHAS isoform of *E. coli* is used (or a variant thereof).

SEQ ID NO: 36 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform III from *Shigella boydii*, whereas SEQ ID NO: 38 is the amino acid sequence of the small subunit of said isoform.

SEQ ID NO: 40 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform III from *Serratia marcescens*, whereas SEQ ID NO: 42 is the amino acid sequence of the small subunit of said isoform.

SEQ ID NO: 44 is the amino acid sequence of the large subunit of acetohydroxyacid synthase isoform III from *Bacillus subtilis*, whereas SEQ ID NO: 46 is the amino acid sequence of the small subunit of said isoform.

As set forth above, it is envisaged that both the activity of an aspartate kinase and of a homoserine dehydrogenase are modulated in the microbial host cell (as compared to the unmodified microbial host cell). In a preferred embodiment, the increase of these two activities is by introducing and expressing a polynucleotide encoding a polypeptide having aspartate kinase and homoserine dehydrogenase activity. Accordingly, said polypeptide is a bifunctional polypeptide. Preferably, said polynucleotide encoding a polypeptide having aspartate kinase and homoserine dehydrogenase activity:

- i) comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 17, and/or
- ii) encodes a polypeptide comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 18.

The term “polynucleotide” as used herein refers to a linear or circular nucleic acid molecule. It encompasses DNA as well as RNA molecules. The polynucleotide of the present invention shall be provided, preferably, either as an isolated polynucleotide (i.e. isolated from its natural context) or in genetically modified form. The term encompasses single as well as double stranded polynucleotides. Moreover, the term comprises chemically modified polynucleotides including naturally occurring modified polynucleotides such as glycosylated or methylated polynucleotides or artificial modified polynucleotides, such as biotinylated polynucleotides. The polynucleotide of the present invention is characterized in that it shall encode a polypeptide as referred to above. The polynucleotide, preferably, has a specific nucleotide sequence as mentioned above. Moreover, due to the degeneracy of the genetic code, polynucleotides are encompassed which encode a specific amino acid sequence as recited above.

Moreover, the term “polynucleotide” as used in accordance with the present invention further encompasses variants of the aforementioned specific polynucleotides. Said variants may represent orthologs, paralogs or other homologs of the polynucleotide of the present invention. The polynucleotide variants, preferably, comprise a nucleic acid sequence characterized in that the sequence can be derived from the aforementioned specific nucleic acid sequences by at least one nucleotide substitution, addition and/or deletion whereby the variant nucleic acid sequence shall still encode a polypeptide having the activity as specified above.

In a preferred embodiment of the present invention, the polynucleotide as set forth herein above shall have at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 3, 5, 7, 11, 13, 15, 17, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45. Further, it is envisaged that the polynucleotide encodes a polypeptide comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 4, 6, 8, 12, 14, 16, 18, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46. The term “at least 40%” as used herein means 40% or more than 40%. In particular, the term means, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity (in increasing order of preference). Moreover, the term encompasses the exact sequence, i.e., 100% sequence identity. Thus, the polynucleotide may have or comprise the nucleic acid sequence as shown in SEQ ID NO: 3, 5, 7, 11, 13, 15, 17, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45 or may encode a polypeptide comprising an amino acid sequence as shown in SEQ ID NO: 4, 6, 8, 12, 14, 16, 18, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46.

Sequence identity between amino acid sequences or nucleic acid sequences as used herein can be assessed by determining the number of identical nucleotides or amino acids between two nucleic acid sequences or amino acid sequences wherein the sequences are aligned so

that the highest order match is obtained. The percent identity values are, preferably, calculated over the entire amino acid or nucleic acid sequence region. A series of programs based on a variety of algorithms is available to the skilled worker for comparing different sequences. In this context, the algorithms of Needleman and Wunsch or Smith and Waterman give particularly reliable results. It can be calculated using published techniques or methods codified in computer programs such as, for example, BLASTP, BLASTN or FASTA (Altschul 1990, J Mol Biol 215, 403). The percent identity values are, preferably, calculated over a comparison window. A comparison window, preferably, is the length of the entire sequence of the shorter sequence to be aligned or at least half of said sequence. To carry out the sequence alignments, the program PileUp (Higgins 1989, CABIOS 5, 151) or the programs Gap and BestFit (Needleman 1970, J Mol Biol 48: 443; Smith 1981, Adv Appl Math 2: 482), which are part of the GCG software packet (Genetics Computer Group 1991, 575 Science Drive, Madison, Wisconsin, USA 53711), may be used. The sequence identity values recited above in percent (%) are to be determined, in another aspect of the invention, using the program GAP over the entire sequence region with the following settings: Gap Weight: 50, Length Weight: 3, Average Match: 10.000 and Average Mismatch: 0.000, which, unless otherwise specified, shall always be used as standard settings for sequence alignments. Preferably, the degree of sequence identity is calculated over the entire length.

In an embodiment, the algorithm of Needleman and Wunsch (see above) is used for the comparison of sequences. The algorithm is incorporated in the sequence alignment software packages GAP Version 10 and wNEEDLE. E.g., wNEEDLE reads two sequences to be aligned, and finds the optimum alignment along their entire length. When amino acid sequences are compared, a default Gap open penalty of 10, a Gap extend penalty of 0.5, and the EBLOSUM62 comparison matrix are used. When DNA sequences are compared using wNEEDLE, a Gap open penalty of 10, a Gap extend penalty of 0.5, and the EDNAFULL comparison matrix are used.

Variants also encompass polynucleotides comprising a nucleic acid sequence which is capable of hybridizing to the aforementioned specific nucleic acid sequences, preferably, under stringent hybridization conditions. These stringent conditions are known to the skilled worker and can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6. Accordingly, it is envisaged that the polynucleotide as set forth above shall be capable of hybridizing to a polynucleotide having a sequence as shown in SEQ ID NO: 3, 5, 7, 11, 13, 15, 17, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45, in particular, the polynucleotide as set forth above shall be the complement of a polynucleotide having a sequence as shown in SEQ ID NO: 3, 5, 7, 11, 13, 15, 17, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45.

The term "hybridization" is well known in art and refers to a process in which substantially homologous complementary nucleotide sequences anneal to each other. The hybridization process can occur entirely in solution, i.e. both complementary nucleic acids are in solution. Hybridization can also occur with one of the complementary nucleic acids immobilized to a matrix such as Sepharose beads. In order to allow hybridization to occur, the nucleic acid molecules are thermally or chemically denatured to melt a double strand into two single strands and/or to remove hairpins or other secondary structures from single stranded nucleic acids.

In accordance with the present invention, the polynucleotide shall be capable of hybridizing under stringent conditions, in particular under high stringency hybridization conditions, to a polynucleotide having a sequence as shown in SEQ ID NO: 3, 5, 7, 11, 13, 15, 17, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45. . The term "stringency" refers to the conditions under which hybridization takes place. The stringency of hybridization is influenced by conditions such as temperature, salt concentration, ionic strength and hybridization buffer composition. In a preferred embodiment high stringency hybridization conditions encompass hybridization at 65°C in 1x SSC or at 42°C in 1x SSC and 50% formamide, followed by washing at 65°C in 0.3x SSC. Further, it is envisaged that the hybridization is followed by washing at 65°C in 0.1x SSC. 1xSSC is 0.15M NaCl and 15mM sodium citrate; the hybridization solution and wash solutions may additionally include 5x Denhardt's reagent, 0.5-1.0% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA, and 0.5% sodium pyrophosphate.

In step (b) of the method of the present invention, the polypeptide of interest is expressed in said microbial host cell. Preferably, said step is carried out under conditions which allow for the expression of the said polypeptide. Suitable cultivation conditions can be determined by the skilled person without further ado. Preferably, the step is carried out under standard conditions. Thus, the host cell is incubated under standard conditions, e.g. under standard conditions described in Example 9.

In an embodiment, the expression is carried out under large-scale conditions in a bioreactor. The term "bioreactor" as used herein refers to a system in which conditions are closely controlled. In an embodiment, said bioreactor is a stirred tank bioreactor. Preferably, the bioreactor is made of a non-corrosive material such as stainless steel. The bioreactor can be of any size. In some embodiments, the bioreactor has a volume of at least 10, at least 100, 500, at least 1000, at least 2500, or at least 5000 liters or any intermediate volume.

In accordance with the present invention, the polynucleotide encoding the polypeptide of interest is operably linked to a promoter. The same applies to the polynucleotide encoding the polypeptide having the enzymatic activity as referred to herein. However, is preferred that the aforementioned polynucleotides are linked to different promoters.

The promoter shall allow for the expression of the polynucleotide. Preferably, the promoter is heterologous with respect to the sequence controlled by it. Thus, the promoter shall be heterologous to the polynucleotide encoding the polypeptide having the enzymatic activity and to the polynucleotide encoding the polypeptide of interest, respectively. Preferably, the promoter is a constitutive or inducible promoter. Constitutive and inducible promoters which allow for the expression in a microbial host cell are well known in the art.

A “constitutive promoter” refers to a promoter that is transcriptionally active during most, but not necessarily all, phases of growth and under most environmental conditions.

An “inducible promoter” refers to a promoter that has increased transcription initiation in response to a stimulus.

In accordance with the present invention, it is particularly contemplated that the promoter is an inducible promoter. Thus, the polynucleotide encoding the polypeptide of interest and/or the polynucleotide encoding the polypeptide having said enzymatic activity shall be operably linked to an inducible promoter. In this case, the enzymatic activity is only increased temporarily (i.e. after induction of the promoters).

In an embodiment, the inducible promoter is the arabinose-inducible araBAD promoter. The promoter is well known in the art. The arabinose-inducible araBAD promoter (PBAD) together with its regulator protein AraC has been described as expression system for high-level recombinant protein production as well as for metabolic engineering purposes since expression is tunable over a broad range of arabinose concentrations (Guzman *et al.*, 1995, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology*, 177(14), 4121-4130). The AraC-PBAD promoter system was also shown to be regulated by catabolite repression (Schleif, R., 2000, Regulation of the L-arabinose operon of *Escherichia coli*. *Trends in Genetics*, 16(12), 559-565; Megerle *et al.*, 2008, Timing and dynamics of single cell gene expression in the arabinose utilization system. *Biophysical journal*, 95(4), 2103-2115). Moreover, the applicability of the AraC-PBAD expression system was reported to be dependent on the *E. coli* strain. Cloning vectors usually contain native sequences of araBAD promoter and *araC* gene from the native araBAD regulatory system.

The sequence of the araBAD promoter is shown in SEQ ID NO: 19. Accordingly, the araBAD promoter preferably comprises a nucleic acid sequence as shown in SEQ ID NO: 19 or is at least 70%, at least 80% or is at least 90% identical thereto.

The nucleic acid sequence of the polynucleotide *araC* is shown in SEQ ID NO: 20. This gene encodes the AraC polypeptide. If the *araBAD* promoter is used, a polynucleotide expressing the AraC polypeptide has to be introduced into the microbial host cell to activate the promoter.

Further inducible promoters are the *rhaBAD* promoter, variants of the *lac* promoter (such as the *lacUV5*, *P_{tac}*, *P_{trc}* and *T7lac* promoter), the *XylSPm* promoter, and *tet* promoter.

The polypeptide having an enzymatic activity as referred to herein and/or the polypeptide of interest may comprise additional sequences. For example, the polypeptide(s) may further comprise a purification tag. The tag shall be operably linked to the polypeptide.

The tag shall allow the purification of the polypeptide. Such tags are well known in the art. The term "purification tag" as used herein preferably refers to an additional amino acid sequence (a peptide of the polypeptide) which allows for purification of the polypeptide. In an embodiment, the purification tag is a peptide or polypeptide which is not naturally linked to the polypeptide as referred to herein. Thus, the purification tag shall be heterologous with respect to the polypeptide.

Preferably, the purification tag is selected from the group consisting of a polyhistidine tag, a polyarginine tag, glutathione- S-transferase (GST), maltose binding protein (MBP), influenza virus HA tag, thioredoxin, staphylococcal protein A tag, the FLAG™ epitope, and the c-myc epitope. In a preferred embodiment, the purification tag is a polyhistidine tag. Preferably, said polyhistidine tag comprises at least 6 consecutive histidine residues.

In a preferred embodiment of the methods of the present invention, the method further comprises the isolation of the polypeptide of interest from the cell, and the purification of said polypeptide. The isolation of the polypeptide can be achieved by well-known means such as via the use of a suitable purification tag.

In an embodiment, the purification comprises the enrichment of polypeptides which do not comprise non-canonical branched-chain amino acids. This can be achieved by well-known methods which are e.g. described in Min, C. K., *et al.*, 2012, Insulin related compounds and identification. *Journal of Chromatography B*, 908, 105-112; Harris, R. P., *et al.*, 2014, Amino acid misincorporation in recombinant biopharmaceutical products. *Current opinion in biotechnology*, 30, 45-50; Cveticic, N., *et al.*, 2016, Proteome-wide measurement of non-canonical bacterial mistranslation by quantitative mass spectrometry of protein modifications. *Scientific reports*, 6, 2863. For example, the enrichment is achieved by chromatography.

The microbial host cell has been defined above. As set forth above, the microbial host cell is preferably an *E. coli* cell. As it is known by the skilled person, *E. coli* cells naturally expresses polypeptides having ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), or L-threonine dehydratase activity (EC 4.3.1.19).

- the *ilvC* gene encodes for a polypeptide having ketol-acid reductoisomerase (NADP(+)) activity
- the *ilvIH* operon encodes for a polypeptide having acetohydroxyacid synthase activity (acetohydroxyacid synthase isozyme III),
- the *ilvBN* operon encodes for a polypeptide having acetohydroxyacid synthase activity (acetohydroxyacid synthase isozyme I),
- the *ilvGM* operon encodes for a polypeptide having acetohydroxyacid synthase activity (acetohydroxyacid synthase isozyme II),
- the *thrA* gene encodes for a bifunctional polypeptide having aspartokinase and homoserine dehydrogenase activity, and
- the *ilvA* gene encodes for a polypeptide having L-threonine dehydratase activity.

Accordingly, the microbial host cell that has been modified may have endogenous enzymatic activities, i.e. activities which are naturally present in the microbial host cell. However, it is in particular envisaged that the microbial host does not express an endogenous polynucleotide encoding polypeptide having the enzymatic activity as referred to herein. Thus, the endogenous enzymatic activity may have been knocked-out. Accordingly, it is envisaged that the endogenous polynucleotide (and thus the endogenous enzymatic activity) has been knocked-out. Thus, the *ilvC* gene, the *ilvIH* operon, the *ilvBN* operon, the *thrA* gene, the *ilvGM* operon, or the *ilvA* gene may have been knocked-out. This can be achieved by methods well known in the art and is described in the Examples section. Further, such knock-outs are already known and can be assessed e.g. from the *E. coli* Genetic Stock Center (CGSC) of Yale University. E.g., *E. coli* K12 BW25113 has a two-base insertion event between base pairs 1250 and 1253 of *ilvG* genetic sequence, resulting in a frameshift mutation. As a consequence, a stop codon is formed resulting in a premature termination of *ilvG* gene expression (*ilvG*⁻). A functional AHASII is then not expressed and distal gene expression of *ilvEDA* operon is impaired (Lawther *et al.*, 1981, Molecular basis of valine resistance in *Escherichia coli* K-12, PNAS 78 (2) 922-925; Parekh, B.S. and Hatfield, G.W., 1997. Growth rate-related regulation of the *ilvGMEDA* operon of *Escherichia coli* K-12 is a consequence of the polar frameshift mutation in the *ilvG* gene of this strain. Journal of bacteriology, 179(6) 2086-2088.). In specific embodiments disclosed herein, knock-outs for $\Delta thrA$, $\Delta ilvA$, and $\Delta ilvC$ were acquired from the *E. coli* Genetic Stock Center (CGSC) of Yale University (see Examples section).

In an embodiment, the endogenous polynucleotide encoding the polypeptide having said enzymatic activity has been deleted. Alternatively, it may have been mutated, thereby deactivating the endogenous enzymatic activity. The endogenous polynucleotide is the naturally occurring polynucleotide which encodes for a polypeptide having an enzymatic activity as referred to herein.

For example, the L-threonine dehydratase activity can be increased in a microbial host cell by introducing and expressing a polynucleotide (i.e. a recombinant polynucleotide) encoding for a polypeptide having L-threonine dehydratase activity. In an embodiment, said microbial host cell expresses an endogenous polynucleotide encoding a polypeptide having L-threonine dehydratase activity (i.e. in addition to the recombinant polynucleotide). In another, more preferred embodiment, the endogenous polynucleotide encoding for the polypeptide having L-threonine dehydratase activity (and thus the *ilvA* gene) has been knocked-out in said microbial cell. Accordingly, said microbial cell does not express an endogenous polypeptide having L-threonine dehydratase activity. Only a recombinant polypeptide having said activity is expressed.

The knock-out of the endogenous gene (i.e. the endogenous enzymatic activity) in connection with the use of inducible promoters as described above allows for an improved regulation of the enzymatic activity in the microbial host cell.

In a preferred embodiment, the produced polypeptide of interest shows lower misincorporation of non-canonical branched-chain amino acids (ncBCAAs) as compared to a polypeptide which has been produced by expression in an unmodified microbial host cell. Accordingly, the misincorporation of ncBCAAs is reduced.

Preferably, the non-canonical branched-chain amino acid is selected from norvaline, norleucine and beta-methylnorleucine. Accordingly, the polypeptide of interest preferably shows lower norvaline, norleucine and beta-methylnorleucine misincorporation as compared to a polypeptide which has been produced by expression in an unmodified microbial host cell.

In some embodiments, misincorporation of non-canonical branched-chain amino acids is reduced in the intracellular soluble protein fraction. In some embodiments, misincorporation of non-canonical branched-chain amino acids is reduced in the inclusion body fraction.

Norvaline is an amino acid with the formula $\text{CH}_3(\text{CH}_2)_2\text{CH}(\text{NH}_2)\text{CO}_2\text{H}$. The compound is an isomer of the more common amino acid valine. The IUPAC name is 2-Aminopentanoic acid. Norvaline can be misincorporated into recombinant proteins in place of leucine residues.

Accordingly, the term “norvaline misincorporation” refers to the incorporation of a norvaline residue in the polypeptide of interest for which a leucine residue is encoded by the corresponding nucleic acid encoding the polypeptide of interest.

Norleucine is an amino acid with the formula $\text{CH}_3(\text{CH}_2)_3\text{CH}(\text{NH}_2)\text{CO}_2\text{H}$. Norleucine is an isomer of the more common amino acid leucine. The IUPAC name is 2-aminohexanoic acid. Norleucine can be misincorporated into recombinant proteins in place of methionine residues. Accordingly, the term “norleucine misincorporation” refers to the incorporation of a norleucine residue in the polypeptide of interest for which a methionine residue is encoded by the corresponding nucleic acid encoding the polypeptide of interest.

Beta-methylnorleucine is an amino acid. Synonyms for this amino acid are beta-methylnorleucine; (2S,3S)-2-Amino-3-methylhexanoic acid and [2S,3S,(+)]-2-Amino-3-methylhexanoic acid. Beta-methylnorleucine can be misincorporated into recombinant proteins in place of isoleucine residues. Accordingly, the term “beta-methylnorleucine misincorporation” refers to the incorporation of a beta-methylnorleucine residue in the polypeptide of interest for which an isoleucine residue is encoded by the corresponding nucleic acid encoding the polypeptide of interest.

Thus, misincorporation of the ncBCAAs as referred to above may occur if the polypeptide of interest comprises at least one leucine residue, at least one methionine residues, and/or at least one isoleucine residue. Accordingly, it is envisaged that the polypeptide of interest comprises at least one leucine residue, at least one methionine residues, and/or at least one isoleucine residue.

In an embodiment of the present invention, the misincorporation of norvaline is reduced.

In an embodiment of the present invention, the misincorporation of norleucine is reduced.

In an embodiment of the present invention, the misincorporation of beta-methylnorleucine is reduced.

In an embodiment of the present invention, the misincorporation of norvaline and norleucine is reduced.

In an embodiment of the present invention, the misincorporation of norleucine and beta-methylnorleucine is reduced.

In an embodiment of the present invention, the misincorporation of norvaline and beta-methylnorleucine is reduced.

In an embodiment of the present invention, the misincorporation of norvaline, norleucine and beta-methylnorleucine is reduced.

In accordance with the present invention, the percent reduction of the ncBCAA content, in particular of the norvaline, norleucine and/or beta-methylnorleucine content, is preferably at 5%, more preferably at least 10%, even more preferably at least 15%, and even more preferably at least 20% (as compared to the content in a polypeptide which has been produced by expression in an unmodified microbial host cell, i.e. the control cell). Thus, the misincorporation of ncBCAAs is preferably reduced by at least 5%, at least 10%, at least 15%, or at least 20%. The percent reduction in the content of ncBCAAs or in the content of norvaline, norleucine and/or beta-methylnorleucine is preferably calculated as a reduction in percentage of polypeptides of interest containing ncBCAAs or norvaline, norleucine and/or beta-methylnorleucine.

The definitions and explanations provided herein above apply *mutatis mutandis* to the following embodiments of the present invention.

The present invention further relates to a method for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest expressed in a microbial host cell, said method comprising

- (a) modulating an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) in the microbial host cell,
- (b) introducing a polynucleotide encoding the polypeptide of interest into said microbial host cell, and
- (c) expressing said polypeptide of interest in said microbial host cell.

Alternatively, step (a) may comprise obtaining a microbial host cell having modulated activity of an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19).

In an embodiment, the enzymatic activity as set forth above shall be modulated (increased or decreased) in the microbial host cell as compared to an unmodified microbial host cell. Methods of modulating the enzymatic activity is described above. Preferably, said enzymatic activity is modulated (e.g. increased) by introducing and expressing a polynucleotide encoding a polypeptide having said enzymatic activity in said microbial host cell. Preferred sequences for the polynucleotide/polypeptide are described above.

Further envisaged by the present invention is a microbial host cell comprising

- (a) a recombinant polynucleotide encoding a polypeptide of interest, and
- (b) a recombinant polynucleotide encoding a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19).

Moreover, the present invention relates to the use of a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), L-threonine dehydratase (EC 4.3.1.19), or of a polynucleotide encoding said polypeptide for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest produced in a microbial host cell.

Finally, the present invention concerns the use of a microbial host cell for producing a recombinant polypeptide of interest, wherein the microbial host cell has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell.

In the following, embodiments of the present invention are summarized. The definitions and explanations provided above apply *mutatis mutandis* to embodiments.

List of embodiments

1. A method for producing a recombinant polypeptide of interest in a microbial host cell, comprising the steps of
 - (a) introducing a polynucleotide encoding the polypeptide of interest into a microbial host cell which has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell, and
 - (b) expressing said polypeptide of interest in said microbial host cell.

2. The method of embodiment 1, wherein the produced polypeptide of interest shows lower misincorporation of non-canonical branched-chain amino acids as compared to a polypeptide which has been produced by expression in an unmodified microbial host cell.
3. The method of embodiments 1 and 2, wherein the amount of the produced polypeptide is increased as compared to the amount of said polypeptide produced in an unmodified microbial host cell.
4. The method of any one of embodiments 1 to 3, wherein the method further comprises the isolation of the polypeptide from the cell, and the purification of the polypeptide.
5. The method of embodiment 4, wherein the purification comprises the enrichment of polypeptides which do not comprise non-canonical branched-chain amino acids.
6. The method of embodiments 1 and 5, wherein said enzymatic activity is increased by introducing and expressing a polynucleotide encoding a polypeptide having said enzymatic activity in said microbial host cell.
7. The method of embodiment 6, wherein,
 - (a) the polynucleotide encodes a polypeptide having ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), and wherein
 - i) the polynucleotide comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 3, and/or
 - ii) the polynucleotide encodes a polypeptide comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 4,or
 - (b) the polynucleotide encodes a polypeptide having L-threonine dehydratase activity (EC 4.3.1.19), and wherein
 - i) the polynucleotide comprises a nucleic acid sequence having at least 40% sequence identity to the nucleic acid sequence as shown in SEQ ID NO: 11, and/or
 - ii) the polynucleotide encodes a polypeptide comprising an amino acid sequence having at least 40% sequence identity to the amino acid sequence shown in SEQ ID NO: 12.

8. The method of any one of embodiments 6 and 7, wherein said microbial host cell does not express an endogenous polypeptide having said enzymatic activity.
9. The method of any one of embodiments 1 to 8, wherein the polypeptide of interest is a therapeutic peptide or polypeptide.
10. The method of any one of embodiments 1 to 9, wherein the polynucleotide encoding the polypeptide of interest and/or the polynucleotide encoding the polypeptide having said enzymatic activity is operably linked to an inducible promoter.
11. The method of any one of embodiments 1 to 10, wherein said microbial host cell is an *Escherichia coli* cell.
12. A method for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest expressed in a microbial host cell, said method comprising
 - (d) modulating an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) in the microbial host cell,
 - (a) introducing a polynucleotide encoding the polypeptide of interest into said microbial host cell, and
 - (b) expressing said polypeptide of interest in said microbial host cell.
13. The method of embodiment 12, wherein the at least one non-canonical branched-chain amino acid is selected from norvaline, norleucine and beta-methylnorleucine.
14. A microbial host cell comprising
 - (a) a recombinant polynucleotide encoding a polypeptide of interest, and
 - (c) a recombinant polynucleotide encoding a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19).
15. Use of a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine

- dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19), or of a polynucleotide encoding said polypeptide for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest produced in a microbial host cell.
16. Use of a microbial host cell for producing a recombinant polypeptide of interest, wherein the microbial host cell has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell.
 17. The method according to any one of embodiments 1 to 13, the microbial cell of embodiment 14, or the use of embodiments 15 or 16, wherein the enzymatic activity is modulated.
 18. The method according to any one of embodiments 1 to 13 and 17, the microbial cell of embodiment 14 or 17, or the use of embodiments 15, 16 or 17, wherein the polypeptide of interest is a proinsulin.

FIGURES

Figure 1: Plasmid map of pSW3_*lacI*⁺, expressing mini-proinsulin. Plasmid pSW3_*lacI*⁺ confers resistance to ampicillin and expresses an 11 kDa recombinant protein (mini-proinsulin) under the control of an IPTG inducible promoter. Plasmid also co-expresses the repressor LacI, which is under the control of a *lacI*⁺ promoter variant. Generated with Snapgene®.

Figure 2: Plasmid map of 16ABZ5NP_1934177, containing fragment 1. Generated with Snapgene®.

Figure 3: Plasmid map of pCP20. Generated with Snapgene®.

Figure 4: Plasmid map of pETcoco1, including ori2 and its elements (*repE*, *sopA*, *sopB*, *sopC*). Generated with Snapgene®.

Figure 5: Plasmid map of the arabinose-tunable plasmid pACG_araBAD. It includes ori2 and its elements (*repE*, *sopA*, *sopB*, *sopC*) which ensure 1 copy plasmid per cell. Plasmid confers

resistance to chloramphenicol. Exogenous genes can be cloned by restriction cloning with enzymes NheI and NotI. Cloned genes are under the control of an arabinose promoter. AraC is necessary for the activation of the arabinose promoter and is also present in the plasmid. Additional unique restriction sites (SmaI, XhoI, XmaI, MssI) are present in order to allow exchange of the origin of replication, antibiotic resistance marker and promoter region. Generated with Snapgene®.

Figure 6: Plasmid map of the arabinose-tunable plasmid pACG_araBAD_ilmH. This plasmid results upon restriction cloning of gene *ilmH* into the original pACG_araBAD plasmid with enzymes NheI and NotI. This plasmid allows regulation of *ilmH* gene expression by addition of arabinose into the medium. Generated with Snapgene®.

Figure 7: Genetic modifications performed to wild type *E. coli* in this work. Genomic DNA contains a knock out of a certain gene (*gene A*). Expression of that gene can then be regulated by L-arabinose thanks to the presence of a tunable expression plasmid, containing such gene (pACG_araBAD_ilmA). In addition, plasmid pSW3_ilmA expresses mini-proinsulin, which allows testing ncBCAA misincorporation.

Figure 8: Plasmid map of pKD46. Generated with Snapgene®.

Figure 9: Plasmid map of pKD3. Generated with Snapgene®.

Figure 10: Plasmid map of pKD4. Generated with Snapgene®.

Figure 11: Molar concentrations of norvaline (A), norleucine (B) and β -methylnorleucine (C) normalized to OD_{600nm} present in the intracellular soluble protein fraction calculated over time after induction of different *E. coli* cultivations in a 15L reactor under standard conditions (STD) and under conditions triggering ncBCAA accumulation, i.e. pyruvate pulsing and oxygen limitation (PYR-O₂). Indicated in the legend, "WT *E. coli*" refers to the wild type strain *E. coli* K-12 BW25113 pSW3_ilmA, "*ilmGM*-tunable *E. coli*" alludes to strain *E. coli* K-12 BW25113 pSW3_ilmA pACG_araBAD_ilmGM and "*ilmH*-tunable *E. coli*" corresponds with strain *E. coli* K-12 BW25113 Δ ilmH pSW3_ilmA pACG_araBAD_ilmH. Arrows indicate time points where 1 g/L pyruvate pulse combined with 5 min O₂ limitation was applied.

Figure 12: Molar concentrations of norvaline (A) and norleucine (B) normalized to OD_{600nm} present in the inclusion body fraction calculated over time after induction of different *E. coli* cultivations in a 15L reactor under standard conditions (STD) and under cultivation conditions triggering ncBCAA accumulation, i.e. pyruvate pulsing and oxygen limitation (PYR-O₂).

Indicated in the legend, “WT *E.coli*” refers to the wild type strain *E. coli* K-12 BW25113 pSW3_*lacI*⁺. “*ilvGM*-tunable *E. coli*” alludes to strain *E. coli* K-12 BW25113 pSW3_*lacI*⁺ pACG_araBAD_*ilvGM* and “*ilvIH*-tunable *E. coli*” corresponds with strain *E. coli* K-12 BW25113 Δ *ilvIH* pSW3_*lacI*⁺ pACG_araBAD_*ilvIH*. Arrows indicate time points where 1 g/L pyruvate pulse combined with 5 min O₂ limitation was applied.

All references referred to above are herewith incorporated by reference with respect to their entire disclosure content as well as their specific disclosure content explicitly referred to in the above description.

The following examples merely illustrate the invention. They should not be construed as limiting the scope of protection in any way.

EXAMPLES

Example 1: Transformation of K12 BW25113 Δ *thrA*, Δ *ilvA*, Δ *ilvC*, Δ *ilvIH* and Δ *ilvBN* mutants with plasmid pSW3_*lacI*⁺

Δ *thrA*, Δ *ilvA*, Δ *ilvC* knock-outs

Strain *E. coli* K12 BW25113 as well as single knock-out mutants *E. coli* K12 BW25113 Δ *thrA*, Δ *ilvA* and Δ *ilvC* were acquired from the *E. coli* Genetic Stock Center (CGSC) of Yale University. Those mutant strains belong to the so-called KEIO collection (Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., ... & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular systems biology, 2(1)). These strains contain plasmid pKD46 (Figure 8) and a kanamycin resistance marker substituting target gene. CGSC identification for each acquired strain is indicated as follows:

Strain	CGSC identification number
<i>E. coli</i> K12 BW25113	7636
<i>E. coli</i> K12 BW25113 Δ <i>thrA</i>	JW0001-1
<i>E. coli</i> K12 BW25113 Δ <i>ilvA</i>	JW3745-2
<i>E. coli</i> K12 BW25113 Δ <i>ilvC</i>	JW3747-2

Plasmid pKD46 was curated from the acquired *E. coli* K12 BW25113 single knock-out mutants and the respective electrocompetent cells were transformed with pCP20 (Figure 3), a temperature-sensitive plasmid encoding a flipase. Plasmid pCP20 was curated and removal

of the antibiotic resistance marker from the mutants was tested by sequencing. The final respective electrocompetent *E. coli* K12 BW25113 mutants were transformed with pSW3_*lacI*⁺ (Figure 1), a high copy plasmid encoding mini-proinsulin.

ΔilvIH and *ΔilvBN* knock-outs

The knock-out strains *E. coli* K12 BW25113 *ΔilvIH* and *ΔilvBN* were not acquired but manually generated. Procedure to generate *E. coli* knock-out mutants described at "Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences, 97(12), 6640-6645" was used as a reference. Electrocompetent *E. coli* K12 BW25113 cells were transformed with pKD46, a temperature-sensitive recombination helper plasmid. Knock-out mutants for the operons *ilvIH* and *ilvBN* were then generated by transformation of electrocompetent *E. coli* K12 BW25113 cells containing pKD46 with the respective deletion cassette, previously obtained by PCR from pKD3 (Figure 9) or pKD4 (Figure 10). PCR-based verification was carried out to test proper integration of the deletion cassette into the genome. Plasmid pKD46 was curated and the respective electrocompetent *E. coli* K12 BW25113 mutants were transformed with pCP20 (Figure 3), a temperature-sensitive plasmid encoding a flipase. Plasmid pCP20 was curated and removal of the antibiotic resistance marker from the mutants was tested by sequencing. The final respective electrocompetent *E. coli* K12 BW25113 mutants were transformed with pSW3_*lacI*⁺ (Figure 1), a high copy plasmid encoding mini-proinsulin.

Example 2: Design and generation of an araC-PBAD tunable expression vector (pACG_araBAD)

An arabinose-based tunable expression plasmid, allowing regulation of genes of study, previously knocked-out, was obtained by the junction of 3 different DNA segments: Fragment 1 contains the araC-PBAD promoter region (Guzman *et al.*, 1995, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of bacteriology, 177(14), 4121-4130), UTRs, T7 terminator, a cloning site allowing gene cloning with rare-cutting restriction enzymes NheI and NotI and a C-terminal 6xhis-tag sequence allowing expression of fusion proteins. Fragment 2 contains a chloramphenicol resistance cassette while fragment 3 includes the ori2 origin of replication and genes *sopA*, *sopB*, *sopC* and *repE*.

Fragment 1 was chemically synthesized and subsequently cloned in plasmid 16ABZ5NP_1934177 (Figure 2). Fragment 1 was then amplified by PCR from plasmid 16ABZ5NP_1934177. Fragments 2 and 3 were directly amplified by PCR from plasmids

pCP20 (Figure 3) and pETcoco1 (Figure 4), respectively. The three DNA segments were joined together according to the In-Fusion cloning strategy (Takara Bio USA; Ref.: 639649) to generate plasmid pACG_araBAD (Figure 5).

Example 3: Cloning of the target genes into the tunable expression vectors

Genes of study (*ilvA*, *ilvC*, *ilvIH*, *ilvBN*, and *thrA*) were amplified by PCR from the *E. coli* K12 BW25113 genomic DNA and they were subsequently cloned into the previously generated tunable expression plasmid (Figure 5) by using restriction enzymes NheI and NotI. *ilvGM* was amplified from an *E. coli* strain where *ilvG* was not mutated. As example, plasmid map of the resulting arabinose-tunable plasmid pACG_araBAD_ *ilvIH* is shown in Figure 6.

Example 4: Transformation of the tunable expression vectors with cloned genes into the respective mutant containing pSW3_ *lacI*⁺

The final respective electrocompetent *E. coli* K12 BW25113 mutants containing pSW3_ *lacI*⁺ were transformed with the tunable expression plasmid expressing the corresponding gene of study.

After all genetic modifications, the generated *E. coli* mutant strains look like as described in Figure 7. A total of 6 *E. coli* mutant strains were then generated:

- *E. coli* K12 BW25113 Δ *ilvC* expressing pSW3_ *lacI*⁺ and pACG_araBAD_ *ilvC*
- *E. coli* K12 BW25113 Δ *ilvIH* expressing pSW3_ *lacI*⁺ and pACG_araBAD_ *ilvIH*
- *E. coli* K12 BW25113 Δ *ilvBN* expressing pSW3_ *lacI*⁺ and pACG_araBAD_ *ilvBN*
- *E. coli* K12 BW25113 Δ *thrA* expressing pSW3_ *lacI*⁺ and pACG_araBAD_ *thrA*
- *E. coli* K12 BW25113 Δ *ilvA* expressing pSW3_ *lacI*⁺ and pACG_araBAD_ *ilvA*
- *E. coli* K12 BW25113 expressing pSW3_ *lacI*⁺ and pACG_araBAD_ *ilvGM*

Example 5: Evaluation of L-arabinose induction effect on cell growth of mutant *E. coli* strains.

The aim of this experiment was to evaluate the effect of addition of different concentrations of L-arabinose on the expression of genes under the control of the araBAD promoter and, as a consequence, the effect on cell growth of the generated mutant *E. coli* strains.

E. coli cells were grown in a defined mineral salt medium containing (per L): 0.67 g Na₂SO₄, 0.82 g (NH₄)₂SO₄, 0.17 g NH₄Cl, 4.87 g K₂HPO₄, 1.2 g NaH₂PO₄ x 2H₂O and 0.33 g (NH₄)₂-H-citrate. The medium was supplemented with 0.67 ml/L trace elements solution and 0.67 ml/L MgSO₄ solution (1.0 M). The trace element solution comprised (per L): 0.5 g CaCl₂ x 2H₂O, 0.18 g ZnSO₄ x 7H₂O, 0.1 g MnSO₄ x H₂O, 16.7g FeCl₃ x 6H₂O, 0.16 g CuSO₄ x 5H₂O, 0.18 g

CoCl₂ x 6H₂O. Additionally, a 0.1 M Na-Phosphate buffer was used for further buffering of the medium.

In order to generate the cultivation, 50 µL of a cryostock containing the corresponding *E. coli* strain were used to inoculate 5 mL of supplemented defined mineral salt medium containing 5 g/L glucose, 100 µg/mL ampicillin, 25 µg/mL chloramphenicol (only for mutant *E. coli* strains) and a given concentration of L-arabinose. The cultivation was incubated at 37 °C and 220 rpm in an orbital shaker for 15h. At the end of the process, OD_{600nm} was measured for each cultivation. Following table summarizes results obtained:

Tested strain	L-arabinose concentration (%)	OD _{600nm}
<i>E. coli</i> K12 BW25113 $\Delta ilvC$ pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvC</i>	0	0.21
	0.05	0.32
	0.1	0.64
	0.4	2.34
	1.6	2.42
<i>E. coli</i> K12 BW25113 $\Delta thrA$ pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>thrA</i>	0	0.16
	0.05	0.17
	0.2	0.37
	0.4	2.88
	0.8	2.79
<i>E. coli</i> K12 BW25113 $\Delta ilvA$ pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvA</i>	1.6	2.69
	0	0.14
	0.025	0.78
	0.05	2.95
	0.2	2.83
<i>E. coli</i> K12 BW25113 $\Delta ilvH$ pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvH</i>	0.8	2.62
	1.6	2.61
	0	2.84
	0.025	3.03
	0.05	3.08
<i>E. coli</i> K12 BW25113 $\Delta ilvBN$ pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvBN</i>	0.2	3.03
	0.8	2.43
	1.6	2.77
	0	2.66
	0.025	2.66
<i>E. coli</i> K12 BW25113 pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvGM</i>	0.05	2.65
	0.2	2.96
	0.8	2.80
	1.6	2.66
	0	3.10
<i>E. coli</i> K12 BW25113 pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvGM</i>	0.025	3.05
	0.05	3.17
	0.2	3.01
	0.8	2.65
	1.6	2.93
<i>E. coli</i> K12 BW25113 pSW3_ <i>lacI</i> ⁺	0	3.10

Example 6: Cultivation conditions for evaluation of L-arabinose induction effect on ncBCAA production in mutant *E. coli* strains at mini-bioreactor level.

- Cultivation medium

E. coli cells were grown in a defined mineral salt medium containing (per L): 0.67 g Na₂SO₄, 0.82 g (NH₄)₂SO₄, 0.17 g NH₄Cl, 4.87 g K₂HPO₄, 1.2 g NaH₂PO₄ x 2H₂O and 0.33 g (NH₄)₂-H-citrate. The medium was supplemented with 0.67 ml/L trace elements solution and 0.67 ml/L MgSO₄ solution (1.0 M). The trace element solution comprised (per L): 0.5 g CaCl₂ x 2H₂O, 0.18 g ZnSO₄ x 7H₂O, 0.1 g MnSO₄ x H₂O, 16.7g FeCl₃ x 6H₂O, 0.16 g CuSO₄ x 5H₂O, 0.18 g CoCl₂ x 6H₂O. Additionally, a 0.1 M Na-Phosphate buffer was used for further buffering of the medium.

- Pre-cultivation

30 µL of a cryostock containing the corresponding *E. coli* strain were used to inoculate 30 mL of supplemented defined mineral salt medium containing 5 g/L glucose, 100 µg/mL ampicillin and 25 µg/mL chloramphenicol (only for mutant *E. coli* strains) in order to generate the pre-cultivation. For each mutant *E. coli* strain, medium also contained the minimum L-arabinose concentration necessary to recover the cell growth of the non-engineered strain, which was previously tested in Example 5. The pre-cultivation was incubated at 37 °C and 220 rpm in an orbital shaker, overnight.

- Main cultivation

OD_{600nm} at the end of the pre-cultivation was measured and a given volume was used to inoculate a 5 mL starting volume Pall Micro24 mini-bioreactor (Microreactor Technologies Inc.) so that initial OD_{600nm} was 0.4. The mini-reactor medium consisted of supplemented defined mineral salt medium containing 4 g/L glucose, 100 µg/mL ampicillin, 25 µg/mL chloramphenicol (only for mutant *E. coli* strains) and 1 µL/mL Desmophen antifoam. Medium was also supplemented with different concentrations of L-arabinose. Cultivation was carried out at 37 °C and the pH was maintained at 7 by automatic control with NH₄OH and CO₂. Stirrer speed was set to 800 rpm and DO set-point to 25 %, maintaining the last by automatically increasing the oxygen flow into the mini-reactor. Batch phase lasted around 4h. After batch phase was finished, 1 mL 400 g/L EnPump 200 solution and 50 µL 3000 U/L amylase solution were manually added into the mini-reactor, hence starting the fed-batch phase. EnPump 200 is a glucose polymer and when amylase is present, it constantly hydrolyses the polymer, thus delivering free glucose molecules over time, ensuring then a glucose-limited fermentation. 30 min after beginning of the fed batch phase, recombinant protein expression was induced by manual addition of an IPTG pulse to a final concentration of 0.5 mM. Fed-batch phase was active for 3.5 h.

Example 7: Amino acid analysis

Intracellular soluble protein fraction and inclusion body fraction were isolated from cell extracts according to protocol provided in “BugBuster Protein Extraction Reagent” kit (Merck, Cat. Nr.: 70584-4). 250 µL of the isolated intracellular soluble protein fraction were mixed with 750 µL 5M HCl. Isolated inclusion body pellets were resuspended with 200 µL dH₂O and 100 µL of the resulting inclusion body suspension were mixed with 900 µL 5M HCl. Resulting solutions were introduced in crystal vials with screw caps and vials were incubated closed for 24h at 80°C for acid hydrolysis. Afterwards, vials were left opened in a heating block for 16-24h at 65°C while rotating until all liquid was evaporated. Amino acid isolation from dried hydrolyzed samples was performed according to protocol provided in “EZ:faast™ for free (physiological) amino acid analysis by GC-FID” kit (Phenomenex, Cat. Nr.: KG0-7165). After isolation process, around 120 µL of the resulting upper layer were introduced into GC vials and 2 µL were then injected into the GC analyzer. The GC was run according to following oven conditions: equilibration time of 0.5 min, 110 °C for 1 min, 30 °C/min heating up to 320 °C and then 320 °C for 1 min. Nitrogen was used as a carrier gas with a constant flow rate of 1.5 mL/min. Injection was carried out with a 1:15 split ratio at 250 °C.

Example 8: Evaluation of L-arabinose induction effect on ncBCAA production in mutant *E. coli* strains at mini-bioreactor level..

Following tables summarize experimental results for each tested protein fraction, ncBCAA and mutant strain under different concentrations of L-arabinose. Bold data presented in tables correspond to the concentration of a given ncBCAA in the *E. coli* BW25113 pSW3_ *lacI*⁺ control strain. Data in percentage format shown in tables correspond to the variation percentage of the ncBCAA concentration obtained in the mutant strain under a given concentration of L-arabinose with respect to ncBCAA concentration obtained for the *E. coli* BW25113 pSW3_ *lacI*⁺ control strain.

Inclusion body fraction**A) Norvaline**

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvC</i> pSW3_ <i>lacI</i> ⁺
0.026 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvC</i>
0.4 % L-ara	6.8 %
0.8 % L-ara	-1.2 %

1.6 % L-ara	-21.5 %
-------------	---------

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>thrA</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>thrA</i>
0.036 (nmol aa/OD _{600nm})	
0.4 % L-ara	-36.4 %
0.8 % L-ara	-44.9 %
1.6 % L-ara	-40.6 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> +	<i>E. coli</i> BW25113 Δ <i>ilvIH</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvIH</i>	<i>E. coli</i> BW25113 Δ <i>ilvA</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvA</i>	<i>E. coli</i> BW25113 Δ <i>ilvBN</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvBN</i>	<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvGM</i>
0.023 (nmol aa/OD _{600nm})				
0.05 % L-ara	-12.3 %	-7.6 %	2.7 %	-61.8 %
0.2 % L-ara	-	4.5 %	48.3 %	-56.0 %
0.8 % L-ara	-40.7 %	-17.1 %	62.9 %	-58.4 %

For strains *E. coli* BW25113 Δ *ilvC* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvC* and *E. coli* BW25113 Δ *ilvIH* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvIH*, norvaline concentration significantly decreases when adding increasing concentrations of L-arabinose into the medium. The opposite behavior is observed for strain *E. coli* BW25113 Δ *ilvBN* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvBN*. For strains *E. coli* BW25113 Δ *thrA* pSW3_ *lacI*⁺ pACG_araBAD_ *thrA* and *E. coli* BW25113 pSW3_ *lacI*⁺ pACG_araBAD_ *ilvGM*, norvaline concentration shows a significant reduction but effect of increasing L-arabinose concentrations does not show a clear trend on variation of norvaline concentration. For strain *E. coli* BW25113 Δ *ilvA* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvA* no significant reduction of norvaline concentration was reported and effect of increasing L-arabinose concentrations does not seem to show a clear trend on variation of norvaline concentration.

B) Norleucine

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvC</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvC</i>
0.091 (nmol aa/OD _{600nm})	
0.4 % L-ara	71.9 %
0.8 % L-ara	8.7 %

1.6 % L-ara	-24.3 %
<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>thrA</i> pSW3_ <i>lacI</i> ⁺
0.237 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>thrA</i>
0.4 % L-ara	-57.2 %
0.8 % L-ara	-58.4 %
1.6 % L-ara	-60.3 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvIH</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvIH</i>	<i>E. coli</i> BW25113 Δ <i>ilvA</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvA</i>	<i>E. coli</i> BW25113 Δ <i>ilvBN</i> pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvBN</i>	<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvGM</i>
0.060 (nmol aa/OD _{600nm})				
0.05 % L-ara	-4.2 %	-22.3 %	67.4 %	-100.0 %
0.2 % L-ara	-	16.9 %	132.2 %	-88.5 %
0.8 % L-ara	-70.8 %	-10.5 %	245.9 %	-100.0 %

For strains *E. coli* BW25113 Δ *ilvC* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvC* and *E. coli* BW25113 Δ *ilvIH* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvIH*, norleucine concentration significantly decreases when adding increasing concentrations of L-arabinose into the medium. The opposite behavior is observed for strain *E. coli* BW25113 Δ *ilvBN* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvBN*. For strains *E. coli* BW25113 Δ *thrA* pSW3_ *lacI*⁺ pACG_araBAD_ *thrA* and *E. coli* BW25113 pSW3_ *lacI*⁺ pACG_araBAD_ *ilvGM*, norleucine concentration shows a significant reduction but effect of increasing L-arabinose concentrations does not seem to show a clear trend on variation of norleucine concentration. For strain *E. coli* BW25113 Δ *ilvA* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvA* no significant reduction of norleucine concentration was reported and effect of increasing L-arabinose concentrations does not seem to have a clear effect on variation of norleucine concentration.

Intracellular soluble protein fraction

A) Norvaline

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvC</i> pSW3_ <i>lacI</i> ⁺
0.300 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvC</i>
0.4 % L-ara	-22.0 %
0.8 % L-ara	-36.0 %
1.6 % L-ara	-39.5 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>thrA</i> pSW3_ <i>lacI</i> ⁺
0.253 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>thrA</i>
0.4 % L-ara	-41.3 %
0.8 % L-ara	-44.3 %
1.6 % L-ara	-41.6 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvIH</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvA</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvBN</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺
0.529 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvIH</i>	pACG_araBAD_ <i>ilvA</i>	pACG_araBAD_ <i>ilvBN</i>	pACG_araBAD_ <i>ilvGM</i>
0.05 % L-ara	-53.1 %	-19.5 %	-2.2 %	-71.9 %
0.2 % L-ara	-	-23.9 %	23.5 %	-73.0 %
0.8 % L-ara	-63.6 %	-20.8 %	211.4 %	-75.9 %

For strains *E. coli* BW25113 Δ *ilvC* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvC* and *E. coli* BW25113 Δ *ilvIH* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvIH*, norvaline concentration decreases when adding increasing concentrations of L-arabinose into the medium. The opposite behavior is observed for strain *E. coli* BW25113 Δ *ilvBN* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvBN*. For strains *E. coli* BW25113 Δ *thrA* pSW3_ *lacI*⁺ pACG_araBAD_ *thrA*, *E. coli* BW25113 Δ *ilvA* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvA* and *E. coli* BW25113 pSW3_ *lacI*⁺ pACG_araBAD_ *ilvGM*, norvaline concentration shows a significant reduction but effect of increasing L-arabinose concentrations does not seem to show a clear trend on variation of norvaline concentration.

B) Norleucine

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvC</i> pSW3_ <i>lacI</i> ⁺
0.180 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvC</i>
0.4 % L-ara	154.8 %
0.8 % L-ara	42.5 %
1.6 % L-ara	2.8 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>thrA</i> pSW3_ <i>lacI</i> ⁺
0.304 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>thrA</i>
0.4 % L-ara	-30.1 %
0.8 % L-ara	-26.6 %
1.6 % L-ara	-36.4 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvIH</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvA</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvBN</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺ pACG_araBAD_ <i>ilvGM</i>
0.241 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvIH</i>	pACG_araBAD_ <i>ilvA</i>	pACG_araBAD_ <i>ilvBN</i>	
0.05 % L-ara	-17.8 %	-58.5 %	2.8 %	-100.0 %
0.2 % L-ara	-	-56.1 %	36.6 %	-81.3 %
0.8 % L-ara	-69.9 %	-54.7 %	127.5 %	-100.0 %

For strains *E. coli* BW25113 Δ *ilvC* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvC* and *E. coli* BW25113 Δ *ilvIH* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvIH*, norleucine concentration decreases when adding increasing concentrations of L-arabinose into the medium. The opposite behavior is observed for strain *E. coli* BW25113 Δ *ilvBN* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvBN*. For strains *E. coli* BW25113 Δ *thrA* pSW3_ *lacI*⁺ pACG_araBAD_ *thrA*, *E. coli* BW25113 Δ *ilvA* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvA* and *E. coli* BW25113 pSW3_ *lacI*⁺ pACG_araBAD_ *ilvGM*, norleucine concentration shows a significant reduction but L-arabinose concentration does not seem to have a clear effect on norleucine concentration.

C) β -methylnorleucine

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvC</i> pSW3_ <i>lacI</i> ⁺
0.459 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvC</i>
0.4 % L-ara	-5.5 %
0.8 % L-ara	-25.2 %
1.6 % L-ara	8.5 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvC</i> pSW3_ <i>lacI</i> ⁺
0.459 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvC</i>
0.4 % L-ara	-5.5 %
0.8 % L-ara	-25.2 %
1.6 % L-ara	8.5 %

<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvIH</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvA</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 Δ <i>ilvBN</i> pSW3_ <i>lacI</i> ⁺	<i>E. coli</i> BW25113 pSW3_ <i>lacI</i> ⁺
0.432 (nmol aa/OD _{600nm})	pACG_araBAD_ <i>ilvIH</i>	pACG_araBAD_ <i>ilvA</i>	pACG_araBAD_ <i>ilvBN</i>	pACG_araBAD_ <i>ilvGM</i>
0.05 % L-ara	-26.4 %	-10.3 %	-42.0 %	-21.3 %
0.2 % L-ara	-	-14.1 %	-23.2 %	-12.8 %
0.8 % L-ara	-12.8 %	-2.2 %	-14.0 %	-18.5 %

For almost all tested strains and L-arabinose concentrations a slight reduction of β -methylnorleucine concentration is reported in the intracellular soluble protein fraction. However, this reduction is not really significant if compared with norvaline and norleucine, with exception of strain *E. coli* BW25113 Δ *ilvBN* pSW3_ *lacI*⁺ pACG_araBAD_ *ilvBN* induced with 0.05 % L-ara, where reduction reached about 42%. In addition, effect of increasing L-arabinose concentrations does not show a clear trend on variation of β -methylnorleucine concentration.

EXAMPLE 9: Screening of potential *ilvGM*- and *ilvIH*-tunable *E. coli* strains in a 15 L reactor under conditions triggering ncBCAA formation

During fermentation in large scale reactors, gradient zones of substrate, dissolved oxygen, pH and other parameters are formed due to inefficient mixing and *E. coli* cells respond to these environmental changes by modulating their metabolism (Schweder (1999). Monitoring of genes that respond to process-related stress in large-scale bioprocesses. *Biotechnology and bioengineering*, 65(2), 151-159). For instance, *E. coli* responds to glucose excess and oxygen limitation by shifting metabolism from oxidative respiration to mixed-acid fermentation, resulting in overflow metabolism (Enfors et al. (2001). Physiological responses to mixing in large scale bioreactors. *Journal of biotechnology*, 85(2), 175-185). Under these conditions, not only the mixed-acid fermentation products accumulate, but also pyruvate (Soini, J. et al. (2008). Norvaline is accumulated after a down-shift of oxygen in *Escherichia coli* W3110. *Microb. Cell Fact.*, 7: 1-14). Pyruvate excess present intracellularly increases the metabolic flux going to ncBCAA biosynthesis through the sequential keto acid chain elongation from pyruvate to α -ketocaproate over α -ketobutyrate and α -ketovalerate by the actuation of the *leu* operon-encoded enzymes (Apostol, I. et al. (1997). Incorporation of norvaline at leucine positions in recombinant human hemoglobin expressed in *Escherichia coli*. *Journal of Biological Chemistry*, 272(46), 28980-28988). This hypothesis is supported by the observations reported by Soini et al. (2011, Accumulation of amino acids deriving from pyruvate in *Escherichia coli* W3110 during fed-batch cultivation in a two-compartment scale-down bioreactor. *Advances in Bioscience and Biotechnology*, 2(05), 336): the combination of oxygen limitation with a constant glucose supply in a two-compartment STR-PFR scale-down bioreactor reported a significant impact on enhancing norvaline biosynthesis due to pyruvate accumulation in a recombinant *E. coli* cultivation. Furthermore, Soini et al. (2008) originally reported accumulation of pyruvate-based amino acids such the ncBCAAs norleucine and norvaline as well as alanine and valine in a standard STR fed-batch *E. coli* cultivation under glucose excess and induced oxygen limitation upon a stirrer downshift.

Concentration gradients happening in large industrial-scale bioreactors due to deficient mixing can be also simulated in small bioreactors in the laboratory. In this investigation scale-up effects are reproduced in a 15L reactor by combining pyruvate pulsing and O₂ limitation. This novel cultivation strategy might represent more accurately the physiological behavior of bacterial cultivations taking place in large scale bioreactors.

According to the mini-bioreactor screening (Example 8), strains *E. coli* K-12 BW25113 pSW3_*lacI*⁺ pACG_araBAD_*ilvGM* (*ilvGM*-tunable *E. coli*) and *E. coli* K-12 BW25113 Δ *ilvIH* pSW3_*lacI*⁺ pACG_araBAD_*ilvIH* (*ilvIH*-tunable *E. coli*) induced with 0.8% L-arabinose showed the best performance among all screened mutants in a 10 mL mini-bioreactor, since they reported the most significant reduction of ncBCAA mis-incorporation into recombinant mini-proinsulin in comparison with the control non-engineered *E. coli* strain. The aim of this

experiment was to verify the performance of the aforementioned potential tunable *E. coli* strains in a 15L reactor under cultivation conditions triggering formation of ncBCAA, i.e. pyruvate pulses and oxygen limitation, in order to confirm its advantage as strain ensuring product quality. For comparison, the control non-engineered *E. coli* host (*E. coli* K-12 BW25113 pSW3_*lacI*⁺) was also cultivated.

- **Cultivation of *E. coli* K-12 BW25113 pSW3_*lacI*⁺ (control strain) under standard conditions**

100 µL of a cryostock containing *E. coli* K-12 BW25113 pSW3_*lacI*⁺ were used to inoculate 500 mL of supplemented mineral salt medium containing 5 g/L glucose and 100 µg/mL ampicillin in order to generate the pre-culture. Composition of the mineral salt medium was as follows: 2 g/L Na₂SO₄, 2.468 g/L (NH₄)₂SO₄, 0.5 g/L NH₄Cl, 14.6 g/L K₂HPO₄, 3.6 g/L NaH₂PO₄·2H₂O and 1 g/L (NH₄)₂-H-citrate. The mineral salt medium was then supplemented with 2 mL/L MgSO₄ solution (1.0 M) and 2 mL/L trace elements solution. The trace element solution comprised (per L): 0.5 g CaCl₂ × 2H₂O, 0.18 g ZnSO₄ × 7H₂O, 0.1 g MnSO₄ × H₂O, 16.7g FeCl₃ × 6H₂O, 0.16 g CuSO₄ × 5H₂O, 0.18 g CoCl₂ × 6H₂O. The pre-culture was incubated at 37 °C and 220 rpm in an orbital shaker for 12 h, using an initial cold-start technique. OD_{600nm} at the end of the pre-culture was measured and a given volume was used to inoculate a 7 L starting volume reactor so that initial OD_{600nm} was 0.4. The reactor medium consisted of supplemented mineral salt medium containing 5 g/L glucose, 2 mL antifoam (Antifoam 2014, Sigma) and 100 µg/mL ampicillin. Cultivation was carried out at 37 °C and the pH was maintained at 7 by automatic control with 25% NH₄OH. Airflow was set to 7 vvm and DO set-point to 20 %, maintaining the last by using a cascade control altering stirrer speed (initial stirrer speed was set to 800 rpm). Batch phase lasted effectively 4h, with an intermediate 13h cold phase at 15 °C. At the end of the batch phase, exponential feeding was started, according to following equation:

$$F(t) = \frac{q_s}{S} \cdot (X \cdot V) \cdot e^{\mu_{set} \cdot t}$$

where $F(t)$ represents the feed flow rate over time (L h⁻¹), q_s the set-point of the specific substrate uptake rate (0.514 gS gX⁻¹ h⁻¹), S the concentration of glucose in the feed solution (442 g/L), X the biomass concentration over time (g/L), V the volume of the reactor over time (L), μ_{set} the set-point of the specific cell growth rate (0.3 h⁻¹) and t the time during the fed-batch phase. The feed solution consisted of TUB mineral salt medium supplemented with 4 mL/L trace elements solution, 2 mL/L MgSO₄ solution (1.0 M), 100 µg/mL ampicillin and 442 g/L glucose.

Exponential fed-batch phase was carried out for 3 hours and afterwards expression of recombinant mini-proinsulin was induced by automatic addition of IPTG to a final concentration of 0.5 mM. Induction time was 30 minutes. During the induction phase no feed was added into the reactor. After induction, a constant feeding phase was started, so that the constant flow rate was equal to the last flow rate achieved in the exponential feeding phase. Constant feed fed-batch phase was active for 5-6 h.

- **Cultivation of *E. coli* K-12 BW25113 pSW3_***lacI*⁺ **(control strain) under conditions triggering ncBCAA formation**

Cultivation was performed as described for the standard cultivation in previous section. However, after the exponential fed-batch phase, 1 g/L pyruvate pulse was automatically added into the reactor. Pyruvate solution was constantly pumped for 5 minutes. During that time period no feed was added, airflow rate was temporary set to 0 and DO cascade control was disconnected. After the first pyruvate pulse, expression of recombinant mini-proinsulin was induced by automatic addition of IPTG to a final concentration of 0.5 mM. Induction time was 30 minutes. During the induction phase no feed was added into the reactor and airflow and DO cascade control were re-established. After induction, sequential 1 g/L pyruvate pulses were performed each 30 min as described above for a total of 4 pulses. Between pulses, constant feeding phase was activated, so that the constant flow rate was equal to the last flow rate achieved in the exponential feeding phase, and airflow and DO cascade control were re-established. Constant feed fed-batch phase was active for 5-6 h.

- **Cultivation of *ilvGM*-tunable *E. coli* under conditions triggering ncBCAA formation**

Cultivation operation was as described in previous section and only minor changes were done in order to adapt the cultivation process to *ilvGM*-tunable *E. coli* strain. Both pre-culture and reactor medium contained additionally 25 µg/mL chloramphenicol. The reactor medium additionally contained 0.8% L-arabinose, necessary to induce expression of gene *ilvGM* hosted in plasmid pACG_araBAD_*ilvGM*. The feeding solution was also additionally supplemented with 25 µg/mL chloramphenicol and 0.8% L-arabinose.

- **Cultivation of *ilvIH*-tunable *E. coli* under conditions triggering ncBCAA formation**

Cultivation operation was as described in previous section.

Example 10: Analysis of ncBCAA

Concentrations of ncBCAA present in the intracellular soluble protein fraction and in the inclusion body fraction over cultivation time for each tested strain in Example 9 are shown in Figure 11 and Figure 12, respectively.

The cultivation of the control *E. coli* strain subjected to pyruvate pulses combined with O₂ limitation ("WT *E. coli*, PYR-O2") reported a progressive accumulation of norleucine and β -methylnorleucine in the intracellular soluble protein fraction over time after induction, being that more significant for norleucine. Furthermore, norvaline concentration also increased progressively under aforementioned cultivation conditions, but only until 3h after induction. From that time point on, norvaline concentration progressively dropped until reaching initial values at 5 h after induction. This might suggest that, after 2h from last pyruvate pulse combined with O₂ limitation, its associated effect triggering norvaline accumulation is not active anymore (Figure 11). As expected, the aforementioned strain reported a higher level of norvaline and norleucine than the control *E. coli* strain cultivated under standard conditions ("WT *E. coli*, STD"). Such concentration difference could not be observed for β -methylnorleucine.

Both tested potential mutants in cultivations "*ilvGM*-tunable *E. coli*, PYR-O2" and "*ilvIH*-tunable *E. coli*, PYR-O2" reported a dramatic reduction of norvaline and norleucine concentrations in the intracellular soluble protein fraction, being such decrease higher for norleucine in "*ilvGM*-tunable *E. coli*, PYR-O2". However, β -methylnorleucine concentrations did not significantly vary with respect to the control cultivation. It is noteworthy to highlight that, for most samples, norvaline could not be properly detected since concentrations were under the limit of detection of the GC-FID equipment (Figure 11).

The cultivation of the control *E. coli* strain subjected to pyruvate pulses combined with O₂ limitation ("WT *E. coli*, PYR-O2") reported a progressive accumulation of norvaline and norleucine in the inclusion body fraction over time after induction. As expected, the aforementioned strain reported a higher level of norvaline and norleucine than the control *E. coli* strain cultivated under standard conditions ("WT *E. coli*, STD"). Again, and similar to reported in the intracellular soluble fraction, both tested potential mutants in cultivations "*ilvGM*-tunable *E. coli*, PYR-O2" and "*ilvIH*-tunable *E. coli*, PYR-O2" reported a dramatic reduction of norvaline and norleucine concentrations in the inclusion body fraction, being this decrease even higher for norleucine in "*ilvGM*-tunable *E. coli*, PYR-O2". Norvaline could not be detected in any case for both tested mutants. β -methylnorleucine could not be detected in any tested samples (Figure 12).

Claims

1. A method for producing a recombinant polypeptide of interest in a microbial host cell, comprising the steps of
 - (d) introducing a polynucleotide encoding the polypeptide of interest into a microbial host cell which has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell, and
 - (e) expressing said polypeptide of interest in said microbial host cell.
2. The method of claim 1, wherein the produced polypeptide of interest shows lower misincorporation of non-canonical branched-chain amino acids as compared to a polypeptide which has been produced by expression in an unmodified microbial host cell.
3. The method of any one of claims 1 and 2, wherein the method further comprises the isolation of the polypeptide from the cell, and the purification of the polypeptide.
4. The method of claim 3, wherein the purification comprises the enrichment of polypeptides which do not comprise non-canonical branched-chain amino acids.
5. The method of claims 1 and 4, wherein said enzymatic activity is increased by introducing and expressing a polynucleotide encoding a polypeptide having said enzymatic activity in said microbial host cell.
6. The method of any one of claim 5, wherein said microbial host cell does not express an endogenous polypeptide having said enzymatic activity.
7. The method of any one of claims 1 to 6, wherein the polypeptide of interest is a therapeutic peptide or polypeptide.
8. The method of any one of claims 1 to 7, wherein the polynucleotide encoding the polypeptide of interest and/or the polynucleotide encoding the polypeptide having said enzymatic activity is operably linked to an inducible promoter.

9. The method of any one of claims 1 to 8, wherein said microbial host cell is an *Escherichia coli* cell.
10. A method for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest expressed in a microbial host cell, said method comprising
 - (a) modulating an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) in the microbial host cell,
 - (b) introducing a polynucleotide encoding the polypeptide of interest into said microbial host cell, and
 - (c) expressing said polypeptide of interest in said microbial host cell.
11. The method of claim 10, wherein at least one non-canonical branched-chain amino acid is selected from norvaline, norleucine and beta-methylnorleucine.
12. Use of a polypeptide having an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19), or of a polynucleotide encoding said polypeptide for reducing misincorporation of at least one non-canonical branched-chain amino acid into a recombinant polypeptide of interest produced in a microbial host cell.
13. Use of a microbial host cell for producing a recombinant polypeptide of interest, wherein the microbial host cell has been modified such that an enzymatic activity selected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19) is modulated in said microbial host cell as compared to the enzymatic activity in an unmodified microbial host cell.
14. The method according to any one of claims 1 to 11, or the use of claims 12 or 13, wherein the enzymatic activity is modulated.

15. The method according to any one of claims 1 to 11 and 14, or the use of claim 12, 13 or 14, wherein the polypeptide of interest is a proinsulin.
16. A microbial host cell comprising
 - (a) a recombinant polynucleotide encoding a polypeptide of interest, and
 - (b) a recombinant polynucleotide encoding a polypeptide having an enzymatic activityselected from the group consisting of ketol-acid reductoisomerase (NADP(+)) activity (EC 1.1.1.86), acetohydroxyacid synthase activity (EC 2.2.1.6), aspartate kinase activity (EC 2.7.2.4), homoserine dehydrogenase activity (EC 1.1.1.3), and L-threonine dehydratase (EC 4.3.1.19).
17. A bioreactor comprising the microbial host cell of claim 16.
18. The bioreactor of claim 17, wherein bioreactor has a volume of at least 10 L.