(vii) can be produced in an amount of more than 0.5 g protein per 100 g cells and is scalable to more than 100 g, preferably more than 200 g, more preferably more than 300 g protein per batch.

Not all of the objectives will be realized by all embodiments of the invention. The scope of the invention is defined by the claims. It is however preferred to meet 2, 3, 4, 5, or 6 of the aforementioned objectives of the invention.

SUMMARY

In a first aspect, the disclosure provides a method for the recombinant expression of a MYDGF protein in a bacterial host cell, comprising (a) providing a host cell that comprises a nucleic acid encoding a protein which after maturation consists of 143 amino acids having the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2; (b) culturing the host cell under conditions that allow the expression of the protein; (c) isolating inclusion bodies containing MYDGF protein from the host cell; and (d) solubilising the inclusion bodies and refolding the MYDGF protein.

In a second aspect, the disclosure provides a use of a composition of the disclosure for the preparation of a pharmaceutical composition.

In a third aspect, the disclosure provides a pharmaceutical composition comprising a composition of the disclosure.

In a fourth aspect, the disclosure provides a composition of the disclosure or pharmaceutical composition of the disclosure for use as a medicament.

In a fifth aspect, the disclosure provides a use of the composition of the disclosure or pharmaceutical composition of the disclosure in the manufacture of a medicament for (i) treating or preventing a disease or condition selected from the group consisting of injury, wounding, ischemia, reperfusion injury, trauma, mechanical overload, intoxication, surgery, primary or acquired cardiomyopathy, postischemic contractile dysfunction, myocardial infarction, preferably acute myocardial infarction, angina pectoris, heart failure, inflammation of the heart, heart insufficiency, hypertrophy, and fibrosis; (ii) promoting or improving heart tissue regeneration, cardiomyocyte proliferation, neovascularisation, heart function or left ventricular systolic

function after myocardial infarction; (iii) protecting cardiomyocyte from death, e.g. through apoptosis or necrosis; or (iv) decreasing infarct size after myocardial infarction, preferably acute myocardial infarction.

In a first example, the disclosure relates to a method for the recombinant expression of a MYDGF protein in a host cell.

In a second example, the disclosure relates to a composition which is obtainable from the method of the first example of the disclosure.

In a third example, the disclosure relates to the use of a composition of the second example of the disclosure for the preparation of a pharmaceutical composition.

In a fourth example, the disclosure relates to a pharmaceutical composition comprising a composition of the second example of the disclosure.

In a fifth example, the disclosure relates to a pharmaceutical composition of the fourth example of the disclosure for use as a medicament.

In a sixth example, the disclosure relates to a composition of the second example of the disclosure or a pharmaceutical composition of the fourth example of the disclosure for use in a method of (i) treating or preventing a disease or condition selected form the group consisting of injury, wounding, ischemia, reperfusion injury, trauma, mechanical overload, intoxication, surgery, primary or acquired cardiomyopathy, postischemic contractile dysfunction, myocardial infarction, preferably acute myocardial infarction, angina pectoris, heart failure, inflammation of the heart, heart insufficiency, hypertrophy, and fibrosis; (ii) promoting or improving heart tissue regeneration, cardiomyocyte proliferation, neovascularisation, heart function or left ventricular systolic function after myocardial infarction; (iii) protecting cardiomyocyte from death, e.g. through apoptosis or necrosis; or (iv) decreasing infarct size after myocardial infarction, preferably acute myocardial infarction.

In a seventh example, the disclosure relates to a protein having the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2, and preferably SEQ ID NO:1.

In an eighth example, the disclosure relates to a nucleic acid encoding a protein of the seventh example of the disclosure.

In a ninth example, the disclosure relates to a vector comprising the nucleic acid of the eighth example of the disclosure.

In a tenth example, the disclosure relates to a host cell comprising a protein of the seventh example of the disclosure, a nucleic acid of the eighth example of the disclosure, or a vector of the ninth example of the disclosure.

In an eleventh example, the disclosure relates to a pharmaceutical composition comprising the protein of the seventh example of the disclosure.

In a twelfth example, the disclosure relates to a protein of the seventh example of the disclosure or a pharmaceutical composition of the eleventh example of the disclosure for use as a medicament.

In a thirteenth example, the disclosure relates to a protein of the seventh example of the disclosure or a pharmaceutical composition of the eleventh example of the disclosure for use in a method of (i) treating or preventing a disease or condition selected from the group consisting of injury, wounding, ischemia, reperfusion injury, trauma, mechanical overload, intoxication, surgery, primary or acquired cardiomyopathy, postischemic contractile dysfunction, myocardial infarction, preferably acute myocardial infarction, angina pectoris, heart failure, inflammation of the heart, heart insufficiency, hypertrophy, and fibrosis; (ii) promoting or improving heart tissue regeneration, cardiomyocyte proliferation, neovascularisation, heart function or left ventricular systolic function after myocardial infarction; (iii) protecting cardiomyocyte from death, e.g. through apoptosis or necrosis; or (iv) decreasing infarct size after myocardial infarction, preferably acute myocardial infarction.

In a fourteenth example, the disclosure relates to a method for the recombinant expression of a protein of the seventh example of the disclosure in a host cell.

In a fifteenth example, the disclosure relates to a composition which is obtainable from the method of the fourteenth example of the disclosure.

Finally, in a sixteenth example, the disclosure relates to the use of a host cell of the tenth example of the disclosure for the recombinant expression of a MYDGF protein.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides proteins having MYDGF activity and exerting a minimal degree of degradation and process-derived post translational modifications upon recombinant expression in a host cell. The invention also provides a method for producing these proteins in large amounts in a cell-based expression system. The method requires significantly reduced purification efforts for providing a homogeneous protein composition that is suitable for being formulated into a pharmaceutical product. If reference is made to SEQ ID NO:1 or SEQ ID NO:2 hereinafter, it must be understood that SEQ ID NO:1 is the preferred alternative.

The proteins disclosed herein in the context with the present invention are depicted in SEQ ID NO:1 and SEQ ID NO:2. Both proteins consist of 143 amino acid building blocks and comprise the complete amino acid sequence of the mature human MYDGF protein. The native human MYDGF protein is expressed as a precursor protein with an N-terminal signal peptide of 31 amino acids and a C-terminal KDEL-like endoplasmic reticulum (ER) retention sequence. The sequence of the MYDGF precursor protein of 173 amino acids is set forth herein as SEQ ID NO:5. Upon cleavage of the N-terminal signal peptide, the mature MYDGF is released. The sequence of the mature human MYDGF protein consists of 142 amino acids and is set forth herein as SEQ ID NO:6. The proteins of the invention differ from mature MYDGF only in a single amino acid that has been added to their N-terminus. Hence, these proteins can be regarded as recombinant variants of the native human MYDGF protein. The protein of SEQ ID NO:1 comprises an additional alanine residue at its N-terminus, which is not present in the mature human MYDGF protein. This variant is referred to as "[+A]" or "the [+A] variant" herein below. The protein of SEQ ID NO:2 differs from the mature human MYDGF protein by an additional serine residue at its N-terminus. This variant is referred to herein as "[+S]" or "the [+S] variant".

The proteins are associated with a particularly low risk of comprising antigenic epitopes that are not derived from human MYDGF. Accordingly, the proteins of the invention exert a minimal risk for generating anti-drug antibodies. The differences between the proteins of the invention and native human MYDGF protein reside in a single amino acid which means that the region added to the native protein is too small to give rise to new epitopes. The absence of anti-

CLAIMS

1. Method for the recombinant expression of a MYDGF protein in a bacterial host cell, comprising

   (a) providing a host cell that comprises a nucleic acid encoding a protein which after maturation consists of 143 amino acids having the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2;

   (b) culturing the host cell under conditions that allow the expression of the protein;

   (c) isolating inclusion bodies containing MYDGF protein from the host cell; and

   (d) solubilising the inclusion bodies and refolding the MYDGF protein.

2. Method of claim 1, wherein step (a) comprises

   (i) providing a host cell that comprises a nucleic acid that contains an open reading frame, flanked by start and stop codon, according to the sequence of SEQ ID NO:11 or SEQ ID NO:12, and preferably according to the sequence of SEQ ID NO:11, operably linked to a promotor; or

   (ii) providing a host cell that comprises a nucleic acid encoding a protein which before maturation consists of 144 amino acids having the amino acid sequence of SEQ ID NO:15 or SEQ ID NO:16, and preferably the amino acid sequence of SEQ ID NO:15.

3. Method of any of claims 1-3, wherein step (a) comprises providing a host cell that comprises a nucleic acid of SEQ ID NO:7 or SEQ ID NO:8, and preferably a nucleic acid of SEQ ID NO:7.

4. Method of any of claims 1-3, wherein said maturation is the removal of the N-terminal methionine residue.

5. Method of claim 4, wherein said removal of the N-terminal methionine residue is effected by one or more host cell-derived aminopeptidases.

6. Method of any of claims 1-5, further comprising (e) obtaining after step (d) a refolded MYDGF protein of 143 amino acids having the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2, and preferably the amino acid sequence of SEQ ID NO:1.

7. Method of any of claims 1-6, wherein refolding of the protein in step (d) comprises the incubation of the protein in the presence of urea.

8. Method of any of claims 1-7, wherein said method further comprises (f) purifying the MYDGF protein.

9. Method of claim 8, wherein step (f) comprises ultrafiltration, diafiltration, hydrophobic interaction chromatography and/or anion ion exchange chromatography.

10. Method of claim 9, wherein anion exchange chromatography or hydrophobic interaction chromatography step is performed by contacting the MYDGF protein to the chromatography resin material under conditions that allow for the adsorption of the MYDGF protein to the resin, optionally washing the resin, and eluting the MYDGF protein from the resin.

11. Method of claim 10, wherein the adsorption of the MYDGF protein to the anion exchange chromatography resin is performed under conditions of low ionic strength.

12. Method of claim 11, wherein adsorption is performed at a conductivity of less than 3 mS/cm, less than 2 mS/cm, less than 1,5 mS/cm or less than 1mS/cm.

13. Any method of any of claims 9-12, wherein desorption of the MYDGF protein from the anion exchange resin is effected by increasing the salt concentration and/or lowering the pH of the liquid phase.

14. Method of claim 13, wherein the bacterial host cell is an *Escherichia coli* cell.

15. Method of claim 14, wherein the *Escherichia coli* cell is an *Escherichia coli* cell of strain BL21 or a derivative strain thereof.

16. Method of any of claims 1-14, wherein said nucleic acid is DNA or RNA.

17. Method of claim 16, wherein said nucleic acid comprises the sequences of SEQ ID NO:11 or SEQ ID NO:12.

18. Method of any of claims 1-17, wherein the nucleic acid is contained in a vector.

19. Method of claim 18, wherein said vector is a prokaryotic expression vector.

20. Method of claim 18 or 19, wherein said vector comprises a T7 promoter.

21. Method of any of claims 18-20, wherein said vector comprises or consists of the sequences of SEQ ID NO:7 or SEQ ID NO:8.

22. Composition obtainable from the method of any of claims 1-21, wherein said composition comprises a protein of 143 amino acids having the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2.

23. Composition of claim 22, wherein said composition comprises less than 1% (w/w) of protein molecules that are shorter than 143 amino acids, as measured by liquid chromatography mass spectrometry (LCMS).

24. Composition of claim 22 or 23, wherein said composition comprises less than 20 pg/mg, preferably less than 15 pg/mg, more preferably less than 10 pg/mg, and most preferably less than 5 pg/mg, less than 3 pg/mg, less than 2 pg/mg or less than 1 pg/mg host cell DNA.

25. Composition of any of claims 22-24, wherein said composition comprises less than 0.2 EU/mg, and preferably less than 0.1 EU/mg or 0.08 EU/mg bacterial endotoxin.

26. Composition of any of claims 22-25, wherein less than 8% (w/w), and preferably less than 7% (w/w), less than 6% (w/w) or less than 5% (w/w), less than 4% (w/w), less than 3% (w/w), or less than 2% (w/w) of the proteins in said composition are carbamoylated.

27. Composition of any of claims 22-26, wherein less than 6% (w/w), and preferably less than 5% (w/w), less than 4% (w/w) or less than 3% (w/w), or less than 2% (w/w) of the proteins in said composition are gluconoylated.

28. Composition of any of claims 22-27, wherein less than 8%, and preferably less than 7%, less than 6% or less than 5% of the MYGDF proteins in the composition of the invention are carbamoylated, wherein the percentage is based on the sum of the peak intensities of unmodified MYDGF protein as well as annotated post-translational modification (PTM) species of MYDGF in a deconvoluted intact mass spectrum of the MYDGF protein in the composition.

29. Composition of any of claims 22-28, wherein less than 6%, and preferably less than 5%, less than 4% or less than 3% of the MYDGF proteins in the composition of the invention are gluconylated, wherein the percentage is based on the sum of the peak intensities of unmodified MYDGF protein as well as annotated post-translational modification (PTM)

species of MYDGF in a deconvoluted intact mass spectrum of the MYDGF protein in the composition.

30. Composition of any of claims 22-29, wherein said composition comprises urea.

31. Composition of any of claims 22-30, wherein said composition comprises a protein of 143 amino acids having the amino acid sequence of SEQ ID NO:1 and the ratio of the signal for the protein according to SEQ ID NO:1 and the signals for shorter variants in liquid chromatography mass spectrometry (LCMS) after reductive dimethylation (stable isotope dimethyl labelling, SIDL) is at least 50, and preferably more than 100, 200, 300 or 400, wherein only signals from non-carbamoylated and non-gluconoylated proteins are used for calculating said ratio.

32. Composition of any of claims 22-31, wherein said composition comprises a protein of 143 amino acids having the amino acid sequence of SEQ ID NO:2 and the ratio of the signal for the protein according to SEQ ID NO:2 and the signals for shorter variants in liquid chromatography mass spectrometry (LCMS) after reductive dimethylation (stable isotope dimethyl labelling, SIDL) is at least 50, and preferably more than 75, 100, 150 or 175, wherein only signals from non-carbamoylated and non-gluconoylated proteins are used for calculating said ratio.

33. Composition of any of claims 22-32, wherein said composition comprises a protein which is folded such that more than 70%, and preferably more than 80%, more than 90%, or more than 95%, of the ^{1}H and/or ^{15}N peaks in the two-dimensional nuclear magnetic resonance spectroscopy (2D-NMR) map result in combined chemical shift deviation (CCSD) values below 0.01 ppm when compared to the corresponding peaks in Table 1.

34. Composition of any of claims 22-33, wherein said composition comprises the MYDGF protein with a monomer content of more than 95%, 96%, 97%, 98% or 99%.

35. Use of a composition of any of claims 22-34 for the preparation of a pharmaceutical composition.

36. Pharmaceutical composition comprising a composition of any of claims 22-34.

37. Pharmaceutical composition of claim 36, further comprising a pharmaceutically acceptable carrier.

38. Pharmaceutical composition of claim 36 or 37, wherein said composition is formulated for parental administration.

39. Pharmaceutical composition of claim 38, wherein said composition is formulated for intravenous, intraarterial or intracoronary administration.

40. Pharmaceutical composition of claim 39, wherein said composition is formulated for intravenous administration.

41. Composition of any of claims 22-34 or pharmaceutical composition of any of claims 36-40 for use as a medicament.

42. Use of the composition of any of claims 22-34 or pharmaceutical composition of any of claims 36-41 in the manufacture of a medicament for

(i) treating or preventing a disease or condition selected from the group consisting of injury, wounding, ischemia, reperfusion injury, trauma, mechanical overload, intoxication, surgery, primary or acquired cardiomyopathy, postischemic contractile dysfunction, myocardial infarction, preferably acute myocardial infarction, angina pectoris, heart failure, inflammation of the heart, heart insufficiency, hypertrophy, and fibrosis;

(ii) promoting or improving heart tissue regeneration, cardiomyocyte proliferation, neovascularisation, heart function or left ventricular systolic function after myocardial infarction;

(iii) protecting cardiomyocyte from death, e.g. through apoptosis or necrosis; or

(iv) decreasing infarct size after myocardial infarction, preferably acute myocardial infarction.

43. The use of claim 44 or 45, wherein said cardiomyopathy is inherited cardiomyopathy or cardiomyopathy caused by spontaneous mutations.

44. The use of claim 42, wherein said cardiomyopathy is acquired cardiomyopathy, preferably ischemic cardiomyopathy caused by atherosclerotic or other coronary artery diseases, cardiomyopathy caused by infection or intoxication of the myocardium, hypertensive heart disease caused by pulmonary arterial hypertension and/or arterial hypertension and diseases of the heart valves.

45. The use of claim 42, wherein said cardiomyopathy is selected from the group consisting of hypertrophic cardiomyopathy (HCM or HOCM), arrythmogenic right ventricular cardiomyopathy (ARVC), isolated ventricular non-compaction mitochondrial myopathy, dilated cardiomyopathy (DCM), restrictive cardiomyopathy (RCM), Takotsubo cardiomyopathy, Loeffler endocarditis, diabetic cardiomyopathy, alcoholic cardiomyopathy, or obesity-associated cardiomyopathy.

46. The use of claim 42, wherein said heart failure is chronic heart failure.

47. The use of claim 46, wherein said heart failure or chronic heart failure is heart failure with preserved ejection fraction (HFpEF), heart failure with reduced ejection fraction (HFrEF), or heart failure with mid-range ejection fraction (HFmrEF).