from >60°C to 80°C with hyperthermophiles capable of optimal growth above 80°C), for example, microorganisms that are used in industrial fermentations or for *in vitro* laboratory processes conducted at elevated temperatures.

To date there is no experimental evidence for active Cas9 proteins in thermophiles. Based on a comparative genome screening by Chylinski *et al*. (2014; Nucleic Acids Research 42: 6091-61-05) on the presence of Cas9 in bacteria it was found that the Type II-C CRISPR-Cas system is only present in approximately 3.3% of all bacterial genomes. Among thermophilic bacteria, the Type II system is underrepresented based on statistical analysis (P=0.0019). In addition, no Type II system has been found in archaea however, this could possibly be due to the absence of the RNase III protein (involved in the Type II system) in archaea. Chylinski, *et al*., (2014; Nucleic Acids Research 42: 6091-6105) did describe the classification and evolution of type II CRISPR-Cas systems, in particular, two species are identified which exhibit these systems, however these species grow maximally at 55°C and do not exhibit strictly thermophilic growth with optimum growth temperature 60-80°C, with hyperthermophiles capable of growing optimally above 80°C.

Despite the rarity of the CRISPR-Cas system in bacterial genomes and in particular the fact that Cas9 has been found only in bacteria (not archaea) with optimal growth temperatures below 45°C, the inventors have surprisingly discovered several thermostable Cas9 variants which enable genome editing to be carried out at elevated temperatures. These Cas9 nucleases provide novel tools for genetic engineering at elevated temperatures and are of particular value in the genetic manipulation of thermophilic organisms; particularly microorganisms.

SUMMARY OF THE INVENTION

In one aspect, the present disclosure provides an isolated Cas protein having an amino acid sequence of 89% - 99.8% identity to SEQ ID NO: 1, or a polypeptide fragment thereof, wherein the Cas protein or polypeptide fragment is capable of cleaving DNA comprising a target nucleic acid sequence at a temperature between 20°C and 100°C, when associated with at least one RNA molecule which recognizes the target sequence, for cleaving or modifying genetic material of an organism.

In another aspect, the present disclosure provides an isolated Cas protein comprising an amino acid sequence of least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof, wherein one or more native nuclease activity sites of the Cas protein or polypeptide fragment are inactivated for modifying genetic material of an organism.

In another aspect, the present disclosure provides a ribonucleoprotein complex comprising a Cas protein or polypeptide as disclosed herein and at least one targeting RNA molecule which recognises a sequence in a target polynucleotidef, wherein the ribonucleoprotein complex is capable of cleaving DNA comprising a target nucleic acid sequence at a temperature of 65°C, for cleaving or modifying genetic material of an organism.

In another aspect, the present disclosure provides a ribonucleoprotein complex comprising a Cas protein and at least one targeting RNA molecule which recognises a sequence in a target polynucleotide; wherein the Cas protein has an amino acid sequence of at least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof, wherein the ribonucleoprotein complex is capable of cleaving DNA comprising a target nucleic acid sequence at a temperature of 65°C, for cleaving or modifying genetic material of an organism, wherein the targeting RNA molecule is a single guide RNA (sgRNA).

In another aspect, the present disclosure provides an isolated Cas protein comprising an amino acid sequence of least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof when used for cleaving or modifying genetic material *in vitro*.

In yet another aspect, the present disclosure provides an expression vector comprising a polynucleotide sequence encoding a Cas protein or polypeptide fragment as disclosed herein, or the ribonucleoprotein complex as disclosed herein.

In yet another aspect, the present disclosure provides a method of modifying a target nucleic acid comprising contacting the nucleic acid *in vitro* with:

a. a ribonucleoprotein complex as described herein; or

b. a protein or protein complex as described herein and at least one targeting RNA molecule as described herein; or

c. an isolated Cas protein comprising an amino acid sequence of least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof and at least one targeting RNA molecule as disclosed herein.

In yet another aspect, the present disclosure provides an *in vitro* method of modifying a target nucleic acid in a cell, comprising transforming, transfecting or transducing the cell with an expression vector as described herein having a polynucleotide sequence encoding a Cas protein or polypeptide, and further comprising a nucleotide sequence encoding at least one targeting RNA molecule; or alternatively transforming, transfecting or transducing the cell with an expression vector as described herein and a further expression vector comprising a nucleotide sequence encoding a targeting RNA molecule as described herein; or alternatively transforming, transfecting or transducing the cell with an expression vector comprising a polynucleotide sequence encoding a clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) protein having an amino acid sequence of at least 89% identity to SEQ ID NO: 1; or a polypeptide fragment thereof and a further expression vector comprising a nucleotide sequence encoding a targeting RNA molecule as disclosed herein.

In yet another aspect, the present disclosure provides an *in vitro* method of modifying a target nucleic acid in a cell comprising transforming, transfecting or transducing the cell with an expression vector as described herein, or with an expression vector comprising a polynucleotide sequence encoding a clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) protein having an amino acid sequence of at least 89% identity to SEQ ID NO: 1; or a

polypeptide fragment thereof, and then delivering a targeting RNA molecule as described herein to, or into the cell.

In yet another aspect, the present disclosure provides an *in vitro* method of silencing gene expression at a desired locus according to any of the methods as described herein.

In yet another aspect, the present disclosure provides an *in vitro* method of modifying or deleting and/or inserting a desired nucleotide sequence at a desired location according to any of the methods as described herein.

In yet another aspect, the present disclosure provides a host cell comprising polynucleotide sequence encoding a Cas protein or polypeptide fragment as disclosed herein, or the ribonucleoprotein complex as disclosed herein or the expression vector as disclosed herein, transformed by a method as disclosed herein.

DESCRIPTION

Accordingly, the present invention provides an isolated clustered regularly interspaced short palindromic repeat (CRISpPR)-associated (Cas) protein or polypeptide comprising;

a. the amino acid motif EKDGKYYC [SEQ ID NO: 2]; and/or
b. the amino acid motif $X_1X_2CTX_3X_4$ [SEQ ID NO: 3] wherein X_1 is independently selected from Isoleucine, Methionine or Proline, X_2 is

CLAIMS

1. An isolated Cas protein having an amino acid sequence of 89% - 99.8% identity to SEQ ID NO: 1, or a polypeptide fragment thereof, wherein the Cas protein or polypeptide fragment is capable of cleaving DNA comprising a target nucleic acid sequence at a temperature between 20°C and 100°C, when associated with at least one RNA molecule which recognizes the target sequence, for cleaving or modifying genetic material of an organism.

2. An isolated Cas protein comprising an amino acid sequence of least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof, wherein one or more native nuclease activity sites of the Cas protein or polypeptide fragment are inactivated for modifying genetic material of an organism.

3. A Cas protein or polypeptide fragment as claimed in claim 1, wherein the Cas protein or fragment is capable of nucleic acid cleavage between 50°C and 75°C, preferably above 60°C; more preferably between 60°C and 80°C; even more preferably between 60°C and 65°C.

4. A Cas protein or polypeptide fragment as claimed in claim 1 or claim 3, wherein the nucleic acid cleavage is DNA cleavage.

5. A Cas protein or polypeptide fragment as claimed in any preceding claim, wherein the Cas protein is obtainable from a bacterium, archaeon or virus; preferably wherein the Cas protein is obtainable from *Geobacillus* sp., more preferably from *Geobacillus thermodenitrificans*.

6. A ribonucleoprotein complex comprising a Cas protein or polypeptide of any of claims 1 to 5 and at least one targeting RNA molecule which recognises a sequence in a target polynucleotide, wherein the ribonucleoprotein complex is capable of cleaving DNA comprising a target nucleic acid sequence at a temperature of 65°C, for cleaving or modifying genetic material of an organism.

7. A ribonucleoprotein complex as claimed in claim 6, wherein the targeting RNA molecule is a single guide RNA (sgRNA).

8. A ribonucleoprotein complex as claimed in claim 6 or claim 7, wherein the length of the at least one RNA molecule is in the range 35 – 135 nucleotide residues.

9. A ribonucleoprotein complex as claimed in claim 6 or claim 7, wherein the target sequence is 31 or 32 nucleotide residues in length.

10. A ribonucleoprotein complex comprising a Cas protein and at least one targeting RNA molecule which recognises a sequence in a target polynucleotide; wherein the Cas protein has an amino acid sequence of at least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof, wherein the ribonucleoprotein complex is capable of cleaving DNA comprising a target nucleic acid sequence at a temperature of 65°C, for cleaving or modifying genetic material of an organism, wherein the targeting RNA molecule is a single guide RNA (sgRNA).

11. A Cas protein or polypeptide fragment as set forth in any of claims 1 to 5 or a ribonucleoprotein complex as set forth in any of claims 6 to 10, wherein the protein or polypeptide is provided as part of a protein complex comprising at least one further functional or non-functional protein.

12. A Cas protein, polypeptide fragment, or ribonucleoprotein complex as claimed in claim 11, wherein the Cas protein or polypeptide, and/or the at least one further protein further comprise at least one functional moiety.

13. A Cas protein, polypeptide fragment as claimed in any of claims 1, 2, 5, 11 or 12, or a ribonucleoprotein complex as claimed in any of claims 6 to 12, wherein the native nuclease activity of the Cas protein or polypeptide is inactive and the Cas protein or polypeptide is fused to at least one functional moiety.

14. A Cas protein or polypeptide fragment, or ribonucleoprotein complex as claimed in any of claims 11 to 13, wherein the at least one functional moiety is fused or linked to the N-terminus and/or the C-terminus of the Cas protein, polypeptide or ribonucleoprotein complex; preferably the N-terminus.

15. A Cas protein or polypeptide fragment, or a ribonucleoprotein complex as claimed in any of claims 11 to 14, wherein the at least one functional moiety is a protein; optionally selected from a helicase, a nuclease, a helicase-nuclease, a DNA methylase, a histone methylase, an acetylase, a phosphatase, a kinase, a transcription (co-)activator, a transcription repressor, a DNA binding protein, a DNA structuring protein, a marker protein, a reporter protein, a fluorescent protein, a ligand binding protein, a signal peptide, a subcellular localisation sequence, an antibody epitope or an affinity purification tag.

16. A Cas protein or polypeptide, or a ribonucleoprotein complex as claimed in claim 15, wherein the native nuclease activity of the Cas protein or polypeptide is inactive and the at least one functional moiety is a nuclease domain; preferably a FokI nuclease domain.

17. A Cas protein or polypeptide, or a ribonucleoprotein complex as claimed in claim 15, wherein the native nuclease activity of the Cas protein or polypeptide is inactive and the at least one functional moiety is a marker protein.

18. An isolated Cas protein or polypeptide fragment as claimed in any of claims 1 to 5 or 11 to 17, or the ribonucleoprotein complex as claimed in any of claims 6 to 17 when used for cleaving or modifying genetic material *in vitro*.

19. An isolated Cas protein comprising an amino acid sequence of least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof when used for cleaving or modifying genetic material *in vitro*.

20. An expression vector comprising a polynucleotide sequence encoding a Cas protein or polypeptide fragment as defined in any of claims 1 to 5 or 11 to 17, or the ribonucleoprotein complex as defined in any of claims 6 to 17.

21. A method of modifying a target nucleic acid comprising contacting the nucleic acid *in vitro* with:

a. a ribonucleoprotein complex as set forth in any of claims 6 to 17; or

b. a protein or protein complex of any of claims 1 to 17 and at least one targeting RNA molecule as defined in any of claims 6 to 9; or

c. an isolated Cas protein comprising an amino acid sequence of least 89% identity to SEQ ID NO: 1, or a polypeptide fragment thereof and at least one targeting RNA molecule as defined in any of claims 6 to 9.

22. An *in vitro* method of modifying a target nucleic acid in a cell, comprising transforming, transfecting or transducing the cell with an expression vector of claim

20 having a polynucleotide sequence encoding a Cas protein or polypeptide, and further comprising a nucleotide sequence encoding at least one targeting RNA molecule; or alternatively transforming, transfecting or transducing the cell with an expression vector of claim 20 and a further expression vector comprising a nucleotide sequence encoding a targeting RNA molecule as defined in any of claims 6 to 9; or alternatively transforming, transfecting or transducing the cell with an expression vector comprising a polynucleotide sequence encoding a clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) protein having an amino acid sequence of at least 89% identity to SEQ ID NO: 1; or a polypeptide fragment thereof and a further expression vector comprising a nucleotide sequence encoding a targeting RNA molecule as defined in any of claims 6 to 9.

23. An *in vitro* method of modifying a target nucleic acid in a cell comprising transforming, transfecting or transducing the cell with an expression vector of claim 20 or with an expression vector comprising a polynucleotide sequence encoding a clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) protein having an amino acid sequence of at least 89% identity to SEQ ID NO: 1; or a polypeptide fragment thereof, and then delivering a targeting RNA molecule as defined in any of claims 6 to 9 to, or into the cell.

24. An *in vitro* method of modifying a target nucleic acid as claimed in any of claims 21 to 23, wherein the at least one functional moiety is a marker protein or reporter protein and the marker protein or reporter protein associates with the target nucleic acid; preferably wherein the marker is a fluorescent protein, for example a green fluorescent protein (GFP).

25. An *in vitro* method as claimed in any of claims 21 to 24, wherein the target nucleic acid is DNA; preferably dsDNA.

26. An *in vitro* method as claimed in any of claims 21 to 24, wherein the target nucleic acid is RNA.

27. An *in vitro* method of modifying a target nucleic acid as claimed in claim 25, wherein the nucleic acid is dsDNA, the at least one functional moiety is a nuclease or a helicase-nuclease, and the modification is a single-stranded or a double-stranded break at a desired locus.

28. An *in vitro* method of silencing gene expression at a desired locus according to any of the methods as claimed in claims 22, 23, 25 or 27.

29. An *in vitro* method of modifying or deleting and/or inserting a desired nucleotide sequence at a desired location according to any of the methods as claimed in claims 22, 23, 25 or 27.

30. An *in vitro* method of modifying gene expression in a cell comprising modifying a target nucleic acid sequence as claimed in a method of any of claims 21 to 25; wherein the nucleic acid is dsDNA and the functional moiety is selected from a DNA modifying enzyme (e.g. a methylase or acetylase), a transcription activator or a transcription repressor.

31. An *in vitro* method of modifying gene expression in a cell comprising modifying a target nucleic acid sequence as claimed in a method of claim 26, wherein the nucleic acid is an mRNA and the functional moiety is a ribonuclease; optionally selected from an endonuclease, a 3' exonuclease or a 5' exonuclease.

32. An *in vitro* method of modifying a target nucleic acid as claimed in any of claims 21 to 31, wherein the method is carried out at a temperature between 50°C and 100°C.

33. An *in vitro* method of modifying a target nucleic acid as claimed in claim 32, wherein the method is carried out at a temperature at or above 60°C, preferably between 60°C and 80°C, more preferably between 60°C and 65°C.

34. A method as claimed in any of claims 21 to 33 wherein the cell is a prokaryotic cell.

35. A method as claimed in any of claims 21 to 33 wherein the cell is a eukaryotic cell.

36. A host cell comprising polynucleotide sequence encoding a Cas protein or polypeptide fragment as defined in any of claims 1 to 5 or 10 to 17, or the ribonucleoprotein complex as defined in any of claims 6 to 17 or the expression vector of claim 20, transformed by a method as claimed in any of claims 21 to 34.

37. A host cell as claimed in claim 36, wherein at least one nucleic acid sequence fused to the nucleic acid molecule encoding the Cas protein or polypeptide, encodes a protein selected from a protein selected from a helicase, a nuclease, a helicase-nuclease, a DNA methylase, a histone methylase, an acetylase, a phosphatase, a kinase, a transcription (co-)activator, a transcription repressor, a DNA binding protein, a DNA structuring protein, a marker protein, a reporter protein, a fluorescent protein, a ligand binding protein, a signal peptide, a subcellular localisation sequence, an antibody epitope or an affinity purification tag.

38. A host cell as claimed in claim 36 or claim 37, further comprising a nucleotide sequence encoding at least one targeting RNA molecule.