

[0005] Therefore, there is a need to provide novel nucleobase editors which facilitate C:G to G:C editing (CGBEs) with high efficiency, specificity, and purity.

SUMMARY

[0005a] In one aspect, provided herein is a fusion protein or a protein complex comprising an RNA-guided DNA binding protein (DnaBP), wherein the RNA-guided DnaBP is a Cas9 protein, a nucleobase modifying protein (NMP), wherein the NMP is a cytosine deaminase domain from an apolipoprotein B mRNA-editing complex (APOBEC) family deaminase, and a Base Excision Repair associated protein (BERAP), wherein the BERAP is XRCC1, a DNA Ligase III (LIG3), or a DNA binding or lyase domain of DNA Polymerase beta; wherein the fusion protein or protein complex does not comprise a Uracil binding protein or a catalytically active DNA polymerase.

[0005b] In another aspect, provided herein is a fusion protein comprising:

- a first amino acid sequence as set forth in any one of SEQ ID NOs: 1-2,
- a second amino acid sequence as set forth in SEQ ID NO: 3,
- a third amino acid sequence as set forth in any one of SEQ ID NO: 4-10.

[0005c] In another aspect, provided herein is a protein complex comprising:

- a first protein comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 1-2,
- a second protein comprising an amino acid sequence as set forth in SEQ ID NO: 3,
- a third protein comprising an amino acid sequence as set forth in any one of SEQ ID NO: 4-11.

[0005d] In yet another aspect, provided herein is a protein-nucleic acid complex comprising a nucleic acid molecule and any one of:

- the fusion protein or protein complex as described herein,
- the fusion protein as described herein; and
- the protein complex as described herein.

[0005e] In yet another aspect, provided herein is an *in-vitro* method of replacing a cytosine with a guanine on a DNA strand in a cell, said method comprises introducing to the cell the fusion protein or protein complex as described herein, the fusion protein as described herein, the protein complex as described herein, or the protein-nucleic acid complex as described herein.

[0005f] In another aspect, provided herein is a use of the fusion protein or protein complex as described herein, the fusion protein as described herein, the protein complex as described

herein, the protein-nucleic acid complex as described herein, the pharmaceutical composition as described herein, the polynucleotide as described herein, or the vector as described herein in the manufacture of a medicament for treating a subject having or suspected of having a disease or disorder.

[0006] In one aspect, the present disclosure refers to a fusion protein or a protein complex comprising a DNA binding protein (DnaBP), a nucleobase modifying protein (NMP), and a Base Excision Repair associated protein (BERAP); wherein the fusion protein or protein complex does not comprise a Uracil binding protein or a catalytically active DNA polymerase.

[0007] In another aspect, the present disclosure refers to a fusion protein comprising:

- a first amino acid sequence that is at least 80% identical to the amino acid sequence of any one of SEQ ID NOs: 1-2,
- a second amino acid sequence that is at least 80% identical to the amino acid sequence of SEQ ID NO: 3,
- a third amino acid sequence that is at least 80% identical to the amino acid sequence of SEQ ID NO: 4-10.

[0008] In another aspect, the present disclosure refers to a fusion protein comprising a sequence of any one of SEQ ID NOs: 42 to 72.

[0009] In yet another aspect, the present disclosure refers to protein complex comprising:

- a first protein comprising an amino acid sequence that is at least 80% identical to the amino acid sequence of any one of SEQ ID NOs: 1-2,
- a second protein comprising an amino acid sequence that is at least 80% identical to the amino acid sequence of SEQ ID NO: 3,
- a third protein comprising an amino acid sequence that is at least 80% identical to the amino acid sequence of SEQ ID NO: 4-11.

[0010] In yet another aspect, the present disclosure refers to a protein-nucleic acid complex comprising a nucleic acid molecule and any one of:

- the fusion protein or protein complex as disclosed herein,
- the fusion protein as disclosed herein; and
- the protein complex of claim 20.

[0011] In yet another aspect, the present disclosure refers to a pharmaceutical composition comprising the fusion protein or protein complex as disclosed herein, the fusion protein as disclosed herein, the protein complex as disclosed herein, or the protein-nucleic acid complex as disclosed herein.

Claims

1. A fusion protein or a protein complex comprising
an RNA-guided DNA binding protein (DnaBP), wherein the RNA-guided DnaBP is a Cas9 protein,
a nucleobase modifying protein (NMP), wherein the NMP is a cytosine deaminase domain from an apolipoprotein B mRNA-editing complex (APOBEC) family deaminase, and
a Base Excision Repair associated protein (BERAP), wherein the BERAP is XRCC1, a DNA Ligase III (LIG3), or a DNA binding or lyase domain of DNA Polymerase beta;
wherein the fusion protein or protein complex does not comprise a Uracil binding protein or a catalytically active DNA polymerase.
2. The fusion protein or protein complex according to claim 1, wherein the DNA binding or lyase domain of DNA Polymerase beta corresponds to a region contained within amino acids 1-140, 1-120, 1-100, or 1-87 of a full DNA Polymerase beta sequence (SEQ ID NO: 12 or 13).
3. The fusion protein or protein complex according to claim 1, wherein the Cas9 protein is a Cas nickase (nCas).
4. The fusion protein or protein complex according to claim 1, wherein the Cas9 protein is a nuclease inactive Cas (dCas).
5. The fusion protein according to any of claims 1-4, wherein the orientation of the DnaBP, NMP and BERAP within the fusion protein is selected from the group consisting of: [NMP]-[DnaBP]-[BERAP], [NMP]-[BERAP]-[DnaBP] and [BERAP]-[NMP]-[DnaBP]; wherein each instance of “]-[” comprises an optional linker.
6. The fusion protein according to claim 5, wherein the orientation of the DnaBP, NMP and BERAP within the fusion protein is [NMP]-[DnaBP]-[BERAP].
7. A fusion protein comprising:
 - a first amino acid sequence as set forth in any one of SEQ ID NOs: 1-2,
 - a second amino acid sequence as set forth in SEQ ID NO: 3,
 - a third amino acid sequence as set forth in any one of SEQ ID NO: 4-10.
8. A fusion protein comprising a sequence of any one of SEQ ID NOs: 42 to 72.

9. The fusion protein of claim 8, wherein the sequence is SEQ ID NO: 44 or SEQ ID NO: 69.
10. A protein complex comprising:
- a first protein comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 1-2,
 - a second protein comprising an amino acid sequence as set forth in SEQ ID NO: 3,
 - a third protein comprising an amino acid sequence as set forth in any one of SEQ ID NO: 4-11.
11. A protein-nucleic acid complex comprising a nucleic acid molecule and any one of:
- the fusion protein or protein complex according to any one of claims 1-4,
 - the fusion protein according to any one of claims 5-9; and
 - the protein complex of claim 10.
12. A pharmaceutical composition comprising the fusion protein or protein complex of any one of claims 1-4, the fusion protein according to any one of claims 5-9, the protein complex of claim 10, or the protein-nucleic acid complex of claim 11.
13. The pharmaceutical composition of claim 12, further comprising a pharmaceutically acceptable excipient.
14. An *in-vitro* method of replacing a cytosine with a guanine on a DNA strand in a cell, said method comprises introducing to the cell the fusion protein or protein complex of any one of claims 1-4, the fusion protein according to any one of claims 5-9, the protein complex of claim 10, or the protein-nucleic acid complex of claim 11.
15. A polynucleotide encoding the fusion protein or protein complex of any one of claims 1-4, the fusion protein according to any one of claims 5-9, the protein complex of claim 10, or the protein-nucleic acid complex of claim 11.
16. A vector comprising the polynucleotide of claim 15.
17. A cell comprising the fusion protein or protein complex of any one of claims 1-4, the fusion protein according to any one of claims 5-9, the protein complex of claim 10, or the protein-nucleic acid complex of claim 11.

18. Use of the fusion protein or protein complex of any one of claims 1-4, the fusion protein according to any one of claims 5-9, the protein complex of claim 10, the protein-nucleic acid complex of claim 11, the pharmaceutical composition of any one of claims 12-13, the polynucleotide of claim 15, or the vector of claim 16 in the manufacture of a medicament for treating a subject having or suspected of having a disease or disorder.

19. The use of claim 18, wherein the disease or disorder comprises one or more C>G and/or one or more G>C mutations.

20. The use of claim 19, wherein the disease or disorder is selected from a group consisting of skin fibrosis, bladder cancer, liver cancer, Myasthenic syndrome, Spondyloepimetaphyseal dysplasia, Parkinson's disease, Deafness, blood disorders, and Schnyder crystalline corneal dystrophy.