CLAIMS

[Claim 1]

A polypeptide, which is any one of the following polypeptides:

(C-1) a polypeptide consisting of any one of the sequences of SEQ ID NOs: 4 to 7;

(C-2) a polypeptide consisting of any one of the sequences of SEQ ID NOs: 4 to 7 having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, and 34, having a cytosine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(C-3) a polypeptide having a sequence identity of at least 70% to any one of the sequences of SEQ ID NOs: 4 to 7, provided that the amino acids at positions 1, 4, 6, and 34 are identical, having a cytosine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(A-1) a polypeptide consisting of the sequence of SEQ ID NO: 8 having a substitution of the amino acid at position 6 with asparagine or aspartic acid;

(A-2) a polypeptide consisting of the sequence of (A-1) having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, and 34, having an adenine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(A-3) a polypeptide having a sequence identity of at least 70% to the sequence of (A-1), provided that the amino acids at positions 1, 4, 6, and 34 are identical, having an adenine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(G-1) a polypeptide consisting of the sequence of SEQ ID NO: 9 having a substitution of the amino acid at position 6 with asparagine or aspartic acid;

(G-2) a polypeptide consisting of the sequence of (G-1) having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, and 34, having a guanine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(G-3) a polypeptide having a sequence identity of at least 70% to the sequence of (G-1), provided that the amino acids at positions 1, 4, 6, and 34 are identical, having a guanine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(U-1) a polypeptide consisting of the sequence of SEQ ID NO: 10 having a substitution of the amino acid at position 6 with asparagine or aspartic acid;

(U-2) a polypeptide consisting of the sequence of (U-1) having a substitution, deletion, or

addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, and 34, having a uracil-binding property, and being a PPR motif represented by the formula 1 mentioned below; and

(U-3) a polypeptide having a sequence identity of at least 70% to the sequence of (U-1), provided that the amino acids at positions 1, 4, 6, and 34 are identical, having a uracil-binding property, and being a PPR motif represented by the formula 1 mentioned below.

[Formula 1]

$$\text{(Helix A)-X-(Helix B)-L} \quad \text{(Formula 1)}$$

wherein, in the formula:

Helix A is a moiety of 12-amino acid length capable of forming an α-helix structure, and is represented by the formula 2;

[Formula 2]

$$A_1\text{-}A_2\text{-}A_3\text{-}A_4\text{-}A_5\text{-}A_6\text{-}A_7\text{-}A_8\text{-}A_9\text{-}A_{10}\text{-}A_{11}\text{-}A_{12} \quad \text{(Formula 2)}$$

wherein, in the formula 2, A_1 to A_{12} independently represent an amino acid;

X does not exist, or is a moiety of 1- to 9-amino acid length;

Helix B is a moiety of 11- to 13-amino acid length capable of forming an α-helix structure; and

L is a moiety of 2- to 7-amino acid length represented by the formula 3;

[Formula 3]

$$L_{vii}\text{-}L_{vi}\text{-}L_{v}\text{-}L_{iv}\text{-}L_{iii}\text{-}L_{ii}\text{-}L_{i} \quad \text{(Formula 3)}$$

wherein, in the formula 3, the amino acids are numbered "i" (-1), "ii" (-2), and so on from the C-terminus side,

provided that L_{iii} to L_{vii} may not exist.

[Claim 2]

A polypeptide, which is any one of the following PPR motifs:

(C-1) a polypeptide consisting of any one of the sequences of SEQ ID NOs: 4 to 7;

(C-2) a polypeptide consisting of any one of the sequences of SEQ ID NOs: 4 to 7 having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, 9, and 34, having a cytosine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(C-3) a polypeptide having a sequence identity of at least 70% to any one of the sequences of

SEQ ID NOs: 4 to 7, provided that the amino acids at positions 1, 4, 6, 9, and 34 are identical, having a cytosine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(A-1) a polypeptide consisting of the sequence of SEQ ID NO: 8 having such substitutions of the amino acids at positions 6 and 9 that any one of the combinations defined below is satisfied;

(A-2) a polypeptide consisting of the sequence of (A-1) having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, 9, and 34, having an adenine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(A-3) a polypeptide having a sequence identity of at least 70% to the sequence of (A-1), provided that the amino acids at positions 1, 4, 6, 9, and 34 are identical, having an adenine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(G-1) a polypeptide consisting of the sequence of SEQ ID NO: 9 having such substitutions of the amino acids at positions 6 and 9 that any one of the combinations defined below is satisfied;

(G-2) a polypeptide consisting of the sequence of (G-1) having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, 9, and 34, having a guanine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(G-3) a polypeptide having a sequence identity of at least 70% to the sequence of (G-1), provided that the amino acids at positions 1, 4, 6, 9, and 34 are identical, having a guanine-binding property, and being a PPR motif represented by the formula 1 mentioned below;

(U-1) a polypeptide consisting of the sequence of SEQ ID NO: 10 having such substitutions of the amino acids at positions 6 and 9 that any one of the combinations defined below is satisfied;

(U-2) a polypeptide consisting of the sequence of (U-1) having a substitution, deletion, or addition of 1 to 9 amino acids other than the amino acids at positions 1, 4, 6, 9, and 34, having a uracil-binding property, and being a PPR motif represented by the formula 1 mentioned below; and

(U-3) a polypeptide having a sequence identity of at least 70% to the sequence of (U-1), provided that the amino acids at positions 1, 4, 6, 9, and 34 are identical, ahaving a uracil-binding property, and being a PPR motif represented by the formula 1 mentioned below.

- a combination of asparagine as the amino acid at position 6 and glutamic acid as the amino

acid at position 9,

- a combination of asparagine as the amino acid at position 6 and glutamine as the amino acid at position 9,

- a combination of asparagine as the amino acid at position 6 and lysine as the amino acid at position 9, and

- a combination of aspartic acid as the amino acid at position 6 and glycine as the amino acid at position 9.

[Formula 4]

(Helix A)-X-(Helix B)-L (Formula 1)

wherein, in the formula:

Helix A is a moiety of 12-amino acid length capable of forming an α-helix structure, and is represented by the formula 2;

[Formula 5]

A_1-A_2-A_3-A_4-A_5-A_6-A_7-A_8-A_9-A_{10}-A_{11}-A_{12} (Formula 2)

wherein, in the formula 2, A_1 to A_{12} independently represent an amino acid;

X does not exist, or is a moiety of 1- to 9-amino acid length;

Helix B is a moiety of 11- to 13-amino acid length capable of forming an α-helix structure; and

L is a moiety of 2- to 7-amino acid length represented by the formula 3;

[Formula 6]

L_{vii}-L_{vi}-L_v-L_{iv}-L_{iii}-L_{ii}-L_i (Formula 3)

wherein, in the formula 3, the amino acids are numbered "i" (-1), "ii" (-2), and so on from the C-terminus side,

provided that L_{iii} to L_{vii} may not exist.

[Claim 3]

The polypeptide according to claim 1 or 2, which is any one of the following PPR motifs:

(C-4) a polypeptide consisting of the sequence of SEQ ID NO: 4;

(A-4) a polypeptide consisting of the sequence of SEQ ID NO: 58;

(G-4) a polypeptide consisting of the sequence of SEQ ID NO: 59; and

(U-4) a polypeptide consisting of the sequence of SEQ ID NO: 60.

[Claim 4]

Use of the polypeptide according to any one of claims 1 to 3 in a PPR protein as the first PPR motif from the N-terminus.

[Claim 5]

The use according to claim 4, which is for reducing aggregation of the PPR protein.

[Claim 6]

A protein capable of binding to a target nucleic acid having a specific nucleotide sequence, which comprises 1 to 30 of PPR motifs represented by the formula 1 mentioned below, and

wherein the A_6 amino acid of the first PPR motif (M_1) from the N-terminus is a hydrophilic amino acid; or

the polypeptide according to any one of claims 1 and 3:

[Formula 7]

$$\text{(Helix A)-X-(Helix B)-L} \quad \text{(Formula 1)}$$

wherein, in the formula:

Helix A is a moiety of 12-amino acid length capable of forming an α-helix structure, and is represented by the formula 2;

[Formula 8]

$$A_1\text{-}A_2\text{-}A_3\text{-}A_4\text{-}A_5\text{-}A_6\text{-}A_7\text{-}A_8\text{-}A_9\text{-}A_{10}\text{-}A_{11}\text{-}A_{12} \quad \text{(Formula 2)}$$

wherein, in the formula 2, A_1 to A_{12} independently represent an amino acid;

X does not exist, or is a moiety of 1- to 9-amino acid length;

Helix B is a moiety of 11- to 13-amino acid length capable of forming an α-helix structure; and

L is a moiety of 2- to 7-amino acid length represented by the formula 3;

[Formula 9]

$$L_{vii}\text{-}L_{vi}\text{-}L_{v}\text{-}L_{iv}\text{-}L_{iii}\text{-}L_{ii}\text{-}L_{i} \quad \text{(Formula 3)}$$

wherein, in the formula 3, the amino acids are numbered "i" (-1), "ii" (-2), and so on from the C-terminus side,

provided that L_{iii} to L_{vii} may not exist.

[Claim 7]

The protein according to claim 6, wherein the A_9 amino acid of M_1 is a hydrophilic amino acid or glycine.

[Claim 8]

The protein according to claim 6 or 7, wherein the A_6 amino acid of M_1 is asparagine or aspartic acid.

[Claim 9]

The protein according to any one of claims 6 to 8, wherein the A_9 amino acid of M_1 is glutamine, glutamic acid, lysine, or glycine.

[Claim 10]

The protein according to any one of claims 6 to 9, wherein the A_6 amino acid of M_1 and the A_9 amino acid of M_1 correspond to any of the following combinations:
- combination of asparagine as the A_6 amino acid and glutamic acid as the A_9 amino acid,
- combination of asparagine as the A_6 amino acid and glutamine as the A_9 amino acid,
- combination of asparagine as the A_6 amino acid and lysine as the A_9 amino acid, and
- combination of aspartic acid as the A_6 amino acid and glycine as the A_9 amino acid.

[Claim 11]

A fusion protein of

at least one selected from the group consisting of a fluorescent protein, a nuclear localization signal peptide, and a tag protein, and

a PPR protein containing the polypeptide according to any one of claims 1 to 3 as the first PPR motif from the N-terminus, or the protein according to any one of claims 6 to 10.

[Claim 12]

A method for reducing aggregation of a PPR protein containing the PPR motif represented by the formula 1 mentioned below and capable of binding to a target nucleic acid having a specific nucleotide sequence, which comprises making the A_6 amino acid of the first PPR motif (M_1) from the N-terminus a hydrophilic amino acid.

[Claim 13]

A method for detecting a nucleic acid, which uses the protein according to any one of claims 6 to 10, or the fusion protein according to claim 11.

[Claim 14]

A nucleic acid encoding the polypeptide according to any one of claims 1 to 3, or the protein according to any one of claims 6 to 10.

[Claim 15]

A vector comprising the nucleic acid according to claim 14.

[Claim 16]

A cell (except for human individual) containing the vector according to claim 15.

[Claim 17]

A method for manipulating a nucleic acid, which uses the polypeptide according to any one of claims 1 to 3, the protein according to any one of claims 6 to 10, or the vector according to claim 15 (implementation in human individual is excluded).

[Claim 18]

A method for producing an organism, which comprises the manipulation method according to claim 17.