

Myotonic Dystrophy Type 1 Therapeutic Drug

Technical Field

[0001]

The present invention relates to a pharmaceutical composition or a method for treating myotonic dystrophy type 1.

Background Art

[0002]

Myotonic Dystrophy type 1 (hereinafter sometimes abbreviated as "DM1") is the most common muscle disease in adults. DM1 is an autosome-dominant multi-organ disorder that influences the skeletal muscle, the smooth muscle, as well as the eyeball, the heart, the endocrine system, and the central nervous system, and its symptoms are diverse, such as muscular atrophy, muscular weakness, myotonia, cataract, insulin resistance, hypogonadism, cardiac conduction disorder, anterior baldness, and intellectual impairment. The cause of DM1 is abnormal elongation of the CTG repeat sequence in the 3' UTR (untranslated region on 3' side) of the myotonin protein kinase (DMPK) gene. Splicing abnormality is evoked when a particular splicing control factor binds to a transcription product (mRNA) having an abnormally elongated CUG repeat sequence, resulting in deficiency of splicing control factors required for normal splicing.

[0003]

Although development of therapeutic drugs for DM1 is currently brought forward with various approaches such as low molecular weight compounds and nucleic acid medicines (such as Non-Patent Literatures 1 and 2), there is no fundamental therapeutic drug at this stage, and treatment of DM1 is confined mainly to palliative therapy.

Citation List

[0004]

[Non-Patent Literature 1] Nakamori et al. Ann Clin Transl Neurol. (2015)

[Non-Patent Literature 2] Overby et al. Drug Discov. Today (2018)

Summary of the Invention

Problems to be Solved by the Invention

[0005]

The object of the present invention is to provide a pharmaceutical composition or a method that is effective for treating myotonic dystrophy type 1 (DM1).

Means for Solving the Problems

[0006]

Pentatricopeptide repeat (PPR) protein is a protein that comprises a repeat of PPR motif that is about 35 amino acids long, and it is known that one PPR motif specifically or selectively binds to one base. In recent years, by virtue of the amino acid that functions when the PPR motif exerts the RNA binding property being identified, and the relationship between the PPR motif structure and the target base being clarified, development of RNA editing/modifying technology that utilizes the PPR protein is under way (such as WO 2013/058404).

[0007]

As a result of searching for novel compounds/substances that may treat DM1, the present inventors found that DM1 may be treated by utilizing a modified PPR protein.

[0008]

The present invention is based on the above knowledge, and may have the following characteristics.

[0009]

[1]: A pharmaceutical composition for treating myotonic dystrophy type 1, wherein

said pharmaceutical composition comprises a nucleic acid encoding a protein that specifically binds to the CUG repeat sequence,

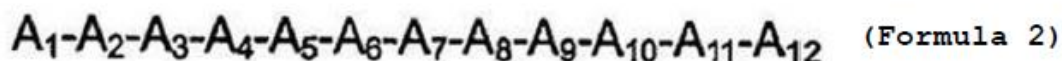
said protein that specifically binds to the CUG repeat sequence comprises at least 6 pentatricopeptide repeat (PPR) motifs consisting of polypeptides of 30 - 38 amino acids long represented by Formula 1,

[Formula 1]

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

[wherein in Formula 1:

Helix A is a 12 amino acids long moiety that is capable of forming an α -helix structure, and is represented by Formula 2,
[Formula 2]



wherein in Formula 2, $A_1 - A_{12}$ each independently represents an amino acid;

X is absent, or is a moiety consisting of a length of 1 - 9 amino acids;

Helix B is a moiety that is capable of forming an α -helix structure consisting of a length of 11 - 13 amino acids;

L is a moiety represented by Formula 3 that is 2 - 7 amino acids long;

[Formula 3]



and wherein in Formula 3, each amino acid is numbered from the C-terminal such as "i" (-1), "ii" (-2),

provided that $L_{iii} - L_{vii}$ may be absent]

and the combination of three amino acids at A_1 , A_4 , and L_{ii} , or the combination of two amino acids at A_4 and L_{ii} in said each PPR motif is selected so that said each PPR motif binds to C, U, or G, and thereby configured so that said protein specifically binds to said CUG repeat sequence.

[0010]

[2]: The pharmaceutical composition according to [1], wherein said protein that specifically binds to the CUG repeat sequence comprises 9 - 30 of said PPR motifs.

[0011]

[3]: The pharmaceutical composition according to [2], wherein said protein that specifically binds to the CUG repeat sequence comprises 12 - 24 of said PPR motifs.

[0012]

[4]: The pharmaceutical composition according to any of [1] to [3], wherein the combination of three amino acids at A₁, A₄, and L_{ii} in said each PPR motif is:

when the base to be the target for the PPR motif is A (adenine), the combination of three amino acids at A₁, A₄, and L_{ii} is, in the order of (A₁, A₄, L_{ii}), (valine, threonine, asparagine), (phenylalanine, serine, asparagine), (phenylalanine, threonine, asparagine), (isoleucine, asparagine, aspartic acid), or (threonine, threonine, asparagine);

when the base to be the target for the PPR motif is G (guanine), the combination of three amino acids at A₁, A₄, and L_{ii} is, in the order of (A₁, A₄, L_{ii}), (glutamic acid, glycine, aspartic acid), (valine, threonine, aspartic acid), (lysine, threonine, aspartic acid), or (leucine, threonine, aspartic acid);

when the base to be the target for the PPR motif is U (uracil), the combination of three amino acids at A₁, A₄, and L_{ii} is, in the order of (A₁, A₄, L_{ii}), (valine, asparagine, aspartic acid), (isoleucine, asparagine, asparagine), (isoleucine, asparagine, aspartic acid), (isoleucine, methionine, aspartic acid), (phenylalanine, proline, aspartic acid), or (tyrosine, proline, aspartic acid); or

when the base to be the target for the PPR motif is C (cytosine), the combination of three amino acids at A₁, A₄, and L_{ii} is, in the order of (A₁, A₄, L_{ii}), (valine, asparagine, asparagine), (isoleucine, asparagine, asparagine), (valine, asparagine, serine), or (isoleucine, methionine, aspartic acid).

[0013]

[5]: The pharmaceutical composition according to any of [1] to [3], wherein the combination of two amino acids at A₄ and L_{ii} in said each PPR motif is:

when the base to be the target for the PPR motif is A (adenine), the combination of two amino acids at A₄ and L_{ii} is, in the order of

(A₄, L_{ii}), (threonine, asparagine), (serine, asparagine), or (glycine, asparagine);

when the base to be the target for the PPR motif is G (guanine), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (threonine, aspartic acid), or (glycine, aspartic acid);

when the base to be the target for the PPR motif is U (uracil), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (asparagine, aspartic acid), (proline, aspartic acid), (methionine, aspartic acid), or (valine, threonine);

when the base to be the target for the PPR motif is C (cytosine), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (asparagine, asparagine), (asparagine, serine), or (leucine, aspartic acid).

[0014]

[6]: The pharmaceutical composition according to any of [1] to [5], characterized in that said nucleic acid encoding a protein that specifically binds to the CUG repeat sequence is integrated into an expression vector.

[0015]

[7]: The pharmaceutical composition according to any of [1] to [5], characterized in that said nucleic acid encoding a protein that specifically binds to the CUG repeat sequence is integrated into a viral vector.

[0016]

[8]: The pharmaceutical composition according to [7], characterized in that said viral vector is a viral vector with tropism towards muscle tissue.

[0017]

[9]: The pharmaceutical composition according to [8], characterized in that said viral vector is an adeno-associated virus (AAV) vector, an adenovirus vector, a retrovirus vector, a lentivirus vector, or a herpes simplex virus vector.

[0018]

[10]: The pharmaceutical composition according to [9], characterized in that said viral vector is an AAV vector.

[0019]

[11]: The pharmaceutical composition according to [10], characterized in that said AAV vector is an AAV1 vector, an AAV2 vector, an AAV6 vector, an AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV10 vector, an AAV11 vector, or an AAV12 vector.

[0020]

[12]: An expression vector comprising the nucleic acid encoding the protein that specifically binds to the CUG repeat sequence as defined in [1].

[0021]

[13]: A cell comprising the expression vector according to [12].

[0022]

[14]: A protein that specifically binds to the CUG repeat sequence that is produced from the cell according to [13].

[0023]

[15]: A viral expression vector comprising the nucleic acid encoding the protein that specifically binds to the CUG repeat sequence as defined in [1].

[0024]

[16]: A cell comprising the viral expression vector according to [15].

[0025]

[17]: A viral vector that is produced from the cell according to [16].

[0026]

[18]: The use of the nucleic acid encoding the protein that specifically binds to the CUG repeat sequence as defined in [1] in the manufacture of myotonic dystrophy type 1 therapeutic medicine.

[0027]

[19]: A method for treating myotonic dystrophy type 1, comprising a step of applying to a subject a therapeutically effective amount of the pharmaceutical composition according to [1].

[0028]

Note that an invention of any combination of one or more of the above characteristics is also encompassed by the scope of the present invention.

Effects of the Invention

[0029]

The symptoms of myotonic dystrophy type 1 (DM1) may be treated or alleviated by the present invention.

Brief Description of the Drawings

[0030]

Figure 1 shows that the PPR protein of the present invention specifically bound to the CUG repeat sequence.

Figure 2 shows that the PPR protein of the present invention suppressed the formation of RNA-foci in DM1 model cells.

Figure 3 shows that the PPR protein of the present invention improved splicing abnormality in DM1 model cells.

Figure 4 shows that the PPR protein of the present invention improved muscle differentiation efficiency in DM1 model animals.

Figure 5 shows that the PPR protein of the present invention suppressed the formation of RNA-foci in DM1 model animals.

Figure 6 shows that the PPR protein of the present invention improved splicing abnormality in DM1 model animals.

Figure 7 shows that the PPR protein of the present invention improved myotonia in DM1 model animals.

Figure 8 shows the relationship between the abnormal splicing improvement effect and the myotonia improvement effect by the present invention.

Figure 9 shows that AAV9-CUG-PPR 1 dose-dependently suppressed the formation of RNA-foci in DM1 model animals. In the Figure, "Ctrl" indicates the PBS administration group, "LD" indicates the AAV9-CUG-PPR 1 low dose administration group, "MD" indicates the AAV9-CUG-PPR 1 medium dose administration group, and "HD" indicates the AAV9-CUG-PPR 1 high dose administration group. The vertical axis shows the RNA-foci positive cell rate.

Figure 10 shows that AAV9-CUG-PPR 1 dose-dependently improved splicing abnormality in DM1 model animals. The left figure shows the results of the abnormal splicing detection system of *Clcn1*, the right figure of *Atp2a1*. In the figures, "Ctrl" indicates the PBS administration group, "LD" indicates the AAV9-CUG-PPR 1 low dose administration group, "MD" indicates the AAV9-CUG-PPR 1 medium dose

administration group, and "HD" indicates the AAV9-CUG-PPR 1 high dose administration group. The vertical axis shows the respective content (%) of the normal splicing isoform.

Figure 11 shows that AAV9-CUG-PPR 1 dose-dependently improved myotonia in DM1 model animals. The vertical axis shows the myotonia score, and the frequency of myotonic discharge decreases in the order of Scores 3, 2, 1, and 0. The frequency of the number of times of myotonic discharge and the severity are correlated. In the Figure, "Ctrl" indicates the PBS administration group, "LD" indicates the AAV9-CUG-PPR 1 low dose administration group, "MD" indicates the AAV9-CUG-PPR 1 medium dose administration group, and "HD" indicates the AAV9-CUG-PPR 1 high dose administration group.

Figure 12 shows the relationship between the administration dosage of AAV9-CUG-PPR 1 and the amount of PPR mRNA expression. GAPDH is used as the internal standard. Each of the individual points show the amount of PPR mRNA expression in the femoral muscle of each mouse individual. The vertical axis is shown in Log representation. In the Figure, "Ctrl" indicates the PBS administration group, "LD" indicates the AAV9-CUG-PPR 1 low dose administration group, "MD" indicates the AAV9-CUG-PPR 1 medium dose administration group, and "HD" indicates the AAV9-CUG-PPR 1 high dose administration group.

Figure 13 shows that AAV9-CUG-PPR 1 time-dependently suppressed the formation of RNA-foci in DM1 model animals. In the Figure, "Ctrl" indicates the PBS administration group. "2 w" "4 w" "8 w," and "16 w" each shows the test duration after administration of AAV9-CUG-PPR 1. The vertical axis shows the RNA-foci positive cell rate.

Figure 14 shows that AAV9-CUG-PPR 1 time-dependently improved splicing abnormality in DM1 model animals. The left figure shows the results of the abnormal splicing detection system of *Clcn1*, the right figure of *Atp2a1*. In the figures, "Ctrl" indicates the PBS administration group. "2 w" "4 w" "8 w," and "16 w" each shows the test duration after administration of AAV9-CUG-PPR 1. The vertical axis shows the respective content of the normal splicing isoform.

Figure 15 shows that AAV9-CUG-PPR 1 time-dependently improved myotonia in DM1 model animals. The vertical axis shows the myotonia

score, and the frequency of myotonic discharge decreases in the order of Scores 3, 2, 1, and 0. The frequency of the number of times of myotonic discharge and the severity are correlated. In the Figure, "Ctrl" indicates the PBS administration group. "2 w" "4 w" "8 w," and "16 w" each shows the test duration after administration of AAV9-CUG-PPR 1.

Figure 16 shows that splicing abnormality is comprehensively normalized after administration of AAV9-CUG-PPR 1. Figure 16a is a figure that shows how splicing events that had a change of 0.2 or more PSI (Percent spliced-in) between WT and PBS-administered HSA-LR mice (left) changed in the PPR-administered HSA-LR mice (right). Figure 16b is a figure that shows change in regard to DM1-related splicing events identified in Tanner et al. (2021). Black circles indicate signals of normal mice, blue rectangles indicate signals of the PBS-administered drug effect mice group, and red triangles indicate signals of the PPR-administered drug effect mice group. Figure 16c shows the result of gene ontology analysis on the genes of each point in Figure a. The P-value indicates the P value of each gene ontology term. The number of genes with exon skipping improvement rate are represented in the 3 stages of greater than 50%, 20 - 50%, and 20% or less for PPR-administered HSA-LR mice.

Figure 17 shows the amount of PPR mRNA expression in each organ of AAV9-CUG-PPR 1-administered mice. The vertical axis shows the relative expression amount when the expression amount of 1×10^{13} vg/kg-administered mice was set as 1 in Log representation. GAPDH mRNA expression amount is the internal standard. The horizontal axis shows the AAV9-CUG-PPR 1 dosage (vg/kg of mouse body weight). The values of each mouse individual are plotted with circles, and the median is connected with a line.

Figure 18 shows the persistency of mRNA expression in each organ of AAV9-CUG-PPR 1-administered mice. The vertical axis presents the relative expression amount when the expression amount of one AAV administration Week 2 (Day 14) mouse individual was set as 1. GAPDH mRNA expression amount is the internal standard. For the horizontal axis, "Ctrl" indicates the PBS group, and "14" "28" "56" "112," and "183" shows the number of days after administration.

The values of each mouse individual are plotted with circles, and the average value is connected with a line.

Description of Embodiments

[0031]

[PPR motif and PPR protein]

"PPR motif " in the present disclosure, unless particularly described, refers to a polypeptide, when analyzing the amino acid sequence with a protein domain search program on the web, composed of 30 - 38 amino acids having an amino acid sequence having an E value less than or equal to a predetermined value (desirably E-03) obtained as PF01535 in Pfam and PS51375 in Prosite. While the position numbers of the amino acids configuring the PPR motif as defined in the present disclosure are almost equivalent to PF01535, they correspond to numbers which are two subtracted from the amino acid positions of PS51375 (e.g. position 1 of the present invention -> position 3 of PS51375), provided that when referring to the amino acid at "ii" (-2) position, it is the second amino acid from the tail end (the C-terminal side) of the amino acids configuring the PPR motif, or two amino acids to the N-terminal side of the first amino acid of the next PPR motif, i.e. the amino acid at position -2. When the next PPR motif is not clearly identified, the amino acid that is two amino acids before the first amino acid of the next helix structure is set at "ii."

[0032]

Although the conserved amino acid sequence of the PPR motif has low conservativeness at the amino acid level, the two α -helices are well conserved in the secondary structure. A typical PPR motif is composed of 35 amino acids, but its length is variable at 30 - 38 amino acids.

[0033]

The basic skeleton of the PPR motif is known to be represented by Formula 1.

[0034]

[Formula 4]

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

[wherein:

Helix A is a 12 amino acids long moiety that is capable of forming an α -helix structure, and is represented by Formula 2,

[0035]

[Formula 5]

A₁-A₂-A₃-A₄-A₅-A₆-A₇-A₈-A₉-A₁₀-A₁₁-A₁₂ (Formula 2)

wherein in Formula 2, A₁ - A₁₂ each independently represents an amino acid;

X is absent, or is a moiety consisting of a length of 1 - 9 amino acids;

Helix B is a moiety that is capable of forming an α -helix structure consisting of a length of 11 - 13 amino acids;

L is a moiety represented by Formula 3 that is 2 - 7 amino acids long;

[0036]

[Formula 6]

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

and wherein in Formula 3, each amino acid is numbered from the C-terminal such as "i" (-1), "ii" (-2),

provided that L_{iii} - L_{vii} may be absent].

[0037]

"PPR protein " in the present disclosure, unless particularly described, refers to a PPR protein having one or more, preferably two or more of the aforementioned PPR motifs. "Protein " in the present disclosure, unless particularly described, refers to a general substance consisting of a polypeptide (a chain having multiple amino acids bound by peptide bonds), and also includes those consisting of polypeptides of relatively low molecular weight. "Amino acid" in the present invention may refer to an ordinary amino

acid molecule, as well as to an amino acid residue that configures the peptide chain. To which it is referring to will be clear to those skilled in the art from the context.

[0038]

"Selective" or "specific" in the present disclosure in regard to the binding ability of the PPR motif to an RNA base, unless particularly described, refers to the fact that the binding activity against any one base of the RNA bases is higher than the binding activity against other bases. This selectivity or specificity can be confirmed by those skilled in the art by planning experiments based on well-known methods, as well as be obtained from calculation by those skilled in the art.

[0039]

RNA base in the present disclosure, unless particularly described, refers to the ribonucleotide base configuring the RNA, and specifically refers to any of adenine (A), guanine (G), cytosine (C), or uracil (U). Note that although the PPR protein may have selectivity against a base in the RNA, it will not bind to the nucleic acid monomer.

[0040]

PPR proteins are abundantly present in plants, and 500 proteins, about 5000 motifs are found in *Arabidopsis thaliana*. PPR motifs and PPR proteins of diverse amino acid sequences are also present in many land plants such as rice, poplar, and selaginella. In the present invention, naturally occurring PPR motifs and PPR proteins may be employed, or PPR motifs and PPR proteins designed based on the method disclosed in for example WO 2013/058404 may be utilized. Specifically, desired PPR motifs and PPR proteins can be designed based on the following information disclosed in WO 2013/058404 .

[0041]

(I) Information related to amino acid positions important for selective binding

The combination of three amino acids at 1, 4, and "ii" (-2) positions (A_1 , A_4 , L_{ii}), or the combination of two amino acids at 4

and "ii" (-2) positions (A_4 and L_{ii}) of the PPR motif is important for selective binding with the RNA base, and which RNA base it binds to can be determined by these combinations.

[0042]

The present invention can utilize the knowledge related to the combination of three amino acids at A_1 , A_4 , and L_{ii} and/or the combination of two amino acids at A_4 and L_{ii} as disclosed in WO 2013/058404.

[0043]

(II) Information related to the correspondence between the combination of three amino acids at A_1 , A_4 , and L_{ii} and the RNA base

(3-1) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is valine, asparagine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, in the next place binding to C, further followed by to A or G.

(3-2) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is valine, threonine, and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, in the next place binding to G, further followed by to C, but not binding to U.

(3-3) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is valine, asparagine, and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to C, in the next place binding to A or U, but not binding to G.

(3-4) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is glutamic acid, glycine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to G, but not binding to A, U, and C.

(3-5) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is isoleucine, asparagine, and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to C, in the next place to U, further followed by binding to A, but not binding to G.

(3-6) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is valine, threonine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to G, in the next place binding to U, but not binding to A and C.

(3-7) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is lysine, threonine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to G, in the next place binding to A, but not binding to U and C.

(3-8) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is phenylalanine, serine, and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, in the next place to C, further followed by binding to G and U.

(3-9) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is valine, asparagine, and serine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to C, in the next place binding to U, but not binding to A and G.

(3-10) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is phenylalanine, threonine, and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, but not binding to G, U, and C.

(3-11) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is isoleucine, asparagine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, in the next place binding to A, but not binding to G and C.

(3-12) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is threonine, threonine, and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, but not binding to G, U, and C.

(3-13) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is isoleucine, methionine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding

strongly to U, in the next place binding to C, but not binding to A and G.

(3-14) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is phenylalanine, proline, and aspartic acid in that order, PPR motifs thereof has selective RNA base binding ability that is binding strongly to U, in the next place binding to C, but not binding to A and G.

(3-15) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is tyrosine, proline, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, but not binding to A, G, and C.

(3-16) When the combination of three amino acids at A_1 , A_4 , and L_{ii} is leucine, threonine, and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to G, but not binding to A, U, and C.

[0044]

(II) Information related to the correspondence between the combination of two amino acids at A_4 and L_{ii} and the RNA base

(2-1) When A_4 and L_{ii} is asparagine and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, in the next place to C, further followed by binding to A and G.

(2-2) When A_4 and L_{ii} is asparagine and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to C, in the next place to U, further followed by binding to A and G.

(2-3) When A_4 and L_{ii} is threonine and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, followed by weakly binding to G, U, and C.

(2-4) When A_4 and L_{ii} is threonine and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to G, followed by weakly binding to A, U, and C.

(2-5) When A_4 and L_{ii} is serine and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, in the next place binding to G, U, and C.

(2-6) When A_4 and L_{ii} is glycine and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to G, in the next place to U, further followed by binding to A, but not binding to C.

(2-7) When A_4 and L_{ii} is asparagine and serine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to C, in the next place to U, further followed by binding to A and G.

(2-8) When A_4 and L_{ii} is proline and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, in the next place binding to G, C, and C, but not binding to A.

(2-9) When A_4 and L_{ii} is glycine and asparagine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to A, in the next place binding to G, but not binding to C and U.

(2-10) When A_4 and L_{ii} is methionine and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, followed by weakly binding to A, G, and C.

(2-11) When A_4 and L_{ii} is leucine and aspartic acid in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to C, in the next place binding to U, but not binding to A and G.

(2-12) When A_4 and L_{ii} is valine and threonine in that order, the PPR motif thereof has selective RNA base binding ability that is binding strongly to U, in the next place binding to A, but not binding to G and C.

[0045]

Note that the above rules mean that the possibility that the PPR motif binds to the target base may be increased with statistical significance by the amino acid combination at (A_1 , A_4 , L_{ii}) or (A_4 and L_{ii}) of the PPR motif, and it does not mean that a PPR motif (or PPR

protein) that binds to the target base (or base sequence) with 100% probability may be manufactured by the above rules. However, those skilled in the art will be able to obtain a PPR motif (or PPR protein) that binds to the target base (or base sequence) with high probability by manufacturing about several to several tens of types of candidate PPR proteins based on the above rules. Note that the selection of the desired PPR protein through manufacture and confirmation of candidate PPR proteins may be performed within the scope of ordinary trial and error by those skilled in the art, and shall not be excessive burden to those skilled in the art.

[0046]

[Utilization of PPR motifs and PPR proteins]

One PPR motif may recognize a particular base of RNA. Further, based on the present invention, by appropriately selecting the amino acids at particular positions, selective PPR motifs can be selected or designed for each of A, U, G, and C. Furthermore, a protein comprising an appropriate series of such PPR motifs may specifically or selectively recognize the corresponding base sequence. Further, a PPR motif that is capable of selectively binding to a desired RNA base, and a protein having multiple PPR motifs capable of sequence-specifically binding to a desired RNA can be designed with the aforementioned knowledge. Upon designing, the sequence information of the native form PPR motif may be referred for portions other than the amino acids at the important positions in the PPR motif. Moreover, designing may be performed by employing the native form PPR motif in general and substituting only the said amino acids at the important positions. The number of repeats of the PPR motif can be set appropriately according to the target sequence, for example it can be two or more, and can be 2 - 30.

[0047]

[PPR protein that specifically binds to the CUG repeat sequence]

The present invention relates to a PPR protein that specifically binds to an mRNA comprising abnormally elongated CUG repeat sequence that is the cause of DM1. The PPR protein according

to the present invention can be produced by linking multiple sets of [a PPR motif that binds to C, a PPR motif that binds to U, a PPR motif that binds to G] based on the aforementioned theory. Note that the first PPR motif of the PPR protein according to the present invention does not necessarily need to be a PPR motif that binds to C, and the first PPR motif may be a PPR motif that binds to U or a PPR motif that binds to G as long as the PPR protein specifically binds to the CUG repeat sequence as a whole. Moreover, naturally, as long as the PPR protein according to the present invention has the configuration to specifically bind to the CUG repeat sequence as a whole, the number of PPR motifs contained in the PPR protein does not need to be a multiple of 3.

[0048]

Non-limiting examples of the C-binding PPR motif that may be employed in the present invention can include PPR motifs having the following sequences.

[0049]

<u>VTYNT</u> LIDGLCKSGKIEEALKLFKEME ^{EE} KGITP <u>SV</u>	(SEQ ID NO. 1)
<u>VTYNT</u> NIDQLCKSGKIEEALKLFKEME ^{EE} KGITP <u>SV</u>	(SEQ ID NO. 2)
<u>VTYNT</u> NIDQLCKSGRIEEALKLFKEME ^{EE} KGITP <u>SV</u>	(SEQ ID NO. 3)
<u>VTYNT</u> LIDGLCKSGRIEEALKLFKEME ^{EE} KGITP <u>SV</u>	(SEQ ID NO. 4)

[0050]

In the present invention, provided that the amino acids at positions A₁, A₄, and L_{ii} of the above sequence or the amino acids at positions A₄ and L_{ii} are conserved, a PPR motif having a sequence homology (or sequence identity) of 80% or more (preferably, 85% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more) with the above sequence, which has the ability to bind to C may be employed.

[0051]

In the present invention, provided that the amino acids at positions A₁, A₄, and L_{ii} of the above sequence or the amino acids at positions A₄ and L_{ii} are conserved, a PPR motif having substitutions, additions, and/or deletions of 7 bases or less (i.e., 1, 2, 3, 4, 5, 6, or 7 bases) to the above sequence, which has the ability to

bind to C may be employed. Note that the amino acid substitution may be for example a conservative substitution of amino acids known in the art.

[0052]

Non-limiting examples of the U-binding PPR motif that may be employed in the present invention can include PPR motifs having the following sequences.

[0053]

VTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDV (SEQ ID NO. 5)

VTYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDV (SEQ ID NO. 6)

[0054]

In the present invention, provided that the amino acids at positions A₁, A₄, and L_{ii} of the above sequence or the amino acids at positions A₄ and L_{ii} are conserved, a PPR motif having a sequence homology (or sequence identity) of 80% or more (preferably, 85% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more) with the above sequence, which has the ability to bind to U may be employed.

[0055]

In the present invention, provided that the amino acids at positions A₁, A₄, and L_{ii} of the above sequence or the amino acids at positions A₄ and L_{ii} are conserved, a PPR motif having substitutions, additions, and/or deletions of 7 bases or less (i.e., 1, 2, 3, 4, 5, 6, or 7 bases) to the above sequence, which has the ability to bind to U may be employed. Note that the amino acid substitution may be for example a conservative substitution of amino acids known in the art.

[0056]

Non-limiting examples of the G-binding PPR motif that may be employed in the present invention can include PPR motifs having the following sequences.

[0057]

<u>V</u> <u>T</u> <u>Y</u> <u>T</u> <u>T</u> <u>L</u> <u>I</u> <u>D</u> <u>G</u> <u>L</u> <u>C</u> <u>K</u> <u>A</u> <u>G</u> <u>K</u> <u>V</u> <u>D</u> <u>E</u> <u>A</u> <u>L</u> <u>E</u> <u>L</u> <u>F</u> <u>D</u> <u>E</u> <u>M</u> <u>K</u> <u>E</u> <u>R</u> <u>G</u> <u>I</u> <u>K</u> <u>P</u> <u>D</u> <u>V</u>	(SEQ ID NO. 7)
<u>V</u> <u>T</u> <u>Y</u> <u>T</u> <u>T</u> <u>L</u> <u>I</u> <u>D</u> <u>G</u> <u>L</u> <u>C</u> <u>K</u> <u>A</u> <u>G</u> <u>K</u> <u>V</u> <u>D</u> <u>E</u> <u>A</u> <u>L</u> <u>E</u> <u>L</u> <u>F</u> <u>D</u> <u>E</u> <u>M</u> <u>K</u> <u>E</u> <u>R</u> <u>G</u> <u>I</u> <u>K</u> <u>P</u> <u>D</u> <u>E</u>	(SEQ ID NO. 8)
<u>V</u> <u>T</u> <u>Y</u> <u>T</u> <u>T</u> <u>L</u> <u>I</u> <u>D</u> <u>G</u> <u>L</u> <u>C</u> <u>K</u> <u>A</u> <u>G</u> <u>R</u> <u>V</u> <u>D</u> <u>E</u> <u>A</u> <u>L</u> <u>E</u> <u>L</u> <u>F</u> <u>D</u> <u>E</u> <u>M</u> <u>K</u> <u>E</u> <u>R</u> <u>G</u> <u>I</u> <u>K</u> <u>P</u> <u>D</u> <u>V</u>	(SEQ ID NO. 9)
<u>V</u> <u>T</u> <u>Y</u> <u>T</u> <u>T</u> <u>L</u> <u>I</u> <u>D</u> <u>G</u> <u>L</u> <u>C</u> <u>K</u> <u>A</u> <u>G</u> <u>R</u> <u>V</u> <u>D</u> <u>E</u> <u>A</u> <u>L</u> <u>E</u> <u>L</u> <u>F</u> <u>D</u> <u>E</u> <u>M</u> <u>K</u> <u>E</u> <u>R</u> <u>G</u> <u>I</u> <u>K</u> <u>P</u> <u>D</u> <u>E</u>	(SEQ ID NO. 10)

[0058]

In the present invention, provided that the amino acids at positions A₁, A₄, and L_{ii} of the above sequence or the amino acids at positions A₄ and L_{ii} are conserved, a PPR motif having a sequence homology (or sequence identity) of 80% or more (preferably, 85% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more) with the above sequence, which has the ability to bind to G may be employed.

[0059]

In the present invention, provided that the amino acids at positions A₁, A₄, and L_{ii} of the above sequence or the amino acids at positions A₄ and L_{ii} are conserved, a PPR motif having substitutions, additions, and/or deletions of 7 bases or less (i.e., 1, 2, 3, 4, 5, 6, or 7 bases) to the above sequence, which has the ability to bind to G may be employed. Note that the amino acid substitution may be for example a conservative substitution of amino acids known in the art.

[0060]

When referring to "identity" herein in relation to a base sequence (sometimes referred to as nucleotide sequence) or an amino acid sequence, unless particularly described, it means the percentage of the number of matched bases or amino acids shared between the two sequences when two sequences are aligned in an optimal aspect. In other words, it can be calculated by identity = (number of matched positions/total number of positions) x 100, and can be calculated with algorithms which are distributed for a charge or for free. Further, such algorithms are integrated within e.g. NBLAST and XBLAST programs described in Altschul et al., J. Mol. Biol. 215 (1990) 403-410. More specifically, search/analysis related to the identity of the base sequence or the amino acid sequence can be performed by algorithms or programs well-known to those skilled in

the art (such as BLASTN, BLASTP, BLASTX, ClustalW). The parameters when employing the programs can be appropriately set by those skilled in the art, and default parameters of each program may also be employed. The specific methods of these analysis methods are also well-known to those skilled in the art.

[0061]

Conservative substitution of amino acids may be for example a substitution to amino acids with similar property. Amino acids with similar property refer to for example an amino acid with similar physical properties such as hydropathy, charge, pKa, and solubility, and for example refer to the following.

- Hydrophobic amino acids: alanine, valine, glycine, isoleucine, leucine, phenylalanine, proline, tryptophan, tyrosine;
- Non-hydrophobic amino acids: arginine, asparagine, aspartic acid, glutamic acid, glutamine, lysine, serine, threonine, cysteine, histidine;
- Hydrophilic amino acids: arginine, asparagine, aspartic acid, glutamic acid, glutamine, lysine, serine, threonine;
- Acidic amino acids: aspartic acid, glutamic acid;
- Basic amino acids: lysine, arginine, histidine;
- Neutral amino acids: alanine, asparagine, cysteine, glutamine, glycine, isoleucine, leucine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine;
- Sulfur-containing amino acids: methionine, cysteine;
- Aromatic ring-containing amino acids: tyrosine, tryptophan, phenylalanine .

[0062]

[Treatment of DM1 by the present invention]

The effects of the present invention are exerted when the PPR protein of the present invention binds to an mRNA comprising an abnormally elongated CUG repeat sequence within the subject cell. The delivery means for the PPR protein into the subject cell is not limited, and a nucleic acid encoding the PPR protein of the present invention may be delivered into the subject cell and allow the expression of the PPR protein of the present invention within the

subject cell, or the PPR protein of the present invention per se may be delivered into the subject cell. The delivery means for the protein or nucleic acid is not limited, and various means known in the art can be employed. Non-limiting examples of the delivery means for the nucleic acid or protein into the subject cell that can be employed in the present invention can include liposomes, lipid nanoparticles (LNP), high molecular weight micelles, emulsions, high molecular weight microcapsules, antibody-nucleic acid complexes, antibody-drug complexes, viruses, and the like.

[0063]

Non-limiting examples of viral vectors that may be employed in the present invention can include adeno-associated virus (AAV) vectors, adenovirus vectors, retrovirus vectors, lentivirus vectors, or herpes simplex virus vectors, and it can be appropriately selected by those skilled in the art depending on the tropism towards cells or tissues, the necessity of integration of nucleic acid into the host genome, and the like.

[0064]

AAV vectors can be particularly favorably employed in the present invention because they have the ability to infect both dividing cells and resting cells, and can transfer genetic material to extremely diverse cell types. Twelve types of AAV serum types (AAV1 - AAV12) have been reported to date, and all of the known serum type can infect various types of tissue cells. The serum type of the AAV vectors that may be employed in the present invention is not limited, and for example an AAV1 vector, an AAV2 vector, an AAV6 vector, an AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV10 vector, an AAV11 vector, or an AAV12 vector that are known to have tropism towards muscle tissue (skeletal muscle, cardiac muscle, and/or smooth muscle) may be employed.

[0065]

Treatment of DM1 employing the present invention may be an aspect of directly applying the present invention to the living body of a subject suffering from DM1 (in vivo), or may be an aspect of treating the cells or tissues outside the living body with the

present invention and then introducing into the living body of a subject (ex vivo).

[0066]

The method for producing an expression vector comprising a nucleic acid encoding the PPR protein of the present invention is not limited, and production can be done by those skilled in the art by conventional means (such as use of commercially available protein expression vector production kits, use of commercially available viral expression vector production kits, and consignment to manufacturing companies).

[0067]

The method for producing the PPR protein of the present invention is not limited, and for example, a desired PPR protein can be obtained by selectively extracting and/or purifying the PPR protein by conventional means from a cell (such as an animal cell, a plant cell, *E. coli*, and yeast) comprising an expression vector (such as a plasmid, a transposon, and a virus) encoding the PPR protein designed by the methods described herein.

[0068]

The terms used herein, except for those that are particularly defined, are employed for describing particular embodiments, and do not intend to limit the invention.

[0069]

Moreover, the term "comprising" as used herein, unless the content clearly indicates to be understood otherwise, intends the presence of the described items (such as components, steps, elements, and numbers), and does not exclude the presence of other items (such as components, steps, elements, and numbers).

[0070]

Unless otherwise defined, all terms used herein (including technical and scientific terms) have the same meanings as those broadly recognized by those skilled in the art of the technology to which the present invention belongs. The terms used herein, unless explicitly defined otherwise, are to be construed as having meanings consistent with the meanings herein and in related technical fields,

and shall not be construed as having idealized or excessively formal meanings.

[0071]

The present invention will now be described in further detail with reference to Examples. However, the present invention can be embodied by various aspects, and shall not be construed as being limited to the Examples described herein.

Examples

[0072]

[Example 1: Production of PPR protein that binds to CUG repeat RNA sequence and binding analysis to RNA]

[0073]

Recombinant proteins were produced for production of PPR protein that binds to the CUG repeat RNA sequence (hereinafter abbreviated as CUG-PPR) and binding analysis thereof, and RNA binding experiment was carried out.

[0074]

<Experimental materials and methods>

(1) Preparation of PPR expression plasmid vector

A gene fused in the order of luciferase, PPR protein, and 3 x hexahistidine tag was cloned into multiple cloning sites of E. coli expression plasmid vector pET-22b. The expression of the fusion gene is controlled by a T7 promoter. Cloning of the gene having the correct size was confirmed by PCR method, and the gene sequence was confirmed by sequencing (SEQ ID NOs. 11 - 14). These were designated Luc-CUG-PPR 1, Luc-CUG-PPR 2, Luc-CUG-PPR 3, and Luc-CUG-PPR 4, and employed in subsequent tests.

[0075]

[Table 1]

[Table 1]

Table 1. Amino acid sequences encoded by the cloned genes

Protein name	Sequence (PPR protein is underlined)
L u c - C U G - P P R 1 (SEQ ID NO. 11)	MAGVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPIQRIVLS GENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVID GVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRV TINGVTGWRLCERILAGGGGSLVPRGSGGGGSSRLHMGNSV <u>VTYNTLIDGLC</u> <u>KSGKIEEALKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGI</u> <u>KPDVVTYTTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSG</u> <u>KIEEALKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPD</u> <u>VVTYTTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEE</u> <u>ALKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVT</u> <u>YTTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEEALK</u> <u>LFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVTYTT</u> <u>LIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEEALKLFK</u> <u>EMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVTYTTLID</u> <u>GLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEEALKLFKEME</u> <u>EKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVTYTTLIDGLC</u> <u>KAGKVDEALELDEMKEGKIPDELTYNTLISGLGKAGRARDPPVLSSGSGS</u> GGGGHHHHHHGR
L u c - C U G - P P R 2 (SEQ ID NO. 12)	MAGVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPIQRIVLS GENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVID GVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRV TINGVTGWRLCERILAGGGGSLVPRGSGGGGSSRLHMGNSV <u>VTYNTNIDQL</u> <u>CKSGKIEEALKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEL</u> <u>GIKPDVVTYTTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCK</u> <u>SGKIEEALKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIK</u> <u>PDVVTYTTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGK</u> <u>IEEALKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDV</u> <u>VVTYTTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEEA</u> <u>LKLFKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVTY</u> <u>TTLIDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEEALKL</u> <u>FKEMEELGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVTYTTL</u> <u>IDGLCKAGKVDEALELDEMKEGKIPDVVTYNTLIDGLCKSGKIEEALKLFKE</u> <u>MEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEELGIKPDVVTYTTLIDG</u> <u>LCKAGKVDEALELDEMKEGKIPDELTYNTLISGLGKAGRARDPPVLSSGSG</u> GSGGGHHHHHHGR

L u c – C U G – P P R 3 (SBQ ID NO. 13)	MAGVFTLEDVFGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPIQRIVLS GENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVID GVTPNMIDYFGRPYEGIAVFDGKKITVTGTWNGNKKIIDERLINPDGSLLFRV TINGVTGWRLCERILAGGGGSLVPRGSGGGGSSRLHMGNSV <u>VTYNTNIDQL</u> <u>CKSGKIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGKLDEAEELLEEMEEK</u> <u>GIKPDVVITYTTLIDGLCKAGKVDEALELFDKMERGIKPDVVITYNTLIDGLCK</u> <u>SGKIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGKLDEAEELLEEMEEKGIK</u> <u>PDVVITYTTLIDGLCKAGKVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGK</u> <u>IEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDV</u> <u>VTYTTLIDGLCKAGKVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGKIEEA</u> <u>LKLFKEMEEKGITPSVVTYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDVVITY</u> <u>TTLIDGLCKAGKVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGKIEEALKL</u> <u>FKEMEEKGITPSVVTYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDVVITYTTL</u> <u>IDGLCKAGKVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGKIEEALKLFKE</u> <u>MEEKGITPSVVTYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDVVITYTTLIDG</u> <u>LCKAGKVDEALELFDKMERGIKPDDELTYNTLISGLGKAGRARDPPVLSSGSG</u> GSGGGGHHHHHHGR
L u c – C U G – P P R 4 (SBQ ID NO. 14)	MAGVFTLEDVFGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPIQRIVLS GENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVID GVTPNMIDYFGRPYEGIAVFDGKKITVTGTWNGNKKIIDERLINPDGSLLFRV TINGVTGWRLCERILAGGGGSLVPRGSGGGGSSRLHMGNSV <u>VTYNTNIDQL</u> <u>CKSGRIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEK</u> <u>GIKPDVVITYTTLIDGLCKAGRVDEALELFDKMERGIKPDVVITYNTLIDGLCKS</u> <u>GRIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIK</u> <u>DVVITYTTLIDGLCKAGRVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGRIE</u> <u>EALKLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVV</u> <u>TYTTLIDGLCKAGRVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGRIEEAL</u> <u>KLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVITYT</u> <u>TLIDGLCKAGRVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGRIEEALKLF</u> <u>KEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVITYTTLI</u> <u>DGLCKAGRVDEALELFDKMERGIKPDVVITYNTLIDGLCKSGRIEEALKLFKE</u> <u>MEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVITYTTLIDG</u> <u>LCKAGRVDEALELFDKMERGIKPDDELTYNTLISGLGKAGRARDPPVLSSGSG</u> GSGGGGHHHHHHGR

[0076]

The details of the sequences of the PPR motifs contained in the PPR protein in the above Table 1 are shown below. All of the PPR proteins comprise 18 PPR motifs.

[Table 2]

CUG-PPR1

Amino Acid Numbering																																					
15															18																						
1	4	6	9																18																		
1	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	PPR code	Target residue
2	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VNS	C
3	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VND	U
4	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VTD	G
5	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VNS	C
6	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VND	U
7	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
8	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U
9	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
10	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
11	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U
12	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
13	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
14	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U
15	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
16	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
17	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U
18	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G

Motifs

CUG-PPR1

Motifs

[Table 3]

Amino Acid Numbering																																		
1	4	6	9	15	11													PPR code	Target residue															
1	V	T	N	I	D	Q	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
2	V	T	N	I	D	G	L	C	K	A	G	R	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
3	V	T	N	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
4	V	T	N	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
5	V	T	N	I	D	G	L	C	K	A	G	R	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
6	V	T	N	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
7	V	T	N	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
8	V	T	N	I	D	G	L	C	K	A	G	R	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
9	V	T	N	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
10	V	T	N	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
11	V	T	N	I	D	G	L	C	K	A	G	R	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
12	V	T	N	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
13	V	T	N	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
14	V	T	N	I	D	G	L	C	K	A	G	R	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
15	V	T	N	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
16	V	T	N	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
17	V	T	N	I	D	G	L	C	K	A	G	R	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
18	V	T	N	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	E	VTD	G

Motifs

CUG-PPR2

Motifs

[Table 4]

CUG-PPR3																																					
Amino Acid Numbering																																					
15															ii																						
1	4	6	9																15															PPR code	Target residue		
1	V	T	Y	N	T	N	I	D	Q	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
2	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	K	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
3	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
4	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
5	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	K	L	D	E	A	E	L	L	E	E	M	E	E	K	G	I	K	P	D	V	VND	U	
6	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
7	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
8	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	K	L	D	E	A	E	L	L	E	E	M	E	E	E	K	G	I	K	P	D	V	VND	U
9	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
10	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
11	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	K	L	D	E	A	E	L	L	E	E	M	E	E	E	K	G	I	K	P	D	V	VND	U
12	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
13	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
14	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	K	L	D	E	A	E	L	L	E	E	M	E	E	E	K	G	I	K	P	D	V	VND	U
15	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
16	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	K	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
17	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	K	L	D	E	A	E	L	L	E	E	M	E	E	E	K	G	I	K	P	D	V	VND	U
18	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	K	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	E	VTD	G

Motifs

[Table 5]

CUG-PPR4	Amino Acid Numbering																																					
	1	4	6	9	15															ii								PPR code	Target residue									
Motifs	1	V	T	Y	N	T	N	I	D	Q	L	C	K	S	G	R	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
	2	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	M	E	E	K	G	I	K	P	D	V	VND	U	
	3	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	R	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
	4	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	R	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
	5	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	M	E	E	K	G	I	K	P	D	V	VND	U	
	6	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	R	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
	7	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	R	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
	8	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	M	E	E	K	G	I	K	P	D	V	VND	U	
	9	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	R	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
	10	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	R	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
	11	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	M	E	E	K	G	I	K	P	D	V	VND	U	
	12	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	R	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
	13	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	R	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
	14	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	M	E	E	K	G	I	K	P	D	V	VND	U	
	15	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	R	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G
	16	V	T	Y	N	T	L	I	D	G	L	C	K	S	G	R	I	E	E	A	L	K	L	F	K	E	M	E	E	K	G	I	T	P	S	V	VNS	C
	17	V	T	Y	N	T	L	I	D	G	L	C	K	A	G	R	L	D	E	A	E	E	L	L	E	M	E	E	K	G	I	K	P	D	V	VND	U	
	18	V	T	Y	T	T	L	I	D	G	L	C	K	A	G	R	V	D	E	A	L	E	L	F	D	E	M	K	E	R	G	I	K	P	D	V	VTD	G

Motifs

[0077]

(2) Preparation of PPR protein

The expression plasmid vector constructed as above was gene transferred into *E. coli* Rosetta (DE3) strain. This *E. coli* was cultured in 2 mL of LB medium containing 100 µg/µL ampicillin at 37°C for 12 hours. The culture medium was transferred to an incubator at 15°C when OD₆₀₀ reached 0.5 to 0.8, and left still for 30 minutes. Then, 100 µL (final concentration 0.1 mM IPTG) was added, and this was cultured at 15°C for 16 hours. Centrifugation was performed at 5,000 x g, 4°C, for 10 minutes, the *E. coli* pellets were collected, 1.5 mL of lysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.5% NP-40, 1 mM MgCl₂, 2 mg/ml lysozyme, 1 mM PMSF, 2 µl of DNase) was added, and this was frozen at -80°C for 20 minutes. Frozen homogenization of cells was carried out at 25°C for 30 minutes with shaking. Subsequently, centrifugation operation was performed at 3700 rpm, 4°C, for 15 minutes, and the supernatant comprising soluble PPR protein (*E. coli* lysate) was collected and employed for the following experiment.

[0078]

(3) Binding test of PPR protein to RNA

The binding test between the PPR protein and RNA was carried out with a binding experiment method between the PPR protein and biotinylated RNA on a streptavidin plate. RNA probes were synthesized to have biotin modified at the 5'-terminal of each of a 30-base RNA comprising the target CUG x 7 sequence (GACACUGCUGCUGCUGCUGCUGCUGAUGCA (SEQ ID NO. 15)) and a 30-base RNA comprising a non-target CAG x 6 (GACAUGCCAGCAGCAGCAGCAGCAGGACUG (SEQ ID NO. 16)) (consigned to Greiner). 2.5 pmol of the biotinylated RNA probes were added to a streptavidin-coated plate (Cat No. 15502, Thermo fisher), and allowed to react for 30 minutes at room temperature. This was washed with probe washing buffer (20 mM Tris-HCl (pH 7.6), 150 mM NaCl, 5 mM MgCl₂, 0.5% NP-40, 1 mM DTT, 0.1% BSA). A well with lysis buffer without adding biotinylated RNA was also prepared for background measurement (designated "- Probe"). Then, blocking buffer (20 mM Tris-HCl (pH 7.6), 150 mM NaCl, 5 mM

MgCl₂, 0.5% NP-40, 1 mM DTT, 1% BSA) was added, and the plate surface was blocked for 30 minutes at room temperature. One hundred microliters of the E. coli lysate comprising luciferase-fusion PPR protein having 1.5×10^8 LU/ μ L of luminescence amount was added to the wells, and binding reaction was performed for 30 minutes at room temperature. Washing with 200 μ L of washing buffer (20 mM Tris-HCl (pH 7.6), 150 mM NaCl, 5 mM MgCl₂, 0.5% NP-40, 1 mM DTT) was performed 5 times, 40 μ L of the luciferase substrate (Promega, E151A) diluted 2500-folds with the washing buffer was added to the wells and allowed to react for 5 minutes, and then the amount of luminescence was measured with a plate reader (PerkinElmer, Cat No. 5103-35).

[0079]

<Results>

The results are shown in Figure 1. All of Luc-CUG-PPR 1, Luc-CUG-PPR 2, Luc-CUG-PPR 3, and Luc-CUG-PPR 4 had binding ability against the CUG probe and no binding ability against the CAG probe. In other words, all of the PPR proteins had high binding specificity against the CUG repeat sequence.

[0080]

[Example 2: Effect of PPR on RNA aggregate formation in DM1 model cells]

PPR protein was allowed to act on DM1 model cells to measure the number of RNA aggregates with FISH (fluorescent in situ hybridization) method.

[0081]

<Experimental materials and methods>

(1) DM1 model cells

A cell having the DMPK gene region comprising 800 CTG repeats introduced into a C2C12 cell (C2C12-DMPK800R, Nucleic Acids Research, 2014, 42(10), 6591-6602) was used as the DM1 cell model. C2C12-DMPK800 was produced with the following procedure. In other words, a plasmid (pLC16) comprising a sequence having 800 CTGs inserted into the 3' untranslated region of the DMPK gene and a

plasmid that expresses PhiC31 integrase were both transfected to mouse myoblast strain C2C12 with Nucleofector (product name, from Lonza), and then stable expression strains were selected with a selection medium supplemented with puromycin. Subsequently, a plasmid that expresses Cre recombinase was transfected to the stable expression strain, and clones in which transcription of mRNA having 800 CUG repeats is induced were selected with a selection medium supplemented with hygromycin.

[0082]

(2) Production of PPR protein expression plasmid vector for cell transfection

A fusion gene in the order of green fluorescent protein mClover3 gene, PPR protein gene, 3 x nucleus local signal, 3 x FLAG epitope tag was cloned into a mammal expression plasmid vector. The expression of the fusion gene is controlled by a CMV promoter and a SV40 poly-A signal. Cloning of the gene having the correct size was confirmed by PCR method, and the gene sequence was confirmed by sequencing. Moreover, as control, a fusion gene that does not comprise the PPR protein gene, i.e., a fusion gene of green fluorescent protein mClover3 gene, 3 x nucleus local signal, and 3 x FLAG epitope tag, was prepared with the same method. These were respectively designated mCLO-CUG-PPR 1, mCLO-CUG-PPR 2, mCLO-CUG-PPR 3, mCLO-CUG-PPR 4, and mCLO-empty, and employed in subsequent tests. Note that the sequences of the PPR proteins (PPR 1 - PPR 4) that are encoded by each expression plasmid vector are identical to the sequences of each of the PPR proteins described in Table 1.

[0083]

[Table 6]

	<u>TTLIDGLCKAGKVDEALELFDKMERGIKPDVVYNTLIDGLCKSGKIEEALKL</u> <u>FKEMEEKGITPSVVYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVYTTTL</u> <u>IDGLCKAGKVDEALELFDKMERGIKPDDELTYNTLISGLGKAGRARDPPVLS</u> PPKKRRKVEDPPKRRKVEDPPKRRKVGSDYKDHDGDYKDHDIDYKDDDDK MAGR
m C L o – C U G – P P R 3 (SEQ ID NO. 19)	MAGVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICT TGKLPVPWPTLVTTFGYGVACFSRYPDHMKQHDFFKSAMPEGYVQERTISF KDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFSHYVYIT ADKQKNCKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSHQ SKLSKDPNEKRDHMLLEFVTAAGITHGMDELYKSRLHMGNSVVTYNTNID <u>QLCKSGKIEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGKLDEAEELLEEME</u> <u>EKGKIPDVVYTTTLIDGLCKAGKVDEALELFDKMERGIKPDVVYNTLIDGL</u> <u>CKSGKIEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGKLDEAEELLEEMEEK</u> <u>GIKPDVVYTTTLIDGLCKAGKVDEALELFDKMERGIKPDVVYNTLIDGLCK</u> <u>SGKIEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGKLDEAEELLEEMEEKGIK</u> <u>PDVVYTTTLIDGLCKAGKVDEALELFDKMERGIKPDVVYNTLIDGLCKSGK</u> <u>IEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDV</u> <u>VYTTTLIDGLCKAGKVDEALELFDKMERGIKPDVVYNTLIDGLCKSGKIEEA</u> <u>LKLFKEMEEKGITPSVVYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDVVY</u> <u>TTLIDGLCKAGKVDEALELFDKMERGIKPDVVYNTLIDGLCKSGKIEEALKL</u> <u>FKEMEEKGITPSVVYNTLIDGLCKAGKLDEAEELLEEMEEKGIKPDVVYTTTL</u> <u>IDGLCKAGKVDEALELFDKMERGIKPDDELTYNTLISGLGKAGRARDPPVLS</u> PPKKRRKVEDPPKRRKVEDPPKRRKVGSDYKDHDGDYKDHDIDYKDDDDK MAGR
m C L o – C U G – P P R 4 (SEQ ID NO. 20)	MAGVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICT TGKLPVPWPTLVTTFGYGVACFSRYPDHMKQHDFFKSAMPEGYVQERTISF KDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFSHYVYIT ADKQKNCKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSHQ SKLSKDPNEKRDHMLLEFVTAAGITHGMDELYKSRLHMGNSVVTYNTNID <u>QLCKSGRIIEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGRLDEAEELLEEME</u> <u>EKGKIPDVVYTTTLIDGLCKAGRVDEALELFDKMERGIKPDVVYNTLIDGL</u> <u>CKSGRIIEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGRLDEAEELLEEMEEK</u> <u>GIKPDVVYTTTLIDGLCKAGRVDEALELFDKMERGIKPDVVYNTLIDGLCKS</u> <u>GRIIEEALKLFKEMEEKGITPSVVYNTLIDGLCKAGRLDEAEELLEEMEEKGIK</u> <u>PDVVYTTTLIDGLCKAGRVDEALELFDKMERGIKPDVVYNTLIDGLCKSGRIE</u> <u>EALKLFKEMEEKGITPSVVYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVV</u>

	<u>TYTTLIDGLCKAGRVDEALELFDEMKERGIKPDVVTYNTLIDGLCKSGRIEEAL</u> <u>KLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYT</u> <u>TLIDGLCKAGRVDEALELFDEMKERGIKPDVVTYNTLIDGLCKSGRIEEALKLF</u> <u>KEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYTTLI</u> <u>DGLCKAGRVDEALELFDEMKERGIKPD</u> <u>ELTYNTLISGLGKAGRARDPPVLSSP</u> KKRKRVEDPKKKRKVEDPKKKRKVGSDYKDHDGDYKDHDIDYKDDDDKM AGR
m C L o - e m p t y (SEQ ID NO. 21)	MAGVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICT TGKLPVPWPTLVTTFGYGVACFSRYPDHMKQHDFFKSAMPEGYVQERTISF KDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFSHYVYIT ADKQKNCIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSHQ SKLSKDPNEKRDHMLVLEFVTAAGITHGMDELYKSRLHMGSGGLSSPKKKR KVEDPKKKRKVEDPKKKRKVGSDYKDHDGDYKDHDIDYKDDDDKMAGR

[0084]

(3) Counting of RNA aggregate formation

DM1 model cells were cultured with DMEM medium containing 10% FBS and penicillin/streptomycin at 37°C under 5% CO₂ condition. Two hundred nanograms of the plasmid DNA constructed as above, 0.6 μL Eugene™-HD (Promega, E2311), 200 μL Opti-MEM were mixed, the whole amount was added to the wells, cultured at 37°C under 5% CO₂ environment for 72 hours, and then fixed with 3% paraformaldehyde for 15 minutes at room temperature. After fixation, this was washed twice with PBS, and then permeabilized with PBS containing 0.5% TritonX-100 for 5 minutes. Subsequently, this was subjected to 10 minutes of prehybridization treatment with 2 x SSC buffer containing 30% formamide. Then, this was subjected to 1 hour of hybridization treatment at 37°C with 2 x SSC buffer containing 30% formamide, 2 μg/mL BSA, 66 μg/mL yeast tRNA, 2 mM vanadyl complex (vanadyl complex), and 1 ng/μL Texas Red CAG probe. After subjecting to 30 minutes of post hybridization treatment with 2 x SSC buffer containing 30% formamide at 42°C, this was washed once with 1 x SSC buffer, and then washed twice with PBS. After fixing with Vectashield with DAPI (product name, from Vector Laboratories), the number of RNA aggregates inside the nucleus was counted with a fluorescence microscope (Keyence PZ-9000).

[0085]

<Results>

The results are shown in Figure 2. As is clear from Figure 2, in DM1 model cells of the untreated group, the rate of RNA aggregate positive cells was greater than 40%, whereas in 4 experiment groups that were applied PPR protein, the rate of RNA aggregate positive cells was significantly reduced.

[0086]

[Example 3: Effect of PPR protein on splicing abnormality in DM1 model cells]

PPR protein was applied to DM1 model cells, and it was verified whether or not the splicing abnormality at exon 22 of the Atp2a1 gene improved.

[0087]

<Experimental materials and methods>

(1) DM1 model cells

The same DM1 model cells as in Example 2 (C2C12-DMPK800R) was used.

[0088]

(2) PPR protein

The same expression plasmid vector as in Example 2 was used.

[0089]

(3) Quantification of splicing product

PPR protein expression plasmid vector was gene transferred into DM1 model cells via the same method as in Example 2, and after 72 hours, total RNA was extracted with RNeasy Mini Plus Kit (product name, from Qiagen). As control, DM1 model cells without gene transfer of PPR protein expression plasmid vector were employed, and total RNA was similarly extracted. Subsequently, cDNA was prepared with SuperScript III First Strand Synthesis System (from Invitrogen). After subjecting the cDNA to RNaseH treatment, RT-PCR was performed with the following Atp2a1 exon 22 RT primers. The RT-PCR product was subjected to electrophoresis with 2% agarose gel, stained with

GelRed, and then normal type PCR products and abnormal type PCR products were each quantified with an image analyzer (ChemIDoc Touch Imaging System, from BioRad).

Atp2a1 exon 22 RT primers

Fw: GCTCATGGTCCTCAAGATCTCAC (SEQ ID NO. 22)

Rv: GGGTCAGTGCCTCAGCTTTG (SEQ ID NO. 23)

[0090]

<Results>

The results are shown in Figure 3. The lower portion is the electrophoresis result of the RT-PCR product, and the upper portion is a graph showing the proportion of the normal type in the RT-PCR product. In the PPR protein untreated group, the proportion of the abnormal type was larger than the normal type. By allowing PPR protein to act, the proportion of the normal type in DM1 model cells significantly increased, showing that the splicing abnormality improved.

[0091]

[Example 4: Effect of PPR protein on muscle differentiation abnormality in DM1 model cells]

PPR protein was allowed to act on DM1 model cells, and it was verified whether or not the muscle differentiation efficiency of DM1 model cells improved.

[0092]

<Experimental materials and methods>

(1) DM1 model cells

The same DM1 model cells as in Example 2 (C2C12-DMPK800R) was used.

[0093]

(2) PPR protein

The same expression plasmid vector of PPR protein as in Example 2 was used.

[0094]

(3) Quantification of muscle differentiation efficiency

PPR protein expression plasmid vector was gene transferred into DM1 model cells via the same method as in Example 2. On the next day, the medium was exchanged to a differentiation medium, and 72 hours after transfection, nucleus staining by DAPI as well as antibody staining with anti-MyHC antibody were carried out. Antibody staining was first carried out, cells having 2 or more nuclei inside the cell were further set as myotube cells, and the total number of nuclei inside the myotube cells confirmed in the observation field was counted (hereinafter referred to as value A). Subsequently, the total number of nuclei in the whole field including not only the myotube cells in the above passage but also undifferentiated cells was counted (hereinafter referred to as value B). The Fusion index represents the percentage of value A divided by value B, and means the muscle differentiation efficiency.

[0095]

<Results>

The results are shown in Figure 4. The empty vector administration group (represented as "no" in the figure) showed a Fusion index of about 3. On the other hand, the PPR administration group all had a Fusion index of about 5 - 15, achieving recognition of significant improvement of muscle differentiation efficiency.

[0096]

[Example 5: Effect of PPR in DM1 model mice]

To DM1 model mice, AAV6 loaded with the CUG-PPR gene was administered in single dose into the anterior tibial muscle, and the effect on myotonia symptoms in the anterior tibial muscle, the effect on the splicing abnormality of the Atp2a1 gene and the skeletal muscle-type chloride channel (Clcn1) gene, as well as the number of RNA aggregate formation were verified.

[0097]

<Experimental materials and methods>

(1) Animals used

As the DM1 model mice, HSA^{LR} mice (granted from Dr. Charles Thornton (University of Rochester) and bred in the laboratory animal facility at University of Osaka) were used. HSA^{LR} mouse is a transgenic animal having the CTG repeat sequence elongated 220 times at the 3' untranslated region of the hACTA gene (human gene constitutively expressed in myocytes) integrated into the genome, and the symptoms of DM1 appear when the mRNA having the CUG sequence elongated is expressed in myocytes (Science, 2000, 289 (5485), 1769-73). As control, wild-type mice (FVB/NJcl mice, purchased from CLEA Japan) were used. Fourteen weeks-old female and male mice were employed for the experiments.

[0098]

(2) PPR protein

(Production of AAV vector)

To the multi-cloning site of an AAV vector (component of AAVproTM Helper Free System (AAV6), Takara, 6651), a gene having green fluorescent protein mCLOver3, PPR protein, 3 x nuclear localization signal, and 3 x FLAG epitope tag fused in order was inserted. Verification of the expression size of the constructed gene by PCR, and verification of the inserted sequence by sequencing was performed. The expression of the fusion gene is controlled by the CMV promoter and the Human growth hormone polyA signal.

[0099]

(Preparation of AAV vector)

Twenty micrograms each of the AAV vector constructed in the preceding section, the pRC6 vector and the phelper vector (both components of AAVproTM Helper Free System (AAV6), Takara, 6651) were gene transferred into 2.5×10^8 cells of the AAVproTM HEK293T cell line (Takara, 632273) with Polyethylenimine, Linear, MW 25000 (Polysciences, Inc., 23966-1). Three days after transfection, AAV producer cells were obtained. AAVproTM Purification Kit Maxi (all serotypes) (Takara, 6666) was employed for AAV purification. Titration was carried out with AAVpro titration kit for real time

PCR ver. 2 (Takara, 6233), and this was diluted to 5×10^{11} vg/mL with D-PBS and set as the administration solution. These were respectively designated AAV6-mCLO-CUG-PPR 1, AAV6-mCLO-CUG-PPR 2, and AAV6-mCLO-empty, and employed in subsequent tests. Note that the sequences of the PPR proteins (PPR 1, PPR 2) that are encoded by each AAV vector are identical to the sequences of each of the PPR proteins described in Table 1.

[0100]

(3) Grouping and administration schedule

The four groups of the AAV6-mCLO-CUG-PPR 1 administration DM1 group, the AAV6-mCLO-CUG-PPR 2 administration DM1 group, the AAV6-mCLO-empty administration DM1 group, and the PBS administration DM1 group were set ($n = 5$ for each group (left and right limbs of 3 males and 2 females were used)). The dosage of AAV was 1×10^{10} vg/leg and administered in single dose. To the PBS administration DM1 group, the same volume of PBS was administered in single dose.

[0101]

(4) Counting of RNA aggregate formation

Fifty-six days after the final administration, the anterior tibial muscle was collected from the mice and sections were produced. These were washed twice with PBS, and then permeabilized with PBS containing 0.5% TritonX-100 for 5 minutes. Subsequently, these were subjected to 10 minutes of prehybridization treatment with $2 \times$ SSC buffer containing 30% formamide. Then, these were subjected to 1 hour of hybridization treatment at 37°C with $2 \times$ SSC buffer containing 30% formamide, $2 \mu\text{g/mL}$ BSA, $66 \mu\text{g/mL}$ yeast tRNA, 2 mM vanadyl complex (vanadyl complex), and $1 \text{ ng}/\mu\text{L}$ Texas Red CAG probe. After subjecting to 30 minutes of post hybridization treatment with $2 \times$ SSC buffer containing 30% formamide at 42°C , these were washed once with $1 \times$ SSC buffer, and then washed twice with PBS. After fixing with Vectashield with DAPI (product name, from Vector Laboratories), the number of RNA aggregates inside the nucleus was counted with a fluorescence microscope (Keyence PZ-9000).

[0102]

(5) Quantification of splicing product

Fifty-six days after the final administration, the anterior tibial muscle was collected from the mice. After extracting total RNA with TRI Reagent (product name, MRC), cDNA was prepared with SuperScript III First Strand Synthesis System (from Invitrogen). After subjecting the cDNA to RNaseH treatment, RT-PCR was performed with Atp2a1 exon 22 RT primers similar to Example 2 and the following Clcn1 exon 7a RT primers. The RT-PCR product was subjected to electrophoresis with 2% agarose gel, stained with GelRed, and then normal type PCR products and abnormal type PCR products were each quantified with an image analyzer (ChemIDoc Touch Imaging System, from BioRad), and the proportion of the normal type in the RT-PCR product was calculated. T-test was employed for statistical analysis.

Clcn1 exon 7a RT primers

Fw: TGAAGGAATACCTCACACTCAAGG (SEQ ID NO. 24)

Rv: CACGGAACACAAAGGCACTG (SEQ ID NO. 25)

[0103]

(6) Measurement of myotonia

Fifty-six days after the final administration, the electrical myotonia phenomenon of the mouse anterior tibial muscle was analyzed with needle electromyography under anesthesia. Specifically, the occurrence frequency of electrical myotonia phenomenon when a needle electrode was inserted 20 times into the mouse anterior tibial muscle was evaluated in the following 4 stages.

Severity 0: No electrical myotonia phenomenon is seen at all.

Severity 1: Electrical myotonia phenomenon is seen in less than 50% of total.

Severity 2: Electrical myotonia phenomenon is seen in 50% or more and less than 90% of total.

Severity 3: Electrical myotonia phenomenon is seen in 90% or more.

[0104]

<Results>

The results of RNA-foci analysis are shown in Figure 5. In the PBS administration group and the AAV6-mCLO-empty administration group, the positive cell rate of the RNA-foci was shown to be about 30%. On the other hand, for the AAV6-mCLO-CUG-PPR 1 and the AAV6-mCLO-CUG-PPR 2 administration groups, significant decrease in the RNA-foci positive rate was recognized.

[0105]

The evaluation results of splicing abnormality are shown in Figure 6. In regard to *Atp2a1*, in the PBS administration DM1 group and the AAV6-mCLO-Empty administration MyoD group, the proportion of the normal type was shown to be about 60%. In the AAV6-mCLO-CUG-PPR 1 administration DM1 group and the AAV6-mCLO-CUG-PPR 2 administration DM1 group, it became clear from the significant increase in the proportion of the normal type RT-PCR product that the splicing abnormality of the *Atp2a1* gene is improved by PPR protein administration. Similarly for *Clcn1*, it was recognized that the splicing abnormality of the *Clcn1* gene was improved by PPR administration.

[0106]

The analytical results of myotonia are shown in Figure 7. For the PBS administration group and the AAV6-mCLO-empty administration group, myotonia of Severity 2 was detected in all five out of five mice. For both the AAV6-mCLO-CUG-PPR 1 administration DM1 group and the AAV6-mCLO-CUG-PPR 2 administration DM1 group, two out of five mice showed Severity 1 and the remaining three mice showed Severity 2. Moreover, the improved two cases were from the experiment group with high improvement effect for *Clcn1* splicing abnormality (Figure 8).

[0107]

From the above results, it was shown that splicing abnormality based on abnormal elongation of the CTG repeat sequence is improved by the present invention, and as a result, myotonia symptoms in DM1 is improved.

[0108]

[Example 6: Dose-dependency test of drug effect with M1 model mice]

In order to evaluate the dose-dependency of the drug effect employing DM1 model mice, AAV9 loaded with the CUG-PPR gene was administered in single dose to DM1 model mice in the tail vein at each amounts of 3×10^{13} , 1×10^{14} , and 3×10^{14} vg/kg, and the effect on myotonia symptoms in the femoral muscle, the effect on the splicing abnormality of the Atp2a1 gene and the skeletal muscle-type chloride channel (Clcn1) gene, as well as the number of RNA aggregate formation were verified.

[0109]

<Experimental materials and methods>

(1) Animals used

As the DM1 model mice, HSA^{LR} mice (granted from Dr. Charles Thornton (University of Rochester) and bred in the laboratory animal facility at University of Osaka) were used. HSA^{LR} mouse is a transgenic animal having the CTG repeat sequence elongated 220 times at the 3' untranslated region of the hACTA gene (human gene constitutively expressed in myocytes) integrated into the genome, and the symptoms of DM1 appear when the mRNA having the CUG sequence elongated is expressed in myocytes (Science, 2000, 289 (5485), 1769-73). As control, wild-type mice (FVB/NJcl mice, purchased from CLEA Japan) were used. Fourteen weeks-old female and male mice were employed for the experiments.

[0110]

(2) PPR protein

(Production of AAV vector)

To the multi-cloning site of a pAAV-CMV vector (component of AAVproTM Helper Free System, Takara, 6651), a gene having PPR protein and 3 x nuclear localization signal fused in order was inserted. Verification of the expression size of the constructed gene by PCR, and verification of the inserted sequence by sequencing was performed. The expression of the fusion gene is controlled by the CMV promoter and the Human growth hormone polyA signal.

[0111]

(Preparation of AAV vector)

The production of the AAV vector constructed in the preceding section was outsourced to SignaGen Laboratories (hereinafter abbreviated as SG Company). Packaging was carried out by co-transfecting the packaging cell strain with the Rep/Cap plasmid and the Helper plasmid from SG Company and the AAV vector sent from the inventors' company. The cell homogenate comprising the packaged AAV was purified by density gradient centrifugation and titrated. The AAV obtained was 1.16×10^{14} vg/ml and the volume was 5 ml. This was employed in the subsequent experiment as AAV9-CUG-PPR 1. Note that the sequences of the PPR proteins (PPR 1) encoded in each AAV vector are identical to the sequences of the PPR proteins described in Table 1.

[0112]

(3) Grouping and administration schedule

The four groups of the AAV9-CUG-PPR 1 low dose administration group (hereinafter abbreviated as Low dose: LD), the AAV9-CUG-PPR 1 medium dose administration group (hereinafter abbreviated as Middle dose: MD), the AAV9-CUG-PPR 1 high dose administration group (hereinafter abbreviated as High dose: HD), and the PBS administration group (hereinafter abbreviated as PBS) were set. $n = 5$ for each group (3 male and 2 female mice). The dosage of AAV was 3×10^{13} vg/kg for LD, 1×10^{14} vg/kg for MD, and 3×10^{14} vg/kg for HD, and administered in single dose in the tail vein. To the PBS administration group, the same volume of PBS was administered in single dose.

[0113]

(4) Counting of RNA aggregate formation

Twenty-eight days after the final administration, the femoral muscle was collected from the mice and sections were produced. These were washed twice with PBS, and then permeabilized with PBS containing 0.5% TritonX-100 for 5 minutes. Subsequently, these were subjected to 10 minutes of prehybridization treatment with $2 \times$ SSC

buffer containing 30% formamide. Then, these were subjected to 1 hour of hybridization treatment at 37°C with 2 x SSC buffer containing 30% formamide, 2 µg/mL BSA, 66 µg/mL yeast tRNA, 2 mM vanadyl complex (vanadyl complex), and 1 ng/µL Texas Red CAG probe. After subjecting to 30 minutes of post hybridization treatment with 2 x SSC buffer containing 30% formamide at 42°C, these were washed once with 1 x SSC buffer, and then washed twice with PBS. After fixing with Vectashield with DAPI (product name, from Vector Laboratories), the number of RNA aggregates inside the nucleus was counted with a fluorescence microscope (Keyence PZ-9000).

[0114]

(5) Quantification of splicing product

Twenty-eight days after the final administration, the femoral muscle was collected from the mice. After extracting total RNA with TRI Reagent™ (MRC), cDNA was prepared with SuperScript III First Strand Synthesis System (from Invitrogen). After subjecting the cDNA to RNaseH treatment, RT-PCR was performed with Atp2a1 exon 22 RT primers similar to Example 2 and the following Clcn1 exon 7a RT primers. The RT-PCR product was subjected to electrophoresis with 2% agarose gel, stained with GelRed, and then normal type PCR products and abnormal type PCR products were each quantified with an image analyzer (ChemIDoc Touch Imaging System, from BioRad), and the proportion of the normal type in the RT-PCR product was calculated. T-test was employed for statistical analysis.

[0115]

Clcn1 exon 7a RT primers

Fw: TGAAGGAATACCTCACACTCAAGG (SEQ ID NO. 24)

Rv: CACGGAACACAAAGGCACTG (SEQ ID NO. 25)

[0116]

(6) Measurement of myotonia

Twenty-eight days after the final administration, the electrical myotonia phenomenon of the mouse femoral muscle was analyzed with needle electromyography under anesthesia. Specifically, the occurrence frequency of electrical myotonia

phenomenon when a needle electrode was inserted 20 times into the mouse femoral muscle was evaluated in the following 4 stages.

Severity 0: No electrical myotonia phenomenon is seen at all.

Severity 1: Electrical myotonia phenomenon is seen in less than 50% of total.

Severity 2: Electrical myotonia phenomenon is seen in 50% or more and less than 90% of total.

Severity 3: Electrical myotonia phenomenon is seen in 90% or more.

[0117]

(7) Measurement of amount of PPR mRNA expression in AAV administration mouse tissue

The resected femoral muscle organ fragment was wet with 800 μ L of the TRIzol Reagent. Two homogenization beads (TOMY, SUB-30) were added, and the organ was homogenized with a beads cell homogenizing device Micro Smash (TOMY SEIKO, MS-100R). Six hundred microliters of the supernatant obtained after centrifugation at 4°C, 15,000 g for 5 minutes were collected, 160 μ L of chloroform solution was added, and this was mixed by inversion at room temperature for 5 minutes. Then, this was centrifugated at 4°C, 15,000 g for 5 minutes, and separated into the aqueous phase and the organic phase. Three hundred microliters of the aqueous phase were collected, and this was subjected to nucleic acid purification device Maxwell RSC Instrument to obtain a purified total-RNA product. The concentration and the purity of the RNA was confirmed by carrying out absorbance measurement with NanoDrop 8000. In order to verify the degree of degradation of the RNA obtained, 20 - 250 ng/ μ L of the total-RNA was used to perform electrophoresis by LabChip GX Touch HT (PerkinElmer, CLS137031J), and it was confirmed from the band strength ratio of 28S and 18S rRNA that degradation had not proceeded. For cDNA synthesis, 20 - 100 ng/ μ L of the RNA obtained was subjected to reverse transcription reaction according to the product protocol of SuperScript III Reverse Transcriptase (Thermo Fisher Scientific (Life Technologies), 18080085). For the amounts of PPR mRNA and GAPDH mRNA, the cDNA produced, Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent, 600882), and the primers shown below were

mixed according to manual, and then amplified with real-time PCR system Aria MX (Agilent) to perform quantitative analysis.

[0118]

Primers for PPR quantitative PCR

Fw: GATGAGGCTTTGGAAGTGTGTTG (SEQ ID NO. 26)

Rv: TCTCTGGCTCTGCCGGCCTTGC (SEQ ID NO. 27)

[0119]

Primers for GAPDH quantitative PCR

Fw: ATCATCCCTGCCTCTACTGG (SEQ ID NO. 28)

Rv: CTGCTTCACCACCTTCTTGA (SEQ ID NO. 29)

[0120]

<Results>

The results of RNA-foci analysis are shown in Figure 9. The femoral muscle organ sections obtained were employed to carry out RNA-FISH, and the RNA-foci formation suppression effect by AAV9-CUG-PPR 1 was verified. As a result, dose-dependent suppression effect on the formation of RNA-foci was confirmed in the AAV9-CUG-PPR 1 administration group. In regard to the RNA-foci positive cell rate, formation suppression effect observed were 39% for the PBS administration group, 18% for the AAV9-CUG-PPR 1 low dose group, 23% for the AAV9-CUG-PPR 1 medium dose group, and 17% for the AAV9-CUG-PPR 1 high dose administration group.

[0121]

The evaluation results of splicing abnormality are shown in Figure 10. Total RNA was collected from the organ obtained, reverse transcription PCR was performed, and abnormal splicing of the Clcn1 gene and the Atp2a1 gene were evaluated. The abnormal splicing phenomenon of Clcn1 is the Cl channel which is known to be the direct cause of myotonic discharge phenomenon. Atp2a1 is a Ca transporter likewise related to myotonic discharge, and is known to be a sensitive marker gene. As a result of splicing analysis, in the detection system of Clcn1, the PBS administration group showed 58% normal splicing, whereas the AAV9-CUG-PPR 1 low dose group showed 72%, the AAV9-CUG-PPR 1 medium dose group showed 78%, and the AAV9-CUG-PPR 1 high dose group showed 82%, confirming a

dose-dependent splicing improvement. In the detection system of Atp2a1, the PBS administration group showed 25% normal splicing, whereas the AAV9-CUG-PPR 1 low dose group showed 56%, the AAV9-CUG-PPR 1 medium dose group showed 60%, and the AAV9-CUG-PPR 1 high dose group showed 64%, showing a dose-dependent splicing improvement.

[0122]

The decrease improvement effect of myotonic discharge by PPR was represented by a score in 4 stages based on the number of myotonic discharge. In other words, Score 0 means detection of myotonic discharge in 0 out of 20 times, Score 1 means 1 - 9 times out of 20, Score 2 means 10 - 19 times out of 20, and Score 3 means 20 out of 20 times. As a result of myotonic discharge analysis employing the femoral muscle, the PBS administration group showed Score 3, the AAV9-CUG-PPR 1 low dose group showed Score 3 in 1/5 case and Score 2 in 4/5 cases, the medium dose group showed Score 2 in 4/5 cases and Score 1 in 1/5 case, and the low dose group showed improvement to Score 2 in 3/5 cases and Score 1 in 2/5 cases (Figure 11).

[0123]

In order to investigate the pharmacokinetics, total RNA was extracted from the analytical organ femoral muscle and quantitative PCR of PPR mRNA was carried out. As a result of RT-qPCR, dose-dependent expression of PPR mRNA was confirmed (Figure 12). For the expression amount of PPR mRNA for each group, the relative value when the value of the individual with the lowest signal in the AAV9-CUG-PPR 1 low dose administration group was set as 1 using GAPDH as the internal standard was used. Taking the median of each group, the PBS administration group was 0.2, the low dose administration group was 0.7, the medium dose administration group was 1.8, and the high dose administration group was 2.6, confirming a dose-dependent PPR expression (Figure 12).

[0124]

From the above results, dose-dependent suppression of myotonic discharge of AAV9-CUG-PPR 1, RNA-foci formation suppression effect, and abnormal splicing improvement effect were confirmed.

[0125]

[Example 7: Time base change in drug effect in DM1 model mice]

In order to evaluate the time base change in the drug effect employing DM1 model mice, AAV9 loaded with the CUG-PPR gene was administered in single dose to DM1 model mice in the tail vein, and the effect on myotonia symptoms in the femoral muscle at each time point, the effect on the splicing abnormality of the Atp2a1 gene and the skeletal muscle-type chloride channel (Clcn1) gene, as well as the number of RNA aggregate formation were verified.

[0126]

<Experimental materials and methods>

(1) Animals used

As the DM1 model mice, HSA^{LR} mice (granted from Dr. Charles Thornton (University of Rochester) and bred in the laboratory animal facility at University of Osaka) were used. HSA^{LR} mouse is a transgenic animal having the CTG repeat sequence elongated 220 times at the 3' untranslated region of the hACTA gene (human gene constitutively expressed in myocytes) integrated into the genome, and the symptoms of DM1 appear when the mRNA having the CUG sequence elongated is expressed in myocytes (Science, 2000, 289 (5485), 1769-73). As control, wild-type mice (FVB/NJcl mice, purchased from CLEA Japan) were used. Twelve weeks-old female and male mice were employed for the experiments.

[0127]

(2) PPR protein

(Production of AAV vector)

To the multi-cloning site of a pAAV-CMV vector (component of AAVproTM Helper Free System, Takara, 6651), a gene having PPR protein and 3 x nuclear localization signal fused in order was inserted. Verification of the expression size of the constructed gene by PCR, and verification of the inserted sequence by sequencing was performed. The expression of the fusion gene is controlled by the CMV promoter and the Human growth hormone polyA signal.

[0128]

(Preparation of AAV vector)

The production of the AAV vector constructed in the preceding section was consigned to SignaGen Laboratories (hereinafter abbreviated as SG Company). Packaging was carried out by co-transfecting the packaging cell strain with the Rep/Cap plasmid and the Helper plasmid from SG Company and the AAV vector sent from the inventors' company. The cell homogenate comprising the packaged AAV was purified by density gradient centrifugation and titrated. The AAV obtained was 1.16×10^{14} vg/ml and the volume was 5 ml. This was employed in the subsequent experiment as AAV9-CUG-PPR 1. Note that the sequences of the PPR proteins (PPR 1) encoded in each AAV vector are identical to the sequences of the PPR proteins described in Table 1.

[0129]

(3) Grouping and administration schedule

The four groups of 14 days after administration (2 w), 28 days after administration (4 w), 56 days after administration (8 w), and 112 days after administration (16 w) of AAV9-CUG-PPR 1 were set. $n = 5$ for each group (2 males and 3 females) of mice was used. The dosage of AAV was 3×10^{14} vg/kg, and administered in single dose in the tail vein.

[0130]

(4) Counting of RNA aggregate formation

14, 28, 56, or 112 days after the final administration, the femoral muscle was collected from the mice and sections were produced. These were washed twice with PBS, and then permeabilized with PBS containing 0.5% TritonX-100 for 5 minutes. Subsequently, these were subjected to 10 minutes of prehybridization treatment with 2 x SSC buffer containing 30% formamide. Then, these were subjected to 1 hour of hybridization treatment at 37°C with 2 x SSC buffer containing 30% formamide, 2 µg/mL BSA, 66 µg/mL yeast tRNA, 2 mM vanadyl complex (vanadyl complex), and 1 ng/µL Texas Red CAG probe. After subjecting

to 30 minutes of post hybridization treatment with 2 x SSC buffer containing 30% formamide at 42°C, these were washed once with 1 x SSC buffer, and then washed twice with PBS. After fixing with Vectashield with DAPI (product name, from Vector Laboratories), the number of RNA aggregates inside the nucleus was counted with a fluorescence microscope (Keyence PZ-9000).

[0131]

(5) Quantification of splicing product

14, 28, 56, or 112 days after the final administration, the femoral muscle was collected from the mice. After extracting total RNA with TRI Reagent™ (MRC), cDNA was prepared with SuperScript III First Strand Synthesis System (from Invitrogen). After subjecting the cDNA to RNaseH treatment, RT-PCR was performed with Atp2a1 exon 22 RT primers similar to Example 2 and the following Clcn1 exon 7a RT primers. The RT-PCR product was subjected to electrophoresis with 2% agarose gel, stained with GelRed, and then normal type PCR products and abnormal type PCR products were each quantified with an image analyzer (ChemIDoc Touch Imaging System, from BioRad), and the proportion of the normal type in the RT-PCR product was calculated. T-test was employed for statistical analysis.

[0132]

Clcn1 exon 7a RT primers

Fw: TGAAGGAATACCTCACACTCAAGG (SEQ ID NO. 24)

Rv: CACGGAACACAAAGGCACTG (SEQ ID NO. 25)

[0133]

(6) Measurement of myotonia

Immediately before autopsy, the electrical myotonia phenomenon of the mouse femoral muscle was analyzed with needle electromyography under anesthesia. Specifically, the occurrence frequency of electrical myotonia phenomenon when a needle electrode was inserted 20 times into the mouse femoral muscle was evaluated in the following 4 stages.

Severity 0: No electrical myotonia phenomenon is seen at all.

Severity 1: Electrical myotonia phenomenon is seen in less than 50% of total.

Severity 2: Electrical myotonia phenomenon is seen in 50% or more and less than 90% of total.

Severity 3: Electrical myotonia phenomenon is seen in 90% or more.

[0134]

<Results>

The results of RNA-foci analysis are shown in Figure 13. The femoral muscle organ sections obtained were employed to carry out RNA-FISH, and the RNA-foci formation suppression effect by AAV9-CUG-PPR 1 was verified. As a result, dose-dependent suppression effect on the formation of RNA-foci was confirmed in the AAV9-CUG-PPR 1 administration group. The value of the PBS administration group employed in the dose-dependency test of Example 6 was used as control. In regard to the RNA-foci positive cell rate, formation suppression effect confirmed were 39% for the control group, 36% for the Day 14 after administration of AAV9-CUG-PPR 1 group, 23% for the Day 28 after administration group, 19% for the Day 56 after administration group, and 18% for the Day 112 after administration group (Figure 13).

[0135]

The evaluation results of splicing abnormality are shown in Figure 14. Total RNA was collected from the organ obtained, reverse transcription PCR was performed, and abnormal splicing of the *Clcn1* gene and the *Atp2a1* gene were evaluated. The abnormal splicing phenomenon of *Clcn1* is the Cl channel which is known to be the direct cause of myotonic discharge phenomenon. *Atp2a1* is a Ca transporter likewise related to myotonic discharge, and is known to be a sensitive marker gene. As a result of splicing analysis, in the detection system of *Clcn1*, the PBS administration group shown in Example 6 showed 58% normal splicing, whereas the experiment group of Day 14 after administration of AAV9-CUG-PPR 1 showed 62%, Day 28 after administration of AAV9-CUG-PPR 1 showed 76%, and Day 56 after administration of AAV9-CUG-PPR 1 showed 79%, and 80% splicing improvement was confirmed for Day 112 after administration of

AAV9-CUG-PPR 1. In the detection system of Atp2a1, the PBS administration group shown in Example 6 showed 25% normal splicing, whereas the experiment group of Day 14 after administration of AAV9-CUG-PPR 1 showed 30%, Day 28 after administration of AAV9-CUG-PPR 1 showed 63%, and Day 56 after administration of AAV9-CUG-PPR 1 showed 76%, and 75% splicing improvement was confirmed for Day 112 after administration of AAV9-CUG-PPR 1 (Figure 14).

[0136]

The decrease improvement effect of myotonic discharge by PPR was represented by a score in 4 stages based on the number of myotonic discharge. In other words, Score 0 means detection of myotonic discharge in 0 out of 20 times, Score 1 means 1 - 9 times out of 20, Score 2 means 10 - 19 times out of 20, and Score 3 means 20 out of 20 times. As a result of myotonic discharge analysis employing the femoral muscle, the PBS administration group shown in Example 6 showed Score 3, the experiment group of Day 14 after administration of AAV9-CUG-PPR 1 showed Score 3 in 4/5 cases and Score 2 in 1/5 case, the experiment group of Day 28 after administration of AAV9-CUG-PPR 1 showed Score 2 in 5/5 cases, the experiment group of Day 56 after administration of AAV9-CUG-PPR 1 showed Score 2 in 3/5 cases and Score 1 in 2/5 cases, and the experiment group of Day 112 after administration showed Score 2 in 3/5 cases and Score 1 in 2/5 cases (Figure 15).

[0137]

In order to investigate the pharmacokinetics, total RNA was extracted from the analytical organ femoral muscle and quantitative PCR of PPR mRNA was carried out. As a result of RT-qPCR, continuous expression of PPR was confirmed even at Day 112 after administration (data not shown).

[0138]

From the above results, it was confirmed that AAV9-CUG-PPR 1-administered mice showed improvement of drug effect until 14 days, 28 days, and 56 days after administration, and the drug effect still continued at Day 112 after administration.

[0139]

[Example 8: Comprehensive improvement of DM1-related splicing abnormality by AAV9-CUG-PPR 1 administration]

<Experimental materials and methods>

(1) Analytical organ

To HSALR mice (14 weeks-old), AAV9-CUG-PPR 1 was administered in single dose at 3×10^{14} vg/kg in the tail vein, and the femoral muscle tissue fragment resected 8 weeks after administration (Day 56) was set as the PPR therapy group. The same volume of PBS was administered in single dose, and the resected femoral muscle was set as the PPR untreated group. The femoral muscle tissue fragment of a wild-type FVB/NJcl mouse having similar genetic background was employed as the wild-type group.

[0140]

(2) Analysis of abnormal splicing events

Following the total RNA extraction method described in Example 6, total-RNA was extracted from the femoral muscle of the PPR therapy group and the PPR untreated group and the wild-type group, and necessary amounts were sent to GENEZWIZ to carry out RNA Sequence (Illumina Hiseq 2 x 150 bp sequence). In the RNA sent, ribosomal RNA in the total RNA was removed by Poly A selection method. RNA was reverse transcribed, cDNA was synthesized, and a library with adaptor was prepared. Sequencing was performed, and adaptor trimming was carried out on the lead information obtained with Trimmomatic (v0.33) 59. Employing GRCm39 (obtained from Ensemble) as the reference genome, alignment was performed on the lead sequence after trimming with STAR (v2.7.9a). Specification of splicing events was detected with rMATS (v3.2.5) based on the STAR-aligned file (.bam). For specification of the varied splicing events, comparison of the wild-type group and the PPR untreated group (HSALR mice administered with PBS) was performed, the presence or absence of variation was judged with "p-adjusted-value" and "InclLevelDifference" as indicators, and only the candidates that were defined to be varied were extracted. To the varied splicing

events, cutoff with a threshold of FDR (False Discovery Rate)-adjusted P value < 0.05 was performed and employed for a comparison test between groups. The Percent spliced-in index (PSI) was defined as the proportion of the amount of spliced-in transcript against the total lead number of the target gene. Metascape (<https://metascape.org/gp/index.html#/main/step1>) was used for gene ontology analysis.

[0141]

<Results>

Total RNA was extracted from the femoral muscle resected from wild-type mice (14 weeks-old) described in the preceding section (hereinafter referred to as the wild-type group), the femoral muscle resected 8 weeks after PBS administration to HSALR mice (hereinafter referred to as the PPR untreated group), and further the femoral muscle resected 8 weeks (Day 56) after AAV9-CUG-PPR 1 administration (hereinafter the PPR treatment group), and exhaustive analysis of expressed RNA was carried out by next-generation sequencing technology.

[0142]

After mapping of the lead sequence obtained, the number of leads of adjacent exon sequences were compared, PSI (percent spliced-in index) was calculated, and the inclusion levels of each exon of each gene were calculated. Among the PSI of the wild-type and the PBS administration groups, those having the difference of 0.2 or more were set to be abnormal splicing events in HSALR mice (n = 455, Figure 16a left). As a result of comparing the PSI of the same splicing events at Week 8 after administration of AAV9-CUG-PPR 1 and the PSI in WT (Figure 16a right), it was found that PSI had generally recovered.

[0143]

In the following literature, improvement was also seen in splicing events reported to have splicing abnormality in HSA-LR other than *Atp2a1* or *Cln1* (Figure 16b).

Reference literature: Matthew K Tanner et al. Nucleic Acids Research, Volume 49, Issue 4 (2021)

[0144]

Gene ontology analysis was performed on the gene group with abnormal splicing, and they were the terms related to muscle function or muscle differentiation. As a result of analyzing the number of genes in each gene ontology that achieved recognition of improvement, it was found that genes achieving recognition of strong improvement (greater than 50%) was about 40%, coming to a total of 90% including those achieving recognition of improvement (greater than 20%) (Figure 16c).

[0145]

From these, it was found that various splicing abnormalities are improved by administration of AAV9-CUG-PPR 1.

[0146]

[Example 9: Analysis of PPR mRNA expression amount in each organ of normal mice that were applied AAV9-CUG-PPR 1]

[0147]

<Experimental materials and methods>

(1) Analytical organ

Four organ fragments (heart, femoral muscle, gastrocnemius muscle, and anterior tibial muscle) resected from normal mice iv-administered with AAV9-CUG-PPR 1 (administration dosage: 1×10^{13} vg/kg, 3×10^{13} vg/kg, or 3×10^{14} vg/kg) were used.

[0148]

(2) Administration schedule and autopsy timing

Eight weeks-old normal female mice were employed. Administration dosages were 1×10^{13} vg/kg, 3×10^{13} vg/kg, or 3×10^{14} vg/kg and administration was performed in the tail vein, and the volume was set to be 10 μ L/g. Observation duration was set to be 28 days after administration. The heart, the femoral muscle, the gastrocnemius muscle, and the anterior tibial muscle was resected from each mouse individual to perform analysis of PPR mRNA.

[0149]

(3) Measurement of amount of PPR mRNA expression in AAV administration mouse tissue

Total-RNA was extracted from the resected organ fragments, cDNA was synthesized, and quantitative PCR was carried out. First, 50 mg ($\pm$ 5 mg) fragments of the organ fragments obtained in the preceding section were wet with 800 μ L of the TRIzol Reagent. Two homogenization beads (TOMY, SUB-30) were added, and the organs were homogenized with a beads cell homogenizing device Micro Smash (TOMY SEIKO, MS-100R), and this was centrifugated at 4°C, 15,000 g for 5 minutes. Six hundred microliters of the supernatant obtained was collected, 160 μ L of chloroform solution was added, and this was mixed by inversion at room temperature for 5 minutes. Then, this was centrifugated at 4°C, 15,000 g for 5 minutes, and separated into the aqueous phase and the organic phase. Three hundred microliters of the aqueous phase were collected, and this was subjected to nucleic acid purification device Maxwell RSC Instrument to obtain a purified total-RNA product. The concentration and the purity of the RNA was confirmed by carrying out absorbance measurement with NanoDrop 8000. In order to verify the degree of degradation of the RNA obtained, 20 - 250 ng/ μ L of the total-RNA was used to perform electrophoresis by LabChip GX Touch HT (PerkinElmer, CLS137031J), and it was confirmed from the band strength ratio of 28S and 18S rRNA that degradation had not proceeded.

[0150]

For cDNA synthesis, 20 - 100 ng/ μ L of the RNA obtained was subjected to reverse transcription reaction according to the product protocol of SuperScript III Reverse Transcriptase (Thermo Fisher Scientific (Life Technologies), 18080085).

[0151]

Quantitative PCR was carried out with real-time PCR system Aria MX (Agilent) according to the general manual of Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent, 600882).

[0152]

Primers for PPR quantitative PCR

Fw: GATGAGGCTTTGGAAGTGTTC (SEQ ID NO. 26)

Rv: TCTCTGGCTCTGCCGGCCTTGC (SEQ ID NO. 27)

[0153]

Primers for GAPDH quantitative PCR

Fw: ATCATCCCTGCCTCTACTGG (SEQ ID NO. 28)

Rv: CTGCTTCACCACCTTCTTGA (SEQ ID NO. 29)

[0154]

<Results>

The results of quantitative PCR are shown in Figure 17. As a result of quantitative PCR, a rise in dose-dependent PPR expression amount in each tissue was confirmed, and a maximum expression amount was shown in the maximally administered 3×10^{14} vg/kg administration group.

[0155]

From the above, it was shown that AAV9-CUG-PPR 1 can be systematically delivered and expressed, and suggests that it is effective for DM1 which is a multi-organ disease that affects the heart in addition to skeletal muscles and smooth muscles.

[0156]

[Example 10: Analysis of PPR mRNA expression persistency in each organ of normal mice that were applied AAV9-CUG-PPR 1]

[0157]

<Experimental materials and methods>

(1) Analytical organ

Four organ fragments (heart, femoral muscle, gastrocnemius muscle, and anterior tibial muscle) resected from normal mice iv-administered with AAV9-CUG-PPR 1 (administration dosage: 1×10^{14} vg/kg) were used.

[0158]

(2) Administration schedule and autopsy timing

Eight weeks-old normal male mice were employed. Administration dosage was 1×10^{14} vg/kg and administration was performed in the tail vein, and the volume was set to be 10 μ L/g. Observation duration was set to be Day 14 after administration, Day

28 after administration, Day 56 after administration, Day 112 after administration, and Day 182 after administration. The heart, the femoral muscle, the gastrocnemius muscle, and the anterior tibial muscle was resected from each mouse individual to perform analysis of PPR mRNA.

[0159]

(3) Measurement of amount of PPR mRNA expression in AAV administration mouse tissue

Total-RNA was extracted from the resected organ fragments, cDNA was synthesized, and quantitative PCR was carried out. First, 50 mg ($\pm$ 5 mg) fragments of the organ fragments obtained in the preceding section were wet with 800 μ L of the TRIzol Reagent. Two homogenization beads (TOMY, SUB-30) were added, and the organs were homogenized with a beads cell homogenizing device Micro Smash (TOMY SEIKO, MS-100R). Six hundred microliters of the supernatant obtained after centrifugation at 4°C, 15,000 g for 5 minutes was collected, 160 μ L of chloroform solution was added, and this was mixed by inversion at room temperature for 5 minutes. Then, this was centrifuged at 4°C, 15,000 g for 5 minutes, and separated into the aqueous phase and the organic phase. Three hundred microliters of the aqueous phase were collected, and this was subjected to nucleic acid purification device Maxwell RSC Instrument to obtain a purified total-RNA product. The concentration and the purity of the RNA was confirmed by carrying out absorbance measurement with NanoDrop 8000. In order to verify the degree of degradation of the RNA obtained, 20 - 250 ng/ μ L of the total-RNA was used to perform electrophoresis by LabChip GX Touch HT (PerkinElmer, CLS137031J), and it was confirmed from the band strength ratio of 28S and 18S rRNA that degradation had not proceeded. For cDNA synthesis, 20 - 100 ng/ μ L of the RNA obtained was subjected to reverse transcription reaction according to the product protocol of SuperScript III Reverse Transcriptase (Thermo Fisher Scientific (Life Technologies), 18080085). Quantitative PCR was carried out with real-time PCR system Aria MX (Agilent) according to the general manual of Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent, 600882).

[0160]

Primers for PPR quantitative PCR

Fw: GATGAGGCTTTGGAAGTGTGTTG (SEQ ID NO. 26)

Rv: TCTCTGGCTCTGCCGGCCTTGC (SEQ ID NO. 27)

[0161]

Primers for GAPDH quantitative PCR

Fw: ATCATCCCTGCCTCTACTGG (SEQ ID NO. 28)

Rv: CTGCTTCACCACCTTCTTGA (SEQ ID NO. 29)

[0162]

<Results>

The results of quantitative PCR are shown in Figure 18. As a result of quantitative PCR, continued expression was confirmed even at Day 183 after administration, and a rise in time-dependent expression was also confirmed. It is shown in the preceding articles that expression can last for a long time (6 months or longer) when AAV9 is administered in the tail vein, and the results of the present experiment also agree with the said knowledge.

[0163]

From the above, it was suggested that the present invention may exert the desired effect for a long time with fewer number of administrations to the subject.

CLAIMS

1. A pharmaceutical composition for treating myotonic dystrophy type 1, wherein

said pharmaceutical composition comprises a nucleic acid encoding a protein that specifically binds to the CUG repeat sequence,

said protein that specifically binds to the CUG repeat sequence comprises at least 6 pentatricopeptide repeat (PPR) motifs consisting of polypeptides of 30 - 38 amino acids long represented by Formula 1,

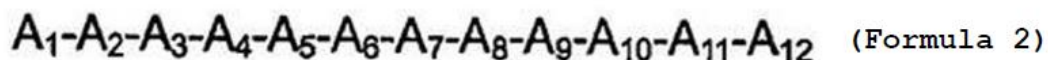
[Formula 1]



[wherein in Formula 1:

Helix A is a 12 amino acids long moiety that is capable of forming an α -helix structure, and is represented by Formula 2,

[Formula 2]



wherein in Formula 2, $A_1 - A_{12}$ each independently represents an amino acid;

X is absent, or is a moiety consisting of a length of 1 - 9 amino acids;

Helix B is a moiety that is capable of forming an α -helix structure consisting of a length of 11 - 13 amino acids;

L is a moiety represented by Formula 3 that is 2 - 7 amino acids long;

[Formula 3]



and wherein in Formula 3, each amino acid is numbered from the C-terminal such as "i" (-1), "ii" (-2),

provided that $L_{iii} - L_{vii}$ may be absent]

and the combination of three amino acids at A_1 , A_4 , and L_{ii} , or the combination of two amino acids at A_4 and L_{ii} in said each PPR motif is selected so that said each PPR motif binds to C, U, or G, and thereby configured so that said protein specifically binds to said CUG repeat sequence.

2. The pharmaceutical composition according to claim 1, wherein said protein that specifically binds to the CUG repeat sequence comprises 9 - 30 of said PPR motifs.

3. The pharmaceutical composition according to claim 2, wherein said protein that specifically binds to the CUG repeat sequence comprises 12 - 24 of said PPR motifs.

4. The pharmaceutical composition according to any one of claim 1 to 3, wherein the combination of three amino acids at A_1 , A_4 , and L_{ii} in said each PPR motif is:

when the base to be the target for the PPR motif is A (adenine), the combination of three amino acids at A_1 , A_4 , and L_{ii} is, in the order of (A_1 , A_4 , L_{ii}), (valine, threonine, asparagine), (phenylalanine, serine, asparagine), (phenylalanine, threonine, asparagine), (isoleucine, asparagine, aspartic acid), or (threonine, threonine, asparagine);

when the base to be the target for the PPR motif is G (guanine), the combination of three amino acids at A_1 , A_4 , and L_{ii} is, in the order of (A_1 , A_4 , L_{ii}), (glutamic acid, glycine, aspartic acid), (valine, threonine, aspartic acid), (lysine, threonine, aspartic acid), or (leucine, threonine, aspartic acid);

when the base to be the target for the PPR motif is U (uracil), the combination of three amino acids at A₁, A₄, and L_{ii} is, in the order of (A₁, A₄, L_{ii}), (valine, asparagine, aspartic acid), (isoleucine, asparagine, asparagine), (isoleucine, asparagine, aspartic acid), (isoleucine, methionine, aspartic acid), (phenylalanine, proline, aspartic acid), or (tyrosine, proline, aspartic acid); or

when the base to be the target for the PPR motif is C (cytosine), the combination of three amino acids at A₁, A₄, and L_{ii} is, in the order of (A₁, A₄, L_{ii}), (valine, asparagine, asparagine), (isoleucine, asparagine, asparagine), (valine, asparagine, serine), or (isoleucine, methionine, aspartic acid).

5. The pharmaceutical composition according to any one of claim 1 to 3, wherein the combination of two amino acids at A₄ and L_{ii} in said each PPR motif is:

when the base to be the target for the PPR motif is A (adenine), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (threonine, asparagine), (serine, asparagine), or (glycine, asparagine);

when the base to be the target for the PPR motif is G (guanine), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (threonine, aspartic acid), or (glycine, aspartic acid);

when the base to be the target for the PPR motif is U (uracil), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (asparagine, aspartic acid), (proline, aspartic acid), (methionine, aspartic acid), or (valine, threonine);

when the base to be the target for the PPR motif is C (cytosine), the combination of two amino acids at A₄ and L_{ii} is, in the order of (A₄, L_{ii}), (asparagine, asparagine), (asparagine, serine), or (leucine, aspartic acid).

6. The pharmaceutical composition according to claim 1, characterized in that said nucleic acid encoding a protein that specifically binds to the CUG repeat sequence is integrated into an expression vector.

7. The pharmaceutical composition according to claim 1, characterized in that said nucleic acid encoding a protein that specifically binds to the CUG repeat sequence is integrated into a viral vector.

8. The pharmaceutical composition according to claim 7, characterized in that said viral vector is a viral vector with tropism towards muscle tissue.

9. The pharmaceutical composition according to claim 8, characterized in that said viral vector is an adeno-associated virus (AAV) vector, an adenovirus vector, a retrovirus vector, a lentivirus vector, or a herpes simplex virus vector.

10. The pharmaceutical composition according to claim 9, characterized in that said viral vector is an AAV vector.

11. The pharmaceutical composition according to claim 10, characterized in that said AAV vector is an AAV1 vector, an AAV2 vector, an AAV6 vector, an AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV10 vector, an AAV11 vector, or an AAV12 vector.

12. An expression vector comprising the nucleic acid encoding the protein that specifically binds to the CUG repeat sequence as defined in claim 1.

13. A cell comprising the expression vector according to claim 12.

14. A protein that specifically binds to the CUG repeat sequence that is produced from the cell according to claim 13.

15. A viral expression vector comprising the nucleic acid encoding the protein that specifically binds to the CUG repeat sequence as defined in claim 1.

16. A cell comprising the viral expression vector according to claim 15.

17. A viral vector that is produced from the cell according to claim 16.

18. The use of the nucleic acid encoding the protein that specifically binds to the CUG repeat sequence as defined in claim 1 in the manufacture of myotonic dystrophy type 1 therapeutic medicine.

19. A method for treating myotonic dystrophy type 1, comprising a step of applying to a subject a therapeutically effective amount of the pharmaceutical composition according to claim 1.